Exploring Roots of Inequality in Latin America and Peru

Feridoon Koohi-Kamali, Editor

PETER LANG
PROMPT

This book explores Latin American inequality broadly in terms of its impact on the region's development and specifically with two country studies from Peru on earnings inequality and child labor as a consequence of inequality for child labor. The first chapter provides substantial recent undated analysis of the critical thesis of deindustrialization for Latin America. The second chapter provides an approach to measuring labor market discrimination that departs from the current treatment of unobservable influences in the literature. The third chapter examines a much-neglected topic of child labor using a panel data set specifically on children.

The book is appropriate for courses on economic development and labor economics and for anyone interested in inequality, development and applied econometrics.

Feridoon Koohi-Kamali studied economics at the University of London and received his doctorate in economics from Oxford University. He has taught at Oxford and a number of US universities including Columbia. Currently Associate Professor at The New School for Social Research, he has also been a consultant to the World Bank and a research associate at the Schuman Center for Advanced Studies, European University Institute. His research interests are in applied econometrics, public policy and development.

www.peterlang.com

Exploring Roots of Inequality in Latin America and Peru

PETER LANG
PROMPT

PETER LANG
New York • Bern • Berlin
Brussels • Vienna • Oxford • Warsaw

Exploring Roots of Inequality in Latin America and Peru

Feridoon Koohi-Kamali, Editor

PETER LANG
New York • Bern • Berlin
Brussels • Vienna • Oxford • Warsaw

Library of Congress Cataloging-in-Publication Data

Names: Koohi-Kamali, Feridoon, editor
Title: Exploring roots of inequality in Latin America and Peru /
edited by Feridoon Koohi-Kamali.
Description: New York: Peter Lang, 2021.
Includes bibliographical references and index.
Identifiers: LCCN 2021034000 (print) | LCCN 2021034001 (ebook)
ISBN 978-1-4331-8989-0 (hardback)
ISBN 978-1-4331-9129-9 (ebook pdf) | ISBN 978-1-4331-9130-5 (epub)
Subjects: LCSH: Equality—Latin America. | Equality—Peru. | Latin
America—Economic conditions—21st century. | Peru—Economic
conditions—21st century. | Labor market—Latin America. | Labor
market—Peru. | Child labor—Latin America. | Child labor—Peru.
Classification: LCC HN110.5.A8 E97 (print) | LCC HN110.5.A8 (ebook) |
DDC 306.098—dc23
LC record available at https://lccn.loc.gov/2021034000
LC ebook record available at https://lccn.loc.gov/2021034001
DOI 10.3726/b18766

Bibliographic information published by **Die Deutsche Nationalbibliothek**.
Die Deutsche Nationalbibliothek lists this publication in the "Deutsche
Nationalbibliografie"; detailed bibliographic data are available
on the Internet at http://dnb.d-nb.de/.

This publication has been peer reviewed and meets
the highest quality standards for content and production.

Table of Contents

List of Tables

CHAPTER ONE

CHAPTER TWO

CHAPTER THREE

List of Figures

CHAPTER ONE

Introduction

This volume deals with inequality. Isolating the roots of inequality in a region as vast and diverse as Latin America requires an undertaking beyond the narrow focus of the present volume. Most studies on Latin America point to the growing diversity in the region's labor markets; the Mexico/Central American region is characterized by informality, outmigration, and low productivity, while the Southern region has benefitted from higher average levels of schooling and less ethnic diversity, though overall, the region has very high levels of income inequality. Nonetheless, this volume demonstrates, both from a macroeconomic perspective of the entire region and via two country-level studies, that a few prominent causes of inequality are pervasive and empirically evident. The evidence presented here, while far from comprehensive, points to three specific causes of inequality that are suggestive of shared features of inequality across Latin America.

First there are sectoral differences in productivity that are closely related to declining shares of labor in agriculture and increases in the share of small firms operating in the informal services sector. Second, the growth of the informal services sector is the result of the migration from rural areas to the urban informal service sector because of a relative decline of manufacturing and deindustrialization. Third, compared to observations from East Asian economies, this trend has prevented the services sector from playing a leading role in Latin American economic growth due to the size of the high-skilled services sector in Latin America being relatively small; that is, the region does not have the average levels of education and the skilled labor to support a dynamic services sector that is capable of absorbing a high proportion of surplus labor. Hence, inequality as employed in this volume either results in sectoral differences in earnings or, at least, has a significant impact on labor market outcomes. The first chapter in this collection is a macroeconomic examination of aspects of Latin American inequality in terms of these factors. Using microdata, the second chapter, a country study, follows with an examination of the causes of earnings inequality. The third chapter, also a country study, focuses on child labor stemming from household shocks and inequality in terms of household wealth, education, etc. While the country studies employ wider sets of explanatory variables, both chapters empirically demonstrate the significant impacts of informal labor, rural migration, and education on labor market-related inequalities. While both country studies examine Peru and, thus, admittedly narrow the diversity of data sources, it is notable that Peru is also a country for which more extensive good quality data are available from international bodies. For that reason, it is often chosen for research on inequality and development change in Latin America. The longitudinal data from Peruvian households that is employed in the final chapter reflects this assessment.

Latin America has suffered a decline in labor productivity during recent decades. This has had significant consequences for the region's economic growth and income inequality. A major issue that is prominent in the field is the cause of the region's productivity decline. In many developing countries, productivity declines in recent decades have corresponded with a major shift to services provision. This is due to their manufacturing sectors having been prematurely exposed to international competition as a result of globalization. It is also in contrast to developed economies that outsourced industrial production abroad and focused on expanding their services sectors as a deliberate policy choice. The chapter by Bezares Calderón investigates this premature deindustrialization hypothesis for Latin America, explores the impact the shift has brought about in the sectoral composition of Latin American economies, estimates the distributional consequences of this shift, and offers evidence of increased income inequality.

The evidence from the existing research that Bezares Calderón examines singles out bulging informal service sectors as the main problem for Latin America's economic development; over 1954–2011, the informal sector accounted for 53% of the jobs in Latin America. The three factors listed above are closely related to Latin America's inequality. Rodrick (2013, 2015) sums up their relationship as follows. Deindustrialization limits the capacity of the manufacturing sector to absorb rural labor migration. One successful alternative, adopted by East Asian economies, is to promote the services sector as the engine of growth. However, for that to happen, the sector must be able to rely on highly skilled labor which requires high levels of education. The Latin America services sector does not have sufficient high-skilled labor to act as a leading engine of economic growth. As a result, rural migration has led to the rapid growth of a large informal sector of self-employed/small firms with low levels of productivity. Due to outmigration, the share of agricultural labor in the region has declined, manufacturing has absorbed less than 20% of these workers whereas the services sector has absorbed most, mainly in informal services. This combination of (a) rural migration, (b) an informal sector that is increasing in size, and (c) relatively low levels of skilled labor and education have resulted in significant sectoral inequality in earnings and incomes in Latin America. The author provides empirical evidence on the contribution of these factors to the Latin American pattern of inequality and development.

Bezares Calderón first spells out the econometric methods employed to test for evidence of joint movement between the time-series of productivity and industrial growth. The chapter then explores premature deindustrialization in terms of the growth rate of labor in manufacturing as a share of total employment of other sectors, a particularly important step for capturing the overall effect of the informal service sector on lagging productivity. Finally, the paper explores the scope of wage inequality by regressing the Gini Index of income inequality on employment for each of the main sectors defined by the population share employed. The author applies this empirical strategy to nine Latin American countries (Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Mexico, Peru and Venezuela) employing, among others, data from the Penn World Table and from the World Income Inequality Database.

The co-integration test suggests there is no long-run joint movement between labor productivity and manufacturing growth. Moreover, separately estimating a two-equation productivity-growth model of deindustrialization for each country reveals that increases in imports reduced the size of the manufacturing sector in Latin America during the 1980s and 1990s. The paper then moves to identify the distributional effects of deindustrialization, presenting empirical evidence that while the shares of labor employed in the service sector and in agriculture correspond with higher levels of sectoral income inequality, with Gini measures of 0.34 and 0.11 respectively, and the Gini Index for the share of labor in manufacturing recording a reduction of 0.4 points. This outcome highlights how Latin America's inequality is generated by differences in the sectoral shares of labor as outlined earlier. The author lists several caveats, including the difficulty of obtaining accurate data on the size of the large informal sector in Latin America.

The next two chapters examine country-specific data and demonstrate the perversive influences of informal migrant labor and education/skills on inequality. Thus, the chapters complement the initial macroeconomic-focused study by providing further evidence on some related influences that are not easily ascertained from macro-level data. For example, the second chapter includes the effect of unionized labor, a feature of the formal sector, on wage inequality, while the third chapter is devoted to child labor, an important type of informal labor.

The chapter by Koohi-Kamali investigates gender and ethnic wage discrimination for Peru. Measuring discrimination requires separate identification of the bias in the labor market between otherwise similarly productive workers from that which is related to differences in productively and generated outside the labor market—for instance, the result of pre-labor market entry differences in school quality. The author argues that many features required for that identification are unobservable in the data. Evidence on the sources of discrimination in Peru's labor market suggest several factors that are potentially unobservable in a typical survey—for example, productivity and the quality of education, or job discrimination by region of birth. This chapter focuses on correcting for estimation bias that stems from unobservable influences on earning differentials by ethnicity and gender.

The available research on Peru's labor market emphasizes the importance of such unobservable influences on pay discrimination, yet research on such influences is surprisingly missing in the literature. One obvious approach is to regress individual wages on observables such as age, education, gender, ethnicity, etc. (Nopo, 2009). The problem with this approach is that it fails to control for unboservables that may cause correlation of the estimated error term with the observable explanatory variables, rendering inconsistent estimates. Of course, not all unobservables are related to labor market discrimination but if measurement of labor market discrimination is the primary research interest, then not controlling for the unobservables may compound the endogeneity problem. The author employs a Heckman (1976) two-step selectivity bias model with a labor market entry equation and a separate wage equation to test and control for the effects of unobservable influences on wage differentials between indigenous and non-indigenous Peruvians and between Peruvian men and women. However, the author warns that this method cannot separate the impact of discrimination on pay from that of non-discrimination-related factors; for instance, differences in the quality of education relating to observable years of education would vary from ethnic discrimination. Moreover, the study points out that effective application of this approach requires an exclusion restriction(s) with variables that appear in the first stage, an indicator dependent variable to represent whether the individual enters the labor market but is excluded from the wage equation. This approach is then applied to two Living Standards Measurement Study surveys of Peru from 1985 to 1994 to estimate a baseline OLS model of wage determination and two versions of the two-step selectivity model, with and without exclusion restrictions in the wage equation. Furthermore, the chapter also employs instruments for potential endogeneity of hours of labor supplied as a key explanatory variable in the wage equation.

The results reject the hypothesis of no selectivity bias with both versions of the selectivity model employed. This implies that endogeneity problems are due to the correlation of the unobeservables, as a part of the error term of the equation, with the baseline model explanatory variables. Employing age and age-squared for exclusion restrictions, the two-step model estimates are well-determined; all variables are statistically significant and of the expected signs, in addition to correction for hours of labor supply endogeneity instrumented by earning from "payment-in-kind". In particular, both gender and indigenous status negatively and significantly affect individuals' earnings. The author computes the gender differences between the earnings of individuals with otherwise identical characteristics (replacing a male worker with a female worker). Without controlling for unobservable selectivity bias, the difference is -13.5% compared to less than -10% with the selectivity control. Similarly, if an otherwise identical white worker is replaced with an indigenous worker, the difference in earnings would be -20% without the selectivity bias control and -16.5% with selectivity control using exclusion restrictions.

It is notable that the evidence for the three fundamental influences on Latin America's inequality from the macro paper are also present in this study. All levels of education are found to have significant positive effects. All different sectors of the economy also have significant positive effects (relative to the agricultural sector, which serves as a base of comparison). Finally, the migration dummy is negative and significant, particularly in the wage equation models with or without unobservable controls. The latter can be explained by the concentration of small informal services in Lima. Hence, the micro evidence on the impact of these factors on the labor market have the expected direction displayed in the first macro chapter, confirming the critical impacts of migration, education, and informal labor on inequality. Whilst there is room for improving these measures of wage discrimination by using more effective exclusion restrictions in a future study—for instance, using spouse earnings or spouse education level—the differences reported in this study are large and suggest that not controlling for unobservable variables produces estimates of earnings discrimination that are inconsistent and misleading.

The chapter by White and Rouleau examines an important form of informal labor, namely child labor. This work focuses on explaining how household shocks affect child labor incidence; not only in terms of the child characteristics but also of household-level inequality. A prominently feature of this study is in highlighting the evidence that an increase in the level of child labor supply is affected, in part, by inequality of household wealth and asset ownership, differences in household quality, and the education level of the household head. The topic of how economic shocks specifically affect child labor has received scant attention in development economics research. However, a few papers have demonstrated how shocks lead to increases in the incidence of child labor due to agricultural losses in Tanzania (Beegle et al., 2006) and Nigeria (Acheampong and Huang, 2018) and, more notably, due to natural disasters in Guatemala (Vasquez and Bohara, 2010). The chapter extends the literature by providing a broad empirical investigation of this relatively neglected aspect of child welfare in Latin America, specifically by addressing four critical questions affecting child labor in Peru. First, whether children in households experiencing shocks are more likely to engage in child labor; second, whether child labor incidence varies significantly by the number of shocks experienced; third, whether incidence of child labor varies significantly by the type of shocks (e.g., economic, environmental, etc.); and, finally, how the incidence of child labor varies with specific observable characteristics of child, parents and household.

White and Rouleau address these questions by estimating four binomial logit models which are applied to data from five rounds of panel surveys for Peru that have been conducted by International Study of Childhood Poverty over the years 2002–2017. Child labor incidence is identified as having resulted from each household shock experienced since the most recent survey date. The authors then define a series of binary variables for the incidence of child labor that is based on that definition. Indicators for whether the child's family has experienced a shock(s), indicators for the number of shocks (up to five), indicators for five different types of shock(s), and how the incidence of child labor is affected by the characteristics of the child, her/his parents, and their household. The results of this study indicate a consistent positive and statistically significant relationship between household shocks and increased child labor incidence in Peru. Bearing in mind that child labor is an important type of informal sector labor, it must be noted that this study confirms the significant impacts of parental education, rural/urban location, and household wealth inequality on the incidence of child labor with both education and wealth highly significant in all models of child labor examined. The coefficient identifying rural household location is also significant in the models where paid child labor and unpaid household tasks are used as the dependent variable series. The effects are all in the expected directions suggested by the earlier studies in this volume that highlight the key influences on inequality.

The chapter reports a significantly positive log-odds coefficient of the number of shocks on child labor of 0.0699 magnitude and a positive and significant coefficient estimate for the effects of differences in the number of shocks of 0.2358. These findings are broadly in line with results reported in earlier studies. Second, the study reports a positive and significant impact of the number of shocks on child labor; for example, taking no shock as the base, a household experiencing only one shock since the last survey round is estimated to realize an increase in the probability of child labor by 2.12%, while two shocks increase that probability by 6.32%. Third, and most important, the authors report significant and positive impacts of economic shocks and family shocks on child labor incidence that are equal to 3.34% and 1.86%, respectively. Moreover, the authors present similar effects for various types of child labor (i.e., paid labor, unpaid labor, domestic care, etc.). Finally, the chapter also reports that 48% of child-specific, and 70% of the parents-specific and household-specific variables are statistically significant in all regressions. All in all, this is a welcome addition to the scant literature on a predominantly informal form of labor in relation to unexpected economic shocks. While it raises further questions that need to be addressed, it also provides a guide for future research on an issue for which an understanding is critical to designing public policy on child welfare in Latin America.

Summing up, this volume provides evidence that sectoral differences related to informality, education, and low productivity have significant impacts on Latin American inequality at a macroeconomic level and that they are also significantly influence labor market earnings inequality and help to explain child labor incidence.

I would like to acknowledge the very helpful comments I received from the two anonymous referees and to thank the authors in this volume for their contributions.

Feridoon Koohi-Kamali

CHAPTER ONE

Labor and Deindustrialization in Latin America: A Look at Productivity, Globalization and Inequality

ALMA A. BEZARES CALDERÓN[1]

1. INTRODUCTION

Industrial activity has traditionally been considered an important determinant of economic growth and productivity. The Industrial Revolution and the Asian miracle are just two examples of how manufacturing is related to fast economic growth. This sector is supposed to provide the necessary elements for growth and unconditional convergence, such as generating economies of scale through technology, the capacity to employ large shares of the population, the opportunities for capital accumulation, and the possibility of creating a comparative advantage (Solow, 1956).

According to the Dual-Sector Theory, the modernization process is characterized by a transfer of workers from the agricultural sector towards the manufacturing industry (Lewis, 1954). With this shift, labor surplus would be absorbed, and there would be gains in productivity that would lead to economic growth. Furthermore, Modernization Theory argues that the manufacturing sector would be the pillar of development. Once productivity gains are high enough (allowing for firms to produce more with fewer workers) and consumers become more affluent, the period of 'mass consumption' would arrive, leading to a decrease in

1 abezares@whittier.edu.

the size of the manufacturing sector (Rostow, 1959; Ramaswamy and Rowthorn, 1997; Krugman, 1996; Rodrik, 2015).

Yet, in recent times, countries that had not reached a high level of economic development have already seen the participation of their manufacturing sector shrank, as attested by Amirapu and Subramanian (2015) for the case of India, and Hamid and Khan (2015) for the case of Pakistan. In this chapter, I focus on the case of Latin America to verify how the evolution of total factor productivity (TFP) in the region, measured by the annual change of output that is not accounted by the growth in inputs, is associated to changes in the demand for labor in the manufacturing sector. If positive changes in productivity are linked to changes in the manufacturing sector, then the modernization theory would be explaining these changes. If however, as posited by authors such as Rodrik (2015), the region has not shown increases in productivity and changes in the participation in the agricultural sector have not been translated into a shift towards industrial activity, then we may be attesting an early deindustrialization process in which changes in productivity are not related to labor demand in the manufacturing sector. Furthermore, I look into how the current global configuration has affected this process and the consequences this has on income inequality across the region.

Previous research on productivity in Latin American countries has linked issues related to the market structure, corruption, and weak property rights to the lack of productivity (Cavalcanti Ferreira et al., 2013). However, little research has focused on how deindustrialization and productivity might be associated, nor has the potential link of this deindustrialization process on income inequality received much attention.

This chapter explores the case of nine Latin American countries with different economic and political configurations[2]. The analysis focuses on the period between 1954 and 2011. The data used comes from the Groningen Growth Development Center, as well as from the Penn World Tables 9.1 and the World Income Inequality Database (WIID). One of the main challenges of this study is the large informal sector that these countries have, accounting for 53.1% of the jobs in the continent, and 49.1% of the jobs in the industrial sector (Latina et al., 2019). This means that there might be informal activities not accounted for in the GDP that, if considered, would change the structural configuration of the economy. Besides, large informal sectors are correlated with low levels of productivity, and the results may be downward biased.

The chapter broadly explores the patterns of productivity, economic growth, and manufacturing activity in the Latin American countries examined. It also examines the role of globalization as a potential cause of the deindustrialization

2 Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Mexico, Peru, and Venezuela.

process by looking if the share of international trade. One of the main arguments on deindustrialization posits that globalization may reduce the manufacturing sector's size, as this process is outsourced to coun tries with a comparative advantage in this sector (Rodrik, 2013; Froibel et al., 1980). Rodrik (2015) argues that globalization exposed many low- and middle-income countries before they could strengthen their comparative advantage in the manufacturing sector, inducing a premature deindustrialization process. Globalization may also cause a shift in the structure of the economy, leading to an increase in labor demand of other sectors, such as the services industry. Still, Latin America is a region of small firms, both in manufacturing sectors and in the other areas in the economy. Just as with informality, this also leads to labor misallocation and low levels of productivity that perpetuates across sectors (IADB, 2010).

The chapter follows different strategies. First, a cointegration analysis verifies the relationship between deindustrialization and productivity. Then, a panel-data analysis ex- amines how international trade affects changes in the share of the manufacturing sector, and how changes in the value-added per worker in the manufacturing and services sector explain the deindustrialization process, following Rodrik's (2013, 2015) premise. Finally, a panel-data specification estimates the relationship between deindustrialization and income inequality.

The study finds no long-run equilibrium relationship between the size of the manufacturing sector[3] and total factor productivity[4]. Contrary to what Modernization theory predicts, there is no close relationship between the evolution of productivity and industrialization. This finding goes in line with the literature that sustains that, for developing countries, deindustrialization was imported from abroad.

When looking at the impact of globalization on deindustrialization, the share of imports in GDP has a negative effect on changes in the manufacturing sector. The effects are more pronounced during the decades of 1980 and 1990. There is a clear relationship between deindustrialization and income inequality, particularly in a context of mix of low- and medium- level productivity firms. Whereas areas with higher shares of labor working in the manufacturing sector show lower levels of income inequality, increases in the share of labor in the services sector are linked to higher levels of inequality. However, the results become less clear once analyzed across time. Still, the results are intuitively sound and they provide important information for public policy.

Learning more about the configuration of the economies in Latin America is crucial for them to secure a path of extensive and sustainable growth. Although the deindustrialization process does not immediately restrict economic growth,

3 Measured as the share of labor that it absorbs.

4 Measured in relative terms, with 2011=1.

it does limit the unconditional convergence advantages that the manufacturing sector provides (Solow, 1956; Rodrik, 2013). Therefore, finding other paths to secure sustainable growth may be necessary (Amirapu and Subramanian, 2015; Hausmann et al., 2007).

The rest of the chapter is structured as follows. The following section introduces a literature review on deindustrialization, emphasizing the research on developing countries. Section 3 introduces the case of Latin America, showing the patterns of productivity and industrialization in the countries analyzed. Section 4 introduces the empirical strategy of the chapter, and Section 5 analyzes the results. Section 6 concludes.

2. ANALYTICAL FRAMEWORK

Traditionally, manufacturing has been the leading engine of growth. In 1928, Young (1928) introduced the concept of *macroeconomies of scale*, referring to the macro spillover effects of the manufacturing sector. These ideas influenced Kaldor (1966, 1967) who indicated that manufacturing had to grow faster to absorb the surplus labor from the agricultural sector, as Lewis theory on economic growth predicted in 1954, and produce faster economic growth (Lewis, 1954).

Kaldor established that faster growth of manufacturing output could fasten manufacturing productivity, creating dynamic economies of scale derived from a 'learning by doing' process' (Arrow, 1962). This process, not usually seen in the services sector or agriculture, could create incentives at the national level, as manufacturing products have a higher in- come elasticity of demand (Dasgupta and Singh, 2007).

In that vein, Rodrik (2013) examines the case of 118 countries and concludes that manufacturing provides an unconditional convergence of productivity that does not depend on geographic, economic, or political factors. In terms of the neoclassical growth theory, this means that unconditional convergence even if not present in the economy as a whole, will be present in the manufacturing sector Solow (1956), and Rodrik (2013). Manufacturing activities are technologically dynamic, provide unconditional labor productivity convergence, absorb unskilled labor, and provide expansion opportunities given their tradeable characteristics (Rodrik, 2015).

Considering the evidence, manufacturing seems to be the sector that is best equipped to promote economic development and lead countries out of poverty. This opportunity, however, may not be available to low- and middle-income countries. Contrary to what Modernization Theory predicts, these countries are showing a reduction of the share of the manufacturing sector before they have reached a high-income level (Rostow, 1959). Thus, these countries do not seem to obtain the inverted-U shaped curve between industrial production and economic development that exists in high-income countries (Rodrik, 2015).

Instead, this turning point on manufacturing activities is starting at lower levels of eco- nomic development. While for developed countries, deindustrialization starts at per capita income levels of US$10,000, in some developing countries, it starts at levels of US$3,000 (Dasgupta and Singh, 2007). In terms of employment, the threshold is located around a share of 30% of total employment (Amirapu and Subramanian, 2015). These results indicate that premature deindustrialization starts at levels of very low industrialization.

Rodrik (2015) argues that whereas deindustrialization in high-income countries represents a change linked to technological improvements, demographic trends, and income, the case in low- and middle-income countries has a different root. In this case, deindustrialization has to do with a premature exposure to global markets. This is, developing countries became exposed to international competition without developing a comparative advantage in the industrial sector. As price takers, they were not ready to compete at low international prices. The region that has struggled the most from this decline of the manufacturing sector is Latin America, particularly in terms of manufacturing employment and value-added.

Different authors have captured this phenomenon in the region. Be´ne´trix et al. (2015) document that, while Latin America was the first region in the periphery to catch up at the end of the 1800s and beginning of the 1900s, the levels of manufacturing activity have decreased since the 1990s and approached levels of activity similar to those of developed countries, risking the growth of the region. Rodrik (2015) also notices the presence of deindustrialization in the Latin American region. He argues that this is closely related to the liberalization policies followed during the 1980s and 1990s.

However, even if there is a generalized agreement regarding the possibilities of the manufacturing sector to be the most important engine for growth, scholars have recognized the possibilities for the services sector to provide growth opportunities. Ghani and O'Connell (2014), for instance, argue that services can also participate as a dynamic tool for growth, creating jobs, and increasing productivity. Indeed, for late developers, services can be the best option, provided that manufacturing has become more technologically intensive and generates fewer jobs.

To promote growth, services need to fulfill specific characteristics that have similar effects to those of the manufacturing sector. Services need to be highly productive to provide un- conditional productive convergence, be compatible with the comparative advantage of the country, and have the capacity to expand, either in the domestic market or through trade.

Some authors argue that these features can only be accomplished by the manufacturing sector (Rodrik, 2013, 2015). In contrast, other scholars believe that services can generate fast economic growth (Amirapu and Subramanian, 2015; Ghani and O'Connell, 2014). In particular, Ghani and O'Connell (2014) indicate that, nowadays, services have become more productive thanks to technology. For instance, services are not limited by transportation costs so they can expand to different markets. This also reduces the environmental footprint and generates opportunities for a more inclusive type of development.

An important point from this literature is that services could appear as a viable alternative to propel growth in developing countries where technology can stimulate, and not hurt, employment creation. This will happen in areas that do not have high automation rates and that require more specialized labor.

