

ecos de  
**Economía**

A Latin American Journal of Applied Economics

ISSN 1657-4206  
e-ISSN 2462-8107

**Vol. 21 - No. 44**  
Enero – Junio de 2017  
PP. 1-104  
Medellín - Colombia

**Ecoss de Economía:**

A Latin American Journal of Applied Economics

Frequency: 2 issues/year

ISSN: 1657 - 4206

e-ISSN: 2462-8107

**Focus and Scope**

Ecoss de Economía is an international journal of applied economics, with a particular interest in the quantitative analysis of economic, financial, and public policy issues and/or themes related to Latin America. Articles that include and analyze national data are particularly welcome, since these data are often unavailable to international researchers outside the region. The journal is not interested in manuscripts that are solely conceptual in nature.

**Abstracting & Indexing Services**

The Journal is found indexed and reported in: Pub-  
lindex "A2" | Latindex | SciELO Citation Index - WoS  
| Scielo\_Colombia | EconLit | Economía y Negocios-  
EBSCOhost | Fuente Académica Premier – EBSCO-  
host | Business Source Corporate Plus - EBSCOhost  
| ABI/INFORM Complete – Proquest | ABI/INFORM  
Global – Proquest | DoTEc | RePEc | Ideas | EconPa-  
pers | E-revist@s | Index Copernicus International |  
Dialnet | PKP – OJS | Informe Académico – Gale |  
Actualidad Iberoamericana | Clase - UNAM | DOAJ |  
DRJI | Google Scholar | JournalTOCs | ResearchBib |  
ProQuest Social Science Database | Business Insights:  
Global | Redalyc | ECONIS | LatAm-Studies

**Editor-in-Chief**

Theodore R. Breton, Universidad EAFIT, Colombia

**Managing Editor**

Oscar Caicedo-Alarcon, Universidad EAFIT, Colombia

**Board of Editors**

**Klaus F. Zimmermann**, Harvard University,  
United States

**Subal Kumbhakar**, Binghamton University - State Uni-  
versity of New York, United States

**Ramazan Gençay**, Simon Fraser University, Canada

**Jorge Martínez-Vazquez**, Georgia State University,  
United States

**Rodrigo Soares**, Escola de Economía de São Paulo -  
Fundação Getúlio Vargas, Brazil

**Enrique López-Bazo**, University of Barcelona, Spain

**Sergio Firpo**, Insper Institute of Education and Re-  
search, Brazil

**Associate Editors**

**Sergio S. Urzúa**, University of Maryland, United States

**Daniel Bergstresser**, Brandeis University, United States

**Félix Rioja**, Georgia State University, United States

**Hugo Ñopo**, GRADE Group for the Analysis of Devel-  
opment, Peru

**Claudio Bravo-Ortega**, Universidad de Chile, Chile

**Luis Díaz-Serrano**, Universitat Rovira i Virgili, Spain

**Luis H. Gutierrez**, Universidad del Rosario, Colombia

**Juan C. Duque**, Universidad EAFIT, Colombia

**Gustavo Canavire Bacarreza**, Universidad  
EAFIT, Colombia

**Carlos Medina**, Banco de la República, Colombia

**Jorge N. Valero-Gil**, Universidad Autónoma Nuevo  
León, Mexico

**Jorge Guillén Uyen**, Universidad ESAN, Peru

**Diego A. Agudelo Rueda**, Universidad EAFIT, Colombia

**Contact Editorial Office:**

Ecoss de Economía: A Latin American Journal of Applied Economics

Universidad EAFIT, School of Economics and Finance

Cra. 49 7sur 50 Oficina 26-206, Medellín

Phone: (57) (4) 261 95 00 ext. 9465 • Fax: (57) (4) 261 9294

A.A. 3300 - ecoseco@eafit.edu.co

# Contents

Unions and Economic Performance in Developing Countries: Case Studies from Latin America <i>Los sindicatos y rendimiento económico en los países en desarrollo: Estudios de Casos de América Latina</i> <i>Fernando Rios-Avila</i> .....	5
Estimación robusta de betas y el ratio de cobertura sobre futuros de índices bursátiles en el Mercado Integrado Latinoamericano (MILA) <i>Robust Estimation of beta and the hedging ratio in Stock Index Futures in the Integrated Latin American Market</i> <i>Andrés Gómez, Astrid K. Gutiérrez y Juan C. Gutiérrez</i> .....	38
Sistema Vial Terciario y Educación Rural en Antioquia (Colombia): un modelo probit ordenado y multivariado <i>The Rural Roads Impact on Education Performance in Antioquia (Colombia): an ordered probit model</i> <i>Guillermo David Hincapié Velez, Ivan Montoya Gomez, John Jaime Bustamante</i> .....	73
Efecto de los aumentos en la escolaridad en el mercado laboral colombiano entre 2008 y 2016 <i>The Effect of Increased Schooling in the Colombian Labor Market Between 2008 and 2016</i> <i>Tomás Aristizábal Lopera, Esteban Ángel López</i> .....	87

---

UNIONS AND ECONOMIC  
PERFORMANCE IN DEVELOPING  
COUNTRIES: CASE STUDIES  
FROM LATIN AMERICA

Los sindicatos y rendimiento  
económico en los países  
en desarrollo: Estudios de  
Casos de América Latina

Fernando Rios-Avila

---

Research Article

## UNIONS AND ECONOMIC PERFORMANCE IN DEVELOPING COUNTRIES: CASE STUDIES FROM LATIN AMERICA\*

### Los sindicatos y rendimiento económico en los países en desarrollo: Estudios de Casos de América Latina

Fernando Rios-Avila<sup>a</sup>

**Key words:** Unions; Productivity; Profitability; Investment; Latin America; Developing Countries

**Palabras clave:** Sindicatos; Productividad; América latina; Países en desarrollo

**JEL Classifications:** J50, L1, O54

**Received:** 02/02/2017

**Accepted:** 10/04/2017

**Published:** 15/06/2017

#### Abstract

This paper analyzes the economic impact of unions on productivity in the manufacturing sector across six Latin American countries: Argentina, Bolivia, Chile, Mexico, Uruguay, and Panama. Using an augmented Cobb-Douglas production function, the paper finds that unions have positive, but mostly small, effects on productivity, with the exception of Argentina, with a large negative effect, and Bolivia, with no effect. An analysis on profitability shows that, in most cases, the positive productivity effects barely offset higher union compensation, and that unions are negatively related to investment in capital and R & D. Different explanations for these effects are discussed.

#### Resumen

Este artículo analiza el impacto económico de los sindicatos sobre la productividad en el sector manufacturero en seis países de América Latina: Argentina, Bolivia, Chile, México, Uruguay y Panamá. Utilizando una función de producción aumentada de Cobb-Douglas, el documento encuentra que los sindicatos tienen efectos positivos, pero sobre todo pequeños, sobre la productividad, con la excepción de Argentina, con un gran efecto negativo, y Bolivia, sin efecto. Un análisis de la rentabilidad muestra que, en la mayoría de los casos, los efectos positivos de la productividad apenas compensan una indemnización sindical más alta, y que los sindicatos están relacionados negativamente con la inversión en capital y en I + D. En el artículo se discuten diferentes explicaciones de estos efectos.

a. Levy Economics Institute, United States.  
e-mail: friosavi@levy.org

\* This paper has benefited from comments and suggestions by Barry T. Hirsch, Julie L. Hotchkiss, Rachana Bhatt, and Myriam Quispe-Agnoli.

## Introduction

One of the most contentious debates in the literature on unions has been on how unions affect a firm's performance. While most of the literature agrees that unions have mostly negative effects on profitability and investment, there is little consensus on the effects that unions have on productivity (Kuhn 1998; Doucouliagos and Laroche 2003, 2009; Aidt and Tzannatos 2002; Hirsch 2004). To some extent, the uncertainty surrounding these issues reflects problems with data limitations, as well as the underlying heterogeneity in union effects across establishments, industries, and countries.

Although there is a reasonably large literature for several developed economies, there is little evidence regarding "what unions do" for establishment productivity in developing economies.<sup>1</sup> Because businesses in developing economies face different types of obstacles, such as restrictions on access to capital, unfavorable institutions, high levels of corruption, less competitive markets, and unstable business environments, compared to those in developed countries, it is not clear how unions affect productivity.

The purpose of this paper is to expand the literature by analyzing the impact of unions on productivity for six countries in Latin America, namely Argentina, Bolivia, Chile, Mexico, Panama, and Uruguay. Differences in their economic settings and adaptation to market reforms provide a mixture of settings which suggests a cross-country study may prove valuable to better understand the relationship between unions and productivity.

Using data from the World Bank Enterprise Survey, modified Cobb-Douglas production functions are estimated to determine the impact of unionization on establishment productivity, controlling for various measures of establishment characteristics. Due to considerable levels of non-reporting in the survey, a "principled" multiple imputation approach is used to improve the completeness and reliability of the data. The preferred model indicates that unions have slightly positive but mostly insignificant effects on productivity, with Chile and Panama showing the largest, but not significant, union-productivity effects. The exception is Argentina, where the estimates are negative and statistically significant across all specifications. Alternative measures of unionization reveal that the union-productivity relationship is not linear in all countries, which has contributed to the low significance of the results.

The analysis of profitability indicates that in most countries the small gains in productivity are not large enough to offset the higher wage costs faced by unionized establishments. In Bolivia, the profit and productivity evidence is not closely aligned, which seems to be driven by substantial differences on capital intensity. Similarly, the evidence suggests that unions are negatively associated with measures investment and innovation.

The rest of the paper is structured as follows. The first section presents a review of the empirical literature, with emphasis on the research done for Latin America. The second section presents a brief description of the history of unions in Latin America and describes the legal framework under which unions operate in each country. The third and fourth sections describe the data and the empirical strategy. The fifth and sixth sections present results on productivity and performance analysis. Section seven concludes.

---

<sup>1</sup> As Freeman (2010) indicates, the research for developing economies is limited because data for this type of research is typically inadequate, and also because unions in developing countries have been weak and unable to fulfill their role as bargaining agents in their economies.

## 1. How do Unions Affect Productivity?

### 1.1. Theoretical background

There is a large theoretical literature that has explored the potential costs and benefits of unions in terms of firm performance and productivity. Examples include [Brown and Medoff \(1978\)](#), [Addison \(1982\)](#), [Addison and Barnett \(1982\)](#), [Freeman and Medoff \(1984\)](#), [Hirsch and Addison \(1986\)](#), [Turnbull \(1991\)](#) and, more recently, [Hirsch \(2004\)](#) and [Kaufman \(2004\)](#). This literature has identified various channels through which unions can have positive and negative effects on productivity.

According to the “two faces” approach, popularized in [Freeman and Medoff \(1984\)](#), union’s effects on productivity can be described using the monopoly face – from microeconomic theory – and the collective voice/institutional response framework – from the industrial organization literature. The traditional approach analyzes unions as monopolistic agents, stressing the negative aspects of unions and the distortions they create compared to the perfect competition model. Within this framework, unions extract monopoly gains from the employers by constraining the labor supply. This translates into compensation above competitive levels for their members while potentially causing temporary negative productivity shocks.

It is also possible that unions might impose the adoption of inefficient contractual work rules and generate reductions in managerial discretion that may increase the cost of reacting to economic shocks in dynamic economic environments ([Hirsch, 2008](#)).<sup>2</sup> Furthermore, union rent-seeking behavior can further reduce long run productivity by imposing a pseudo “union tax” on capital returns, limiting the adoption of new technology and investment ([Connolly, Hirsch, and Hirschey 1986](#); [Hirsch 1991, 2004](#)).

Constraints in labor supply might also generate wage/price distortions that could force firms to shift toward suboptimal mix of inputs, possibly causing (small) deadweight welfare loss and lower overall labor productivity. These distortions might spuriously increase production per worker if firms shift their input mix toward higher capital intensity and/or higher skilled workers, without generating gains of technical efficiency. This is less likely to be observed to the extent that unions tax the quasi-rents from capital, reducing incentives to increase investment.<sup>3</sup> Besides, although high union wages opens the possibility to employ workers with higher skills, such outcome is unlikely, given repeated bargaining ([Wessels 1994](#); [Hirsch 2004](#)).

The “collective voice/institutional response” face of unions, as described in [Freeman and Medoff \(1984\)](#), puts more emphasis on the positive aspects of unions and their potential roles enhancing operations and labor relationships within establishments. Legally protected unions can freely express their members’ preferences in the workplace, improving communication between employers and employees, inducing managers to alter methods of production and adopt more efficient personnel policies. In turn they can also reduce potential transaction costs associated with turnover, training and recruiting, monitoring and enforcement in the workplace ([Kuhn 1985](#); [Allen 1984](#)). The presence of unionization and pressure for higher wages can increase productivity through shock effects,

---

2 Although there is anecdotal evidence regarding inefficient union work rules, it seems unlikely that such inefficiencies would be long lived in markets with high levels of competition. There is no systematic evidence relating the interactions between union governance, dynamism, and productivity.

3 For theory on unions and quasi-rents, see [Grout \(1984\)](#) and [Baldwin \(1983\)](#). For the earliest empirical test, see [Connolly, Hirsch and Hirschey \(1986\)](#).

reducing the so called “X-inefficiency” through improved operations which could have persisted in the absence of unions ([Hirsch and Addison 1986](#); [Addison and Hirsch 1989](#); [Kaufman 2004](#)).

As [Freeman and Medoff \(1984\)](#) and [Kuhn \(1998\)](#) emphasize, the positive outcomes from the union’s collective voice are constrained to positive and cooperative relationship between management and organized labor. For instance, [Kleiner \(2002\)](#) finds that in the Aircraft industry, overall productivity was considerably lower during periods of conflicts between management and union leaders. While [Black and Lynch \(2001\)](#) and [Bloom and Van Reenen \(2011\)](#) find that negative union productivity effects are driven by unionized plants with traditional management systems, while positive effects are found among those (few) union establishments that adopt “best-method” human resources practices such as incentive pay.

### 1.2. Empirical evidence

As in other aspects of the literature on unions, the inherited endogeneity of the unionization process has made the identification of causal effects difficult ([Lewis 1963](#); [Freeman and Medoff 1984](#); [Hirsch and Addison 1986](#); [Hirsh 2004](#)). The strategy in most studies has been to compare unionized versus nonunionized firms, using cross-section or panel data, to identify the impact that unions had on productivity and performance. More recent studies, using event study and regression discontinuity approaches, have been able to provide estimates closer to causal effects ([DiNardo and Lee 2004](#); [Lee and Mas 2012](#)).

The seminal paper on unions and productivity by [Brown and Medoff \(1978\)](#) is one of the few studies finding a large and positive effect on productivity (22-24%). These results, however, were not supported by subsequent reviews of the literature, which attributed the results to serious data limitations ([Freeman and Medoff 1984](#); [Hirsch and Addison 1986](#); [Hirsch 2004](#)). The rough consensus on U.S. studies is that union productivity effects are, on average, small and non-significant and highly variable across different economic settings ([Doucouliagos and Laroche 2003](#)). When positive, they are too small to fully offset union wage effects, consistent with the findings of lower profitability among union companies ([Fuchs, Krueger and Poterba 1998](#); [Hirsch 2004](#); [Doucouliagos and Laroche 2009](#)). Regression discontinuity analysis of [DiNardo and Lee \(2004\)](#) find unions have an insignificant effect on productivity or wages, although the even study of [Lee and Mas \(2012\)](#) find a negative effect on firm’s stock value.

International evidence for other developed countries likewise finds unions have negative effects on profitability, but also that unions have mostly negative effects on productivity, except for industries with high competition or good relationships between management and unions ([Aidt and Tzannatos 2002](#); [Doucouliagos and Laroche 2003](#)).

Beyond the scope of developed countries, the literature on the economic effects of unions is limited, and little is known about how unions affect productivity in developing countries. A brief summary of the relevant literature in developing Latin American countries is provided.<sup>4</sup>

[Fairris \(2006\)](#) finds, for Mexico, that unions have a positive effect on productivity, but not profitability, which is possibly attributed to higher training rates. [Menezes-Filho, et al. \(2005\)](#) finds that unions in the manufacturing sector in Brazil are correlated with lower levels of profitability and

.....  
<sup>4</sup> A more comprehensive review of the literature can be found in Freeman (2010)



investment, but that some level of unionization could have a positive impact on productivity, particularly in larger firms with profit sharing.

[Saavedra and Toledo \(2005\)](#) finds evidence for Peru that union firms earn lower profits and have lower productivity, but that such effects are strongly related to the firm characteristics. The authors partially attribute this to conflicts between labor and management. [Cassoni, et al. \(2005\)](#) finds a positive effect on productivity and productivity growth in the manufacture sector in Uruguay, with mixed evidence regarding profitability. The authors argue that improvements in productivity might be explained by increased labor stability and lower turnover, and to a lesser extent improved cooperation and labor morale. Finally, [Urizar and Lee \(2005\)](#), studying coffee producers in Guatemala, find evidence that becoming unionized decreases productivity, although the estimates seem to be tied to firm characteristics.

## 2. Unions in Latin America: Background

There is a substantial literature focused on the development of unions in Latin America, most of which has taken a historical and legal approach describing the evolution of the labor movements in these countries. This section provides a brief overview of important features in the development of unions and legal framework in Argentina, Bolivia, Chile, Mexico, Panama and Uruguay.<sup>5</sup>

### 2.1. History

Most unions in Latin America have been characterized as playing strong roles in the political arena, both in opposition to and as supporters of the political parties in power ([Carrière, Haworth and Roddick 1989](#)). The economic and political development of unions in Latin America is, to no small degree, a story of union alliances and unions strongly influenced by government. Argentina, Bolivia, Chile, Mexico, Panama and Uruguay are no exception.

The alliances between unions and the Peronist Party in Argentina and the Partido Revolucionario Institucional (PRI) in Mexico became long-lasting relationships that benefited the unions for decades. In Bolivia and Chile, alliances between the government and unions were more fragile, with strong unions that played important roles as government allies and opposition. In Panama, during the government of Omar Torrijos (from 1968-1978), the president promoted reforms and encouraged alliances to empower the formation of stronger and more active unions. In Uruguay, where little if any coordination existed between unions and the government, unions played a major role in the democratization process of the 1980s. While the strong alliances between unions and government became the pillar of unions' bargaining power, they weaken their role as collective bargaining in the private sector [O'Connell \(1999\)](#).

The era of dictatorships in Latin America, between the 1970s and late 1980s, produced a major setback for unions. With the exception of Panama, unions were dissolved and persecuted. In Argentina, Bolivia and Mexico, where unions were declared illegal, they remained active opposing the dictatorship. In Chile, while unions were initially disbanded, but following the *Plan Laboral* (late 1970s), the right of association was reestablished, reforming the role of unions as a decentralized unit operating in a newly-adopted neoliberal economy. Finally, in Uruguay, the military regime outlawed union activity

<sup>5</sup> This overview does not pretend to be exhaustive. For a more comprehensive analyses see [Alexander and Parker \(2005\)](#), [Anner \(2008\)](#), [Carrière, Haworth, and Roddick \(1989\)](#), [Cassoni, Allen, and Labadie \(2004\)](#), [Cassoni, et al., \(2005\)](#), [Hudson \(1994\)](#), [Hudson and Hanratty \(1991\)](#), [Hudson and Meditz \(1992\)](#), [Meditz and Hanratty \(1989\)](#), [Merrill and Miró \(1997\)](#), [Murillo \(2000\)](#), [Murillo and Schrank \(2005\)](#), [O'Connell \(1999\)](#), [OECD \(1996\)](#) and [Ulloa \(2003\)](#).

and granted rights of dismissal to employers in case of strikes, producing massive layoffs. These actions effectively eliminated substantive union activities until early 1980s, when unions were allowed to resurface, with close control from the government to reduce the politicization of their activities.

With the return of democracy, unions resurfaced throughout the region. The debt crisis that affected Latin America in the 1980s, however, marked a change in the economic system for most of the countries in the region. In an attempt to overcome the crisis, many countries attempted a series of reforms moving from a centralized and protectionist market, to an open and more flexible market environment with smaller governments. These changes greatly reduced the leverage that unions had on government policies, forcing them to rely on their role in the private sector (O'Connell 1999). Although policies were implemented to favor more flexible labor markets, a series of union-friendly reforms were also adopted during the 1980s and 1990s (Murillo and Schrank 2005). Such reforms were insufficient to strengthen the role of unions as bargaining agents, fragmenting further an already weakened institution (Anner 2008).

## 2.2. Legal Background

As described by Murillo (2000), Murillo and Schrank (2005), Anner (2008) and O'Connell (1999), the aftermath of the debt crisis, the return to democracy, economic reform and subsequent union friendly reforms transformed the role unions had in the labor market. Unions across the region had to adapt to a more flexible labor market, transitioning from a state-union to a firm-union relationship. The combination of legal responses, coupled with a union's background, brought considerable heterogeneity in the way unions operated. Table 1 summarizes important characteristics that describe the conditions under which unions operate in selected countries.

Similar to the experience worldwide, Bolivia, Mexico and Panama have shown a decline in unionization rates, while Argentina and Uruguay have shown a slight increase in unionization rates. According to Anner (2008), the market-oriented reforms weaken unions by eliminating protectionist policies, reducing the public sector, and contributing to the growth of informal labor.<sup>6</sup> The union-friendly reforms were limited and unable to counteract the increasing employment flexibility, and were incapable of providing adequate protection and enforcement systems from anti-union discrimination.

The bargaining system in Argentina, Bolivia and Mexico are characterized by substantial state intervention. In Argentina, while unions can be formed freely, only the union with the most members in a specific industry and/or area is recognized. In Bolivia and Mexico, unions can be formed freely, but need to be authorized by the Department of Labor to be recognized and engage with employers. In Chile and Panama, there is little intervention of the state on the formation of unions other than notification of the authorities. In Uruguay there are no formal regulations on the formation and activities of unions. In most countries freedom of association is guaranteed for all workers except for public officials or government workers.

---

<sup>6</sup> This doesn't imply that workers in the informal sector do not form other types of labor organizations similar to unions. Those organizations, also referred as unions, have characteristics that differ from the traditional role of unions, and are not considered in this research.

**Table 1. Descriptive Evidence on Unions and the Bargaining System**

	Argentina	Bolivia	Chile
GDP per capita 2006 (in US\$)	5485.5	1230.5	8912.2
%manufacture	22.27	14.38	13.20
Union density			
1990-1995	28.7	30.9	13.6
1995-2000	25.6	16.4	11.3
2000-2006	37.6	12.9	13.8
ILO conventions			
C87:Freedom of association	1960	1965	1999
C98:Right to organize and Collective bargain	1956	1973	1999
Freedom of association	<ul style="list-style-type: none"> <li>All workers except for Military personnel</li> </ul>	<ul style="list-style-type: none"> <li>All workers but public administration</li> </ul>	<ul style="list-style-type: none"> <li>All workers but public administration</li> </ul>
Restrictions	<ul style="list-style-type: none"> <li>One union per industry and geographical area recognized.</li> <li>Most representative union is recognized</li> </ul>	<ul style="list-style-type: none"> <li>One union per establishment.</li> <li>Needs government authorization.</li> </ul>	<ul style="list-style-type: none"> <li>More than one union per establishment allowed</li> <li>Unions are automatically recognized</li> </ul>
Union formation	<ul style="list-style-type: none"> <li>Needs to represent at least 20% of the workers</li> </ul>	<ul style="list-style-type: none"> <li>At least 20 workers are needed for professional unions. And 50% for firm unions.</li> </ul>	<ul style="list-style-type: none"> <li>Small firms (less than 50 wrks) need 8 workers to form a union. Otherwise, at least 25 workers are needed.</li> </ul>
Collective bargain	<ul style="list-style-type: none"> <li>Allowed at regional, provincial or firm level</li> <li>Contracts need to be approved by the Ministry of Labor</li> </ul>	<ul style="list-style-type: none"> <li>Allowed for Unions, Federations and Confederations.</li> </ul>	<ul style="list-style-type: none"> <li>Firm level bargain is recognized.</li> <li>National level bargaining is voluntary.</li> <li>Worker associations (nonunionized) can engage into collective bargaining.</li> </ul>
Access to financial information	<ul style="list-style-type: none"> <li>Yes</li> </ul>	<ul style="list-style-type: none"> <li>No</li> </ul>	<ul style="list-style-type: none"> <li>No</li> </ul>
Right to strike	<ul style="list-style-type: none"> <li>Right to strike is recognized</li> <li>Only unions that are registered have the right to strike</li> </ul>	<ul style="list-style-type: none"> <li>Right to strike is recognized. Requires 3/4 support.</li> <li>Strikes in public sector, general strikes and solidary strikes are illegal.</li> </ul>	<ul style="list-style-type: none"> <li>Right to strike is recognized, except in public sector. Requires simple majority support.</li> </ul>
Protection	Adequate	Inadequate	Adequate

Notes: GDP per capita and Manufacture as % of GDP were obtained from the World Bank Indicators (2012). Union Density Information is obtained from Household surveys, [Anner \(2008\)](#), [Hayter and Stoevska \(2011\)](#), [Cassoni et al. \(2005\)](#), and information from the OIT. Characteristics of the bargaining systems were obtained from the countries labor codes, [O'Connell \(1999\)](#), [Anner \(2008\)](#); [Murillo and Schrank \(2005\)](#), [Ronconi \(2012\)](#), [Anner \(2008b\)](#) and [Murillo \(2000\)](#).

**Table 1. Descriptive Evidence on Unions and the Bargaining System (continued)**

	Mexico	Panama	Uruguay
GDP per capita 2006 (in US\$)	8830.8	5201.6	5907.3
%manufacture	18.7	7.1	17.0
Union density			
1990-1995	22.4	14.2	17.3
1995-2000	21.0	11.0	14.7
2000-2006	16.4	12.0	19.0
ILO conventions			
C87:Freedom of association	1950	1958	1954
C98:Right to organize and Collective bargain	Not recognized	1966	1954
Freedom of association	<ul style="list-style-type: none"> <li>No prior authorization is required to create a trade union.</li> </ul>	<ul style="list-style-type: none"> <li>All workers but public officials administration</li> </ul>	<ul style="list-style-type: none"> <li>There exist few regulations on unions</li> </ul>
Restrictions	<ul style="list-style-type: none"> <li>Unions require officially authorization.</li> <li>More than one union per firm allowed, but only the most representative is recognized</li> </ul>	<ul style="list-style-type: none"> <li>Only one union per establishment.</li> </ul>	<ul style="list-style-type: none"> <li>No noticeable restrictions</li> </ul>
Union formation	<ul style="list-style-type: none"> <li>Unions need at least 20 workers</li> </ul>	<ul style="list-style-type: none"> <li>Unions require 40 members.</li> </ul>	<ul style="list-style-type: none"> <li>There are no minimum of requirements</li> </ul>
Collective bargain	<ul style="list-style-type: none"> <li>Employers have the obligation to engage into collective bargaining with unions at request.</li> </ul>	<ul style="list-style-type: none"> <li>Worker associations (nonunionized) can engage into collective bargaining.</li> </ul>	<ul style="list-style-type: none"> <li>Collective bargaining usually at industry level.</li> </ul>
Access to financial information	<ul style="list-style-type: none"> <li>No</li> </ul>	<ul style="list-style-type: none"> <li>Yes</li> </ul>	<ul style="list-style-type: none"> <li>No</li> </ul>
Right to strike	<ul style="list-style-type: none"> <li>Right to Strike recognized. Requires simple majority.</li> <li>Strikes in the Public sector requires 2/3 support</li> </ul>	<ul style="list-style-type: none"> <li>Right to Strike recognized in case of working conditions improvements. Requires simple majority.</li> </ul>	<ul style="list-style-type: none"> <li>No noticeable restrictions</li> </ul>
Protection	Inadequate	Mostly adequate	Adequate

Notes: GDP per capita and Manufacture as % of GDP were obtained from the World Bank Indicators (2012). Union Density Information is obtained from Household surveys, [Anner \(2008\)](#), [Hayter and Stoevska \(2011\)](#), [Cassoni et al. \(2005\)](#), and information from the OIT. Characteristics of the bargaining systems were obtained from the countries labor codes, [O'Connell \(1999\)](#), [Anner \(2008\)](#); [Murillo and Schrank \(2005\)](#), [Ronconi \(2012\)](#), [Anner \(2008b\)](#) and [Murillo \(2000\)](#).

The restrictions on minimum requirements for formation and nature of the bargaining relationship also show significant heterogeneity. Uruguay, in absence of a legal framework, does not have restrictions on union formation. In contrast, Panama requires at least 40 workers to form a union, the strongest restriction in the sample. Argentina, Bolivia, and Mexico, have similar requirements, with more flexibility for smaller establishments in Chile.

The bargaining systems range from a highly centralized system in Argentina to a decentralized system in Chile. In Mexico, the system is highly centralized due to considerable coordination between unions and the state ([O'Connell 1999](#)). Uruguay, which historically had a centralized system, has slowly transition toward bargaining at the firm level ([Cassoni et al., 2005](#)). Chile and Panama have decentralized system that also allows nonunion workers to collectively bargain. In Bolivia, while a mixed

bargaining system is allowed, collective bargaining at centralized levels is common. Only unions in Argentina and Panama have access to financial information before they engage in collective bargaining.

According to an Organisation for Economic Co-operation and Development (OECD) (1996) evaluation, and more recently the Annual Survey of Violations of Trade Union Rights (ITUC 2007), Argentina, Chile and Uruguay have an adequate system to protect unions, while Bolivia and Mexico are considered inadequate. In the case of Panama, except for the requirements to form unions, protection is deemed adequate except in export-processing zones. In terms of enforcement capacity, the information reported in Ronconi (2012), shows that in the 2000s, Bolivia and Mexico had the lowest enforcement capacity, while Chile, Panama and Uruguay, have by far the best enforcement capabilities in the region.

### 3. Data and Imputation Strategy

The present analysis uses data from the Enterprise Survey (ES) 2006, concentrating on the manufacturing sector in six selected Latin American countries (Table 2).<sup>7</sup> The Enterprise Survey provides a standardized establishment level data set, with a representative sample of establishments in the non-agricultural, private sector.<sup>8</sup> The survey provides rich information that can be used to identify aspects of establishment performance, market competition, managerial characteristics and labor force structure, among other things.

**Table 2. Sample Size by Country and Eligibility**

Economy Size	Countries	Sample size (manufacture)	Eligible	Complete
Small	Bolivia	359	298	215
	Panama	238	185	112
	Uruguay	317	251	136
Middle	Argentina	623	540	294
	Chile	627	564	409
Large	Mexico	1,113	974	802
	All	3,277	2,812	1,968

Notes: Prepared from the information of the Enterprise Survey 2006: Eligibility is assessed on whether or not the observation reports information on Sales. Complete data refers to cases when all the basic information (sales, capital and labor) is available for analysis.

A limitation of the ES data is the relatively high non-response incidence regarding sensitive information.<sup>9</sup> Table 2 presents a summary availability of information based on alternative criteria for data completeness. The minimum eligibility criteria is to restrict the data to establishments with no more than 500 permanent workers, reported total sales last year, and provided information on

<sup>7</sup> The six countries in this analysis were selected for having a large enough presence of unionized (and not unionized) establishments in the sample. Other countries in the region, for example, Ecuador, El Salvador, Nicaragua, Honduras, Paraguay and Peru, have less than 5% of the interviewed unionized establishments.

<sup>8</sup> Details on the implementation and survey structure can be found in the implementation notes for the Latin America Enterprise Surveys Data Set, and the methodological notes found at <http://www.enterprisesurveys.org/Methodology>.

<sup>9</sup> Missing information is coded as follows: the subject's refusal to answer, they did not know the information asked or the question was not applicable for the establishment.

unions.<sup>10</sup> This reduces the sample by 14% (call this the “eligible sample”). The sample is reduced by 40% when one requires establishments to have complete information on sales, union status, capital and production cost.

Even though the analysis could be conducted by removing the incomplete cases, case-wise deletion provides valid inferences only if the share of deleted cases is small or if the data are missing completely at random (MCAR).<sup>11</sup> If this is not the case, the inferences obtained from a complete set might be significantly biased. In this paper, a Multiple Imputation (MI) approach is used to account for the missing information. The imputation process is based on the assumption that all the missing information are “missing at random” (MAR) in the sense of [Little and Rubin \(2002\)](#). This implies that the process that characterizes missingness of the information is entirely explained by the observed information. Under this assumption, the approach uses available information to create multiple independent imputed samples for the unobserved data that can be independently analyzed and the results combined to provide a single MI result.

This procedure advantage over simple imputation approach is that MI introduces new information to the system, by using the empirical distribution of the missing variables. Because this strategy uses all available information, observations with partial missing information are still considered to characterize the missingness and imputation equations. The next section describes the specification and implementation details of the MI strategy used in the paper.<sup>12</sup>

### 3.1. Multiple imputation: Implementation

The working sample is restricted to establishments with complete information on labor, union status and sales variables. Establishments owned by the public sector (more than 50%) are excluded. To reduce the bias caused by new or large establishments, the sample is restricted to those with at least 3 years of operation and at most 500 permanent workers. Finally, in order to avoid biases due to data errors and inconsistencies, some minor edits are implemented.<sup>13</sup> This reduces the working sample from 3,277 to 2,812 enterprises across the 6 countries.

For consistency, imputation models are kept constant across countries, except for the characteristics of region and industry.<sup>14</sup> The imputation model includes variables capturing market competition, establishment structure, infrastructure characteristics, investment, labor force characteristics, and level of unionization.<sup>15</sup> Imputation models are estimated using weights to obtain results representative at the national level. Iterative chained equations (ICE) are used to obtain imputed values given the

---

10 Only 98 observations in the sample correspond to establishments with more than 500 permanent workers. These observations are excluded because there are not enough observations to compare union and nonunion establishments. Some of estimations are sensitive to their inclusion in the sample.

11 In the nomenclature of [Little and Rubin \(2002\)](#), data are missing completely at random if the probability of being missing does not depend on any observed or unobserved data. A weaker condition is missing at random (MAR) or ignorable non-response, which means that the distributions for missing and non-missing observations are similar after conditioning on measurable covariates.

12 Details on the MI process are explained in appendix A

13 In some instances, information such as wages, sales or costs are either too high or too low, to be consistent with other information within the establishment and other similar establishments (typos). Depending on each case, the values were inflated or deflated (reducing the excess of zeroes), or change the value to missing data.

14 The regions with major economic activity are selected for interviews in each region. The industry fixed effects correspond to the ISIC codes 15-37 (ISIC Rev.3.1).

15 A complete list of the variables that are used in the imputation process can be found in the [appendix C](#).

observed data.<sup>16</sup> While one cannot rule out the possibility that part of the information in the dataset is “missing not at random” (MNAR), i.e. missing is a function of unobserved characteristics, [Graham, et al. \(1997\)](#) shows that the sensitivity to unobserved factors is frequently small, and that even under such circumstances, the MI approach provides better inferences than working with complete reported data.

**Table 3. Multiple Imputation Summary**

Variable	Method	Complete	Imputed	% Imputed	Total
Nr of workers in t-1	PMM	2623	189	6.7%	2812
Cost of labor as % of sales	PMM	2563	249	8.9%	2812
Cost of electricity as % of sales	PMM	2572	240	8.5%	2812
Cost of communications as % of sales	PMM	2570	242	8.6%	2812
Cost of materials and inputs as % of sales	PMM	2479	333	11.8%	2812
Cost of fuel as % of sales	PMM	2441	371	13.2%	2812
Cost of transportation as % of sales	PMM	2460	352	12.5%	2812
Cost of water as % of sales	PMM	2408	404	14.4%	2812
Cost of rentals as % of sales	PMM	2453	359	12.8%	2812
Log Nr of workers in t-1	OLS	2623	189	6.7%	2812
Log sales in t-1	OLS	2288	524	18.6%	2812
Log wages production workers	OLS	2721	91	3.2%	2812
Log wages non production workers	OLS	2589	223	7.9%	2812
Log capital (book value)	OLS	1961	851	30.3%	2812
Log capital (market value)	OLS	2346	466	16.6%	2812
Log materials and Inputs	OLS	2441	371	13.2%	2812
Log salaries	OLS	2574	238	8.5%	2812

Note: the complete set of the variables and imputations are shown in appendix C. OLS imputation uses linear predictions to obtain the imputed values. PMM is a predictive mean matching algorithm that uses the value of the closest observation (using predicted means) to impute missing information.

Table 3 presents a summary of the imputations for the most important variables in the study. As one can observe, information regarding capital, a fundamental variable in the analysis, has one of the largest incidences of missing information (30.3% book value of capital and 16.6% market value). Among production costs, the costs of electricity and communication have the lowest missing rates (8.5% and 8.6%), while costs of fuel and water have the highest rates of missing information (13.2% and 14.4%).

### 3.2. Summary Statistics

[Table 4](#) presents the summary statistics of the combined imputed samples. Most countries’ establishment unionization rates are between 20-30%.<sup>17</sup> The exceptions are Bolivia, with the lowest shares of unionized establishments (13.4% or 7.0% weighted), and Argentina, with more than 90% of establishments being unionized. Unionized establishments are larger (in terms of number of workers) and operate for longer hours per week. With the exception of Chile, unionized establishments show a more intensive use of their installed capital. Both hours of operation and use of capital are positively correlated to the capital intensity.

<sup>16</sup> While the literature recommends 5-10 imputed samples to obtain appropriate inferences ([Rubin 1987](#)), 50 imputed samples are obtained for the analysis in order to obtain stable results ([Horton and Lipsitz 2001](#)). Following the literature, 20 iterations are used for the burn-in period to achieve convergence on the system ([van Buuren 2007](#)).

<sup>17</sup> An establishment is classified as unionized if any share of their workforce is considered to be part of a union.

Except for Argentina, unionized establishments have higher levels of sales per capita than their counterparts, and, with the exception of Argentina and Mexico, unionized establishments have higher levels of capital intensity.<sup>18</sup> Unionized establishments are on average older.<sup>19</sup> With the exception of Argentina, companies owned by foreign capital are more likely to be unionized, but establishments with a single majority shareholder are less likely to be unionized. In all countries but Panama, unionized establishments are more likely to have some type of certification for production quality. Similarly, unionized establishments invest more in physical capital, and research and development. In terms of workforce characteristics, unionized establishments are more likely to have an ongoing training program, with larger shares of the workforce trained. Apart from Chile, unionized establishments have larger shares of production workers.

