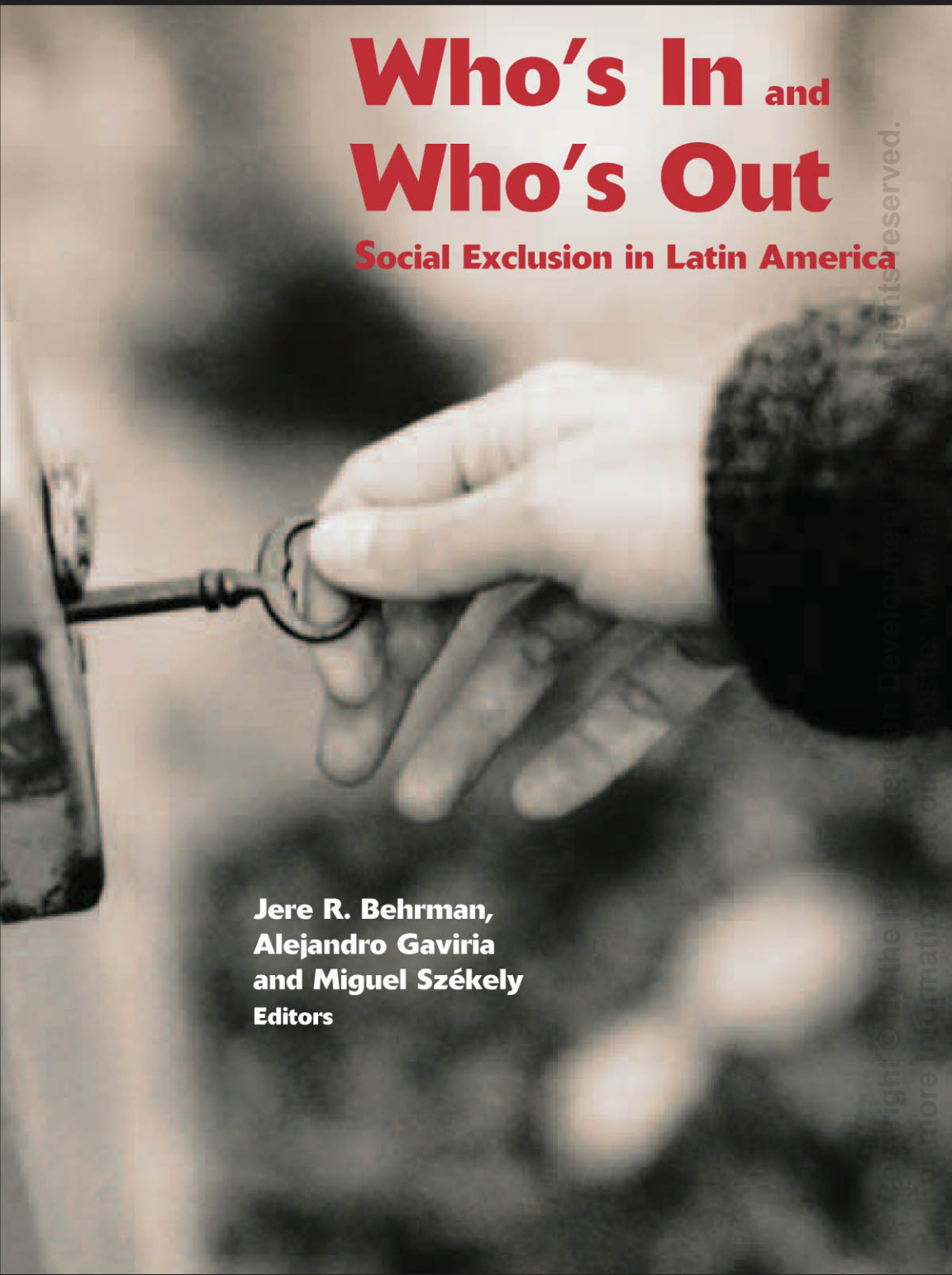
Inter-American Development Bank

Who's In and Who's Out

Social Exclusion in Latin America



Jere R. Behrman,
Alejandro Gaviria
and Miguel Székely
Editors

LATIN AMERICAN RESEARCH NETWORK

This page intentionally left blank

Who's In and Who's Out
Social Exclusion in Latin America

Jere R. Behrman,
Alejandro Gaviria
and Miguel Székely
Editors

Published by the Inter-American Development Bank
Distributed by the Johns Hopkins University Press

Washington, D.C.
2003

Produced by the IDB Publications Section

To order this book, contact:
IDB Bookstore
Tel: 1-877-PUBS IDB/(202) 623-1753
Fax: (202) 623-1709
E-mail: idb-books@iadb.org
www.iadb.org/pub

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Who's in and who's out : social exclusion in Latin America / Jere R. Behrman, Alejandro Gaviria, Miguel Székely, editors.

p. cm. Includes bibliographical references.

1. Marginality, Social—Latin America. 2. People with social disabilities—Latin America—Economic conditions. 3. Discrimination in housing—Bolivia. 4. Discrimination in medical care—Brazil. 5. Nicaraguans—Costa Rica—Social conditions. 6. Labor market—El Salvador. 7. Discrimination in education—Mexico. I. Behrman, Jere R. II. Gaviria Trujillo, Alejandro. III. Székely, Miguel. IV. Inter-American Development Bank.

ISBN: 1-931003-42-4 LCNN: 2002116830
305.56 W122—dc21

TABLE OF CONTENTS

ACKNOWLEDGMENTS

The authors are indebted to many people for their contributions at various stages in the preparation of this book. We would particularly like to thank Rodolfo Islas, Allan Pasalagua, Fabiana Tito and Ignez Tristao for their outstanding research assistance. Paul Schultz, Chris Udry and participants in the Yale University Development Lunch and the IDB seminars associated with this research initiative provided valuable comments and advice. Further thanks are due to Norelis Betancourt, Rita Funaro and Raquel Gómez for their efforts in coordinating the projects that led to this book and the production of the book itself. Finally, we would like to thank John Dunn Smith for his editorial efforts in helping to make the style and substance of this work accessible to the widest possible audience.

PREFACE

Social exclusion is closely linked with a wide range of economic problems in Latin America. Individuals who lack access to the opportunities enjoyed by others in health care, education and employment are prevented from reaching their full productive potential, which in turn holds down growth and public revenues. These same individuals are also more likely to incur public costs through health and social service expenses and, in some cases, through crime. Moreover, when people lack the material means to improve their circumstances, or the knowledge of how to do so, social exclusion can persist for generations.

These problems are further complicated by the subtle ways in which social exclusion takes place. Seldom if ever is there a "Keep Out" sign. Nor is the exclusion of certain groups or individuals, in general, explicitly intended. Instead, self-perpetuating networks of association and prohibitive housing prices often prevent some people from gaining access to all the benefits of society. In other cases, misperceptions among both mainstream society and members of excluded groups lead to self-fulfilling prophecies.

Who's In and Who's Out makes a unique contribution to the literature. Based on the results of an Inter-American Development Bank Research Network project, the book explores several forms of social exclusion, based on a variety of criteria, in five countries of the region. The studies show that in whatever forms social exclusion takes, it affects the overall well-being of both individuals and society as a whole.

The implications of these studies are vast. Identifying the causes, mechanisms and effects of social exclusion is essential to formulate policies that will enable the greatest number of people to lead productive lives and enjoy the economic and psychological benefits of fully participating in society. Without this background knowledge, even the best-intentioned and most lavishly funded initiatives can do little for marginalized and vulnerable sectors of the population.

Much work remains to be done in this field, particularly in Latin America, but *Who's In and Who's Out* represents a necessary first step in defining the terms of debate and setting the policy agenda in this insufficiently understood aspect of development.

Guillermo Calvo
Chief Economist
Inter-American Development Bank

This page intentionally left blank

ABOUT THE AUTHORS

Denisard Alves is a professor at the University of São Paulo and a researcher with the Fundação Instituto de Pesquisas Econômicas.

Jere R. Behrman is the William R. Kenan, Jr. Professor at the University of Pennsylvania.

Lissette Calderón Martínez is a research economist with the Fundación Salvadoreña para el Desarrollo Económico y Social (FUSADES).

Edward Funkhouser is a professor of economics at the University of California, Santa Barbara.

Alejandro Gaviria is assistant director of Fedesarrollo in Bogota, Colombia.

George Gray-Molina is a professor of economics at the Universidad Católica Boliviana in La Paz.

Wilson Jiménez is a researcher with the Fundación Diálogo in La Paz, Bolivia.

Anabella Lardé de Palomo is a research economist with the Fundación Salvadoreña para el Desarrollo Económico y Social (FUSADES).

Susan W. Parker is a research economist with the Centro de Investigación y Docencia Económicas in Mexico City.

Ernesto Pérez de Rada is a research economist with the Fundación Diálogo in La Paz, Bolivia.

Juan Pablo Pérez Sáinz is a researcher at the Facultad Latinoamericano de Ciencias Sociales (FLACSO) in Costa Rica.

Luis Rubalcava is a research economist with the Centro de Investigación y Docencia Económicas in Mexico City.

Carlos Sojo is director of the Facultad Latinoamericano de Ciencias Sociales (FLACSO) in Costa Rica.

Miguel Székely is assistant secretary for planning and evaluation at the Social Development Secretariat of the Office of the President of Mexico.

Graciela Teruel is a professor of economics at the Universidad Iberoamericana in Mexico City.

Christopher Timmins is a professor of economics at Yale University.

Ana Regina Vides de Andrade is a research economist with the Fundación Salvadoreña para el Desarrollo Económico y Social (FUSADES).

CHAPTER ONE

Social Exclusion in Latin America: Perception, Reality and Implications

Jere R. Behrman
Alejandro Gaviria
Miguel Székely

Income inequality is higher in Latin America and the Caribbean than in any other region of the world, and precisely because of this skewed distribution, the region's absolute poverty rates are much higher than one would predict on the basis of average income.

To what extent, however, is this inequality driven by individual differences in ability and work ethic rather than by differences in opportunities? If some individuals prefer to work more hours or to invest more energy in their work than others, the resulting income inequality would not necessarily be a policy issue. In fact, reducing this type of inequality through policy interventions could well lead to *reductions* rather than increases in welfare.

This type of "efficient" inequality, however, probably does not explain the extent of inequalities in Latin America. Rather, inequality in the region appears to arise largely from the absence of opportunities for large segments of the population. The outright (or implicit) exclusion of some groups on the basis of their gender, ethnic origin, place of residence or social status may in turn explain inequality of opportunity.

This book attempts to start to fill an important gap—the paucity of research on social exclusion—in the literature on the causes of poverty and inequality in Latin America. It aims not only to generate some evidence that may be useful for policy design, but also to illustrate the challenges posed by such a difficult topic. The focus is on particular forms of exclusion that are important for the determination of income—and thus poverty and income inequality—and that lend themselves to quantitative analysis. It is hoped that this research will help shed some light on the mechanisms of social exclusion and provide policymakers with guidance for corrective policies.

The collection of studies also illustrates the problems of finding adequate data, as well as how some of the main methodological problems may be addressed, and shows that it is possible to address this important issue rigorously.