Nevertheless, not every service will be able to use technology as an advantage to foster growth. Some sectors are, by their configuration, not very productive. For instance, certain commercial activities are not very productive, and, according to Amirapu and Subramanian (2015), neither are public administration activities. Latin America is a region characterized by the presence of small firms, which tend to have lower levels of productivity. According to the Inter-American Development Bank, in Mexico 97% of retail establishments and 94% in the services sector employ fewer than ten employees, limiting the productive gains of this sector and leading to labor misallocation (IADB, 2010).

Additionally, in many developing countries, and notably in Latin America, the informal sector is substantial and often concentrated in services activities. Informality can prevent services for becoming more productive, and thus, they cannot be used as a tool to promote fast economic growth.

2.1 Deindustrialization and Income Inequality

Even if the services sector could provide the advantages of the manufacturing sector and promote economic growth, a deindustrialization process combined with a productivity di- vergence of the services sector could have important social consequences and exacerbate the social disparities in the region.

First, a shrinkage of the manufacturing sector reduces the bargaining power of a large share of workers in industries that are more exposed to international competition and automation. Second, as the services sector becomes the motor for economic development, there will be an increasing divergence among the different types of services provided. As mentioned before, Latin America has a large informal sector that is concentrated in service activities that show low levels of productivity and that remain small, since they are credit- constrained and because of the lack of incentives for investing in new, better capital. This is also the case for non-informal firms, remain smaller because of the lack of incentives and access to gain exposure and grow. At the same time, Latin American countries also host a large share of highly productive firms and workers in the services sector. Therefore, even if the service sector has the potential to become a motor for economic growth, there is a risk that this type of growth widens the income gap even more, in the region that presents the largest income inequality in the world.

Although research on this topic is somewhat recent, for the United States, Kollmeyer (2018) states that the interaction of trade union decline in a context of deindustrialization exacerbates distributional effects. This finding is supported by Doussard et al. (2009) who show that, for the case of Chicago, deindustrialization has exacerbated the pattern of wage inequality.

Baymul and Sen (2020) look into the effects of deindustrialization on income inequality for a range of low-, middle- and high-income countries between 1960 and 2012. The authors find that an increase in the share of workers in the services sector increases income inequality, whereas an increase in the share of workers in the agricultural or manufacturing sector leads to a decrease in income inequality, measured by the Gini coefficient.

3. THE LATIN AMERICAN CASE

It has been long acknowledged that productivity in Latin America is stagnant (Cavalcanti Ferreira et al., 2013; Latina et al., 2019). Furthermore, recent research has shown that, although Latin American productivity was closer to that of the United States until the mid-seventies, there has been an acute decrease in productivity since the 1980s (Latina et al., 2019).This section looks into the evolution of productivity across different Latin American countries, to verify if there is any correlation between this evolution and the configuration of the manufacturing sector.

3.1 Productivity

Using a Cobb-Douglas production function of the form:

$$Y_{it} = A_{it}K_{it}^{\alpha}(H_{it}L_{it})^{1-\alpha} \tag{1}$$

where Y represents the gross domestic product of country *i* during year *t*, K stands for the capital stock, H is human capital and L is the aggregate labor. α represents the capital share of income. *A* represents the productivity of a country. This is, the increase in production that cannot be accounted from an increase in inputs.

Using the Solow model's residual to proxy for productivity, we can calculate TFP and the TFP growth rate, as follows:

$$ln(A_{it}) = ln(Y_{it}) - \alpha ln(K_{it}) - (1-\alpha)ln(L_{it}) - (1-\alpha)ln(H_{it}) \tag{2}$$

$$\Delta A_{it} = ln(A_{it}) - ln(A_{it-1}) \tag{3}$$

Graph 1 illustrates the evolution of TFP in a set of Latin American countries. The graph on the left measures the relative productivity of country *i* with respect that of the United States (*k*) computed by:

$$(A_{it}A_{kt})^{1/2} = \frac{\frac{GDP_{it}}{GDP_{kt}}}{Q_{ikt}} \tag{4}$$

where Q_{ijt} represents the To¨ rnqvist quantity index of factor inputs (capital and labor) at a given point in time and is calculated as:

$$Q_{ikt} = \frac{1}{2}(\alpha_i + \alpha_k)ln\frac{K_i}{K_k} + [1 - \frac{1}{2}(\alpha_i + \alpha_k)]ln\frac{L_i}{L_k} + [1 - \frac{1}{2}(\alpha_i + \alpha_k)]ln\frac{H_i}{H_k} \tag{5}$$

The graph on the right uses data from the Penn World Tables 9.1 to calculate the growth of TFP using the Solow residual method.

Although productivity across different countries in the continent varies, in general, it is clear that there has been a decrease in productivity across most countries and that the gap in productivity between Latin American countries and the U.S. has been widening across time[5], with countries such as Mexico and Venezuela going from levels of productivity higher than those of the U.S. in the late 1970s to early 1980s to productivity that is much lower. Figure 1.2 shows the cumulative growth between 1950 and 2011 in TFP in Latin America. Other than Brazil and Colombia, most countries show a net decrease of TFP across time.

Also interesting is to see the evolution of TFP compared to that of the U.S. across time. Graph 3 illustrates the average growth of TFP relative to that of the U.S., whereas Table 1.1 shows the average TFP of Latin American countries across decades relative to U.S. and the technological progress during that period. Although Latin America's productivity has been consistently lower than that of the U.S., the gap has been widening across time for the whole region Araujo et al. (2014), Bastos and Nasir (2004), and Kim and Loayza (2019).

5 The U.S. is considered as the technological frontier.

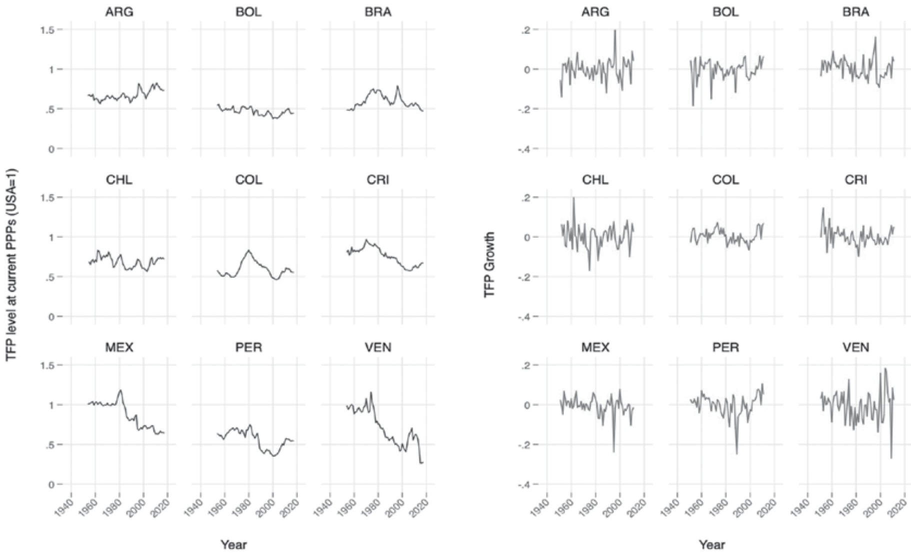

Figure 1.1: Productivity Latin America, 1950–2017

Source: Feenstra et al., 2015

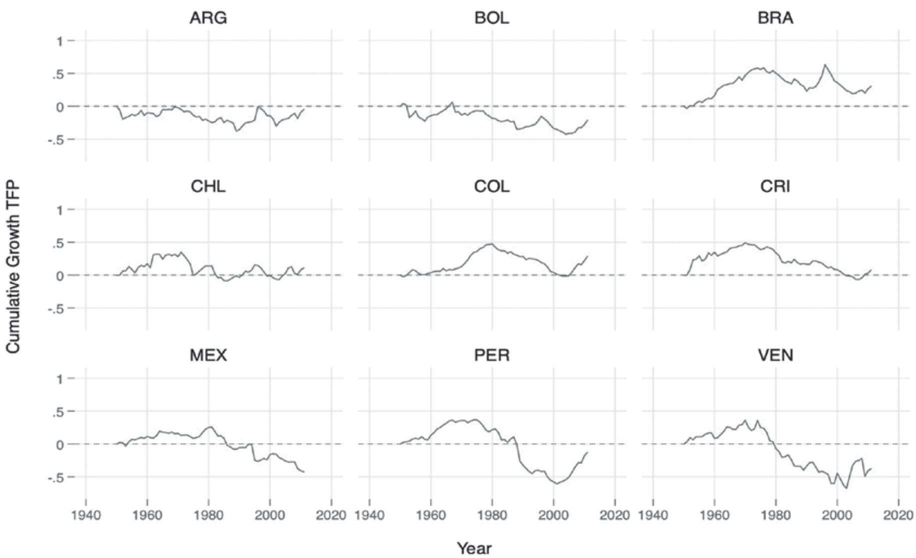

Figure 1.2: Cumulative TFP Growth, 1950–2017

Source: Feenstra et al., 2015

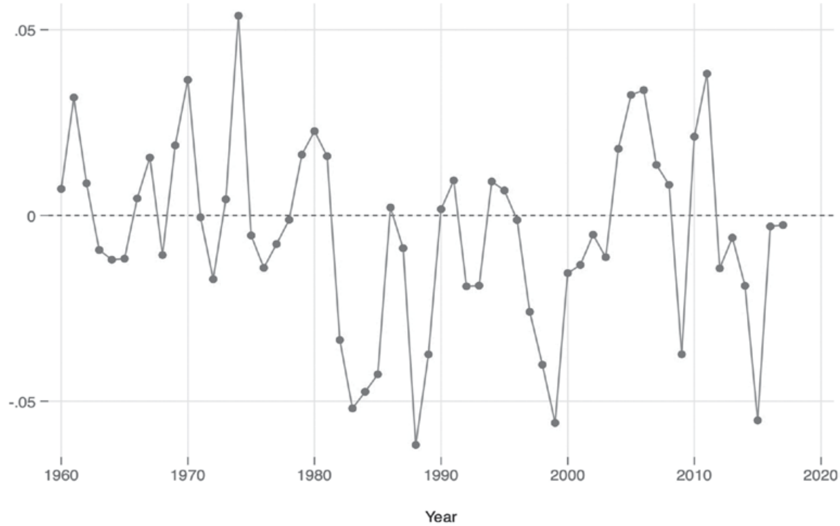

Figure 1.3: Productivity Growth in Latin America, 1960–2017
Source: Feenstra et al., 2015

Table 1.1: TFP Evolution, 1950–2011

Decade	Proximity (US=1)	Average Growth TFP
1950–59	0.699	0.007
1960–69	0.708	0.015
1970–79	0.751	-0.004
1980–89	0.690	-0.026
1990–99	0.599	-0.006
2000–11	0.574	0.008
Average	0.665	-0.001

3.2 Deindustrialization

In Latin America, the evolution of the industrial sector has been somewhat inconsistent. As Bénétrix et al. (2015) show, from 1870 until the end of World War II, the region experienced a movement of industrial convergence with higher levels of industrial growth than the three major economies at that time: the U.S., Germany and the U.K. During the period in which Import Substitution Industrialization was implemented, even if industrial growth did not surpass that of more developed economies, it recorded growth rates of 5.28% in the largest economies in Latin America[6] during the 1952–1972 period (Bénétrix et al., 2015).

After this, the oil boom and the debt crisis during the 1980s caused a structural transformation in most economies. The measures imposed by the Washington Consensus included the privatization of multiple sectors, and the openness of the goods and financial sectors to trade and private investment, leading to a decrease in the industrial activity in the region.

This pattern is visible in Graph 4, which fits a quadratic equation to show that an inverted U-shape relationship exists between manufacturing productivity and income (GDP) in Latin America, signaling a deindustrialization process in the region. This process, however, has not been consistent across decades, as manufacturing productivity has decreased at different levels of development (see Figure 1.5).

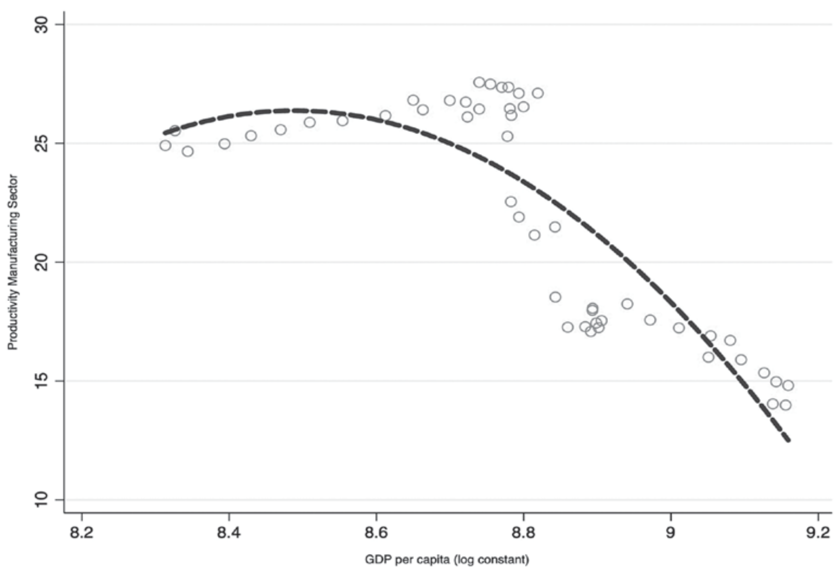

Figure 1.4: Manufacturing Productivity in Latin America, 1960–2016
Source: The World Bank, 2020

6 Argentina, Brazil, Chile, Mexico, and Uruguay.

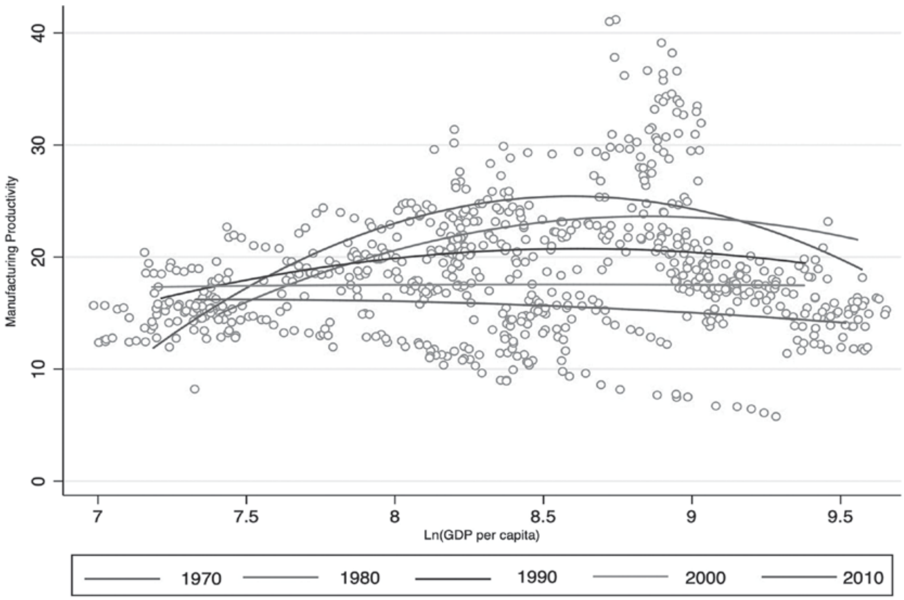

Figure 1.5: Manufacturing Productivity in Latin America across Decades
Source: The World Bank, 2020

As mentioned by Rodrik (2013, 2015) and Be´ne´trix et al. (2015), trade openness and the globalization process accelerated deindustrialization in Latin America. This process took place before the countries in the region generated a comparative advantage in most manufacturing activities[7]. Table 1.2 shows a correlation analysis of the different sectors and trade.

Interestingly, whereas productivity in the manufacturing and agricultural sectors is negatively correlated with trade, the services sector and trade are positively correlated.

Table 1.2: Trade and Value-Added in Economic Sectors

Variables	Trade (Share of GDP)
Agriculture (VA Share of GDP)	-0.937***
Industry (VA Share of GDP)	-0.544***
Services (VA Share of GDP)	0.838***

* p<0.05, ** p<0.01, *** p<0.00.

7 The automotive and electronics sectors are exceptions to this, with Brazil and Mexico having strong industries.

3.3 Dutch Disease

Globalization seems to be one of the driving mechanisms behind deindustrialization. Many countries in Latin America rely on the production and export of commodities. Since it is safe to assume that, for the most part, these countries are price takers in the international market, an increase in the exploitation of these resources driven by an increase in international prices may lead to the appreciation of the local currency and to a Dutch disease that reduces the competitiveness of the industrial sector, which in turn leads to a deindustrialization process.

Graph 6 shows the revealed comparative advantage of a subset of Latin American countries for three sectors: agricultural commodities, ores and metals, and manufacturing products. The revealed comparative advantage is calculated as:

$$RCA_{ipt} = \frac{\frac{X_{ipt}}{\sum_{j \epsilon P} X_{ijt}}}{\frac{X_{wpt}}{\sum_{j \epsilon P} X_{wjt}}} \tag{6}$$

where X_{ipt} is the country i's exports of product *p* at time *t*. (*w* denotes the world's exports of product *p*), and $\sum j \varepsilon P$ represents country i's total exports (of all products j in P).

From the graph, it seems that, except for Mexico, Latin American countries have a revealed comparative advantage in exporting agricultural commodities and ores and metals. In the case of Mexico, Graph 7 shows, with more detail, that the country has a comparative advantage in machinery and equipment, as well as in the production of manufactured goods. In recent years, Costa Rica has also increased its comparative advantage in the production of medical equipment. Brazil has a comparative advantage in the production of certain manufactured goods (UNCTAD, 2020).

This shows that, with a few exceptions, Latin American countries have not developed a strong manufacturing sector and instead rely on the exploitation of commodities that is linked to a Dutch disease.

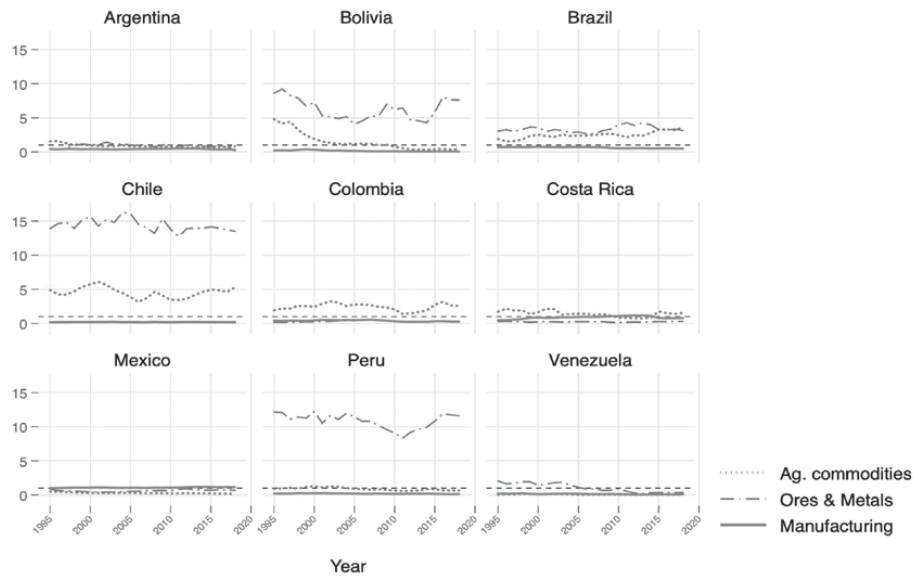

Figure 1.6: Revealed Comparative Advantage in Latin America
Source: UNCTAD, 2020

3.4 The Role of the Services Sector

The economic structure of Latin American countries has shifted during the period examined in the study (see Figure 1.8). Whereas the share of the labor force absorbed by the agricultural sector has decreased, there has been no transfer towards the manufacturing sector, which absorbs less than 20% of the labor force in the region. In the meantime, the services sector has become more relevant for these economies, absorbing large shares of the formal labor force.

Although the services sector could become a motor for economic growth, this depends on the activities that compose the sector, the structure of the local market, and its tradeable characteristics. Hausmann et al. (2007) argue that the sectors that are tradable provide most of the productivity benefits. Although data is not available for many years, trade in services has seen a rapid increase, even if, on average, the region still is a net importer of services (see Figure 1.9).

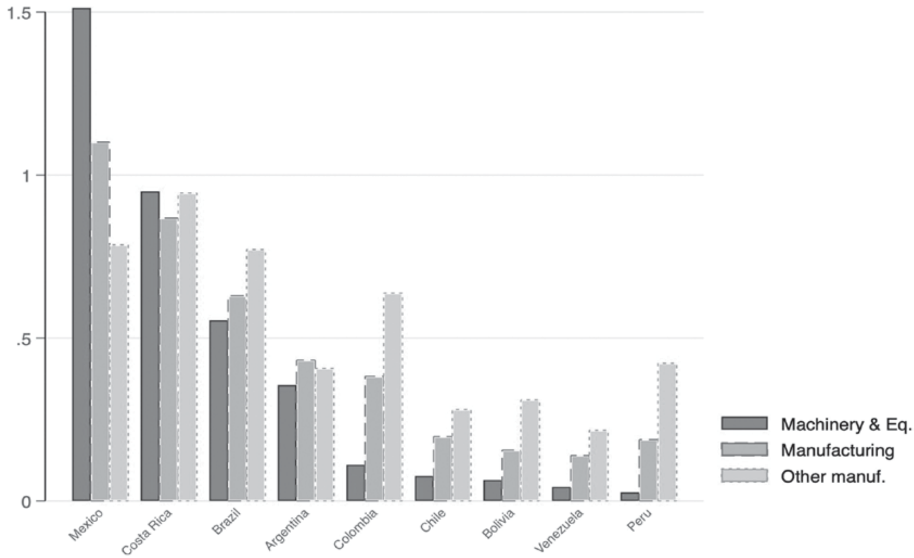

Figure 1.7: Manufacturing Comparative Advantage in Latin America
Source: UNCTAD, 2020

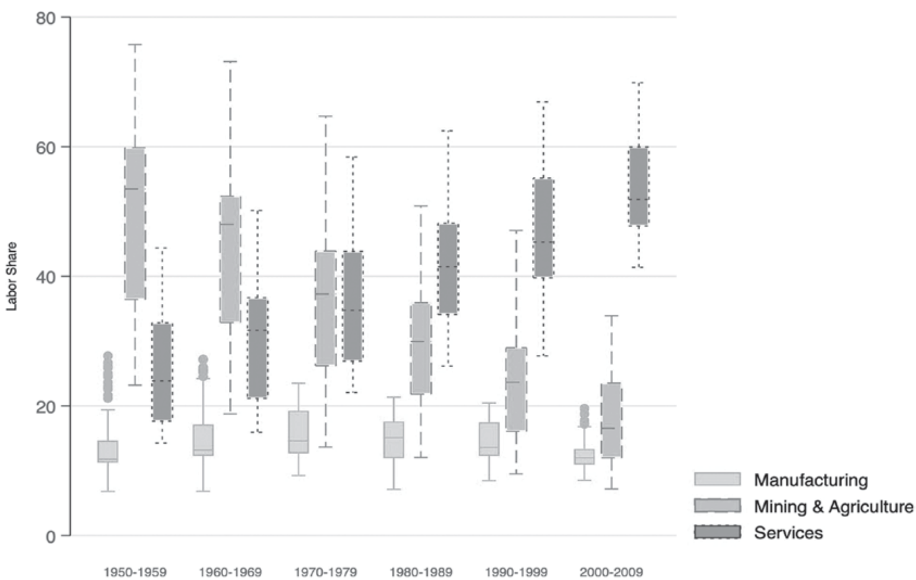

Figure 1.8: Employment by Sector in Latin America
Source: Timmer et al., 2015

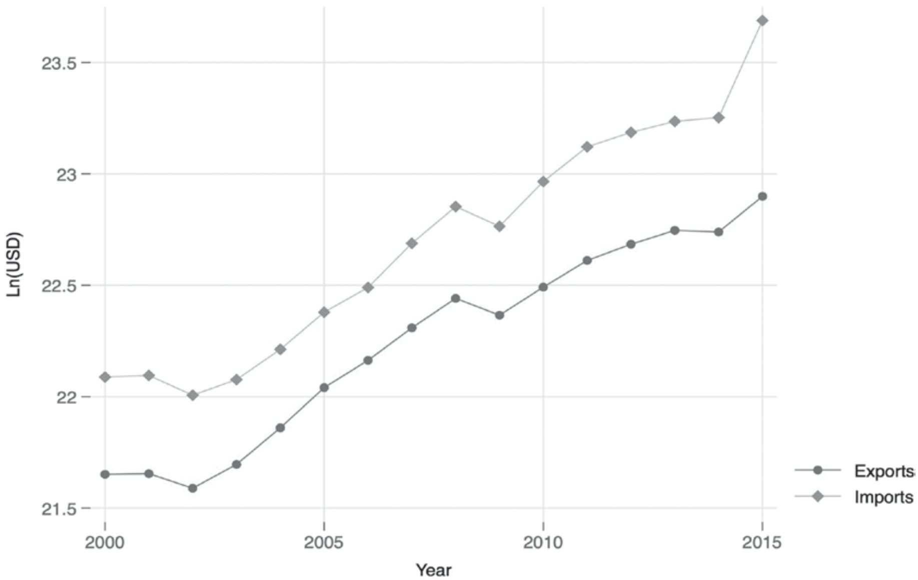

Figure 1.9: Average Trade in Services in Latin America, 2000–2015
Source: WTO, 2020

Supporting economic development through the services sector could have important distributional effects, in a region where inequality is already high (see Figure 1.10). The services sector in Latin America is composed by activities that are very dissimilar in terms of productivity. On one side, the region has a large informal sector and small businesses that focus on the provision of services that have low levels of productivity, and limited innovation. On the other side, there are activities that are completed by highly skilled workers, that are highly productive and have the potential to promote innovation and sustained economic growth. As the manufacturing sector declines, people will have to move to a different sector.