**Table 4. Summary Statistics**

	Argentina		Bolivia		Chile	
	Nonunion	Union	Nonunion	Union	Nonunion	Union
Nr of establishments	46	494	258	40	415	149
Share	8.5	91.5	86.6	13.4	73.6	26.4
Share weighted	11.0	89.0	93.0	7.0	71.9	28.1
Union density (% unionized)		68.50%		5.80%		13.00%
Log sales per capita	10.58	10.46	8.79	9.66	10.41	10.77
Nr of equivalent permanent workers	22.73	58.27	28.07	143.03	40.69	115.24
Log capital (market value) per worker	9.13	8.99	7.75	8.83	8.84	9.22
Cost of labor as share of sales	19.1%	24.3%	23.1%	19.9%	24.7%	21.9%
Cost of inputs as share of sales	37.3%	42.7%	39.0%	37.1%	42.5%	42.1%
% Level of utilization of facilities	66.6%	70.2%	62.5%	63.7%	72.3%	68.5%
Avg hrs. of operation per Week	70.03	62.33	59.12	82.05	64.23	81.51
%Sales coming from manufacture	93.5%	93.4%	97.4%	98.7%	95.1%	98.1%
%Sales subcontracted	17.4%	8.7%	12.1%	13.5%	6.9%	4.9%
Age of establishment	25.01	35.19	21.76	29.60	25.94	41.30
Exp. top manager	30.39	27.65	21.15	19.73	26.53	24.14
Owned by foreign capital	5.5%	5.0%	3.9%	28.7%	2.1%	7.1%
>50% own by largest shareholder	93.8%	75.2%	84.8%	65.2%	90.1%	80.6%
Quality certification	16.5%	24.8%	8.6%	38.4%	19.8%	37.1%
New production or process	79.4%	80.1%	83.7%	93.7%	78.3%	74.9%
Investment in R&D or capital	75.2%	75.4%	62.0%	79.4%	77.9%	82.6%
% with no training program	72.4%	47.3%	42.2%	30.8%	60.7%	41.4%
% with 1-33% trained wf	0.9%	13.5%	20.6%	11.1%	21.0%	20.4%
% with 34-66% trained wf	0.2%	8.6%	17.0%	22.3%	8.2%	17.8%
% with 67-100% trained wf	26.5%	30.5%	20.1%	35.7%	10.1%	20.5%
Share prod Workers	63.2%	73.7%	65.0%	72.3%	70.7%	65.8%
Share skill Workers	53.6%	55.7%	68.6%	61.2%	54.7%	64.1%
% with 7+ yrs avg worker education	99.7%	97.6%	81.6%	70.0%	98.5%	94.0%
Share temporary workers <sup>a</sup>	10.5%	5.6%	28.3%	14.9%	9.0%	5.8%

Note: The averages are calculated using survey weights and all imputed data.<sup>a</sup> The Share of temporary workers is defined as number of total temporary workers divided by total number of permanent and temporary workers.

<sup>18</sup> Per capita measures are calculated dividing the variables of interest by the total number of permanent workers plus equivalent temporary workers in the establishment. Capital per capita is calculated using the hypothetical or "market value" of capital. It is the establishment estimation of fair value of their machinery.

<sup>19</sup> There is no information available on when unions formed within the establishment.



**Table 4. Summary Statistics (Continued)**

	Mexico		Panama		Uruguay	
	Nonunion	Union	Nonunion	Union	Nonunion	Union
Nr of Establishments	639	335	148	37	181	70
Share	65.6	34.4	80.0	20.0	72.1	27.9
Share weighted	72.6	27.4	78.9	21.1	79.4	20.6
Union Density (% unionized)		21.00%		12.30%		10.80%
Log sales per capita	9.69	9.95	10.15	10.53	9.94	10.30
Nr of equivalent permanent workers	24.72	69.59	32.94	80.79	21.90	45.88
Log capital (market value) per worker	8.31	8.27	8.90	10.36	8.64	8.91
Cost of labor as share of sales	25.6%	24.5%	20.8%	23.5%	21.3%	19.0%
Cost of inputs as share of sales	26.9%	26.7%	34.3%	39.9%	47.4%	46.4%
% Level of utilization of facilities	73.9%	73.8%	71.8%	79.6%	65.9%	64.9%
Avg hrs. of operation per Week	60.69	67.45	55.81	67.79	70.35	88.36
%Sales coming from manufacture	96.4%	99.0%	93.4%	91.0%	96.8%	96.5%
%Sales subcontracted	8.5%	13.2%	5.6%	10.1%	9.4%	8.0%
Age of establishment	17.61	22.71	23.30	35.52	27.41	33.70
Exp. top manager	16.90	19.02	22.08	26.14	25.29	26.70
Owned by foreign capital	1.6%	6.6%	8.0%	9.2%	2.7%	13.1%
>50% own by largest shareholder	84.5%	73.2%	89.2%	76.4%	82.7%	64.6%
Quality certification	9.8%	25.9%	11.2%	7.7%	6.6%	14.5%
New production or process	26.3%	57.0%	73.8%	65.9%	70.9%	82.0%
Investment in R&D or capital	22.6%	48.9%	63.9%	78.8%	56.0%	64.6%
% with no training program	87.2%	49.5%	58.3%	38.9%	76.8%	47.5%
% with 1-33% trained wf	1.9%	5.1%	15.2%	14.6%	10.1%	24.7%
% with 34-66% trained wf	4.4%	14.3%	10.7%	23.7%	4.9%	6.7%
% with 67-100% trained wf	6.5%	31.2%	15.7%	22.8%	8.3%	21.1%
Share prod Workers	72.8%	74.4%	66.2%	71.1%	72.8%	74.0%
Share skill Workers	85.3%	78.1%	74.4%	78.7%	56.4%	57.7%
% with 7+ yrs avg worker education	13.3%	10.6%	94.8%	93.7%	64.9%	58.2%
Share temporary workers <sup>a</sup>	4.1%	6.0%	16.6%	18.9%	7.8%	7.0%

Note: The averages are calculated using survey weights and all imputed data.<sup>a</sup> The share of temporary workers is defined as number of total temporary workers divided by total number of permanent and temporary workers.

#### 4. Econometric Strategy

To determine the effects of unions on productivity, the starting point is the model developed by [Brown and Medoff \(1978\)](#). This is a variant of a Cobb-Douglas production function that distinguishes between two types of workers (nonunion and union). Assuming constant returns to scale, the production function can be written as follows:

$$Q_i = A_i K_i^a (L_{in} + cL_{iu})^{1-a} \quad (1)$$

where  $Q$  is a measure of output or productivity,  $K$  is the level of capital,  $L_n$  and  $L_u$  are nonunion and union workers respectively, all measured at the establishment level  $i$ ;  $A$  is the constant of proportionality that depends on the measurement units of capital, labor and output, and accounts for other characteristics that determine productivity; and “ $a$ ” and “ $1-a$ ” are the output elasticities with respect to capital and labor, assuming constant returns to scale. In this framework, “ $c$ ” reflects the productivity differences between union and nonunion labor. After some manipulation, [equation \(1\)](#) can be linearized and written as:

$$\log q_i = \log A + a * \log k_i + \delta * P_i \quad (2)$$

where  $q = \frac{Q}{L}$  and  $k = \frac{K}{L}$  are measures of labor productivity and capital per capita,  $P = \frac{L_u}{L_u + L_{nu}}$  is the share of unionized workers in the establishment, and  $L$  is total number of workers in the establishment. Here  $\delta$  represents the overall impact that unions have on establishment labor productivity, once we control for capital intensity.

Because [equation \(2\)](#) is rather restrictive a more flexible specification is used, following a general form of a translog specification for the production function ([Christensen, Jorgenson, and Lau, 1973](#)). After including an error term, and additional controls for productivity, the specification to be estimated can be written as follows:

$$\log q = a_0 + a_1 * \log k + a_2 * \log L + a_3 \log K^2 + a_4 \log L^2 + a_5 \log K \log L + \delta P + X' \beta + e \quad (3)$$

Although similar specifications have been widely used in the literature, there are limitations that need to be discussed (for details on the discussion see [Brown and Medoff 1978](#) and [Hirsch and Addison 1986](#)). First, while physical production is preferable, in its absence, this paper uses value added.<sup>20</sup> The potential problem is that this measure might confound effects of both prices and quantity changes, as firms might shift some of the increasing cost (wages) to the consumers. Following [Hirsch \(2004\)](#), this problem is mitigated by controlling for industry fixed effects and measures of market competition.

A second problem is that the specification assumes union and nonunion establishments share the same production function, except for the productivity parameter associated to unions. Although this problem could be alleviated by introducing different set of interactions, it may require more information than what is available. In addition, the flexibility obtained using a translog production function ([equation 3](#)) should help reduce the severity of the problem.

The most vexing problem is the potential endogeneity of establishment unionization. According to [Clark \(1984\)](#), one might not expect unionization and sales to be simultaneously determined, since unionization should have happened long before the survey interview. Concerns with respect to the

<sup>20</sup> Value added is defined as annual sales minus production costs on materials, electricity and water, divided by total labor force.

inter-temporal effects of unionization remain. If union negatively affects profits, businesses will be less likely to survive. This survivor bias should lead to overstate union productivity effects since businesses with detrimental union effects on performance are least likely to remain in the sample (Addison and Hirsch 1989). Further, omitted variables in the specification can also generate inconsistent estimates if the unobserved characteristics have systematic variations between union and nonunion establishments.

Taking advantage of the rich information contained in the survey, the preferred specifications controls for different characteristics such as age of the establishment, manager experience and ownership characteristics, among others, that provides a flexible specification, reducing the impact of otherwise unmeasured characteristics. Nevertheless, because of potential survivor bias and the presence of additional unobserved factors, the estimations could be upward biased, should be considered with care, and should not be interpreted as causal effects. To test the sensitivity of the productivity estimates to the controls, different specifications are used to control for aspects related to market competition, establishment characteristics and organization, and innovation policies. Though the estimates here presented are “descriptive” in nature, the evidence is informative, as they are the first step toward identifying how unions affect productivity in developing countries.

## 5. Results

Following the specification shown in equation (3), the natural logarithm of value added per worker is used as the productivity measure. For observations where production costs exceed the value of total sales, total cost is constrained and the cases controlled using a dummy variable.<sup>21</sup> For the production factors, employment is measured as the total number of permanent workers plus the equivalent number of seasonal workers, while capital is measured as log of the market value of machinery and equipment (including vehicles).<sup>22</sup> In addition, the basic model includes controls for region and broad industry fixed effects. The main variable of interest, union density, is included as a share between 0 and 1, which indicates what share of the permanent labor force in the establishment is unionized.

**Table 5. Effect of Unions on Establishment Productivity, by Country**

	Argentina	Bolivia	Chile	Mexico	Panama	Uruguay
Avg. % unionization in union establishments	77.0%	82.5%	46.2%	76.8%	58.2%	52.2%
%Workforce unionized	-0.389+ (0.026)	0.160 (0.695)	0.167 (0.353)	0.095 (0.555)	0.256 (0.467)	0.178 (0.448)
Log capital pc	-0.777 (0.306)	0.316 (0.421)	-0.45 (0.194)	-0.086 (0.650)	-0.248 (0.490)	-0.114 (0.717)
Log total labor force	-0.427 (0.425)	-0.792 (0.150)	-0.443 (0.357)	0.19 (0.548)	0.045 (0.951)	0.522 (0.549)
Log K log L	-0.012 (0.878)	0.161* (0.004)	-0.005 (0.914)	-0.015 (0.595)	-0.001 (0.986)	-0.034 (0.520)
Log K <sup>2</sup>	0.033 (0.380)	-0.027 (0.250)	0.027 (0.122)	0.017 (0.115)	0.012 (0.546)	0.008 (0.606)

<sup>21</sup> Overall, only 1.3% of the observations fall within these characteristics.

<sup>22</sup> Equivalent seasonal workers are measured as the total number of temporary workers multiplied by the average time a temporal worker participates in the establishment in a year. As described in the survey manual, information collected on the market value (hypothetical value) of capital is recommended to be used as the best approximation for capital intensity in the establishment.

	Argentina	Bolivia	Chile	Mexico	Panama	Uruguay
Log L <sup>2</sup>	-0.001 (0.983)	-0.124 (0.162)	0.000 (0.998)	-0.031 (0.395)	-0.031 (0.646)	0.029 (0.813)
Constant	13.648* (0.001)	6.917* (0.000)	11.143* (0.000)	7.987* (0.000)	10.390* (0.000)	8.291* (0.001)
$\delta V A_{pc} / \delta K$ , at means	-0.006	0.207	0.203	0.250	0.043	-0.026
$\delta V A_{pc} / \delta L$ , at means	0.196	-0.104	-0.055	-0.068	0.093	0.409
Observations	540	298	564	974	185	251

Note: ^ p<0.1, + p<0.05, \* p<0.01. P-values in parentheses. All models are estimated using sample weights. All models include region and broad industry fixed effects.

Table 5 presents the results of the basic specification which controls only for production factors, and region and industry fixed effects. These basic results show that for most countries in our sample, unions have a positive but weak correlation with productivity. In the absence of endogeneity, the point estimates indicate that, on average, if a nonunion establishment unionizes it could increase productivity per worker between 0.077 log points (approximately 8%) to up to 0.149 log points (15%).<sup>23</sup> Although the point estimates are sizable, the results also indicate that there is considerably heterogeneity across establishments, shown by the low significance levels of the parameters. The only exception is Argentina; the estimate shows that unions have a large negative and significant impact on productivity (-0.389\*77%=0.299 log points). It should be kept in mind that in contrast to other countries in the sample, most of Argentina's manufacture establishments are unionized.

Given the type of production function used here, the coefficients for capital and labor are more difficult to interpret than in the standard Brown and Medoff model. To facilitate interpretation, at the bottom of Table 5, the derivatives with respect to labor and capital are obtained and evaluated at the mean.<sup>24</sup> The parameters are consistent with the expectations for Bolivia, Chile and Mexico. In Argentina, Panama and Uruguay, however, the estimates are somewhat unexpected as the marginal effect of capital is almost zero, with a positive marginal effect from labor. Although these results are worrisome, they remain consistent across different specification, and should not to be crucial for the main focus of the analysis.

### 5.1. Sensitivity to Additional Controls

The basic model estimated in Table 5 does not take into account other characteristics that can explain productivity or that can differ between unionized or nonunionized establishments. Tables 6 and 7 present estimations using richer specifications. Table 6 presents results of the union productivity effect only, to show how sensitive the estimates are to additional controls, while Table 7 presents the results of the full specification model.

Controlling for the level of competition should improve the estimates by partially accounting for differences in prices caused by union effects on labor costs. The second row of in Table 6 provides the estimates after controlling the number of competitors the in the market. In Chile and Uruguay, the productivity estimates are greatly reduced, with the estimates for Uruguay becoming negative.

23 The average marginal effect is obtained by multiplying the union coefficient estimate by the average unionization rate among union establishments (i.e. Marginal effect in Bolivia: 82.5%\*0.16=0.132 log points).

24 The estimates of the basic specification using the basic Brown and Medoff (1978) model are shown in appendix C. The results are comparable to the estimations of Table 5, except for Bolivia, where the Brown and Medoff basic model predicts a much larger productivity relationship.

In Chile, these results seem to be driven by the impact that some competition (2-5 competitors in the market) has on sales and productivity (see [Table 7](#)). In Uruguay, while partial model (not shown here) suggests a story similar to Chile, after controlling for other factors, competition seems to have a small and not significant impact on productivity.

An implicit assumption in the model is that all sales (production) come from the establishment's own manufactured goods produced using full capacity of the available resources. Longer hours of operation or greater use of production capacity can spuriously increase productivity estimates if they are not controlled for. Likewise, if establishments subcontract production to smaller units, or generate sales via services (rather than manufacture), it may bias the productivity estimates. Results in [Table 7](#) show the results controlling for level of utilization of capital, weekly hours of operation, and sales structure (services and subcontracts). Controlling for these productivity adjustments suggests that union productivity enhancements are not coming from more intensive use of production factors, showing larger estimates increase for all countries. Chile and Uruguay show the most drastic changes on productivity, with modest changes elsewhere. The results in [Table 7](#) show that the parameters for these controls are consistent with the expected signs.

Other factors correlated with union status and productivity are linked to establishment characteristics and management quality. To account for establishment characteristics, variables controlling for years of operation, whether the establishment is part of a larger firm, and if it the establishment is owned by foreign capital are included in the specification. Controlling for these factors reduces the productivity relationship in Bolivia, and increases it in Chile to almost significant levels ( $p=12.5\%$ ). According to the results in [Table 7](#), the fall in the union-productivity relationship in Bolivia is explained because most of the positive relationship was driven by foreign owned companies, which are typically unionized. This is consistent with the fact that being part of a larger firm and being owned by foreign capital is correlated with better managerial policies ([Aitken, et al. 1996](#)). Alternatively, in Chile, not accounting for younger establishments, which are more productive and less likely to be unionized, was understating union productivity effects.

To control for management quality and organization, variables describing the ownership structure and management characteristics are included in the specification. Argentina and Panama show the largest positive change in the estimates of the union-productivity relationship. These results suggest that unionized establishments have relatively more inefficient management, which puts downward pressure on productivity if management characteristics are not accounted for. The results in [Table 7](#) suggest that the strong relationship between productivity and decision strength of the largest shareholder is the main factor explaining the impact on the union-productivity estimates in Panama.

Although some of the previous controls are arguably not directly affected by unions, aspects such as investment, training, and workforce structure could be affected their presence. In the interest of disentangling the direct effect unions on productivity, additional controls are introduced in the specification. Under the assumption that unions have a negative impact on investment (rent seeking behavior), controlling for investment should have a positive impact on the union-productivity relationship. The results on [Table 6](#), however, indicate that controlling for investment and innovation have little impact on the union estimates.

Because unions are often associated with lower turnover, unionized establishments might be more likely to provide more training, since they can benefit from the returns of such investment

through higher productivity (Kuhn 1985; Allen 1984). Although human capital upgrades are legitimate sources of productivity enhancements, they may not necessary generate improvements in technical efficiency. Including training as a control variable in the estimations see Table 7) shows that, on average, training has a positive, mostly not significant, impact on productivity. Adding these controls to the specification has the expected negative effect on the union-productivity relationship for all countries, especially for Panama, indicating that training is an important channel through which unions improve productivity.

Unions can also influence the structure and composition of the workforce within establishment, changing the mix of production factors and affecting productivity. The last estimates in Table 6 control for two aspects of workforce characteristics: workers' average education and workforce structure. Except for Chile and Panama, controlling for these effects increases the impact of unions on productivity. In Bolivia, Mexico and Uruguay, similar to the evidence at the micro level (Rios-Avila and Hirsch, forthcoming), unionized establishments are associated with low-educated, low-skill workers, and not controlling for it understates productivity. The evidence, however, is insufficient to detect the effect of other workforce characteristics on union-productivity effect.

Although there is limited evidence on the productivity effects of unions for these countries in the literature, the evidence that does exist appears consistent with the results found here. Using information from a national survey of manufacturing in Mexico in 1999, Fairris (2006) finds that unionized establishments are about 11% more productive than their counterparts (compared to the 9% estimate here). In the case of Uruguay, Cassoni, et al. (2005) finds a modest 5.7% effect, qualitatively similar to the 12% (0.12 log points) found here. These similarities increase the confidence in the results shown in Table 6.

**Table 6. Effect of Unions on Establishment (Est.) Productivity, Sensitivity to Specifications**

	Argentina	Bolivia	Chile	Mexico	Panama	Uruguay
%Workforce unionized	-0.389+	0.160	0.167	0.095	0.256	0.178
Basic model	(0.026)	(0.695)	(0.353)	(0.555)	(0.467)	(0.448)
%Workforce unionized	-0.388^	0.190	0.038	0.114	0.286	-0.083
+Competition	(0.054)	(0.662)	(0.832)	(0.447)	(0.401)	(0.763)
%Workforce unionized	-0.328+	0.232	0.156	0.145	0.280	0.117
+Productivity adjustments	(0.028)	(0.609)	(0.316)	(0.382)	(0.419)	(0.578)
%Workforce unionized	-0.322+	-0.013	0.251	0.101	0.305	0.072
+Est. characteristics and ownership	(0.034)	(0.976)	(0.125)	(0.535)	(0.393)	(0.755)
%Workforce unionized	-0.264^	-0.034	0.219	0.073	0.407	0.114
+ Management and organization	(0.058)	(0.930)	(0.170)	(0.620)	(0.242)	(0.570)
%Workforce unionized	-0.273^	-0.05	0.254	0.049	0.461	0.092
+Investment policy and technology	(0.090)	(0.891)	(0.117)	(0.732)	(0.178)	(0.683)
%Workforce unionized	-0.279^	-0.101	0.251	0.026	0.378	0.073
+Training	(0.097)	(0.797)	(0.118)	(0.863)	(0.272)	(0.787)
%Workforce unionized	-0.257+	-0.009	0.173	0.093	0.349	0.117
+Labor force structure	(0.049)	(0.981)	(0.281)	(0.539)	(0.295)	(0.558)
Observations	540	298	564	974	185	251

Note: ^ p<0.1, + p<0.05, \* p<0.01. P-values in parentheses. All models are calculated using all controls specified in the previous model. All models are estimated using the sample weights, and include region and broad industry fixed effects.

**Table 7. Effect of Unions on Establishment (Est.) Productivity, Full Specification**

	Argentina	Bolivia	Chile	Mexico	Panama	Uruguay
%Workforce unionized	-0.257+ (0.049)	-0.009 (0.981)	0.173 (0.281)	0.093 (0.539)	0.349 (0.295)	0.117 (0.558)
<b>Competition</b>						
Establishment has 2-5 competitors	0.044 (0.850)	-0.342 (0.331)	0.443+ (0.013)	-0.359* (0.002)	-0.009 (0.970)	-0.425 (0.268)
Establishment has 5 or more competitors	-0.105 (0.718)	-0.209 (0.509)	0.154 (0.323)	-0.287+ (0.016)	0.138 (0.645)	-0.368 (0.217)
Establishment faces international market	-0.047 (0.822)	-0.206 (0.510)	0.029 (0.908)	0.341 (0.261)	0.127 (0.765)	0.046 (0.876)
<b>Capacity and Utilization</b>						
Level of utilization of capital	0.008* (0.004)	0.006^ (0.060)	0.002 (0.236)	0.002 (0.210)	-0.001 (0.890)	0.016* (0.003)
Log hours of operation per week	-0.245 (0.230)	0.148 (0.454)	-0.128 (0.165)	0.062 (0.643)	0.152 (0.569)	-0.239^ (0.086)
%Sales not from manufacture	-0.014 (0.972)	1.203 (0.288)	0.927+ (0.041)	0.796 (0.240)	1.326^ (0.058)	-1.012 (0.418)
%Sales that are subcontracted	-0.246 (0.224)	0.275 (0.408)	0.190 (0.357)	0.116 (0.619)	0.206 (0.614)	-0.430 (0.201)
<b>Est. Characteristics</b>						
Owned by foreign capital (>50%)	0.227 (0.472)	0.727+ (0.013)	0.239 (0.274)	-0.117 (0.540)	0.335 (0.280)	0.182 (0.734)
Establishment part of larger firm	0.150 (0.343)	-0.311 (0.307)	0.073 (0.541)	0.084 (0.440)	0.107 (0.750)	0.220 (0.620)
Age of the establishment (Years of operation)	0.009 (0.236)	0.028 (0.107)	-0.010 (0.111)	0.027* (0.000)	0.002 (0.896)	0.006 (0.524)
Age2/100	-0.006 (0.101)	-0.009 (0.587)	0.007 (0.173)	-0.029* (0.000)	-0.007 (0.711)	-0.0004 (0.960)
<b>Management and Organization</b>						
>50% own by largest shareholder	0.128 (0.439)	-0.234 (0.378)	0.031 (0.751)	-0.036 (0.653)	0.541^ (0.051)	-0.034 (0.845)
Any of the main owners female	-0.043 (0.683)	-0.331^ (0.066)	0.052 (0.615)	-0.214* (0.004)	0.366^ (0.071)	-0.264 (0.223)
Experience top manager	0.018 (0.254)	0.028^ (0.069)	0.007 (0.456)	0.000 (0.982)	-0.039 (0.253)	-0.009 (0.780)
Experience2/100	-0.042+ (0.013)	-0.097* (0.010)	-0.023 (0.170)	-0.005 (0.760)	0.076 (0.246)	-0.002 (0.965)
Public or private	-0.057 (0.686)	0.303 (0.280)	0.150 (0.160)	0.173^ (0.080)	0.283 (0.263)	0.311 (0.138)
<b>Investment and Innovation</b>						
Uses foreign company technology	0.312+ (0.044)	-0.064 (0.829)	0.162 (0.255)	0.180 (0.308)	0.422 (0.233)	0.005 (0.989)
Product quality certification	0.027 (0.877)	0.165 (0.578)	0.150 (0.114)	0.356* (0.001)	0.514 (0.110)	0.567+ (0.022)
Introduced new process or product	0.366+ (0.018)	-0.409+ (0.028)	0.211+ (0.036)	0.055 (0.510)	0.130 (0.545)	0.137 (0.436)

	Argentina	Bolivia	Chile	Mexico	Panama	Uruguay
Invested in capital or R&D	0.144 (0.216)	0.053 (0.755)	-0.027 (0.821)	0.156 <sup>^</sup> (0.088)	-0.003 (0.989)	-0.038 (0.811)
<b>Training</b>						
1-33% workforce trained	-0.006 (0.982)	-0.275 (0.308)	0.070 (0.519)	0.137 (0.471)	0.472 (0.123)	0.201 (0.472)
34-66% workforce trained	0.588* (0.001)	-0.249 (0.284)	0.191 (0.293)	-0.008 (0.960)	0.043 (0.873)	0.322 (0.409)
67-100% workforce trained	0.200 (0.127)	-0.219 (0.109)	0.121 (0.375)	0.067 (0.597)	-0.174 (0.543)	-0.280 (0.143)
<b>LF characteristics</b>						
Avg education 4-6 yrs		0.174 (0.631)		0.167 <sup>^</sup> (0.094)		
Avg education 7-12 yrs		0.126 (0.614)		0.365+ (0.045)		0.341 (0.169)
Avg education 13+ yrs	-0.088 (0.385)	0.006 (0.984)	0.055 (0.689)		-0.271 (0.199)	
Share of production workers	-0.882+ (0.016)	-0.448 (0.161)	-0.179 (0.632)	-0.291 (0.261)	0.377 (0.503)	-1.518* (0.000)
Share of skill workers	-0.088 (0.461)	-0.148 (0.536)	0.098 (0.449)	0.340+ (0.022)	-0.146 (0.675)	-0.118 (0.604)
Share of temporary workers	0.868 (0.419)	-0.843+ (0.017)	-0.499 (0.335)	-1.240 (0.119)	-0.257 (0.826)	-0.831 (0.477)
Share of female workers	-0.670+ (0.037)	-0.601+ (0.020)	-0.483+ (0.035)	0.182 (0.320)	0.090 (0.796)	-0.313 (0.569)
Constant	13.614* (0.000)	7.422* (0.000)	11.527* (0.000)	7.361* (0.000)	7.495* (0.001)	10.519* (0.000)
Observations	540	298	564	974	185	251

Note: <sup>^</sup> p<0.1, + p<0.05, \* p<0.01. P-values in parentheses. The base categories for competition are those establishments facing none or 1 other establishment in the market. For training, the base category corresponds to establishments with no training programs. For education, the base category corresponds to workers with 3 or fewer years of education for Bolivia and Mexico, 7 to 12 years for Argentina, Chile and Panama, and 4-6 years of education for Uruguay. All models are estimated using sample weights, and include region and broad industry fixed effects.

## 5.2. Interpretation

Taken as a whole, the results shown in [Table 6](#) suggest that unions are associated with positive union-productivity effects, but that there is a lot of heterogeneity in the relationship both within and across countries (the former seen by the low significance levels of results). The results also provide some evidence that the union-productivity estimates are reasonably robust in richer specifications that take into account typically unobserved establishment characteristics. On one end of the spectrum, using the preferred specification (all controls), it can be observed that unions in Bolivia have no effect on productivity (practically zero), while a negative and significant estimate is seen in Argentina (-0.284). On the other side, large positive, albeit insignificant, effects are observed in Chile and Panama, while estimates for Mexico and Uruguay are positive and consistent across specifications, but not significant.

The evidence presented in [Table 6](#) also brings some light on determinants of the union-productivity relationship. Controlling for productivity adjustments increases union productivity estimates.



This suggests unions might be improving technical efficiency, adjusting for differences in sales and production structures across establishments, consistent with the reduction of “X-inefficiencies” (Hirsch and Addison 1986; Addison and Hirsch 1989; Kaufman, 2004). Part of the positive productivity effects, however, seem to be related to large unionization rates across large establishments, foreign owned, with typically better management. Consistent with the hypothesis in Fairris (2006), unions seem to be improving productivity by improving job training. Finally, lower levels of worker human capital seem to have contributed to an understatement of the productivity effects of unions.

Taken at face value, the estimate found for Argentina indicates that output per worker in a fully unionized establishment is 25 log points lower than in an equivalent nonunion establishment. Were we to have obtained such a result in other countries, it would have raised the question how union establishments could survive given lower productivity and (presumably) higher compensation. Argentina, however, is a special case compared to other countries in the region. There are very few establishments in the sample that are not unionized, and those nonunionized are rather different (relatively younger and smaller).

As Kuhn (1998) and others have argued, negative productivity effects have been typically found in environments of union and management conflict, which might explain the results observed in Argentina. In Table 8, two aspects of labor regulations and the perception of management are presented. About 44% of establishments in Argentina consider that labor regulations are serious or very serious obstacles to the operation of the establishment. Similarly, 60% of establishments declare that labor regulations have affected their hiring and firing decisions, which is almost twice as high as in other countries.<sup>25</sup>

**Table 8. Perception of Labor Regulations**

	Argentina	Bolivia	Chile	Mexico	Panama	Uruguay
Labor regulations and H/F decisions						
Hire decision affected	8.4	2.2	5.4	0.7	4.3	13.6
Fire decision affected	7.2	3.3	12.6	1.9	6.9	2.1
Hire and Fire affected	44.5	9.3	19.4	3.5	6.1	17.3
Labor regulation as obstacles						
No obstacle	8.1	52.3	29.8	49.6	56.8	38.7
Minor obstacle	12.6	13.8	24.0	25.4	15.9	9.1
Moderate obstacle	34.4	20.2	26.2	19.1	18.8	30.4
Serious obstacle	26.9	7.5	15.6	3.8	7.5	14.0
Very serious obstacle	17.9	3.7	4.4	0.7	0.7	7.0

Note: All information reflects the weighted share of establishments within each category for each country.

To assess the extent that perceptions of labor regulations affect the union-productivity relationship, Table 9 presents two alternative specifications that show the interaction between unionization rates and the perception of regulations. The direction of the estimates suggests unions have a smaller (or larger and negative) effect on productivity in cases of negative perception of labor regulation. For instance, except for Mexico or Uruguay, the union-productivity effect is smaller (or more negative) if establishments perceive that labor regulations have affected their decisions on hiring or firing workers.

25 For more details on establishment perceived obstacles, taxes, regulations, and other topics, including comparisons to countries in the region can be found in the Country Profiles, and accessed at <http://www.enterprisesurveys.org/Reports>.

Regarding the perception of labor regulations as obstacles to production, the results are similar. Establishments that do not consider labor regulations as obstacles to production, less conflictive environments, show a more positive impact on productivity compared to establishments that consider labor regulations a moderate or serious obstacle to production.

**Table 9. Labor Regulations and effect on Union Productivity effects**

	Argentina	Bolivia	Chile	Mexico	Panama	Uruguay
<b>Labor regulations and H/F decisions</b>						
%Workforce unionized	-0.119	0.15	0.188	0.0862	0.377	-0.206
x No H/F decision affected	(0.414)	(0.736)	(0.399)	(0.578)	(0.288)	(0.533)
%Workforce unionized	-0.304+	-0.638	0.148	0.129	0.237	0.614^
x H/F decision affected	(0.031)	(0.275)	(0.411)	(0.570)	(0.697)	(0.084)
<b>Labor regulation as obstacles</b>						
%Workforce unionized	-0.344	0.216	0.0975	0.0795	0.456	-0.006
x No or minor obstacle	(0.266)	(0.681)	(0.673)	(0.595)	(0.298)	(0.982)
%Workforce unionized	-0.407+	0.051	0.428	0.0437	-0.0505	-0.649
x Moderate obstacle	(0.046)	(0.933)	(0.160)	(0.877)	(0.938)	(0.325)
%Workforce unionized	-0.175	-0.886	-0.112	0.289	0.791	0.669
x Serious or v. serious obstacle	(0.122)	(0.134)	(0.660)	(0.195)	(0.110)	(0.209)
<b>N</b>	540	298	564	974	185	251

Note: ^ p<0.1, + p<0.05, \* p<0.01. P-values in parentheses. All models are estimated using the full specification shown in Table 14.

An alternative explanation is rooted on the settings under which unions operate. [O'Connell \(1999\)](#) describes Argentinian unions to be highly centralized and protected by the government, which provides unions with fewer incentives to internalize the costs of bargaining demands, potentially harming productivity. As noted in [Anner \(2008\)](#), unions in Argentina also have access to financial information of the employers, which can be used during bargaining. Access to this information may allow unions to better internalize the cost of increasing wages, providing incentives to engage contracts that benefit both workers and employers. In an environment of conflict, it is more likely that financial information is used to redirect resources toward higher wages (rent seeking behavior), reducing establishment's flexibility to transfer resources to investment or innovation, reducing productivity. Given the unique situation of Argentina, there is little reason to suspect that unionized establishments will be at a disadvantage in the domestic market, as unionized establishments primarily compete with each other.

The positive, albeit insignificant, union productivity effects seen in Chile can be associated with the decentralized collective bargaining system in this country. As [Campero \(2001\)](#) and [Vergara \(1998\)](#) describe, the decentralized bargaining system in Chile has allowed establishments to negotiate wage and benefits linked to specific productivity targets. This is consistent with [O'Connell's \(1999\)](#) hypothesis, which indicates that decentralized collective bargaining systems can increase productivity by allowing unions to internalize their externalities, facilitating their input in production process decisions. Our evidence shows, however, that once labor force structure characteristics are taken into account, the union-productivity relationship becomes statistically insignificant.

In the remaining countries, unions appear to have a positive but insignificant effect on productivity (Bolivia is an exception). It can be argued that the decentralized bargaining system in Panama, as in Chile, explains its large positive union productivity estimates, but large heterogeneity within Panama makes these estimates insignificant. In Bolivia, Mexico and Uruguay, where union-productivity

estimates are smaller, there are no specific reasons that might lead one to expect a highly positive or negative union effect.

Uruguay has transitioned toward a decentralized bargaining system that tries to incorporate productivity clauses in their contracts (O’Connell 1999; Cassoni, et al. 2005), similar to the Chilean model. In Mexico, the combination of a centralized bargaining system, and conflicts between employers and unions, reflected in the violation of union rights, should have generated a negative productivity impact, which is not observed. In the case of Bolivia, the results show that unions have no net effect on productivity, and positive relationships are mainly driven by establishment characteristics, in particular the type of ownership. Then again, the presence and strength of unions might be too limited (compared to the other countries) to generate substantive pressure for productivity enhancements.

### 5.3. Robustness to Unionization measurements

Union density might contain measurement errors because employers, who may not have perfect information on union membership, report “guestimate” of union density in the establishment. In this case, a categorical union measure (say from 1 to 50; 51 to 100 percent, with zero unionization the omitted base group), may contain less measurement error than a continuous measure of union density does. Table 10 presents estimates using two alternative measures of unionization within establishments, using the same controls as in the full specification (Table 7). Although these results are informative regarding nonlinear effects of unions, the interpretation of some coefficients might be difficult as some countries, like Bolivia and Panama, might not have sufficient observations to identify the corresponding union effects.

**Table 10. Effect of Unions on Establishment Productivity: Alternative Union Measures**

	Argentina	Bolivia	Chile	Mexico	Panama	Uruguay
(1)						
%Workforce unionized	-0.257+	-0.009	0.173	0.093	0.349	0.117
	(0.049)	(0.981)	(0.281)	(0.539)	(0.295)	(0.558)
% unionization	77.0%	82.5%	46.2%	76.8%	58.2%	52.2%
Avg effect on productivity	-0.194	-0.007	0.082	0.071	0.198	0.066
(2)						
Union dummy	-0.347*	0.0331	0.017	0.094	0.138	0.138
	(0.008)	(0.922)	(0.851)	(0.423)	(0.585)	(0.465)
(3)						
Less than 50% unionized	-0.205	0.213	-0.055	0.156	0.010	0.226
	(0.232)	(0.773)	(0.609)	(0.374)	(0.978)	(0.362)
More than 50% unionized	-0.379*	0.003	0.177	0.074	0.290	0.032
	(0.005)	(0.994)	(0.153)	(0.558)	(0.391)	(0.873)

Note: ^ p<0.1, + p<0.05, \* p<0.01. P-values in parentheses. All models are estimated using the full specification shown in Table 14.

In the first part of Table 10, the results from the preferred model are used as reference. It includes information on the average unionization rate within union establishments to estimate the average union effect on productivity. The first alternative measure uses a dummy variable for unionization, and its coefficient can be interpreted as an average union-productivity effect. The estimates are somewhat consistent with the estimated average union impact observed in row 1. In Argentina, the average productivity effects using dummy variables are larger, in absolute value, than in the preferred specification. In Chile, on the other hand, while the previous estimate shows a large and positive

impact on productivity, the estimates using a dummy variable indicate that the effect is practically zero. This shows that there is some heterogeneity (nonlinearity) in the effects of unions across different levels of unionization.

An alternative measure is to estimate the union effects using dummy variables for different levels of unionization density in the establishment. While some cells are difficult to identify due to insufficient observations, this measurement is better capturing non-linear effects of unions. In Argentina, regardless of the level of unionization, the estimates are strongly negative, and increasing for establishments with higher unionization rates. In Mexico and Uruguay, the results suggest that some levels unionization (less than 50%) can have a positive impact on productivity. In contrast, for Chile and Panama, unionization is positively correlated with productivity for establishments with high rates of union density. Among establishments with low unionization rates, however, unions have either a negative or no relationship with productivity.

## 6. Conclusions

This paper has aimed to fill some gaps in the literature by studying the effects that unions have on productivity and performance at the establishment level in selected countries in Latin America. These Latin American countries have been historically known for the strong role unions have played in their political and economic histories. Following periods of dictatorship, debt crisis and economic recovery, however, these countries developed in ways that have produced substantial heterogeneity in their collective bargaining systems and the roles that unions play in their economies.