Attitudes towards Social Exclusion in Latin America

The reality of social exclusion in Latin America has many faces and dimensions. Latin Americans themselves perceive it in different ways, and those perceptions are captured here by Latinobarómetro, a public opinion survey carried out annually in 17 Latin American countries since 1995.[1]

The 2000 Latinobarómetro asked respondents to name the groups most discriminated against in their countries. Answers could include groups according to ethnic background, nationality, class, gender, sexual orientation and political affiliation.[2] From the 25 groups listed, most answers clustered around three types of people: blacks, Indians, and the poor. In short, race and class are perceived as the basis for exclusion and discrimination in Latin America.

Perceptions of the groups most discriminated against, however, vary widely from country to country. In Brazil, for example, half of the respondents stated that blacks are the group most discriminated against; in Guatemala, almost 60 percent stated that Indians face the greatest discrimination; and in El Salvador, 70 percent said that the poor are the most discriminated against. Table 1.1 shows that Latin American countries can be classified into two groups according to their citizens' perceptions of discrimination. The first group includes all countries where most respondents singled out a racial group as the most discriminated against; the second includes the rest of

[1] Latinobarómetro interviews approximately 1,000 people in each country each year. The sampling method varies slightly from country to country, as implementation is contracted out to national polling firms. However, in most cases, selection procedures include some quotas to ensure representation across gender, socioeconomic status, and age. The survey is restricted to urban populations, and the emphasis is on political perceptions and attitudes. See IDB (2000).

[2] Respondents chose from the following menu of options: blacks, Indians, whites, mulattos, mestizos, Asians, Arabs, Jews, the poor, immigrants, homosexuals, AIDS victims, women, the elderly, and the illiterate, as well as people belonging to 10 other political and religious groups.

Table 1.1. Groups Most Discriminated Against
(In percent)

Country	Blacks	Indians	The poor
Panama	32.0	13.3	21.2
Mexico	3.8	46.9	25.5
Bolivia	1.5	46.9	26.5
Guatemala	2.3	58.7	26.9
Peru	22.1	26.6	28.7
Brazil	49.8	0.7	29.5
Ecuador	21.1	31.8	30.6
Uruguay	18.6	0.1	30.9
Costa Rica	4.4	11.0	31.5
Honduras	4.6	6.9	35.3
Chile	1.2	22.2	36.6
Colombia	17.3	11.3	39.8
Argentina	4.9	3.8	40.0
Paraguay	0.4	18.9	44.8
Venezuela	10.9	23.1	45.4
Nicaragua	5.4	3.1	60.3
El Salvador	0.5	0.5	69.1

Source: Latinobarómetro 1999–2000.

the countries. This classification makes it possible to separate countries where social divisions are based on race from those where social divisions are based on class.

The first group includes Bolivia, Brazil, Ecuador, Guatemala and Panama, all countries where a high percentage of the population is either Indian or black. The remainder of the countries of the region fall into the second group. Not surprisingly, the data suggest that countries where Indians or blacks represent a high percentage of the population are also countries where these groups are perceived as the most discriminated against. Hence, Indians are perceived as the most discriminated against in Bolivia and Guatemala, while black people have that distinction in Brazil and Panama. In Argentina and Uruguay, where the racial makeup is much more homogenous, the poor are perceived as being the most discriminated against.

Latinobarómetro also includes questions on the extent of discrimination against Indians and blacks. Specifically, people were asked to rank the extent of discrimination against these groups on a scale from 1 to 10, where 1 means the absence of discrimination and 10 means outright discrimination. Questions addressed the extent of discrimination at the workplace, at school, in political parties, and by the police and the courts.

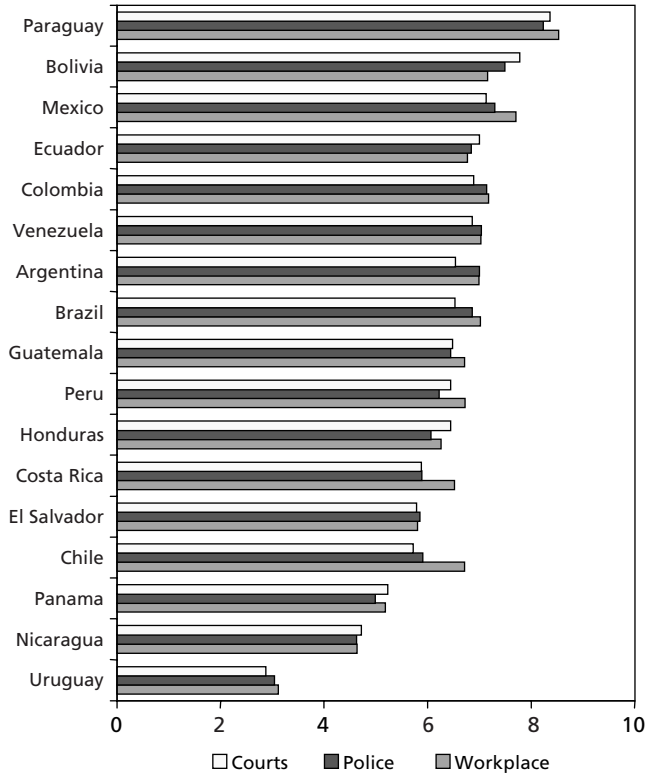
Individual responses show a high correlation among different forms of discrimination. Those who feel there is discrimination in the workplace also feel that there is discrimination at school, in political parties and by the police and the courts. In other words, few individuals appear to be able (or willing) to discern different degrees of discrimination in different venues and institutions. In addition, the data show that the mean of the responses on discrimination against Indians is 6.5 (with a large variance), and the mean against blacks is 6.0 (also with a large variance). In sum, the data indicate that while most Latin Americans do believe that there is discrimination against Indians and blacks, they hardly agree on the extent of the problem. The differences are large not only across countries, but also among citizens of the same country.

Figure 1.1 presents countries' mean perceptions of discrimination against Indians. These perceptions are very high in Paraguay, Bolivia and Mexico, and much lower in Panama, Nicaragua and Uruguay. Figure 1.2 does the same for discrimination against blacks. In this case, perceptions are the highest in Brazil, Ecuador and Peru, and the lowest in Paraguay, Nicaragua and Uruguay. Overall, the higher the extent of perceived discrimination against a racial group, the higher the share of that group in a country's population.

Perceptions of discrimination against Indians and blacks vary across racial groups in a predictable fashion. Figure 1.3 shows that Indians perceive more discrimination against themselves than do other racial groups. Likewise, blacks perceive greater levels of discrimination. In general, people of European descent report lower levels of perceived racial discrimination than either Indians or blacks.

The data also show that perceived discrimination against Indians and blacks is higher among the educated and the young, and, to a lesser extent, among women. Once differences in education and age are taken into account, perceptions of racial discrimination tend to be higher among middle-class individuals than among either the very poor or the very rich. This

Figure 1.1. Discrimination against Indians by Country
(Index)

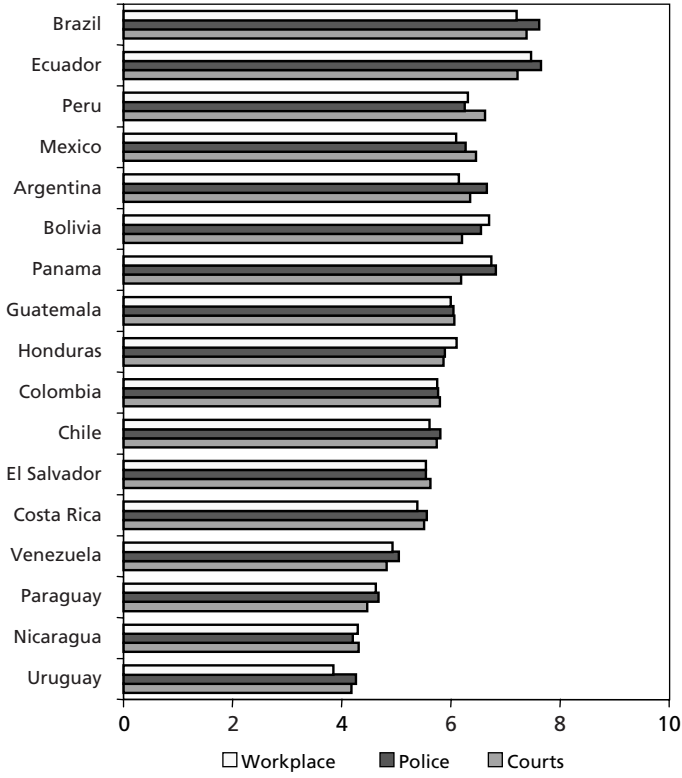


Source: Latinobarómetro 1999–2000.

suggests that the core support for policies against discrimination can be found in the middle classes.

In spite of widespread perceptions of racial discrimination, Latin Americans are not fully supportive of racially based social investments. When asked whether they would prefer a social policy aimed at improving the living conditions of the poor not targeted by race, or a similar policy aimed at improving the living conditions of Indian and black communities, most respondents opted for the first choice. Nonetheless, many Latin Americans support strong policies against discrimination. Most respondents agreed that the passage of laws guaranteeing either the same salary for the same job for

Figure 1.2. Discrimination against Blacks by Country
(Index)



Source: Latinobarómetro 1999–2000.

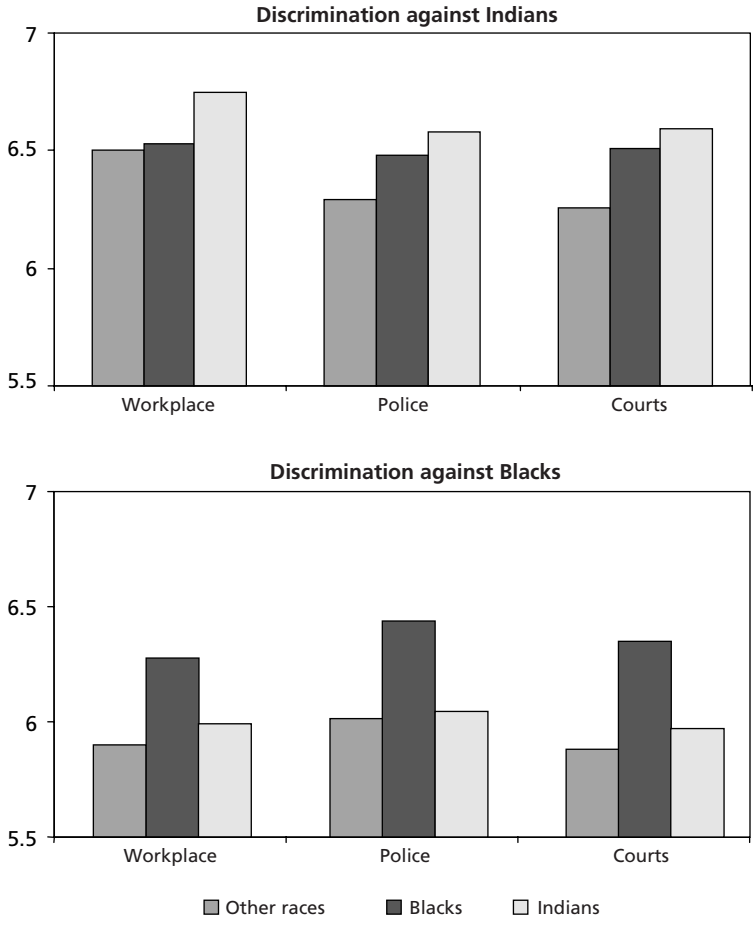
all racial and ethnic groups or the harsh punishment of those who commit discriminatory acts is crucial to solving issues of discrimination.