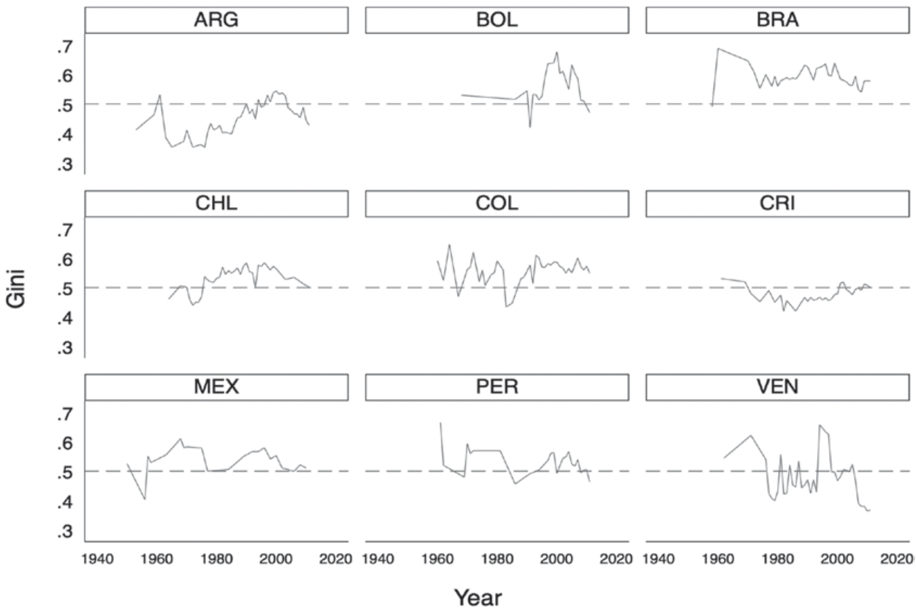

Figure 1.10: Gini Coefficients in Latin American Countries, 1950–2011
Source: WIID, 2020

As seen in Figure 1.8, most people are transferring towards the services sector. The type of specific activity will largely depend on the skills and knowledge of different workers. This will increase the income differential among the population, provided that the skills that are needed to work in a highly productive service activity are higher than those required to work in the manufacturing sector (i.e. doctor vs. worker in a car manufacturing plant). Figure 1.11 shows the labor force distribution by level of productivity. Low-productivity sectors consider agriculture, commerce and low-productivity services. Medium-productivity sectors comprehend construction, manufactures and transportation. High-productivity sectors include finance, mining, and electricity (ECLAC, 2021).

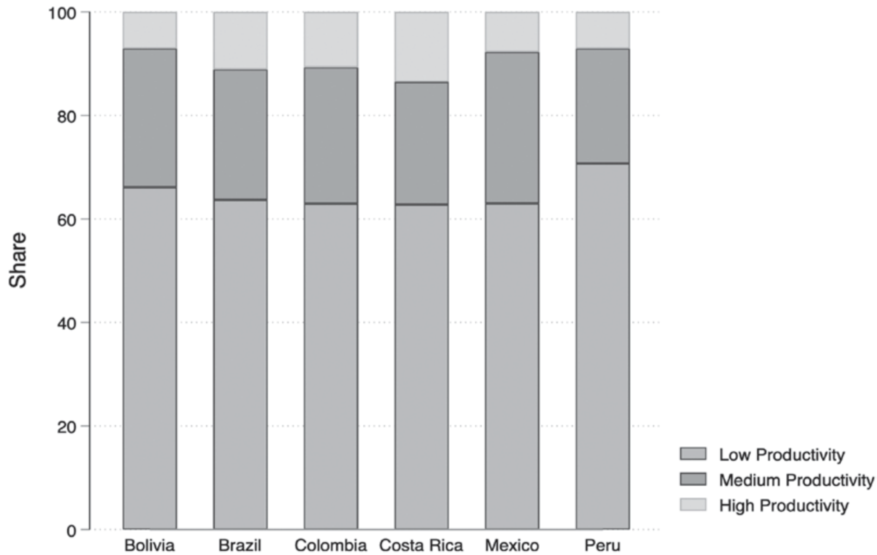

Figure 1.11: Labor Force by Productivity Level, Selected Countries in Latin America, 2018
Source: ECLAC, 2021

However, it is also important to acknowledge that the levels of income inequality in Latin America have been persistently high across the years and have been linked to the distribution of wealth across the society since the colonization period Daron Acemoglu Simon Johnson and Robinson (2002). Therefore, unless the restructuring of the economy exacerbates some of these relations, a shift towards the services sector may only contribute to the status quo in terms of the socioeconomic configuration across the continent.

The following section empirically looks into the evolution of the manufacturing sector, searching if changes in productivity are linked to it. Then, the study analyzes if globalization, measured by trade openness has impacted the size of the manufacturing sector in Latin American countries. After looking at this, the chapter examines if increases in the share of the services sector are linked to increases in income inequality, measured by the Gini coefficient.

4. EMPIRICAL STRATEGY

4.1 Deindustrialization and Productivity

To verify how deindustrialization and productivity may be related and considering that different countries have different structural configurations, it is important to look if these two series are co-integrated. To do this, the nine countries are separately examined over the period between 1954 and 2011. Two I(1) series are co-integrated if their linear combination is stationary. In that case, the two series are said to have a long-run relationship.

To verify this, the Engle-Granger procedure is followed, in which the long-run equation is estimated:

$$y_t = \alpha_o + \beta x_t + \epsilon_t \tag{7}$$

and then, $\in_t$ is tested for stationarity using the Augmented Dickey-Fuller test[8]. After verifying if the series are co-integrated, the study proceeds to look if the evolution of productivity can predict (Granger-cause) deindustrialization.

Granger causality is used in multivariate time series data and, although it does not determine causality, it verifies if, after controlling for the history of the dependent variable, the history of the variable of interest helps predict the dependent variable. Causality is not determined since there might be omitted variables that could be driving the effect, but it explores the predictive power of one time series on another one.

The procedure requires to run a restricted regression of the form:

$$y_t = \alpha_o + \sum_{r=1}^{n} \beta_r y_{t-r} + \epsilon_t \tag{8}$$

and then run an unrestricted regression:

$$y_t = \alpha_o + \sum_{r=1}^{n} \beta_r y_{t-r} + \sum_{s=1}^{k} \gamma_s x_{t-s} + \epsilon_t \tag{9}$$

Variable x is said to predict y_t if the inclusion of x_s reduces the variance of the prediction error, which is evaluated using a F-test. It is important to indicate that this analysis assumes that all the series used in the analysis are endogenous, and therefore, each variable has an equation modeling its evolution across time.

4.2 Globalization and Deindustrialization

Once the cointegration analysis shows if the evolution of productivity predicts changes in the manufacturing sector, the study proceeds proceed to examine the

8 Critical values are adjusted for the number of variables.

effect of globalization on deindustrialization, using a panel data specification to calculate the following equations:

$$manuf_{it} = \beta_1 export_{it-1} + \beta_2 import_{it-1} + \beta_3 \mathbf{X}_{it-1} + \epsilon_{it} \tag{10}$$

$$manuf_{it} = \beta_1 cupg_{it-1} + \beta_2 cupgsq_{it-1} + \beta_3 \mathbf{X}_{it} + \epsilon_{it} \tag{11}$$

To account for variance for each panel differs, feasible generalized least squares is used. The study measures deindustrialization ($manuf_{it}$) using the growth rate of labor in manufac turing as a share of total sector employment. A decrease in the share of employment would signal that, at the very least, other sectors are absorbing a higher share of the labor force. Two different sets of variables serve as explanatory values for globalization.

Equation 10 measures trade openness using exports as a share of GDP ($export_{it-1}$) and imports as a share of GDP ($import_{it-1}$) at current PPPs. Equation 11, explores how the growth of the value-added per worker of the manufacturing sector vis-a`-vis the growth of the value- added per worker of the services sector affects the share of employment in the manufacturing sector ($cupg_{it}$ and $cupgsq_{it}$). By doing this, it is possible to examine how unbalanced productivity growth between the manufacturing and services sectors could affect deindustrialization in Latin America. According to Baumol (1967), there are limits in productivity imposed by the structural characteristics of specific sectors (in this case, the services industry). In contrast, the industrial sector has more space to increase productivity. If this is the case, as the gap between the value-added per worker widens between these two sectors, the services sector could absorb a larger share of the labor force.

4.3 Income Inequality

The next step in the analysis requires to explore how the structure of the Latin American economies examined in this chapter and income inequality are linked, using the following equation:

$$\text{gini}_{it} = \beta_1 manuf_{it} + \beta_2 services_{it} + \beta_3 agric_{it} + \beta_4 mining_{it} + \beta_5 utilities_{it} + \beta_6 const_{it} + \beta_7 \mathbf{X}_{\text{it}} + \epsilon_{it} \tag{12}$$

where $gini_{it}$ represents the Gini coefficient in country i and each of the sectors denotes the percentage of the labor force that is occupied in that sector in a given country. Although for the most part, agriculture, manufacture and services employ the largest share of the population, the rest of the economic sectors are also used to consider potential changes in the economic structure. Moreover, since the objective is to identify if the changes in the structure of the economy lead to changes in income inequality in the countries examined, Equation 12 is also calculated using the growth rate of the different sectors included. The model is specified using FGLS.

4.4 Data

The study uses three main sources of data. First, to measure productivity, TFP is calculated as the Solow residual using data from the Penn World Tables 9.1 (PWT). Although the PWT has two TFP variables, the method used in this study fits better the model. Data on international trade also comes from the PWT. Employment in manufacturing as a share of total sector employment is used to measure deindustrialization. This data comes from the 10-Sector Database published by the Groningen Growth and Development Centre (Timmer et al., 2015). This source is also used to calculate the share of employment of other sectors. To measure the cumulative unbalanced productivity growth (CUPG) between the manufacturing and services industry, two steps are completed: first, a variable is created using the 10-Sector Database, including trade services, transportation services, business services, and personal services. The variable measures total employment in these sectors and obtain the share as total employment. Then, another variable is created to calculate the difference in the growth rate of employment in the manufacturing sector, minus the growth rate of employment in the services sector. With these two variables, a third variable is constructed to calculate the cumulative differential across time. This allows measuring the cumulative effects of this unbalanced growth across sectors (Kollmeyer, 2009).

The Gini coefficient information comes from the World Income Inequality Database (WIID), published by UNU-Wider (UNU-WIDER, 2020). Although the data shows wide gaps across time, it is one of the most reliable and complete data sources with data on income inequal- ity. As control variables, the study uses the natural logarithm of the real GDP per capita (at chained PPPs 2011 USD) and its squared value to proxy for economic development and the natural logarithm of the number of persons employed as a proxy for the size of the labor market. The first lagged value of all the independent variables is used[9].

9 Firms usually make decisions about their operations yearly based on the macroeconomic and firm condi tions of the previous period.

Table 1.3 includes the summary statistics of the variables used in the study. Statistics for the variables used in the cointegration analysis are shown at the country level[10]. These variables are non-stationary, but their first difference is (I(1)).

Table 1.3: Summary Statistics

N		Mean	SD	Min	Max
Cointegration Analysis					
TFP (2011=1)	557	1.19	0.36	0.63	2.49
Manuf. Labor Share	546	14.31	4.05	6.81	27.73
Globalization					
CUPG	558	0.53	0.37	-0.16	1.57
Share of exports	557	0.14	0.09	0.02	0.46
Share of imports	557	-0.12	0.07	-0.48	-0.03
Income Inequality					
Gini Coefficient	296	51.97	6.78	35.2	68.9
Services Labor Share	546	40.06	13.84	14.27	69.87
Agricul. Labor Share	546	31.51	16.04	6.48	72.56
Mining Labor Share	546	1.38	1.11	0.08	5.92
Utilities Labor Share	546	0.69	0.43	0.07	2.47
Const. Labor Share	546	6.00	1.93	1.80	12.69
Control Variables					
Ln(GDP per capita)	557	8.62	0.58	7.32	9.90
Ln(employed persons)	557	1.84	1.25	-1.24	4.59

5 RESULTS

5.1 Cointegration Analysis

Table 1.4 reports the results of the cointegration test conducted to verify if employment and productivity show a long-term relationship[11]. Examining the different tests, it seems clear that there is no cointegration between TFP and the share of employment in the manufacturing sector. To be more precise, in each

10 The level of (relative) TFP is calculated considering 2011=1 and using the growth rates of TFP to calculate the rest. Technological progress is calculated using variables at chained PPPs 2011USD.

11 The Dickey Fuller test used to test that the residuals are I(0) fits a random walk model.

test, the null hypothesis is that there is no cointe gration relationship between the series analyzed. The tests are specific for data containing many panels.

Table 1.4: Cointegration Test

Statistic	Kao C. Test	Pedroni C. Test	Westerlund C. Test
Modified DF Test	0.94		
	(0.174)		
Augmented DF Test	0.95	1.79**	
Phillips-Perron Test Modified Phillips-Perron Test Variance ratio	(0.171)	(0.047) 1.35* (0.089) 1.50* (0.068)	-0.3864
			(0.345)
Panels	9	9	9
Avg. Periods	58.56	59.56	60.56

H0: Panels are cointegrated
p-values in parentheses
$^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$

These conclusions remain valid when doing the analysis on a country-by-country basis. After looking at the number of cointegrating equations, only Costa Rica seems to have a stochastic trend, with the rest of the countries not showing a long-term equilibrium between the share of employment of the manufacturing sector and TFP. For the case of Costa Rica, based on the Wald test, it seems that the share of employment in the manufacturing sector Granger-causes TFP (see Table 1.5).

Table 1.5: Cointegration by Country

Country	Cointegrating Equations	Granger C. Wald Test (1)	Granger C. Wald Test (2)
Argentina	0	0.308	0.324
Bolivia	0	0.609	0.083
Brazil	0	0.373	0.896
Chile	0	0.219	0.225
Colombia	0	0.823	0.365
Costa Rica	1	0.485	0.005
Mexico	0	0.066	0.076
Peru	0	0.131	0.191
Venezuela	0	0.083	0.463

(1) Equation: Share employment manufacturing
(2) Equation: TFP (2011=1)
H0: VAR1 does not Granger-cause VAR2

These results align with Rodrik's argument on the effect of technological progress on deindustrialization in developing countries. The countries included in the analysis are, over- all, price takers. This is particularly true for the prices in the manufacturing industry since, as shown before, almost no country has a comparative advantage in the sector. Unlike what we can expect in high-income countries, where changes in domestic prices due to increases in productivity impact the structure of the manufacturing sector, developing economies import deindustrialization.

5.2 Deindustrialization and Globalization

To examine how an increase in trade openness is linked to deindustrialization, this section looks at the differentiated role that exports and imports may play on this process. A differential effect can be expected if the structure of imports vis-a`-vis exports reduces the possibility of domestic firms to compete internationally. This is, if the terms of trade are not favorable to the local economy, an increase in imports will further deteriorate the capacity of local firms to compete and will fasten deindustrialization. To rule out the case that deindustrialization may be the effect of a cumulative differential between the value-added per worker in the manufacturing and services sector increases, this mechanism is also examined.

Table 1.6 shows the results. Model 1 calculates Equation 10, while Model 2 focuses on unbalanced productivity growth (Equation 11). Model 3 includes all the relevant variables.

Table 1.6: Deindustrialization and the Structure of the Economy

	(1) Trade	(2) VA	(3) Trade + VA
Share Exports	-0.0017		0.0139
	(0.0305)		(0.0334)
Share Imports	-0.1252*** (0.0374)		-0.1081*** (0.0390)
CUPG		-0.0368	-0.0255
		(0.0215)	(0.0217)
$(\text{CUPG})^2$		0.0137	0.0112
		(0.0159)	(0.0158)
N	450	449	449
Wald χ^2	27.20***	22.14***	29.82***

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1.12 includes the results desegregated by decade. With exception of the 1970s decade, imports have reduced the size of the manufacturing sector in the Latin American countries analyzed. It seems that deindustrialization in the region is not linked by the differential in productivity between the manufacturing and services sectors, as it is the case in developed economies.

These results fall in line with the theory and with empirical evidence that shows that during the 1980s and 1990s, as Latin American countries globalized, the deindustrialization process became more acute. Interestingly, the increase in exports has no statistically significant effects over the manufacturing sector, which seems to be at odds with the Dutch disease hypothesis. However, it may reflect differences in the structure of the different economies of the region.

5.3 Deindustrialization and Inequality

Even if the services sector could become a motor for economic development in Latin America, it is crucial to take into account the distributional effects that this economic configuration may have. Table 1.7 contains the results of estimating Equation 12. The presence of a larger manufacturing sector in terms of employment is linked to lower levels of inequality. A 1 percentage point increase in the share of manufacturing labor is linked to a reduction of 0.4 points of the Gini coefficient. At the same time, the agricultural sector is correlated with higher levels of income inequality. In both cases, the results are statistically significant at the 1% level.

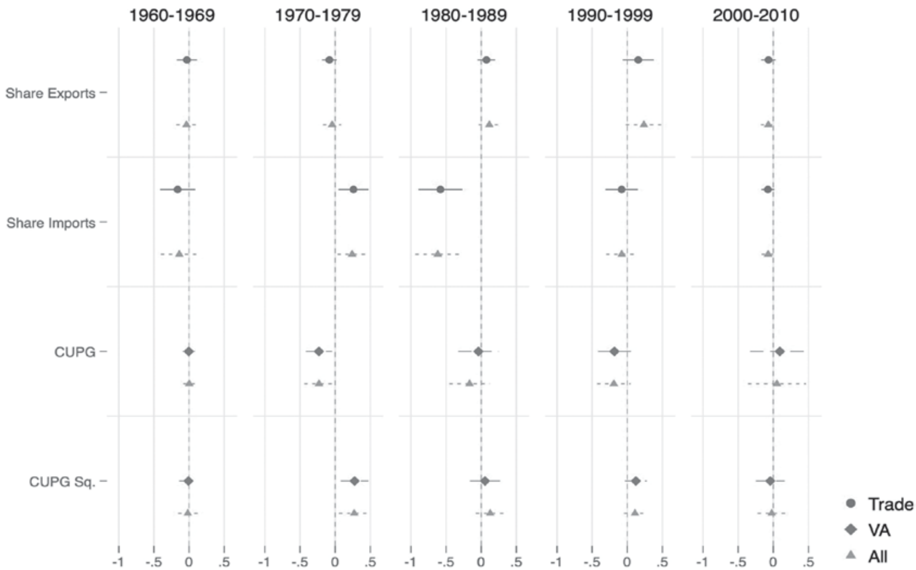

Figure 1.12: Deindustrialization and the Structure of the Economy by Decade

Interestingly, when looking at how changes in the economic structure are connected to different levels of income inequality, it seems that it has been the transition towards the services sector that seems to be related to more income disparities: places with a larger services sector present higher levels of income inequality, though the effect is moderate.

These results must be taken with a grain of salt. Data from the beginning of the series may not be entirely reliable and it presents many gaps. As time has passed, data has become more accessible and accurate. In fact, Figures 1.13 and 1.14 look at the effects by decade and the results seem less clear. Whereas there seems to be an association between an increase in the participation of the labor force in the services sector and income inequality, the relationship is not strong. Although the data on employment is constructed using labor force surveys and business surveys, in addition to census data, it may be the case that a pro- portion of the informal sector is not accounted for and this may compromise the consistency of the results. As mentioned before, the informal sector is mostly composed by smaller firms that, because of structural factors have little incentives to growth and are capital-constrained (IADB, 2010).

Table 1.7: Deindustrialization and Income Inequality

	(1) Gini	(2) Gini
Agriculture Share	0.3423*** (0.0683)	
Manuf. Share	-0.4034*** (0.1434)	
Services Share	0.1113* (0.0667)	
Gr. Agriculture		9.7849
		(10.2160)
Gr. Manuf.		-0.6963
		(7.4260)
Gr. Services		38.6615** (18.4221)
N	279	279
Wald χ^2	318.31	173.83

Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Still, it is interesting to see that there is an inverse relationship between income inequality and the manufacturing and services sector. Further research on this issue is important to determine paths of future public policy.

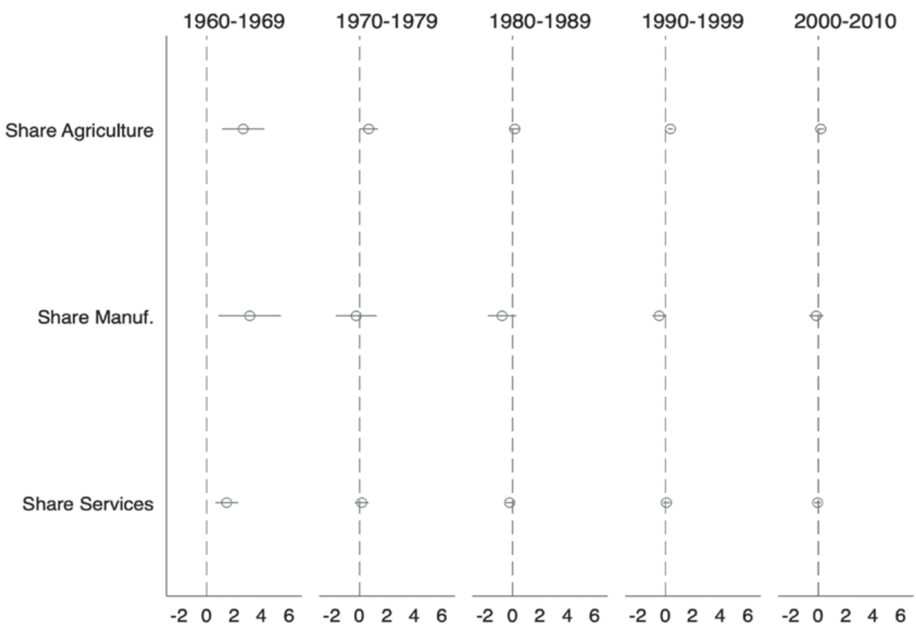

Figure 1.13: Economic Composition and Income Inequality

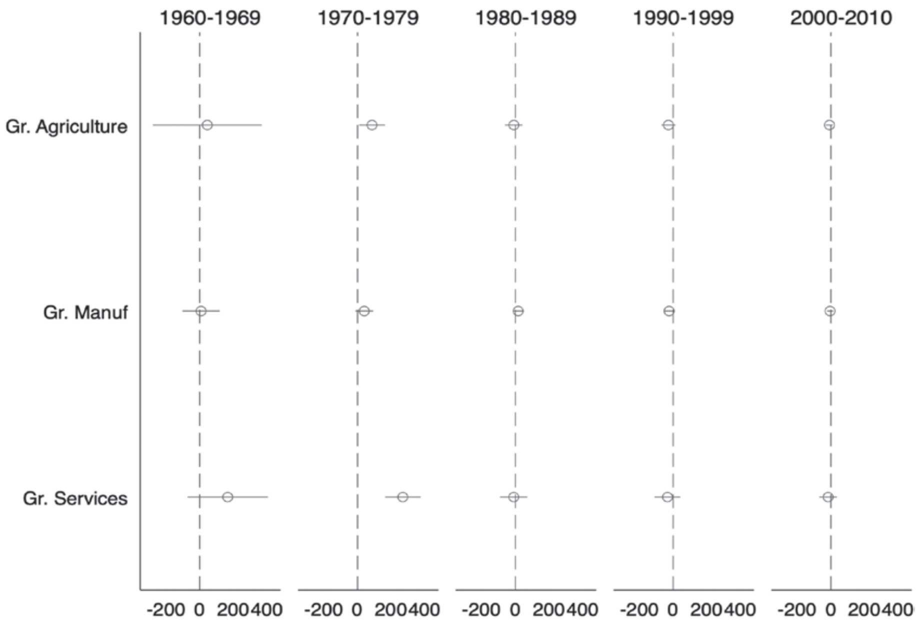

Figure 1.14: Change Economic Composition and Income Inequality

6. CONCLUSION

This chapter introduces relevant evidence of the deindustrialization process that has taken place in Latin America. Although further, country-specific, research is needed to examine specific patterns, this chapter provides suggestive evidence on the regional trends that Latin America is exhibiting. Although deindustrialization seems to be the result of multiple, unavoidable, forces, finding services that can provide the productivity gains that the manufacturing sector offers and have similar labor-absorbing capacity may provide a new avenue to support the sustainable economic growth of the region. Still, particular attention needs to be paid to the potential distributional effects that a service-oriented economic structure may generate. This is particularly true taking into account that the services sector is in a large degree composed by smaller, informal firms.

REFERENCES

Amirapu, A., and A. Subramanian. June 2015, 7. "Manufacturing or Services? An Indian Illustration of a Development Dilemma.", *Center for Global Development*, Working Paper 409, *SSRN Electronic Journal*.

Araujo, J. A., D. G. Feitosa, and A. B. da Silva. 2014, 12. "Latin America: Total Factor Productivity and its Components." *Cepal Review* 2014 (114): 51–66.

Arrow, K. 1962. "Economic Welfare and the Allocation of Resources for Invention." *Review of Economic Studies* 29 (3): 609–626.

Bastos, F., and J. Nasir. 2004. "Productivity and the Investment Climate: What Matters Most?" *World Bank Policy Research Working Paper* (3335).

Baumol, W. J. 1967. "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Cri- sis." *The American Economic Review* 57 (3): 415–426.

Baymul, C., and K. Sen. 2020, 9. "Was Kuznets Right? New Evidence on the Relationship between Structural Transformation and Inequality." *Journal of Development Studies* 56 (9): 1643–1662.

Be´ne´trix, A. S., K. H. O'Rourke, and J. G. Williamson. 2015. "The Spread of Manufacturing to the Poor Periphery 1870–2007." *Open Economics Review* 26, 1–37.

Cavalcanti Ferreira, P., S. de Abreu Pessoˆa, and F. A. Veloso. 2013. "On the Evolution of TFP in Latin America." *Economic Inquiry* 51 (1): 16–30.

Daron Acemoglu Simon Johnson, and J. A. Robinson. 2002, 11. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 117: 1231–1294.