According to the results presented here, unions appear to have small but positive effects on productivity across all countries in the sample, with the notable exception of Argentina, where a strong negative productivity effect is found, and Bolivia, where no effect is found. The positive relationships between unions and productivity appear to be sufficient to offset higher labor costs, translating in small and insignificant profitability estimates. In addition, with few exceptions, unions seem to have either a null or negative effect for current investment decisions.

While some of the observed effects can be linked to labor regulations, unions and managements conflicts, economic structure, or unionization organizations, no single narrative can readily explain all results presented here. Even at the establishment level, a precisely estimated union- productivity effect of zero is difficult to interpret, being consistent with unions having either no effect or having offsetting positive and negative effects. The empirical analyses presented here, however, provide an important step toward a better understanding of the role of unions in developing countries in Latin America, an area where there has been little prior evidence. Given the nature of the data and the limitations they place on modeling, the results presented here cannot be strictly represented as causal effects. Instead, they represent the best estimates of partial correlations that capture a portrait of the net outcomes resulting from unions and collective bargaining in these Latin American economies. With the development of new data, similar analysis can be extended to different regions and time frames, and can open the opportunity for future research that provides a more detailed analysis of the effects and channels through which unions affect establishment performance.

## APPENDIX A

### Multiple Imputation using Chain Equations

In cases when the missing data structure follows an arbitrary missing pattern and simultaneous imputations of multiple variables are required, two standard imputation approaches can be used: multivariate normal imputation and imputations using chained equations (ICE). While the first approach estimates a model that tries to identify the underlying multivariate distribution, it imposes restrictions on the estimation, assuming the system follows a normal multivariate distribution. The ICE approach, by contrast, lacks a formal theoretical justification, but provides more flexibility in the specification of the imputation models, being consistent with different types of underlying distributional assumptions.

The idea of the ICE approach is to construct univariate imputation models for each variable with missing information, using a fully conditional specification where all variables, other than the one being imputed, are used as independent variables. These conditional models are used to obtain predictions for the missing information, and can be used in subsequent iterations. In cases where the missing data structure follows an arbitrary pattern, an iterative imputation process is needed to account for possible dependence of the estimated parameters to the imputed data. Formally, the procedure can be described as follows.

Let  $X_1, X_2, \dots, X_p$  be a set of variables with missing information (imputed variables), and let  $Z$  be a set of complete predictors. For each imputed variable, it is possible to construct a univariate imputation model, where each model can be a different distribution function (normal, logistic, etc), that best identifies the specific underlying distribution of the variable :

$$X_1^{t+1} = f_1(X_1^t | X_2^t, \dots, X_p^t, Z, \theta_1)$$

$$X_2^{t+1} = f_2(X_2^t | X_1^t, \dots, X_p^t, Z, \theta_2)$$

$$X_p^{t+1} = f_p(X_p^t | X_1^t, X_2^t, \dots, X_{p-1}^t, Z, \theta_p)$$

Once the imputation models are specified for the first iteration, only complete observations are used for each individual model. Based on the imputation models, random draws using the empirical distribution of the imputed variables are obtained and used in the next iteration of the imputation until convergence is obtained. Although there is no specific rule on the number of imputations needed to obtain convergence of the system, the literature suggests that 10 iterations are typically sufficient to achieve convergence ([van Buuren, 2007](#)). However, depending on the complexity of the imputation system, more iterations may be needed. Once convergence is achieved, a random draw of the empirical process is obtained and used to create an imputed sample. This process is repeated for each additional set of imputed samples needed until M different imputed samples are created.

Although the validity of the MI approach relies on the asymptotic properties of the imputation procedure with M approaching infinity, in practice fewer imputations are needed to obtain consistent and stable results. According to [Rubin \(1987\)](#), M=5 imputations should be sufficient to obtain valid inferences for most procedures, but depending on the amount of information missing and the type of analysis required, a larger set of imputed samples could be required.

Once M completed samples are obtained, each of them can be used to obtain M independent estimations for the desired model. Assume the model to be estimated can be written as:

$$y_m = X'_m \beta_m + e_m \text{ for } m = 1, \dots, M$$

such that we have a set of parameters  $\beta_m$  and a variance covariance matrix  $\sum_m$  for each imputed sample m. Following Rubin's rule (Rubin, 1987) the results for the parameters and variance covariance matrices can be combined as follows:

$$\tilde{\beta}_M = \frac{1}{M} \sum \tilde{\beta}_m, \text{ and}$$

$$\tilde{\Sigma}_M = \frac{1}{M} \sum \tilde{\Sigma}_m + \left(1 + \frac{1}{M}\right) * \sum \frac{(\tilde{\beta}_m - \tilde{\beta}_M)(\tilde{\beta}_m - \tilde{\beta}_M)'}{M-1}$$

Here,  $\tilde{\beta}_M$  and  $\tilde{\Sigma}_M$  are the parameters and variance covariance matrix corresponding to the combination of models across the M imputed samples. See Rubin (1987) for more details.

## APPENDIX B

### Specification of Imputation Model

**Table B1. Variables with Complete Information:**

Variable	Definition
Part larger	Indicates if the establishment is part of a larger firm
Public or private shareholding	Indicates if the establishment has stocks in private hands or public stock exchange.
Foreign owned	Indicates if more than 50% of the establishment is owned by foreign capital
% largest shareholder	Indicates if the largest shareholder owns more than 50% of the establishment
Age	Number of years since the establishment began operations
Manager experience	Number of years of experience of top manager
Quality certification	Indicates if establishments have an ISO quality certification
Electric problems	Indicates if establishments have suffered 2 or more outages
Has a generator	Indicates if establishments possess a generator
Electricity request	Indicates if establishments have submitted a request for electricity connection
Water request	Indicates if establishments have submitted a request for water connection
Water obstacle	Indicates if establishments consider access to water as a major obstacle of production
Electricity obstacle	Indicates if establishments consider access to electricity as a major obstacle of production
Mono production	Indicates if all production comes from the main product
Sales export	% of sales that come from export
Inputs from small establishments	%Inputs bought from smaller firms
Foreign input	%Inputs imported
Principal buyer	Indicates if consumers are main buyers from production
Customs and trades	Indicates if Customs and trades regulations are an obstacle for operations
Own transport	Indicates if establishment possess its own transportation system
Transport problem	Indicates if transportation is considered a major obstacle for operation
Subcontract production	%Sales that are subcontracted to other firms
Competition	Indicates the level of competition the establishment faces: None or one competitor (no competition), 2-5 competitors (medium competition), 5 or more competitors (high competition), operates on international market

Variable	Definition
Sales change	Indicates if sales of main product have gone up or down in the last year
Prices change	Indicates if prices of main product have gone up or down in the last years
Domestic Competition	Pressure from domestic competitors on production costs is important
International Competition	Pressure from international competitors on production costs is important
Foreign Technology	Establishment uses foreign technology for their production
New product or New process	Indicates if the establishment introduced a new or significantly improved product/service or production process
Informality	Indicates if informal markets are a consider a major obstacle for establishment operations
Share of Production Workers	Share of Production workers as % of total permanent workers
Share of skill workers	Share of Skill workers as % of total permanent workers
Share of seasonal workers	Share of seasonal workers as % of total permanent equivalent workers
Share of female workers	Share of Production workers as % of total permanent workers
Level of utilization of capital	%of current output compared to maximum output possible under normal circumstances
Hours of operation per week	Normal weekly hours of operations of the establishment, Includes the variable in levels and logs
New buildings	Indicates if establishment submitted an application to obtain a construction-related permit in last 2 years
Land problem	Indicates if access to land is considered a major obstacle for operation
Government problem	Indicates if government regulations are considered major obstacles for operation
Investment	Indicates if establishment has investment any resources on machinery or vehicles during last fiscal year
Hires seasonal workers	Indicates if establishment hires seasonal workers at all.
Industry fixed effects	Includes industry fixed effects using ISIC Rev.3.1 classification to 2 digits.
Nr of permanent workers	Total number of permanent workers, including its logarithm, logarith squared and interaction with a union Dummy
Nr workers on t-1	Total number of permanent workers 3 years ago
Nr production workers	Total number of workers directly engaged in the production process. Includes its log
Nr non production Workers	Total number of workers not engaged in the production process. Includes its log
Zero production workers	Indicates if there are no production workers in the establishment
Nr of seasonal workers	Nr of workers that are hired for a short-term (i.e. for less than a fiscal year), with no guarantee of renewal of employment contract. Includes its log and interaction with union dummy
Labor regulations	Indicates if establishments consider labor regulations as major obstacles for operations
Inadequate Education	Indicates if establishments consider inadequate education as major obstacles for operations
Manufacture production	% of sales that come from manufacture
Refusal capital	Indicate if the establishment refused to provide information on book or market capital values
Refusal land	Indicate if the establishment refused to provide information on book or market land value
Log sales	Logarithm of total sales in last fiscal year. Includes its square.
Sales in t-1 dummy	Indicates if the establishment didn't provide information on sales 3 years ago.
Union	Variables indicating if the establishment is unionized, the union density within the establishment and a dummy if more than 50% of the establishment is unionized.
Information quality flags	Two dummies indicating if the interviewer perceives the information provided is true, or if the data was taken from administrative records.
Workers avg education	Average education attainment of typical worker. 0-3 yrs of education, 4-6 yrs of education, 7-12 yrs of education and 13+ yrs of education
Training	Indicators of training among permanent workers: No active training program in the establishment, 0-33% of workers trained, 34-66% of workers trained and 67-100% workers trained.
Owner female	Indicates if any of the main owners of the establishment is female.
Region	Fixed effects using region dummies survey in each country. Varies across countries.

**Table B2. Imputed Variables:**

Variable	Method	Definition
Nr of workers in t-1	PMM	Total number of permanent workers 3 years ago
Cost of labor	PMM	Total annual cost of labor as share of sales
Cost of electricity	PMM	Total annual cost of electricity as share of sales
Cost of communications	PMM	Total annual cost of communications as share of sales
Cost of materials and inputs	PMM	Total annual cost of materials and inputs as share of sales
Cost of fuel	PMM	Total annual cost of fuel as share of sales
Cost of transportation	PMM	Total annual cost of transportation as share of sales
Cost of water	PMM	Total annual cost of water as share of sales
Cost of rentals	PMM	Total annual cost of rent of equipment, building and land as share of sales
Log Nr of workers in t-1	OLS	Log Total number of permanent workers 3 years ago
Log Sales in t-1	OLS	Log Sales 3 years ago
Log wages production workers	OLS	Log average wage of production workers
Log wages non production workers	OLS	Log average wage of non-production workers
Log capital (book value)	OLS	Log of net book value of machinery
Log capital (market value)	OLS	Log of hypothetical cost of purchase of machinery
Log materials and inputs	OLS	Log of total cost of material and inputs
Log salaries	OLS	Log of total cost of salaries

OLS: This method uses linear predictions (plus the empirical standard error) to impute the values of the missing values.

PMM: This method uses linear predictions to match observations with missing values to those with complete information. The observed values are then used for the imputation.

**Table B3. Other Measurements:**

Variable	Definition
Log value added pc	Value added is defined as sales minus costs on materials and inputs, electricity, fuel and water. It is divided by total equivalent permanent workers.
Log total equivalent permanent workers	Total equivalent workers are estimated as total permanent workers plus equivalent seasonal workers. Equivalent seasonal workers are estimated as total number of temporary/seasonal workers multiplied by the average time (in months) a temporal worker participates in the establishment in a year.
Log capital per capita	Log of hypothetical value of capital divided by total number of equivalent workers. Hypothetical value captures the market value of capital, or how much the establishment would pay for it in current state.
Profit	Price cost margin, defined as total sales minus total production costs, divided by total costs.
Sales growth	Defined as the difference between current log sales, and log sales three years ago
Investment in R&D	Indicates if the establishment has spent on research and development
Investment in capital	Indicates if the establishment has bought any fixed assets in the previous period



## APPENDIX C

### Effect of Unions on Establishment Productivity, by Country

#### Basic Brown and Medoff Specification

	Argentina	Bolivia	Chile	Mexico	Panama	Uruguay
%Workforce unionized	-0.377 <sup>^</sup> (0.057)	0.488 (0.166)	0.227 (0.212)	0.061 (0.675)	0.304 (0.367)	0.188 (0.409)
Log capital pc	-0.009 (0.914)	0.260* (0.000)	0.232* (0.000)	0.249* (0.000)	0.017 (0.814)	-0.03 (0.678)
Log total labor force	0.201* (0.000)	0.077 (0.426)	0.172* (0.001)	0.164* (0.000)	0.126 (0.101)	0.396* (0.000)
Constant	9.308* (0.000)	5.668* (0.000)	6.949* (0.000)	6.680* (0.000)	9.188* (0.000)	8.166* (0.000)
Observations	540	298	564	974	185	251

Notes: <sup>^</sup> p<0.1, <sup>+</sup> p<0.05, \* p<0.01. P-values in parentheses. Models include region and broad industry fixed effects

## REFERENCES

- Aidt, T., and Tzannatos, Z. (2002). *Unions and Collective Bargaining: Economic Effects in a Global Environment*. Washington, D.C.: World Bank.
- Addison, J. T., and Hirsch, B. T. (1989). Union Effects on Productivity, Profits, and Growth: Has the Long Run Arrived? *Journal of Labor Economics*, 7(1), 72-105.
- Aitken, B., Harrison, A., and Lipsey, R. E. (1996). Wages and foreign ownership A comparative study of Mexico, Venezuela, and the United States. *Journal of International Economics*, 40 (3-4), 345-371.
- Alexander, R. J., and Parker, E. M. (2005). *A History of Organized Labor in Bolivia*. Westport, CT: Praeger Pub Text.
- Allen, S. G. (1984). Unionized Construction Workers Are More Productive. *The Quarterly Journal of Economics*, 99 (2), 251-274.
- Anner, M. (2008). Meeting the Challenges of Industrial Restructuring: Labor Reform and Enforcement in Latin America. *Latin American Politics and Society*, 50(2), 33-65.
- Baldwin, C. Y. (1983). Productivity and Labor Unions: An Application of the Theory of Self-Enforcing Contracts. *The Journal of Business*, 56 (2), 155-185.
- Black, S. E., and Lynch, L. M. (2001). How to Compete: The Impact of Workplace Practices and Information Technology on Productivity. *Review of Economics and Statistics*, 83 (3), 434-445.
- Bloom, N., and Reenen, J. V. (2011). Chapter 19 Human Resource Management and Productivity. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics* (Vol. Volume 4, Part B, pp. 1697-1767): Elsevier.
- Brown, C., and Medoff, J. (1978). Trade Unions in the Production Process. *The Journal of Political Economy*, 86 (3), 355-378.
- Campero, G. (2001). *Trade union responses to globalization: Chile*. Discussion Paper DP/126/2001. International Institute for Labour Studies.
- Carrière, J., Haworth, N., and Roddick, J. (1989). *The State, Industrial Relations and the Labour Movement in Latin America*. New York: St. Martin's Press.
- Cassoni, A., Allen, S. G., and Labadie, G. J. (2004). Unions and Employment in Uruguay. In J. Heckman and C. Pages (Eds.), *Law and Employment: Lessons from Latin American and the Caribbean*: University of Chicago Press.
- Cassoni, A., Labadie, G. J., and Fachola, G. (2005). The Economic Effects of Unions in Latin America: Their Impact on Wages and the Economic Performance in Uruguay. In P. Kuhn and G. Márquez (Eds.), *What Difference Do Unions Make? Their Impact on Productivity and Wages in Latin America*: Inter-American Development Bank.
- Christensen, L. R., Jorgenson, D. W., and Lau, L. J. (1973). Transcendental Logarithmic Production Frontiers. *The Review of Economics and Statistics*, 55 (1), 28-45.
- Connolly, R. A., Hirsch, B. T., and Hirschey, M. (1986). Union Rent Seeking, Intangible Capital, and Market Value of the Firm. *The Review of Economics and Statistics*, 68 (4), 567-577.
- DiNardo, J., and Lee, D. S. (2004). Economic Impacts of New Unionization on Private Sector Employers: 1984-2001. *The Quarterly Journal of Economics*, 119(4), 1383-1441.
- Doucouliagos, C., and Laroche, P. (2003). What do Unions do to Productivity? A Meta-Analysis. *Industrial Relations: A Journal of Economy and Society*, 42 (4), 650-691.
- Doucouliagos, C., and Laroche, P. (2009). Unions and Profits: A Meta-Regression Analysis. *Industrial Relations: A Journal of Economy and Society*, 48 (1), 146-184
- Doucouliagos, C., and Laroche, P. (2013). Unions and Innovation: New Insights From the Cross-Country Evidence. *Industrial Relations: A Journal of Economy and Society*, 52 (2), 467-491.

- Edwards, S. (1989). The Debt Crisis and Economic Adjustment in Latin America. *Latin American Research Review*, 24(3), 172-186.
- Fairris, D. (2003). Unions and Wage Inequality in Mexico. *Industrial and Labor Relations Review*, 56 (3), 481-497.
- Fairris, D. (2006). Union Voice Effects in Mexico. *British Journal of Industrial Relations*, 44 (4), 781-800.
- Falaris, E. M. (2008). A Quantile Regression Analysis of Wages in Panama. *Review of Development Economics*, 12 (3), 498-514.
- Freeman, R. B. (2010). Chapter 70 Labor Regulations, Unions, and Social Protection in Developing Countries: Market Distortions or Efficient Institutions?. In D. Rodrik and M. Rosenzweig (Eds.), *Handbook of Development Economics* (Vol. Volume 5, pp. 4657-4702): Elsevier.
- Freeman, R. B., and Medoff, J. L. (1984). *What Do Unions Do?* (Vol. 23): Basic Books New York.
- Fuchs, V. R., Krueger, A. B., and Poterba, J. M. (1998). Economists' Views About Parameters, Values, and Policies: Survey Results in Labor and Public Economics. *Journal of Economic Literature*, 36(3), 1387-1425.
- Graham, J. W., Hofer, S. M., Donaldson, S. I., MacKinnon, D. P., and Schafer, J. L. (1997). Analysis with missing data in prevention research *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 325-366).
- Hayter, S., and Stoevska, V. (2011). *Social Dialogue Indicators: International Statistical Inquiry 2008-09*. Technical Brief. International Labour Office.
- Hirsch, B. T. (1991). *Labor Unions and the Economic Performance of Firms*. Kalamazoo, Michigan: Upjohn Institute for Economic Research.
- Hirsch, B. T. (2004). What do Unions do for Economic Performance?. *Journal of Labor Research*, 25 (3), 415-455.
- Hirsch, B. T. (2008). Sluggish Institutions in a Dynamic World: Can Unions and Industrial Competition Coexist?. *Journal of Economic Perspectives*, 22(1), 153-176.
- Hirsch, B. T., and Addison, J. T. (1986). *The Economic Analysis of Unions: New Approaches and Evidence*: Allen and Unwin.
- Hirsch, B. T., and Connolly, R. A. (1987). Do Unions Capture Monopoly Profits?. *Industrial and Labor Relations Review*, 41 (1), 118-136.
- Horton, N. J., and Lipsitz, S. R. (2001). Multiple Imputation in Practice. *The American Statistician*, 55 (3), 244-254.
- Hudson, R. A. (Ed.). (1994). *Chile: A Country Study*. Washington, D.C.: Federal Research Division, Library of Congress.
- Hudson, R. A., and Hanratty, D. M. (Eds.). (1991). *Bolivia: A Country Study*. Washington, D.C. : Federal Research Division, Library of Congress.
- Hudson, R. A., and Meditz, S. W. (Eds.). (1992). *Uruguay: A Country Study*. Washington, D.C. : Federal Research Division, Library of Congress.
- ITUC. (2007). Annual Survey of Violations of Trade Union Rights. <http://survey07.ituc-csi.org/getcontinent.php?IDContinent=0&IDLang=EN>
- Kaufman, B. E. (2004). What Unions Do: Insights From Economic Theory. *Journal of Labor Research*, 25(3), 351-382.
- Kleiner, M. M., Leonard, J. S., and Pilarski, A. M. (2002). How industrial relations affects plant performance : the case of commercial aircraft manufacturing. *Industrial and Labor Relations Review*, 55 (2), 195-218.
- Kuhn, P. (1985). Union Productivity Effects and Economic Efficiency. *Journal of Labor Research*, 6 (3), 229.
- Kuhn, P. (1998). Unions and the Economy: What We Know; What We Should Know. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 31 (5), 1033-1056.
- Lee, D. and Mas A. (2012). Long-Run Impacts of Unions on Firms: New Evidence from Financial Markets, 1961-1999. *The Quarterly Journal of Economics*, 127 (1) , 333-378.

- Little, R. J. A., and Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed. ed.). New York: John Wiley and Sons.
- Machin, S., and Wadhvani, S. (1991). The Effects of Unions on Investment and Innovation: Evidence From Wirs. *The Economic Journal*, 101 (405), 324-330.
- Mansilla, H. C. F. (1993). Apogeo y Declinacion del Movimiento Sindical Boliviano. *Revista de Estudios Politicos* (79), 227-245.
- Meditz, S. W., and Hanratty, D. M. (Eds.). (1989). *Panama: a country Study*: Federal Research Division, Library of Congress.
- Menezes-Filho, N., Zylberstajn, H., Chahad, J. P., and Pazello, E. (2005). Unions and the Economic Performance of Brazilian Establishments. In P. Kuhn and G. Márquez (Eds.), *What Difference Do Unions Make? Their Impact on Productivity and Wages in Latin America*: Inter-American Development Bank.
- Merrill, T., and Miró, R. (Eds.). (1997). *Mexico: A Country Study*. Washington, D.C.: Federal Research Division, Library of Congress.
- Murillo, M. V. (2000). From Populism to Neoliberalism: Labor Unions and Market Reforms in Latin America. *World Politics*, 52 (2), 135-174.
- Murillo, M. V., and Schrank, A. (2005). With a Little Help from my Friends. *Comparative Political Studies*, 38 (8), 971-999.
- O'Connell, L. D. (1999). Collective Bargaining Systems in 6 Latin American Countries: Degrees of Autonomy and Decentralization. *Inter-American Development Bank, Office of the Chief Economist, Working paper*, 399.
- OECD. (1996). *Trade, Employment and Labour Standards: A Study of Core Workers' Rights and International Trade*. Paris France: OECD Publishing.
- Rios-Avila, F. and Hirsch B.T. (forthcoming). Unions, Wages Gaps and Wage Dispersion: New Evidence from the Americas. *Industrial Relations: A Journal of Economy and Society*.
- Ronconi, L. (2012). Globalization, Domestic Institutions, and Enforcement of Labor Law: Evidence from Latin America. *Industrial Relations: A Journal of Economy and Society*, 51(1), 89-105.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons, Inc.
- Saavedra, J., and Torero, M. (2005). Union Density Changes and Union Effects on Firm Performance in Peru. In P. Kuhn and G. Márquez (Eds.), *What Difference Do Unions Make? Their Impact on Productivity and Wages in Latin America*: Inter-American Development Bank.
- Turnbull, P. J. (1991). Trade Unions and Productivity: Opening the Harvard "Black Boxes". *Journal of Labor Research*, 12 (2), 135-150.
- Ulloa, V. (2003). *El Movimiento Sindical Chileno del Siglo XX Hasta Nuestros Días*: OIT, Oficina Internacional del Trabajo. Textos de Capacitacion.
- Urizar H., C., and Lee, S. (2005). The Effects of Unions on Productivity: Evidence From Large Coffee Producers in Guatemala. In P. Kuhn and G. Márquez (Eds.), *What Difference Do Unions Make? Their Impact on Productivity and Wages in Latin America*: Inter-American Development Bank.
- van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, 16 (3), 219-242.
- Vergara, M. (1998). *Incremento de Remuneraciones Asociados a Aumentos de Productividad Aportes al Debate Laboral 5*. Department of Labor: Chile.
- Wessels, W. J. (1994). Do Unionized Firms Hire Better Workers? *Economic Inquiry*, 32 (4), 616-629.
- World Bank (2006). Latin American Enterprise Surveys Data set. World Bank, Washington, DC. Available at <http://www.enterprisesurveys.org/Data>

---

ESTIMACIÓN ROBUSTA DE BETAS  
Y EL RATIO DE COBERTURA SOBRE  
FUTUROS DE ÍNDICES BURSÁTILES  
EN EL MERCADO INTEGRADO  
LATINOAMERICANO (MILA)

Robust Estimation of beta  
and the hedging ratio in Stock  
Index Futures In the Integrated  
Latin American Market

Andrés Gómez, Astrid K. Gutiérrez  
y Juan C. Gutiérrez

---

## Research Article

# ESTIMACIÓN ROBUSTA DE BETAS Y EL RATIO DE COBERTURA SOBRE FUTUROS DE ÍNDICES BURSÁTILES EN EL MERCADO INTEGRADO LATINOAMERICANO (MILA)

## Robust Estimation of beta and the hedging ratio in Stock Index Futures In the Integrated Latin American Market

Andrés Gómez<sup>a</sup>, Astrid K. Gutiérrez<sup>b</sup> y Juan C. Gutiérrez<sup>c\*</sup>

**Palabras clave:** estimación de beta, método robusto MM (RMM), método mínimos cuadrados ordinarios (MCO), cobertura con futuros sobre índices MILA.

**Key words:** estimation of beta, robust statistics MM (RMM), ordinary least squares (OLS), hedging ratio with stock MILA market index futures.

**JEL Classifications:** G12, G17, C14, C18

**Received:** 15/12/2016

**Accepted:** 18/04/2017

**Published:** 15/06/2017

### Resumen

El presente trabajo tiene por objeto estudiar el efecto que ejercen los datos atípicos en el parámetro beta de acciones pertenecientes al Mercado Integrado Latinoamericano (MILA), estimado por dos diferentes métodos: mínimos cuadrados ordinarios (MCO) y método robusto MM (RMM). Adicionalmente, para ilustrar la relevancia empírica de las betas calculadas, se efectuó una aplicación de cobertura con futuros sobre índices. Los resultados indican que las estimaciones realizadas por el método RMM, ofrecen un mejor ajuste y una mayor eficiencia de la cobertura cuando existe presencia de datos atípicos en la ventana de estimación de la beta.

### Abstract

This paper examines the effect exerted by outliers in the equity betas in the Integrated Latin American Market (MILA), estimated by two different methods: ordinary least squares (OLS) and robust estimation (RMM). To illustrate the empirical relevance of the estimated betas, we evaluate the hedging ratio using stock index futures. The results indicate that the estimates made by the RMM method provide a better fit and increase the efficiency of a hedging strategy when there are outliers in the estimation window of beta.

a, b. Empresas Públicas de Medellín, Colombia.

c. Universidad EAFIT, Escuela de Economía y Finanzas, Departamento de Finanzas, Grupo de Investigación en Finanzas y Banca (GIFyB).

\* Autor para correspondencia:  
Correo electrónico: [jgutie31@eafit.edu.co](mailto:jgutie31@eafit.edu.co)

## 1. INTRODUCCIÓN

### Efectos del mercado y estimación de beta

El crecimiento sostenido que se ha venido presentando en los mercados de capitales en países emergentes latinoamericanos y el nivel de integración que se ha logrado por acuerdos como el Mercado Integrado Latinoamericano (MILA), han atraído mayor atención de inversionistas de portafolio de la región, quienes monitorean permanentemente medidas de riesgo y rendimiento de los activos del mercado para la toma de decisiones de inversión.

En particular, muchos inversionistas fijan su atención en el parámetro beta, el cual se constituye en una medida del riesgo sistemático debido a que permite capturar el movimiento del precio de las acciones atribuible a los movimientos del mercado accionario, por lo tanto, el empleo de métodos confiables para la estimación de betas resulta de gran interés para la comunidad financiera.

La formulación de beta tiene su raíz en el Modelo de Valoración de Activos de Capital (CAPM) introducido por [Sharpe \(1964\)](#), [Lintner \(1965\)](#) y [Mossin \(1966\)](#), y desde el punto de vista práctico su aplicación en la determinación de los rendimientos esperados de los activos con relación a su nivel riesgo, se ha extendido en forma generalizada entre los directivos de empresas y profesores de finanzas, tal como lo soportan los hallazgos empíricos de [Graham y Harvey \(2001\)](#), para Estados Unidos, y [Fernández y Bermejo \(2009\)](#), para Europa.

Desde un punto de vista empírico, el modelo CAPM presenta las limitaciones propias de la idealización de un portafolio de mercado, supone homogeneidad en las expectativas de los inversionistas y considera la rentabilidad esperada y no la rentabilidad real o realizada. Por su parte, la implementación se ha efectuado mediante el modelo de mercado que utiliza el índice bursátil como proxy representativo del portafolio de mercado y emplea la rentabilidad realizada histórica.

En la práctica, la técnica más utilizada para la estimación del parámetro beta es el método de mínimos cuadrados ordinarios (MCO) basado en la utilización de información histórica, lo cual se justifica en el hecho que el coeficiente resultante es el mejor estimador lineal insesgado óptimo y provee una distribución teórica conveniente para realizar inferencias estadísticas, cuando los errores tienen una distribución normal.

Sin embargo, distintas investigaciones como las realizadas por [Mandelbrot \(1963\)](#), [Fama \(1965\)](#), [Roll \(1988\)](#), entre otras, han generado evidencia empírica que indica que las series de retornos de los activos financieros contienen valores atípicos, generando distribuciones de colas pesadas; que corresponden a eventos intrínsecos de la firma, eventos extremos del mercado, o errores de medición. Si estas observaciones atípicas no son tenidas en cuenta de manera apropiada, el parámetro beta resultante estará sesgado, derivando en una interpretación equivocada de las características de riesgo y rendimiento de la firma.

En este sentido, [Bekaert, Erb, Harvey, & Viskanta \(1998\)](#) estudian los mercados accionarios emergentes y encuentran marcadas características de curtosis, asimetría y colas pesadas, que alejan la distribución de los retornos de la normalidad. Bajo estas condiciones, se pierde la eficiencia del estimador de MCO dado que es altamente sensible a la presencia de valores extremos ([Huber, 1964; 1973; 1981; Chan & Lakonishok, 1992; Knez & Ready, 1997; Martin & Simin, 1999 y 2003; Bailer, 2005; Maronna, Martin & Yohai, 2006](#)).

Específicamente, al evaluar los rendimientos históricos de los mercados financieros del MILA, se puede validar lo indicado anteriormente, encontrando tendencias marcadas de sesgo y colas pesadas en sus distribuciones. En consecuencia, el manejo estadístico de los datos históricos disponibles del MILA requiere de enfoques complementarios con el fin de alcanzar valores adecuados para la estimación de betas.

Al respecto, algunos practicantes de las finanzas han optado por emplear diversos métodos para eliminar los efectos de los datos extremos que inducen distorsiones, tales como la winsorización, entre otros. Sin embargo, aunque estos métodos permiten mitigar el problema de estimación, no logran proteger el modelo contra distorsiones en la estimación de la beta por MCO.

Teniendo en cuenta estas limitaciones, en las últimas décadas se incrementó el interés de los académicos por encontrar métodos de estimación resistentes a la presencia de los valores extremos comunes en los mercados financieros, que arrojen parámetros de betas confiables bajo distribuciones no normales, y que conserven su eficiencia aun cuando la distribución de los retornos sea normal.

De aquí que, en la presente investigación, se propone el uso de un método alternativo y complementario al MCO fundamentado en la estadística robusta para la estimación de betas, considerando que sus estimadores conservan adecuadas propiedades de eficiencia y estabilidad tanto en presencia de datos atípicos como en ausencia de ellos.

En particular, se hace uso del estimador robusto MM (RMM), considerando que dentro del amplio espectro de estimadores robustos desarrollados desde los años 70's, éste no sólo alcanza simultáneamente las propiedades de alta eficiencia y estabilidad, sino también la evidencia empírica lo señala como el de mejor ajuste en el cálculo de la beta ([Martin & Simin, 1999 y 2003](#); [Bailer, 2005](#); [Bailer, Maravina, & Martin, 2011](#)).

### **Betas y su uso en estrategias de cobertura con futuros sobre índices bursátiles**

Revisando la relevancia del parámetro beta en las finanzas, se encuentran múltiples usos que denotan la importancia de afirmar el método de estimación.

Un caso de estudio corresponde al uso de las Betas en estrategias de cobertura con futuros sobre índices bursátiles. Para este aparte, los inversionistas buscan administrar el riesgo sistemático asociado a sus inversiones, encontrando el tamaño de la posición corta que debe tomar el inversionista en el mercado de futuros, compensando la posición larga sostenida en el portafolio de activos riesgosos, de tal forma que se maximice la utilidad esperada del agente.

Este problema corresponde a la estimación del ratio de cobertura óptima, el cual puede ser definido a partir de betas cuando el portafolio que se desea cubrir no replica de manera exacta el índice subyacente del contrato de futuros ([Hull, 2009](#)). Así, la beta se constituye en un importante parámetro de decisión de inversión y por tanto, la metodología que se emplee para estimarla debe proveer parámetros confiables que permitan el logro de los objetivos de cobertura.

Con el objetivo de revisar los efectos generados en la estimación de la beta por el método alternativo robusto RMM seleccionado, en la presente investigación se efectúa una aplicación cuyo propósito es medir la mejora en la efectividad de una estrategia de cobertura con futuros sobre índices, cuando



se emplean betas estimadas por el método robusto seleccionado MM (RMM) en lugar de betas estimadas a partir de MCO

De esta manera, esta investigación se constituye en una importante contribución a la literatura sobre el tema, dado que el estudio de la estimación del parámetro beta por el método RMM y su aplicación en cobertura no ha sido estudiado previamente en los mercados integrantes del MILA.

El presente artículo está organizado de la siguiente manera: el [numeral 2](#) relata las consideraciones teóricas y metodológicas que inician con un breve contexto acerca de la fundamentación de la estadística robusta, posteriormente, esboza las técnicas desarrolladas para la estimación robusta, a continuación, expone la relevancia de la estadística robusta en las finanzas, también, profundiza en el método robusto para la estimación de la beta y finalmente, detalla el estimador robusto RMM seleccionado. El [numeral 3](#) describe la metodología, la muestra y las medidas de eficiencia empleadas en la estimación de las betas por el método clásico y por el método robusto; y adicionalmente se describen las estrategias de cobertura con futuros sobre índices bursátiles, construidas a partir de las betas estimadas, y las medidas de desempeño aplicadas para medir su efectividad. Finalmente, en los [numerales 4 y 5](#), se exponen los resultados y conclusiones obtenidas en las distintas aplicaciones.

## 2. CONSIDERACIONES TEÓRICAS Y METODOLÓGICAS

### Fundamentos de la estadística robusta

Para caracterizar cualquier tipo de distribución de probabilidad se usan medidas tanto de localización como de escala o dispersión, por ejemplo, la media y la desviación estándar para la distribución normal, respectivamente. Al respecto, [Tukey \(1960\)](#), al evidenciar la inestabilidad de la media y la desviación estándar cuando la muestra se encuentra contaminada con un pequeño porcentaje de datos atípicos, propuso la mediana y la desviación absoluta de la mediana (MAD) como medidas alternativas para la media y la desviación estándar, por poseer mejores propiedades de eficiencia bajo este tipo de distribución.

Posteriormente, [Huber \(1964; 1981\)](#) introdujo los fundamentos de las técnicas de estimación robusta desarrollando la estadística matemática subyacente. Luego, [Hampel \(1968\)](#) complementó este desarrollo con los conceptos de punto de ruptura y curva de influencia para medir el grado de robustez. De forma general, para el caso de estimadores robustos lo que se busca es minimizar la distancia entre la variable aleatoria y una medida de localización o escala equivalentes a las usadas para la distribución normal, en forma tal, que incorporen las siguientes tres características ([Wilcox R. R., 2005a](#)):

- **Robustez cualitativa:** Cuando la medida de localización o escala no se ve afectada ante pequeños cambios en la función de probabilidad, se parte de la continuidad de esta función como punto inicial para garantizar esta propiedad.
- **Robustez infinitesimal:** Está ligada al concepto de eficiencia de la estimación y busca que la medida de localización o escala no presente grandes cambios ante pequeñas variaciones de la muestra. Para garantizar esta condición, se requiere que la función, también conocida como función de influencia, que describe los cambios de la medida, sea limitada.

- **Robustez cuantitativa:** Está definida con base en el punto de ruptura; el cual corresponde a la máxima fracción de datos atípicos que puede contener una muestra antes de que la medida de estimación de localización o escala tienda a infinito.

En síntesis, un estimador puede considerarse robusto si cumple con tener un punto de ruptura y eficiencia relativa alta, entendiendo esta última como el nivel de desempeño que tiene un estimador alternativo frente a otro que se evalúe.

### Regresión Robusta

Para el análisis de regresión, se han desarrollado múltiples técnicas, sin embargo, las más usadas en las aplicaciones financieras son las siguientes:

- **Estimador de Mínimas Desviaciones Absolutas:** (conocido también como regresión L1 o LAD): la regresión por este método se obtiene minimizando la sumatoria de los errores en valor absoluto. Aunque su aplicación es quizá, la más antigua, se reconocen las siguientes restricciones en su uso que pueden sesgar los resultados de la regresión: la regresión L1 resiste la presencia de valores atípicos en la variable explicada Y, pero es poco efectiva para valores anómalos en las variables explicativas X; el vector de coeficientes estimados no es único; la eficiencia disminuye a medida que aumenta el número de casos atípicos y la regresión asigna gran peso a residuales pequeños ([Wilcox R. R., 2005b](#)).
- **Estimador M:** Este estimador introducido por [Huber \(1973\)](#), se define por una función de minimización de residuales estimados y puede ser construido seleccionando una función de ponderación que asigne bajas ponderaciones a los datos atípicos de las series analizadas. La limitación del método se encuentra en que los estimadores se pueden alterar por la presencia de valores atípicos en las variables explicativas, siendo su punto de ruptura nulo.
- **Estimadores L:** Están basados en combinaciones lineales de orden estadístico y no implican, en general, la realización de ningún supuesto (implícito o explícito) sobre la distribución de los residuales, y pueden ser definidos como el promedio ponderado de los parámetros de la regresión lineal en cada cuantil de la regresión, organizada por tamaño de los residuales. La desventaja del método es que su eficiencia relativa es baja.
- **Estimador Mínimos Cuadrados Recortados (LTS, por sus siglas en inglés):** Este método introducido por [Rousseeuw \(1984\)](#) está basado en la minimización de la mediana del cuadrado de los residuos e incorpora un punto de ruptura alto. Sin embargo, tiene una baja eficiencia relativa cuando la distribución de los datos es normal ([Wilcox R. R., 2005b](#)).
- **Estimador S:** Fue introducido por [Rousseeuw y Yohai \(1984\)](#) y tiene la propiedad de minimizar la mediana de los valores absolutos de los residuos. Este estimador tiene un punto de ruptura alto, inclusive mayor al del método LTS, sin embargo, tiene la desventaja al no poder ser estable y eficiente al mismo tiempo. Por otra parte, ofrece la ventaja de que puede calcularse a partir de los datos sin necesidad de estimadores iniciales de localización y escala. Por lo anterior, los estimadores S son utilizados en los algoritmos de cálculo de otros estimadores robustos ([Zamar, 1994](#)).
- **Estimador MM:** Los estimadores MM introducidos por [Yohai \(1987\)](#), corresponden a estimadores de máxima verosimilitud que tienen simultáneamente la propiedad de ofrecer alta eficiencia, cuando

los errores de la regresión se distribuyen normalmente, y ofrecer un alto punto de ruptura. Por lo anterior, su uso es adecuado tanto en presencia de datos atípicos como en ausencia de ellos.