Race and Socioeconomic Status

A high and persistent correlation between race and socioeconomic status is usually although not always considered a sign of discrimination and social exclusion. If only for this reason, it is worthwhile to consider the evidence on the association between racial affiliation and status.

Those who responded to the Latinobarómetro survey were asked to report their racial or ethnic affiliation. They were also asked about posses-

Figure 1.3. Discrimination against Blacks and Indians by Race
(Index)



Source: Latinobarómetro 1999–2000.

sions of durable goods and the main features of their dwellings, which can be used to infer socioeconomic status. Three racial groups (blacks, Indians and "others") were distinguished, and all respondents were divided into quintiles of socioeconomic status.[3] Overall, 12.5 percent of the respondents

[3] Computing quintiles of socioeconomic status entails three main steps. First, principal components are used to compute a weighted average of the relevant household variables. All households are then ranked on the basis of this average. Finally, the corresponding ranking is used to compute quintiles of socioeconomic status. See Gaviria and Pagés (2001) and Filmer and Pritchett (1998).

classified themselves as Indians and 8.9 percent as blacks. Guatemala, Mexico and El Salvador have the highest percentage of Indians, and Brazil and Panama the highest percentage of blacks. With some caution, these numbers can be taken as representative of the urban populations of the countries under analysis.

Figure 1.4 shows that Indians and blacks are disproportionately represented in the lower quintiles of socioeconomic status.[4] Indians represent 12.5 percent of all respondents and 16.6 percent of those belonging to the first quintile. Blacks represent 8.9 percent of all respondents and 11.6 percent of those in the first quintile. There are also sizable differences in education among races. Average schooling is almost a full year lower among Indians and blacks than among other respondents. Not surprisingly, then, blacks and Indians are more likely to complain about their economic well-being: while 38 percent of blacks and 29 percent of Indians reported their economic situation as either bad or very bad, only 25 percent of those from other races did so.

Other Differences According to Race

There may be many other relevant differences between races that can provide important clues both on the extent of exclusion and discrimination as well as on the mechanisms whereby these problems affect socioeconomic outcomes. Relevant dimensions include political participation, social capital and general perceptions about the role of the state and access to opportunities.[5]

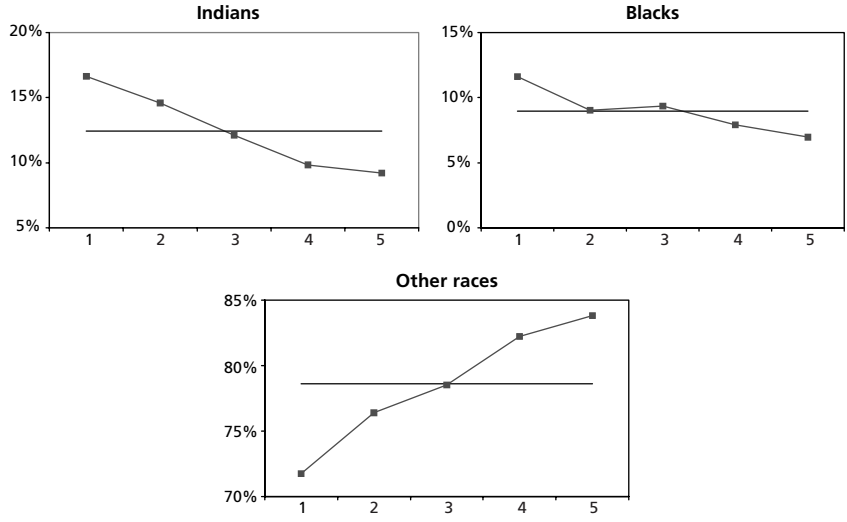
Differences in political participation, for example, may help explain political biases in favor of one group and against another. If people from one racial group participate in politics less assiduously than others do, social decisions would be biased against them. Moreover, low political participation by one group may be self-reinforcing; that is, people from this group do not participate in politics because they have been regularly left out, and they have been left out precisely because they do not actively participate.

However, Figure 1.5 shows no sizable differences in mean political participation among individuals from different races. No differences among

[4] Figure 1.4 and subsequent figures are based solely on differences among residents of the same country (i.e., cross-national differences were eliminated from the data).

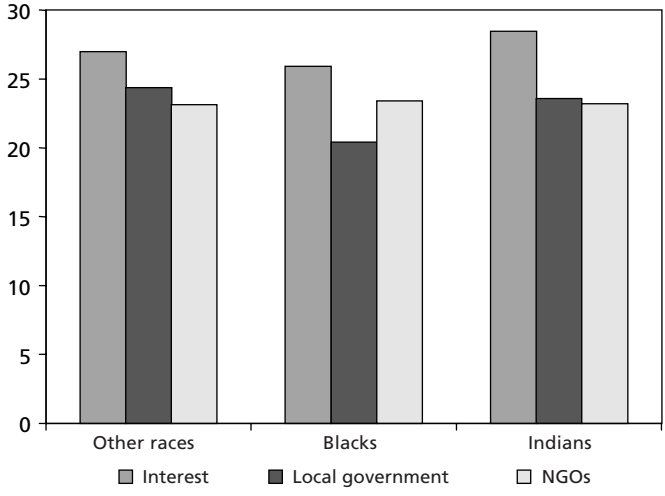
[5] Social commentators in the United States often cite the differences in the opinions between blacks and whites on a whole range of issues—from the advantages of affirmative action programs to the fairness of the justice system—as symptomatic of that country's racial divide.

Figure 1.4. Population Shares by Quintile and by Race



Source: Latinobarómetro 1999–2000.

Figure 1.5. Participation in Politics by Race
(In percent)



Source: Latinobarómetro 1999–2000.

races are apparent in the share of individuals that reported being interested in politics (26 percent), or in the share that reported that they regularly contact local governments (24 percent) or nongovernmental organizations (23 percent). These results do not depend on whether differences among races in education and socioeconomic status are controlled for.[6] Taken together, these results cast serious doubt on any attempt to explain political biases on the grounds of participation differentials across races.

Differences in the extent and density of social networks (that is, differences in social capital) may help explain differences among racial and ethnic groups, not only in socioeconomic outcomes but also in life satisfaction and other indicators of subjective well-being. As many fashionable theories have it, social capital (or the lack thereof) can explain why some people are richer, happier and healthier.[7]

However, Figure 1.6 shows that differences in social capital among racial and ethnic groups are insignificant, at least where social capital is measured by self-reported participation in civic organizations. After differences in education and socioeconomic status are taken into account, the proportion of individuals participating in at least one civic organization is two percentage points greater among Indians and blacks than among individuals from other racial or ethnic groups. In this instance, however, one should not put all the emphasis on participation, if only because participation in some organizations can promote social isolation, thus adversely affecting socioeconomic outcomes.

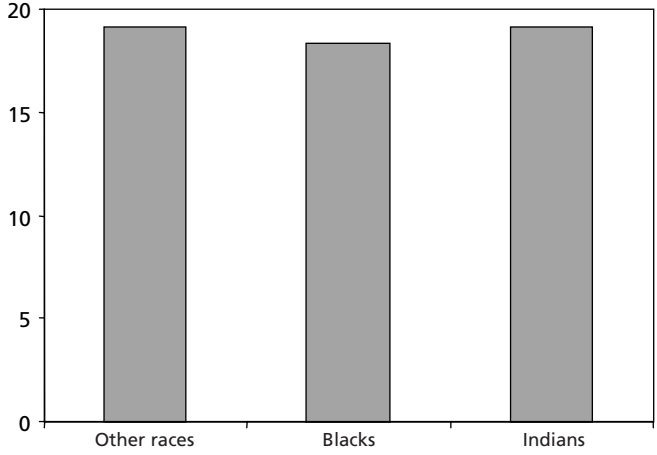
Differences in subjective well-being can also be of interest in their own right, as they complement the objective indicators mentioned above. Figure 1.7 shows that the share of individuals reporting that they are satisfied with their lives is at least five percentage points lower among Indians and blacks than among individuals from other races. This difference decreases only marginally after controlling for differences among races in education and socioeconomic status. This suggests that the lower levels of satisfaction with life among Indians and blacks apparently go well beyond what their relatively lower socioeconomic outcomes would indicate.

---

[6] IDB (2000) shows that political participation in Latin America is also similar among the different socioeconomic levels.

[7] See Putnam (2000), who argues that the decline of social capital is at the heart of many social problems affecting the United States.

Figure 1.6. Membership in Civic Associations by Race
(In percent)



Source: Latinobarómetro 1999–2000.

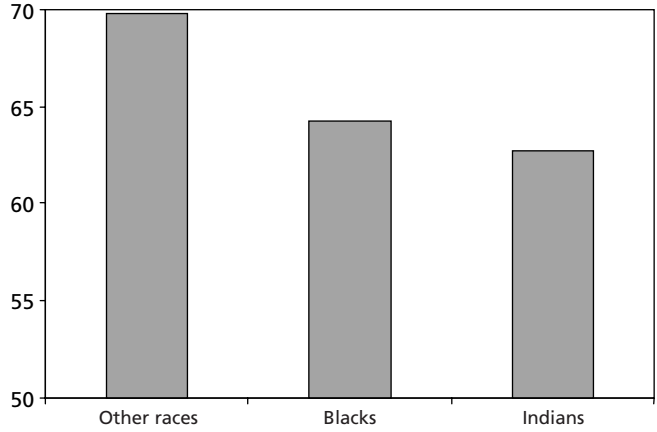
Although it may be tempting to interpret these differences as reflecting the psychological costs of exclusion, they may also be driven by unobserved differences in material possessions, occupational status or social mobility.

Analytical Framework

Social exclusion may be clear to see, but it can be difficult to define. For the purposes of this book, social exclusion is defined as "the denial of equal access to opportunities imposed by certain groups of society upon others." Such groups can be defined on the basis of religious beliefs, geographic location, ethnic origin, race, nationality, socioeconomic status, legal status, or other characteristics. The opportunities on which this book focuses include those related to schooling and labor and credit markets.

Some efforts to deny access to opportunities can be explicit. For example, homeowners enact zoning regulations so as to restrict access to their neighborhood; alumni control admissions to elite universities; bank officials discriminate against individuals on the basis of race. Other mechanisms are implicit. Housing prices prevent disadvantaged groups from

Figure 1.7. Percentage of Individuals Satisfied with their Lives, by Race



Source: Latinobarómetro 1999–2000.

moving to better neighborhoods or attending better schools, or health service or health insurance prices prevent excluded groups from obtaining better health care.

According to the definition used here, social exclusion occurs if the following two conditions apply: (i) social interactions occur predominantly within groups; and (ii) group membership has a sizable impact on access to opportunities for socioeconomic advancement. As an example, therefore, a definition of an exclusionist society would be one in which individuals interact mainly with others of the same race, and in which interactions conditional on race are key to access to jobs, credit, schooling opportunities and health care options.

Exclusion may take other very different forms. Perhaps the most evident is institutionalized exclusion, in which some groups are denied a voice and representation in public decisions. Women and minorities, for example, were not allowed to vote in many countries as late as the 1960s. Equality before the law, though, does not mean the absence of exclusion. Their legal right to vote notwithstanding, many social groups in Latin America have long been ignored by politicians and governmental officials and, as a result, have had little representation, despite their legal rights to participate in politics. As important and interesting as this type of exclusion may be, though,

it lies beyond the scope of this book, which instead focuses on exclusion as related to human capital investments such as schooling and health, and as related to income, such as that related to labor and capital markets.