Dasgupta, S., and A. Singh. 2007. "Manufacturing, Services and Premature Deindustrialization in Developing Countries: A Kaldorian Analysis." In *Advancing Development*, 435–454. Palgrave Macmillan, London UK

Doussard, M., J. Peck, and N. Theodore. 2009. "After Deindustrialization: Uneven Growth and Economic Inequality in "Postindustrial" Chicago." *Economic Geography* 85 (2): 183–207.

ECLAC. 2021. CEPALSTAT. https://estadisticas.cepal.org/cepalstat/portada.html?idioma=english

Feenstra, R. C., R. Inklaar, and M. P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105 (10): 3150–3182.

Froibel, F., J. Heinrichs, and O. Kreye. 1980. *The New International Division of Labor: Structural Unemployment in Industrialized Countries and Industrialization in Developing Countries*. Cambridge: Cambridge University Press.

Ghani, E., and S. D. O'Connell. 2014, 7. *Can Service Be a Growth Escalator in Low-Income Countries?* The World Bank, Washington D.C.

Hamid, N., and M. Khan. 2015. "Pakistan: A Case of Premature Deindustrialization." *Lahore Journal of Economics* 20 (Special Edition): 107–141.

Hausmann, R., J. Hwang, and D. Rodrik. 2007. "What you Export Matters." *Journal of Eco- nomic Growth* 12 (1): 1–25.

IADB. 2010. *The Age of Productivity: Transforming Economies from the Bottom Up*. New York: Palgrave Macmillan US.

Kaldor, N. 1966. *Causes of the Slow Rate of Economic Growth of the United Kingdom*. Cambridge: Cambridge University Press.

Kaldor, N. 1967. *Strategic Factors in Economic Development*. Ithaca: New York State School of Industrial and Labor Relations, Cornell University.

Kim, Y. E., and N. V. Loayza. 2019. "Productivity Growth: Patterns and Determinants across the World." *Economıa* 42 (84), 36–93.

Kollmeyer, C. 2009. "Explaining Deindustrialization: How Affluence, Productivity Growth, and Globalization Diminish Manufacturing Employment." *American Journal of Sociology* 114 (6): 1644–1674.

Kollmeyer, C. 2018. "Trade Union Decline, Deindustrialization, and Rising Income Inequality in the United States, 1947 to 2015." *Research in Social Stratification and Mobility* 57: 1.

Krugman, P. 1996. "Domestic Distortions and the Deindustrialization Hypothesis." *NBER Working Papers Series 5473*.

Latina, A., J. Manuel, S. Xirinachs, J. Chacaltana, and J. Chacaltana. 2019. Formalization Policies in Latin America . (April), Conference Presentation, ILO.

Lewis, A. 1954. "Economic Development with Unlimited Supplies of Labour." *Manchester School* 22 (2), 139-191.

Ramaswamy, R., and R. Rowthorn. 1997. "Deindustrialization: Causes and Implications." *IMF Working Papers* 97 (42): 1.

Rodrik, D. 2013. "Unconditional Convergence in Manufacturing." *Quarterly Journal of Economics* 128 (1): 165–204.

Rodrik, D. 2015. "Premature Deindustrialization." *Journal of Economic Growth* 21 (1): 1–33.

Rostow, W. 1959. "The Stages of Economic Growth." *The Economic History Review* 12 (1), 1-16.

Solow, R. 1956. "A Contribution to the Theory of Economic Growth." *The Quarterly Journal of Economics* 70 (1): 65–94.

The World Bank. 2020. *World Development Indicators*. Technical report, The World Bank.

Timmer, M., G. d. Vries, and K. d. Vries. 2015. Patterns of Structural Change in Developing Countries. In *Routledge Handbook of Industry and Development*, edited by M. Tribe and J. Weiss, 65–83. Routledge, London.

UNCTAD. 2020. Merchandise Trade Matrix.

UNU-WIDER. 2020. World Income Inequality Database.

Young, A. A. 1928. "Increasing Returns and Economic Progress." *The Economic Journal* 38 (152): 527–542.

CHAPTER TWO

Measuring Discrimination in Peru's Labor Market

FERIDOON KOOHI-KAMALI*

1. INTRODUCTION

Separating gender and racial discrimination from human capital and other effects on earnings is critical for designing effective employment policy. Empirical evidence suggests gender and ethnicity are major influences on wage differentials in Latin America. Quantification of ethnic and gender effects on wage poses a estimation bias problem in the presence of sources of discrimination unobservable to the research such as quality of education and productivity. Such unobservable influences affect both the entry of individuals into the labor market and their wage rates once in employment The residuals in the model for market entry and that for wages would then be correlated, and that correlation, as we argue below, leads to estimation bias that distort the measured impact of ethnicity and gender on pay differentials.

In this paper we examine the residual correlation problem and employ a two-step model of selectivity designed to deal with unobservable sources of bias in the context of estimating an earnings function for Peru, a country with a large informal economy and a history of discrimination by gender and ethnicity. We first estimate a base-line Mincerian log of weekly wage specification for a wage function with education, experience, race and gender as key explanatory variables. In

* fkoohikf@newschool.edu.

addition, we test this model for the number of hours worked for possibly endogeneity. We attempt to account for the effects of unobservable factors on earnings that the literature tends to identify with discrimination in the labor market and discrimination out of the labor market (exclusion), by estimating a measure of sample selectivity bias and employ that as control variable for unobservable elements in the wage equation. We apply this procedure to the World Bank Peru LSMS surveys of 1985 and 1994. Our results suggest problems with the base-line OLS estimates of gender and ethnic effects on wages. After correction for unobservable influences in term of a control variable obtained from the two-step model and included as an additional variable to the wage equation, we obtain estimates that suggest strong negative effects of gender and ethnicity on wage rate in Peru's labor market. Furthermore, we also present some evidence on the broader theme of this volume on how the evidence presented on labor market outcomes relate to individual productivity, sectoral differences, and rural-urban changes over time.

Section 1 examines the literature on gender and ethnic discrimination in Peru's labor market and the key influences that drive such discrimination, esp. the potential pitfalls of disregarding the impact of the unobservable. Section 2 outlines the two-step econometric model of wage determination with correction for bias due to unobservable factors. Section 3 discusses the data employed and some of its descriptive features. Section 4 examines the results obtained and a final Section 5 highlights the relevance of the approach adopted in this paper for econometric modeling of discrimination and briefly sums up our findings and its implications for public policy.

2. LITERATURE

Research on Peru's labor market has highlighted race and gender as two principal influences on earning disparities. Some of the research employs the popular Blinder-Oaxaca decomposition method and its extensions to measure race-gander effects counterfactually if individuals possessed the same characteristics of their opposites. For example, MacIssac and Patrinos (1995) applies this approach to the 1991 LSMS of Peru and finds that between 50 and 70% of the ethnic pay differentials are due to observable differences in productivity, while the rest, including the influences of unobservable factors, remain unexplained. Argumedo and Pimbert (2010) point to the growth of the informal and barter economy in Peru in the 1990s that disproportionately accounted for the economic activities of the indigenous population; the study suggests exclusion from the formal labor market is a principal form of discrimination. Field (2007) finds that the government legalization of property titles to 1.2 million urban squatters between 1996 and 2003 resulted in a substantial increase in hours of labor supply and a shift away from work at home to outside work; suggesting the absence of property rights as an important determinant of extensive and intensive labor market participation by a segment of Peru's population that are primarily indigenous. Differences in human capital has also received attention in research on Peru's wage discrimination. Figueroa (2010) develops an approach to Peru's racial wage inequality that attempts to explain the relative lack of correlation between improving the inter-ethnic education gap and the inter-ethnic pay gap. The paper notes that differences in pay are determined by access to high wage sectors of the economy, and by differences in human capital, not years of education, given that quality of public/private schooling is sharply different in Peru. He argues that inter-ethnic wage disparity is the outcome of both discrimination and *exclusion*. The prediction of this approach, that the rise in the education of the Peru's indigenous population will have limited impact on closing its relative racial wage gap, is consistent with the empirical evidence presented in the study.

A major issue to consider is how to control for the effect of those working age individuals who are not observed as participants in the labor market, on the wage function parameter estimates. Nopo (2009) has argued that analyzing Peru's highly gender-segmented labor market requires extending the standard Blinder-Oaxaca decomposition-based comparison of the average female-male wage differential to one based on the gender differences in the distribution of wages. The study is based on matching identical individual males and females with the same observable characteristics, and then obtaining the expected value of the earning differences, or the wage gap. The wage gap is then decomposed into explained observable and unexplained residual components. The paper breaks down the gender wage gap into those components and reports 28% of the average gender pay differences are due to a combination of the unobservable/random components. Moreover, the study maintains that unobservable/random factors relate to gender differences in access to the labor market rather than to wage differences; those effects play an even larger role in labor market participation rates. Finally, the study finds that while during 1990's, Peruvian females achieved higher years of education, that did not result in more jobs or better pay. The implicit assumption in this approach is that the mean of the characteristics across the two components are linearly independent of each other; hence can be added to obtain the mean wage gap. That rules out the impact of the unobservable on the observable variables. Without that assumption, the decomposition of the wage gender gap into explained and unexplained components would be incorrectly estimated. Some of these themes in Peru's ethnic wage disparities are also explored. Nopo et al. (2004) focus is on correct identification of ethnic status in a model of ethnic wage gap. They note that the Peru's population are neither a predominantly indigenous nor white but rather a mixture of ethnicities different in the degrees of mixture, and that makes ethnic status hard to identify from the self-reported ethnicity data. The majority of Peruvians self-report their ethnicity as "Mestiza" that are perceived by others as very different. As the difference in ethnicity status data between self-reported and reported by pollsters trained for the research plan, turned out to be significant, the authors employed the pollster ethnicity data to estimate ethnic wage gap discrimination by the Blinder-Oaxaca decomposition. The study reports ethnically related earnings differences mostly in favor of white workers. Once again, the question of control for variables observable to the firms but unobservable to the pollsters is left unanswered.

However, there are bound to be many unobservable influences on productivity and other relevant individual covariates, Hackman (1998). Suppose we compare two types of individuals by ethnicity, or gender, who otherwise appear to have similar observable characteristics. Since the method makes the comparison at market-level entry, the estimate assumes the difference in market outcome is a measure of discrimination, ignoring productivity differences gained outside the labor market before entry. The assumption that all differences in labor market outcomes between otherwise similar applicants are due to hiring discrimination distorts the identification of the sources of differences in the labor market outcomes. That distinction is critical for the adoption of effective public policy i.e. whether policy should target reinforcing equal treatment of similar workers regardless of ethnicity or gender in the job market, or focus on improvement of schooling quality and productivity, etc. before market entry.

Yet another alternative approach to the unobservable determinants of the wage gap estimate is by audit method. The audit method controls for differences between two job applicants who are assumed to be otherwise similar in terms of education, skill and personal features, and then compare the outcomes of the applications to the same firm, and average the outcomes over all firms for the same pair of individual applicants, in order to obtain an aggregate estimate of discrimination. Moreno et. al. (2012) applies this method to Peru; Nopo (2012) more extensively to Latin America by both ethnicity and gender. Hackman (1998) discusses a number of implicit audit method assumptions are required to interpret such an estimate as a consistent measure of discrimination. Suppose the expected labor market outcomes based on the observable characteristics are identical between the audit pair. Then even if unobservable covariates are independent from the observable ones, the mean productivity will not be equal if individuals have different mean unobserved characteristics; an unbiased estimate would have to assume equality of such unobserved characteristics. Hence, it is possible for the audit approach to conclude no discrimination based on the observables where there may be discrimination; the likely dependence of the two sets of observable and unobservable productivity would reinforce estimation bias due to unobservable differences. Furthermore, the method assumes a linear model in characteristics; non-linearity due to different thresholds adopted by firms, typically for the level of applicants qualifications, gives rise to another problem to detecting discrimination from the data by audit method. Suppose both ethnic groups have the same mean characteristics and same productivity but one ethnic group has more heterogeneity in unobserved productivity characteristics required by employers. Then even if the more heterogenous group contained more highly qualified individuals, with a longer right-hand distributional tail of unobserved characteristics, the firms will hire more from the less heterogenous group if the unobserved threshold is set relatively low. That is, despite equal observed and unobserved mean productivity, difference in unobserved variances can erroneously lead to a conclusion of no discrimination where the two groups may have unequal probability of selection by employers. The opposite will accrue if the threshold is set too high.

The unobservable group of individual characteristics includes not only those gained outside the labor market by choice, but also of those characteristics that lead to exclusion from having access to that market. One solution to this problem is to employ a sample that includes observations on all the labor market active population, namely, both those with reported positive monetary wage pay and those with zero wage observations, and estimate a wage function by OLS, as in O'Neil and O'Neil (2005). This would not be a satisfactory solution, especially for Peru, since there is evidence of substantial labor market participation without monetary payment, (see below). This type of employment results in a large percentage of the sample observations on wages concentrated at zero. Treating all such observations as though they were genuinely zero would be to ignore that the sample would not be random, and randomness (of the error term) is a requirement of valid OLS estimates. A better alternative is to specify two separate equations for market entry and for earnings conditional on entry. Barron (2006) examines two issues for an econometric estimation of ethnic pay differences based on such a two-equation approach. The share of unpaid labor in Peru is substantial; many receive payment in consumption goods only. For these workers, the distribution of monetary pay is truncated at zero. Hence the OLS estimates of the wage equations would be biased, particularly with respect to the estimates for the impact of ethnicity on earnings in Peru[1]. He employs a two-stage model with the first stage equation based on a dummy dependent variable for participation, and the second stage a log wage equation conditional on positive participation. This approach is known as the *hurdle* model since an observation has to pass over the zero threshold to be observed as positive. It should be noted that the estimation employs for both equations the *same* set of explanatory exogenous variables. The main finding highlighted in this paper is that exclusion from lack of access to the formal

1 The unpaid workers are mostly indigenous laborers; according to Barron 23% of the unpaid are indigenous compared to only 8% of non-indigenous.

labor market is more important than discrimination by wage differential within the labor market. Despite some interesting results, there is, however, a problem with the estimated standard errors obtained from the application of the two-stage model when it is based on the same set of explanatory variables in both market entry and wage equations. We shall address these in the next section.

The key points that emerge from this literature are: that Peru's wage gap has significant gender and ethnic components, and these cannot be analyzed without taking into account both labor participation and the amount of labor supplied conditional on participation. The pay gaps are affected by important unobserved factors that must be taken into account; pay differences are affected by lack of access to the formal labor market, not solely by discrimination faced in that market. The econometric model outlined in the next section will address these issues.

3. ECONOMETRIC MODELS

A challenging issue in estimation of an earnings function for an economy like Peru is that the wage rate as dependent variable may be systematically affected by omitted factors that influence both the decision to enter the labor market and the number of hours worked. For example, the quantity of human capital measured by years of education and experience is often observable from data as explanatory variables in a wage equation; while other factors, such as ability, or quality of human capital typically remain unobservable to the researcher. Such unobservable factors would then become a component of the equation's residual, violating the OLS assumption of a mean zero error term. More generally, whenever a disproportionately large number of zero observations of the dependent variable are not quite consistent with a base-line model using only the positive values of that variable, then randomness of the sample must be tested as the OLS estimates will suffer from *selectivity bias*, a particularly important problem in the context of a labor market analysis. The Heckman selectivity method is designed to correct for this type of bias by specifying a model based on a wage equation and a separate labor market participation equation, (see appendix for more details on the issues discussed below).

$$D_h = x_1\beta_1 + \mu_1 \text{ if } D_h = 1;\ D_h = 0 \text{ otherwise}$$

$$\log(\text{wage})_h = x_2\beta_2 + \mu_2$$

The assumptions are that the expectation of μ_2 conditional on x_2 is zero: $(\mu_2|x_2)=0$; all x_1 are exogenous and μ_1 is normally distributed, independent of x_2: $E(\mu_2|x_2, x_1)=0$. Taking expectation of *log(wage)* shows that the correlation between μ_2 and μ_1 is the cause of the estimation bias from sample selectivity. Because (μ_1, μ_2) is independent of x_1, we write this as

$$E[log(wage)|\ x_1, \mu_1]=x_2\beta_2+E(\mu_2|x_1,\mu_1)$$

That is, the expected value of *log(wage)* is equal to $x_2\beta_2$ plus an additional term $\lambda= E(\mu_2|x_1,\mu_1)$ that is a function of $x_1\beta_1$ from participation equation. λ can be estimated by the ratio of *pdf* to *cdf* from a probit equation using *all* sample observations on working age individuals (those employed in the labor market and those actively searching work), given $\mathbf{y}.\gamma$ from estimation of the first equation, and is known as the *inverse Mills ratio* (*imr*). Heckman (1976) argued adding *imr* as an additional variable to the wage equation provides a test of whether the coefficient values of the wage equation suffer from sample selectivity bias: a statistically significant *t*-ratio for λ rejects the null hypothesis of no selectivity bias. Heckman also demonstrated that the presence of *imr* among the explanatory variables of the wage equation controls for omitted unobservable factors to ensure unbiased and consistent estimates. The application of the Heckman sample selectivity model involves two steps, and provides a test of selectivity bias and a method of correcting for it. First, we estimate the first equation; then we use *imr,* as a variable added to the right hand of the wage equation along with all other explanatory variables.

$$\log(\text{wage})= x_2\beta_2 +\delta.\lambda+\nu$$

x_1 and x_2 usually have many variables in common, but for this method to work, it is critical that x_2 is a *subset* of x_1, hence x_1 must contain additional explanatory variables that are unique to the participation equation.

Two important consequences follow if x_1=x_2. First, inclusion of λ in the log wage equation introduces very high collinearity among the regressors of the wage equation (4), leading to large β_2 standard errors, for details see Wooldridge (2010, chapter 19). Second, given the joint normality assumption of μ_2 and μ_1, one can estimate their covariance σ^{12}. However, σ_{12} is not estimable if x_1=x_2, hence the two-step model is applied based on the assumption that σ_{12}=0, see Heckman (1990); also Cameron and Trivedi (2005) for details. The selectivity model would then reduce to the hurdle model. However, if there are unobservable factors that are components of μ_1 and μ_2, and these are correlated with y and x, as it is very likely, then one cannot reasonably assume σ_{12} =0 [2].

However, two-step equations may contain *unobservable* factors that influence both participation and wage equations. The literature examined earlier suggests

2 This is the model employed in Barron (2006) based on assuming only fully observable independent variables.

several omitted variables that would create correlation between μ_1 and μ_2. A typical example is differences in ability; more able workers are more likely to find a job and also more likely to earn more, other things being equal. More relevant unobservable factors that employers may use to discriminate against indigenous workers are region of birth, or the name identity of workers, or insecurity of property rights as suggested in the literature discussed. Moreover, although the data set allow control for quantity of education, *quality* of education remains unobservable and as evidence for return on education in Peru shows, that may be a powerful tool of labor market entry against indigenous workers: poor quality of education manifests in lower productivity and that makes it harder to gain access to the formal labor market and likely to result in lower wage rates once in the labor market[3]. Just like ability, such unobservable factors are likely to be present and common to residuals of both participation and wage equations, affecting access to the formal labor market; that is discrimination by exclusion, and discrimination by lower payment.

Thus, one would have to establish the absence of selectivity bias by conducting a selectivity test based on *imr* first without assuming $\sigma_{12}=0$, and control for it if the test suggests its presence. This is the version of the Heckman test employed in this study, with an exclusion restriction involving one or more variables appearing in the market entry selection equation but excluded from the wage equation. As a preliminary exercise, we rely mainly on *age* to be one such variable; it is usually not included in a wage equation based on years of experience due to its high collinearity affecting the standard errors of the estimates.

We adopt the above approaches for the specification of an earnings function for Peru bearing in mind the particular features of Peru's labor market highlighted in the literature. The most common approach to the measurement of labor market gender and ethnic discrimination is a estimation of a model with log wage per unit time regressed on education, experience, and experience squared, hours worked per unit time, ethnicity, and gender dummies.

$$ln(wage)=\alpha_0 +\alpha_1.exduc+\alpha_2.exper+\alpha_3.(exper)^2+\alpha_4.hoursw+\alpha_5.gender+ \alpha_6.race+\Sigma\beta.z+\varepsilon \tag{1}$$

where z is a vector of additional vector variables, e.g. regions, year fix-effects if observations also vary over time., interactive terms, etc.; (1) a classical, normally distributed random error term. Provided (1) is correctly specified and contains no endogenous variable, estimation by OLS produces unbiased and efficient estimates for measures of discrimination by gender and ethnicity.

3 Chapter One examines the macroeconomic consequences low quality schooling and low productivity on inequality, informal sector and development.

There are two important problems to face in obtaining unbiased and efficient estimates for such a wage function. The first is the potential endogenous nature of hours worked since some individuals may choose the amount of labor supply, particularly in rural areas with subsistence, own production farming. The potential endogeneity of an observable variable can be tested by a Hausman test provided at least one effective proxy is available for it, and as long as the sample employed is randomly selected. At the first stage, each endogenous variable is regressed on its proxies and all exogenous variables, the error term from this regression is then added to the earnings equation as an additional variable; a significant coefficient estimate for this variable would indicate the presence of an endogenous variable. Number of hours worked (*hoursw*) in (1) is one such variable and will be tested for endogeneity. Moreover, the probit model in this paper follows two specifications: first, specification without any exclusion restriction, and a second with *age*, *age squared* and *number of children* added in the probit equation but not the wage equation.

$$D_h = \alpha_0 + \alpha_1.hours\hat{} + \alpha_2.exduc + \alpha_3.exper + \alpha_4.(exper)^2 + \alpha_5.gender + \alpha_6.race + \Sigma\delta.x + \mu \tag{2}$$

$$D_h = \gamma_0 + \gamma_1.hours\hat{} + \gamma_2.exduc + \gamma_3.exper + \gamma_4.(exper)^2 + \gamma_5.gender + \gamma_6.race + \varsigma\delta.x + \beta_1 age + \beta_2 age^2 + \beta_3 chd\# \tag{3}$$

where $Imr = normal\ density\ of\ (\hat{D}_h\ /\ cumulative\ normal\ of\ \hat{D}_h)$

$$ln(wage) = \sigma_0 + \sigma_1.\mathrm{exduc} + \sigma_2.\mathrm{exper} + \sigma_3.(\mathrm{exper})^2 + \sigma_4.hours\hat{} + \sigma_5.imr + \sigma_6.gender + \sigma_7.\mathrm{race} + \Sigma\beta.z + \eta \tag{4}$$

The two stage selectivity test by *imr* assumes joint normal distribution of $(\mu, \eta) \sim N(0, 1)$. The vector of x variables includes year dummy fix-effects, and interaction between ethnicity, gender and other independent variables in (4).

This estimation procedure has some bearing on the issues raised in the last section. First, it controls the effect of unobservable factors other than the random error that may affect gender and ethnic discrimination in the wage equation. Second, since estimation of *imr* affects the proportion of working age persons outside the labor market, the procedure accounts for the impact of an important aspect of exclusion, namely lack of access to the formal labor market, on earnings. As the exclusion is likely to affect mainly the indigenous, and/or female labor force of Peru, this approach goes some way to address the limitations of measuring the effect of race on earnings using a language-based definition of indigeneity and its interaction with gender. *Imr* can therefor provide a compound measure for control of discriminatory and non-discriminatory unobservables, for example differences in propensity to work, though with its break down between the two components unknown.

4. DATA

Data for this study comes from the World Bank LSLM surveys of Peru available for 1985, 1990, 1991, and 1995. The 1990 survey covers only Lima, while the 1991 survey left out important parts of the country where a survey could not be safely conducted. In this study we employ the 1985 and 1994 surveys since both have national coverage. Excluding the individuals who are not in the economically active age group of 14–65 results in a sample size of 25,487 from combining the 1985 and 1994 surveys, and it is the sample employed in this study. Table 2.1 presents the descriptive data for this sample for the variables employed, while Table 2.2 shows the distribution of the sample by gender and indigenous status for the full sample of 25487 down to the level of the sample employed in the wage equation using only the observations for which wage>0, a sample size of 2,396. Note that there are 202 wage observations equal to zero that include 200 individuals whose working hours are greater than zero. We consider these 202 persons out of the labor market and treat them the same as those individuals who do not report their wages information. Note that a very low share of indigenous workforce is indicative of poor identification indigenous by language. Note also the higher share of women in the labor force but their lower share in the labor market (opposite holds for the indigenous labor). Finally, it should be mentioned that the survey defines experience by the number of years in employment in the *same* category of work.

We draw attention to a number of features of this sample that relate closely to the issues of exclusion and discrimination by gender and indigenous status in the labor market.

a. The adequate information for the identification of the indigenous population is not available in the data sets. The three common methods of identification are: by last names of job applicants, by rural district of birth, and by mother tongue. The first method, evidently commonly used to identify indigenous status, results in a significantly fewer call of job applicants back, but the information is not available from our data set. The other two both understate the size of the indigenous labor force. District of birth method does not count as indigenous the children of immigrants in urban centers; the mother langue method does not count as indigenous the generation who have lost their mother tongue (school instruction in Peru is only in Spanish). The information for district of birth is unavailable, hence we employ mother tongue to identify the indigenous workers.

On this basis, the indigenous active population was 3.4% of the 1985 sample, 18.5% of the sample in 1994, and 13% in the combined sample. This figure compares to 14.5% as the indigenous population share in Peru's labor market reported in MacIssac and Patrinos (1995, table 1). This is far below the indigenous share of the population in Peru. Other sources suggest a much larger share of indigenous persons, at least in the population, if not in the labor force. Since many among Peru's indigenous population, especially in the urban areas, receive education in Spanish, the identification of ethnicity by language is in any case inadequate. We also note that we focus on Peru's indigenous population since the share of the Afro-Caribbean population is relatively small.

b. The surveys provide data on education only in terms of six different levels of education achieved rather than number of years of schooling. However, the lowest category of illiterates constitutes only less than 10%, and yet 41.3% of individuals have missing values for education. A large percentage of the indigenous population 14–65 years of age falls into this category. In this study all such missing cases are allocated to the educationally illiterate category. Hence, the high correlation of low wage, informal sector employment with indigeneity status should be borne in mind in interpreting the indigenous coefficient estimates below; the same applies to the correlation between gender and informality, Nopo (2009) presents similar evidence between gender and informality, see Nopo (2009) presented similar evidence on the relationship between the gender-informality.