En el siguiente aparte, se explora la aplicación de la estadística robusta en las finanzas en el cálculo de la beta, y posteriormente, se detallan las características del estimador seleccionado en la presente investigación.

### Relevancia de la estadística robusta en finanzas

En el ámbito financiero y específicamente en el contexto de la estimación de beta, se han realizado investigaciones que brindan evidencia empírica acerca de las propiedades teóricas de los estimadores robustos. Así, [Sharpe \(1971\)](#), y [Cornell & Dietrich \(1978\)](#) emplearon el estimador LAD<sup>1</sup> para estimar la beta de grandes firmas pertenecientes al Dow Jones Industrial y al S&P 500, respectivamente, con la expectativa de encontrar un método que tuviera un mejor desempeño al obtenido por MCO dado que los retornos algunas veces no eran normales. Sin embargo, concluyeron que esta alternativa robusta ofrecía una leve mejora de la capacidad explicativa de este estimador frente al de MCO. No obstante, sus resultados no fueron concluyentes acerca de la bondad de emplear estimadores robustos.

Posteriormente, [Chan & Lakonishok \(1992\)](#) implementaron nuevos estimadores robustos tipo L para el cálculo de la beta, concluyendo que estos ofrecen una mejor eficiencia que los estimadores de MCO, debido a la distorsión incorporada por los valores atípicos asociados con los anuncios de las firmas. Por su parte, [Knez y Ready \(1997\)](#) aplican el modelo multifactorial usando estimadores robustos LTS, encontrando que con sólo aislar el 1% de los datos de rendimientos mensuales anormales, la relación negativa del tamaño de la firma con los rendimientos promedio pasa a ser positiva; esto se explica por la asimetría de la distribución estadística de los retornos, especialmente en las firmas nacientes y pequeñas.

Luego, un análisis comparativo de los diferentes estimadores robustos fue introducido por [Bowie & Bradfield \(1998\)](#), concluyendo que el estimador de influencia limitada de la familia de estimadores M es el que presenta una mayor eficiencia en la estimación, en comparación con el estimador MCO. Los autores también indican que las diferencias sistemáticas entre MCO y los estimadores robustos, motivan a los investigadores a profundizar en la relación que existe entre el retorno de las acciones y el retorno del mercado cuando se presentan condiciones extremas derivadas de la baja liquidez de algunos activos, tal como se presenta en mercados financieros de poca profundidad y desarrollo.

A partir de allí, aparecen los esfuerzos de [Martin & Simin \(1999\)](#), [Martin et al. \(2003\)](#), quienes brindan evidencia empírica que indica el logro de un mejor ajuste de la beta mediante el uso del estimador RMM, dado que incorpora una ponderación adecuada de los datos atípicos. Al mismo tiempo, estos autores aclaran que las divergencias entre los resultados obtenidos a partir de los estimadores MCO y de los estimadores robustos M son característicos en compañías pequeñas de baja capitalización, lo cual contrasta los resultados obtenidos en los estudios previos de [Sharpe \(1971\)](#) y [Cornell et al. \(1978\)](#) que fueron obtenidos con base en firmas de gran tamaño. Resultados similares obtuvieron ([Bailer, 2005](#)) y ([Bailer et al., 2011](#)), al emplear el mismo estimador robusto.

En cuanto a la evidencia empírica asociada con el uso de betas en estrategias de cobertura con futuros sobre índices bursátiles, se destaca la investigación de [Butterworth & Holmes \(2001\)](#), quienes motivados por encontrar un ratio de cobertura óptima, utilizan el método robusto de mínimos cuadra-

1 LAD: Least Absolute Deviation

dos recortados (LTS) para la estimación del parámetro beta, que se emplea en el cálculo de la razón de cobertura. Los autores verifican la efectividad de esta técnica en comparación con los resultados obtenidos a partir de otros tres modelos para la construcción del ratio de cobertura, buscando verificar si la exclusión de los valores extremos genera cambios en el resultado de la estrategia de cobertura; para tal efecto, utilizan futuros sobre el índice Mid250 y FT-SE100 en el período comprendido entre febrero de 1994 a diciembre de 1996. Ellos encuentran que la regresión LTS tiene un pequeño impacto en el rendimiento de la cobertura en el contexto de los fondos de inversión.

Por su parte, [Harris & Shen \(2003\)](#) implementan una alternativa robusta que consiste en construir estrategias de cobertura dinámica a partir del contraste entre metodologías de regresión de promedios móviles ponderados exponencialmente (EWMA), que luego son comparadas con métodos estándar de las mismas. Los resultados señalan que la razón de cobertura robusta genera una menor varianza del portafolio cubierto en comparación con la que se obtiene con la razón de cobertura estimada por métodos estándar.

El aporte metodológico de estas investigaciones permite explorar el uso de otros estimadores robustos que mejores propiedades respecto a escenarios con presencia y en ausencia de datos atípicos.

Teniendo en cuenta las propiedades teóricas de los distintos estimadores robustos y la evidencia empírica, el presente trabajo de investigación empleará el método de estimador robusto RMM. En el siguiente aparte, se describe la estimación de la beta robusta y finalmente, se describe el proceso iterativo de regresión en el estimador robusto RMM seleccionado.

### Estimación de la beta robusta

Para el cálculo de la beta de un activo, se parte del modelo de índice único:

$$r_{t,i} = \alpha_i + \beta_i r_{m,t} + \mu_{t,i} \quad t = 1, \dots, n \quad (1)$$

Donde  $r_{t,i}$  define el rendimiento del activo  $i$  en exceso de la tasa libre de riesgo,  $r_{m,t}$  es el rendimiento del mercado en exceso de la tasa libre de riesgo,  $\alpha_i$  corresponde al rendimiento en exceso del activo  $i$  si el rendimiento en exceso del mercado es nulo,  $\beta_i$  es la sensibilidad del rendimiento en exceso del activo  $i$  con relación al rendimiento en exceso del mercado, y  $\mu_{t,i}$  caracteriza el componente específico de la firma, el cual no se encuentra correlacionado con los retornos del mercado.

Como se mencionó previamente, convencionalmente se ha utilizado el método de MCO para la estimación de las betas, el cual fija un ajuste lineal del conjunto de pares de datos, rendimiento en exceso del activo y rendimiento en exceso de mercado, minimizando la suma de los cuadrados de los errores de los datos con respecto a la línea de ajuste, según la siguiente expresión (Martin et al., 2003):

$$\hat{\beta}_{MCO,i} = \operatorname{argmin}_{\beta_i} \sum_{t=1}^n f(\mu_{t,i}) = \operatorname{argmin}_{\beta_i} \sum_{t=1}^n (\mu_{t,i})^2 \quad (2)$$

Donde  $\mu_{t,i} = r_{t,i} - \alpha_i - \beta_i r_{m,t}$ , y  $f(\mu_{t,i}) = (\mu_{t,i})^2$  es la función objetivo, la cual trata los errores cuadráticamente causando que los valores atípicos generen una desfavorable influencia sobre el ajuste de la regresión lineal, derivando en una disminución de la eficiencia del estimador.

Por su parte los estimadores robustos de la familia de estimadores M, buscan minimizar el sesgo máximo y la varianza máxima sin detrimento de la eficiencia del estimador, mediante la exclusión

o ponderación de los residuales de valores atípicos causantes del sesgo y exceso de curtosis en la distribución de la muestra.

Para lograr este efecto, este tipo de estimadores consideran, en primer lugar, una función objetivo  $\rho$  dentro del proceso de minimización que difiere de la función cuadrática empleada en el proceso de estimación de MCO, y, en segundo lugar, utilizan los residuales estandarizados por una medida robusta de escala (Bailer, et al., 2011):

$$\hat{\beta}_{Mi} = \operatorname{argmin}_{\beta_i} \sum_{t=1}^n f(\mu_{s,t}) = \operatorname{argmin}_{\beta_i} \sum_{t=1}^n \rho\left(\frac{\mu_{s,t}}{s}\right) \quad (3)$$

Así, al aplicar la condición de primer orden con respecto a la beta, el estimador robusto de localización se obtiene de resolver la siguiente ecuación, en la cual la función  $\psi$  es la derivada de la función  $\rho$  ( $\psi = \rho'$ ):

$$\sum_{t=1}^n r_{m_t} \cdot \psi\left(\frac{r_{s,t} - \hat{\alpha}_i - \hat{\beta}_i r_{m_t}}{s}\right) = \sum_{t=1}^n r_{m_t} \cdot \psi\left(\frac{r_{s,t} - \hat{\alpha}_i - \hat{\beta}_i r_{m_t}}{\operatorname{MADN}(r_{m_t})}\right) \quad (4)$$

Donde la variable que define la desviación absoluta normalizada de la mediana de los rendimientos del mercado,  $s$ , es una medida robusta de escala que permite hacer el estimador invariante con respecto a la escala. Esta variable se encarga de caracterizar los residuales, de forma tal, que se pueda determinar cuando éstos son grandes o pequeños, evitando que los valores atípicos puedan no ser detectados quedando enmascarados en la muestra. Esta medida de escala se rige por la siguiente expresión (Maronna et al., 2006):

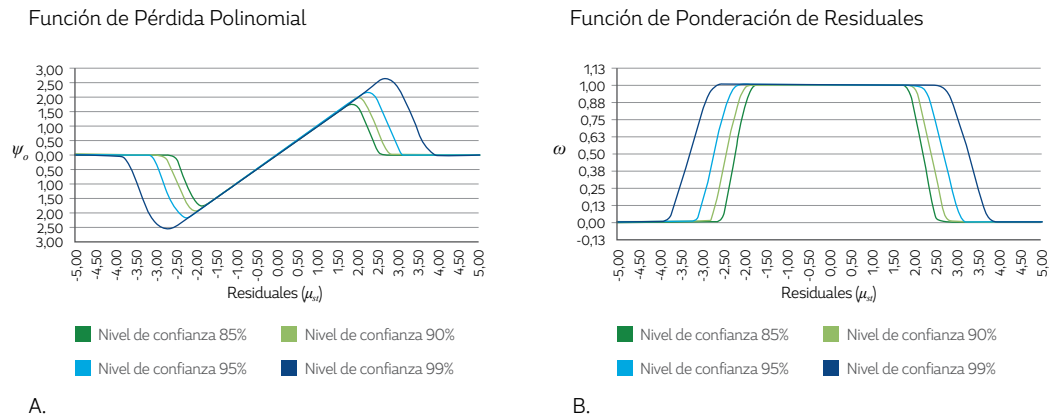
$$\operatorname{MADN}(r_{m_t}) = \frac{\operatorname{MAD}(r_{m_t})}{Z_{0,75}} = \frac{\operatorname{MEDIANA}|r_{m_t} - \operatorname{MEDIANA}(r_{m_t})|}{0,6745} \quad (5)$$

La constante 0,6745 corresponde al cuantil 75 de la distribución normal y hace que la medida de escala MADN, sea aproximadamente un estimador insesgado de  $\sigma$  cuando la distribución de los residuales es normal y el tamaño de la muestra ( $n$ ) es grande. Por su parte, la función objetivo  $\rho$ , es una función de pérdida simétrica que, según sus características, imprime una mayor eficiencia al estimador robusto de localización.

Si bien son diversas las funciones de pérdida que se han propuesto, fueron Martin et al. (1999), quienes en su estudio de betas del mercado estadounidense introdujeron una función de pérdida polinomial. Los autores definieron dicha función de forma tal, que su derivada ( $\psi_o$ ) fuera una función limitada, suavizada y continua por tramos, en los que se incluye una zona de descenso a cero que permite evitar problemas de convergencia durante el proceso de optimización.

La principal característica de estas funciones es que no sólo minimizan el sesgo asintótico garantizando una mayor eficiencia relativa, si no que permiten al usuario establecer, mediante el parámetro  $c$ , el nivel de confianza deseado para el estimador robusto de localización. Este valor  $c$ , determina una región por fuera de la cual se rechazan los residuales en el proceso de estimación, lo cual se evidencia con mayor claridad en la función de ponderación de residuales ( $\omega$ ). La siguiente figura muestra las funciones óptimas  $\psi_o$  y  $\omega$ , cuya expresión matemática se presenta en el Apéndice A.

**Figura 1. Forma funcional de  $\psi_o$  y  $\omega$**



Se puede apreciar en el panel A. de la [Figura 1.](#), que la función  $\psi$  está determinada por tres tramos; el primero, delimitado por el intervalo  $|\mu_{s_i}| \leq 2c$ , donde la función es lineal, equivalente a su similar en MCO. En este caso, los residuales tienen el mismo peso que en MCO, es decir que se ponderan por 1, como lo muestra el panel B.; el segundo, es un tramo de descenso delimitado por el intervalo  $2c < |\mu_{s_i}| \leq 3c$ , donde la función pondera el efecto de los residuales, disminuyendo su efecto en la medida en que los valores de los residuales se acercan al límite del intervalo (panel B.); y finalmente, existe un tramo de rechazo delimitado por el intervalo  $|\mu_{s_i}| > 3c$ , en el cual los residuales no influyen en el proceso de estimación porque se ponderan por cero, como lo señala el panel B.

Es así, como el parámetro  $c$  permite controlar el sesgo causado por los valores atípicos cuando la distribución no es normal, y permite alcanzar un nivel de confianza deseado, cuando los retornos son normalmente distribuidos. Por lo tanto, para los niveles de confianza equivalentes al 85%, 90%, 95% y 99%, estudios previos de [Martin & Simin \(1999\)](#), [Bailer \(2005\)](#) y [Bailer et. al \(2011\)](#), han determinado valores para  $c$  de 0.868, 0.944, 1.06 y 1.29, respectivamente.

Finalmente, cabe resaltar que para lograr que un estimador  $M$  de localización tenga un grado de robustez adecuado, se deben solucionar básicamente tres problemas ([Wilcox, 2005c](#)): utilizar una función de pérdida adecuada, seleccionar una medida de escala apropiada y aplicar el proceso de estimación más eficiente. Atendiendo a esta necesidad, [Yohai \(1987\)](#) propone el uso del estimador robusto MM que se describe a continuación y que será aplicado en el presente trabajo.

**Estimador robusto MM**

Este estimador incorpora las características de un estimador  $M$  de escala y un estimador  $M$  de localización, con lo cual se garantiza un punto de ruptura alto y una eficiencia relativa apropiada.

Su aplicación consta de las siguientes tres etapas, que se desarrollan efectuando un proceso iterativo con base en el algoritmo de mínimos cuadrados iterativamente ponderados (IRWLS<sup>2</sup>), cuyo desarrollo inicial se atribuye a [Beaton & Tukey \(1974\)](#).

En la primera y segunda etapas, se calcula un estimador inicial  $\hat{\Theta}_0 = (\hat{\alpha}_0, \hat{\beta}_0)$  mediante un modelo de regresión tipo S ([Rousseeuw & Yohai, 1984](#)), que exige el siguiente proceso de minimización:

.....  
 2 IRWLS: Iteratively Reweighted Least Squares

$$\hat{\theta}_0 = \underset{\theta}{\operatorname{argmin}} s(\theta) \quad (6)$$

Donde es una medida robusta de escala con la que se logra un punto de ruptura alto y p los grados de libertad del estimador, la cual se obtiene de resolver la siguiente ecuación utilizando los residuales calculados para :

$$\frac{1}{n-p} \sum_{t=1}^n \rho \left[ \frac{r_{it} - \hat{\alpha}_i - \hat{\beta}_i r_{m_t}}{s(\hat{\theta})} \right] = 0.5 \quad (7)$$

Así, efectuando un proceso de optimización bidimensional, se encuentra  $\hat{\theta}_0$  que minimiza el valor de  $s(\theta)$ , obteniendo a su vez el estimador inicial de escala  $\hat{s}_0 = s(\hat{\theta}_0)$ .

Luego, en la tercera etapa, se calcula el estimador óptimo de localización  $\hat{\theta}_{MM}$ , a partir del proceso de estimación IRWLS, la medida de escala óptima  $\hat{s}_0$  y la función de pérdida óptima  $\rho_0$ . Su valor final se determina mediante el siguiente proceso de minimización, cuyo parámetro inicial de iteración es  $\hat{\theta}_0$ :

$$\hat{\theta}_{MM} = \underset{\theta}{\operatorname{argmin}} \sum_{t=1}^n \rho_0 \left[ \frac{r_{it} - \hat{\alpha}_i - \hat{\beta}_i r_{m_t}}{s(\theta)} \right] \quad (8)$$

La rutina de la estimación de RMM se encuentra disponible en la librería “robustbase package” del software R, y para el presente estudio se aplica un 95% de nivel de confianza.

### 3. METODOLOGÍA

#### Datos

La presente investigación emplea rendimientos de las firmas listadas en los mercados accionarios del MILA como son Chile, Perú y Colombia, durante el periodo muestral comprendido entre enero 1 de 2008 y marzo 28 de 2013, considerando el tiempo disponible en los índices bursátiles de capitalización IPSA, IGBVL y COLCAP, elegidos como proxy del portafolio de cada mercado, respectivamente.

Se seleccionaron acciones que integran los índices bursátiles en cada mercado, garantizando de esta forma la alta bursatilidad y liquidez. Adicionalmente, para evitar sesgo en las estimaciones, se eligieron activos que tuvieran al menos el 80% de los rendimientos diarios diferentes a cero. En el caso de las tasas libres de riesgo, se optó por usar las tasas de interés interbancarias overnight en cada mercado. La información de los precios de las acciones e índices bursátiles fue tomada del sistema de información de Bloomberg, mientras que la información de las tasas interbancarias se tomó de las publicaciones estadísticas de los bancos centrales de cada país: Banco Central de Chile, Banco Central de Reserva del Perú y Banco de la República de Colombia.

En cuanto a la selección del periodo de estimación y la frecuencia de los datos existe una discusión en el escenario académico, debido a que el modelo CAPM considera que la beta es invariante con respecto a la longitud del periodo, lo cual difiere de las investigaciones realizadas por [Hawawini \(1977\)](#), [Smith \(1978\)](#), [Cohen, Hawawini, Maier, Shwartz, & Whitcomb \(1983\)](#), entre otras, que aportan evidencia empírica de que las betas de activos individuales cambian cuando la longitud del intervalo varía. Particularmente, [Daves, Ehrhardt, & Kunkel \(2000\)](#), apoyan el uso de retornos diarios en la estimación de betas dado que tienen menor error de estimación y por tanto ofrecen mayor precisión. En cuanto a la ventana de estimación, este estudio concluye que elegir un periodo de dos a tres años permite lograr una gran reducción del error estándar de estimación.

En este orden de ideas y teniendo en cuenta que la literatura aborda diferentes posturas en cuanto a la selección de la frecuencia de los retornos y ventanas de estimación, la presente investigación examina los resultados de estimación de betas en intervalos de retornos diario y semanal con periodos de uno, dos y tres años, con el fin de establecer cuál medida tiene mayor precisión y en consecuencia, es más apropiada en los mercados objeto de estudio para la toma de decisiones de inversión.

En el análisis se emplearon rendimientos logarítmicos en las frecuencias diarias y semanales, a partir de los precios de cierre de cada día hábil; así mismo, se eliminaron los días festivos para cada país. Para el caso de los rendimientos semanales de los activos, se realizó el cálculo a partir de los precios de cierre de cada jueves, y en la definición de la tasa libre de riesgo de frecuencia semanal, se calculó la tasa efectiva semanal teniendo en cuenta los cinco días hábiles de negociación.

La [Tabla 1](#) resume las variables tomadas en cada caso, incluyendo las firmas seleccionadas:

**Tabla 1. Datos**

País	Tasa libre de riesgo	Índice	Acciones	Nemotécnicos
Colombia	Tasa interbancaria overnight	COLCAP	Cementos Argos, Corporación Financiera Colombiana, Ecopetrol, Éxito, Isa, Isagen, Nutresa, Preferencial Bancolombia	CEMARGOS, CORFICOL, ECOPETROL, ÉXITO, ISA, ISAGEN, NUTRESA, PFBCOLO
Chile	Tasa interbancaria overnight	IPSA	AES Gener, Embotelladora Andina, Antar Chile, Banco de Crédito e Inversiones, Banco Santander Chile, Compañía de Acero del Pacífico, Compañía Cervecería Unidas, Cencosud, Banco de Chile, Colbún, Viña Concha y Toro, Copec, Corpbanca, Endesa, Enersis, Entel, Falabella, Inversiones Aguas Metropolitanas, Lan, Parque Arauco, Ripley, Salfacorp, Sonda, Sociedad Química y Minera, Compañía Sudamericana de Vapores	AESGENER, ANDINAB, ANTAR, BCI, BSAN, CAP, CCU, CENCOSUD, BANCHILE, COLBUN, CONCHA, COPEC, CORPBANC, ENDESA, ENERSIS, ENTEL, FALAB, IAM, LAN, PARAUCO, RIPLEY, SALFACOR, SONDA, SQM/B, VAPORES
Perú	Tasa interbancaria overnight	IGBVL	Credicorp, Compañía de Minas Buenaventura, Casa Grande, Compañía Minera Cerro Verde, Minsur, Volcan Compañía Minera	BAP, BVN, CASAGR, CVERDE, MINSUR, VOLCAB

### Medidas de eficiencia en la estimación

Si bien el uso de la beta histórica de un activo parte de suponer que la beta es constante durante el período de estimación, es amplia la literatura que proporciona evidencia empírica de que la beta de un activo presenta un comportamiento inestable en el tiempo y, que esta variable puede describirse como una serie de tiempo.

En el presente estudio, se hará uso de una técnica de estimación que modela en forma simple un patrón de comportamiento variable del parámetro beta en el tiempo. Dicha técnica es conocida como regresiones móviles con ventana fija (rolling regression), y consiste en realizar regresiones sucesivas bajo un mismo tipo de estimador (normalmente el MCO), empleando en cada regresión un período muestral levemente modificado con relación a la regresión anterior. Para tal efecto, se elimina la observación inicial de la muestra y se adiciona una observación con igual frecuencia al final de la



misma, manteniendo así inalterada la ventana de estimación de cada regresión; la especificación de esta técnica es la siguiente:

$$r_{t,j,i} = \alpha_{j,i} + \beta_{j,i} rm_{t,j} + \mu_{t,j,i} \quad t = 1, \dots, n; j = 1, \dots, z \quad (9)$$

Donde, “i” denota el activo en estudio, “n” corresponde al número de observaciones que componen la ventana móvil dependiendo si se trata de frecuencia diaria o semanal y ventanas de estimación de uno, dos o tres años; y “z” determina el número de regresiones sucesivas que componen la regresión móvil, el cual se tomó como el equivalente a un período de dos años.

El uso de la técnica de regresiones móviles, obedece en primer lugar a que en el proceso de estimación utiliza las betas históricas y, en segundo lugar, a que no disipa la acción de los datos atípicos ya que, al mantener una ventana de estimación con el mismo número de datos, le da igual peso a este tipo de observaciones.

De otro lado, con el propósito de evidenciar la eficiencia de cada estimador en forma independiente, RMM y MCO, se emplearán las siguientes medidas introducidas por Genton & Ronchetti (2008):

$$Q(p)_{i,j} = \left( \frac{1}{n} \sum_{t=1}^n |Y_{ti} - \hat{Y}_{ti}|^p \right)^{1/p} \quad j = 1, 2, \dots, z \quad (10)$$

De aquí, se emplean tres alternativas para la variable p que caracterizan los siguientes estadísticos a utilizar: Q(p=2) corresponde a la raíz del error cuadrático medio (RMSE), Q(p=1) al error absoluto medio (MAE) y Q(p=1/2) define el error absoluto raíz cuadrático medio (STAE).

Estas medidas se usarán para determinar el nivel de error de cada estimación de la regresión móvil, así “n” será el número de observaciones que componen la ventana móvil j (j=1,2,...,z),  $Y_{ti}$  indica el rendimiento observado del activo i, y  $\hat{Y}_{ti}$  define el rendimiento estimado del activo i, con lo cual se logra establecer una distribución de este estadístico para toda la regresión móvil (z datos) y comparar las distribuciones que se obtienen al usar los dos estimadores (RMM y MCO) paralelamente.

Una vez obtenidas estas medidas de eficiencia, se construye un indicador de ganancia relativa de la estimación RMM en comparación con su similar de MCO, a partir de los resultados de la mediana de la distribución de cada uno de los estadísticos de error descritos anteriormente, usando la siguiente expresión:

$$GR(p)_i = \frac{(Med[Q(p)_{j=1}]_{MCO}^2)}{(Med[Q(p)_{j=1}]_{RMM}^2)} \quad (11)$$

De esta forma, se espera que la medida de error sea menor para el estimador RMM, y, en consecuencia, el indicador de ganancia relativa en eficiencia sea mayor a la unidad.

### Cobertura con futuros sobre índices bursátiles

Para ilustrar la relevancia empírica de las betas calculadas, se empleará una estrategia de cobertura con futuros sobre índices, que puede ser usada por administradores de portafolio cuando esperan en el corto plazo, que las condiciones del mercado accionario se deterioren, y, por tanto, deseen reducir la exposición al riesgo sistemático del portafolio.

La estrategia consiste en tomar posiciones cortas en contratos de futuros sobre índices para cubrir un portafolio de acciones, de forma tal, que se reduzca a cero la beta del nuevo portafolio cubierto.

De esta manera, si se presentan movimientos adversos del mercado, se compensa la pérdida que se obtiene en el portafolio de acciones no cubierto con la utilidad de las posiciones en futuros, según la siguiente expresión:

$$\beta_p V_0 + N_F \beta_f m f_0 = \beta_N V_0 \quad (12)$$

Donde  $\beta_p$  corresponde a la beta del portafolio accionario a cubrir,  $V_0$  señala el valor de mercado del portafolio de acciones en el momento de la cobertura,  $N_F$  señala el número de contratos de futuros que deben tomarse para realizar la cobertura,  $\beta_f$  indica la beta de las posiciones de futuros,  $m$  determina el tamaño del contrato del futuro en términos de puntos del índice expresados en unidades monetarias,  $f_0$  indica el precio teórico del futuro al inicio de la estrategia y  $\beta_N$  indica la sensibilidad a los movimientos de mercado del portafolio cubierto o la beta deseada al realizar la estrategia.

Para llevar a cabo lo anterior, se conformará un portafolio para cada país constituido por las acciones más representativas del índice bursátil, asignando la misma ponderación a cada activo, debido a las pocas acciones liquidas encontradas en cada uno de los mercados y con el propósito de evitar soluciones de esquina o concentración del portafolio en unas pocas acciones, dada la condición de los mercados latinoamericanos de ser altamente ineficientes contraria a los fundamentos de la teoría de optimización de portafolios. Posteriormente, se calcularán las betas del portafolio empleando la siguiente forma funcional lineal, a partir de las betas individuales de las acciones estimadas tanto por MCO como por el método robusto RMM:

$$\beta_p = \sum_{i=1}^n \beta_i W_i \quad (13)$$

Donde  $\beta_i$  corresponde a la beta de la acción estimada por MCO y por el método robusto RMM, y  $W_i$  corresponde a la ponderación asignada, que para el caso específico será  $1/n$  donde  $n$  es el número total de acciones del mercado en estudio.

En la [ecuación \(12\)](#), el número de contratos de futuros se estima a partir de la razón de cobertura empleando betas. La ecuación para estimar el número de contratos que permite alcanzar la beta deseada es la siguiente:

$$N_F = \left( \frac{\beta_N - \beta_p}{\beta_f} \right) \left( \frac{V_0}{m f_0} \right) \quad (14)$$

Teniendo en cuenta que el objetivo de la estrategia es disminuir la beta del portafolio cubierto a cero ( $\beta_N = 0$ ),  $N_F$  será negativo, indicando que se debe tomar posición corta en futuros.

Para propósitos del presente estudio se ha adoptado un monto de inversión,  $V_0$ , de mil millones de unidades monetarias para cada mercado, y, un tamaño de contrato de futuros sobre índice ( $m$ ) tomado de las condiciones propias de cada mercado, a excepción de Perú que no posee este tipo de productos, así:  $m_{CHILE} = 3000 \text{ CLP}$ ,  $m_{COL} = 25000 \text{ COP}$  y  $m_{PERÚ} = 1000 \text{ PEN}^3$ .

En cuanto a la variable beta del futuro ( $\beta_f$ ) mide el grado de exposición que tiene la posición en futuros al riesgo sistemático, el cual, teóricamente, converge a uno por la relación existente entre el

.....  
3 Para efectos del presente estudio se adoptó una constante que permitiera escalar el precio del contrato hipotético del futuro sobre el índice IGBVL de Perú, en forma tal, que la cantidad de contratos requeridos en la cobertura fuera de 2 cifras siendo así consistente con sus similares de Chile y Colombia

índice en el mercado de contado y en el mercado de futuros. Para efectos del ejercicio de cobertura que se realiza en la presente investigación, se asume  $\beta_f$  igual a uno.

Con respecto a la variable *Precio del futuro* ( $f_o$ ), para efectos de la verificación empírica, se realizará la proyección del futuro teórico empleando el modelo de valoración de futuros sobre índices. Lo anterior, considerando la reducida información de precios y la poca liquidez de los contratos de futuros en los mercados estudiados.

El precio teórico ( $f_{t,T}$ ) en el tiempo  $t$  de un contrato de futuros sobre índices con vencimiento en un período  $T$ , es igual al costo de oportunidad de conservar una cartera que replique el índice subyacente en el mercado al contado, entre el tiempo  $t$  y  $T$ :

$$f_{t,T} = S_t e^{(r_f - d)(T-t)} \quad (15)$$

Donde  $S_t$  es el valor del índice en el momento  $t$ ,  $r_f$  es la tasa libre de riesgo y  $d$  es la tasa de dividendo que recibiría el tenedor de las acciones que componen el portafolio. Al respecto de esta última variable, su exclusión en trabajos empíricos ha sido generalmente aceptada por los investigadores. Por ejemplo, el trabajo de [Figlewski \(1984\)](#) concluye que la exclusión de los dividendos de la tasa de retorno, parece tener un efecto muy pequeño en el desempeño de la cobertura. Siguiendo este hallazgo, no se tendrán en cuenta los dividendos en el presente estudio.

La anterior expresión considera los siguientes supuestos: en primer lugar, la existencia de igualdad entre la varianza de los retornos en el mercado de contado y la varianza de retornos en el mercado de futuros, y, en segundo lugar, que la tasa de retorno del índice bursátil subyacente y el contrato de futuros están correlacionados positivamente. Sin embargo, aunque se reconoce que los mercados del MILA distan de estos supuestos por la existencia de asimetría en la información y costos de transacción, el uso del modelo teórico de valoración del futuro brinda los elementos necesarios para la verificación comparativa que se pretende realizar.

Se considera un futuro teórico a tres meses, el cual fue valorado cada día hasta su vencimiento para obtener las distintas medidas de desempeño de la cobertura. Tal valoración se realizó diariamente por un periodo de evaluación de dos años, comprendido entre el 30 de diciembre de 2010 y el 28 de diciembre de 2012, con lo cual se obtuvieron series de tiempo de las medidas de desempeño de la cobertura construidas a partir de la reducción del riesgo, la razón de sortino y número de contratos requeridos.

El análisis realizado incluye la comparación de resultados de las medidas de desempeño obtenidas mediante la estrategia de cobertura construida teniendo en cuenta las betas estimadas por MCO y RMM, a partir de ventanas de estimación de uno, dos y tres años.

Con el propósito de identificar el efecto derivado de la crisis de 2008, para la ventana de estimación de tres años, su periodo de evaluación se fraccionó en dos subperíodos, antes y después de que este evento se excluyera de la ventana de estimación de la beta que se usa como parámetro de entrada de la cobertura, subperíodos que se denominaron Pre crisis y Poscrisis. Dicho evento se vio reflejado en los rendimientos atípicos de los activos, ocurridos entre los meses de septiembre y octubre del 2008.

### Medidas de desempeño en la cobertura

El desempeño de la cobertura aplicada usando betas estimadas por MCO y por el método robusto RMM, se evaluó a partir de las siguientes tres medidas.

En primer lugar, se verificó el nivel de reducción de riesgo que se obtuvo al implementar la estrategia empleando las betas estimadas por ambos métodos, con relación a un portafolio descubierto. El indicador se construyó para una ventana trimestral hacia adelante, mediante la evaluación de la volatilidad de los retornos del portafolio cubierto y el no cubierto, según la siguiente expresión ([Butterworth & Holmes, 2001](#)):

$$\% \text{ Reducción riesgo} = \frac{\sigma_{desc}^2 - \sigma_{cub.}^2}{\sigma_{desc}^2} \times 100 \quad (16)$$

Donde  $\sigma_{desc}^2$  señala la varianza de los retornos del portafolio sin cobertura, calculados como la suma equiponderada de los retornos de los activos que componen el portafolio; mientras que  $\sigma_{cub.}^2$  indica la varianza del portafolio cubierto, calculada mediante la siguiente ecuación ([Hull, 2009](#)):

$$\sigma_{cub.}^2 = \sigma_{desc.}^2 + \beta_p^2 \cdot \sigma_f^2 - 2 \cdot \beta_p \cdot COV(desc., f) \quad (17)$$

En esta [ecuación \(17\)](#),  $\sigma_f^2$  indica la varianza de los retornos del contrato de futuros durante el trimestre de vigencia, calculado como la diferencia logarítmica del precio teórico consecutivo del futuro; y  $COV(desc., f)$  señala la covarianza entre los retornos del portafolio no cubierto y los retornos del futuro sobre el índice. El rendimiento del portafolio cubierto se calculó como la sustracción de los rendimientos del portafolio no cubierto y el rendimiento de la cobertura, lo cual es equivalente al rendimiento del futuro multiplicado por la beta del portafolio.

La segunda medida empleada para verificar el desempeño corresponde a la razón de Sortino, la cual de manera similar a la razón de Sharpe, es usada para la selección de portafolios teniendo en cuenta la recompensa o exceso de retorno por unidad de riesgo. Así, cuanto mayor sea el valor de la razón de Sortino, mejor será el portafolio cubierto seleccionado.

De acuerdo con [Estrada \(2006\)](#), la diferencia entre la razón de Sharpe y la razón de Sortino consiste en que el denominador de este último, incluye la semidesviación de las rentabilidades en lugar de la desviación estándar, lo cual permite incorporar la volatilidad negativa (downside risk) que recoge la principal preocupación del inversionista, y no la volatilidad total con respecto al rendimiento medio. Entiéndase esta volatilidad negativa, como el riesgo que afronta el inversionista de no alcanzar un objetivo determinado de rentabilidad (benchmark), el cual para la presente aplicación corresponde a la tasa libre de riesgo. La razón de Sortino es descrita por la siguiente ecuación:

$$\text{Razón de Sortino} = \frac{E[R_t] - B}{\sum_B} \quad (18)$$

Donde, B representa el rendimiento de referencia,  $\sum_B$  representa la semidesviación con respecto al rendimiento de referencia (volatilidad por debajo de dicho rendimiento), formalmente definida como:

$$\sum_B = \sqrt{\left(\frac{1}{n}\right) \cdot \sum_{t=1}^n \{ \min(R_t - B, 0) \}^2} \quad (19)$$

Una vez calculada la razón de Sortino, usando paralelamente la beta estimada por el método MCO y por el método RMM, se construyó un indicador para determinar el incremento de la eficiencia

cuando se realiza o no la cobertura. En este caso, se espera que el indicador supere la unidad, para la estrategia construida con base en la beta estimada tanto por MCO como por RMM. La expresión matemática del indicador es la siguiente:

$$\% \text{ Incremento eficiencia Sortino} = \frac{\text{Razón de Sortino CUB}}{\text{Razón de Sortino NO CUB}} \times 100 \quad (20)$$

Como tercera medida de desempeño, se utilizó el número de contratos requeridos para la cobertura cuando se emplea la beta estimada por MCO o la beta estimada por RMM, la cual está descrita por la [ecuación \(14\)](#).