Objectives

Given the definition of social exclusion set out above, this book has three specific objectives. The first is to measure the consequences of exclusion, especially as they pertain to the size and persistence of income inequalities and poverty. In the United States, for instance, there is ample empirical evidence that group membership has sizable associations with personal income, even after individual characteristics have been controlled; typically, the individual human capital characteristics that are observed in socioeconomic data sets account for less than 35 percent of the variance in income. This book attempts to document whether similar group effects are present in Latin America, and the extent to which group effects can account for the observed inequalities.

A second objective is to understand the main mechanisms of social exclusion in Latin America. While there are broad mechanisms of exclusion, as mentioned above, much uncertainty remains regarding the main channels through which this exclusion operates. Social exclusion, for example, may operate mainly by means of cultural norms that are transmitted through social interactions at the community level and allow the exclusion of those who do not speak, dress or behave in certain ways. Alternatively, exclusion may operate through informal networks that provide (or deny) access to job and educational opportunities. Similarly, colleges and schools can also be important means of social exclusion. Finally, exclusion may operate indirectly through prices.

The third objective is to identify policy interventions. The policies that can be derived from this framework may complement traditional prescriptions, such as increasing access to education, health services and labor and capital markets. Such policies usually aim to increase societal integration by ameliorating the cleavages that allow social exclusion to take place. Examples include charter schools, scholarship programs to private elite universities, and improved employment information. The ultimate goal is to understand the scope and importance of such policies in Latin America and their effectiveness in the region.

Methodology

Many studies have approached social exclusion descriptively. Ethnographic studies describing the mechanisms of exclusion and the community norms that usually prevent the poor from improving their lives have long been a staple of sociology. Among recent studies, the influential work of Wilson (1992) has painstakingly described the ways in which black youth from Chicago ghettos are denied opportunities for advancement.

In the economics literature, previous studies have tried to estimate the effect of group membership on socioeconomic performance. The general idea behind most of these studies is that estimating some variant of the following linear approximation provides information on the extent of social exclusion:

$$Y = c + X_1\beta_1 + X_2\beta_2 + C_1\gamma_1 + C_2\gamma_2 + \alpha Z + e \qquad (1)$$

where Y is an indicator of socioeconomic performance for an individual (e.g., school enrollment, school attainment, health clinic usage, employment, wage rate, formal sector job), X_1 is a set of observable personal characteristics (age, sex, etc.) for that individual, X_2 is a set of unobservable characteristics (e.g., ability, work ethic) for that individual, C_1 is a set of observable characteristics for the community in which the individual lives (e.g., market prices, climate and other exogenous conditions), C_2 is a set of unobservable community characteristics, Z is an indicator of membership in a group or of some relevant attribute of the group of which the individual is a member, and e is a stochastic term to reflect chance events. The group can be defined by geographical proximity (e.g., a neighborhood, a city), demographic characteristics (e.g., religion, ethnicity, migrant status), class, or affiliation with some institution such as a school or a firm.[8]

The coefficient α can be interpreted as a measure of the strength of group effects. High values of α point to the presence of social exclusion, as they indicate the importance of group membership for access to economic opportunities. With good estimates of equation (1), the relative importance of group membership in the overall variance of the socioeconomic index of interest (Y) can be determined by variance decomposition of equation (1)

[8] See Case and Katz (1991), Crane (1991), Borjas (1995a), Cutler and Glaeser (1997) and Kremer (1997).

to find var(αZ)/var(Y). Interactions can be added to equation (1) to explore, for example, whether group membership interacts with individual characteristics such as sex or with community characteristics such as the nature of schools and job markets.

This approach, however, faces the following challenges:

(1) Finding variables that accurately represent Z. A pragmatic approach is to investigate whether different categories of group memberships that are available in the data have significant effects. But there is no guarantee that particular data sets include indicators of membership in the most relevant groups.

(2) Obtaining estimates of the effects of Z that are not contaminated by unobserved variable biases from unobserved individual (X_2) or community (C_2) variables. To lessen such possibilities, it is desirable that as many as possible of the relevant individual and community variables be controlled in the estimates (e.g., through expanded efforts at measuring them or through fixed effects).

(3) Disentangling true group effects from the aggregation of a number of individual effects if the dependent variable is identical to or closely related to the group indicator (e.g., individual schooling, group average schooling). Manski (1993, 1995) has called this the "reflection problem." For example, if all members of one group, defined by geography or by some demographic characteristics, perceive there to be relatively low returns to schooling relative to those perceived by others, the component of Z representing average schooling for the group is likely to be significant in econometric estimates—not because the average schooling for the group *causes* low individual schooling, but because schooling for all members of the group is responding to the perceived low returns to schooling.

(4) Assessing the impact of group membership if there is correlation between group membership and individual or community characteristics. The expression var(αZ)/var(Y) ignores such covariances. One alternative is to present both this expression and the one in which all the covariances between group memberships and individual and community characteristics are included in the numerator in order to see how sensitive the calculation of the con-

tribution of the group effect is to the two extreme treatments of these covariances.

The most desirable strategy would be to estimate an equation such as (1)—including variants with interactions—where the dependent variables are measures of socioeconomic performance (including both human resource investments and market outcomes) and the independent variables include information on group membership as well as all relevant correlated individual and community characteristics. Such estimates would make it possible to assess: (i) the extent to which membership in these groups affects opportunities to achieve an adequate standard of living through human resource investments and aspects of market access; (ii) which types of group membership are most important empirically; (iii) what proportions of the variations in the socioeconomic variables investigated are accounted for by group memberships; (iv) whether the impacts of group membership are similar across the various socioeconomic indicators, as would be the case if there are general patterns of social exclusion for particular groups; and (v) whether group membership, and thus social exclusion, affects access to social services and other policy-related indicators.

Readily available data sets, such as typical household surveys, permit only limited exploration of these issues. This book has benefited from researchers and data sets that permit more extensive examination of these issues as well as a multi-country perspective.

Country Studies

The particular combination of country studies chosen for this book is interesting for several reasons. Not only do most of them use very different types of non-conventional data sets, but each study also uses different econometric techniques for avoiding some of the interpretation problems discussed above. The studies further point out that the most obvious policy responses are not always the best options. There is no standard recipe for fighting exclusion.

Residential Segregation in Bolivian Cities

Even though this is the type of exclusion most studied in the literature, George Gray-Molina, Ernesto Pérez de Rada and Wilson Jiménez use an innovative

approach to explore the effects of residential segregation in Bolivian cities. The question they address is whether living in certain geographic areas negatively affects incomes and schooling attainment. Since indigenous groups constitute a large proportion of the Bolivian population, one important challenge is to disentangle the economic effects of ethnicity and other personal characteristics from the effects of living in specific neighborhoods.

The authors reach two main conclusions: first, living in specific geographic areas within Bolivian cities has a negative and significant effect on incomes; and second, individuals living in segregated geographic areas and belonging to certain racial groups have lower incomes and educational attainment, probably because of social exclusion.

The chapter first discusses an analytical framework, which consists of adapting an existing model to the particular case of segregation in Bolivian cities. Within this discussion, the three main econometric problems of the analysis are highlighted. The first of these problems is reverse causality, which is a standard problem in this context, since residential segregation might be the result of poor economic outcomes rather than a cause of them (that is, exclusion determines location and not vice versa). To address this problem, the authors propose three different types of instrumental variables.[9] The first consists of data on residential settlements 25 years ago. The second consists of data on changes in residential location after a drought-induced migratory shock, which generated a pattern of migration into several Bolivian cities characterized by an even distribution of migrants into low-income and high-income neighborhoods. The third uses data on population density in order to include information on the geographic features of the largest cities, which are expected to be important determinants of location.

The second econometric problem is omitted variable bias. This problem originates because income and schooling (which are the main variables of inter-

[9] In instrumental variable estimates, right-side variables are replaced by estimated values (based on "instruments") that hopefully are independent of the disturbance term (i.e., of the unobserved individual and community variables on the right-side of equation 1 above). If the instruments themselves are independent of the disturbance term, if they do not belong in the relation of interest being estimated, and if they account for sufficient variance in the right-side variable(s) being instrumented, the result will be consistent estimates of the effects of right-side variables. This method may help eliminate biases due to right-side variables that reflect current or past behaviors or omitted variable biases. Finding instruments that satisfy the three conditions noted above is often difficult, and not all studies that purport to have such instruments are persuasive in this regard (see Rosenzweig and Wolpin, 2000).

est) may be influenced by unobserved parental and community characteristics that can be correlated with the neighborhood variables. Parental and community-level attributes are included in the econometric estimations in order to reduce this bias, but, as argued by the authors, some biases may persist.

A third problem arises because, if individuals are able to change location at will, econometric analysis may not capture the effect of neighborhood on outcomes, even if segregation does have an adverse effect on incomes and schooling. The authors propose using information on younger cohorts, which presumably have not yet made their location choices. Additionally, they propose incorporating information on the migrant or non-migrant situation of the household to reinforce controls on mobility. These two additional variables help assure that the effects captured by the econometric estimations can be interpreted as evidence of social exclusion.

One interesting feature of this chapter is the data used for the analysis. The authors perform tests on the information from the 1999 standard household survey for Bolivia, which includes data on self-reported ethnicity and geographic location, as well as a set of socioeconomic variables and personal characteristics. However, since the survey does not contain enough information for constructing adequate instrumental variables or to address omitted variables bias, the authors conducted a new survey for the purposes of this book. They collected detailed information for two neighborhoods in the cities of La Paz and El Alto. The sample, consisting of 801 households in 43 neighborhoods, is representative of the three main socioeconomic strata (high-, middle- and low-income households), and of the entire metropolitan areas of each city. The survey questionnaire is used for collecting basic socioeconomic characteristics, but additional questions are included to collect information on parental ethnic background and language, human and social capital formation, and perceptions of segregation.

The econometric analysis leads the authors to conclude that living in a segregated neighborhood adversely affects labor income and educational attainment. This effect remains after controlling for personal characteristics and community background. On the whole, the evidence hints at the existence of geographically based exclusion in the main Bolivian cities.[10]

[10] One surprising result of the analysis is that residential segregation apparently tends to have a stronger negative effect on income for second-generation migrants. First-generation migrants seem to benefit from social capital networks, which have a positive effect on income and minimize the effect of spatial segregation.

Social Exclusion and the Two-Tiered Health Care System of Brazil

Denisard Alves and Christopher Timmins focus on the implicit exclusion that occurs in the Brazilian health care system. By means of differential pricing and quality of services, some sectors of the population are effectively excluded from obtaining adequate heath care.

The authors first describe the Brazilian health care system in detail and argue that its quality is very low, with high shadow prices explained by long waiting times and travel costs. These "hidden" costs often discourage poor households from using the public system. Since the poor are not able to afford higher-quality private services, they end up under-using health care, if they use it at all. By contrast, richer households pay higher up-front costs for private medical care and are able to obtain better services.

The authors then carry out a more formal analysis using the 1998 national household survey (PNAD). This particular round of the survey permits such an analysis because it includes a special supplement with information on health conditions, health care consumption, and the types of health services used by the population.