5. RESULTS

Table 2.3 presents the results obtained from implementing the above model. All estimates, except for probit equations (2) and (3)[4], have (White) heteroskedastic-corrected standard errors. The first column shows the OLS estimates for (1). The additional variables are year (fix-effects), a small (0–6 years) child dummy; a married status dummy, a labor union membership dummy, and three sector dummies, with agriculture as reference. Education level has six category dummies, with the lowest illiterate group acting as the reference.

All variables have the expected signs, Lima and the small child dummy are statistically significant at α=5%; all the rest at α=1%. In particular, being a woman and being an indigenous both have significant negative impacts on wage rate. Note that all individuals with more than the lowest education level earn more than that reference group. The indigenous labor force is mainly concentrated in

4 Variance of binary variable D_b is $p(1-p)$ where $p=-x'_1\beta_1$. If the model is correctly specified, correction for heteroscedasticity has no benefits; corrected and uncorrected standard errors remain the same, see Cameron and Trivedi (2005, p. 469)

the agricultural sector (see Table 2.2) and since this is the reference base for sectors, the evidence suggests workers in all other sections receive higher pay.

The evidence also suggests much the same for gender bias with regard to education and sectors since the shares of women is likely to be significant in the lowest (reference) educational level, and the lowest agriculture/service sectors compared to men. Column 2 repeats column 1 estimates with a Hausman endogeneity test for number of hours worked *(roh)* as discussed above, using *additional-payment-in-kind* (*payinkind*) as an instrument. The test fails to reject the null hypothesis of no endogeneity ($t_{\_roh}$=0.34). Otherwise columns 1 and 2 are very similar. In the rest of this paper, we treat *working hours* as an exogenous variable.

Next, we examine how much this evidence is affected by sample selectivity bias. Columns 3 and 4 present the results obtained from the estimation of the (2)–(4) model with control for selectivity bias but without imposing any exclusion restrictions, that is the LH variables are identical in both equations. Note that none of the education-level effects in the probit equation (2) are significant, but experience and sector effects are significant determinants of the probability of labor market entry. From this equation we obtain *imr* by (2), and add it as an additional variable to (4) to control for selectivity bias; column 4 shows the outcome.

First, note that the test suggests selectivity bias resulting from potential lack of randomness in the earning sample given that *imr* is significant at α=5%, with a *t*-ratio of -2.09. Note that the model (2)–(4) accounts for bias from potential correlation between the residuals of the two equations affected by the compound measure of the omitted (unobservable) influences, some of which are those that affect ethnic and gender discrimination, although we cannot control separately for the unobservable effects of discriminatory and non-discriminatory influences, for example, due to individual differences in ability. Since *Dh* in (2) is defined over the economically active population aged 14–65 years with zero values for wage payment, *imr* in part controls for exclusion effects that stem from lack of access to the formal labor market, and other unobservable influences related to name or region of birth, the effects of insecurity of property rights, and others of this type.

Finally, columns 5 and 6 address the impact of imposing exclusion restrictions on the two-step Heckman model employed in columns (3) and (4). As a preliminary exercise, we employ *age, age-squared* and *number of children* as additional variables in the probit equation (3), and include *imr* obtained from (3) as a variable in (4). The estimates by (3) given in column 5 show *age* and *age-squared,* but not number of children, to have significant effects; otherwise, the results between column (2) and column (4) are very similar. However, the estimated standard error on *imr* increases with the exclusion restriction imposed, and the *t*-ratio falls (from -2.09 with no restriction to -1.9) as the exclusions imposed on (4) leads to more accurate estimates with reduced multi-collinearity between *imr* and other variables in (4). However, the new *imr t*-ratio still rejects the hypothesis of no selectivity bias, suggesting that with control for unobservable factors, the estimates in columns (4) and (6) are unbiased while those in columns (1) and (2) are not. Column (6) incorporate the best estimates of this paper, although a more effective list of exclusions might secure larger reductions in collinearity, and more accurate estimates. The remaining variables in column 6 are similar in size and significance to those in column 4. The model is well determined and has the statistically significant expected signs. In particular, the final estimates point to a significant reduction in earnings if the worker in the labor market is a female or indigenous, after control for unobservable effects, including discriminatory practices.

Quantification of the marginal effects of replacing an indigenous or a female worker with a non-indigenous or a male worker must account for the interaction estimates. The interaction parameter estimates should be assessed at selected sample values for the gender and ethnic indicator dummies. In this case, the sample means would produce large distortions, given the number of zero observations, especially for indigenous status dummy, see Table 2.2. Hence, we use the first non-zero percentile, equal to 1 for both gender and indigenous dummies (for gender, this happens at the median, and for ethnicity at the 90th percentile). An approximation for the gender and indigenous marginal values can thus be obtained by simple summation of each parameter estimate and the interaction estimates. Applying this method to gender and interaction estimates in column (1) gives a fall of -13.6% compared to a fall of -9.7% with column (6) estimates for gender. Similarly, addition of the indigenous and interaction estimate provides an approximate percentage of fall in weekly wages by indigenous status. The calculation in this case leads to a fall of 20.2% based on column (1) estimates and but -16.6% based on column (6) estimates. Hence, while both models suggest large falls, there are notably smaller declines obtained by the two-step selectivity model compared to the single equation model[5]. Some further evidence on the impact of time dummy for 1985–94 and rural-urban migration and sectoral changes in employment are given in Appendix 2.

5 The exact percentage changes in earnings by the difference of 1 from the exponential values of the above calculations, in percentage terms, are all somewhat lower, but the key differences between the OLS values and the selectivity model values remain much the same as above.

6. CONCLUSION

This paper examined ethnic and gender discrimination in weekly wage earnings for Peru employing the LSMS of 1985 and 1994. We estimated a base-line, single equation linear wage function with an observable set of explanatory variables, and tested that model for whether hours of labor supply is endogenous. The model appeared well-determined with exogenous labor supply and with gender and ethnic estimates consistent with discrimination. However, the literature attributes a large share of pay differential to a mix of observable and unobservable factors. In this study we attempted to isolate the latter by estimation of a sample selectivity bias control, and the preliminary results suggest that once a portion of variation in log of wage due to unobservable factors is accounted for, the null hypothesis of no estimation bias is rejected. We obtained estimates of ethnic and gender effects on wage rates with control of unobservable factors, though the method does not offer separate estimates for discriminatory and non-discriminatory components of that control. Our best results, obtained with a two-step selectivity model with exclusion restrictions imposed, demonstrate statistically significant, strong negative effects of gender and ethnicity on individual wage rates. We also explored the bearing in our approach to some key influences regarded as root causes of Latin America's inequality in Appendix 2.

The method for test and control of unobservable factors employed in this study are standard and widely applied in many fields. However, its application to an analysis of labor market discrimination in Peru has not been explored despite the focus in the literature on sources of exclusion and discrimination. The contribution of this paper are two folds. First, it demonstrates the potential of the approach for a developing Latin America economy with a large informal sector where it is critical to model unobservable factors to obtain far more credible estimates of ethnicity and gender effects on wage rates. Its second contribution is in obtaining empirical evidence that labor market entry in Peru is not affected by productivity (education) differences but mainly by the sectors of the economy; hence exclusion appears to be the outcome of limited labor absorbing capacity between available skills and available jobs. However, the evidence suggest public policy should pay attention to the significant impacts of differences in productivity and education on earnings, and the presence of significant degree of gender and indigeneity discrimination faced by employed Peruvians.

Table 2.1: List of Mean and Standard Deviation Values of Each Variable

Variables	Mean	Standard Deviation
Logwage	4.227612	1.297949
Experience (years in same job)	0.7605838	3.87881
experience2	15.62306	122.6202
Working hours	0.7790246	2.499058
Lima (1 if living in Lima)	0.280339	0.4491736
unionmemb (1 if unionized)	0.0133794	0.1148951
indigenous (1 if indigenous)	0.1074666	0.3097116
Gender dummy (1 if female)	0.5195982	0.4996256
married (1 if married)	0.6227096	0.484718
payinkind (1 if pay-in-kind)	0.0072978	0.0851168
Age	33.40264	13.99608
smallchild (1 if child < 6 years)	0.562012	0.4961493
Childnumber	0.9997646	1.137146
edu1 (1 if illit. or basic educ.)	.0576372	.2330607
edu2 (1 if primary diploma)	.320791	.4667898
edu3 (1 if secondary diploma)	.3535528	.4780818
edu4 (1 if technical diploma)	.0247577	.1553889
edu5 (1 if post-2nd diploma)	.0475144	.2127406
edu6 (1 if univ. degree or +)	.095382	.2937476
agriculture (1 if in agri. sector)	.2103818	.4075878
comm_finan (1 if com. or fin)	.1049947	.3065526
indust_const (1 if ind. or cons)	.079413	.2703876
service (1 if in service sector)	.0896928	.2857467
sample size	25,487	

Table 2.2: Sample Distribution of Zero and Missing Wage Observations by Indigenous Status, Gender & Sector (Percentages in Brackets)

wage missing or wage=0 & (hours=0 or >0)	wage>0 and hours>0
23,091 (90.6)	2,396 (9.4)
Wage missing or Wage=0 & Hours>0	Wage missing or Wage=0 & Hours=0
202 (0.9)	22,889 (99.1)
(Wage missing or Wage=0 & Hours=0 or >0) & (wage>0 and hours>0) Sample (25487):	
Indigenous	Non-Indigenous
2739 (10.8)	22748 (89.2)
Man	Woman
12244 (48.0)	13243 (52.0)
Wage>0 & Hours>0 Sample (2396):	
Indigenous	Non-Indigenous
179 (7.5)	2217 (92.5)
Man	Woman
1441 (60.1)	955 (39.9)
Agriculture	Commerce & Finance
503 (21.0)	623 (26.0)
Construction & Industry	Service
613 (26.0)	655 (27.0)48.0)

Table 2.3: Estimation of Log of Weekly Wage in Peruvian Labor Market, 1985– 1994 (standard errors in brackets are corrected for heteroscedasticity in columns 1–2, 4 & 6)

	eq. 1	eq. 1 + *rho*	eq. 2 probit	eq.4+eq.2 *imr*	eq.3probit	eq. 4+eq.3 *imr*
Hourswork	.057 (.008)**	.052 (.014)**	.559(,016)**	.038(.012)**	.559(.016)**	.039(.012)**
educ2	.367 (.120)**	.368 (.121)**	.043(.109)	.372(.094)**	.029(.109)	.371(.094)**
educ3	.668 (.122)**	.669 (.122)**	.014(.114)	.670(.099)**	-.017(.117)	.670(.099)**
educ4	.627 (.154)**	.629 (.155)**	.0157(.158)	.630(.141)**	.136(.161)	.630(.141)**
educ5	.898 (.142)**	.896 (.142)**	-.119(.169)	.893 (.131)**	-.128(.172)	.894 (.131)**
educ6	1.26 (.133)**	1.26(.133)**	.108(.132)	1.26(.115)**	.105(.135)	1.26(.115)**
Experience	.041(.007)**	.044(.011)**	.119(.011)**	.037(.007)**	.123(.011)**	.038(.007)**
$(\text{experrien})^2$	-.001(.000)**	-.001(.000)**	-.003(.000)**	-.001(.000)**	-.003(.000)**	-.001(.000)**
indust_con	.535 (.075)**	.542 (.077)**	1.43(.096)**	.503 (.071)**	1.45 (.098)**	.505 (.071)**
finan_com	.661 (.076)**	.668 (.079)**	1.29(.092)**	.631(.073)**	1.31 (.093)**	.633 (.073)**
Service	.513 (.078)**	.520 (.080)**	1.48(.098)**	.484 (.073)**	1.52 (.100)**	.486(.073)**
Smallchild	-.095(.046)*	-.094(.046)*	-.090(.056)	-.095(.046)*	-.031(.088)	-.095(.046)*
Married	.393(.050)**	.390(.050)**	-.075(.062)	.40 (.052)**	.036(.082)	.400(.052)**
Lima	-.191(.066)**	-.190(.066)**	.027 (.090)	-.200(.062)*	.023(.090)	-.200(.062)*
Union	.534(.051)**	.550(.070)**	2.85(.562)**	.508 (.068)**	2.84 (.558)**	.510 (.068)**
Year	-.083(.007)**	-.083(.007)**	-.129(.014)**	-.082(.006)**	-.130(.014)**	-.082(.006)**
Gender	-.525(.049)**	-.524(.049)**	.206(.058)**	-.529(.043)**	.200(.058)**	-.529(.048)**
Indigen	-.591(.118)**	-.591(.118)**	.136 (.155)	-.598(.101)**	.148 (.155)	-.598 (.101)**
gend*indig	.389(.194)*	.388(.194)*	-1.03(.218)*	.435(.182)*	-1.03(.217)**	.432(.182)*
*rho*_hours		.004 (.013)				
Age					-.030(.009)*	
$(\text{age})^2$					.0003(.0002)*	
children #					.022(.022)	
Imr				-.130(.0623)*		-.120 (.0624)*
Cons	168.(13.0)**	166.(14.0)**	253(28.0)**	165(12.7)**	256(28.1)**	165. (12.7)**

Notes:

* indicates significance at α=5%, and ** indicates significant at α=1%. **col.1**-OLS estimates by (1), R^2=0.345; **col.2**-equ. (1) plus *rho* for a Hausman test of *hoursw* with *pay-in-kind* as instrument (2sls 1st stage=6.04 with *t*-ratio= 23.38); **col. 3**. by probit equ. (2) with dep. variable=1 if observed to be in the market; =0 otherwise; **col. 4**. equ. (4) plus *imr* obtained from equ. (4)/col. 3- *t*-ratio on coefficient of *imr*=-2.09; corrected standard errors also adjusted for $\hat{\lambda}$, **col. 5** by probit equ. (3) plus *age, age2, and no. of children*, **col. 6**. equ. (4) plus *imr* obtained from equ. (3)/col. 5- *t*-ratio on coefficient of *imr*=-1.93; corrected standard errors also adjusted for $\hat{\lambda}$.

REFERENCES

Argumedo, A, and M. Pimbert. 2010. "Bypassing Globalization: Barter Markets as a New Indigenous Economy in Peru." *Development* 53: 343–9.

Barron, M. 2006. "Exclusion and Discrimination as Sources of Inter-Ethnic Inequality in Peru." Documento De Trabajo *253*, http://www.pucp.edu.pe/economia/pd/DD233.pdf

Cameron, A. C., and P. K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Field, E. 2007. "Entitled to Work: Urban Property Rights and the Labor Supply in Peru." *Quarterly Journal of Economics* 122 (4): 1561–1602.

Figueroa, A. 2010. "Is Education Income-Equalizing? Evidence from Peru." *Cepal Review* 102, (Dec.), 113–133

Heckman, J. 1976. "The Common Structure of Statistical Models of Truncation, Selection and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5 (4): 475–92.

Heckman, J. 1990. "Variety of Selection Bias." *American Economic Review* 88 (2) (papers & proc.): 313–318.

Heckman, J. 1998. "Detecting Discrimination." *Journal of Economic Perspectives* 12 (4): 101–116.

MacIssac, D. J., and H. A. Patrinos. 1995. "Labour Market Discrimination against Indigenous People in Peru." *Journal of Development Studies* 32 (2): 218–233.

Moreno, M., H. Nopo, J. Saavedra, and M. Torero. 2012. "Detecting Gender and Racial Discrimination in Hiring Through Monitoring Intermediation Services: The Case of Selected Occupations in Metropolitan Lima, Peru." *World Development* 40 (2): 315–328.

Nopo, H. 2009. "The Gender Wage Gap in Peru 1986–2000: Evidence from a Matching Comparisons Approach." *Working Paper # 675 RG_N3338, Inter-American Development Bank*.

Nopo, H. 2012. *New Century, Old Disparities: Gender and Ethnic Earnings Gaps in Latin America and the Caribbean*. Washington D.C.: Inter-American Development Bank.

Nopo, H, J. Saavedra, and M. Torero. 2004. Ethnicity and Earnings in Urban Peru. ***IZA*** Discussion Paper No. 980.

O'Neill, J. E., and D. M. O'Neill. 2005. "What Do Wage Differentials Tell Us About Labor Market Discrimination." *Working Paper # 11240, National Bureau of Economic Research*.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge: MIT Press.

6.1-Appendix 1: Test and Control of Sample Selectivity Bias by Two-step Model

A major shortcoming of a single-equation wage discrimination model is that it ignores the impact of the working age population outside the labor market on the parameter estimates obtained. There are, however, different possible mechanisms that transmit the effects of market entry to the wage equation. For instance, more able workers are also more likely to be observed in the labor market *and* earn higher wages; exclusion of such effects leads to biased estimates. This bias will be more pronounced if there are more zero observations for entry than is consistent with the wage model. In the sample working age individuals of this study, zero wage observations are nearly 10 times larger than observations with positive wage values (see Table 2.2). One common approach to this problem is to adopt a two-step model with separate specifications for the market entry decision and for the amount earned conditional on that decision. One estimates the probability of entry given the workers characteristics such as age or education. Obvious choices are, for the entry equation, a probit specification and, for the wage equation, a lognormal specification.

$$D_h = x_{1h}\beta_1 + \mu_{1h} \tag{1}$$

$$\ln w_h = x_{2h}\beta_2 + \mu_{2h} \tag{2}$$

and

$$D_h = 1 \text{ if } x_{1h}\beta_1 + \mu_{1h} > 0 \;\&\; D_h = 0 \text{ otherwise} \tag{3}$$

where x_1 and x_2 are vectors of explanatory variables for individual h, including the intercept, and it is further assumed that μ_1 and μ_h are each normally distributed for the sample individuals. This is the *hurdle model* since a positive value for entry is observable only if it crosses a threshold set for D_h. The hurdle model is useful if x_{1h} and x_{2h} are fully *observable* and it is typically applied by employing the same set of explanatory variables in both (1) and (2). The key assumption, however, is that μ_{1h} and μ_{2h} *independent* from each other, see Wooldridge (2010, pp. 694–6); therefore, the expected value of earnings is conditioned solely by observables that appear in x_{1h} but not through any correlation in the error terms of (1) and (2). Since logarithmic specification requires using positive w_h>0 observations, expectation of $\ln w$ conditional on D_h=1 is

$$E(\ln w) = E[(x_2\beta_2) + \mu_2 \mid x_1\beta_1 + \mu_1 > 0 = x_2\beta_2 + E[\mu_2 \mid \mu_1 > -x_1\beta_1].$$

Given independence of μ_1 and μ_2, $E(\mu_2)$=0 regression of ln$w$ on x_2 will produce consistent estimates ofβ_2. However, (1) may contain *unobservable* factors that influence both (1) and (2). Such a common list of omitted variables would create correlation between μ_1 and μ_2. For example, differences in ability but also ethnic and gender discriminatory influences, such as absence of property rights, and poor quality of education. Such influences would affect the unobservable components of the gender and ethnic gaps in both market entry and in wage payment, although the observable components of gender and ethnicity gaps can be further controlled in (1) and (2) by an addition of an interactive term between gender and indigenous variables. However, joint variation of μ_1 and μ_2 violates the classical assumption of a zero mean μ_2 error tem in (2); that assumption is required to ensure sample randomness. Lack of randomness implies biased and inconsistent estimates obtained from the selected sample in question, namely, more able workers disproportionately have selected themselves in the sample. Heckman proposed a solution to overcome this problem by assuming μ_1 and μ_2 have joint normal distribution so their dependence can be expressed linearly by

$$\mu_2=\sigma_{12}\mu_1 + \xi \tag{4}$$

where σ_{12} measures the covariance and where ξ is a mean zero classical error term independent of μ_1. The two-step model with jointly normal error terms provides the basis of the most popular variety of Heckman's sample selectivity bias model, Heckman (1976), and that is also the model briefly outlined below.

The expected mean for (2), given w>0, is now

$$E(\ln w \mid x_1, D>0)=E[(x_2\beta_2) +\mu_2| \; x_1\beta_1+\mu_1>0]=x_2\beta_2 +E[\mu_2| \; \mu_1 > - \; x_1\beta_1]$$

$$=x_2\beta_2 +E[\sigma_{12}\mu_1 + \xi \mid \mu_1 > - \; x_1\beta_1]=x_2\beta_2 +\sigma_{12}\mathrm{E}[\mu_1| \; \mu_1 > - \; x_1\beta_1] \tag{5}$$

Division of μ_1 by σ_1 turns (μ_1 /σ_1) into a standard normal distribution variable $[\mu_1/\sigma_1 > (- \; x_1\beta_1)/ \; \sigma_1]$. Using the property for the expected mean of a truncated distribution[6], we can obtain the conditional expectation of ln$w$ as

$E(\ln w \mid x_1, D>0)=x_2\beta_2 +\sigma_{12}\lambda$

where φ stands for *pdf* and ϕ for *cdf*, and

6 For z~N(0, 1), $E[z|z>c]= \varphi(c) / 1- \phi(c)$; $\phi(c)$ measures probability by the area in the standard normal ϕ to the felt of c, $\Pr[z \leq c]$, hence $\Pr[z>c]=1-\phi(c)$ for a constant c. In this case, $\varphi(-c)= \varphi(c)$, and

$1-\phi(-c)= \phi(c)$, hence $E[z|z> -c]= \varphi(c) /\phi(c)$.

$$\lambda = \frac{\phi(-\mathrm{x}_1'\beta_1 / \sigma_1)}{\Phi(-\mathrm{x}_1'\beta_1 / \sigma_1)} \tag{6}$$

λ is known as the *inverse Mills ratio* (**imr**). The Heckman selectivity model obtains estimates of (6) from a probit equation with a binary dependent variable, and adds that as an additional exogenous variable to the log wage equation:

$$\ln w_h = x_{2h}\beta_2 + \delta\lambda_h(-x_1\beta_1) + \nu_h \tag{7}$$

(7) now replaces (2) to control for omitted unobservable factors, and ν_h is a normally distributed error term independent from the explanatory variables in (7), including λ. This model provides a direct test for the null hypothesis of no selectivity bias based on the *t*-ratio obtained for its coefficient estimate, $\hat{\delta}$. A statistically significant value rejects the null hypothesis; while adding a significant *imr* in (7) will control for unobservable sources of estimation bias, some differences such as ability are non-discriminatory, and others, such as differences in quality of education, are discriminatory in terms of impact on exclusion and pay differential. Therefore, estimation of (7) by OLS will provide unbiased and consistent parameter values. Heteroskedasticity correction for standard errors for this model is complicated since it relies on an estimate of λ; therefore, adjustment must be made for the fact that λ itself is an estimate, in addition to any correction necessary for homoscedasticity.

The above model will be poorly identified if an identical vector of variables is employed in entry and in wage equations; therefore, the simplifying assumption that x_1=x_2 is far more critical for the selectivity model than for the two-step model with all independent variables fully observable. This results fromx_1=x_2, λ(.) in (7) which would be highly collinear with another set of explanatory variables in that equation, namely x_2. Given x_1=x_2, it would be very likely that point estimates obtained for (1)–(3) will have larger standard errors. This highlights the need for additional variation unique to $x_1\beta_1$ in the probit equation for more accurate estimates. There is therefore, a critical role for employment of one or more *exclusion restrictions* on variables that are included in the probit equation but excluded from the wage equation. With $x_1 \neq x_2$, exclusion restrictions will introduce greater variability across $x_1\beta_1$ observations to enable the model to tell apart the effects of participants and nonparticipants on earnings.

Of course, even with x_1=x_2, we still have some variation unique to the probit model of entry by the nature of its nonlinear functional form and in contrast to the log linearity of the wage equation. With this type of exclusion restriction, parameter identification relies solely on functional form restriction. If however, non-linearity is limited rather than pronounced, then identification by functional form would still be rather poor; therefore, effective selectivity bias test and control often rely on the critical role played by at least one exclusion variable blocked from the wage equation and unique to the entry probit equation, for example, fixed costs of entering the labor market such as time or distance to and from work, or age and age squared.

6.2-Appendix 2-Impact of Time Dummy on Rural-urban and Sectoral Employment Differences

Some aspects of the results presented above have implications for the broader themes of inequality in Latin America. Perhaps the most striking features of the two-equation model employed in columns 3–6 are insignificant educational levels and indigenous dummy; only gender is significant in both market entry and wage equations. This suggests, given control for the unobservable, education as a proxy for productivity differences plays little part in exclusion from the labor market, while education has highly significant positive impact on earning differences once in employment; both gender and indigeneity also affect earnings significantly, and in the expected negative direction. Some other estimates also provide indirect suggestive evidence on related issues of labor migration, and the relative size of the service sector; Lima dummy is significant only in the wage equation; insignificant in the market entry equation, though the three sectoral dummies remain highly significant in both the entry and wage equations.

Despite the lack of direct information on Peru's outmigration and informal employment in the data set, it is possible to obtain some idea about the root causes of Latin American inequality, using interactive terms for the consistently significant variable year in Table 2.3 for change over 1985–1994. The results are reported in Table 2.4 using the same entry-wage equations as in columns 5 & 6 of Table 2.3, but this time once with date and Lima interactive in columns 1 and 2, and once with date and sector interactives in columns 3 and 4. The outcome is broadly similar to that evident from Table 2.3, columns 5 and 6. Entry is significantly affected by Lima*date interactive (column 2) only in the wage equation, the main influences on entry are sectors and formal sector union membership, among others. We note that exclusion restrictions by age and age squared, required for identification of the parameters of the wage equation, are both significant, as is *imr* control for the unobservable. The same conclusion applies to the results in Table 2.4, columns 3 and 4 based on statistically significant date*sector interactives, with the important difference that *imr* is no longer significant. It is an open question whether this outcome is due to the effective control for the unobservable, given significant age exclusion restriction, or due to a limited number of restrictions. The answer may require a longer list of restricted variables such as age, education and employment status of the individuals' spouses to explore in a future study. Thus, Table 2.4 provides some indirect evidence of the influences of education, productivity, migration and informality on earnings differences.