Así, tomando estos indicadores de desempeño, se construyeron las siguientes tres medidas de ganancia relativa, que permitieron comparar los resultados obtenidos para la estrategia de cobertura basada en betas estimadas por RMM, en comparación con la estrategia de cobertura basada en betas estimadas por MCO, estas son: ganancia relativa en efectividad (reducción de riesgo), ganancia relativa en eficiencia (razón de Sortino) y ganancia relativa en costo (número de contratos). Estos indicadores se describen como:

$$GR (\text{Efectividad}) = \frac{\% \text{ Reducción riesgo}_{\text{MCO}}}{\% \text{ Reducción riesgo}_{\text{RMM}}} \quad (21)$$

$$GR (\text{Eficiencia}) = \frac{\% \text{ Incremento eficiencia Sortino}_{\text{MCO}}}{\% \text{ Incremento eficiencia Sortino}_{\text{RMM}}} \quad (22)$$

$$GR (\text{costo}) = \frac{N_{\text{MCO}}}{N_{\text{RMM}}} \quad (23)$$

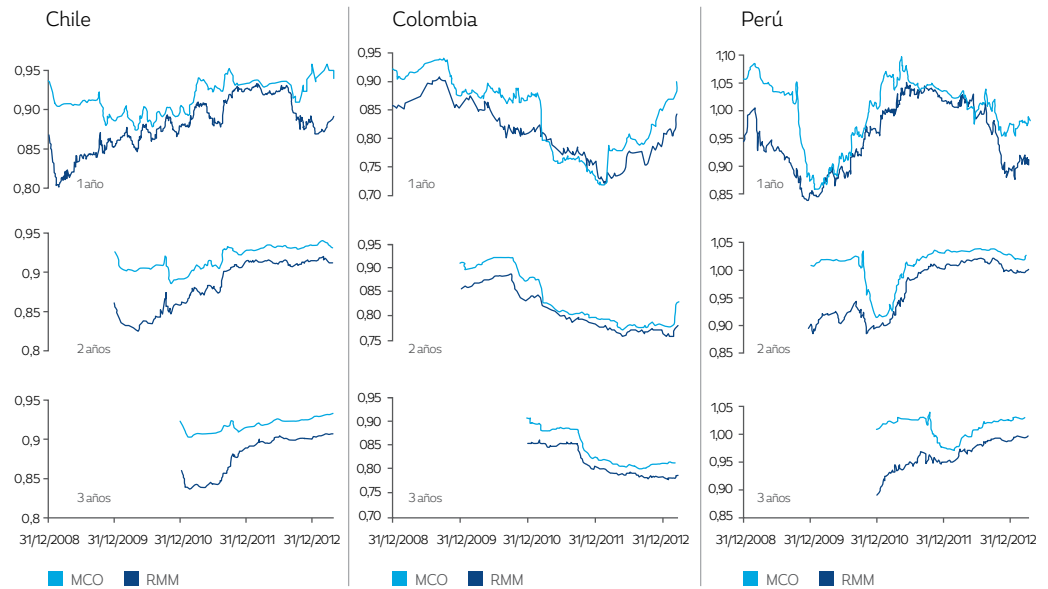
Se espera entonces obtener valores mayores a la unidad para los indicadores de efectividad y costo, y valores menores a la unidad para el indicador de eficiencia, dado que en este último caso se espera un mayor incremento en la eficiencia al utilizar las betas estimadas por el método RMM.

#### 4. RESULTADOS

Al analizar los resultados de las betas obtenidas para los activos individuales y para un portafolio de cada mercado, construido a partir de la participación equiponderada de las acciones que lo constituyen, se encontró que la beta calculada por el método robusto RMM en general, es inferior a la beta calculada por el método MCO. Adicionalmente, se pudo apreciar que cuando existen datos extremos en la ventana de estimación, tal como el evento de la crisis financiera de 2008, la divergencia existente entre las betas de ambos métodos se incrementa de manera sustancial.

Para ilustrar lo anterior, en la [Figura 2](#), se puede visualizar la serie de tiempo de betas estimadas por el método MCO como por el método RMM, con base en una frecuencia diaria de los datos; allí se grafican los resultados obtenidos para cada ventana de estimación: un año, desde el 31 de diciembre de 2008; 2 años, desde el 30 de diciembre de 2009; y finalmente, 3 años, con resultados desde el 29 de diciembre de 2010. Para las tres ventanas de estimación, se obtuvieron resultados hasta el 28 de marzo de 2013. Cuando se incluye en la ventana periodos de alta volatilidad, como es el periodo Pre crisis, las estimaciones por MCO y RMM divergen ampliamente.

**Figura 2. Betas de portafolio estimadas por los métodos MCO y RMM**



A continuación, se presentan los resultados obtenidos, tanto en las medidas de eficiencia de la estimación de las betas como en las estrategias evaluadas de cobertura.

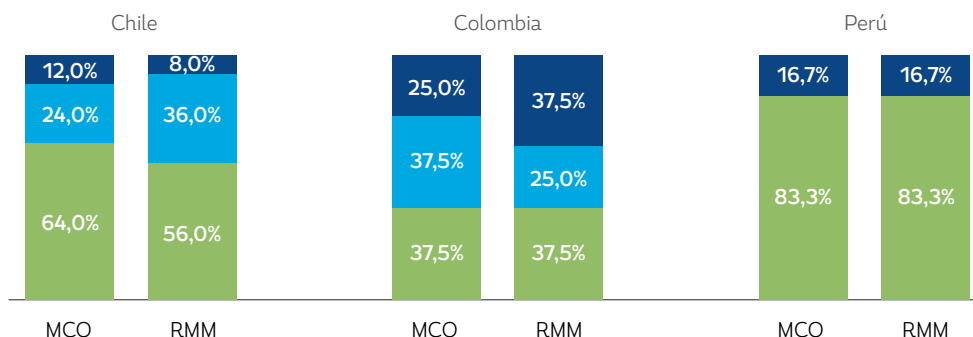
#### Ganancia relativa en eficiencia de la estimación de betas

Inicialmente se evaluó la sensibilidad de los estadísticos de los errores acumulados RMSE, MAE y STAE, con relación al cambio en la longitud de la ventana de estimación de betas. Para tal efecto, se determinó la estructura porcentual del número de activos que presentaron menores errores acumulados en cada ventana de estimación, para cada mercado. Para Chile y Perú predominó un mejor ajuste de los diferentes activos para la ventana de estimación de un año, y se aprecia que la estructura porcentual indicada no presenta variaciones comparativas marcadas entre el uso de los métodos de estimación MCO y RMM.

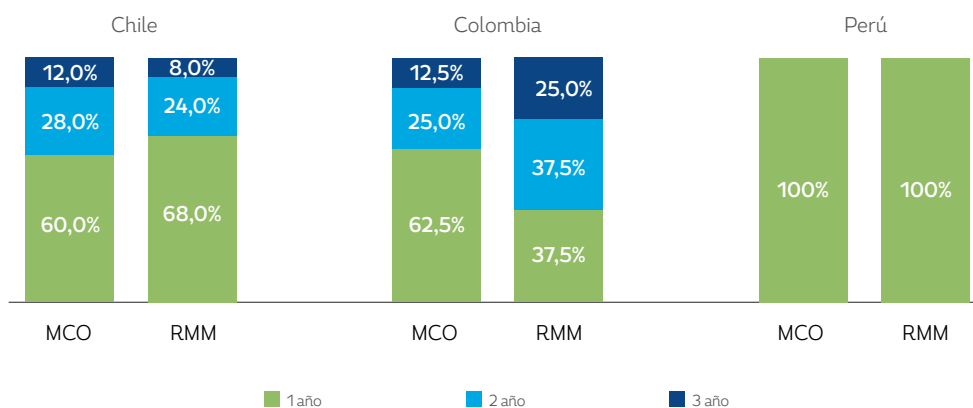
Para Colombia, los resultados no permiten derivar una relación entre la longitud de la ventana de estimación y las medidas de error de las regresiones móviles realizadas tanto por MCO como por RMM. Los resultados anteriores se pueden apreciar en la [Figura 3](#).

**Figura 3. Sensibilidad del RMSE según la longitud de la ventana de estimación**

Frecuencia diaria



Frecuencia semanal



Paralelamente, se evaluó la sensibilidad de los estadísticos de error con respecto a la frecuencia de los datos, obteniendo para los tres mercados del MILA, un menor valor de los estadísticos de error cuando se emplea una frecuencia de datos diaria con ambos métodos de estimación.

De otro lado, los resultados de la medida de ganancia relativa se encuentran en todos los casos por encima de la unidad, alcanzando un incremento de la eficiencia relativa a un valor cercano al 30% en promedio ( $GR=1,30$ ) para frecuencias de datos diarias cuando se usa el estimador RMM en reemplazo del estimador MCO, y un aumento de la eficiencia cercana al 20% en promedio ( $GR=1,20$ ) para frecuencias de datos semanales.

Este resultado soporta el planteamiento según el cual el método RMM permite alcanzar un mejor ajuste de los datos en comparación con el método MCO para los tres mercados del MILA, dado que efectúa un tratamiento ponderado de los datos atípicos de la muestra, tal como se describió anteriormente.

Para mayor ilustración, en la Tabla 2 se presentan los resultados obtenidos para la medida RMSE de los activos pertenecientes a los mercados de Chile, Colombia y Perú, con base en una frecuencia diaria de los datos.

Resultados complementarios para las medidas de eficiencia de la estimación basados en MAE y STAE calculados con base en datos de frecuencias semanales, presentan similar comportamiento. Estos se presentan en el [Apéndice B](#).

**Tabla 2. Ganancia relativa de eficiencia de estimación de betas para datos con frecuencia diaria**

	ACTIVO	RMSE <sub>MCO</sub>			RMSE <sub>RMM</sub>			GR (m) = RMSE <sub>MCO</sub> / RMSE <sub>RMM</sub>		
		1 Año	2 Años	3 Años	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años
COLOMBIA	CEMARGOS	0.012	0.010	0.011	0.010	0.008	0.008	1.201	1.283	1.261
	CORFICOL	0.009	0.010	0.010	0.006	0.007	0.007	1.392	1.543	1.422
	ECOPETROL	0.010	0.009	0.009	0.008	0.007	0.007	1.262	1.262	1.288
	EXITO	0.012	0.013	0.013	0.009	0.009	0.009	1.427	1.454	1.458
	ISA	0.012	0.010	0.010	0.010	0.008	0.008	1.194	1.292	1.270
	ISAGEN	0.010	0.010	0.010	0.007	0.007	0.007	1.359	1.444	1.453
	NUTRESA	0.010	0.010	0.010	0.008	0.008	0.008	1.311	1.324	1.342
	PFBCOLO	0.011	0.011	0.011	0.009	0.009	0.009	1.260	1.203	1.166
CHILE	AESGENER	0.010	0.010	0.011	0.008	0.009	0.009	1.185	1.206	1.185
	ANDINAB	0.012	0.012	0.012	0.010	0.010	0.010	1.178	1.184	1.219
	ANTAR	0.010	0.010	0.011	0.009	0.009	0.009	1.156	1.151	1.125
	BANCHILE	0.010	0.011	0.011	0.008	0.008	0.009	1.282	1.305	1.238
	BCI	0.012	0.012	0.012	0.010	0.010	0.010	1.144	1.175	1.172
	BSAN	0.012	0.013	0.013	0.009	0.010	0.011	1.331	1.305	1.242
	CAP	0.011	0.011	0.012	0.009	0.009	0.010	1.166	1.184	1.210
	CCU	0.013	0.013	0.013	0.011	0.010	0.011	1.162	1.248	1.219
	CENCOSUD	0.012	0.012	0.012	0.010	0.010	0.010	1.243	1.200	1.188
	COLBUN	0.011	0.010	0.011	0.008	0.008	0.009	1.349	1.262	1.211
	CONCHA	0.012	0.013	0.013	0.010	0.011	0.011	1.209	1.215	1.188
	COPEC	0.010	0.010	0.010	0.008	0.008	0.008	1.184	1.210	1.209
	CORPBANC	0.013	0.014	0.013	0.010	0.011	0.010	1.308	1.250	1.295
	ENDESA	0.008	0.007	0.008	0.006	0.006	0.007	1.178	1.155	1.147
	ENERGIS	0.009	0.009	0.010	0.008	0.008	0.008	1.155	1.145	1.288
	ENTEL	0.011	0.011	0.011	0.009	0.008	0.008	1.291	1.403	1.295
	FALAB	0.011	0.011	0.011	0.009	0.009	0.009	1.256	1.240	1.248
	IAM	0.012	0.013	0.013	0.009	0.009	0.009	1.354	1.426	1.423
	LAN	0.011	0.011	0.011	0.008	0.009	0.009	1.248	1.225	1.207
	PARAUCO	0.013	0.013	0.014	0.011	0.011	0.011	1.203	1.191	1.238
RIPLEY	0.013	0.013	0.014	0.010	0.010	0.011	1.311	1.244	1.255	
SALFACOR	0.016	0.016	0.016	0.013	0.013	0.013	1.219	1.228	1.240	
SONDA	0.013	0.013	0.012	0.010	0.010	0.009	1.251	1.286	1.311	
SQM_B	0.011	0.011	0.012	0.009	0.010	0.010	1.197	1.171	1.191	
VAPORES	0.027	0.025	0.024	0.015	0.013	0.013	1.812	1.879	1.820	
PERÚ	BAP	0.015	0.015	0.016	0.013	0.013	0.014	1.134	1.153	1.159
	BVN	0.020	0.021	0.025	0.018	0.018	0.020	1.154	1.172	1.229
	CASAGR	0.014	0.016	0.018	0.010	0.012	0.013	1.338	1.383	1.421
	CVERDE	0.013	0.014	0.015	0.012	0.012	0.012	1.144	1.232	1.275
	MINSUR	0.014	0.014	0.015	0.010	0.011	0.011	1.350	1.322	1.330
	VOLCAB	0.013	0.013	0.015	0.010	0.010	0.012	1.295	1.263	1.212



### Propiedades estadísticas del método robusto en la estimación de betas

Para verificar las propiedades estadísticas de las betas estimadas, se construyeron gráficos que permiten apreciar en primer medida,  $\Delta\beta$  calculado como la diferencia entre las betas obtenidas por el método MCO y por el método RMM, mediante las regresiones móviles para las tres ventanas de estimación (uno, dos y tres años); las series calculadas de curtosis y sesgo de las estimaciones de betas por RMM para cada una de las regresiones móviles, que al graficarlas en una escala temporal contra medidas de normalidad, permitieron encontrar indicios que las estimaciones realizadas por el método RMM para los activos de los mercados del MILA alcanzan niveles de eficiencia similares a los indicados por los fundamentos teóricos. Ver [figura 4](#).

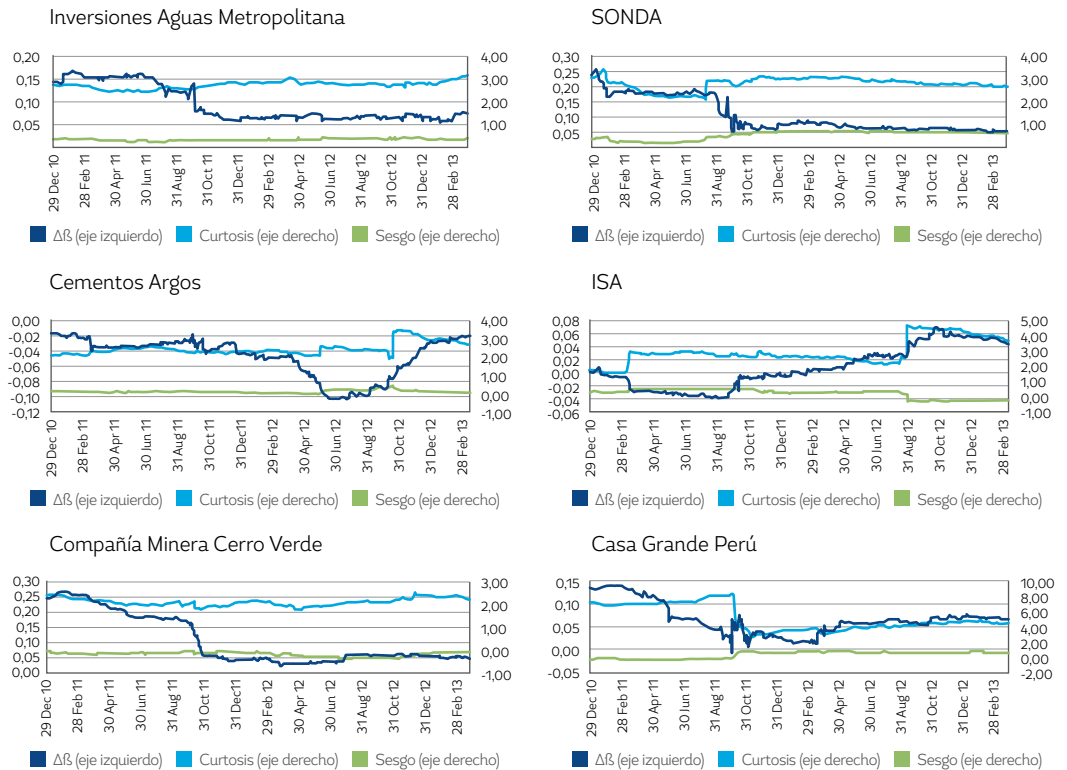
Es decir, cuando los errores de la estimación presentan una distribución cercana a la normal (en ausencia de datos atípicos), el método RMM se comporta tan bien como el método MCO; en este caso cuando la curtosis se mantiene en un valor cercano a 3 y el sesgo se acerca a cero, las betas convergen al mismo valor. Esta condición sólo se pudo apreciar en períodos discretos de algunos activos como Inversiones Aguas Metropolitana, Sonda, Copec y Colbun en el mercado accionario de Chile, ISA y Cementos Argos en Colombia, y Minsur, Compañía Minera Cerro Verde y Casa Grande en Perú.

Alternativamente, bajo la existencia de datos atípicos, e incluso con una pequeña cantidad de ellos, los errores de la estimación por MCO presentan un sesgo en la distribución que es minimizada de una manera óptima por el estimador RMM. Así, cuando la medida de normalidad se distorsiona, con curtosis diferentes a 3 y sesgos diferentes de cero, la diferencia entre betas crece alejándose del valor cero. Para ilustrar esta apreciación, en la Figura 4 se presentan los resultados obtenidos en las estimaciones móviles realizadas para una ventana de tres años con una frecuencia diaria de la muestra, para dos activos de cada mercado; resultados similares se obtuvieron para los demás activos de los tres mercados.

De otro lado, el registro temporal de la diferencia entre las betas MCO y RMM presentados en la [Figura 4](#), evidencia la forma como hacia mediados de septiembre del 2011 se presenta un cambio de tendencia de este indicador en forma generalizada para los activos de Chile y Perú, no siendo notorio para los activos del mercado colombiano, salvo para los casos de las acciones Preferencial Bancolombia y Corficolombiana.

Cambio que es consistente, si tenemos en cuenta que a partir de esta fecha las betas presentan una convergencia en la mayoría de los activos de Chile y Perú, denotando una disminución sustancial en la presencia de datos atípicos y, en consecuencia, de volatilidad para estos mercados. Esto se explica porque a partir del 16 de septiembre de 2011 el evento de la crisis subprime, manifestada en el mercado de EEUU el 16 de septiembre del año 2008, sale de la ventana de estimación de 3 años que fue definida para la regresión móvil. En el caso de las ventanas de estimación de 1 y 2 años este evento pierde relevancia ya que no es tenido en cuenta durante el período de estimación de la regresión móvil. No obstante, en el caso colombiano, si bien este mercado no fue ajeno a la crisis económica mundial de 2008, su impacto no fue de la misma magnitud que para los demás mercados latinoamericanos, evidenciando que los mecanismos de transmisión de volatilidad no fueron tan intensos como en el caso de Chile y Perú, debido a la normatividad del sector financiero y a la efectividad de las medidas de política económica de Colombia, entre otros aspectos.

**Figura 4. Relación entre beta, curtosis y sesgo de regresiones móviles**



**Ganancia relativa en estrategias de cobertura usando betas**

Al evaluar la medida de reducción del riesgo de la cobertura, obtenida cuando se usan betas estimadas por MCO o por RMM, se puede evidenciar que por ambos métodos se genera una alta reducción de riesgo con respecto al portafolio descubierto; por lo cual, para ambos métodos, se observan valores cercanos al 90% para Chile y al 80% para Colombia y Perú.

Lo anterior, conlleva a una ganancia relativa cercana a 1 volviendo indiferente el empleo de uno u otro método cuando se evalúa la efectividad de la cobertura. Tal situación se presenta tanto en el período de evaluación completo, como cuando se evalúa en períodos fraccionados en pre crisis y poscrisis.

Al evaluar el segundo indicador de desempeño, la ganancia relativa en eficiencia (Razón de Sortino) presenta resultados ambiguos cuando se evalúa para el período completo. Así, si se considera el mayor valor porcentual que se obtiene al usar RMM en comparación con MCO, Chile presenta una pérdida de efectividad que crece con la ventana de estimación de la beta, obteniendo así valores entre 1,15% y 5,13%. Colombia en cambio, presenta una ganancia en efectividad para ventanas de estimación de uno y dos años entre 18,03% y 6,88%, y una pérdida del 3,20% efectividad para una ventana de estimación de tres años. Adicionalmente, para el caso de Perú se presenta una ganancia en efectividad que crece con la ventana de estimación entre 3,27% y 18,43%. Se aclara que las pérdidas y ganancias se calculan con valor base 100.

Por su parte, cuando se evalúan los periodos aislados se encuentra lo siguiente: para el periodo pre crisis, los resultados muestran de manera consistente que cuando se emplean betas robustas estimadas a partir de una ventana de estimación que contiene datos atípicos (Pre crisis), se obtienen mayores

ganancias relativas en la Razón de Sortino; y en el periodo poscrisis (sin datos atípicos), los resultados señalan que no pareciera haber ganancias relativas al emplear la beta estimada por el método robusto.

En este sentido, para el mercado chileno se obtiene una ganancia relativa en la razón de sortino en el periodo pre crisis del 25,52% y una pérdida relativa del 16,68% en el período poscrisis. Por su parte, para Colombia se encontró una ganancia relativa del 3,96% para el periodo pre crisis y una pérdida relativa del 6,51% en el período poscrisis. De otro lado, en Perú se obtienen ganancias relativas para los períodos de pre crisis y poscrisis de 22,50% y 16,42%, respectivamente, lo cual puede derivarse de las características de este mercado en términos de su tamaño y profundidad, que agudizan la presencia de datos atípicos.

En términos de la ganancia relativa en números de contratos de cobertura, se evidencia que ésta siempre es positiva indiferentemente de la ventana de estimación de beta y del período de evaluación que se tome; es decir, la beta robusta deriva en un menor número de contratos a adquirir para la cobertura, en comparación a la valoración realizada usando la beta calculada por MCO, aspecto que se hace más notorio en el período de pre crisis que en el de poscrisis. En las [Tablas 3 y 4](#), se presenta un resumen de los resultados obtenidos en el ejercicio de cobertura.

**Tabla 3. Ganancia relativa en eficiencia y efectividad de cobertura con futuros sobre índices**

MEDIDA DE EFICIENCIA	COLOMBIA			CHILE			PERÚ		
	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años
% RED. RIESGO <sub>MCO</sub>	80.93%	80.64%	80.40%	92.38%	92.39%	92.37%	81.21%	81.00%	81.42%
% EFECT. SORTINO <sub>MCO</sub>	1.95%	4.27%	36.66%	-26.28%	-0.16%	-1.94%	-13.33%	-6.47%	-7.35%
# CONTRATOS <sub>MCO</sub>	18.55	18.96	20.07	70.87	70.71	69.95	48.32	49.37	48.43
% RED. RIESGO <sub>RMM</sub>	80.61%	80.61%	80.57%	92.29%	92.18%	92.01%	81.40%	81.24%	81.53%
% EFECT. SORTINO <sub>RMM</sub>	-4.58%	-6.03%	19.98%	-0.81%	-0.53%	2.72%	-3.04%	-0.99%	10.23%
# CONTRATOS <sub>RMM</sub>	18.30	18.62	19.26	68.80	69.26	67.78	47.16	48.21	46.54
% GR. RED. RIESGO	100.12%	100.03%	99.70%	100.06%	100.09%	100.25%	99.97%	99.95%	100.11%
% GR. EFECT. SORTINO	81.97%	93.12%	103.20%	101.15%	103.25%	105.13%	96.73%	96.90%	81.57%
% GR. # CONTRATOS	101.75%	101.56%	103.30%	101.92%	102.28%	102.90%	101.77%	102.37%	103.71%

**Tabla 4. Ganancia relativa en eficiencia y efectividad de cobertura con futuros sobre índices en el periodo pre crisis y poscrisis, con una ventana de tiempo de 3 años.**

MEDIDA DE EFICIENCIA	COLOMBIA			CHILE			PERÚ		
	COMPLETO	PRECRISIS	POSCRISIS	COMPLETO	PRECRISIS	POSCRISIS	COMPLETO	PRECRISIS	POSCRISIS
% RED. RIESGO <sub>MCO</sub>	80.40%	81.62%	79.79%	92.37%	96.86%	91.24%	81.42%	91.19%	77.85%
% EFECT. SORTINO <sub>MCO</sub>	36.66%	69.69%	21.09%	-1.94%	34.52%	-55.80%	-7.35%	-33.52%	6.48%
# CONTRATOS <sub>MCO</sub>	20.07	20.78	18.86	69.95	64.86	71.44	48.43	48.68	48.41
% RED. RIESGO <sub>RMM</sub>	80.57%	82.64%	80.03%	92.01%	96.42%	91.12%	81.53%	90.56%	77.75%
% EFECT. SORTINO <sub>RMM</sub>	19.98%	44.14%	3.47%	2.72%	68.14%	-43.56%	10.23%	11.09%	6.72%
# CONTRATOS <sub>RMM</sub>	19.26	19.88	18.46	67.78	60.10	69.57	46.54	44.59	46.87
% GR. RED. RIESGO	99.70%	98.77%	100.36%	100.25%	100.92%	100.18%	100.11%	100.85%	99.86%
% GR. EFECT. SORTINO	103.20%	96.04%	106.51%	105.13%	74.48%	116.68%	81.57%	77.50%	83.58%
% GR. # CONTRATOS	103.30%	104.03%	102.59%	102.90%	107.77%	102.74%	103.71%	109.21%	103.08%

## CONCLUSIONES

El presente trabajo de investigación aporta evidencia de la influencia que ejercen los datos atípicos en la estimación de betas por MCO en los mercados del MILA, indicando una superioridad estadística de las betas estimadas por RMM ante la presencia de los mismos. Entre los resultados que apoyan esta tesis, se menciona que el método RMM permitió una ganancia relativa del orden del 30% en el nivel de ajuste de los datos en comparación con el método de estimación MCO, siendo consistente con los principios teóricos del método.

De otro lado, los resultados del estudio revelan diferencias entre las betas por MCO y las betas por RMM en algunos periodos, en particular, en aquellos donde se presenta turbulencias de los mercados financieros y alta volatilidad. Tal es el caso de la crisis de 2008, el cual, al ser incluido en la ventana de estimación de la beta, adicionó datos atípicos a las series de rendimientos de los mercados, en mayor magnitud para los casos de Chile y Perú. Al correr la ventana y excluir el valor extremo que desató la crisis en el tercer trimestre de 2008, dichas diferencias se mitigaron e inclusive llegaron a ser casi nulas en algunos activos cuando el periodo no mostraba sobresaltos. Es así, como pudo evidenciarse en la muestra analizada que para los mercados integrantes del MILA, los principios de eficiencia relativa del método RMM sobre el MCO se satisfacen. De esta manera, cuando la distribución es normal el ajuste se torna tan eficiente como el que se obtiene con el método MCO, y cuando no es normal la eficiencia en el ajuste es muy superior.

Específicamente, el presente trabajo presenta evidencia de la mayor efectividad que puede obtenerse cuando se aplican estrategias de cobertura con futuros sobre índices, empleando betas estimadas por RMM, cuando se trata de periodos de alta volatilidad donde existe una alta influencia de los datos atípicos en la distribución (periodo pre crisis), llegando a valores de eficiencia relativa del 25,52% para el mercado de Chile, del 3,96% para el mercado de Colombia y del 22,5% para el mercado de Perú.

Lo anterior, sumado a la ganancia relativa obtenida en el valor de los contratos de futuros a adquirir, permite afirmar que para periodos de crisis el uso del método robusto RMM deriva en una relación

beneficio/costo eficiente, que sería complementaria al uso del método MCO para aquellos períodos de baja volatilidad.

Para finalizar, reconociendo las limitaciones que tiene la aplicación del modelo MCO para la estimación de beta en los mercados MILA, y la importancia que tiene para los agentes del mercado obtener medidas de betas confiables y robustas que se ajusten de manera más precisa a la realidad del mercado, futuras investigaciones podrán continuar explorando alternativas y áreas de aplicación de la estadística robusta en Finanzas. De esta manera, futuras investigaciones podrían extender las aplicaciones de los métodos de estadística robusta en: el cálculo de portafolios óptimos bajo condiciones de valores extremos, y cálculo de razones de cobertura óptimas bajo criterios de mínima varianza, estimación de modelos multifactoriales de valoración de activos.

## REFERENCIAS

- Annema, A., & Goedhart, M. H. (2003). A better beta. *The McKinsey Quarterly*(1), 6-9.
- Bailer, H. M. (2005). Robust Estimation of Factor Models in Finance Ph. D. University of Washington.
- Bailer, H. M., Maravina, T. A., & Martin, R. D. (2011). Robust Betas in Asset Management. *The Oxford Handbook of Quantitative Asset Management*, 203-242.
- Beaton, A. E., & Tukey, J. W. (1974). The Fitting of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data. *Technometrics*, 16(2), 147-185.
- Bekaert, G., Erb, C., Harvey, C. R., & Viskanta, T. E. (1998). Distributional Characteristics of Emerging Markets Returns and Asset Allocation. *The Journal of Portfolio Management*, 102-116.
- Blume, M. (1971). On the Assessment of Risk. *The Journal of Finance*, 26(1), 1-10.
- Blume, M. (1975). Betas and Their Regression Tendencies. *The Journal of Finance*, 30(3), 785-795.
- Blume, M. E. (1979, Marzo). Betas and their Regression Tendencies: Some Further Evidence. *The Journal of Finance*, XXXIV(1), 265-267.
- Bowie, D. C., & Bradfield, D. J. (1998). Robust Estimation of Beta Coefficients: Evidence from a small Stock Market. *Journal of Business Finance & Accounting*, 439-454.
- Butterworth, D., & Holmes, P. (2001). The hedging effectiveness of stock index futures: evidence for the FTSE-100 y FTSE-mid250 indexes traded in the UK. *Applied Financial Economics*, 11, 57-68.
- Celik, S. (2013). Testing the Stability of Beta: A Sectoral Analysis in Turkish Stock Market. *Journal of Economics and Behavioral Studies*. (ISSN: 2220-6140), 5(1), 18-23.
- Chan, L. K., & Lakonishok, J. (1992, Junio). Robust Measurement. *The Journal of Financial and Quantitative Analysis*, 27(2), 265-282.
- Cohen, K. J., Hawawini, G. A., Maier, S. F., Shwartz, R. A., & Whitcomb, D. R. (1983). Estimating and Adjusting for the Intervalling-Effect Bias in Beta. *Management Science*, 29(1), 135-148.
- Cornell, B., & Dietrich, J. K. (1978). Mean-Absolute-Deviation Versus Least-Squares Regression Estimation of Beta Coefficients. *Journal of Financial and Quantitative Analysis*, 13, 123-131.
- Daves, P. R., Ehrhardt, M. C., & Kunkel, R. A. (2000). Estimating Systematic Risk: The Choice of Return Interval and Estimation Period. *Journal of Financial and Strategic Decisions*, 13(1), 7-13.
- Elgers, P. T., Haltiner, J. R., & Hawthorne, W. H. (1979, Marzo). Beta Regression Tendencies: Statistical and Real Causes. *The Journal of Finance*, XXXIV(1), 261-263.
- Estrada, J. (2006). Downside Risk in practice. *Journal of Applied Corporate Finance*, 18(1), 117-125.
- Estrada, J., & Vargas, M. (2012). Black Swan, Beta, Risk, and Return. *Journal of Applied Finance*(2), 77-89.
- Fama, E. (1965). The Behavior of Stock- Market Prices. *Journal of Business*, 38(1), 34-105.
- Fama, E., & French, K. (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance*(47), 427-465.
- Faraway, J. J. (2004). *Linear Models with R*. United Kingdom: Chapman & Hall/CRC Text in Statistical Science.

- Fernández, P., & Bermejo, V. J. (2009, Septiembre). *B=1 does a better job than calculated betas*. Working Paper WP-825. (U. o. IESE Business School, Ed.) Retrieved Enero 26, 2013, from <http://www.iese.edu/research/pdfs/DI-0825-E.pdf>
- Fernández, P., & Bermejo, V. J. (2009). Betas utilizadas por Directivos y Profesores: encuesta europea 2009. *Revista Española de Capital de Riesgo*, 3, 55-78.
- Figlewski, S. (1984). Hedging performance and basis risk in stock index futures. *Journal of Finance*(39), 657-669.
- Fong, W. M. (1997). Robust Beta Estimation: Some Empirical evidence. *Review of Financial Economics*, 6(2), 167-186.
- Genton, M. G., & Ronchetti, E. (2008). Robust Prediction of Beta. *Computational Methods in Financial Engineering*, 147-161.
- Giraldo Gómez, N. (2005, enero-junio). Predicción de Betas y VaR de Portafolios de Acciones mediante el Filtro de Kalman y los Modelos Garch. *Cuadernos de Administración*, 18(29), 103-119.
- Graham, J. R., & Harvey, C. R. (2001). The theory and practice of corporate finance: evidence from the field. *Journal of Financial Economics*(60), 187-243.
- Gray, S., Hall, J., & Klease, D. (2006). Bias, Stability and Predictive Ability in the Measurement of Systematic Risk. *UQ Business School*, 1-40.
- Hampel, F. (1968). *Contribution to the theory of robust estimation*. Berkeley: PhD thesis, University of California.
- Harris, R., & Shen, J. (2003). Robust Estimation of the Optimal Hedge Ratio. *The Journal of Futures Markets*, 24(8), 799-816.
- Hawawini, G. A. (1977). *On the Time Behavior of Financial Parameters: An Investigation of the Intervalling Effect*. New York : Ph D. Dissertation, New York University.
- Hooper, V. J., Ng, K., & Reeves, J. J. (2008). Quarterly beta forecasting: An evaluation. *International Journal of Forecasting*(24), 480-489.
- Huber, P. J. (1964). Robust estimation of a local parameter. *The Annals of Mathematical Statistics*(35), 73-101.
- Huber, P. J. (1973). Robust regression: asymptotics, conjectures and Monte Carlo. *The Annals of Statistics*(1), 799-821.
- Huber, P. J. (1981). *Robust Statistics* (Primera Edición ed.). New York: Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc.
- Hull, J. C. (2009). *Options, Futures, and other Derivatives* (Seventh Edition ed.). New Jersey, United States of America: Pearson Prentice Hall.
- Jiménez Triviño, J. A. (2012). *Betas Contables*. Tesis de Maestría Finanzas Corporativas. Colegio de Estudios Superiores en Administración CESA. Retrieved Enero 2013, 15, from <http://repository.cesa.edu.co/bitstream/10726/534/1/TMFO0237.pdf>
- Knez, P. J., & Ready, M. J. (1997). On the Robustness of Size and Book-to-Market in Cross-Sectional Regressions. *The Journal of Finance*, 52(4), 1355-1382.
- Koenker, R., & Bassett Jr, G. (1978, Enero). *Regression Quantiles*. *Econometrica*, 46(1), 33-50.
- Lally, M. (1998). An Examination of Blume and Vasicek Betas. *Financial Review*, 33(3), 183-198.
- Lee, C.-F., Wang, K., & Chen, Y. L. (2009). Hedging and Optimal Hedge Ratios for International Index Futures Markets. *Review of Pacific Basin Financial Markets and Policies*, 12(4), 593-610.

- Lintner, J. (1965). The Valuation of Risk Asset and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *Review of Economics and Statistics*, 47(1), 13-37.
- Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*(36), 394-419.
- Marin, J. M., & Rubio, G. (2011). El comportamiento del riesgo Beta. In *Economía Financiera* (pp. 389-412). Antoni Bosch.
- Maronna, R. A., Martin, R. D., & Yohai, V. J. (2006). *Robust Statistics: Theory and Methods*. England: John Wiley & Sons Ltd.
- Martin, R. D., & Simin, T. (1999). Estimates of Small - Stock Betas are Often Very Distorted by Outliers. (U. o. Washington, Ed.) *Technical Report*(351), 1-19.
- Martin, R. D., & Simin, T. (1999, Marzo). Robust estimation of beta. (U. o. Department of Statistics, Ed.) *Technical Report*(No. 350), 1-40.
- Martin, R. D., & Simin, T. T. (2003, Sep. - Oct.). Outliers-Resistant Estimates of Beta. *Financial Analysts Journal*, Vol. 59(No. 5), 55-69.
- Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34(4), 768-783.
- Pereiro, L. A. (2010, Autumn ). The Beta Dilemma in Emerging Markets. (U. T. Tella, Ed.) *Journal of Applied Corporate Finance*, Vol. 22(Issue 4), 110-122.
- Quintana, F. A., Iglesias, P. L., & Galea-Rojas, M. (2005). Bayesian robust estimation of systematic risk using product partition models. *Applied Financial Economics Letters*, 1(5), 313-320.
- Reeves, J. J., & Wu, H. (2013). Constant versus Time-Varying Beta Models: Further Forecast Evaluation. *Journal of Forecasting*(32), 256-266.
- Rodríguez, W. R. (2009, Julio-Diciembre). Análisis del efecto día de semana en los mercados accionarios Latinoamericano. *Lecturas de Economía*(71), 189-208.
- Roll, R. (1988). R2. *Journal of finance*, 43(3), 541-566.
- Rousseeuw, P. J., & Yohai, V. (1984). Robust Regression by Means of S estimators. In W. H. J. Franke (Ed.), *Robust and Nonlinear Time Series Analysis* (pp. 256-274). New York: Lecture Notes in Statistics 26, Springer Verlag.
- Rousseeuw, P. J. (1984). Least Median of Squares Regression. *Journal of the American Statistical Association*(79), 871-880.
- Sarker, M. R. (2013). Forecast Ability of the Blume's and Vasicek's Technique: Evidence from Bangladesh. *Journal of Business and Management*, 9(6), 22-27.
- Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *Journal of Finance*, 19(3), 425-442.
- Sharpe, W. F. (1971). Mean-Absolute-Deviation Characteristic Lines for Securities and Portfolios. *Management Science*, 18(2), B1-B13.



- Smith, R. (1978). The Effect of Intervaling on Estimating Parameters of the Capital Asset Pricing Model. *Journal Financial Quantitative Analysis*, 13, 313-344.
- Tukey, J. W. (1960). A survey of sampling from contaminated distributions. In S. G. I. Olkin (Ed.), *Contribution to Probability and Statistics* (pp. 448-485). Stanford: Stanford University Press.
- Vasicek, O. (1973). A Note on Using Cross - Sectional Information in Bayesian Estimation of Security Betas. *Journal of Finance*(28), 1.233-1.239.
- Wilcox, R. R. (2005a). "A Foundation for Robust Methods". In *Introduction to Robust Estimation and Hypothesis Testing. Segunda Edición* (pp. 19-38). San Diego, USA: Elsevier Academic Press.
- Wilcox, R. R. (2005b). "Robust Regression". In *Introduction to Robust Estimation and Hypothesis Testing. Segunda edición* (pp. 413-466). San Diego, USA: Elsevier Academic Press.
- Wilcox, R. R. (2005c). Estimating Measures of Location and Scale. In R. Wilcox, *Introduction to Robust Estimation and Hypothesis Testing. Segunda edición* (pp. 43-102). San Diego, USA: Elsevier Academic Press.
- Yeo, J. (2001). Modelling Time-Varying Systematic Risk in Australia. *International Congress on Modelling & Simulation, Western Australia, Modelling & Simulation Society of Australia & New Zealand Inc*(3), 1565-1570 .
- Yohai, V. J. (1987). High breakdown - point and high efficiency robust estimates for regression. *The Annals of Statistics*, 15(20), 642-656.
- Yohai, V., Stahel, W., & Zamar, R. A. (1991). A procedure for robust estimation and inference in linear regression. (a. S. In W.A. Stahel, Ed.) *Directions in Robust Statistics and Diagnostics*. New York, NY:Springer, Part II.
- Zamar, R. H. (1994). Estimación Robusta. *Estadística Española*, 36(137), 327-387.