The first approach to identifying mechanisms of social exclusion consists of estimating a set of probit regressions that identify the population subgroups that are more likely to use public health services. The limitation of this approach, noted by the authors, is that it does not provide a way of quantifying the welfare costs of this type of exclusion. Therefore, the authors develop a formal model of health care and insurance choice. One important feature of the model is that it incorporates measures of the shadow price of accessing the public health system. The data allow the estimation of such measures for each individual in the sample, which makes the empirical analysis feasible.

The authors use the model, and the estimated shadow prices, to examine the welfare effects (reduced access) of increasing the price of public health care. They conclude that individuals living in the north, belonging to black or mixed racial groups, above 60 years of age, and with lower levels of education will be more vulnerable to price increases than the rest of the population. These effects support the view that these are the groups in Brazilian society that suffer the most from inadequate health services.

To explore the policy implications of their analysis, the authors also perform simulations of the welfare effects of subsidizing private health care

services. Surprisingly, the population groups that are currently excluded from the system would *not* benefit most from this measure, since they still have to pay high shadow prices for accessing them. This type of policy would benefit higher-income individuals in the most developed regions of the country and therefore would imply a transfer of rents to the rich.

Finally, the analysis suggests that expanding public infrastructure, and therefore, decongesting the current public health care system, might be the best way to "include" the excluded groups in the benefits of health care, in the hope that this will permit them to live longer and healthier lives.

Legal Status and Social Exclusion: Nicaraguans in Urban Costa Rica

Edward Funkhouser, Juan Pablo Pérez Sáinz and Carlos Sojo address whether Nicaraguans migrating to Costa Rica have lower socioeconomic status because of social exclusion due to nationality. The question is highly relevant for Costa Rica, because an estimated 350,000 to 450,000 Nicaraguans migrated there during the 1990s, and the growing presence of Nicaraguans has transformed Costa Rican society.

Interestingly, there are relatively few ethnic, language or even cultural differences between Nicaraguans and Costa Ricans that could account for differences in labor market outcomes. The study argues that the reason why Nicaraguans have lower socioeconomic status is *not* because of their nationality, but because of their legal status. Most Nicaraguan migrants have entered the country illegally and therefore are likely to receive different legal treatment. This finding has important policy implications, as it directs public action towards a set of interventions aimed at easing legal conditions rather than at creating jobs for Nicaraguans or creating mechanisms for preventing discrimination.

As in the Bolivian study, the standard household survey available for Costa Rica has many limitations for the task at hand, mainly because it was not designed to study social exclusion. Therefore, the authors conducted a survey for this study in the metropolitan area of San José. It comprises 398 households and uses a questionnaire especially designed to examine social exclusion. One feature of the new survey is that it includes sampling units from neighborhoods with a high, medium and low presence of Nicaraguans. Comparisons between the aggregate values of socioeconomic variables in the national survey and the new survey assure that the data collected for this study are reliable.

The data show that Nicaraguans in Costa Rica are not so much an "excluded" group as an illegal one. For instance, Nicaraguans do not live in segregated neighborhoods: even in neighborhoods with large proportions of Nicaraguans, Costa Ricans are a majority. Moreover, Nicaraguans have access to labor markets, and Nicaraguan women display even higher participation rates than their Costa Rican counterparts.

As the main explanation for the gap in socioeconomic level is legal status, the main policy implication is that increasing the probability of legal residence may help to improve the standard of living of Nicaraguans living in Costa Rica. This can be accomplished by eliminating passport or other document requirements, or simply through an amnesty that changes the legal status of the Nicaraguan migrants.

Geographic Isolation and Labor Market Outcomes in Rural El Salvador

Ana Regina Vides, Anabella Lardé de Palomo and Lissette Calderón analyze the effects of spatial isolation on labor force participation, sector of employment and labor income levels in rural areas of El Salvador.

The study uses a rural household survey conducted in 1999 by the Salvadoran Foundation for Economic and Social Development. This survey contains extensive information about access to markets and other measures of spatial isolation.[11] The main argument is that people living in isolated areas are excluded from the mainstream economy and therefore have lower socioeconomic status and fewer employment opportunities. Lack of roads and transportation infrastructure in general is therefore the main mechanism through which social exclusion operates in rural El Salvador.

Lack of infrastructure creates a combination of security hazards and transaction and moving costs that reduce labor force participation and force workers into jobs with low productivity. These outcomes in turn prevent individuals from moving away from isolated areas, thus completing a vicious circle. The evidence provided by the authors suggests that social isolation is particularly detrimental to women. Women living in isolated areas generally have much lower labor force participation rates and lower incomes, and they tend to concentrate in the sectors with the lowest productivity.

[11] Questions for access to markets include "travel time to closest paved road" and other similar indicators. The location index has two components, one that measures access to all urban jobs, and another that measures access to jobs in free trade zones.

The main policy implication of the analysis is that building new roads and expanding public transportation and household services such as water and electricity may have a larger impact on the socioeconomic conditions of isolated individuals than standard poverty alleviation programs, or even public health and education programs. Well-educated individuals would fare better, even under the unfavorable conditions of isolation, but the evidence suggests that if traditional social spending is not complemented with policies that reduce isolation, such spending will have only limited impact on the standard of living.

Language Barriers and Schooling Inequality of the Indigenous in Mexico

The only case study in the book to focus primarily on social exclusion based on ethnicity is the study for Mexico. This type of exclusion is highly relevant for many Latin American countries, where the indigenous fare much worse than other groups of society. Even though the most common mechanism of social exclusion in terms of ethnicity is outright discrimination, the study by Susan Parker, Luis Rubalcava and Graciela Teruel shows that there are more subtle mechanisms of exclusion that are strong determinants of key socioeconomic characteristics.

Exploring the extent of social exclusion through differences in schooling attainment, the authors find that children who only speak indigenous languages fare much worse in school than similar indigenous children who differ only in terms of knowledge of Spanish. After controlling for family and community characteristics, the authors provide evidence that language barriers and cultural factors faced by monolingual indigenous children prevent them from benefiting from the school system to the same extent as bilingual indigenous children. Therefore, it is not access to schools per se, but to specific types of education, that acts as a mechanism of social exclusion.

This case study requires detailed data on ethnic background, language spoken, and parental socioeconomic characteristics. Such information is only available in Mexico from ENCASEH, a survey of household socioeconomic characteristics undertaken especially for evaluating the effects of the Program for Education, Health and Nutrition (PROGRESA). The ENCASEH data used in the chapter are supplemented with school-level information and variables from the Ministry of Education characterizing the quality and quantity of infrastructure. This makes it possible to link detailed personal data with characteristics of available schools.

While the chapter shows that, in general, monolingual indigenous children have lower schooling outcomes than their bilingual counterparts, the authors undertake a more formal approach to their study, given the limitations of the descriptive analysis. The main challenge is to distinguish between the effects of cultural and language barriers versus social and economic factors affecting schooling outcomes of indigenous children. The central question is whether there is evidence that the poorer performance of monolingual children is due to their worse economic condition, or to other factors such as language barriers. This issue is highly relevant from the policy point of view because, if economic variables explain schooling lags, then the introduction of anti-poverty programs would perhaps be the best response. On the other hand, if language barriers are more important, a different policy approach is indicated.

Empirical estimations for addressing these questions face several problems. The main issue is endogeneity of language choice. Households may choose not to learn Spanish or not to attend school at all due to cultural reasons, and these decisions may be reflected in failure to achieve the standards set by the schooling system. Instrumental variables are used to minimize this potential problem.

In order to improve the understanding of the differential effect of language barriers versus unobserved cultural factors, the authors examine the impact that bilingual education may have on schooling outcomes for indigenous children.[12] This is an important question, since the Mexican government recently embarked on an ambitious initiative to expand bilingual education. The conclusion from this exercise is that bilingual schools significantly improve schooling outcomes of monolingual children, and they actually contribute to narrowing the schooling gap between these children and the bilingual indigenous.

Thus, the results support the view that policy interventions such as expanding access to bilingual schools for indigenous monolingual children may have a strong positive effect on their ability to benefit from the public school system. Bilingual schooling, then, may be an important instrument for "including" excluded groups of society so that they might obtain the benefits of development.

[12] To implement this analysis, data from the Ministry of Education on the location of bilingual schools are combined with the ENCASEH data.

This page intentionally left blank

CHAPTER TWO

Residential Segregation in Bolivian Cities

George Gray-Molina
Ernesto Pérez de Rada
Wilson Jiménez

Rapid population growth and urbanization are changing Bolivia's poverty profile. Between 1992 and 2000, an estimated 1.7 million people moved from rural to urban areas, causing the number of poor urban households to surpass the number of poor rural households.[1] A side effect of rapid urbanization, often neglected by current research on poverty in Bolivia, is a measurable increase in residential segregation by income, ethnicity and other social and cultural identifiers.[2] Neighborhoods that link first- and second-generation migrants through economic as well as cultural ties also increase pressure on available labor markets, housing and public services.

Qualitative evidence on rural-to-urban migration in La Paz and El Alto suggests that, like middle-class residential segregation, much of the new segregation is self-selected (see Albó and Preiswerk, 1986). Migrants seek out neighborhoods and social networks established by relatives and peers. Established networks are then used to get a foothold in formal and informal labor markets. Although qualitative studies largely skirt identification issues, they do raise concerns regarding current analyses of urban poverty in Bolivia. By describing the social context of demographic change, they point to neighborhood effects, both positive and negative, on income generation, labor and educational achievement. Beyond the disadvantages of interpersonal discrimination, residential segregation poses a secondary tier of potential social exclusion.

[1] See UDAPE (2000) and INE (2000a).

[2] Recent studies of urban poverty based on household surveys include Jiménez and Yañez (1997); Pérez de Rada (1997); Vos, Lee and Mejía (1998) and Gray-Molina et al. (1999).

Analytical Framework

A growing literature on urban poverty provides a basic analytical framework to assess the effects of residential segregation on income, labor and educational achievement (see Borjas, 1997; Cutler and Glaeser, 1997; Glaeser, Kahn and Rappaport, 2000). Two types of questions have been considered in recent empirical work. The first set of studies focuses on estimating the marginal effect of ethnicity-based residential segregation on an array of outcomes (see Holzer, 1991; Kain, 1992). Most studies use intra-city comparisons that consider whether minorities living in predominantly minority neighborhoods fare worse than those who live in mixed neighborhoods. A serious identification problem, however, complicates intra-city comparisons. To the extent that residence is self-selected, any direct measure of segregation is likely to be biased. Instruments correlated with segregation but unrelated to outcomes are required to disentangle dual causality.

The second question considered by recent work is whether outcomes for minorities as a whole are better or worse in cities that are more racially segregated compared with cities that are less segregated (Cutler and Glaeser, 1997; Glaeser, Kahn and Rappaport, 2000). This approach avoids intra-city selection problems, but faces other problems linked to the potential for reverse causality—where poor outcomes lead to increased segregation—and the potential bias from sorting of more or less successful minorities across a city. As with intra-city comparisons, across-city comparisons require instruments to address selection problems. This chapter focuses on both intra-city and across-city comparisons of residential segregation by collecting new data that correct for potential self-selection and omitted variable biases.