Table 2.4: Log of Wage Estimates Inclusive of Interactives with Date, 1985– 1994 (Standard Errors in Brackets are Corrected for Heteroscedasticity)

	eq.3probit	eq.3probit	eq. 4+eq.3 *imr*	eq. 4+eq.3 *imr*
Hourswork	.559(.016)**	.040(.012)**	.555(.016)**	.048(.011)*
Educ2	.028(.109)	.375(.095)**	.075(.108)	.387(.094)**
Educ3	-.018(.116)	.671(.099)**	.019(.114)	.705(.098)**
Educ4	.135(.161)	.636(.141)**	.185(.160)	.669(.141)**
Educ5	-.129(.172)	.892(.132)**	-.148 (.177)	.934 (.131)**
Educ6	.104(.135)	1.26(.115)**	.152 (.133)	1.31 (.114)**
Smallchild	-.031(.088)	-.098(.046)*	-.022(.089)	-.098(.046)**
Experience	.123(.011)**	.038(.007)**	.136(.011)**	.038(.007)**
$(\text{Experrien})^2$	-.003(.000)**	-.001(.000)**	-.003(.000)**	-.001(.000)**
Indust_Con	1.45(.098)**	.504(.071)**	_	_
Finan_Com	1.31(.093)**	.631(.073)**	_	_
Service	1.52(.100)**	.487(.073)**	_	_
Indust*date	_	_	1.15 (.075)**	.336 (.048)**
Finan*date	_	_	.995(.064)**	.370(.048)**
Service*date	_	_	1.24(.078)**	.272(.051)**
Lima	_	_	-.116(.096)**	-.234(.062)**
Lima*date	.007(.052)	-.064(.035)*	_	_
Year	-.130(.014)**	-.085 (.007)**	-.185(.051)**	-.110(.007)**
Union	2.84(.558)**	.516 (.066)**	2.82(.568)**	.523(.067)**
Married	.035(.082)	.402 (.052)**	.042(.084)	.408(.052)**
Gender	.197(.058)**	-.528 (.048)**	.170(.059)**	-.504(.048)**
Indigen	.148 (.155)	-.596 (.101)**	.247 (.154)	-.574 (.102)**
Gend*Indig	-1.03(.217)**	.426 (.182)**	-1.03(.220)**	.387(.182)*
Age	-.028(.014)*		-.029 (.014)*	
$(\text{Age})^2$	.0003(.0001)**		.0003(.0002)	
Children #	-.044(.039)		-.041(.039)	
Imr		-.114 (.062)*		-.060 (.061)
Cons	255(28.7)**	172(13.5)**	366 (29.9)**	221 (13.8)**

Notes:* indicates significance at α=5%, and ** indicates significant at α=1%. **Col. 1**. by probit equ. (3) with dep. variable=1 if observed to be in the market; =0 otherwise, plus interactive terms of date and sectors, **col. 2**. equ. (4) plus *imr* obtained from equ. (3) plus interactive terms of date with sectors-*t*-ratio on coefficient of *imr*=-1.83; corrected standard errors also adjusted for $\hat{\lambda}$, **col. 3** by probit equ. (3) plus *age, age2, and no. of children* and plus interactive terms of date with Lima; **col. 4**. equ. (4) plus *imr* obtained from equ. (3) plus interactive terms of date with sectors—*t*-ratio on coefficient of *imr*=-0.97; corrected standard errors also adjusted for $\hat{\lambda}$.

CHAPTER THREE

Household Shocks and Child Labor Incidence: Evidence from Peru

ROGER WHITE[1] AND FORREST ROULEAU

1. INTRODUCTION

The impetus for the work presented in this chapter is the continued prevalence of child labor and the associated ethical dilemmas that cut across numerous social, economic, and cultural environments. One explicit consequence of child labor is its potentially negative influence on child development and, thus, on child welfare[2]. Unfortunately, especially in developing economies, households sometimes find it necessary to rely on a child's earnings to have sufficient income on which to live[3]. Although more restrictive labor laws in advanced economies serve to limit child labor, simply declaring child labor an illegal activity in a developing economy is often neither a practical nor an economically-feasible impediment to child labor. To reduce child labor incidence, it is critical to address the factors that contribute to household reliance on children's earnings. Necessary steps in this process include identifying and understanding the factors that result in parents

1 Corresponding author. Email: rwhite1@whittier.edu. 13406 E. Philadelphia Street, Whittier, CA 90602 USA. Phone: (562) 907–4908. Fax: (562) 907–4956.

2 Heady (2003) and Rosati and Rossi (2003), for example, report that engagement in child labor is negatively related to primary school test scores and, thus, learning achievement. Similarly, Saddik et al. (2005) finds that exposure to solvents adversely affect the neurobehavioral performance, memory, and motor dexterity of working children.

3 While our data does allow for the identification of child labor activity, it is not sufficiently detailed to permit us to identify hazardous forms of child labor (i.e., work performed in dangerous and/or unhealthy conditions that may result in a child's death, injury, or illness).

sending their children to work. Towards this end, we augment the existing literature on child labor by examining a potential relationship between household shocks and child labor incidence.

Few studies analyze the influences of household shocks on child labor. Of those that do, Beegle, et al. (2006) reports that household shocks in Tanzania, especially those related to crop losses, increase the estimated likelihood of child labor by 30%. Acheampong and Huang (2018) provide similar findings, noting that agricultural shocks increase child labor hours within households in Nigeria. Expanding beyond agricultural shocks, Vasquez and Bohara (2010) find that natural disasters are more likely to increase child labor incidence in Guatemala, as compared to socioeconomic shocks such as inflation. While these three studies examine several facets of the relationship between household shocks and child labor activity, the associated literature is quite limited, perhaps due to a paucity of data/information on shocks; thus, many related empirical questions remain open.

To estimate the potential influences of household shocks on the probability that the typical child (5–14 years of age) engages in labor activity, we estimate a series of binomial logit specifications while controlling for associated child-, parent-, and household-specific influences. We adopt the definition provided by the International Labour Organization (ILO) and identify work activities that are detrimental to a child's development as child labor. Our data are from longitudinal household surveys that were conducted in Peru during the period from 2002 through 2017 as part of the International Study of Childhood Poverty (UK Data Service, 2019)[4]. The uniqueness of our data permits consideration of whether and to what extent household shocks influence the probability that a child engages in labor activity. We also determine whether the impacts of household shocks on child labor vary based on the number of shocks experienced, and we consider whether the relationship between household shocks and child labor differs based on the type of shock(s) that the household experiences.

Our results consistently indicate a positive and statistically significant relationship between household shocks and the likelihood that a child will engage in one or more labor activity. For example, for an otherwise typical child, each household shock experienced since the most recent survey date increases the predicted probability of engagement in child labor by 1.26 percentage points. When we employ a binary measure of household shocks (i.e., no shocks or one or more

4 The data used in this publication come from Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India (Andhra Pradesh and Telangana), Peru and Vietnam (www.younglives.org.uk). Young Lives is funded by UK aid from the Department for International Development (DFID) and co-funded by Irish Aid. The views expressed here are those of the authors. They are not necessarily those of Young Lives, the University of Oxford, DFID or other funders.

shocks), we find that experiencing one or more household shocks increases the estimated probability of child labor incidence by 4.12 percentage points. Considering variation in the relationship across the number of shocks a household experiences, we find a distinct and largely consistent pattern of statistical significance and coefficient magnitudes. Experiencing a single shock since their last survey does not significantly increase the probability of child labor activity. However, if the household experienced two shocks since their most recent survey, the estimated probability of child labor increases by 4.2 percentage points. Interestingly, the difference between the "one shock" and "two shocks" coefficient estimates is significantly different from zero; however, the estimated coefficients that correspond with three, four, and five or more household shocks do not differ significantly from the coefficient obtained in the "two shocks" case. Further, when employing separate variables to represent seven different forms of household shocks, we find that experiencing an economic shock or a family shock increases the probability of child labor incidence by 3.34 percentage points and 1.86 percentage points, respectively. Lastly, we also report statistically significant relationships between specific forms of child labor activity (i.e., paid labor, home care, and household tasks) and the incidence of household shocks, the number of shocks experienced (i.e., one shock, two shocks, etc., up to five or more shocks), and the type(s) of shock(s) experienced (e.g., crime shocks, economic shocks, environmental shocks, housing shocks, etc.).

The next section reviews the related literature and includes discussions of the child-, parent-, and household-specific characteristics that we consider as potential determinants of child labor incidence. Particular focus is placed on studies that examine the influence(s) of shocks on child labor. This is followed in Section 3 by the explicit statement of our research questions, the presentation of our empirical model, and the introduction of our data, variables, and econometric methodology. Our results are detailed in Section 4, and Section 5 concludes.

2. REVIEW OF RELATED STUDIES

The theoretical works that examine the varying facets of child labor generally follow household production theory or a bargaining model approach. Household production theory assumes a common utility function that members of the household seek to collectively maximize subject to the constraint of total household income. The bargaining model approach allows each member of the household to have a different utility function that they individually seek to maximize subject to a common/household income constraint. Both approaches emphasize intrahousehold resource allocation. Given the prior establishment of theoretical bases, and as our study is an empirical work, we forgo a detailed discussion of the related theoretical literature in this section and instead discuss the bases for our study in greater depth in Section 3. Discussions of the theoretical modeling that serves as a foundation for empirical studies on child labor can be found in Ray (2002), Basu (1999), Basu and Van (1998), Grootaert and Kanbur (1995), Moehling (1995), and Bourguigon and Chiappori (1994), among others.

To provide context for our study, we focus our review on the empirical works that analyze child labor. These studies can generally be categorized as either macro-oriented or micro-focused. The macro-oriented works analyze the relationships between child labor incidence and macroeconomic variables such as national income, income inequality, economic development/growth, international trade, foreign direct investment, etc. The literature that is micro-focused examines child labor incidence while emphasizing poverty, returns to education, imperfections in land and labor markets, fertility, and gender-based differences. These two perspectives are certainly connected. For example, macroeconomic changes often impact decisions at the micro/household level. Similarly, such changes—especially when negative—are likely amplified in relatively less developed countries where it is more common for households to have a reduced ability to counter or negate change(s). Central to the topic of our study, shocks may influence household decisions including those related to child labor activity. Even so, as noted in Acheampong and Huang (2018), few studies have examined the relationship between household shocks and child labor incidence.

One of the most frequently researched determinants of child labor incidence is household income. If a family is unable to meet its basic needs, children may be employed to earn additional income; otherwise, children will generally engage in developmental activities. In their seminal paper, Basu and Van (1998) classify this as the luxury axiom, within which parents are assumed to exhibit altruistic behavior towards their children. The substitution axiom states that in competitive labor markets firms view child and adult labor as substitutable (Basu and Van, 1998). Thus, when adult labor is insufficient to meet the basic needs of a household, incentives at the household-level and the conditions facing firms are conducive to children working to supplement the family income. These two fundamental assumptions have been essential to all subsequent studies. Related models have proposed inefficiencies in labor markets or limited returns to education as underlying factors that affect household income and, thus, as reasons for the greater prevalence of child labor (Bharadwaj, 2015; Emerson and Knabb, 2006).

Several studies extend from the literature on household income and child labor incidence to examine the relationship between child labor and education/schooling[5]. The connection between these two strands of the literature is that child labor activity presents a trade-off where immediate earnings are valued over

5 See, for example, Doepke and Zilibotti (2005), Fan (2004a and 2004b), Hazan and Berdugo (2002), Ranjan (2001), Baland and Robinson (2000), and Basu and Van (1998).

the child's potential future earnings[6]. Often, researchers treat child labor and education as substitutes; children are classified as either attending school or as working. Increases in the number of hours worked is, thus, expected to decrease the amount of time devoted to education and, thus, to limit investments in formal education. This, in turn, results in lower levels of human capital formation which leads to a diminished capacity for innovation and potentially hinders technological advancement of a country. However, when returns to education are low, children may be compelled to work. In other words, only when the potential for upward socioeconomic mobility is sufficiently high would we expect to see poor households emphasizing education rather than work (Fors, 2012).

In addition to expectations of the probability of upward mobility of an educated child, parental attitudes towards education significantly impact the decision of whether a child works. Because attitudes are often informed by personal experiences, the relationship between engagement in child labor and educational attainment likely depends on the educational background of the parents (as well as other household characteristics) (Bornstein and Putnick, 2015; Soares, et al., 2012; Ersado, 2005; and Bhalotra and Tzannatos, 2003). Generally, a negative relationship has been found between parents' educational attainment and child labor activity (Strauss and Thomas, 1995). One example comes from Vasquez and Bohara (2010) who find that the probability that a child is employed increases over time but does so at a decreasing rate, with the educational attainment of the mother being a statistically stronger indicator (relative to the father's educational attainment) that a child will attend school rather than engaging in labor activity. Supporting this finding, Kurosaki, et al. (2006) find that the educational attainment of mothers in rural India has a greater effect in reducing child labor incidence than does the educational attainment of fathers. However, a contrary example is provided by Emerson and Souza (2007) who examine data for Brazil and find that child labor is reduced to a greater extent by the educational attainment of fathers as compared to that of mothers.

Several other factors have been identified as additional determinants of child labor incidence. Some studies note gender-related differences in child labor activity. For example, Bhalotra (2007) extends from Basu and Van (1998) to examine whether poverty is the impetus for child labor. Using survey data from Pakistan, Bhalotra estimates wage elasticities and identifies differences across genders. A negative estimated wage elasticity is found for boys, while the wage elasticity estimate for girls is insignificant from zero. This suggests that household poverty leads boys to engage in labor activities but the same may not be said for girls. Overall, Bhalotta reports that the probability of child labor is greater for boys

6 Beegle, et al. (2006) show that this tradeoff between child labor and education is a non-linear relationship.

than it is girls. In contrast, the impact of household size on child labor incidence is not entirely clear. On the one hand, Basu and Van (1998) found that increases in household size and higher fertility rates lead to an increase in child labor. This was reaffirmed by Vasquez and Bohara (2010). However, Acheampong and Huang (2018) found that household size is negatively correlated to the number of hours a child works. Finally, buttressing the results of earlier studies, Acheampong and Huang (2018) find that parental characteristics are significant determinants of child labor incidence.

2.1 Household Shocks and Child Labor

Beegle, et al. (2006) analyzes the impact of household income shocks on the prevalence of child labor in Tanzania. Tanzania is an agrarian economy, and household income shocks—such as crop loss—are found to produce substantial increases in child labor. In rural regions, where agricultural technology and capital are lacking, income shocks can be quite impactful for households. The authors find evidence that in response to these shocks children are often removed from school to assist with household activities. The estimated probability that a child will begin to engage in labor activity is increased by approximately 30% in households that experience a crop loss. Beegle, et al. also reports that household assets function as a buffer between the severity of the shock and incidence of child labor. This means that households often look to alternative strategies to cope with shocks before relying on child labor. Curiously, the authors find that the incidence of child labor increased with the level of household assets. Specifically, a one standard deviation increase in the log value of per capita assets is found to correspond with a 6% increase in child labor hours. This is an important finding that notes that within agricultural economic settings, household wealth (i.e., land in most instances) and child labor are positively correlated. This contradiction of the luxury axiom (Basu and Van, 1998) is labeled the "wealth paradox" by Bhalotra and Heady (2003)[7]. The paradox stipulates that land-rich households are more likely to have children employed than households that are poor and own little if any land at all.

Bandara, et al. (2015) also examines data from Tanzania to estimate the potential influences of household shocks on child labor activity. Considering both income shocks and non-income shocks, the authors report that agricultural shocks lead to statistically significant increases in both the number of total hours worked by children and in the hours that children spend working in agricultural

7 While Dumas (2007), Kruger (2007), and Kambhampati and Rajan (2006) support the finding of Bhlotra and Heady (2003) by presenting evidence in support of a wealth paradox for Burkina Faso, Brazil, and India, respectively, Basu, et al. (2010) document an inverted U-shaped relationship between land holdings and child labor.

activities. Further, the authors consider whether access to credit allows households to counter the influence of agricultural shocks on child labor and find that the number of hours worked by girls is significantly decreased by access to a bank account. No significant influence of bank account access is found for the number of hours worked by boys. Similarly, Guarcello, et al. (2010) note that unexpected shocks, when occurring in conjunction with credit market imperfections, contribute to high costs in terms of human capital accumulation.

Studying household responses to natural disasters and socioeconomic shocks, Vasquez and Bohara (2010) reaffirms the finding of household reliance on child labor that is presented in Beegle, et al. (2006). Examining survey data collected from Guatemalan households during the year 2000, the authors report that child labor is more likely to be observed in response to natural disasters than as a counter or a remedy for socioeconomic shocks. The most common socioeconomic shocks faced by households who participated in the survey were inflation and the incurrence of a disability by a wage earner in the household. Vasquez and Bohara report that households differ in their incidence of child labor depending on household income levels; the authors find that, in response to natural disasters, extremely poor households were less likely to rely on child labor than poor households. The authors speculate that this could be due to restricted access to labor markets for extremely poor households.

In a similar study that focuses on Nigerian households, Acheampong and Huang (2018) notes a significant relationship between agricultural shocks and child labor. In response to an additional agricultural shock, child labor hours per week rise by an estimated 22% for the typical household. Household shocks, such as the illness of a father, are reported to produce an 11% increase in child labor hours per week; however, the illness of a mother does not appear to significantly affect child labor hours. Dendir (2007) examines data collected in Malawi during 1995. Identifying household shocks as accidents, death, and illnesses involving household members, the author reports that no statistically significant increase is found between household shocks and child labor incidence or intensity. Dendir does, however, report statistically significant decreases in child labor activity in response to some shocks. The author hypothesizes that these inverse relationships are due to general decreases in economic activity in households that are dealing with an illness or death of a household member.

In summary, the literature highlights a negative relationship between household income and child labor activity. The persistence of child labor is much more common in developing economies, especially within rural and agricultural communities. Household characteristics such as parental educational attainment, child gender, etc. are statistically significant determinants of whether a child engages in labor activities. Nevertheless, details on whether the impact of household wealth on child labor varies across geographic regions remain limited. Likewise, scant information exists on the relationship between household shocks and child labor. Accordingly, we consider whether the number and type(s) of shocks that a household experiences are significantly related to child labor incidence to be an open empirical question.

3. DATA, EMPIRICAL MODEL AND ECONOMETRIC STRATEGY

3.1 Research Questions and Data Overview

The literature on household shocks and child labor incidence, particularly the studies that examine the influence(s) of shocks at the extensive margin, informs our estimation strategy and provides important information that shapes our expected findings. Based on the literature and informed intuition, we have identified four research questions that can be evaluated using the available data. We wish to clarify these questions before discussing our empirical model and data in greater detail. The questions are:

1. Are children who live in households that experience one or more shocks significantly more likely, relative to children in comparable households that do not experience a shock, to engage in non-developmental labor activities?
2. Do the influences of household shocks on child labor incidence vary significantly based on the number of shocks experienced?
3. Do the influences of household shocks on child labor incidence vary significantly based on the type of shock experienced?
4. Regardless of whether a household experiences a shock, does the probability of child labor incidence vary based on measurable child-, parent-, and household-specific characteristics?

We address these questions by estimating a series of binomial logit specifications[8]. Doing so allows us to isolate the potential influences of household shocks on the probability that the typical child engages in labor activities, while controlling for any influences associated with our remaining child-, parent-, and household-specific control variables.

8 We also estimated our battery of econometric specifications using the random-effects panel logit model technique. In all instances, likelihood ratio tests indicate no significant differences between results from the random-effects panel logit estimations and results from the pooled logit estimations.

Our data are from household surveys that were conducted in Peru during the period from 2002 through 2017 as part of the International Study of Childhood Poverty (Boyden, et al. , 2016). We restrict our data to include only children who were 5–14 years of age as of the date of their survey, and we only include observations that represent children who were not engaged in labor activity at the time of their immediately prior survey[9]. Restricting the data in this manner produces a cohort of children who were not working at the time of the prior survey who may or may not be working as of the date of the survey for which data are included in our sample. Thus, the data set facilitates examination of the influence of household shocks on the labor activities of children who were previously not working. Our data set includes 4,934 observations that provide information for a potentially non-representative sample of 2,292 children[10 11].

3.2 Empirical Model and Econometric Strategy

As noted, to estimate the potential relationship(s) between household shocks and child labor incidence, we employ an ad hoc empirical specification in which we control for child-, parent-, and household-specific characteristics that may influence the likelihood that a child engages in non-developmental labor activities. Equation (1) presents our general form empirical model.

$$Labor_{ijkt} = \alpha_0 + \beta_C C_{it} + \gamma_P P_{jt} + \delta_H H_{kt} + \theta_S S_{kt} + \varepsilon_{ijkt} \quad (1)$$

In equation (1), the dependent variable series $Labor_{ijkt}$ is a vector that includes four dichotomous variables. We estimate equation (1) employing each dependent variable series in turn. These variables indicate whether child i of parents j in household k at time t was engaged in one or more labor activities as of their survey date. Specifically, $Labor_{ijkt}$ includes $Any\ Labor_{ijkt}$ which is equal to one if the child is engaged in $Paid\ Labor_{ijkt}$, $Home\ Care_{ijkt}$, and/or $Household\ Tasks_{ijkt}$ and is equal

9 The minimum age for legal work in Peru is 14 years; thus, we limit our analysis to those 5–14 years of age.

10 Surveys were conducted in five rounds that span the following periods: August-December 2002 (Round 1), October 2006-August 2007 (Round 2), July 2009-March 2010 (Round 3), June 2013-March 2015 (Round 4), and May 2016-February 2017 (Round 5). Because we include only observations that represent children who were at least 5 years of age and who were not working at the time of their immediately prior survey, our data sample includes only observations from Rounds 2 through 5.

11 Of the 2,292 children represented in the sample, 719 appear only once, 504 appear twice, and 1,069 appear three times.

to zero otherwise. Each of these three specific activities is also included in $Labor_{ijkt}$.

The Paid Labor variable is equal to one if the child is employed for wages. Similarly, the Home Care variable is equal to one if the child provides care for a household member (e.g., older adult, sibling, etc.), and the Household Tasks variable is equal to one if the child works for the family business or on the family farm. Each of these variables is equal to zero if the child is not engaged in the corresponding activity.

We consider the performance of Paid Labor, Home Care, and Household Tasks as engagement in non-developmental activities. The ILO defines child labor as follows:

> "The term "child labour" is often defined as work that deprives children of their childhood, their potential and their dignity, and that is harmful to physical and mental development. It refers to work that is mentally, physically, socially or morally dangerous and harmful to children; and interferes with their schooling by: depriving them of the opportunity to attend school; obliging them to leave school prematurely; or requiring them to attempt to combine school attendance with excessively long and heavy work."

The surveys that were conducted as part of the International Study of Childhood Poverty elicited information on the typical number of hours spent each day on a variety of activities. Relying on the ILO definition of child labor, we categorize these activities into two broad groups: developmental activities and non-developmental activities. Developmental activities include hours spent at school, studying outside school, doing household chores, pursuing leisure/play activities, and sleeping. Non-developmental activities include hours spent in paid activities (i.e., Paid Labor), hours spent caring for household members (i.e., Home Care), and performing domestic tasks such as farming and/or working at the family business (i.e., Household Tasks).

Because the ILO also notes that not all work performed by children is detrimental to physical and mental development and that the definition of child labor is conditional, we categorize household chores as a developmental activity[12]. Specifically, the ILO states that:

> "Not all work done by children should be classified as child labour that is to be targeted for elimination. Children's or adolescents' participation in work that does not affect their

12 Household tasks include farming and/or working at a family business. While the ILO defines "assisting in a family business" "outside school hours and during school holidays" as developmental activity, farming and working at a family business during school hours is not developmental activity according to the ILO. Since the household tasks variable does not differentiate between farming and working at a family business and does not indicate when the work is performed (i.e., during or outside school hours or only on school holidays), a decision was made to include household tasks as non-developmental activity.

health and personal development or interfere with their schooling, is generally regarded as being something positive. This includes activities such as helping their parents around the home, assisting in a family business or earning pocket money outside school hours and during school holidays. These kinds of activities contribute to children's development and to the welfare of their families; they provide them with skills and experience, and help to prepare them to be productive members of society during their adult life."

"Whether or not particular forms of "work" can be called "child labour" depends on the child's age, the type and hours of work performed, the conditions under which it is performed and the objectives pursued by individual countries. The answer varies from country to country, as well as among sectors within countries."

Looking to our data and comparing engagement in developmental activities during the typical day reveals that the average child that was engaged in labor activity spent less time at school (0.40 hours), sleeping (0.27 hours), and playing (1.19 hours) relative to the average child who was not engaged in labor activity. The average child engaged in labor activity also spent more time doing household chores (0.63 hours) as compared to the average child who was not engaged in such activity. T-tests of differences in mean values for the specific activities indicate that all differences are statistically significant ($p < 0.01$). Aggregating across all activities categorized as developmental and extrapolating to produce an annual value, we see that engagement in child labor results in a loss of 447.25 hours (i.e., nearly 19 days, which is equal to 5.11% of a year) of developmental activity per year. Extrapolating further, we can say that the typical child, if engaged in labor activity throughout the full decade during which they are 5–14 years of age, will have had about six fewer months (i.e., 0.51 fewer years) of developmental activity relative to a child who does not engage in labor activity. This may have a substantial effect on the child's mental and physical development.

Equation (1) posits that the determinants of child labor incidence include child-, parent-, and household-specific characteristics and household shocks. These are identified by the vectors C_{it}, P_{jt}, H_{kt}, and S_{kt}, respectively. β_C, γ_P, δ_H, and θ_S are coefficients to be estimated that correspond to the variables within each vector. ε_{ijkt} is an assumed stochastic error term.

Turning our attention to the variables that comprise the vectors and beginning with the child-specific explanatory variables, the vector includes the child's age, the squared age value (to capture potential non-linearity with respect to labor activity), a dummy variable equal to one if the child is female, and a dummy variable that is equal to one if the child is Mestizo. The vector of parent-specific explanatory variables includes the highest level of educational attainment among the child's parents, a dummy variable equal to one to identify single-parent households, the head of household's age, the squared age value (to capture potential non-linearity), and a dummy variable equal to one if the household head is female.