## APÉNDICE A

Expresiones matemáticas para las funciones óptimas  $\rho_0$ ,  $\psi$  y  $\omega$

$$\rho_0(\mu_{s,t}; c) = \begin{cases} 3,25 \cdot c^2 & |(\frac{\mu_{s,t}}{c})| > 3 \\ c^2 \cdot [1,792 - 0,972 \cdot (\frac{\mu_{s,t}}{c})^2 + 0,423 \cdot (\frac{\mu_{s,t}}{c})^4 - 0,052 \cdot (\frac{\mu_{s,t}}{c})^6 - 0,002 \cdot (\frac{\mu_{s,t}}{c})^8], & 2 < |(\frac{\mu_{s,t}}{c})| \leq 3 \\ 0,5 \cdot \mu_{s,t} & |(\frac{\mu_{s,t}}{c})| \leq 2 \end{cases}$$

$$\psi_0(\mu_{s,t}; c) = \begin{cases} 0 & |(\frac{\mu_{s,t}}{c})| > 3 \\ c^2 \cdot [-1,944 \cdot (\frac{\mu_{s,t}}{c}) + 1,728 \cdot (\frac{\mu_{s,t}}{c})^3 - 0,312 \cdot (\frac{\mu_{s,t}}{c})^5 - 0,016 \cdot (\frac{\mu_{s,t}}{c})^7], & 2 < |(\frac{\mu_{s,t}}{c})| \leq 3 \\ 0,5 \cdot \mu_{s,t} & |(\frac{\mu_{s,t}}{c})| \leq 2 \end{cases}$$

$$\omega(\mu_{s,t}; c) = \psi_0(\mu_{s,t}; c) / \mu_{s,t}$$

## APÉNDICE B

En el presente apéndice se muestran las tablas de la ganancia relativa de eficiencia de estimación de betas empleando las medidas de error RMSE, MAE y STAE para los datos con frecuencia diaria y semanal. Se exceptúa la tabla RMSE con datos con frecuencia diaria dado que se incorporó en el numeral 4

**Tabla 1 B. Ganancia relativa de eficiencia de estimación de betas para datos con frecuencia semanal, utilizando el RMSE**

ACTIVO	RMSE <sub>MCO</sub>			RMSE <sub>RMM</sub>			GR(m) = RMSE <sub>MCO</sub> / RMSE <sub>RMM</sub>			
	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años	
COLOMBIA	CEMARGOS	0,020	0,021	0,022	0,020	0,019	0,019	1,011	1,068	1,136
	CORFICOL	0,015	0,019	0,019	0,015	0,017	0,016	1,019	1,133	1,170
	ECOPETROL	0,020	0,020	0,020	0,017	0,016	0,016	1,194	1,255	1,257
	EXITO	0,024	0,024	0,025	0,020	0,020	0,021	1,182	1,250	1,196
	ISA	0,020	0,018	0,018	0,014	0,014	0,014	1,497	1,322	1,251
	ISAGEN	0,016	0,016	0,018	0,016	0,016	0,016	1,019	1,032	1,158
	NUTRESA	0,021	0,020	0,020	0,019	0,019	0,018	1,088	1,077	1,135
	PFBICOLO	0,017	0,020	0,020	0,015	0,016	0,018	1,166	1,246	1,087
CHILE	AESGENER	0,020	0,022	0,023	0,017	0,019	0,020	1,186	1,202	1,159
	ANDINAB	0,026	0,026	0,026	0,023	0,024	0,024	1,122	1,086	1,063
	ANTAR	0,020	0,019	0,019	0,020	0,018	0,018	1,005	1,067	1,074
	BANCHILE	0,020	0,023	0,025	0,014	0,017	0,019	1,404	1,391	1,339
	BCI	0,023	0,026	0,027	0,022	0,022	0,024	1,052	1,202	1,127
	BSAN	0,025	0,025	0,026	0,021	0,021	0,022	1,203	1,175	1,182
	CAP	0,022	0,024	0,028	0,022	0,023	0,024	1,016	1,050	1,167
	CCU	0,028	0,028	0,028	0,025	0,025	0,026	1,107	1,089	1,100
	CENCOSUD	0,020	0,025	0,026	0,018	0,021	0,023	1,089	1,177	1,133
	COLBUN	0,021	0,021	0,022	0,018	0,018	0,019	1,195	1,132	1,152
	CONCHA	0,026	0,027	0,027	0,024	0,024	0,024	1,082	1,100	1,114
	COPEC	0,018	0,019	0,019	0,016	0,016	0,016	1,127	1,144	1,158
	CORPBANC	0,026	0,027	0,027	0,024	0,025	0,025	1,061	1,081	1,089
	ENDESA	0,017	0,016	0,017	0,014	0,014	0,014	1,163	1,159	1,200
	ENERGIS	0,018	0,018	0,020	0,017	0,017	0,017	1,034	1,064	1,142
	ENTEL	0,022	0,023	0,022	0,017	0,017	0,017	1,348	1,312	1,238
	FALAB	0,025	0,026	0,026	0,019	0,019	0,019	1,359	1,397	1,317
	IAM	0,022	0,023	0,022	0,020	0,019	0,019	1,114	1,205	1,157
	LAN	0,023	0,024	0,025	0,020	0,021	0,022	1,154	1,152	1,112
	PARAUCO	0,025	0,027	0,029	0,024	0,024	0,025	1,040	1,124	1,175
RIPLEY	0,027	0,028	0,028	0,021	0,022	0,024	1,283	1,257	1,165	
SALFACOR	0,035	0,036	0,038	0,028	0,028	0,031	1,231	1,267	1,249	
SONDA	0,024	0,025	0,024	0,019	0,020	0,019	1,246	1,233	1,223	
SQM_B	0,022	0,023	0,026	0,021	0,022	0,023	1,075	1,038	1,120	
VAPORES	0,057	0,058	0,056	0,036	0,033	0,033	1,583	1,729	1,682	
PERU	BAP	0,031	0,031	0,033	0,028	0,029	0,029	1,090	1,082	1,108
	BVN	0,042	0,045	0,049	0,040	0,044	0,049	1,055	1,027	1,001
	CASAGR	0,030	0,036	0,042	0,026	0,029	0,031	1,164	1,236	1,353
	CVERDE	0,027	0,028	0,032	0,023	0,024	0,025	1,171	1,188	1,300
	MINSUR	0,028	0,029	0,033	0,024	0,026	0,028	1,140	1,117	1,165
	VOLCAB	0,028	0,028	0,033	0,023	0,023	0,025	1,228	1,213	1,294

**Tabla 2 B. Ganancia relativa de eficiencia de estimación de betas para datos con frecuencia diaria, utilizando el MAE**

ACTIVO	MAE <sub>MCO</sub>			MAE <sub>RMM</sub>			GR(m) = MAE <sub>MCO</sub> / MAE <sub>RMM</sub>			
	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años	
COLOMBIA	CEMARGOS	0,009	0,008	0,008	0,008	0,006	0,007	1,143	1,190	1,181
	CORFICOL	0,006	0,007	0,007	0,005	0,005	0,006	1,290	1,347	1,266
	ECOPETROL	0,007	0,007	0,007	0,006	0,006	0,006	1,175	1,175	1,195
	EXITO	0,009	0,009	0,009	0,007	0,007	0,007	1,325	1,340	1,316
	ISA	0,009	0,008	0,008	0,008	0,006	0,007	1,135	1,185	1,172
	ISAGEN	0,007	0,007	0,007	0,006	0,005	0,005	1,258	1,298	1,308
	NUTRESA	0,007	0,008	0,007	0,006	0,006	0,006	1,223	1,230	1,246
	PFBCOLO	0,008	0,008	0,008	0,007	0,007	0,007	1,196	1,133	1,112
CHILE	AESGENER	0,007	0,008	0,008	0,007	0,007	0,007	1,085	1,146	1,124
	ANDINAB	0,009	0,009	0,009	0,008	0,008	0,008	1,121	1,134	1,167
	ANTAR	0,008	0,008	0,008	0,007	0,007	0,008	1,101	1,101	1,088
	BANCHILE	0,007	0,008	0,008	0,006	0,007	0,007	1,152	1,188	1,149
	BCI	0,009	0,009	0,009	0,008	0,008	0,008	1,098	1,107	1,114
	BSAN	0,009	0,009	0,010	0,007	0,008	0,008	1,289	1,216	1,163
	CAP	0,008	0,008	0,009	0,007	0,007	0,008	1,112	1,127	1,150
	CCU	0,010	0,010	0,010	0,009	0,008	0,009	1,109	1,170	1,154
	CENCOSUD	0,009	0,009	0,010	0,008	0,008	0,008	1,178	1,155	1,137
	COLBUN	0,008	0,008	0,008	0,006	0,007	0,007	1,226	1,165	1,131
	CONCHA	0,009	0,010	0,010	0,008	0,009	0,009	1,125	1,131	1,136
	COPEC	0,007	0,007	0,007	0,007	0,006	0,006	1,115	1,148	1,136
	CORPBANC	0,010	0,010	0,010	0,008	0,009	0,008	1,216	1,186	1,223
	ENDESA	0,006	0,006	0,006	0,005	0,005	0,005	1,117	1,115	1,102
	ENERSIS	0,007	0,007	0,007	0,006	0,006	0,006	1,121	1,103	1,117
	ENTEL	0,008	0,008	0,008	0,007	0,006	0,006	1,219	1,295	1,198
	FALAB	0,008	0,008	0,008	0,007	0,007	0,007	1,187	1,158	1,157
	IAM	0,009	0,009	0,009	0,007	0,007	0,007	1,294	1,312	1,320
	LAN	0,008	0,008	0,008	0,007	0,007	0,007	1,173	1,146	1,143
	PARAUACO	0,010	0,010	0,011	0,009	0,009	0,009	1,153	1,144	1,188
RIPLEY	0,009	0,010	0,010	0,008	0,008	0,009	1,219	1,165	1,178	
SALFACOR	0,012	0,012	0,012	0,011	0,010	0,011	1,137	1,152	1,159	
SONDA	0,009	0,010	0,009	0,008	0,008	0,007	1,168	1,210	1,224	
SQM_B	0,008	0,009	0,009	0,008	0,008	0,008	1,128	1,120	1,131	
VAPORES	0,016	0,015	0,015	0,012	0,010	0,010	1,375	1,490	1,431	
PERU	BAP	0,012	0,012	0,012	0,011	0,011	0,011	1,095	1,100	1,104
	BVN	0,015	0,016	0,018	0,014	0,014	0,016	1,116	1,117	1,166
	CASAGR	0,010	0,011	0,013	0,008	0,009	0,010	1,230	1,248	1,297
	CVERDE	0,010	0,011	0,011	0,009	0,009	0,009	1,098	1,164	1,200
	MINSUR	0,010	0,010	0,011	0,008	0,009	0,009	1,268	1,223	1,247
	VOLCAB	0,009	0,010	0,011	0,008	0,008	0,010	1,181	1,162	1,148

**Tabla 3 B. Ganancia relativa de eficiencia de estimación de betas para datos con frecuencia semanal, utilizando el MAE**

ACTIVO	MAE <sub>MCO</sub>			MAE <sub>RMM</sub>			GR(m) = MAE <sub>MCO</sub> / MAE <sub>RMM</sub>			
	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años	
COLOMBIA	CEMARGOS	0,017	0,016	0,017	0,017	0,016	0,016	1,005	1,022	1,087
	CORFICOL	0,012	0,015	0,015	0,012	0,013	0,013	1,054	1,120	1,102
	ECOPETROL	0,015	0,015	0,015	0,014	0,013	0,013	1,148	1,182	1,177
	EXITO	0,018	0,018	0,019	0,016	0,015	0,017	1,131	1,187	1,150
	ISA	0,015	0,013	0,014	0,011	0,011	0,011	1,403	1,181	1,185
	ISAGEN	0,013	0,013	0,014	0,013	0,013	0,013	1,022	1,019	1,112
	NUTRESA	0,016	0,016	0,016	0,015	0,015	0,014	1,043	1,048	1,091
	PFBCOLO	0,013	0,015	0,015	0,012	0,012	0,015	1,143	1,180	1,060
CHILE	AESGENER	0,016	0,017	0,018	0,014	0,014	0,016	1,143	1,145	1,110
	ANDINAB	0,020	0,021	0,021	0,019	0,019	0,020	1,062	1,071	1,039
	ANTAR	0,017	0,015	0,016	0,016	0,014	0,015	1,003	1,039	1,069
	BANCHILE	0,014	0,017	0,018	0,011	0,013	0,014	1,258	1,295	1,266
	BCI	0,018	0,019	0,020	0,018	0,017	0,019	1,045	1,105	1,076
	BSAN	0,019	0,019	0,020	0,016	0,017	0,017	1,173	1,116	1,142
	CAP	0,018	0,020	0,022	0,018	0,019	0,020	1,004	1,050	1,103
	CCU	0,023	0,022	0,022	0,021	0,021	0,021	1,083	1,048	1,064
	CENCOSUD	0,016	0,019	0,020	0,015	0,017	0,018	1,059	1,114	1,102
	COLBUN	0,016	0,016	0,017	0,014	0,014	0,015	1,187	1,129	1,142
	CONCHA	0,021	0,021	0,021	0,020	0,020	0,020	1,037	1,059	1,073
	COPEC	0,014	0,015	0,015	0,013	0,013	0,013	1,120	1,105	1,130
	CORPBANC	0,021	0,022	0,021	0,019	0,020	0,020	1,088	1,061	1,068
	ENDESA	0,013	0,012	0,013	0,011	0,011	0,011	1,141	1,142	1,146
	ENERSIS	0,015	0,014	0,015	0,014	0,014	0,014	1,082	1,041	1,060
	ENTEL	0,017	0,017	0,016	0,014	0,014	0,014	1,238	1,248	1,174
	FALAB	0,018	0,019	0,019	0,015	0,015	0,016	1,226	1,275	1,219
	IAM	0,018	0,018	0,017	0,016	0,015	0,016	1,111	1,141	1,112
	LAN	0,018	0,019	0,020	0,016	0,017	0,019	1,113	1,101	1,057
	PARAUCO	0,020	0,021	0,022	0,019	0,019	0,020	1,040	1,101	1,110
RIPLEY	0,019	0,021	0,021	0,016	0,017	0,019	1,168	1,195	1,126	
SALFACOR	0,026	0,027	0,029	0,022	0,022	0,024	1,180	1,197	1,175	
SONDA	0,018	0,019	0,018	0,015	0,016	0,015	1,154	1,186	1,174	
SQM_B	0,018	0,018	0,021	0,017	0,018	0,019	1,052	1,028	1,083	
VAPORES	0,040	0,038	0,038	0,029	0,026	0,026	1,389	1,460	1,445	
PERU	BAP	0,024	0,025	0,025	0,023	0,024	0,024	1,063	1,052	1,071
	BVN	0,033	0,036	0,039	0,032	0,036	0,039	1,035	1,017	1,001
	CASAGR	0,024	0,027	0,030	0,021	0,023	0,024	1,167	1,172	1,247
	CVERDE	0,021	0,022	0,024	0,018	0,019	0,020	1,150	1,150	1,209
	MINSUR	0,021	0,022	0,025	0,019	0,021	0,023	1,117	1,075	1,105
	VOLCAB	0,021	0,021	0,024	0,018	0,018	0,019	1,144	1,143	1,236

**Tabla 4 B. Ganancia relativa de eficiencia de estimación de betas para datos con frecuencia diaria, utilizando el STAE**

ACTIVO	STAE <sub>MCO</sub>			STAE <sub>RMM</sub>			GR(m) = STAE <sub>MCO</sub> / STAE <sub>RMM</sub>			
	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años	
COLOMBIA	CEMARGOS	0,007	0,006	0,006	0,006	0,005	0,005	1,137	1,168	1,173
	CORFICOL	0,005	0,005	0,006	0,004	0,004	0,005	1,284	1,353	1,250
	ECOPETROL	0,006	0,006	0,006	0,005	0,005	0,005	1,152	1,154	1,177
	EXITO	0,007	0,007	0,007	0,005	0,005	0,006	1,341	1,334	1,321
	ISA	0,007	0,006	0,006	0,006	0,005	0,005	1,126	1,160	1,147
	ISAGEN	0,006	0,006	0,006	0,005	0,004	0,004	1,211	1,282	1,287
	NUTRESA	0,006	0,006	0,006	0,005	0,005	0,005	1,186	1,205	1,234
	PFBCOLO	0,007	0,007	0,007	0,006	0,006	0,006	1,176	1,132	1,102
CHILE	AESGENER	0,006	0,007	0,007	0,006	0,006	0,006	1,045	1,122	1,107
	ANDINAB	0,008	0,007	0,007	0,007	0,007	0,006	1,117	1,113	1,152
	ANTAR	0,006	0,006	0,007	0,006	0,006	0,006	1,088	1,086	1,078
	BANCHILE	0,006	0,006	0,006	0,005	0,005	0,006	1,129	1,160	1,130
	BCI	0,007	0,008	0,008	0,007	0,007	0,007	1,080	1,096	1,100
	BSAN	0,007	0,007	0,008	0,006	0,006	0,007	1,290	1,206	1,148
	CAP	0,007	0,007	0,008	0,006	0,006	0,007	1,099	1,114	1,138
	CCU	0,008	0,008	0,008	0,007	0,007	0,007	1,101	1,152	1,146
	CENCOSUD	0,007	0,008	0,008	0,006	0,007	0,007	1,180	1,150	1,130
	COLBUN	0,006	0,006	0,007	0,005	0,006	0,006	1,206	1,144	1,114
	CONCHA	0,007	0,008	0,008	0,007	0,007	0,007	1,109	1,113	1,138
	COPEC	0,006	0,006	0,006	0,006	0,005	0,005	1,115	1,139	1,118
	CORPBANC	0,008	0,008	0,008	0,007	0,007	0,007	1,221	1,184	1,202
	ENDESA	0,005	0,005	0,005	0,004	0,004	0,004	1,101	1,107	1,093
	ENERGIS	0,006	0,006	0,006	0,005	0,005	0,005	1,116	1,091	1,094
	ENTEL	0,007	0,006	0,006	0,005	0,005	0,005	1,221	1,286	1,181
	FALAB	0,007	0,007	0,007	0,006	0,006	0,006	1,193	1,144	1,143
	IAM	0,007	0,007	0,007	0,006	0,005	0,005	1,331	1,313	1,326
	LAN	0,006	0,006	0,007	0,005	0,006	0,006	1,163	1,131	1,129
	PARAUACO	0,008	0,008	0,009	0,007	0,007	0,007	1,137	1,142	1,203
RIPLEY	0,008	0,008	0,008	0,006	0,007	0,007	1,197	1,154	1,165	
SALFACOR	0,010	0,010	0,010	0,009	0,009	0,009	1,106	1,136	1,149	
SONDA	0,008	0,008	0,007	0,007	0,006	0,006	1,166	1,198	1,218	
SQM_B	0,007	0,007	0,008	0,006	0,006	0,007	1,110	1,113	1,117	
VAPORES	0,012	0,012	0,011	0,009	0,008	0,008	1,345	1,436	1,355	
PERU	BAP	0,009	0,010	0,010	0,009	0,009	0,009	1,075	1,088	1,096
	BVN	0,013	0,013	0,015	0,011	0,012	0,013	1,131	1,111	1,155
	CASAGR	0,008	0,009	0,010	0,007	0,007	0,008	1,208	1,231	1,281
	CVERDE	0,008	0,009	0,009	0,008	0,008	0,008	1,091	1,154	1,192
	MINSUR	0,008	0,008	0,009	0,007	0,007	0,007	1,251	1,212	1,240
	VOLCAB	0,008	0,008	0,009	0,007	0,007	0,008	1,177	1,151	1,142

**Tabla 5 B. Ganancia relativa de eficiencia de estimación de betas para datos con frecuencia semanal, utilizando el STAE**

ACTIVO	STAE <sub>MCO</sub>			STAE <sub>RMM</sub>			GR(m) = STAE <sub>MCO</sub> / STAE <sub>RMM</sub>			
	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años	1 Año	2 Años	3 Años	
COLOMBIA	CEMARGOS	0,015	0,014	0,014	0,015	0,014	0,013	1,009	1,008	1,066
	CORFICOL	0,010	0,013	0,012	0,010	0,011	0,011	1,040	1,122	1,088
	ECOPETROL	0,013	0,013	0,013	0,012	0,011	0,011	1,114	1,160	1,155
	EXITO	0,015	0,015	0,015	0,014	0,012	0,013	1,125	1,181	1,144
	ISA	0,012	0,011	0,011	0,009	0,009	0,009	1,380	1,149	1,180
	ISAGEN	0,011	0,011	0,011	0,011	0,011	0,011	1,029	1,016	1,075
	NUTRESA	0,013	0,013	0,013	0,013	0,013	0,012	1,044	1,046	1,086
	PFBCOLO	0,011	0,012	0,013	0,010	0,010	0,012	1,166	1,151	1,057
CHILE	AESGENER	0,013	0,013	0,015	0,011	0,012	0,013	1,137	1,145	1,092
	ANDINAB	0,018	0,018	0,017	0,016	0,016	0,017	1,098	1,086	1,034
	ANTAR	0,014	0,012	0,013	0,014	0,012	0,012	1,003	1,033	1,053
	BANCHILE	0,011	0,014	0,015	0,009	0,011	0,012	1,210	1,268	1,243
	BCI	0,015	0,015	0,016	0,014	0,015	0,015	1,033	1,063	1,062
	BSAN	0,016	0,016	0,016	0,013	0,014	0,014	1,210	1,118	1,140
	CAP	0,016	0,017	0,018	0,016	0,016	0,017	1,011	1,056	1,076
	CCU	0,019	0,018	0,019	0,018	0,018	0,018	1,075	1,032	1,050
	CENCOSUD	0,014	0,016	0,017	0,014	0,015	0,016	1,035	1,097	1,092
	COLBUN	0,013	0,014	0,014	0,011	0,012	0,012	1,237	1,157	1,157
	CONCHA	0,018	0,018	0,018	0,017	0,017	0,017	1,022	1,042	1,067
	COPEC	0,011	0,012	0,012	0,011	0,011	0,011	1,078	1,086	1,135
	CORPBANC	0,017	0,018	0,018	0,016	0,017	0,017	1,075	1,065	1,070
	ENDESA	0,011	0,010	0,010	0,009	0,009	0,009	1,128	1,155	1,131
	ENERSIS	0,013	0,012	0,013	0,012	0,012	0,012	1,114	1,039	1,040
	ENTEL	0,014	0,014	0,013	0,012	0,011	0,012	1,167	1,243	1,168
	FALAB	0,015	0,015	0,016	0,013	0,012	0,013	1,195	1,248	1,196
	IAM	0,016	0,015	0,015	0,014	0,013	0,013	1,083	1,110	1,111
	LAN	0,015	0,016	0,017	0,014	0,015	0,016	1,113	1,106	1,047
	PARAUCO	0,017	0,018	0,019	0,016	0,016	0,017	1,033	1,098	1,118
RIPLEY	0,015	0,016	0,017	0,013	0,014	0,015	1,161	1,180	1,123	
SALFACOR	0,022	0,022	0,023	0,018	0,018	0,020	1,203	1,202	1,168	
SONDA	0,015	0,015	0,015	0,013	0,013	0,013	1,145	1,202	1,194	
SQM_B	0,015	0,016	0,018	0,014	0,015	0,016	1,053	1,037	1,080	
VAPORES	0,032	0,030	0,030	0,023	0,021	0,021	1,404	1,399	1,419	
PERU	BAP	0,021	0,021	0,021	0,019	0,020	0,020	1,101	1,047	1,064
	BVN	0,028	0,030	0,033	0,027	0,030	0,033	1,020	1,008	1,001
	CASAGR	0,021	0,022	0,025	0,017	0,019	0,019	1,221	1,169	1,270
	CVERDE	0,017	0,018	0,019	0,015	0,016	0,016	1,153	1,154	1,184
	MINSUR	0,017	0,018	0,021	0,015	0,017	0,019	1,133	1,057	1,099
	VOLCAB	0,017	0,017	0,019	0,015	0,015	0,016	1,131	1,128	1,237

---

SISTEMA VIAL TERCIARIO Y  
EDUCACIÓN RURAL EN ANTIOQUIA  
(COLOMBIA): UN MODELO PROBIT  
ORDENADO Y MULTIVARIADO

The Rural Roads Impact on  
Education Performance in  
Antioquia (Colombia): an  
ordered probit model

Guillermo David Hincapie Velez, Ivan Montoya  
Gomez y John Jaime Bustamante

---



Research Article

## SISTEMA VIAL TERCIARIO Y EDUCACIÓN RURAL EN ANTIOQUIA (COLOMBIA): UN MODELO PROBIT ORDENADO Y MULTIVARIADO

### The Rural Roads Impact on Education Performance in Antioquia (Colombia): an ordered probit model.

Guillermo David Hincapie Velez<sup>a\*</sup>, Ivan Montoya Gomez<sup>b</sup>, John Jaime Bustamante<sup>c</sup>

**Palabras clave:** Logro educativo; sistema de vías terciarias; modelos probit y logit ordenados; políticas educativas

**Keywords:** Educational performance; Rural Vial System; Probit and Logit Ordered Models; Education

**Clasificación IEL:** R1, I21, C01, C25

**Received:** 05/05/2017

**Accepted:** 12/06/2017

**Published:** 20/06/2017

#### Resumen

El presente artículo tiene el propósito de determinar el impacto del sistema vial del departamento de Antioquia (Colombia), en particular su sistema de vías terciarias, sobre la probabilidad de ascenso en el logro educativo de estudiantes de básica secundaria que viven en zonas rurales del departamento. Para tal efecto, se estima un modelo probit ordenado del rendimiento educativo en la prueba de matemáticas, definiendo, como una de las variables independientes, una transformación de las densidades viales terciarias en relación con la oferta de alcantarillado, energía, servicio de teléfono, entre otros, con el fin de capturar la variabilidad y dispersión de las zonas rurales en cada municipio. Los resultados muestran que las vías terciarias inciden sobre la probabilidad de ascender en el rendimiento educativo para los estudiantes rurales.

#### Abstract

This paper aims to determine the impact of the Antioquia department's road system, particularly the rural roads system, on the probability of advancement in the educational achievement of students from rural areas in high schools in Antioquia, Colombia. For this we propose an ordered logistic model of educational performance, defining as a independent variable a transformation of the rural road densities in relation to the supply of sewage, energy, telephone service, among others, with the aim of capturing the variability and dispersion of rural areas in each municipality. The results show that rural roads influence the probability of increasing educational performance for rural students.

a, b, c. Universidad Pontificia Bolivariana, Medellín, Colombia.

\* Autor para correspondencia:  
guillermo.hincapie@upb.edu.do

## 1. Introducción

Las áreas rurales y urbanas de países en vías de desarrollo, suelen presentar diferencias significativas en términos del desarrollo económico y de la calidad de vida. Por lo tanto, estas diferencias han constituido un objeto de estudio frecuente en la literatura económica de la cual se han derivado recomendaciones de política ([Shilpa Aggarwal, 2013](#)). Entre las diferencias exploradas en la literatura y señaladas como determinantes comunes a la configuración de círculos viciosos de pobreza, se encuentran: la problemática de empleo y la productividad laboral ([Reardon, 1997](#)), ([Reardon, Berdegué, & Escobar, 2001](#)), la inseguridad y la violencia ([Osgood & Chambers, 2000](#)) y ([Spencer & Bryant, 2000](#)), la reducida oferta institucional del Estado en la provisión de bienes públicos ([Castro-Leal, Dayton, Demery, & Mehra, 1999](#)), y el bajo desempeño de las empresas rurales ([Datt & Ravallion, 1998](#)) y ([de Janvry & Sadoulet, 2000](#)).

La nueva geografía económica junto con las teorías de las aglomeraciones económicas, sugieren como factores comunes a estas diferencias entre lo rural y lo urbano, la falta de acceso y la distancia a los centros económicos ([Ribot & Peluso, 2003](#)); por lo tanto, las vías rurales constituirían, en gran parte, dicho factor común. En términos generales, una mayor red de vías rurales permite el abaratamiento de los costos de transporte, lo cual impacta de distintas formas las posibilidades de mayor desarrollo económico-social de las áreas rurales ([Duranton, 2015](#)). En efecto, una mayor red de vías rurales permite una mayor conectividad de las familias a los centros de mercado mejorando sus posibilidades de consumo, también facilita el flujo de trabajadores a centros urbanos, eleva los flujos migratorios, posibilita una mayor presencia del Estado y mejora la productividad de las empresas rurales producto del acceso a los canales de inversión y del crédito ([Campos, 2017](#)), ([Vasconcellos, 1997](#)) y ([Huisman & Smits, 2009](#)).

Las vías rurales tienen también un impacto sobre las posibilidades de educación de la población rural de países en vías de desarrollo, ya que la mayor oferta educativa de calidad se encuentra en los centros económicos o en los llamados cascos urbanos municipales<sup>1</sup>, haciendo que los estudiantes tengan que incurrir en largos viajes para acceder a educación en la medida en que la población rural esté muy dispersa y lejana de dichos centros ([Campos, 2017](#)) y ([Falch, Lujala, & Strøm, 2013](#)). Una mayor red de vías terciarias implica, entonces, una mayor conectividad de los estudiantes rurales con los centros educativos en la medida en que sea paralela a una reducción de costos de transporte, así como oferta de transporte público; permitiendo así que estos estudiantes puedan experimentar un mayor rendimiento educativo y, además, una mayor tasa de permanencia escolar. La literatura al respecto ha establecido que una mayor inversión en infraestructura vial rural tiene efectos positivos sobre una reducción de brechas entre lo rural y urbano, en especial respecto a las posibilidades del capital humano, ([Villar & Ramírez, 2014](#)).

A pesar de que la literatura sobre la relación entre vías rurales y desarrollo económico en general es relativamente amplia, es escasa en cuanto al estudio concreto del impacto de la infraestructura rural sobre el logro educativo de estudiantes rurales. La mayoría de trabajos se concentran en estudiantes urbanos ([Falch et al., 2013](#)), ([Jeremy Atack & Elisabeth Perlman, 2012](#)) y ([Baschieri & Falkingham, 2006](#)) y, y además en estudiantes universitarios ([Gibbons & Vignoles, 2012](#)). Respecto al caso de estudiantes rurales de básica secundaria, se destacan los trabajos de ([Campos, 2017](#)), donde se realiza un análisis, a nivel de municipios en Colombia, sobre el efecto de las densidades viales terciarias municipales sobre un indicador de rendimiento promedio municipal de educación, considerando variables binarias de ruralidad municipal; por otro lado, el trabajo de ([Baschieri & Falkingham, 2006](#)), tiene un propósito

1 Por lo general en el caso de la educación básica secundaria más que en la educación primaria.

similar pero adelanta una comparación entre países de la región de América del Sur aunque su escala de análisis es a nivel de municipio. Estos trabajos concluyen, en general, que una mayor densidad vial tiene un impacto directo sobre el logro y rendimiento educativo de estudiantes de las áreas rurales, así como un impacto inverso sobre las tasas de deserción estudiantil.

Cuando se pretende estudiar el impacto de la distancia, o de los costos de transporte y tiempos de viaje que están asociados a la infraestructura vial, sobre el rendimiento educativo, la necesidad de contar con datos geo-referenciados que den cuenta cabal de dicha distancia se torna en uno de los obstáculos principales para llevar a cabo investigaciones de este tipo. Este problema habitualmente se agrava cuando el investigador se concentra en zonas rurales donde por lo general la adquisición de datos es significativamente más compleja. Es por esta razón, que en general la literatura ha abordado este problema desarrollando análisis más agregados, como a nivel de países, Estados o Municipios. Sin embargo, con tales niveles de agregación se torna difícil capturar las heterogeneidades subyacentes del logro educativo de estudiantes rurales que en general exhiben altos niveles de dispersión en términos de su ubicación en el espacio rural, implicando potenciales sesgos en las estimaciones realizadas. El aporte del presente trabajo va en dicha línea y en el caso concreto de la literatura en Colombia, constituye un aporte adicional al trabajo realizado por ([Campos, 2017](#)).

En el presente artículo se estima el impacto de las densidades viales sobre la probabilidad de ascenso en el logro o rendimiento educativo de estudiantes rurales en el departamento de Antioquia (Colombia), para el año 2012. Dado que no se cuenta con información geo-referenciada respecto a la distancia de los estudiantes a los cascos urbanos, se construye en primer lugar un indicador proxy de las densidades viales rurales teniendo en cuenta la información socioeconómica a nivel de micro-dato aportado por el *Instituto Colombiano para la Evaluación de la Educación* —en adelante ICFES—, de tal manera que puedan ser capturadas las heterogeneidades espaciales subyacentes que son imputables a datos socioeconómicos, para en segundo lugar estimar un modelo probit ordenado multivariado. La elección de un modelo probit responde, en lo fundamental, a que las brechas regionales de los estudiantes en Colombia han presentado históricamente una distribución multimodal, de tal manera que dicho modelo nos aproxima también a estas rigideces distribucionales, en mayor grado presentes en las áreas rurales estable ([Quintero, Leandro, & Hincapié Vélez, 2016](#)).

## 2. Metodología

Con el fin de estimar la probabilidad de que un estudiante ascienda en su rendimiento educativo, fueron consideradas las siguientes fases de un procedimiento general: en primer lugar, fueron recopilados los datos de densidades viales terciarias junto con datos relacionados al logro educativo en la base de datos SABER 11 del ICFES, base de datos a nivel micro que contiene información socioeconómica de los estudiantes. En segundo lugar, se estimó un modelo probit ordenado sobre estudiantes rurales con el fin de aproximarse al efecto promedio de la conectividad vial rural sobre el logro educativo.

### Datos

En este artículo fueron consideradas como variables del rendimiento educativo, a los resultados arrojados por los estudiantes de básica secundaria en las llamadas pruebas estatales conocidas como SABER 11 realizadas por el ICFES; el logro de los estudiantes en estas pruebas, se divide en 6 áreas del conocimiento que van desde Matemáticas y Lenguaje, hasta Filosofía y Biología. Al tener una estructura de micro dato, la base de datos contiene información relativa a las condiciones socioeconómicas del estudiante, así como información referente a su colegio; además, la base de datos distingue a los

estudiantes que residen en los cascos urbanos de sus municipios, de los estudiantes que residen en las zonas rurales. Para este estudio, se utilizó como variable del logro educativo los resultados en la prueba de Matemáticas debido a sus estrechas correlaciones con los resultados de las pruebas restantes, como se muestra en los anexos y como en general se practica en la literatura

Como variable indicativa de las vías, fueron consideradas las densidades viales de los municipios del departamento de Antioquia para el año 2012, información reportada por el Instituto Agustín Gadaži- IGAC- y construida por el Ministerio de Minas del país. Esta densidad es construida computando la razón entre el número de kilómetros de vías y el área en kilómetros del municipio en el cual se encuentran. En el país, son consideradas tres tipos de vías: las vías primarias, que constituyen las autopistas que conectan a los municipios con los centros económicos del país; por su parte se encuentran las vías secundarias, que conectan a los municipios entre si y, por último, las redes de vías terciarias que tienen el propósito de conectar a los municipios con sus veredas y zonas rurales. De esta manera, las densidades viales se calculan para cada tipo de vía en cada municipio. Las vías, como se mostrará más adelante, guardan estrechas correlaciones espaciales entre ellas, de tal manera que estas redes conforman todo un sistema de vías.

### Especificación del Modelo Probit Ordenado

Pretender estimar el impacto del sistema de vías sobre el logro educativo de los estudiantes presenta un problema inicial debido a que la base de datos del ICFES es de corte transversal alrededor de cada estudiante, en tanto que sólo se dispone de información agregada a nivel de municipio sobre las densidades viales. Esta falta de información detallada sobre la ubicación concreta de los estudiantes y por lo tanto la distancia de su hogar al colegio supone desde luego un problema. Además, no se dispone de información referente a la calidad de las vías terciarias, aspecto también crucial para aproximarse a los tiempos de viaje y por lo tanto los costos de transporte conexos al desplazamiento. El trabajo de (Campos, 2017) evita este problema definiendo un indicador promedio del rendimiento educativo por municipio y, de esta manera, lleva las conclusiones del estudio a escala de municipio. Aunque este derrotero es aceptable en términos generales, si se pretenden estudiar las dinámicas rurales sin tener en cuenta esta heterogeneidad se aumentan las probabilidades de incurrir en sesgos.

Con el fin, entonces, de aproximarse a la heterogeneidad *intra-municipal* de los estudiantes en relación con las vías terciarias, se adelanta una reconsideración del indicador de densidad vial terciaria, adaptándolo con respecto a variables control indicativas de la dispersión y de la lejanía relativa del estudiante a los cascos urbanos. Concretamente, se parte de la ecuación (1),

$$\widehat{d}_{3i} = d_{3i} (\text{Telefono}_i + \text{Internet}_i + \text{vehiculo}_i) \quad (1)$$

Donde el término  $\widehat{d}_{3i}$  hace referencia al nuevo indicador de densidad vial terciaria,  $\text{Telefono}_i$  es una variable dummy que toma el valor de 1 si la casa del estudiante tiene servicio de teléfono,  $\text{Internet}_i$  es una variable dummy que indica 1 si la casa del estudiante tiene servicio de internet y, por último,  $\text{vehiculo}_i$  contiene el número de vehículos con que cuenta la familia del estudiante. La elección de estas variables responde a dos criterios: en primer lugar, el criterio de lejanía, debido a que en las viviendas rurales más cercanas al casco urbano, es más probable encontrar servicio de teléfono y servicio de internet en las viviendas; en segundo lugar, el criterio de calidad de las vías terciarias, al que nos aproximamos considerando la variable de vehículos de la familia, que si bien es indicativa del nivel

de ingresos familiar, también se aproxima al estado relativo de las vías terciarias, ya que, en un caso extremo, las familias rurales no podrían tener vehículos-automóviles- en ausencia de vías terciarias<sup>2</sup>.