Basic Model

The point of departure is a spatial equilibrium model developed by Cutler and Glaeser (1997) to analyze the effects of residential segregation across and within cities. The Cutler-Glaeser model, adapted to the analysis of segregation in Bolivian cities, takes the form:

$$Outcome = X'\beta + \beta_1\, segregation + \beta_2\, segregation^{*}indigenous + e \quad (1)$$

where *outcome* is measured at the individual level, $X'\beta$ represents an array of individual and household controls, and s*egregation* is a neighborhood

measure of the separation of ethnic groups. The coefficient β_1 measures the effect of segregation on non-indigenous, and β_2 is the differential effect for indigenous relative to non-indigenous. The variable *segregation* is measured at the level of census section tracts, contiguous groups of households settled in an area equivalent to eight to 10 city blocks. Indexing tracts by *i*, residential segregation within a city is defined as:

$$Segregation = {}^{1}/_{2} \sum_{i=1} | \text{indigenous}_i/\text{indigenous} - \text{nonindigenous}_i/\text{nonindigenous} | \quad (2)$$

where $indigenous_i$ is the number of people speaking an indigenous language (Aymara, Quechua or other) in tract *i*, and *indigenous* is the total number of indigenous-speaking people in each city. If indigenous-speaking residents are evenly distributed throughout metropolitan areas, the term in absolute value brackets will be zero for each census tract and zero for the metropolitan area as a whole. If indigenous and non-indigenous speakers never reside in the same census tract, the measure of residential segregation will be one. This measure estimates the share of the indigenous population that would need to change census tracts to even out the spatial distribution of ethnic groups within a metropolitan area.

Addressing Self-selection Bias

The Cutler-Glaeser model requires three adjustments to account for self-selection, omitted variable bias and mobility across neighborhoods. The first adjustment concerns self-selection bias. Residential segregation might be the result rather than the cause of poor economic outcomes. In this case, successful indigenous-speaking residents will choose to live in richer and non-indigenous neighborhoods, rather than settle randomly. To instrument for segregation, three sets of variables are proposed. The first instrument uses historical data on residential settlements 25 years ago. While historical segregation is likely to be partially correlated to income-based selection, the objective here is to test whether pre-migration settlement patterns vary systematically from post-migration settlements that today make up half of the population in capital cities.

The second instrument uses information on residential sorting before and after a drought-induced migratory shock. The residential sorting of rural

migrants from the 1982–85 highland drought tends to be evenly distributed among high- and low-income neighborhoods. The third instrument uses historical figures for population density to approximate geographic features that account for natural occurrences of segregation in the cities of La Paz and El Alto—neighborhoods separated by geographic fault lines, rivers and gorges.

Addressing Omitted Variable Bias

The second adjustment leads to a reformulated model controlling for omitted parental and community characteristics correlated with segregation. In this case, the measure of success for indigenous-speaking residents might overstate the effect of individual-level characteristics. Questions on parental and community-level human and social capital are proposed to avoid this bias. Omitted variables that account for intergenerational and peer advantages and disadvantages are likely to play a key role in explaining proposed segregation effects.

Controlling for Geographic Mobility

A third adjustment concerns intra-city mobility. Even if segregation does lead to poor outcomes, intra-city comparisons may not pick up the negative effect if mobility is assumed. In other words, with intra-city mobility, outcomes for similar people are expected to be the same regardless of place of residence. One way to control for mobility is to contrast young age cohorts, which presumably have not yet made their residential decision, with older ones. Another way is to control for migrant versus non-migrant households. Both age cohorts and migration are controlled for in the intra-city analysis.

Key Hypotheses

In theory, residential segregation might lead to positive or negative effects on an array of outcomes. Recent research on peer and neighborhood effects suggests that the mechanisms that link segregation to outcomes might include human capital, social capital and ethnic capital externalities.

Positive Effects

Recent economic and sociological work on social capital is converging toward individual-based models of social capital formation.[3] The new approach contrasts with the earlier wave of group-based research led by Coleman (1990) and Putnam (1993 and 2000), which focused on networks and communities. Two consequences of the shift are a more careful account of causality and clearer linkages with household and individual decision-making models.

Among positive social capital effects, three transmission mechanisms are identified that link residential segregation and individual-level outcomes. First posited are positive intergenerational effects accruing from parental social capital, in the form of expanded social networks and access to an array of intergenerational "weak ties." Parental social capital is likely to be most valuable for second- or third-generation migrants who have cultivated a network of neighborhood and extra-neighborhood contacts. Labor, credit and political ties are mobilized by neighborhood contacts.

Second, income and labor effects are posited that accrue from individual membership in local groups, associations and networks. Personal membership captures unobserved characteristics that proxy for mobility, thrift and outward-orientation. Not all membership affiliations, however, will be outward-looking, and discriminating between outward and inward-oriented groups is intended to capture positive "weak tie" effects rather than potentially negative "strong tie" effects. For instance, sports, political and labor affiliations might proxy outward-looking ties, while ethnic, cultural and religious ties might proxy inward-looking ties (see Gray-Molina et al., 1999).

Third, positive political effects through more effective neighborhood voice are posited. Segregated neighborhoods might initiate collective action more readily than non-segregated neighborhoods by mobilizing ethnic or cultural ties. Under decentralized decision-making, neighborhood associations have a significant impact on public investment and public service decisions (see Gray-Molina, 2000). Over the past five years, annual participatory planning rounds have mobilized over 17,000 territorial grassroots organizations,

[3] Notable contributions include Glaeser, Laibson and Sacerdote (2000), Alesina and LaFerrara (2000), Portes (1998), and Portes and Landolt (2000).

1,200 of which are from the La Paz/El Alto metropolitan area alone. Although political effects might be indirect and lagged, they are included in order to account for neighborhood voice since 1995 (see Gray-Molina et al., 1999a).

Negative Effects

The recent literature on urban poverty suggests a number of negative effects through peer or neighborhood interactions.[4] Places of residence are important arenas for learning valuable individual and social skills, and residential segregation tends to limit the social and labor interactions to restricted pools of human capital and public goods.

Two main transmission mechanisms are posited. The first involves negative peer and intergenerational effects. In segregated neighborhoods, human capital endowments among poor and first- or second-generation rural migrants are likely to be significantly lower than in non-segregated neighborhoods. This limits both present and future outcomes through intergenerational effects, and those combined effects can result in damaging externalities to both individuals and households. Such externalities are of particular concern in light of recent research that suggests the strong persistence of ethnic segregation.

The second mechanism involves negative neighborhood-wide effects through public goods externalities. Residential segregation is likely to hurt outcomes of the poor if the quality of neighborhood-specific public goods, such as schools or health centers, varies along segregated lines. Heterogeneous public goods and services will influence human capital formation and reinforce peer and parental effects. Available district-level school and health data on service performance are proposed to account for neighborhood differences.

The hypotheses focus on individual, parental and neighborhood-wide mechanisms for human capital and labor market exclusion. Alternative specifications of the rural segregation measure and interactions with selected variables described above should provide a reliable estimate of the effects of residential segregation on income, labor and educational achievement outcomes. A shortcoming of this framework is that across-city estimates are

[4] Notable contributions include Borjas (1995a and 1997), Cutler and Glaeser (1997), and Glaeser, Kahn and Rappaport (2000).

likely to omit key parental and community characteristics. Intra-city comparisons will attempt to correct for these but will also fall short in capturing the externality effects of ethnicity-based segregation in non-segregated areas. Segregation is likely to hurt indigenous citizens *regardless* of their place of residence. It is assumed that observed and unobserved individual characteristics of indigenous residents in mixed neighborhoods will pick up citywide externalities, but the extent of this effect cannot be established.

Residential Segregation in 10 Bolivian Cities

This section considers the effect of residential segregation on income across 10 Bolivian cities. Available data from the 1999 MECOVI[5] survey is used to construct residential segregation variables from 1976 and 1992 census-tract data. Self-selection and omitted variable bias is controlled for with available data on labor insertion and parental ethnicity and schooling. Individual-level ethnicity is measured in terms of self-identification and neighborhood-level ethnicity through language spoken.

Data

The MECOVI survey includes data from 3,800 households and 13,131 individuals collected in November/December 1999. The survey is nationally representative and also representative of highland, valley and lowland subregions. The survey includes expanded sections on ethnicity (language and self-identification) and demographic indicators. It also includes modules on health and morbidity, education, labor characteristics, labor and non-labor income, food and non-food expenditures, housing conditions, and durable and non-durable assets, as well as special modules on agricultural production and consumption in rural areas. The survey is designed to collect panel data on a yearly basis.

Table 2.1 shows the distribution of residential segregation across Bolivian cities in 1992. The Duncan index presented in the last column measures the proportion of the population that would have to move from a

[5] MECOVI is the acronym for the *Programa de Mejoramiento de las Encuestas y Medición de Condiciones de Vida—Program to Improve Surveys and Measurements of Living Conditions.*

Table 2.1. Duncan Index, Residential Segregation, 1992

	Population age 5 or older	Speaks Spanish (%)	Speaks native language (monolingual and bilingual) (%)	Duncan index[1]
Sucre	110,327	39.0	61.0	0.203
La Paz	614,529	55.1	44.9	0.225
Cochabamba	319,730	28.7	71.3	0.584
Oruro	155,415	49.5	50.5	0.141
Potosí	92,638	30.6	69.4	0.198
Tarija	75,603	84.7	15.3	0.251
Santa Cruz	572,235	84.9	15.1	0.239
Trinidad	46,656	91.4	8.6	0.163
Cobija	8,299	88.7	11.3	0.065
El Alto	332,065	36.0	64.0	0.214

[1] See Cutler and Glaeser (1997).
Source: 1992 census.

segregated to a non-segregated neighborhood in order for the population to be equally distributed within cities. The most segregated cities, according to the index, include La Paz, El Alto, Santa Cruz and Cochabamba, also the largest urban settlements in the country.

Results

The data used in this section focus on urban households in nine departmental capitals and the city of El Alto. The analysis is based on 2,059 urban heads of household. Table 2.2 shows two OLS regressions and one IV regression. The first regression presents basic OLS results for a human capital specification. The expected signs are found for education, experience, experience squared and sex. The second specification shows OLS results for an augmented human capital specification that includes the 1992 residential sorting proxy for segregation. This shows that indigenous-language speakers are significantly worse off than non-indigenous speakers and that living in an indigenous neighborhood is negatively correlated with income. A third

Table 2.2. Results on Residential Segregation in 10 Cities

	(1) OLS 1992 census	(2) OLS 1992 census	(3) IV Migrant shock
Schooling	0.112	0.104	0.105
	23.57	20.99	19.54
Experience	0.033	0.035	0.032
	7.75	8.56	6.44
Experience squared	–0.000	–0.000	–0.000
	–4.37	–5.01	–4.68
Female	–0.310	–0.317	–0.339
	–7.65	–7.85	–7.37
Bilingual indigenous (a)		–0.242	–0.106
		5.31	–0.31
Monolingual indigenous (a)		–0.319	–0.010
		–0.67	–0.02
Residential segregation		–0.002	–0.046
		–2.63	–2.38
Residential segregation squared			0.000
			2.75
Language*Segregation			0.007
			0.62
Indigenous self-identification			–0.109
			–1.96
Constant	1.467	1.732	2.626
	20.37	21.39	5.48
No. of observations	2,057	2,048	1,904
R^2	0.28	0.29	0.27

Source: Authors' calculations based on 1999 MECOVI and 1992 census.

specification instruments for segregation with a measure of drought-induced settlement and adds an interaction term for individual and neighborhood-level ethnicity. Indigenous speakers who live in predominantly indigenous-speaking neighborhoods (language*segregation) are worse off than those who live in less-segregated neighborhoods. This result coincides with much of the literature on residential segregation. The third specification includes an expanded measure of ethnicity to test whether language or self-

identification ethnicity measures have a different impact. Ethnic self-identification is also negatively and significantly correlated with income.