We also include a vector of household-specific explanatory variables that includes the number of the child's siblings that live in the household, the total number of individuals who live in the household, a measure of household quality, a dummy variable equal to one if the household owns any animals, four dummy variables that identify the wealth quintile into which the household falls (i.e., bottom 20%, next 20%, middle 20%, and fourth 20%), and three geography-related dummy variables, one that is equal to one if the household is located in a rural area and two others that are equal to one if the household is located in the Selva or the Sierra regions, respectively[13][14].

Lastly, and most relevant for our examination, our empirical model includes a vector, S_{kt}, that includes several measures of household shocks[15]. More specifically, we include a measure of the total number of household shocks experienced by each child since their most recent survey. This measure is used to generate a related dummy variable that is equal to one if one or more household shocks were experienced. Additional dummy variables identify whether the household has experienced one shock, two shocks, etc., up to five or more shocks. Finally, a separate set of variables indicate the numbers of each type of shock that to five or more shocks. Finally, a separate set of variables indicate the numbers of each type of shock that were experienced by a given household. The types of household shocks include crime, economic, environmental, family housing, regulatory, and "other" shocks. We employ these measures in turn, as appropriate, to address our research questions.

Panel A of Table 3.1 reports the frequency of shocks experienced by households in our sample during the period since the most recent survey. We see that the majority of households (63.8%) experienced at least one shock with slightly more than a third of all households (33.9%) having experienced two or more shocks. As might be expected, fewer households experienced four or more shocks (7.7%) as compared to those who experienced two shocks (17.2%) or three shocks (9%). In Panel B of the table, we see that the two most common

13 The household quality index variable is constructed based on the average number of sleeping rooms per person and the construction material of the house (i.e., walls, roof and floor), and the wealth index is composed of three equally-weighted sub-indexes: the housing quality index; an access to services index, and a consumer durables index (Azubuike and Briones 2016).

14 Peru is comprised of the Costa, Sierra, and Selva regions. The Costa region is a narrow coastal region that borders the Pacific Ocean and stretches the length of Peru from the border with Ecuador in the north to the borders with Chile and Bolivia in the south. The Sierra is the highlands. Spanning the north-south distance between Peru's borders with Ecuador and Bolivia, the region stretches eastward from the Costa region over the Andes mountain range to the Selva region. The Selva region includes the jungle that lies east of the Andes, extending to Peru's borders with Ecuador, Colombia, and Brazil.

15 The Appendix includes a complete listing of each household shock type and the various events that fall within each category.

types of shocks are environmental and family shocks which accounted for nearly two-thirds (62.4%) of all shocks experienced. Economic shocks (16.3%) and crime shocks (12.9%) were less common, and housing, regulatory, and other shocks were much less frequently experienced. Lastly, in Panel C, we provide a breakdown of the types of shocks experienced by the total number of shocks experienced. For example, among households that experienced a single shock, family shocks were most common by a considerable margin. Similarly, among households that two or more shocks we see the most common form of shock is environmental.

Table 3.1: Household Shocks, Numbers and Types Experienced

Panel A: Number of Shocks Experienced Since Prior Survey							
		0	1	2	3	4	5 or more
Distribution of Shocks (%)		36.20	29.91	17.19	9.02	4.01	3.67
Panel B: Relative Frequencies of Household Shock Types							
	Crime	Economic	Environ.	Family	Housing	Other	Regulatory
N Shocks	828	1,042	1,997	1,999	100	404	39
% of All Shocks	12.92	16.26	31.16	31.19	1.56	6.30	0.61
Panel C: Distributions of Household Shock Types (%) by Number of Shocks Experienced							
N Shocks	Crime	Economic	Environ.	Family	Housing	Other	Regulatory
1	9.81	11.20	10.98	20.46	0.00	6.30	0.58
2	2.68	4.09	9.11	8.08	1.56	0.00	0.03
3	0.42	0.70	5.48	2.34	0.00	0.00	0.00
4	0.00	0.19	3.25	0.31	0.00	0.00	0.00
5 or more	0.00	0.08	2.34	0.00	0.00	0.00	0.00

Inserting the specific explanatory variables into equation (1) provides our baseline estimation equation which is presented as equation (2).

$$\begin{aligned} Labor_{ijkt} = {} & \alpha_0 + \beta_1 AGE_{it} + \beta_2 AGE_{it}^2 + \beta_3 FEMALE_i + \beta_4 MESTIZO_i + \gamma_1 PARED_{jt} \\ & + \gamma_2 SINGLE\ PAR_{jt} + \gamma_3 HHEAD\ AGE_{jt} + \gamma_4 HHEAD\ AGE_{jt}^2 + \gamma_5 HHEAD\ FEM_{jt} \\ & + \delta_1 SIBLINGS_{kt} + \delta_2 HH\ SIZE_{kt} + \delta_3 HH\ QUALITY_{kt} + \delta_4 ANIMALS_{kt} \\ & + \delta_5 WEALTH\ Q1_{kt} + \delta_6 WEALTH\ Q2_{kt} + \delta_7 WEALTH\ Q3_{kt} \\ & + \delta_8 WEALTH\ Q4_{kt} + \delta_9 RURAL_k + \delta_{10} SELVA_k + \delta_{11} SIERRA_k + \theta_S S_{kt} \\ & + \varepsilon_{ijkt} \end{aligned} \tag{2}$$

Note that, similar to $Labor_{ijkt}$ which includes our set of dependent variables, the vector S_{kt} contains several variables that represent the incidence, quantity, and specific forms of household shocks. Also similar to the dependent variable series, our estimations include different variables to represent household shocks as necessary to address our research questions.

Table 3.2 presents three sets of summary statistics. Mean values and standard deviations for the full sample are provided in column (a). Columns (b) and (c) provide summary statistics for the observations where the household did not experience any shocks and for those where one or more shocks were experienced, respectively. The comparison of sub-samples permits the determination of significant differences in mean values of the listed variables across the two cohorts. Finally, column (d) indicates the results of t-tests of differences in mean values between the cohorts.

Beginning with the full sample, we see that more than one-quarter (28.4%) of all observations were engaged in one or more labor activities as of their survey date. The most common form of child labor (23.6%) is home care. A considerably smaller proportion of the sample (11.7%) provides labor in the form of household tasks, and relatively few observations (1.8%) report being paid labor[16]. Looking to columns (b) and (c) for these same variable series we see that children who live in households that experienced one or more shocks since their most recent survey (column (c)) are significantly more likely, relative to children in households that did not experience a shock (column (b)), to engage in some form of labor activity: 30.8% as compared to 24% ($p < 0.01$). We also see that, in comparison to children who did not face a household shock, those children in households who were exposed to shocks were significantly more likely to begin providing labor in the forms of providing household care and the performance of household tasks.

Table 3.2: Descriptive Statistics

	All	Any Shocks=0	Any Shocks=1	(c)–(b): Stat. Significant?
N=	4,934	1,786	3,148	
	(a)	(b)	(c)	(d)
Any labor	0.2835	0.2396	0.3084	***
	(0.4508)	(0.427)	(0.4619)	
Paid labor	0.0184	0.0162	0.0197	..
	(0.1346)	(0.1264)	(0.1390)	

16 The sum of individual labor activity incidences (37.2%) is greater than the mean value for the Any Labor variable (28.4%) due to some children engaging simultaneously in multiple forms of labor.

Table 3.2: *Continued*

	All	Any Shocks=0	Any Shocks=1	(c)–(b): Stat. Significant?
Home care	0.2361	0.2027	0.2552	***
	(0.4247)	(0.4021)	(0.4360)	
Household tasks	0.1173	0.0851	0.1356	***
	(0.3219)	(0.2791)	(0.3425)	
Child's age	8.9944	8.8745	9.0625	**
	(2.8043)	(2.8116)	(2.7983)	
Female child	0.4834	0.4821	0.4841	..
	(0.4998)	(0.4998)	(0.4998)	
Mestizo child	0.9191	0.9132	0.9225	..
	(0.2727)	(0.2816)	(0.2674)	
Number of siblings	1.8687	1.6573	1.9886	***
	(1.4042)	(1.3107)	(1.4411)	
Parents' highest educational	9.246	9.9199	8.8637	***
attainment	(3.9825)	(3.8149)	(4.0252)	
Single parent (female or male)	0.0807	0.0549	0.0953	***
	(0.2723)	(0.2278)	(0.2937)	
Female household head	0.0675	0.0521	0.0762	***
	(0.2509)	(0.2222)	(0.2654)	
Household size	5.4459	5.2032	5.5835	***
	(1.8717)	(1.7679)	(1.9147)	
Household quality index	0.4355	0.4726	0.4144	***
	(0.2481)	(0.2516)	(0.2436)	
Household owns any animals	0.6336	0.5571	0.6769	***
	(0.4819)	(0.4969)	(0.4677)	
Household wealth index	0.5472	0.584	0.5263	***
	(0.2126)	(0.2103)	(0.2112)	
Rural household	0.2639	0.1865	0.3078	***
	(0.4408)	(0.3896)	(0.4617)	
Costa region	0.3985	0.4485	0.3701	***
	(0.4896)	(0.4975)	(0.4829)	
Selva region	0.152	0.1461	0.1553	..
	(0.3591)	(0.3533)	(0.3623)	
Sierra region	0.4495	0.4054	0.4746	***
	(0.4975)	(0.4911)	(0.4994)	

Standard deviations in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Moving to the explanatory variable of primary interest, we see that the typical child (column (a)) experienced 1.3 household shocks since their last survey. This is consistent with the fact that more children experienced household shocks (63.8%) than did not. Because nearly two-in-five children lived in households that did not experience even one shock, the overall mean value understates the frequency of shocks experienced by children in households that did experience at least one shock. The household of the average child in this cohort experienced more than two shocks since the last survey.

We now look to the remaining explanatory variables and focus on the subsamples of children who lived in households that did not experience a shock (column (b)) and those that lived in households that experienced one or more shocks (column (c)). Comparing across the cohorts, we can say that the typical child who experienced a household shock tends to be slightly older, and they tend to have more siblings. This corresponds with a larger average household size. Additionally, the highest level of educational attainment among their parents is generally lower and they are more likely to live in a single-parent household. We also see that the typical household head is slightly older and significantly more likely to be female. These households tend to be of lower overall quality and to have lower levels of wealth. They are much more likely to be located in rural areas and to own animals. They are also more likely to be located in the Sierra region and are less likely to be located in the Costa region.

The significant differences across cohorts of children categorized based on whether their household experienced one or more shocks since the last survey date suggests that child-, parent-, and household-specific factors explain at least some proportion of the observed variation in child labor incidence. That the mean values for the labor variables are significantly higher for those children who have experienced a shock further suggests that household shocks may, in fact, contribute to an increased likelihood of child labor incidence.

Consideration of the pairwise correlation coefficients among our set of explanatory variables reveals a maximum correlation of 0.83 between the household quality and household wealth index series. Likewise, the correlation coefficient between the variables indicating household size and the number of siblings is equal to 0.80. Otherwise, in absolute values, all pairwise correlation coefficients are less than 0.80 with 88.7% of the values being less than 0.30 and with 97.8% of the values being less than 0.50. Although the relatively high correlation values between the household quality and household wealth index variables and between the variables that indicate household size and number of siblings are somewhat concerning, each of these variables potentially represents unique household attributes that may bear on child labor incidence. Accordingly, we include the variables in the empirical model but proceed with caution when interpreting the associated coefficient estimates.

4. HOUSEHOLD SHOCKS AND CHILD LABOR INCIDENCE

We address each of our research questions, in turn, through the estimation of variants of equation (2)[17]. Table 3.3 presents the results from a pair of binomial logit estimations. The dependent variable series employed in the estimations is a dummy variable that is equal to one if the child is engaged in one or more forms of labor (i.e., paid labor, home care, and/or household tasks) and is equal to zero if the child is not engaged in any form of labor. Other than the measures of household shocks employed, the sets of explanatory variables are identical across the estimations. Column (a) presents results obtained when we use the number of shocks that the child's household experienced since the immediately prior survey date as our household shock variable. The household shock measure employed in our second estimation (column (b)) is a dummy variable that is equal to one if the household experienced any shocks and is otherwise equal to zero. Collectively, the results of this pair of estimations address our first research question: *Are children who live in households that experience one or more shocks significantly more likely, relative to children in comparable households that do not experience a shock, to engage in non-developmental labor activities?*

Focusing first on the estimated log-odds coefficients of the household shock variables, we find a positive and statistically significant relationship between the number of shocks experienced and child labor incidence (i.e., $\hat{\beta}$=0.0699 with $p < 0.01$, reported in column (a)). This positive relationship is consistent with the findings of prior studies on the topic[18]. Holding all else held constant and evaluating the results while employing the mean values for all explanatory variables, the predicted probability that a child will be engaged in labor activity is equal to 23.08%[19]. Experiencing one shock above the corresponding mean value, again with all else held constant, increases the predicted probability of labor incidence to 24.34%[20].

17 A caveat is in order. As noted in the previous section, we employ an ad hoc empirical model. Additionally, our analysis is undertaken using self-reported (i.e., survey) data and, thus, may be prone to measurement error.

18 See, for example, Acheampong and Huang (2018), Bandara, et al. (2015), Vasquez and Bohara (2010), and Beegle, et al. (2006).

19 Predicted probability values are calculated by first obtaining the predicted logit (L). This is done by summing the products of the estimated log-odds coefficients and the mean values of the corresponding explanatory variables and then adding the estimated constant term (i.e., $L = \alpha_0 + \beta_1\bar{X}_1 + \ldots + \beta_n\bar{X}_n$). The probability value is then given as $P = 1/\left(1+e^{-L}\right)$.

20 The change in the predicted probability is equal to $P_1 - P_0$ (i.e., $\left(1/\left(1+e^{-L_1}\right)\right) - \left(1/\left(1+e^{-L_0}\right)\right)$. Here, $L_1 = -1.13425261$ and $L_0 = -1.20419751$, with the difference between predicted logits being equal to the estimated log-odds coefficient of the "shocks" variable: 0.0699449.

Thus, we estimate the effect of experiencing one shock is about a 1.26 percentage point increase in the predicted probability[21].

A somewhat similar result is found when we employ our binary measure of household shocks (i.e., no shocks or one or more shocks). We again find a positive and statistically significant increase in the predicted probability of engagement in child labor activity (i.e., $\hat{\beta}$=0.2358 with $p < 0.01$, reported in column (b)). In response to experiencing one or more household shocks, the predicted probability that the typical child will begin to engage in child labor increases by 4.12 percentage points, from 20.51% when no shocks are experienced to 24.63%. This represents a 20.09% increase in the estimated probability of child labor activity, a value that is similar to the 30% increase in the estimated probability of child labor reported by Beegle et al. (2006) when focusing on the impact of crop loss in Tanzania.

The estimated coefficients of the remaining explanatory variables that are presented in Table 3.3 provide additional interesting information. We begin with the results presented in column (a), focusing particularly on the coefficient estimates that correspond to the variables representing child-specific characteristics. As we have controlled for both the age of each child and the squared value of the age variable, we see that, all else held constant, the estimated probability of child labor incidence decreases from age 5 through 8 before increasing through age 14[22]. We also find negative and positive coefficient estimates, respectively, for the variables that identify the child as female or as Mestizo; however, these two coefficients are not statistically significant from zero.

Looking to the estimated coefficients of the parent-specific explanatory variables, we see statistically significant decreases in the estimated probabilities of child labor incidence that correspond with higher levels of parents' educational attainment, the age of the household head, and when the household head is female. These findings fit within the results found throughout the literature as a whole. Parents with higher educational attainment are more likely to value childhood education over child labor. A positive, yet statistically insignificant, coefficient estimate is reported education over child labor. A positive, yet statistically insignificant, coefficient estimate is reported for the variable that identifies single-parent households. The values presented in column (b), where we substitute a dummy variable that identifies whether any shocks have been experienced since the prior survey for the variable that represents the number of shocks experienced, are largely consistent with the results presented in column (a). However,

21 The 1.26 percentage point increase represents a 5.46% increase in the estimated probability of child labor activity.

22 All probability estimates are generated by allowing the variable(s) of interest to vary while holding all other variables constant at their mean values.

the coefficient of the variable which represents parent's educational attainment is no longer statistically significant from zero and the coefficients of the variables that identify single-parent households and households with female heads differ in sign relative to the coefficients reported in column (a).

Table 3.3: Household Shocks and Child Labor Incidence

Dependent Variable:	Any Labor	Any Labor
	(a)	(b)
Shocks (number)	0.0699***	
	(0.0247)	
Any shock (dummy)		0.2358***
		(0.0776)
Child's age	-2.1972***	-2.1848***
	(0.1521)	(0.1519)
Child's age squared	0.1403***	0.1396***
	(0.0085)	(0.0085)
Female child	-0.0686	-0.0667
	(0.0719)	(0.0719)
Mestizo child	0.0926	0.0895
	(0.1436)	(0.1436)
Number of siblings	0.2931***	-0.0279**
	(0.0498)	(0.0115)
Parents' highest educational attainment	-0.0277**	0.2315
	(0.0115)	(0.1713)
Single parent (female or male)	0.2502 (p=0.141)	-0.0831***
	(0.1697)	(0.0228)
Household head's age	-0.0819***	0.0007***
	(0.0229)	(0.0002)
Household head's age squared	0.0007***	-0.5028**
	(0.0002)	(0.1982)
Household size	-0.0713*	-0.0728*
	(0.0378)	(0.0379)
Household quality index	0.8841***	0.8933***
	(0.2733)	(0.2735)
Household owns any animals	-0.0350	-0.0299
	(0.0933)	(0.093)
Household wealth index: Second 20%	0.8188***	0.8272***
	(0.2037)	(0.2037)

Continued

Table 3.3 *Continued*

Dependent Variable:	Any Labor	Any Labor
Household wealth index: Middle 20%	0.4844***	0.4831***
	(0.1678)	(0.1678)
Household wealth index: Fourth 20%	0.3705***	0.367***
	(0.1319)	(0.1318)
Rural household	0.1091	0.1110
	(0.1061)	(0.1059)
Selva region	-0.1581	-0.1629
	(0.1158)	(0.1157)
Sierra region	0.1934**	0.2027**
	(0.0888)	(0.0887)
Constant	7.2202***	7.1341***
	(0.8685)	(0.8673)
Log pseudolikelihood	-2,387	-2,386
Count R-squared	0.767	0.767

N=4,934. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Finally, when considering the influences of household-specific variables, we see the probability of child labor incidence decreases significantly with increases in the total number of household members. These results reaffirm the findings on household size of Bandara et al. (2015). The estimated probability that a child will engage in labor activity decreases significantly and consistently with the level of household wealth. As discussed in Basu and Van (1998), household income and the incidence of child labor are negatively related. This was noted as the luxury axiom. Indicative of geographic variation in child labor incidence, children who live in the Sierra region are significantly more likely, relative to those who live in the Costa and Selva regions, to engage in labor activity. Finally, although not significantly different from zero, coefficient estimates reveal that children who live in rural households are more likely to engage in child labor activity, that higher levels of household quality correspond with an increased likelihood of engagement in child labor, and that children who live in households that own animals are less likely to engage in child labor. Our second research question asks: *Do the influences of household shocks on child labor incidence vary significantly based on the number of shocks experienced?* To answer this question, we estimate four variants of equation (2) using the binomial logit technique. The results are provided in Table 3.4. The estimation equations vary only by choice of dependent variables. Column (a) presents coefficients obtained when the dependent variable indicates whether the child is engaged in any form of labor activity (i.e., paid labor, home care, and/or household tasks). Columns (b), (c), and (d) present results obtained when the dependent variable series indicates engagement in specific types of labor activity—paid labor, home care, or household tasks, respectively. As noted, the binomial logit technique is employed; thus, each dependent variable is a dummy variable equal to one if the child is reported to be engaged in the corresponding labor activity and is equal to zero otherwise.

The explanatory variables in each of the estimations include five household shock-related dummy variables that identify the number of shocks the child's household has experienced since their last survey (i.e., one shock, two shocks, etc., to five or more shocks). The null/excluded category is zero shocks; thus, the coefficient estimates are interpreted relative to the scenario where no shock is experienced. By specifying our model in this fashion, we can estimate the probabilities of child labor incidence given various numbers of household shocks having been experienced and we can perform post-estimation Wald tests to determine whether the differences in corresponding coefficient estimates are statistically significant from zero.

Table 3.4: Number of Household Shocks and Specific Forms of Child Labor Activity

	Any Labor	Paid Labor	Home Care	Household Tasks
Dependent Variable:	(a)	(b)	(c)	(d)
One shock	0.1261	0.1042	0.0816	0.1345
	(0.0912)	(0.2794)	(0.0952)	(0.1371)
Two shocks	0.3519***	-0.1647	0.3911***	0.2922**
	(0.1051)	(0.3462)	(0.1083)	(0.1474)
Three shocks	0.3649***	-0.3551	0.2813**	0.4494***
	(0.1322)	(0.4256)	(0.1375)	(0.173)
Four shocks	0.1749	-0.0275	0.2151	0.0138
	(0.1822)	(0.568)	(0.1956)	(0.2442)
Five or more shocks	0.3564*	0.0151	0.1036	0.5331**
	(0.2007)	(0.5433)	(0.22)	(0.2531)
Child's age	-2.1973***	2.0752 (p=0.148)	-2.0781***	-1.1257***
	(0.1522)	(1.4358)	(0.1575)	(0.2629)
Child's age squared	0.1403***	-0.0538	0.1321***	0.0868***
	(0.0086)	(0.0692)	(0.0088)	(0.0142)
Female child	-0.0701	-0.2533	0.0397	-0.2469**
	(0.072)	(0.2339)	(0.075)	(0.1025)

Continued

Table 3.4 *Continued*

Dependent Variable:	Any Labor (a)	Paid Labor (b)	Home Care (c)	Household Tasks (d)
Mestizo child	0.0856	-0.0894	0.0519	0.1679
	(0.1437)	(0.4874)	(0.1496)	(0.2098)
Number of siblings	0.2945***	-0.1975	0.4099***	0.0525
	(0.0499)	(0.138)	(0.0537)	(0.0681)
Parents' highest educational	-0.0279**	-0.0672**	-0.0099	-0.0382**
attainment	(0.0116)	(0.0344)	(0.0122)	(0.0159)
Single parent (female or male)	0.2423	-1.0215**	0.2520	-0.2621
	(0.1714)	(0.4624)	(0.1841)	(0.2672)
Household head's age	-0.0826***	-0.1786***	-0.1087***	0.0437
	(0.0231)	(0.0625)	(0.0264)	(0.039)
Household head's age squared	0.0007***	0.0017***	0.0008***	-0.0005
	(0.0003)	(0.0006)	(0.0003)	(0.0004)
Female household head	-0.5121***	1.0323**	-0.4495**	-0.3007
	(0.1979)	(0.4556)	(0.2104)	(0.2961)
Household size	-0.0742*	0.2304**	-0.0631$^{(p=0.122)}$	-0.0404
	(0.038)	(0.099)	(0.0407)	(0.0531)
Household quality index	0.8984***	2.7103***	0.7725***	0.8137**
	(0.2736)	(0.866)	(0.2889)	(0.393)
Household owns any animals	-0.0363	-0.0256	-0.194**	0.5558***
	(0.0934)	(0.3284)	(0.0973)	(0.1465)
Household wealth	1.1755***	3.9452***	1.0376***	1.4023***
index: Lowest 20%	(0.2311)	(0.818)	(0.2415)	(0.3206)
Household wealth	0.821***	2.7618***	0.6804***	0.8917***
index: Second 20%	(0.2038)	(0.7508)	(0.2116)	(0.2901)
Household wealth	0.4852***	1.7454***	0.4321**	0.3509
index: Middle 20%	(0.1678)	(0.6628)	(0.1745)	(0.2518)
Household wealth	0.3753***	1.1215**	0.3165**	0.2863
index: Fourth 20%	(0.132)	(0.5226)	(0.1356)	(0.2152)
Rural household	0.1026	-0.8223***	-0.0302	0.5291***
	(0.1064)	(0.3036)	(0.1125)	(0.1348)
Selva region	-0.1610	-0.2926	-0.3109**	0.3245*
	(0.1159)	(0.3925)	(0.1225)	(0.1779)
Sierra region	0.1992**	0.3102	0.0772	0.7108***
	(0.0888)	(0.2869)	(0.0921)	(0.1366)

Table 3.4 *Continued*

Dependent Variable:	Any Labor (a)	Paid Labor (b)	Home Care (c)	Household Tasks (d)
Constant	7.1879***	-19.2853**	7.234***	-2.8763*
	(0.8725)	(7.7823)	(0.9418)	(1.4828)
Log pseudolikelihood	-2,383	-334	-2,238	-1,320
Count R-squared	0.766	0.981	0.779	0.889

See Table 3.3 notes.

Beginning with the results that are presented in column (a), we find that all coefficient estimates are positive; however, only the estimated coefficients that correspond with having experienced two shocks, three shocks, and five or more shocks are statistically significant from zero. The estimated probability that a child whose household did not experience any shocks since their most recent survey is engaged in child labor activity is equal to 20.48%. Experiencing only one shock since the last survey, all else constant, increases this probability to 22.60%, an increase of 2.12 percentage points. If the child's household experienced two shocks since their most recent survey is engaged in child labor activity is equal to 20.48%. Experiencing only one shock since the last survey, all else constant, increases this probability to 22.60%, an increase of 2.12 percentage points. If the child's household experienced two shocks since the most recent survey, the estimated probability of child labor incidence increases to 26.8%, a 6.32 percentage point increase relative to the scenario in which no shocks are experienced. Similarly, if three shocks are experienced, the estimated probability of child labor incidence rises to 27.05%. Evaluating the influences of having experienced four household shocks or five or more shocks results in estimates of child labor incidence equal to 23.47% and 26.89%, respectively. It should be noted that while the difference between the "one shock" and "two shocks" coefficient estimates is significantly different from zero (i.e., Wald test statistic=4.47, $p < 0.05$), the differences between the "two shocks" and "three shocks" coefficient estimates, the "three shocks" and "four shocks" coefficient estimates, and the "four shocks" and "five or more shocks" are not significantly different from zero.