Las aglomeraciones económicas se configuran a partir de la interacción conjunta entre fuerzas centrípetas, como la necesidad de ubicarse en el casco urbano, y las fuerzas centrífugas que tienen que ver en general con los precios de la tierra los cuales tienden al alza debido a la competencia por un recurso escaso. De esta manera, la ruralidad más pobre tendería a ubicarse más lejana de las áreas urbanas. Si se acepta que el objetivo de una red vial es la de reducir los tiempos de viaje y abaratar los costos de transporte de los estudiantes, se desprende entonces que una mayor densidad vial afecta al rendimiento educativo. De tal suerte, las dimensiones socioeconómicas y de infraestructura vial que tiene el indicador propuesto permite ir en esta línea.

Con base pues en la ecuación (1), se estimó un modelo probit ordenado para la probabilidad de que un estudiante ascienda en su rendimiento educativo, ecuación (2), de acuerdo a factores familiares, socioeconómicos, -matriz X-, y el estado de su conectividad con el sistema vial rural,  $-\widehat{d}_{3i}$ .

$$p_i = \alpha_0 + \beta X + \delta \widehat{d}_{3i} + e_i = \alpha_0 + \beta \widehat{X} + e_i \quad (2)$$

Donde el término  $\widehat{X}$  constituye la matriz de regresores teniendo en cuenta el nuevo indicador de densidad vial. De esta manera, la variable latente tomará los rangos establecidos por los cuartiles de la distribución de logro educativo en la prueba de matemáticas para el año 2012, de tal manera que

$$y_i = j \quad \text{si } a_{j-1} < y_i^* \leq a_j \quad (3)$$

La probabilidad de que el evento “i” se seleccione en la alternativa “j” será la de la ecuación (4), una vez se tenga en cuenta la diferencia en términos de las densidades de distribución de probabilidad de un modelo probit.

$$p_{ij} = p(y_i = j) = p(a_{j-1} < y_i^* \leq a_j) = F(a_j - \widehat{X}_i \beta) - F(a_{j-1} - \widehat{X}_i \beta) \quad (4)$$

El efecto, entonces, de las variables regresoras sobre la probabilidad de que un estudiante ascienda en los cuartiles de la distribución del logro educativo deberá calcularse nuevamente de acuerdo a los efectos marginales computados mediante la expresión (5),

$$\frac{\partial p_{ij}}{\partial X_{ri}} = (F'(a_j - \widehat{X}_i \beta) - F'(a_{j-1} - \widehat{X}_i \beta)) \beta_r \quad (5)$$

De acuerdo a estos efectos marginales computados, la hipótesis central del estudio consiste en la referencia del efecto positivo de las densidades viales terciarias, en la medida en que estas densidades vienen acompañadas de una mayor oferta de bienes públicos, sobre la probabilidad de que un estudiante de las zonas rurales aumente su rendimiento académico. En línea con el trabajo de [Campos \(2017\)](#), se realizó el test de Hausman para determinar problemas de endogeneidad asociados a la variable de densidades viales terciarias no encontrando problemas de tal tipo en las estimaciones realizadas como el lector puede encontrar en los anexos.

2 La base de datos SABER 11 cuenta también con información relativa a si la familia del estudiante dispone de motocicletas, comúnmente utilizadas en las zonas rurales. Sin embargo, no hay registro de esta información para el departamento de Antioquia en donde se llevó a cabo el presente estudio para dicho año.

Finalmente, dado que el estudio se concentra fundamentalmente en el problema del logro educativo, *score*, y no sobre el problema de deserción escolar, las implicaciones sobre posibles sesgos de autoselección muestral son menores. Tal como lo determina [Ocampo \(2017\)](#), las vías inciden en la deserción escolar; sin embargo, la implicación que pueda tener sobre el rendimiento educativo depende de las características distribucionales de las habilidades innatas de los estudiantes, esto es, si los que deciden salir del sistema escolar pertenecen a la cola izquierda de las habilidades, entonces las estimaciones serán sesgadas sobre el logro, ocurriendo lo mismo si los estudiantes que salen pertenecen a la cola derecha de la distribución. La implicación, entonces, de la autoselección será determinante si se pretende hacer una estimación agregada a escala de municipio que implicaría el cálculo del rendimiento promedio a escala de municipio, como ocurre en la mayoría de trabajos en la literatura. La mayor variabilidad a nivel de estudiante que muestra el indicador propuesto en el estudio permite admitir una mayor representatividad muestral.

### 3. Análisis de Resultados

Los resultados para el año 2012 del rendimiento educativo en básica secundaria de los estudiantes en Antioquia muestra unas regularidades en buena medida asociadas con la distribución subregional del ingreso y de la calidad de vida. Por ejemplo, de acuerdo a la tabla (1), los estudiantes que residen en el casco urbano de los municipios presentan un mayor rendimiento educativo promedio que los estudiantes que viven en las zonas rurales; así mismo, de los estudiantes rurales, en promedio, los estudiantes hombres presentan un mayor rendimiento que las mujeres, situación que también ocurre en el caso de los estudiantes urbanos.

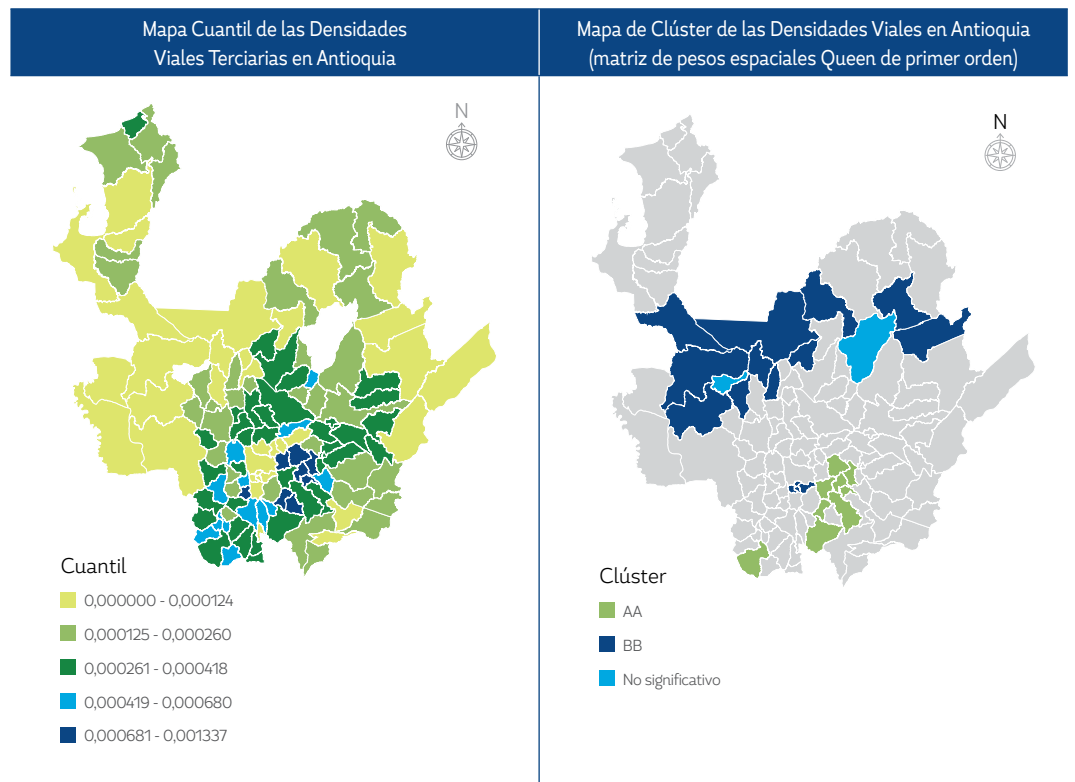
**Tabla (1). Estadísticas Descriptivas:  
Variable Dependiente (Logro Educativo en Matemáticas)**

Estadístico	Estudiantes Rurales	Estudiantes Urbanos	Rurales Hombres	Rurales Mujeres
Min	0.00	0.00	0.00	0.00
1st. Q	35.00	39.00	38.00	35.00
Median	43.00	46.00	43.00	42.00
<u>Mean</u>	<u>42.65</u>	<u>46.05</u>	<u>43.79</u>	<u>41.76</u>
Ds				
3rd.Q	49.00	52.00	50.00	49.00
Max	85.00	126.00	85.00	84.00

Fuente: Cálculos Propios con base en el ICFES pruebas SABER 11.

Como puede apreciarse en la tabla (2), se ilustra la distribución espacial de las densidades viales terciarias a nivel municipal para el año 2012, tanto en el mapa cuantil como el mapa de clúster. Los municipios con mayores niveles de densidad vial terciaria se ubican en general en el centro del departamento de Antioquia, particularmente en los municipios cercanos al Valle de Aburrá, municipios pertenecientes a las subregiones del Oriente y el suroeste. Preciso resulta indicar que el centro del departamento, concretamente en los municipios del Valle de Aburrá y subregiones cercanas, agrupan los municipios de mayor actividad económica del departamento y, en general, presentan los mayores niveles de calidad de vida. Este patrón *centro-periferia* de la actividad económica del departamento también ocurre, en promedio, con el rendimiento educativo en básica secundaria, como el trabajo de [Quintero et al., 2016](#) establece en relación con un patrón centro periferia también del rendimiento educativo en las pruebas de matemáticas y lenguaje.

**Tabla (2). Densidad vial terciaria en Antioquia**



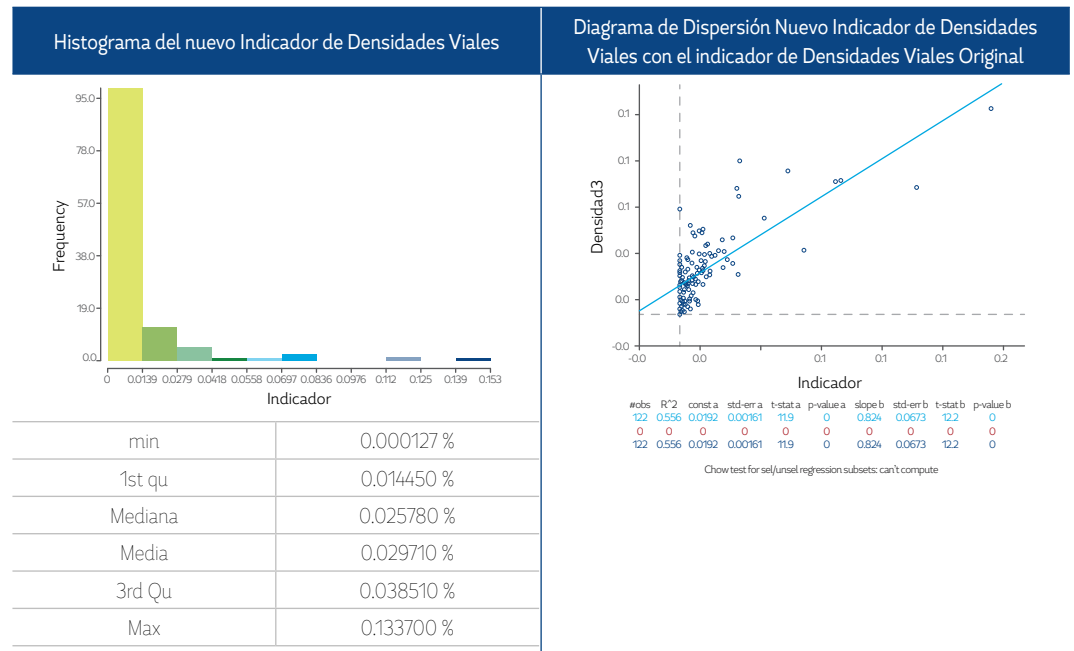
Fuente: Cálculos Propios con base en IGAC

En la tabla (3) se presentan, por su parte, algunas estadísticas descriptivas relevantes del nuevo indicador de densidades viales terciarias, ecuación (1), haciendo una agregación del mismo a escala de municipio. Como puede notarse, las densidades viales terciarias representan un bajo porcentaje del área total de los municipios cuando se considera también una conectividad socio-económica, en promedio del 0.03% del departamento de Antioquia, con un máximo de 0.13%. De acuerdo al histograma, la distribución de esta conectividad rural es estructuralmente desigual, con la mayoría de municipios del departamento con una conectividad socio-económica rural poco representativa y otros, con una mayor conectividad de sus áreas rurales. En la tabla (3) se presenta también el diagrama de dispersión entre el indicador de densidades terciarias municipales y el indicador propuesto; este diagrama establece evidencia de una relación directa entre ambos indicadores, lo cual sugiere que una mayor densidad vial terciaria tiende a corresponderse con una provisión mayor de bienes y servicios públicos en estas áreas rurales como cabría esperarse. Sin embargo, observando con más detalle dicho diagrama, se encuentra una concentración en los niveles inferiores del indicador propuesto — eje horizontal—; esta concentración sugiere que a estos niveles del indicador propuesto hay una importante variabilidad de los municipios en términos de sus densidades viales terciarias solamente. Así pues, el indicador propuesto capta aspectos adicionales distintos a los puramente socio-económicos, que terminan desde luego impactando al rendimiento educativo en términos de la dispersión de las áreas rurales que connota.

En cuanto a la distribución espacial del nuevo indicador, tabla (4), surgen algunas diferencias reseñables. *Prima facie*, el comportamiento espacial de este indicador es similar, en general, al comportamiento de las densidades terciarias. Sin embargo, en cuanto a los municipios que integran los clústeres, el patrón presenta cambios. En el clúster alto-alto (verde), el número de municipios es menor,

pero sigue ubicándose en el centro del departamento, cercano a la subregión del Valle de Aburrá; la disminución en el número de municipios de este clúster es indicativo de las diferencias en términos socio-económicos de los estudiantes rurales. El clúster bajo-bajo (rojo) agrupa también un número menor de municipios, pero adiciona al municipio de Remedios en la subregión del Nordeste del departamento. En general, el patrón espacial antes mencionado del tipo *centro-periferia* se sigue manteniendo en el caso del nuevo indicador de tal manera que los municipios que tienen más conectada sus áreas rurales, tanto desde el punto de vista de las vías como desde el punto de vista socio-económico, están generalmente ubicados en el centro del departamento en las subregiones, cercanos a la subregión del Valle de Aburrá.

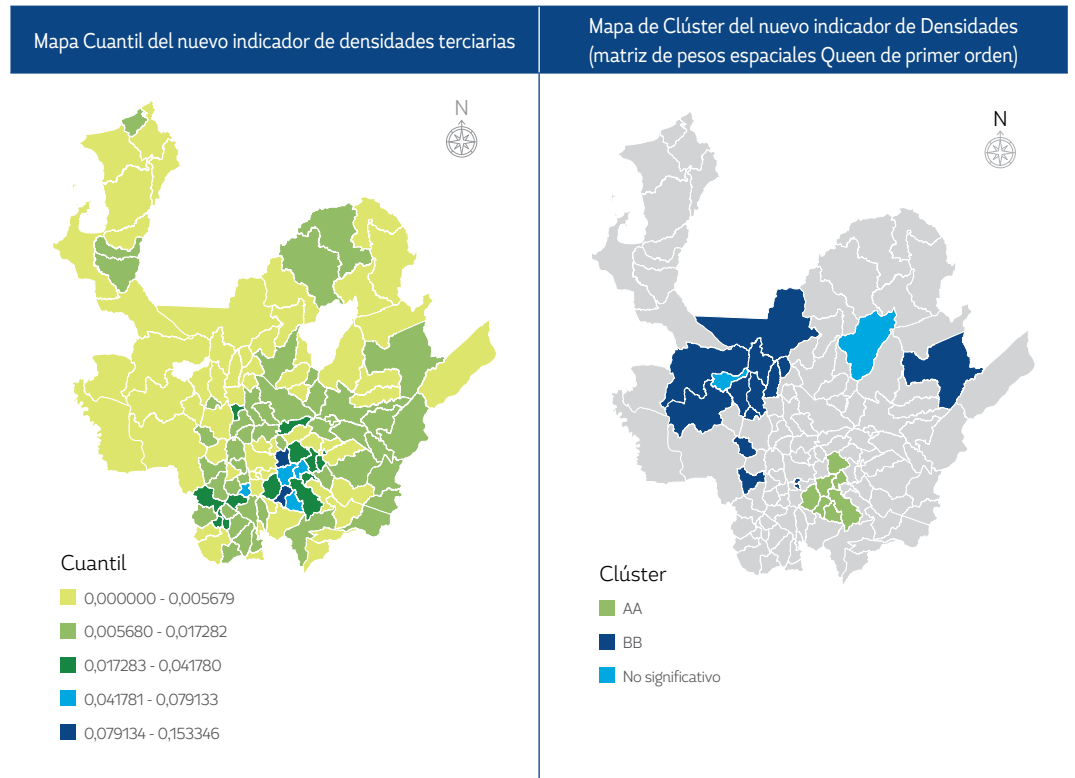
**Tabla (3). Histograma y Diagrama de Dispersión del Nuevo Indicador de Densidades Viales**



Fuente: Cálculos Propios con base en IGAC



**Tabla (4). Patrón Espacial del Nuevo Indicador de Densidades Viales Terciarias.**



Fuente: Cálculos Propios con base en IGAC

Cuando ya se pretende examinar el comportamiento a escala de micro dato, la tabla (5) presenta los resultados del modelo probit Ordenado, así como los efectos marginales. Como puede observarse, si el estudiante se encuentra en un municipio con una alta conectividad de sus zonas rurales con el casco urbano y con la oferta de bienes institucionales y públicos, su probabilidad de aumentar el rendimiento educativo es mayor. Concretamente, sus efectos marginales son ilustrativos: un punto básico de crecimiento de la densidad vial terciaria disminuye la probabilidad de ubicarse en el cuartil más bajo del rendimiento educativo en cerca del 20%; ese mismo aumento de la densidad disminuye la probabilidad de ubicarse en el rendimiento educativo promedio, y por lo tanto ascender, en cerca del 4%; finalmente, el mismo incremento de la densidad genera un aumento de la probabilidad de permanecer en el cuartil más alto del rendimiento educativo en cerca del 29% para los estudiantes rurales.

**Tabla (5). Modelo Probit Ordenado del Logro Educativo.**

Variable	Coficiente	Std. Error	Valor p	Efectos Marginales Bajo Rendimiento	Efectos Marginales Rendimiento Medio	Efectos Marginales Alto Rendimiento
Sexo	.3458149	.0562395	0.00***	-.0738476 (0.000***)	-.0116594 (0.000***)	.085507 (0.000***)
Edad	-.0256565	.0066224	0.00***	.0054789 (0.000***)	.000865 (.0000***)	-.0063439 (0.000***)
Nivel Sisben	.1305097	.0272785	0.00***	-.0278699 (0.000***)	-.0044002 (0.000***)	.0322701 (0.000***)
Dormitorios	.0370141	.0286515	0.196	-.0079042 (0.196)	-.001248 (0.200)	.0091522 (0.196)
Celular	.165647	.0819218	0.043***	-.0353734 (0.043***)	-.0055849 (0.047***)	.0409583 (0.043***)
Jornada Mañana	.2625453	.0953561	0.006***	-.0560657 (0.006***)	-.0088519 (0.008***)	.0649176 (0.006***)
Jornada Tarde	.3080328	.1087968	0.005***	-.0657794 (0.005***)	-.0103855 (0.006***)	.0761649 (0.005***)
Jornada Completa	.232426	.0829169	0.005***	-.0496338 (0.005***)	-.0078364 (0.007***)	.0574702 (0.005***)
Trabaja	-.0249337	.0120073	0.038***	.0053245 (0.038***)	.0008407 (0.042***)	-.0061652 (0.038***)
Densidad Vial Terciaria (nuevo indicador)	<b>1.177898</b>	<b>.6675734</b>	<b>0.078**</b>	<b>-.251536</b> <b>(0.078**)</b>	<b>-.0397137</b> <b>(0.082**)</b>	<b>.2912498</b> <b>(0.078**)</b>
Ingreso Familiar Mensual	.1494207	.0390961	0.000***	-.0319083 (0.000***)	-.0050378 (0.000***)	.0369461 (0.000***)

Fuente: Cálculos Propios con base en el ICFES e IGAC

En lo que atañe, entonces, a la conectividad vial terciaria reformulada mediante la ecuación (1), un avance del sistema vial implica un desplazamiento del logro educativo de la distribución del rendimiento, aunque en mayor medida impacta a los extremos de la distribución, esto es, a cuartil de bajo rendimiento educativo, y al cuartil de alto rendimiento educativo con impactos similares. Una política, entonces, que aumente la densidad vial terciaria de los municipios menos conectados tiene el potencial de reducir las brechas del rendimiento educativo de estudiantes rurales con respecto a los estudiantes urbanos, mejorando generando condiciones para que al largo plazo se fomente la movilidad social y las brechas económicas también puedan eventualmente reducirse.

De acuerdo también a los resultados de la tabla (5), puede concluirse que en las zonas rurales tiene un mayor impacto sobre el rendimiento educativo la práctica de la jornada de la tarde, en los extremos de la distribución, en cerca de un 6% para el cuartil de bajo rendimiento educativo, y del 7% para el cuartil de alto rendimiento educativo; en tanto que la jornada única los resultados de estos efectos marginales son de cerca del 5% en los dos extremos. En las zonas urbanas, como otros estudios lo confirman, la jornada única tienen un impacto mayor. De igual manera, el celular se convierte en una importante herramienta de estudio para los estudiantes rurales.<sup>3</sup>

3 La variable referente a la propiedad de computadoras resultó estadísticamente no significativa para los estudiantes rurales.

## 5. Discusión y Conclusiones

En el presente trabajo, se estimó el impacto del sistema vial terciario sobre la probabilidad de ascenso de los estudiantes rurales en su logro educativo en la prueba de matemáticas. Tal propósito comportó problemas metodológicos en relación con la falta de datos concretos sobre la distancia, y por tanto dispersión, de los estudiantes rurales a los centros urbanos donde se encuentran sus colegios. Con base en la consideración de micro datos, se estableció un indicador proxy del grado de dispersión de los estudiantes rurales teniendo en cuenta la densidad vial terciaria y la oferta de bienes y servicios públicos. Ante la escasez, todavía latente, de datos relacionados con la geo-referenciación de estudiantes, la alternativa metodológica de este artículo tiene el potencial de constituir un referente para estudios en países en desarrollo donde la escasez de datos es mayor.

Los estudios que hasta el momento se han realizado para Colombia, como el de ([Campos, 2017](#)), y para otros países, tienen en la mayoría de casos la característica de ser generales, a escala de municipio o de Estado. Las conclusiones de este trabajo se basan fundamentalmente en la aceptación a nivel micro de los datos y, por lo tanto, la econometría realizada tiene la capacidad de captar las heterogeneidades subyacentes de una manera más directa que en los anteriores aportes de la literatura.

Los resultados permitieron inferir que una mayor densidad vial terciaria junto con una mayor oferta institucional de bienes y servicios públicos, tienen un impacto positivo en la probabilidad de ascender, en su rendimiento educativo, de los estudiantes rurales. Además, este efecto es mayor en los extremos de la distribución del logro educativo, dado que, tanto en el cuartil de bajo rendimiento educativo como en el cuartil superior, el efecto marginal de las densidades viales terciarias es de cerca del 20% sobre dicha probabilidad. Los resultados tienen su utilidad para la puesta en marcha de planes de mejoramiento de vías terciarias con el fin de conocer sus impactos sobre distintas dimensiones del desarrollo económico.

Como sugerencias a trabajos posteriores en este campo y que se derivan de las aportaciones del presente artículo, se encuentra la posibilidad de llevar a cabo estimaciones comparativas entre estudiantes rurales por medio de metodologías como la descomposición de Oaxaca-Blinder, teniendo como referente indicadores del sistema vial. De tal suerte, se tomarían medidas acordes con el problema de auto-selección muestral para estudiar el caso de deserción escolar. Por otro lado, la base de datos utilizada en este estudio permite emprender el estudio también de la problemática del trabajo infantil y su afectación sobre el rendimiento educativo de los estudiantes rurales.

## Bibliografía

- Baschieri, A., & Falkingham, J. (2006). Staying in school: assessing the role of access, availability and cost [Monograph]. Recuperado 27 de marzo de 2017, a partir de <https://eprints.soton.ac.uk/45358/>
- Campos, A. (2017). *Vías Para La Educación: Efecto De La Infraestructura Vial En Los Resultados Educativos (2005 - 2015) (Roads to Education: Impact of Road Infrastructure on Education (2005 - 2015))* (SSRN Scholarly Paper No. ID 2921472). Rochester, NY: Social Science Research Network. Recuperado a partir de <https://papers.ssrn.com/abstract=2921472>
- Castro-Leal, F., Dayton, J., Demery, L., & Mehra, R. (1999). Public Social Spending in Africa: Do the Poor Benefit? *The World Bank Research Observer*, 14(1), 49-72. <https://doi.org/10.1093/wbro/14.1.49>
- Datt, G., & Ravallion, M. (1998). Farm productivity and rural poverty in India. *The Journal of Development Studies*, 34(4), 62-85. <https://doi.org/10.1080/00220389808422529>
- de Janvry, A., & Sadoulet, E. (2000). Rural poverty in Latin America: Determinants and exit paths. *Food Policy*, 25(4), 389-409. [https://doi.org/10.1016/S0306-9192\(00\)00023-3](https://doi.org/10.1016/S0306-9192(00)00023-3)
- Durantón, G. (2015). Roads and trade in Colombia. *Economics of Transportation*, 4(1-2), 16-36. <https://doi.org/10.1016/j.ecotra.2014.11.003>
- Falch, T., Lujala, P., & Strøm, B. (2013). Geographical constraints and educational attainment. *Regional Science and Urban Economics*, 43(1), 164-176. <https://doi.org/10.1016/j.regsciurbeco.2012.06.007>
- Gibbons, S., & Vignoles, A. (2012). Geography, choice and participation in higher education in England. *Regional Science and Urban Economics*, 42(1-2), 98-113. <https://doi.org/10.1016/j.regsciurbeco.2011.07.004>
- Huisman, J., & Smits, J. (2009). Effects of Household- and District-Level Factors on Primary School Enrollment in 30 Developing Countries. *World Development*, 37(1), 179-193. <https://doi.org/10.1016/j.worlddev.2008.01.007>
- Jeremy Atack, R. A. M., & Elisabeth Perlman. (2012). The Impact of Railroads on School Enrollment in Nineteenth Century America. *Working Paper*.
- Osgood, D. W., & Chambers, J. M. (2000). Social Disorganization Outside the Metropolis: An Analysis of Rural Youth Violence\*. *Criminology*, 38(1), 81-116. <https://doi.org/10.1111/j.1745-9125.2000.tb00884.x>
- Quintero, L., Leandro, O., & Hincapié Vélez, D. (2016). A study of regional gaps in education quality in Colombia: 2000-2012. *Ensayos sobre POLÍTICA ECONÓMICA*, 34(SPE79), 3-20. <https://doi.org/10.1016/j.espe.2016.01.001>
- Reardon, T. (1997). Using evidence of household income diversification to inform study of the rural non-farm labor market in Africa. *World Development*, 25(5), 735-747. [https://doi.org/10.1016/S0305-750X\(96\)00137-4](https://doi.org/10.1016/S0305-750X(96)00137-4)
- Reardon, T., Berdegué, J., & Escobar, G. (2001). Rural Nonfarm Employment and Incomes in Latin America: Overview and Policy Implications. *World Development*, 29(3), 395-409. [https://doi.org/10.1016/S0305-750X\(00\)00112-1](https://doi.org/10.1016/S0305-750X(00)00112-1)
- Ribot, J. C., & Peluso, N. L. (2003). A Theory of Access\*. *Rural Sociology*, 68(2), 153-181. <https://doi.org/10.1111/j.1549-0831.2003.tb00133.x>
- Shilpa Aggarwal. (2013). Do Rural Roads Create Pathways out of Poverty? Evidence from India. *Working Paper*.
- Spencer, G. A., & Bryant, S. A. (2000). Dating violence: a comparison of rural, suburban, and urban teens. *Journal of Adolescent Health*, 27(5), 302-305. [https://doi.org/10.1016/S1054-139X\(00\)00125-7](https://doi.org/10.1016/S1054-139X(00)00125-7)
- Vasconcellos, E. A. (1997). Rural transport and access to education in developing countries: policy issues. *Journal of Transport Geography*, 5(2), 127-136. [https://doi.org/10.1016/S0966-6923\(96\)00075-0](https://doi.org/10.1016/S0966-6923(96)00075-0)
- Villar, L., & Ramírez, J. M. (2014). Infraestructura regional y pobreza rural. Recuperado a partir de <http://www.repository.fedesarrollo.org.co/handle/11445/234>

## Anexos

Test de Hasuman para determinar Exogenidad sobre la variable de Densidades Viales Terciarias

Estadísticos	valor	p-valor
Durbin (score) chi cuadrado	0.165571	0.6841
Wu-Hausman F (,14921)	0.165174	0.6845

Fuente: Cálculos propios con base en datos del ICFES e IGAC

---

EFECTO DE LOS AUMENTOS EN LA  
ESCOLARIDAD EN EL MERCADO LABORAL  
COLOMBIANO ENTRE 2008 Y 2016

The Effect of Increased  
Schooling in the Colombian  
Labor Market Between  
2008 and 2016

Tomás Aristizábal Lopera y Esteban Ángel López

---

Research Article

## EFFECTO DE LOS AUMENTOS EN LA ESCOLARIDAD EN EL MERCADO LABORAL COLOMBIANO ENTRE 2008 Y 2016

### The Effect of Increased Schooling in the Colombian Labor Market Between 2008 and 2016

Tomás Aristizábal Lopera<sup>a\*</sup>, Esteban Ángel López<sup>b</sup>

**Palabras clave:** mercado laboral colombiano, salarios colombianos, Colombia, modelo de Mincer

**Keywords:** Colombian labor market, Colombian salaries, Colombia, Mincer model

**Clasificación JEL:** R1, I21, C01, C25

**Received:** 05/04/2017

**Accepted:** 10/06/2017

**Published:** 20/06/2017

#### Resumen

Examinamos los cambios en la oferta de obreros colombianos con niveles diferentes de escolaridad y estimamos el efecto de estos cambios en los salarios entre 2008 y 2016 con un modelo de Mincer. La proporción de la fuerza laboral con escolaridad de primaria o menos bajó de 46% a 32,5%, la proporción con escolaridad de secundaria aumentó de 36% a 39,5%, y con terciaria aumentó de 18% a 28%. Encontramos que la escolaridad presenta efectos positivos sobre el salario para cada nivel de educación, donde el nivel terciario es el que presenta mayor efecto. El efecto marginal de un año de escolaridad bajó en cada nivel entre 2008 y 2016. En términos reales encontramos un aumento de 21% en los salarios promedios de los obreros con primaria entre estas fechas. Encontramos pocos cambios en los salarios promedios de los obreros más educados. Puesto que la fracción de obreros más educados aumentó, hubo un aumento sustancial en el ingreso laboral total.

#### Abstract

We examine the changes in the supply of Colombian workers with different levels of schooling and estimate the effect of these changes on salaries between 2008 and 2016 using a Mincer model. The share of the work force with primary schooling or less declined from 46% to 32.5%, the share with secondary schooling rose from 36% to 39.5%, and the share with tertiary schooling rose from 18 to 28%. We find that schooling had positive effects at every level of schooling, but particularly at the tertiary level. The marginal effect of a year of schooling declined at every level. In real terms we find an increase of 21%

a, b. Universidad EAFIT, Medellín - Colombia.

\* Autor para correspondencia:  
taristiz@eafit.edu.co

in average salaries for workers with primary schooling between 2008 and 2016 but few changes in average salaries among more educated workers. Since the share of more educated workers increased, total labor income in Colombia increased substantially over these years.

## 1. Introducción

Colombia ha venido evidenciando una transformación en el sistema educativo durante los últimos veinte años, convirtiendo la educación como una de las cuentas más privilegiadas en el presupuesto nacional, que para el 2017 se le destinó un presupuesto de \$39,4 billones de pesos ([Portafolio, 2016](#)), centrándose en políticas que permitan aumentar el acceso a la educación. Los resultados ya se han empezado a ver pues la esperanza de vida escolar ha aumentado dos años en el período del 2001 al 2010, pasando de 11,4 a 13,5 años; sin embargo, el país debe seguir trabajando en este objetivo para alcanzar el promedio de la OCDE que equivale a 16,5 años de vida escolar para el mismo período ([MinEducación, 2016](#)).

El propósito de este artículo es de examinar como los aumentos en la escolaridad de la fuerza laboral ha afectado la distribución de obreros calificados y los salarios en el mercado laboral colombiano de 2008 a 2016. Escogemos este periodo para el estudio porque estas fechas corresponden al primer año y el último año en el cual los datos sobre los obreros son disponibles en la Encuesta de Calidad de la Vida en una forma consistente. Esta consistencia nos permite hacer una comparación válida de lo que ha ocurrido entre estos dos años. Este periodo es relativamente corto, pero muestra las tendencias recientes en el mercado colombiano.

Para hacer esta comparación vamos a usar una variación del modelo de Mincer en el cual examinamos el efecto de la escolaridad y la experiencia separadamente para obreros con los niveles diferentes de escolaridad. Interpretamos los cambios en los salarios dentro de un análisis de los cambios en oferta y demanda para obreros con los niveles diferentes de escolaridad.

El artículo se organiza a lo siguiente. [Sección 2](#) presenta la metodología. [Sección 3](#) presenta el análisis de los efectos de la escolaridad en 2008 y 2016. [Sección 4](#) examina los cambios en la oferta laboral y en los salarios entre las dos fechas. [Sección 5](#) concluye.

## 2. Metodología

### Modelo de Mincer

Mincer (1974) propuso un modelo que integra en conjunto la escolaridad y la experiencia laboral para definir los salarios de las personas:

$$\ln(W_i) = \alpha + \beta_1 S_i + \beta_2 X_i + \beta_3 X_i^2 + \mu_i \quad (1)$$

De (1) tenemos que  $\ln(W_i)$  es el logaritmo del salario que devenga el individuo en un determinado período;  $S_i$  es el nivel de educación de la persona, medido como los años que ha estudiado durante su vida;  $X_i$  son los años de experiencia laboral de la persona;  $X_i^2$  es la experiencia elevada al cuadrado; por último,  $\mu_i$  es el término de perturbación aleatoria para las muestras no observables.

Este modelo trabaja con el supuesto de que el efecto de la escolaridad es la misma para cualquier nivel de educación, donde según [Sapelli \(2009\)](#), no es lo mismo una primaria, un bachillerato, un



pregrado, una técnica, un posgrado, etc..., y, además, es importante el hecho de poseer el título o no del estudio realizado. Es por esto que [Montenegro y Patrinos \(2014\)](#) proponen el modelo ampliado de Mincer donde la educación se divide en nivel primario (primaria), secundario (bachillerato) y terciario (pregrado, técnica, posgrado...). Para este caso, según el sistema educativo colombiano se va a tener que el nivel primario va de 1 a 5 años de educación; el nivel secundario de 6 a 11 años de educación y el nivel terciario más de 11 años de educación.

$$\ln(W_i) = \alpha + \beta_p Dp_i + \beta_s Ds_i + \beta_t Dt_i + \beta_1 X_i + \beta_2 X_i^2 + \mu_i \quad (2)$$

Para este modelo se tiene que  $Dp_i$  serán los años en el nivel primario de educación,  $Ds_i$  en el nivel secundario y  $Dt_i$  en el nivel terciario. Según [Psacharopoulos \(1979\)](#), las betas (coeficientes) obtenidos en este modelo serán los efectos de cada nivel de educación sobre el salario.

Sin embargo, en este modelo se trabaja con el supuesto de que la experiencia tiene el mismo coeficiente para todos los niveles de educación, el cual no se ajusta a la realidad. [Heckman, Lochner, y Todd \(2008\)](#) muestran evidencia de que el ingreso de los trabajadores que reciben título, incrementa más con los años de escolaridad, que el de los trabajadores que no lo reciben. También muestran que a veces el efecto de un año de escolaridad y un año de experiencia no son independientes. Adicionalmente la experiencia puede tener mejores efectos sobre el salario de los trabajadores en los niveles más altos de escolaridad.

Es por esto que se decide realizar una regresión independiente para cada nivel, utilizando el modelo de la ecuación de [Mincer \(1975\)](#) con el fin de tener un coeficiente de experiencia según la escolaridad de cada individuo.

(3) Para el nivel primario:  $\ln(W_i) = \alpha + \beta_1 p_i + \beta_2 Xp_i + \beta_3 Xp_i^2 + \mu_i$

(4) Para el nivel secundario:  $\ln(W_i) = \alpha + \beta_1 s_i + \beta_2 Xs_i + \beta_3 Xs_i^2 + \mu_i$

(5) Para el nivel terciario:  $\ln(W_i) = \alpha + \beta_1 t_i + \beta_2 Xt_i + \beta_3 Xt_i^2 + \mu_i$

### Datos

Los datos del estudio son tomados de la Encuesta de Calidad de Vida (ECV) realizada en hogares por el DANE en los años 2008 y 2016. Esta encuesta es realizada según el DANE "para el total nacional, cabecera y centros poblados<sup>1</sup> - rural disperso por grandes regiones (Antioquia, Valle, Atlántico, Pacífica, Central, Oriental); solo cabecera para Orinoquia-Amazonia; Bogotá D.C. y San Andrés" (DANE, 2016).

De acuerdo con los modelos ya especificados, las variables son el salario mensual, los años de escolaridad, y la experiencia. Los datos para estas variables se toman o se calculan de las respuestas a las cuatro preguntas de la encuesta notadas en Tabla 1.

1 Cabecera son ciudades principales del departamento.

**Tabla 1. Variables de la Encuesta de Calidad de Vida (ECV)**

Nombre de variable	Pregunta
<b>Fuerza laboral</b>	
TRSALARIO	¿Cuánto fue el salario del último mes antes de descuentos?
TRBUSCA	¿En los últimos 12 meses ha hecho algo para conseguir trabajo?
<b>Educación</b>	
EDNIVEL	Nivel educativo más alto alcanzado
EDAÑO	Cual fue el último año aprobado en este nivel
<b>Composición del hogar</b>	
EDAD	Edad del individuo

Es importante resaltar que no se tomaron preguntas con respuesta abierta, cada una de las preguntas tiene sus respuestas posibles enumeradas y clasificadas, por ejemplo, las respuestas de “Si” o “No” toman valores de 1 o 2.