While the across-city specifications suggest initial negative effects of segregation on income, the MECOVI surveys do not permit testing additional instrumental variables or controlling for potential omitted variables biases. In particular, it is impossible to control for intergenerational human capital and social capital effects that may account for lagged effects on income generation.

Hypothesis Testing

Table 2.3 tests two hypotheses on the causal mechanism behind peer effects in segregated neighborhoods. First considered are the effects of residential segregation on migrant and non-migrant populations. For migrants from birth, residential segregation has a smaller (when contrasted to non-

Table 2.3. Results by Migrant Status and Age Group

	By place of birth		By residence last 5 years		Age group	
	Non-migrant	Migrant	Non-migrant	Migrant	25 years or less	25 or older
Schooling	0.115	0.095	0.103	0.114	0.120	.0888
	12.45	12.73	16.34	7.88	10.52	14.67
Experience	0.033	0.033	0.033	0.067	0.020	.0232
	4.33	5.39	6.85	5.04	0.79	3.687
Experience squared	–0.000	–0.000	–0.000	–0.001	0.003	–.0002
	–1.61	–3.90	–3.95	–3.74	2.02	–2.32
Female	–0.184	–0.413	–0.286	–0.443	–0.208	–.2957
	–2.77	–6.00	–5.64	–3.27	–2.65	–6.33
Segregation	–0.005	–0.003	–0.004	–0.002	–0.001	–.001
	–3.87	–2.78	–4.74	–0.94	–1.36	–1.69
Segregation*	–0.003	–0.003	–0.003	–0.007	–0.0005	–.002
Indigenous language[1]	–1.78	–2.99	–2.57	–2.36	–0.27	–1.71
Constant	1.595	1.886	1.756	1.551	1.225	1.743
	11.98	14.31	17.13	7.02	6.85	12.10
No. of observations	972	1,085	1,863	194	450	1,609
R^2	0.29	0.32	0.29	0.46	0.36	0.39

[1] Indigenous language includes both monolingual and bilingual indigenous speakers.
Source: Authors' calculations based on 1999 MECOVI and 1976 census.

migrants) but significant negative correlation with income. For five-year migrants, residential segregation is statistically insignificant. Both results bolster findings suggested by Cutler and Glaeser (1997) and Borjas (1995a) that the effects of residential segregation work through intergenerational transmission of income-generating opportunities. For the 10 cities studied, segregation conspires against income generation over a long period of time but has no significant impact in the short run. Subsequently considered are the effects of segregation on different age groups. The results are inconclusive, as no significant differences emerge between young and old age groups. This hypothesis will be discussed further below.

Residential Segregation in La Paz and El Alto

Data

The survey collects income, labor, human and social capital data from 801 households in 43 neighborhoods in the cities of La Paz (24 neighborhoods) and El Alto (19 neighborhoods). The sample is representative of the entire metropolitan area and is further stratified by high, average and low levels of ethnic-based residential segregation. Together, sampled neighborhoods (*zonas censales*) hold a population of close to 200,000 (in a metropolitan area of approximately 1.4 million), with an average of approximately 4,000 residents per neighborhood.

The contents of the survey largely draw from the 1999 MECOVI household survey. Questions on demographic characteristics, educational achievement, labor insertion and income are identical to those on the MECOVI survey. The survey also collects detailed information on parental ethnicity, language, and human and social capital and includes expanded questions on social capital and perceptions of segregation from heads of household and spouses.

Measures of Intra-City Segregation

Four measures of intra-city of segregation are used in the regressions. The first consists of 1976 census residential sorting patterns, as described above. The 1976 data show a photograph of La Paz/El Alto prior to the wave of

rural to urban migration that doubled the population in the period between censuses. The pre-migration data allow for testing whether longstanding residential segregation is systematically different from migration-induced segregation. The 1976 data show a wide variation between neighborhoods. The average neighborhood had an Aymara-language speaking population of 74 percent, with a maximum of 100 percent and a minimum of 33 percent. Neighborhoods in El Alto are noticeably more Aymara-speaking than in La Paz. Second, 1992 census residential sorting patterns are used. The period between censuses is marked by rapid demographic growth, the creation of new neighborhoods, and the loss of the Aymara language. The average neighborhood in 1992 included an Aymara-speaking population of approximately 56 percent, 18 percentage points less than in the previous census period. The distribution of Aymara speakers across neighborhoods also decreases to a maximum of 81 percent and a minimum of 25 percent. Seven neighborhoods that did not exist in 1976 are included in the 1992 sample. Third, demographic proxies are used for geographic segregation. Data on population density provide an instrument for geographic and topographical accidents that separate neighborhoods by mountain ravines, gorges and rivers in the La Paz/El Alto metropolitan area. Highly dense areas are found in both high- and low-income neighborhoods across the metropolitan area. Fourth, a shock-induced proxy is used for randomized residential settlements in the 1980s. The 1982–85 drought induced a wave of rural migrant households that settled in 36 of the 43 neighborhoods sampled in the survey. The impact of this scattered residential settlement pattern on income and schooling outcomes is analyzed.

Income Outcomes

Table 2.4 presents four specifications for the basic residential segregation model. The first specification presents basic OLS results using a 1992 residential sorting proxy for segregation. The expected signs are found for education, experience, experience squared and sex variables. It is also found that Aymara speakers are significantly worse off than non-Aymara speakers, and that living in an Aymara neighborhood is negatively correlated with income. Aymara speakers who live in predominantly Aymara-speaking neighborhoods (Aymara*segregation) are significantly worse off than those who live in less segregated neighborhoods.

Table 2.4. Income OLS and IV Results

	(1)	(2)	(3)	(4)
	OLS 1992 Segregation pattern	IV1 1976 Segregation pattern	IV2 Geographic segregation proxy	IV3 Randomized settlement proxy
Var dep.	**lny**	**lny**	**lny**	**lny**
Schooling	0.063	0.130	0.089	0.124
t-stat	4.966	8.217	3.25	2.884
Experience	0.033	0.073	0.020	0.085
t-stat	1.905	2.407	1.93	2.271
Experience sq.	–0.000	–0.001	–0.000	–0.001
t-stat	–1.568	–1.369	–2.01	–1.102
Female	0.015	0.025	–0.401	0.048
t-stat	1.59	2.014	–3.24	3.663
Aymara	–0.009	–0.037	–0.041	–0.078
t-stat	–9.247	–2.146	–2.691	–5.618
Aymara*Segreg	–0.001	–0.046	–0.002	–0.040
t-stat	–2.146	–5.418	–1.67	–2.975
Segregation	–0.024	–0.008	–0.034	–0.029
t-stat	–1.874	–3.096	–2.09	–1.978
Parental social capital				0.011
t-stat				1.269
Parental schooling				0.148
t-stat				0.256
Own social capital				0.269
t-stat				2.780
Constant	1.246	0.003	3.569	0.591
t-stat	6.899	0.022	(2.76)	0.735
No. of observations	488	488	546	489
F	154.09	21.73	30.43	30.01
Prob > F	0	0	0	0
R^2	0.2258	0.1937	0.305	0.289
Root MSE	0.636	0.934	0.960	0.932

Source: Authors' survey, 2001.

The second specification partially controls for self-selection by instrumenting segregation with a historical measure of segregation from 1976. The resulting historical measure is even more significant than the present-day measure and exhibits the same signs. Individuals living in what was a

highly segregated neighborhood in 1976 are significantly worse off than those living in a historically non-segregated neighborhood. Aymara speakers living in what was a segregated neighborhood in 1976 are also significantly worse off than those living in a historically non-segregated neighborhood. The 1976 segregation data provide a useful but limited instrument for present-day segregation. Historical segregation is likely to be correlated to the historical income distribution that is behind the present-day self-selection problem. It is therefore necessary to test for additional instruments of residential segregation.

The third specification instruments segregation with neighborhood level population density, a proxy of geographic segregation. In the La Paz/El Alto metropolitan area, population density is relatively spread out across high- and low-income neighborhoods, and geographical segregation is negatively and significantly related to income.

The fourth specification instruments segregation with a proxy for random residential settlement. Drought-induced migrants resided in 36 of 43 neighborhoods in the La Paz/El Alto metropolitan area. To the extent that drought migrants settled equally in high- and low-income neighborhoods, they provide an interesting control group against non-migrants; the results indicate that migrants' residence in segregated neighborhoods is negatively and significantly correlated with income. Also included are own and parental human and social capital variables that control for intergenerational endowments of income-generating skills and networks. Controlling for omitted variables confirms the previous findings: residential segregation has a negative and significant effect on income.

Education Outcomes

Table 2.5 shows results for three specifications on determinants of schooling. The first regression is a standard OLS specification, which shows that income, parental education and being female are significantly and negatively correlated with schooling attainment. Speaking Aymara and living in a segregated neighborhood are also significantly and negatively correlated with schooling attainment. The second specification instruments segregation with a historical measure of segregation. The significance of language segregation and historical segregation are slightly lower than in the OLS specification, but they are also negative and significant. Finally, a third speci-

Table 2.5. Schooling OLS and IV Results

	(1)	(2)	(3)
	OLS 1992 Segregation pattern	IV1 1976 Segregation pattern	IV2 Randomized settlement proxy
Var dep.	**Schooling[1]**	**Schooling[1]**	**Schooling[1]**
ln y of household	0.002	0.002	0.011
t-stat	3.617	3.553	10.715
Father's schooling	0.031	0.031	0.033
t-stat	16.107	15.425	6.047
Mother's schooling	0.002	0.000	0.017
t-stat	0.598	0.22	0.696
Sex	0.024	0.023	–0.018
t-stat	1.056	1.007	–1.227
Currently employed	–0.013	–0.013	–0.081
t-stat	–4.32	–4.181	4.576
Language	–0.050	–0.017	–0.013
t-stat	–10.69	–8.229	–8.445
Segregation	–0.001	–0.001	–0.064
t-stat	–2.839	–2.37	–2.672
Language*Segreg.	–0.000	–0.000	–0.003
t-stat	–0.667	–0.174	–1.964
Constant	–0.098	–0.127	–0.706
t-stat	–1.385	–1.57	–0.692
No. of observ.	1,976	1,976	1,976
Fisher	24.42	23.94	28.16
Adjusted R^2	0.245	0.274	0.287

[1] Where education is the education adequation index: EAI = (years of schooling / (age – 6)).
Source: Authors' survey, 2001.

fication instruments segregation with a proxy for randomized settlement. This specification also supports the results observed in the first two specifications. In each case, segregation is a significant barrier to schooling attainment, controlling for parental and own characteristics.

Hypothesis Testing

In order to control for intra-city mobility and test hypotheses concerning positive and negative effects of residential segregation, regressions are run

on young and old cohorts and migrants and non-migrants. The basic rationale for dividing the sample by age and mobility is to analyze whether the negative link between segregation and income generation holds across different subgroups of the population, particularly those who have not yet made the decision to reside in a certain neighborhood.