Table 3.5: Number of Specific Household Shocks and Specific Forms of Child Labor Activity

	Any Labor	Paid Labor	Home Care	Household Tasks
Dependent Variable:	(a)	(b)	(c)	(d)
Crime shocks	-0.0296	-0.1515	-0.0514	-0.0320
	(0.0846)	(0.2482)	(0.0878)	(0.1177)
Economic shocks	0.1831**	0.2132	0.144*	0.1906*
	(0.0754)	(0.2537)	(0.0794)	(0.0977)
Environmental shocks	0.0284	-0.1591	0.0328	0.0701
	(0.0475)	(0.157)	(0.0505)	(0.0566)
Family shocks	0.1045*	0.0690	0.1103*	-0.0749
	(0.0567)	(0.1655)	(0.0591)	(0.0838)
Housing shocks	0.2095	0.3782	0.0274	0.4977***
	(0.1663)	(0.3116)	(0.187)	(0.1762)
Other shocks	0.0082	-0.2074	-0.1069	0.2313
	(0.1387)	(0.47)	(0.1454)	(0.1952)
Regulatory shocks	-0.0134	0.4240	-0.2198	0.6939$^{(p=0.145)}$
	(0.384)	(1.016)	(0.415)	(0.4762)
Child's age	-2.1802***	1.9471	-2.0616***	-1.1222***
	(0.1533)	(1.3788)	(0.1587)	(0.2675)
Child's age squared	0.1395***	-0.0480	0.1312***	0.087***
	(0.0086)	(0.067)	(0.0089)	(0.0145)
Female child	-0.0677	-0.2430	0.0443	-0.2555**
	(0.0719)	(0.2345)	(0.0749)	(0.1026)
Mestizo child	0.0935	-0.0825	0.0510	0.1933
	(0.1444)	(0.4832)	(0.1495)	(0.2092)
Number of siblings	0.2932***	-0.1952	0.4082***	0.0550
	(0.0499)	(0.1409)	(0.0535)	(0.0685)
Parents' highest educational	-0.0281**	-0.0665*	-0.0094	-0.0386**
attainment	(0.0116)	(0.035)	(0.0122)	(0.016)
Mestizo child	0.0935	-0.0825	0.0510	0.1933
	(0.1444)	(0.4832)	(0.1495)	(0.2092)
Number of siblings	0.2932***	-0.1952	0.4082***	0.0550
	(0.0499)	(0.1409)	(0.0535)	(0.0685)
Parents' highest educational	-0.0281**	-0.0665*	-0.0094	-0.0386**
attainment	(0.0116)	(0.035)	(0.0122)	(0.016)

Table 3.5 *Continued*

Dependent Variable:	Any Labor (a)	Paid Labor (b)	Home Care (c)	Household Tasks (d)
Single parent (female or	0.2409	-1.0246**	0.2413	-0.2027
male)	(0.1709)	(0.4591)	(0.1836)	(0.2658)
Household head's age	-0.0819***	-0.1756***	-0.1068***	0.0385
	(0.0228)	(0.0639)	(0.0261)	(0.0392)
Household head's age	0.0007***	0.0017***	0.0008***	-0.0004
squared	(0.0002)	(0.0006)	(0.0003)	(0.0004)
Female household head	-0.5144***	0.9805**	-0.4456**	-0.2749
	(0.1978)	(0.4568)	(0.2104)	(0.2943)
Household size	-0.0734*	0.2204**	-0.0628(p=0.122)	-0.0365
	(0.038)	(0.1014)	(0.0406)	(0.0533)
Household quality index	0.8824***	2.6633***	0.7489***	0.7851**
	(0.2735)	(0.8568)	(0.2887)	(0.3933)
Household owns any animals	-0.0298	-0.0191	-0.1886*	0.5651***
	(0.0936)	(0.3191)	(0.0972)	(0.1475)
Household wealth	1.167***	3.9148***	1.0275***	1.3514***
index: Lowest 20%	(0.2317)	(0.8103)	(0.2419)	(0.321)
Household wealth	0.8283***	2.7503***	0.6864***	0.8702***
index: Second 20%	(0.204)	(0.7404)	(0.2113)	(0.2906)
Household wealth	0.4848***	1.74***	0.4327**	0.3160
index: Middle 20%	(0.1678)	(0.6477)	(0.1745)	(0.2509)
Household wealth	0.3719***	1.1359**	0.3129**	0.2635
index: Fourth 20%	(0.1322)	(0.5204)	(0.1356)	(0.2143)
Rural household	0.1184	-0.7894**	-0.0218	0.5372***
	(0.1071)	(0.3099)	(0.1132)	(0.1361)
Selva region	-0.1518	-0.2601	-0.3**	0.3281*
	(0.1157)	(0.3953)	(0.1224)	(0.1781)
Sierra region	0.2094**	0.3422	0.0881	0.6866***
	(0.0898)	(0.2938)	(0.0932)	(0.1383)
Constant	7.1343***	-18.671**	7.1772***	-2.7587*
	(0.8716)	(7.4801)	(0.9407)	(1.4999)
Log pseudolikelihood	-2,384	-333	-2,241	-1,317
Count R-squared	0.768	0.982	0.777	0.890

See Table 3.3 notes.

Turning attention to columns (b) through (d) reveals several interesting results. Results presented in column (b) indicate that regardless of the number of shocks experienced there is no statistically significant relationship between household shocks and child engagement in paid labor. This may appear unusual, especially if paid labor is the form of labor that comes to mind when one thinks of child labor. While we cannot provide a definitive explanation for this result, we can offer a plausible explanation. In Table 3.2, we see that the mean value for the paid labor dummy variable is 0.0184; thus, only 91 of the 4,934 observations in our data set reported being engaged in paid labor. The corresponding low variation in the dependent variable series when potential determinants of paid labor are considered severely reduces the likelihood of observing statistically significant coefficient estimates. It is also noteworthy that the Peruvian government amended the *Child and Adolescent's Code of 2000* to increase the minimum working age in Peru from 12 years of age to 14 years (US DOL, 2007). Consequently, as we consider children ages 5 to 14 in this study and our reference period spans the years 2002–2017, it may be that paid labor is infrequent among children below the age of 15 and/or that survey respondents were reluctant to admit to violating the minimum working age law.

We also see that the pattern of coefficient magnitude and statistical significance of the coefficients of the household shock variables in columns (c) and (d), where the provision of home care services and the performance of household tasks, respectively, are the specific forms of child labor considered, largely mirror the results that are reported in column (a). Specifically, in both columns (c) and (d), we observe positive and statistically significant coefficient estimates for the "two shocks" and "three shocks" categories. In column (d), we also find a positive and statistically significant coefficient for the variable that identifies having experienced five or more shocks. Looking first to the results reported in column (c), we see a statistically significant difference in the magnitude of the coefficient representing instance where two household shocks were experienced as compared to when only one shock was experienced (i.e., Wald test statistic=7.99 with $p < 0.01$). However, there is no statistically significant change in coefficient magnitudes of the variables that identify cases where three shocks, four shocks, and five or more shocks are experienced. We find no statistically significant differences in coefficient magnitudes across the coefficients that are reported in column (d).

Thus far, we have identified a positive relationship between household shocks and child labor incidence. Further, we have reported variation in this relationship based on the number of household shocks experienced. Our third research question addresses a similar notion: *Do the influences of household shocks on child labor incidence vary significantly based on the type of shock experienced?* To address this question, we estimate an additional variant of equation (2). In this instance, however, we decompose our measure that represents the number of household shocks experienced to include seven variables that separately represent the numbers of specific types of household shocks. The results are presented in Table 3.5.

Beginning with column (a), we report statistically significant coefficient estimates for the economic shocks and family shocks variables. We see that all else held constant, increases in either of these variables correspond with a statistically significantly increase the estimated probability that a child will begin to engage in one or more forms of labor activity. Specifically, evaluating the results at the mean values of the explanatory variable series, we find that experiencing an economic shock or a family shock is estimated to increase the probability of child labor incidence by 3.34 percentage points and 1.86 percentage points, respectively. Results presented in column (b) indicate no statistically significant relationship between household shocks and paid labor activity. Similar to the results presented in column (a), in column (c) we find that economic shocks significantly increase the estimated probability that a child will engage in the provision of home care by 2.23 percentage points while family shocks are estimated to increase the estimated probability by 1.67 percentage points. Finally, in column (d) we report that economic shocks and housing shocks are positively related to the estimated probability that a child will begin to engage in household tasks with the former shock increasing the estimated probability by 1 percentage point and the latter increasing the probability by 3.09 percentage points. While our results suggest that different types of shocks correspond with engagement in different labor activities, we also note that for each form of labor considered, post-estimation Wald tests reveal there are no statistically significant differences between coefficient estimates that are both positive and statistically significant from zero.

Finally, the earlier discussion of the estimated coefficients of the non-household shock variables that are presented in Table 3.3 and the following summary of estimated coefficients of such variables addresses our fourth research question: Regardless of whether a household experiences a shock, does the probability of child labor incidence vary based on measurable child-, parent-, and household-specific characteristics? Consistently, we find statistically significant relationships between these variables and the estimated probabilities of child labor activity. Across Tables 3.3, 3.4, and 3.5, we find 68% of the estimated coefficients of the variables unrelated to household shocks are statistically significant from zero (i.e., 136 of 200 cases). We find generally similar frequencies of statistical significance among the coefficient estimates that correspond with child-, parent-, and household-specific explanatory variables. Specifically, 24 of the 50 coefficients (48%) of the child-specific variables are statistically significant from zero. Likewise, 35 of the 50 coefficients (70%) that correspond with the parent-specific variables are statistically significant, and 70 of the 100 coefficients (70%) of the household-specific variables are statistically significant.

5. SUMMARY AND CONCLUSIONS

We examine data from longitudinal household surveys that were conducted in Peru during the period from 2002 through 2017. Estimation of a series of binomial logit specifications leads to the identification of the potential influences of household shocks on the probability that the typical child (5–14 years of age) engages in labor activity. We restrict our data to only include observations that represent children who were not working at the time of their previous survey to facilitate the examination of the influence of household shocks on engagement in child labor activity. Further, we control for the influences of child-, parent-, and household-specific attributes. In addition to documenting the relationship between household shocks and the probability of child labor incidence, we also determine whether the impacts of household shocks on child labor vary based on the number of shocks experienced, and we consider whether the relationship between household shocks and child labor differ based on the type of shock(s) that the household experiences.

We report positive relationships between household shocks and child labor incidence. Namely, the predicted probability that an otherwise typical child will begin to engage in any form of labor activity is estimated to increase with the number of household shocks experienced. This is documented both when we employ the number of household shocks experienced since the prior survey date as an explanatory variable and when we employ a binary measure of household shocks (i.e., no shocks or one or more shocks). We also report statistically significant increases, of varying magnitudes, in the estimated likelihood of child labor activity in response to the number of household shocks experienced. Decomposing our measure of household shocks into seven separate types, we report statistically significant coefficient estimates for the economic shocks, family shocks, and housing shocks variables. We also report statistically significant effects between specific forms of child labor activity (i.e., paid labor, home care, and household tasks) and the incidence of household shocks, the number of household shocks experienced (i.e., one shock, two shocks, etc., up to five or more shocks), and the type(s) of household shock(s) experienced (e.g., economic shocks, family shocks, housing shocks, etc.).

Our hope is that the information provided here aids in the identification and understanding of the many factors that contribute to parents' decisions to send their children to work and, by doing so, informs public policy such that household reliance on children's earnings is reduced and child engagement in non-developmental labor activity is diminished or, ideally, eliminated.

6. APPENDIX

Survey Question	Type of Shock	Corresponding Shocks
Has the household been the victim of any crime since previous round?	Crime shocks	Destruction/theft of tools for production; Theft of cash; Theft of crops; Theft of livestock; Theft/destruction of housing/consumer goods; and Crime that resulted in death/disablement.
Have any regulations or actions had negative impact on the household since previous round?	Regulatory shocks	Land redistribution; Resettlement of forced migration; Forced contributions; Eviction; and Invasion of property.
Have any changes to economic conditions affected household since previous round?	Economic shocks	Increase in input prices; Decrease in output prices; Death of livestock; Closure place of employment; Loss of job/source of income/family enterprise; Industrial action; Contract disputes (purchase of inputs); Contract disputes (sale of output); Disbanding credit; Confiscation of assets; Disputes with family about assets; and Disputes with neighbors about assets.
Have you experienced any natural disasters since previous round?	Environmental shocks	Drought; Flooding; Erosion; Frost; Pests on crops; Crop failure; Pests on storage; Pests on livestock; Earthquake; Forest fire; Pollution caused by mining; and Storm.
Has anything happened since previous round that has affected the building you live in?	Housing shocks	Fire affecting house; House collapse; and Fire or collapse of building.

Continued

Survey Question	Type of Shock	Corresponding Shocks
Have there been any changes within the family since previous round?	Family shocks	Death of father; Death of mother; Death of other household member; Illness of father; Illness of mother; Illness of other household member; Divorce or separation; Birth of new household member; Enrollment of child in school; Imprisonment; Conscription, abduction or draft; Political imprisonment; Political discrimination; Ethnic/social discrimination; and Illness of non-household member.
Has anything else happened since previous round that has affected the economic situation of the household?	Other shocks	.

REFERENCES

Acheampong, K. O., and Huang, W. 2018. "Household Shocks and Child Labor: Evidence from a Panel Survey of Nigerian Households" Working paper.

Azubuike, O. B., and Briones, K. 2016. Technical Note 35: Young Lives Rounds 1 to 4 Constructed Files. Young Lives.

Baland, J. M., and Robinson, J. A. 2000. "Is Child Labor Inefficient?" *Journal of Political Economy* 108 (4): 663–679.

Bandara, A., Dehejia, R., and Lavie-Rouse, S. 2015. "The Impact of Income and Non-Income Shocks on Child Labor: Evidence from a Panel Survey of Tanzania." *World Development* 67: 218–237.

Basu, K. 1999. "Child Labor: Cause, Consequence, and Cure, with Remarks on International Labor Standards." *Journal of Economic Literature* 37 (3): 1083–1119.

Basu, K., Das, S., and Dutta, B. 2010. "Child Labor and Household Wealth: Theory and Empirical Evidence of an Inverted-U." *Journal of Development Economics* 91 (1): 8–14.

Basu, K., and Van, P. H. 1998. "The Economics of Child Labor." *American Economic Review* 88 (3): 412–427.

Beegle, K., Dehejia, R. H., and Gatti, R. 2006. "Child Labor and Agricultural Shocks." *Journal of Development Economics* 81 (1): 80–96.

Bhalotra, S. 2007. "Is Child Work Necessary?" *Oxford Bulletin of Economics and Statistics* 69 (1): 29–55.

Bhalotra, S., and Heady, C. 2003. "Child Farm Labor: The Wealth Paradox." *The World Bank Economic Review* 17 (2): 197–227.

Bhalotra, S. R., and Tzannatos, Z. 2003. *Child Labor: What Have We Learnt?* Social Protection Discussion Paper Series; No. SP 0317. Washington, D.C.: World Bank.

Bharadwaj, P. 2015. "Fertility and Rural Labor Market Inefficiencies: Evidence from India." *Journal of Development Economics* 115 (C): 217–232.

Bornstein, M. H., and Putnick, D. L. 2015. "Is Child Labor a Barrier to School Enrollment in Low-and Middle-income Countries?" *International Journal of Educational Development* 41: 112–120.

Bourguignon, F., and Chiappori, P. A. 1994. "The Collective Approach to Household Behavior." In *The Measurement of Household Welfare,* edited by R. Blundell, I. Preston, and I. Walker. Cambridge, UK: Cambridge University Press.

Boyden, J., Woldehanna, T., Galab, S., Sanchez, A., Penny, M., Duc, L. T. (2016). Young Lives: an International Study of Childhood Poverty: Round 4, 2013–2014. [data collection]. UK Data Service. SN: 7931, http://dx.doi.org/10.5255/UKDA-SN-7931-1.

Dendir, Seife. 2007. "Household Shocks and Child Labor in Rural Malawi." *Economics Bulletin* 4 (1): 1–8.

Doepke, M., and Zilibotti, F. 2005. "The Macroeconomics of Child Labor Regulation." *American Economic Review* 95 (5): 1492–1524.

Dumas, C. 2007. "Why Do Parents Make Their Children Work? A Test of the Poverty Hypothesis in Rural Areas of Burkina Faso." *Oxford Economic Papers* 59 (2): 301–329.

Emerson, P. M., and Knabb, S. 2006. "Opportunity, Inequality and the Intergenerational Transmission of Child Labor." *Economica* 73 (291): 413–434.

Emerson, P. M., and Souza, A. P. 2007. "Child Labor, School Attendance and Intra-Household Gender Bias in Brazil." *The World Bank Economic Review* 21 (2): 301–316.

Ersado, L. 2005. "Child Labor and Schooling Decisions in Urban and Rural Areas in Nepal." *World Development* 33 (3): 455–458.

Fan, C. S. 2004a. "Relative Wage, Child Labor and Human Capital." *Oxford Economic Papers* 56 (4): 687–700.

Fan, C. S. 2004b. "Child Labor and the Interaction between the Quantity and Quality of Children." *Southern Economic Journal* 71 (1): 21–35.

Fors, H. C. 2012. "Child Labour: A Review of Recent Theory and Evidence with Policy Implications." *Journal of Economic Surveys* 26 (4): 570–593.

Grootaert, C., and Kanbur, R. 1995. "Child Labor, A Review" Policy Research Working Paper No. WPS 1454. Washington, D.C.: World Bank.

Guarcello, L., Mealli, F., and Rosati, F. C. 2010. "Household Vulnerability and Child Labor: The Effect of Shocks, Credit Rationing and Insurance." *Journal of Population Economics* 23 (1): 169–198.

Hazan, M., and Berdugo, B. 2002. "Child Labor, Fertility, and Economic Growth." *Economic Journal* 112 (482): 810–828.

Heady, C. 2003. "The Effect of Child Labor on Learning Achievement." *World Development* 31 (2): 385–398.

Kambhampati, U. S., and Rajan, R. 2006. "Economic Growth: A Panacea for Child Labor?" *World Development* 34 (3): 426–445.

Kruger, D. I. 2007. "Coffee Production Effects on Child Labor and Schooling in Rural Brazil." *Journal of Development Economics* 82 (2): 448–463.

Kurosaki, T., Ito, S., Fuwa, N., Kubo, K., and Sawada, Y. 2006. "Child Labor and School Enrollment in Rural India: Whose Education Matters?" *Developing Economies* 44 (4): 440–464.

Moehling, C. M. 1995. *The Intrahousehold Allocation of Resources and the Participation of Children in Household Decision-Making: Evidence from Early Twentieth Century America*. Mimeo. Evanston, IL: Northwestern University.

Ranjan, P. 2001. "Credit Constraint and the Phenomenon of Child Labor." *Journal of Development Economics* 64: 81–102.

Ray, R. 2002. "The Determinants of Child Labor and Child Schooling in Ghana." *Journal of African Economies* 11 (4): 561–590.

Rosati, F. C., and Rossi, M. 2003. "Children's Working Hours and School Enrollment: Evidence from Pakistan and Nicaragua." *The World Bank Economic Review* 17 (2): 283–295.

Saddik, B., Williamson, A., Nuwayhid, I., and Black, D. 2005. "The Effects of Solvent Exposure on Memory and Motor Dexterity n Working Children." *Public Health Reports* 120 (6): 657–663.

Soares, R., Kurger, D., and Barthelon, M. 2012. "Household Choices of Child Labor and Schooling, A Simple Model with Application in Brazil." *Journal of Human Resources* 47 (4): 1–31.

Strauss, J., and Thomas, D. 1995. "Human Resources: Empirical Modeling of Household and Family Decisions." In *Handbook of Development Economics,* edited by J. Behrman and T. Srinivasan, vol. 3, 1883–2023. Amsterdam, Netherlands: North-Holland.

US Department of Labor, Bureau of International Labor Affairs (US DOL). 2007. *Labor Rights Report: Peru*. Washington, D.C.: US Department of Labor.

Vasquez, W. F., and Bohara, A. K. 2010. "Household Shocks, Child Labor, and Child Schooling." *Latin American Research Review* 45 (3): 165–186.

Contributors

Alma A. Bezares Calderón is an Assistant Professor of Economics at Whittier College. She obtained her Ph.D. in Economics and Politics from Claremont Graduate University. Her research focuses on the political economy of development of low- and middle-income countries, particularly in Latin America and Sub-Saharan Africa.

Feridoon Koohi-Kamali studied economics at University of London and received his doctorate in economics from Oxford University. He has taught at Oxford and a number of US universities including Columbia. Currently Associate Professor at The New School for Social Research, he has also been a consultant to the World Bank and research associate at the Schuman Center for Advanced Studies, European University Institute. His research interests are in applied econometrics, public policy and development.

Forrest Rouleau is a graduate of Whittier College where he earned his Bachelor of Arts degree in Economics.

Roger White is a Professor of Economics and the Douglas Ferguson Chair of International Economics at Whittier College. He is the author or editor of eight books and has authored more than forty published research articles and book chapters. His research largely focuses on topics relating to cross-societal cultural differences/distance, the immigrant-trade link, immigration policy/history, public opinion towards economic globalization, and the consequences of increased international trade for domestic labor.

Index

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

R

PETER LANG
PROMPT

Peter Lang Prompts offer our authors the opportunity to publish original research in small volumes that are shorter and more affordable than traditional academic monographs. With a faster production time, this concise model gives scholars the chance to publish time-sensitive research, open a forum for debate, and make an impact more quickly. Like all Peter Lang publications, Prompts are thoroughly peer reviewed and can even be included in series.

For further information, please contact:

editorial@peterlang.com

To order, please contact our Customer Service Department:

peterlang@presswarehouse.com (within the U.S.)
orders@peterlang.com (outside the U.S.)

Visit our website: www.peterlang.com

Prompts include:

Claudia Aburto Guzmán, *Poesía reciente de voces en diálogo con la ascendencia hispano-hablante en los Estados Unidos: Antología breve*. ISBN 978-1-4331-5207-8. 2020

William Robert Adamson, *Mine Own Familiar Friend: The Relationship between Gerard Hopkins and Robert Bridges*. ISBN 978-1-80079-485-6. 2021

Tywan Ajani, *Barriers to Rebuilding the African American Community: Understanding the Issues Facing Today's African Americans from a Social Work Perspective*. ISBN 978-1-4331-7681-4. 2020

Macarena Areco, *Bolaño Constelaciones: Literatura, sujetos, territories*. ISBN 978-1-4331-7575-6. 2020

Desrine Bogle. *The Transatlantic Culture Trade: Caribbean Creole Proverbs from Africa, Europe, and the Caribbean*. ISBN 978-1-4331-5723-3. 2020

Jean-François Caron. *Irresponsible Citizenship: The Cultural Roots of the Crisis of Authority in Times of Pandemic*. ISBN 978-1-4331-8908-1. 2021

Marcilio de Freitas and Marilene Corrêa da Silva Freitas, *The Future of Amazonia in Brazil: A Worldwide Tragedy*. ISBN 978-1-4331-7793-4. 2020

Mihai Dragnea. *Christian Identity Formation Across the Elbe in the Tenth and Eleventh Centuries*. Christianity and Conversion in Scandinavia and the Baltic Region, c. 800–1600, vol. 1. ISBN 978-1-4331-8431-4. 2021

Janet Farrell Leontiou, *The Doctor Still Knows Best: How Medical Culture Is Still Marked by Paternalism*. Health Communication, vol. 15. ISBN 978-1-4331-7322-6. 2020

Clare Gorman (ed.), *Miss-representation: Women, Literature, Sex and Culture*. ISBN 978-1-78874-586-4. 2020

Eva Marín Hlynsdóttir. *Gender in Organizations: The Icelandic Female Council Manager*. ISBN 978-1-4331-7729-3. 2020

Micol Kates, *Towards a Vegan-Based Ethic: Dismantling Neo-Colonial Hierarchy Through an Ethic of Lovingkindness*. ISBN 978-1-4331-7797-2. 2020

Matt Qvortrup, *Winners and Losers: Which Countries are Successful and Why?*. ISBN 978-1-80079-405-4. 2021

Peter Raina, *Doris Lessing – A Life Behind the Scenes: The Files of the British Intelligence Service MI5*. ISBN 978-1-80079-183-1. 2021

Peter Raina (trans.), *Heinrich von Kleist Poems*. ISBN 978-1-80079-043-8. 2020

Josiane Ranguin, *Mediating the Windrush Children: Caryl Phillips and Horace Ové*. ISBN 978-1-4331-7424-7. 2020

Dylan Scudder, *Coffee and Conflict in Colombia: Part of the Pentalemma Series on Managing Global Dilemmas*. ISBN 978-1-4331-7568-8. 2020

Dylan Scudder, *Conflict Minerals in the Democratic Republic of Congo: Part of the Pentalemma Series on Managing Global Dilemmas*. ISBN 978-1-4331-7561-9. 2020

Dylan Scudder, *Mining Conflict in the Philippines: Part of the Pentalemma Series on Managing Global Dilemmas*. ISBN 978-1-4331-7632-6. 2020

Dylan Scudder, *Multi-Hazard Disaster in Japan: Part of the Pentalemma Series on Managing Global Dilemmas*. ISBN 978-1-4331-7530-5. 2020

Wesley A. Stroud, *Education for Liberation, Education for Dignity: The Story of St. Monica's School of Basic Learning for Women*. ISBN 978-1-4331-7911-2. 2021

Geanneti Tavares Salomon, *Fashion and Irony in «Dom Casmurro»*. ISBN 978-1-78997-972-5. 2021

Shai Tubali, *Cosmos and Camus: Science Fiction Film and the Absurd*. ISBN 978-1-78997-664-9. 2020

Angela Williams, *Hip Hop Harem: Women, Rap and Representation in the Middle East*. ISBN 978-1-4331-7295-3. 2020

Ivan Zhavoronkov (trans.), *The Socio-Cultural and Philosophical Origins of Science* by Anatoly Nazirov. ISBN 978-1-4331-7228-1. 2020