En la sección de educación la variable EDNIVEL hace referencia al último nivel alcanzado en la división que tiene el DANE para clasificar los estudios alcanzados por los individuos, y la variable EDAÑO pregunta por el último grado aprobado dentro del nivel que se escogió.

En la base de datos de 2008 toda la encuesta está unificada, mientras que en 2016 hay una base de datos independiente para cada tema (Fuerza laboral, Educación y Composición del hogar); para dar solución a esto se unieron las tres bases de datos de 2016 guiándose por el número de directorio y el orden, con esta información se pudo diferenciar qué hogar estaba siendo encuestado, y en orden de mayor a menor, cual persona de ese hogar respondió a las preguntas. Tabla 2 presenta la organización de la información sobre escolaridad en las dos encuestas.

**Tabla 2. Clasificación de los Niveles de Educación según DANE**

Niveles de Educación de la ECV		
	2016	2008
1	Ninguno	Ninguno
2	Preescolar	Preescolar
3	Básica Primaria (1º - 5º)	Básica primaria (1º a 5º)
4	Básica secundaria (6º - 9º)	Básica secundaria y media (6º a 13º)
5	Media (10º - 13º)	Técnico
6	Técnico sin título	Tecnológico
7	Técnico con título	Universitaria sin título
8	Tecnológico sin título	Universitaria con título
9	Tecnológico con título	Postgrado sin título
10	Universitario sin título	Postgrado con título
11	Universitario con título	
12	Postgrado sin título	
13	Postgrado con título	

### Ajustes de los datos y cálculos

Luego de tener las bases de datos para ambos años unificadas se sacaron los individuos que no eran objeto de estudio en el trabajo; primero se eliminaron todos los menores de 16 años puesto que no devengaban un salario; luego salieron las personas que no devengan salario, ya que en el trabajo solo sirven los individuos que tengan ingresos para evaluar como la educación o la falta de esta afectan su nivel. Después se eliminaron aproximadamente 200 individuos que reportaron un salario de 99\$ al mes, lo cual se consideró como información poco veraz que afectaba la regresión.

Por otro lado, en el tema de la educación, las personas que tenían como máximo nivel de estudio el preescolar (2), se unieron con las personas que no tenían ningún nivel de educación (1), debido a que es un nivel insignificante de estudio. Otro ajuste que se hizo fue eliminar aproximadamente 80 personas que cursaron en la escuela secundaria hasta los grados 12º y 13º, esto se hizo teniendo en cuenta que el sistema educativo colombiano establece que el bachillerato es hasta el grado 11º, pero que hay unas pocas instituciones que el nivel secundario va hasta 12º o 13º. Por último, los salarios de 2008 se ajustaron a pesos de 2016, usando una tasa de inflación anual promedio de estos ocho años de 3,75%.

La experiencia laboral de cada persona se calculaba de los datos sobre la edad y los años de escolaridad. La Tabla 3 explica el proceso.

**Tabla 3. Metodología de cálculo de los años de experiencia para cada individuo.**

Personas con Educación Terciaria	Personas sin Educación Terciaria
$Edad < 18: X_i = 0$	$Edad < 18: X_i = 0$
$18 < Edad < 22: X_i = edad - 18$	$Edad > 18: X_i = edad - 18$
$Edad > 22^*: X_i = edad - S_i - 6$	

\* $Edad > 22^*: X_i = edad - S_i - 6$

Luego de tener todas las variables organizadas se procede con la creación de los tres niveles que aparecen en el modelo ampliado de Mincer; el primario va desde el grado 1º hasta 5º, que es el equivalente al nivel 3 del DANE (educación primaria); el secundario va desde 6º hasta 11º que son los niveles 4 y 5 en 2016 y 4 en 2008 en la encuesta del DANE (educación secundaria o bachillerato), finalmente el nivel terciario hace referencia a todo estudio superior, incluye técnico, tecnológico, pregrado y postgrado; al tener todos los datos necesarios se procede con las regresiones.

Para estimar la oferta laboral se contaron las personas que se encontraban trabajando y las que habían buscado empleo en los últimos 12 meses para cada nivel de educación. Dado que el tamaño de la muestra del DANE para 2016 ha incrementado respecto al 2008, se utilizan porcentajes de la fuerza laboral, agrupando por años de experiencia, con el fin de poder evidenciar como ha cambiado la oferta laboral en valores comparables.

Vale la pena resaltar que los datos pueden presentar un error de medida dado que las personas no siempre responden con total veracidad, ya que pueden creer que sus respuestas afectarían su situación frente al estado lo cual puede traer mayores impuestos o tributos. Además, las encuestas no son realizadas a las mismas personas de 2008 al 2016, lo cual genera problemas de exactitud en los resultados al momento de analizar el impacto del efecto de la educación a través del tiempo. Sin embargo, se asume que la muestra tomada para esta encuesta es confiable, al igual que la información suministrada.

2 Propuesto por Merlo (2009)

### 3. Resultados de las regresiones

Las regresiones mostradas a continuación son estimadas por Mínimos Cuadrados Ordinarios (MCO) aplicando las ecuaciones (3), (4) y (5) mencionadas anteriormente, con el fin de calcular el efecto de cada año de educación, la experiencia y su cuadrado, sobre el salario desagregado por nivel.

**Tabla 4:** Efecto de la Escolaridad y Experiencia en el Salario

	2008	2016
<b>PRIMARIO</b>		
$D_{p_i}$	0,0656 (0,0109)	0,0491 (0,0099)
$X_{p_i}$	0,0321 (0,0031)	0,0276 (0,0030)
$X_{p_i}^2$	-0,00049 (0,00005)	-0,00042 (0,00004)
Constante	12,182 (0,063)	12,535 (0,067)
<b>SECUNDARIO</b>		
$D_{p_i}$	0,0731 (0,0066)	0,0562 (0,0050)
$X_{p_i}$	0,0369 (0,0031)	0,0273 (0,0021)
$X_{p_i}^2$	-0,00051 (0,00008)	-0,00041 (0,00005)
Constante	12,455 (0,044)	12,754 (0,033)
<b>TERCIARIO</b>		
$D_{p_i}$	0,2002 (0,0073)	0,1931 (0,0045)
$X_{p_i}$	0,0311 (0,0054)	0,0226 (0,0029)
$X_{p_i}^2$	-0,00046 (0,00014)	-0,00024 (0,00069)
Constante	12,919 (0,051)	12,993 (0,0285)

Al observar de manera general los resultados de las regresiones, se evidencia una tendencia creciente del efecto de la escolaridad sobre los salarios a medida que aumenta el nivel; para el año 2008 un año adicional en el nivel primario aumenta en 6,6% el salario, un año adicional del nivel secundario aumenta el 7,3%, y un año adicional en el nivel terciario tiene un efecto positivo del 20,0%; por otro lado, se observa que el efecto de la experiencia varía para cada nivel de escolaridad, tanto para el efecto lineal como el efecto al cuadrado. Con estos resultados se demuestra que el supuesto de experiencia igual para todos los niveles que utiliza [Mincer \(1975\)](#) en su modelo no se ajusta a la realidad.

Sin embargo, no es claro que implica para los salarios en términos absolutos porque las constantes son diferentes en cada nivel.

Analizando los datos que arroja el modelo para 2016 se observa que, aunque el patrón de efectos entre niveles es semejante, el efecto de un año de escolaridad ha bajado en cada nivel. Un año adicional en el nivel primario aumenta en 4,9% el salario, un año adicional del nivel secundario aumenta el 5,6%, y un año adicional en el nivel terciario tiene un efecto positivo del 19,3%.

#### 4. Cambios en el mercado laboral

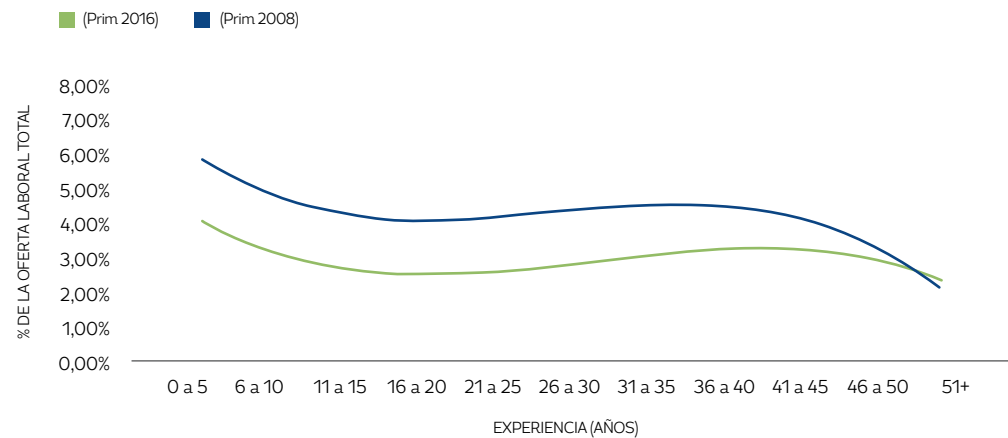
Dado que tanto los coeficientes como la constante cambian entre 2008 y 2016, no es posible concluir el efecto neto en los salarios de las regresiones. Por tanto, más adelante mostramos los patrones de salarios en cada nivel de escolaridad para las dos fechas. También miramos como la oferta de trabajo ha cambiado en cada nivel de escolaridad.

##### Mercado Laboral – Escuela Primaria o Menos

Puesto que el nivel promedio de escolaridad ha aumentado, el número de obreros en el mercado laboral con niveles diferentes de escolaridad ha cambiado sustancialmente. Dado que las encuestas en 2008 y 2016 tienen números diferentes de obreros, analizamos como ha cambiado el mercado laboral por mostrar el porcentaje de la fuerza laboral con los niveles diferentes de escolaridad en las dos fechas.

La gráfica 1 muestra cómo ha cambiado el tamaño relativo de la fuerza laboral con 0 a 5 años de escolaridad entre 2008 y 2016. Vemos que la proporción de los obreros colombianos con este nivel de escolaridad ha disminuido sustancialmente sobre estos ocho años.

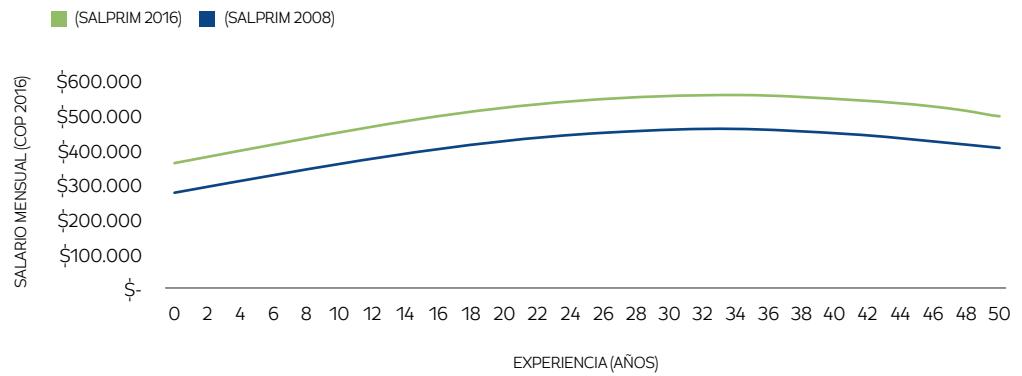
**Gráfica 1. Oferta de Obreros con Primaria o Menos**



Fuente. ECV 2008/2016 DANE. Cálculos propios

En la gráfica 2 se muestran los salarios promedios de obreros según su experiencia que tienen cinco años de escuela primaria a partir de la regresión de la Tabla 4. Se puede observar que los niveles de salario para la gente que hizo estudios del nivel primario han aumentado con los años; a pesar de que el efecto de un año de escolaridad sobre los salarios era superior en 2008 (6,6% vs 4,9% en el 2016), los salarios son mayores en 2016, debido a que el salario base es mayor.

### Gráfica 2. Salarios de Obreros con Primaria 2008 - 2016<sup>3</sup>



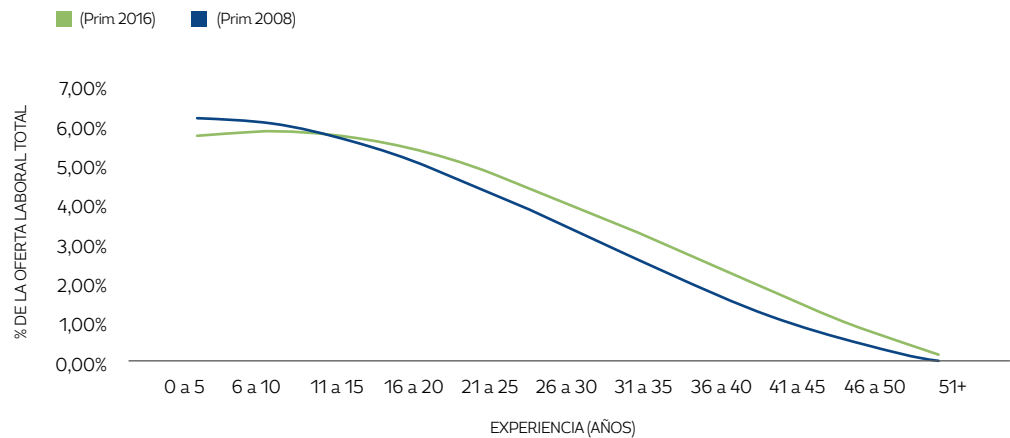
Fuente. ECV 2008/2016 DANE. Cálculos propios

Puesto que la proporción de obreros en la fuerza laboral con escuela primaria o menos ha disminuido, atribuimos el aumento en salarios de estos obreros a la reducción en la oferta de este tipo de obrero en el mercado laboral. Aparentemente, aunque estos obreros no han conseguido mucha educación, se están beneficiando del aumento en la escolaridad de los más educados porque hay menos competencia para trabajos que exige poca educación en 2016 que en 2008.

### Mercado Laboral – Escuela Secundaria

La Gráfica 3 muestra cómo ha cambiado el tamaño relativo de la fuerza laboral con 6 a 11 años de escolaridad entre 2008 y 2016. Vemos que la proporción de los obreros colombianos con este nivel de escolaridad ha aumentado considerablemente desde el 2008 al 2016, sin embargo, este efecto se empieza a evidenciar alrededor de los 11 años de experiencia.

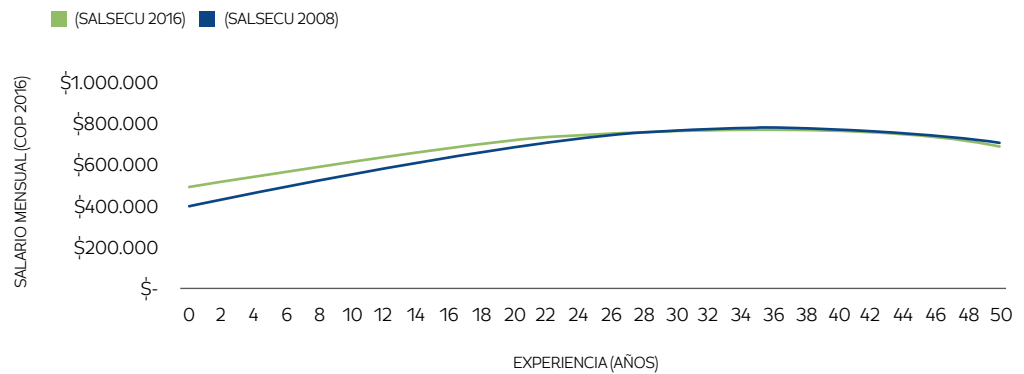
### Gráfica 3. Oferta de Obreros con Secundaria 2008 – 2016



Fuente. ECV 2008/2016 DANE. Cálculos propios

3 La muestra de datos está conformada por la población que posee entre 1 y 5 años de escolaridad. Para el cálculo de las curvas de salarios, se realiza la estimación de la ecuación (3) con 5 años de educación en el nivel primario

### Gráfica 4. Salarios de Obreros con Secundaria 2008 - 2016<sup>4</sup>



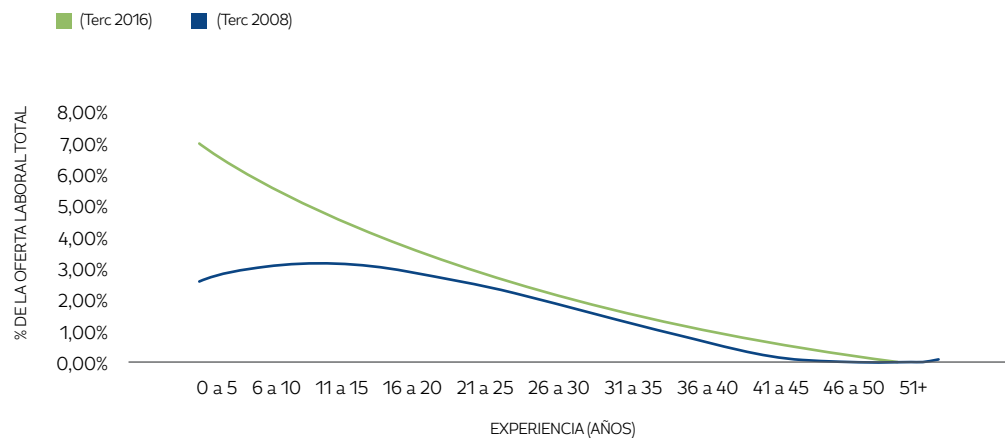
Fuente. ECV 2008/2016 DANE. Cálculos propios

Al observar la gráfica 4, se tiene que entre los 0 y los 26 años de experiencia, es mayor el salario para el 2016, ya que en este año se tiene un salario base mayor que en el 2008; en los años siguientes, se igualan los salarios en ambos períodos ya que es mayor tanto el efecto de la educación como el de la experiencia para el año 2008, lo cual se refleja en un crecimiento más rápido de la curva.

### Mercado Laboral – Escuela Terciaria

La gráfica 5 muestra cómo ha cambiado el tamaño relativo de la fuerza laboral con 11 o más años de escolaridad entre 2008 y 2016. Vemos que la proporción de los obreros colombianos con este nivel de escolaridad ha aumentado en mayor medida que en los niveles anteriores en estos ocho años, siendo el cambio más notorio en los obreros que tienen de 0 a 10 años de experiencia, con lo cual se puede concluir que son las nuevas generaciones las que se están educando.

### Gráfica 5. Oferta de Obreros con Estudios Terciarios 2008 - 2016

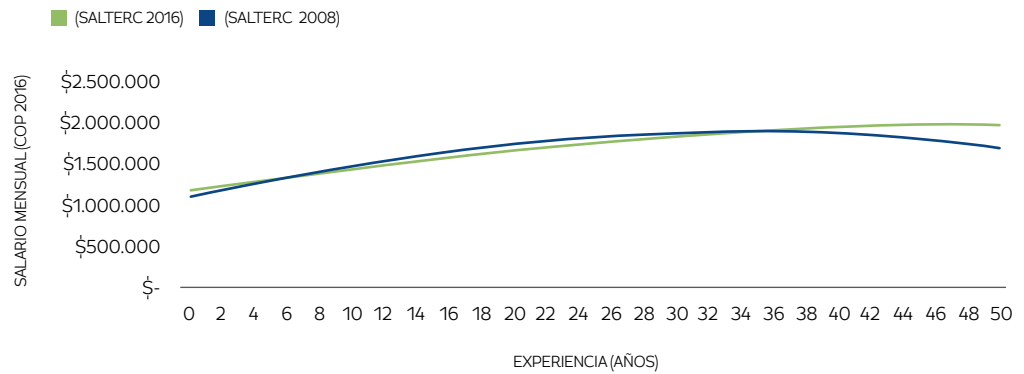


Fuente. ECV 2008/2016 DANE. Cálculos propios

En este nivel terciario, a partir de la gráfica 6, el comportamiento de la curva salarial para ambos años se puede explicar a partir de los salarios que empiezan con un crecimiento muy significativo al principio de la experiencia laboral, donde alcanza un valor muy alto, pero luego en 2008 las empresas prefieren despedirlos y contratar a jóvenes con igual o mejor educación a un menor salario, lo cual podría explicar como a los 34 años de experiencia, la curva empiece a decaer significativamente para 2008.

<sup>4</sup> La muestra de datos está conformada por la población que posee entre 6 y 11 años de escolaridad. Para el cálculo de las curvas de salarios, se realiza la estimación de la ecuación (4) con 6 años de educación en el nivel secundario.

**Gráfica 6. Salarios de Obreros con Estudios Terciarios 2008 - 2016<sup>5</sup>**

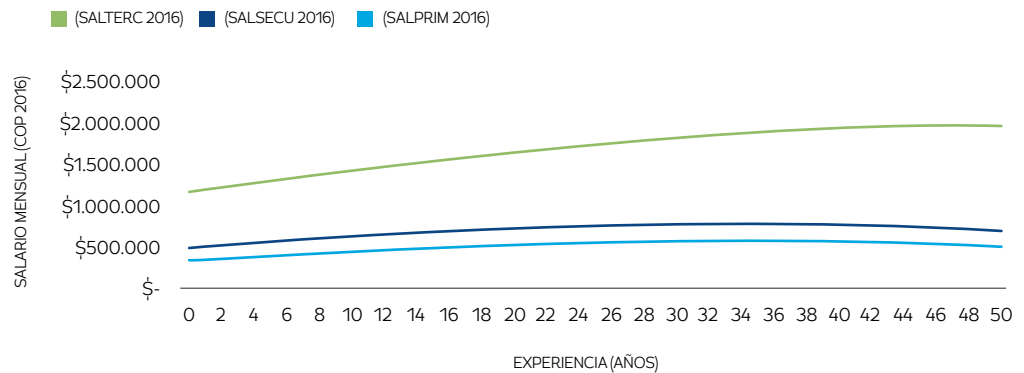


Fuente. ECV 2008/2016 DANE. Cálculos propios

Por otro lado, se tiene que la oferta laboral para este nivel en el 2008 es de 17,84% de la oferta laboral total, mientras que para 2016 es de 27,94%. Con esto queda claro que con en un periodo corto la población colombiana ha aumentado sus estudios para conseguir mejores salarios.

Es importante resaltar que, como determinó [Hernández \(2010\)](#) para este nivel de educación los salarios van a depender en gran medida de la carrera estudiada o la institución educativa de la cual se graduó.

**Gráfica 7. Modelo ampliado de Mincer 2016**



Fuente. ECV 2008/2016 DANE. Cálculos propios

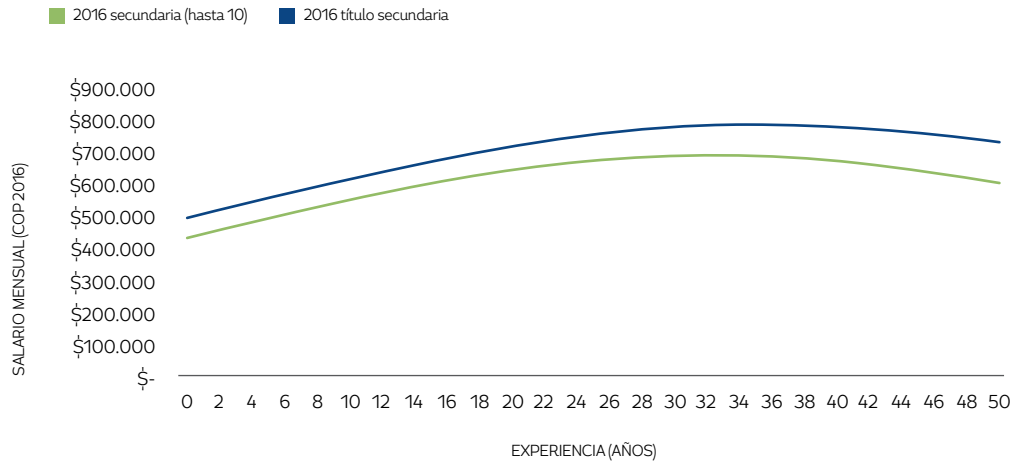
La gráfica 7 muestra el comportamiento de los tres niveles de educación para el año 2016. Se observa claramente que el nivel dominante es el nivel terciario donde se tienen los salarios más grandes de la muestra; luego, más abajo se encuentran el nivel secundario y el nivel primario respectivamente, los cuales no presentan una diferencia muy significativa entre ellos. Por lo tanto, actualmente el efecto de completar la escuela primaria y secundaria ya no tienen un efecto muy grande en el salario de un obrero, mientras que el nivel terciario marca la clara diferencia en los salarios.

<sup>5</sup> La muestra de datos está conformada por la población que posee 12 o más años de escolaridad. Este nivel incluye técnica, tecnológico, pregrado y posgrado. Para el cálculo de las curvas de salarios, se realiza la estimación de la ecuación (5) con 5 años de educación en el nivel terciario.



Ahora, se analizará el peso que tiene el título de secundaria comparando con la población que estudió secundaria pero que no la terminó. Para este análisis se hicieron dos regresiones; la primera, contando únicamente las personas que tienen como máxima escolaridad 11 años (2016 título secundaria), y la segunda la población que terminó sus estudios en 10 años de escolaridad, es decir, en 10° de bachillerato (2016 secundaria hasta 10).

**Gráfica 8. Secundaria con y sin título 2016<sup>6</sup>**

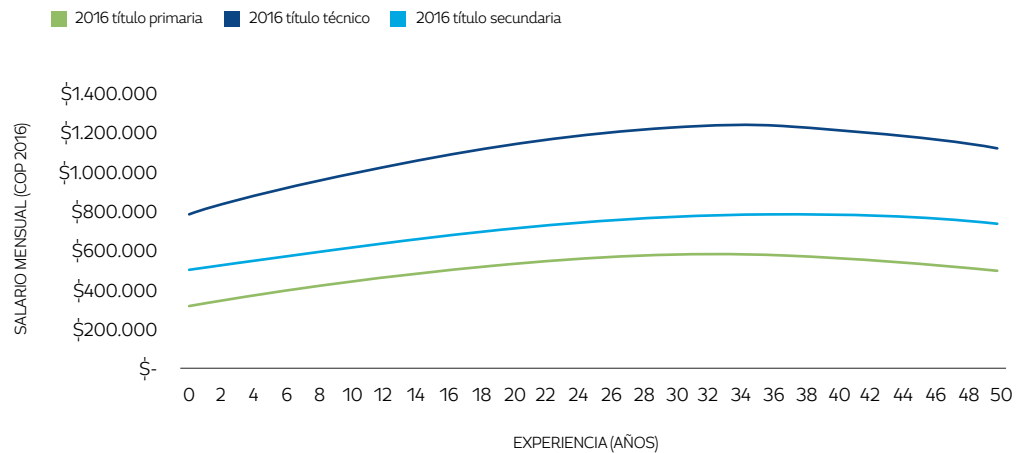


Fuente. ECV 2008/2016 DANE. Cálculos propios.

Observando la gráfica 8 se demuestra que los individuos que terminaron secundaria y obtuvieron su título, ganan más que las personas que iniciaron secundaria pero no la terminaron; para la muestra de los que tienen título se tiene un salario máximo de \$780.000, mientras que los que no tienen el título presentaron un salario máximo de \$690.000 a precios del 2016. Con esto se concluye que a pesar de que cada año de educación en el nivel secundario tenga un efecto positivo sobre el salario, el hecho de obtener el título del nivel de educación también presenta un efecto positivo en los salarios.

<sup>6</sup> “2016 título secundaria” son las personas con 11 años de escolaridad que obtuvieron el título de bachillerato. “2016 secundaria (hasta 10)” son las personas que poseen entre 6 y 10 años de escolaridad, los cuales no consiguieron el título de bachillerato, pero hicieron años de secundaria. Para el cálculo de las curvas de salarios, se realiza la estimación de la ecuación (4) con 6 años del nivel secundario.

### Gráfica 9. Título de Primaria, Secundaria y Técnica 2016<sup>7</sup>



Fuente. ECV 2008/2016 DANE. Cálculos propios.

A nivel general, al comparar el efecto de poseer un título en primaria con secundaria y técnica en el salario, queda claro la importancia de poseer educación superior ya que este nivel muestra una gran superioridad.

Por medio de la gráfica 9, se tiene que la curva de las personas con título de primaria y secundaria tiende a ser más plana con un crecimiento relativamente lento, ya que las personas que no tienen estudios superiores no tendrán variaciones muy significativas de salario en su vida laboral. Por su parte, la curva de título de técnica comienza con una pendiente más inclinada hasta los 26 años de experiencia, lo cual refleja que cuando la persona se encuentra en este rango de experiencia es que tendrá los mayores efectos de la educación sobre el salario; cuando esta curva se aplanan alrededor de 34 años de experiencia, se podría explicar porque las personas ya culminan su ciclo laboral donde sus salarios tornan a ser constantes, o las personas se pensionan, o comienzan a trabajar medio tiempo devengando así un salario menor.

## 5. Conclusiones

La escolaridad presenta efectos positivos sobre el salario para cada nivel de educación tanto para el 2008 como para el 2016, donde el nivel terciario es el que presenta mayor efecto, luego el nivel secundario y por último el nivel primario. Además, se concluye que en la realidad el efecto de la experiencia es diferente para cada nivel de educación, demostrando así que el supuesto utilizado por Mincer en su modelo de que el efecto de la experiencia es igual para todo nivel de educación no se ajusta a la realidad; a medida que la persona aumenta su escolaridad, el efecto de la experiencia en los salarios va a ser mayor. Por otro lado, se observó que en el 2016 la oferta de empleados educados es mucho más grande que en el 2008, lo cual evidencia que actualmente los salarios aumentan en menor medida con el tiempo.

<sup>7</sup> “2016 título primaria” está conformado por las personas con 5 años de escolaridad, obteniendo el título de primaria. “2016 título secundaria” está conformado por las personas con 11 años de escolaridad, obteniendo el título de bachiller. “2016 título técnica” está conformado por las personas que obtuvieron título de técnica (nivel 7 en la tabla 2), los cuales poseen 13 o 14 años de escolaridad. Para el cálculo de la curva de título primaria, se realiza la estimación de la ecuación (3) con 5 años del nivel primario; para título secundaria, se realiza la estimación de la ecuación (4) con 6 años de nivel secundaria; para el título de técnica, se realiza la estimación de la ecuación (5) con 4 años de nivel terciario.

El mayor efecto de la educación superior sobre los salarios se da hasta los 34 años de experiencia de la persona, reflejado en la gráfica 9, donde a partir de los 26 años de experiencia el salario tiende a estabilizarse o a disminuir dependiendo de lo que elija la persona en su futuro laboral.

El hecho de terminar y conseguir el título en el nivel de educación tiene un efecto positivo adicional en los salarios, donde las personas que obtuvieron el título de bachiller tienen un mayor salario que las personas que iniciaron secundaria pero que no la terminaron. Sin embargo, como estableció [Hernández \(2010\)](#), los salarios en el nivel terciario van a depender en gran medida de la carrera estudiada y la institución educativa a la cual asistió.

La proporción de obreros con primaria o menos bajó de 46% a 32,5%, con esto se puede deducir que las personas que tienen este nivel como su máxima escolaridad están disminuyendo debido a las altas expectativas del mercado laboral. Por otro lado, la proporción de obreros con escolaridad de secundaria aumentó de 36% a 39,5%, y con terciaria aumentó de 18% a 28%, especialmente en el nivel terciario se puede apreciar una gran diferencia en las personas jóvenes, con lo cual se puede concluir que las nuevas generaciones están llevando sus niveles educativos hasta el nivel terciario lo cual los hará más atractivos para el mercado laboral.

Al comparar los resultados obtenidos en el 2008 con los del 2016, se tiene que los efectos de la escolaridad sobre el salario disminuyen a través del tiempo para cada nivel de escolaridad, siendo mayor la caída en el nivel primario, luego en el secundario y por último en el terciario. Del mismo modo, el efecto de la experiencia disminuyó del 2008 al 2016 para cada uno de los niveles de educación. Sin embargo, el nivel base de los salarios aumentó para todos los niveles en el 2016.

Se nota que hay muchos cambios de oferta y demanda por el nivel de escolaridad de los obreros entre 2008 y 2016. La oferta de obreros con poca escolaridad ha caído en el mercado proporcionalmente (fracción de obreros), causando un aumento en sus salarios. En cambio, la oferta de obreros con Bachillerato o pregrado ha aumentado. Sin embargo, no vemos una caída en los salarios de obreros con estas capacidades superiores, implicando que la demanda para obreros con estas calificaciones también ha subido.

## Referencias

- Heckman, J. J., Lochner, L. J., & Todd, P. E. (2008). Earnings functions and rates of return. *Journal of human capital*, 2(1), 1-31.
- Hernández, G. A. (2010). ¿Cuán rentable es la educación superior en Colombia?. *Lecturas de Economía*, 73, 181.
- Merlo, J. J. (2009). Retornos a la educación durante una depresión económica. Evidencia empírica para la Argentina. *Santiago de Chile: Instituto de Economía Universidad Católica de Chile*.
- Mincer, J. (1975). Education, experience, and the distribution of earnings and employment: an overview. In *Education, income, and human behavior* (pp. 71-94). NBER.
- Mineducación. (2016). Educación en Colombia. OCDE.
- Montenegro, C. E. & Patrinos, H. A. (2014). *Comparable Estimates of Returns to Schooling Around the World*. Banco Mundial.
- Psacharopoulos, G. (1979). On the weak versus the strong version of the screening hypothesis. *Economics letters*, 4(2), 181-185.

# Instructions for Authors

## Aims and Scope

Ecos de Economía is an international journal of applied economics, with a particular interest in the quantitative analysis of economic, financial, and public policy issues and/or themes related to Latin America. Articles that include and analyze national data are particularly welcome, since these data are often unavailable to international researchers outside the region. The journal is not interested in manuscripts that are solely conceptual in nature.

## Preferred and Acceptable Languages

Manuscripts in English are preferred, but manuscripts in Spanish are acceptable. Authors should be aware that articles in English have an international audience, while articles in Spanish have a regional audience. Communication with authors and reviews of manuscripts will proceed in the language in which the authors initially contact the journal and/or in the language used in the manuscript.

## Publication Schedule

Ecos de Economía is published in June and December, but articles will be posted on-line shortly after acceptance.

## Submission of a Manuscript Precludes Submission to Another Journal

Authors sending manuscripts to the journal must provide exclusive rights of publication to Ecos de Economía. As a consequence, submission to the journal precludes the authors from submitting the manuscript to any other journal for review until they are notified by Ecos de Economía that their manuscript is no longer under consideration for publication. Authors will be required to provide a copyright to the journal prior to publication.

## Procedure for the Review of Submitted Manuscripts

The journal follows a two-part process to provide an efficient review of submitted manuscripts. The Editor-in-Chief, in consultation with the appropriate Associate Editor(s), first reviews the manuscript to determine whether it potentially could provide a valuable contribution to the applied economics literature. In making the initial decision he considers the pertinence of the topic, the validity of the analysis, and the quality of the presentation. If he finds the manuscript promising, he sends it to a referee in the appropriate field for a thorough, written review.

In the event that the Editor-in-Chief rejects the manuscript after the initial review, the corresponding author is notified of the decision and is not provided with a written review. If the manuscript is sent for a thorough review, this review will be provided to the authors regardless of the decision on publication following this review.

## Acceptable Format for Submitted Manuscripts

Articles should be provided in electronic form in WORD or PDF and sent to journal's email address. If the article includes photos or graphics, they should be sufficiently clear to permit evaluation of the manuscript. Better quality photos or graphics may be required for publication. When the articles possess mathematical information and/or econometrics these should be done directly in letter Swis 721 LtCnBt, size 10.5 with the formulas solely in Word (Microsoft Editor in equations 3.0) to avoid problems of presentation. Tables should be with tabs and not have `insert table`. The article should not exceed 10,000 words, and shorter articles are preferred. Manuscripts should be provided in 1.5 spacing and in 12 point font. In special cases the editorial committee will accept initial manuscripts in a different format.

## Each article should contain the following:

- A title in English. If the manuscript is in Spanish, a title in English must also be provided).
- An abstract in English of no more than 100 words describing the objective, methodology and principal conclusions. If the article is in Spanish, the abstract must be provided in Spanish and English.
- Key words (Spanish and English); minimum three, maximum five.
- JEL classification code(s); at least one
- Details of the author: name, academic level, email address and postal address, institutional affiliation (current institution of work, faculty, department, office, etc). If belonging to a research group, the name of the group should be noted.
- The article should contain at least five sections that include an introduction, the methodology, the results, the conclusions, and references. Articles presenting model results should consider including the relevant data in an appendix.

## Quotes and References

Ecos de Economía adopts the APA 6ta. ed. system for quotations and references.

- Short textual quotes of 3 lines or more are incorporated in the text using quotation marks to indicate them. Longer textual quotes are separated from the text and tabbed from the left margin without need for quotation marks. In both cases, the source should be cited completely, eg. author, year and number of page (Ortiz, 1999, p.2). Quotes are eliminated from the foot of the page, except in using them as clarifying notes. At the end of the article are listed all bibliographical references used in alphabetical order.
- Paraphrases should contain only the author's name and year of publication (Ortiz, 1999).
- If the work has more than three authors, they are cited the first time with all the surnames, and thereafter with the surname of the first author only, followed by the phrase et al.

- If there are more than six authors, et al. is used from the first mention.
- At the end of the article should appear the bibliographical references (only those used inside the text). They are organised in alphabetical order while works by the same author are ordered chronologically using French indentation with double spacing.

### **Books**

Surname, Name. (Year). Title. City: Editorial.

### **Capital of books or works of reference**

Surname, N. & Surname, N. (Year). Title of chapter or entry. In Surname, N. (Ed.), Title of book. (pp.xx-xx). City: Editorial.

### **Books in electronic form**

Surname, N. (Year). Title. Sourced from <http://www.xxxxxxxx.xxx>

### **Articles from printed journals**

Surname, N. (Year). Title of article. Title of publication, volume (number), pp. xx-xx.

### **Articles from the Web**

Surname, N. (Year). Title of article. Title of publication, volume (number), pp. xx-xx. Sourced from <http://www.xxxxxxxx.xxx>

### **Articles of daily publication, from the Web**

Surname, N. (Year, day of the month). Title of article. Title of publication. Sourced from <http://www.xxxxxxxx.xxx>

Please submit papers online to: [www.eafit.edu.co/ecos](http://www.eafit.edu.co/ecos)

**UNIVERSIDAD**  
**EAFIT**<sup>®</sup>