Two surprising results are presented in Table 2.6. First, residential segregation has a *positive* and significant effect on migrant income. The relationship reverses for second-generation migrants and non-migrants. It is

Table 2.6. Effects of Residential Segregation by Age and Migrant Status

	Migrant status		Age groups	
	Non-migrant	Migrant	25 and younger	Older than 25
Dep. var.	lny	lny	lny	lny
Schooling	0.134	0.088	0.137	0.097
t-stat	4.232	3.976	3.535	5.098
Experience	0.032	0.045	0.406	0.069
t-stat	–2.223	1.801	2.074	2.374
Experience sq.	–0.002	–0.006	–0.006	–0.000
t-stat	–1.226	–1.232	–1.082	–1.424
Female	0.019	0.017	0.027	0.0138
t-stat	3.878	2.026	1.682	4.961
Aymara	–0.207	–0.052	–2.892	–0.041
t-stat	–1.905	–4.157	–3.041	–5.389
Aymara*Segreg	–0.034	0.064	–0.216	–0.054
t-stat	–2.643	2.083	–2.637	–4.108
Segregation	–0.059	0.045	–0.018	–0.031
t-stat	–3.167	2.057	–1.237	–3.34
Parental social capital	–0.005	0.011	0.012	–.0909
t-stat	–1.009	5.207	0.995	–3.206
Parental schooling	0.107	0.121	0.146	0.116
t-stat	2.917	2.148	3.28	3.401
Own social capital	0.0743	0.128	–0.054	0.115
t-stat	4.132	1.677	–1.449	2.056
_cons	–0.604	–0.147	–0.558	–0.285
t-stat	–1.454	–0.554	–0.381	–1.32
No. of observ.	249	191	120	320
F	7.96	18.29	170.33	29
Prob > F	0	0	0	0
R^2	0.317	0.227	0.135	0.238
Root MSE	0.877	1.023	0.871	0.969

Source: Authors' survey, 2001.

hypothesized that this positive effect works through social capital networks established by recent migrants, which provide access to informal labor markets. As migrants settle, and the positive effects of informal networks wear out, the negative peer effects identified by Cutler and Glaeser (1997), along with low levels of human capital and social networks linked to lower skilled labor markets, may reverse the impact of residential segregation over time.

Second, ethnic-based residential segregation also has a *positive* effect on younger age groups (younger than 25 years old). The sign switches once the regression is run on older groups. As with the previous specification, it is hypothesized that residential segregation might provide young workers with ethnic-based peer advantages over young workers from non-segregated neighborhoods. In addition, parental schooling is strongly and positively correlated with young workers' income. Controlling for intergenerational human capital effects is intended to isolate the net effect of residential segregation on the income-generating capacity of younger age groups.

Research and Policy Implications

Although research on poverty and income distribution in Bolivia has reported differential wage effects between indigenous and non-indigenous households,[6] few studies have focused on the effects of residential or neighborhood segregation. This chapter has considered the marginal effects of living in a segregated indigenous neighborhood for urban indigenous residents.

Three results stand out in the analysis. First, there are significant and negative effects on income generation in both across-city and intra-city comparisons. Second, individual and neighborhood-level interactions between ethnicity and segregation by ethnicity are significantly and negatively correlated with income. Finally, positive "social capital effects" are found for migrants and young workers and "negative human capital effects" for second-generation migrants, non-migrants and older workers living in segregated neighborhoods.

[6] See Chiswick, Patrinos and Hurst (2000), Rivero and Jiménez (1999), Vos, Lee and Mejía (1998), Pérez de Rada (1997), and Jiménez and Yañez (1997).

The analysis has aimed to test whether residential segregation has any measurable effect on income outcomes and to explore alternative causal mechanisms whereby segregation results in costs and benefits for urban households. The first finding is that residential ethnic segregation has a negative and significant effect on personal income. To control for potential self-selection biases, this relationship is tested with three types of instruments: a measure for historical residential sorting patterns, a proxy for geographic segregation, and a proxy for randomized residential patterns. The three instruments are significantly and negatively related to personal income. The second finding, determined through three additional specifications run on schooling attainment, is that parental education, household income and being female are significant determinants of schooling attainment. Segregation, as measured by the OLS and IV specifications, correlates negatively and significantly with schooling. In addition, the effect of potential omitted variables such as intergenerational human and social capital endowments is tested, and both have significant positive effects on income and schooling attainment.

Alternative explanations for these results are then considered. Residential segregation has a *positive* and significant effect on income for recent migrants and a *negative* and significant effect on second-generation migrants or non-migrants. Positive social capital may explain the beneficial effects of ethnic segregation for first-generation migrants who gain a foothold by moving to Aymara neighborhoods close to relatives and migrant peers. This advantage disappears over time, as first-generation migrants either leave the neighborhood or stay and face negative peer effects. Residential segregation also has a *positive* and significant effect on younger age cohorts, particularly workers under the age of 25. This might be explained by differences in labor force participation between poor and non-poor groups. Labor participation tends to be high for the poor, which might explain why the income gap lessens for young and older workers as a whole. Both of these findings are premised on the possibility of controlling for self-selection, omitted variable bias, and identification problems that might emerge from intra-city and across-city mobility. A careful examination of the available data has made it possible to correct for the most important omitted variable and self-selection biases. From a research perspective, this quantitative analysis contributes to a literature that has almost exclusively depended on sociological and anthropological evidence in the past.

From a policy perspective, the analysis of residential segregation provides insights into the current debate on Bolivia's shifting poverty profile. First, the urban poor now outnumber the rural poor. As urbanization accelerates and the demographic profile changes, a window of opportunity arises to achieve gains in savings and productivity. However, a growing and younger population requires complementary opportunities to ensure productive insertion into labor markets. Of particular importance are services in human capital, training and public services. Residential segregation may decrease opportunities for human capital development and lower the quality of public service access. The measurement of segregation effects will provide a yardstick for assessing the costs of policy-induced segregation. Second, the analysis of residential segregation provides a policy handle on the more general problem of exclusion and discrimination. While social and political discrimination has been hard to identify, let alone address, residential segregation is amenable to an array of policy interventions, including urban planning and development norms and school and health district mapping.

This page intentionally left blank

CHAPTER THREE

Social Exclusion and the Two-Tiered Health Care System in Brazil

Denisard Alves
Christopher Timmins

In Brazil's two-tiered system of health care, those with sufficient means, or whose employers provide health coverage, have access to a private health system that provides quality treatment on demand. The rest of the population relies on a system of public clinics and hospitals. Like most public health care systems around the world, the Brazilian system is characterized by long waiting times and questionable quality, with the practical implication that those who are forced to rely on the system spend more time being sick.

Examining this two-tiered system is particularly relevant in Brazil in light of recent changes in the country's socio-demographic structure. In 1990, only 6.7 percent of Brazil's population was over age 60, but by 2010 this is expected to be 9.7 percent and by 2030, 16.9 percent (World Bank, 1994 and 2000). During the last twenty years, family sizes among the poorer segments of Brazilian society (i.e., those who typically rely most on the public provision of health care), have been larger than in wealthier segments of society. This large population group has been aging, and it is nearing a time when its health care needs will grow rapidly (Cutler and Meara, 1998). Concerns have been raised that the Brazilian public health system will not be up to meeting this growing demand. In particular, already-long waiting times for treatment will continue to increase, which essentially will mean that many of the poor will receive no health care at all. This mechanism of social exclusion of the poor, elderly and rural population will increase at the rate at which this segment of the Brazilian population grows.

Long-run feedback effects may result from these trends. Growing health care demand due to increasing numbers of poor and elderly, as well as the

increasing cost of treatment for a limited supply of public health services, means that the poorest segments of Brazilian society will begin to lose access to health care. This will diminish the health stock of the poor, reinforcing their socioeconomic position. To the extent that the poor continue to have larger families (e.g., as a retirement-insurance mechanism or a source of labor for subsistence agriculture), this will lead to further strains on the public health care system in the future and increase the likelihood of further failures.

The Brazilian Health Care System

The health care system in Brazil is based on the premise that individuals and households should, at the most basic level, be protected by the public sector. Therefore, while many Brazilian citizens rely on private sources of health care provision, the system also includes a very large public component, which is intended to act as a "safety net." The public system, which has its own hospital facilities and is federally financed by the Unified Health System (Sistema Único de Saúde—SUS), is intended to provide resources to meet the demands of those who do not have private health care coverage.

The need for such a system in Brazil is very real. Public health figures show the persistence of endemic diseases, leading to an annual mortality rate of 0.6 percent (approximately 1 million people). According to a United Nations estimate, the SUS provided for 12.6 million hospitalizations in about 2,000 public and private hospitals in 1995, and 1.2 billion consultations in out-patient clinics. Similarly, there are some 507,000 hospital beds in Brazil (about 1 per 300 population), of which about one-third are in public establishments. Alves (2000) shows that, in the state of São Paulo, 48.9 percent of the health services used in 1998 were paid for by the SUS, 6 percent paid for directly by users, and 45.1 percent pre-paid.[1] Among the total population of the state of São Paulo, 44.2 percent have some type of health insurance coverage, while 55.8 percent have none. Those without health insurance coverage have to rely entirely on the SUS. For the city of São Paulo, located in the

[1] The principal pre-paid health care services in Brazil are group-medicine plans, medical cooperatives, or corporate plans. All offer pre-paid services for individuals, families and other entities such as businesses and unions.

more developed southern area of Brazil, close to 40 percent of the population is cared for by the public health care system. In the rest of Brazil, over 50 percent of the population relies on this system.

In terms of its structure, the SUS is a decentralized system that provides health care by districts down to the municipal level. The Municipal Health Secretary has jurisdiction over SUS health centers, hospitals, laboratories and services, and administers funds provided by the federal, state and municipal governments. The Municipal Health Council consists of organization members, citizens and school personnel. As it is the council that approves health programs, establishing this type of entity is necessary for the municipality to receive federal funds. Districts and municipalities can cooperate in sharing expenses and resources with the aim of enhancing health care efficiency.

The history of the SUS goes back to the Brazilian Constitution of 1988, which, as a result of a longstanding social movement to counter the inequitable health policies of past governments, mandated free and universal health care. Previous military regimes in Brazil had left a health care system that was highly centralized and with little capacity—in effect, a system that was extremely unresponsive to Brazil's local needs and regional diversity. In the 1970s, inflation was rampant and the Brazilian economy was in crisis. A severe recession followed in the early 1980s, and public health care expenditures fell substantially, driving down the quality of both health services and infrastructure. In response to this, a council consisting of federal ministry representatives undertook the *Reforma Sanitaria,* a health care reform effort. The first phase of this reform was the Integrated Health Actions, an effort to improve coordination and decentralize service delivery from the Ministry of Health and INAMPS (Instituto Nacional de Assistência Médica da Previdência Social), the main financing mechanism, to the state and municipal levels. The second phase of the reform was the creation in 1987–88 of SUDS (Sistema Unificado e Descentralizado de Saúde), which completed the process of decentralization. A new constitution in 1988 paved the way for the creation of the SUS, the third and final phase of the reform. Since then, the SUS has begun to contract out most patient care to a network of private and philanthropic hospitals, clinics, and other facilities. Alves (2000) estimates that 20 percent of the available hospital beds in Brazil used by the SUS in 1998 belonged to private hospitals. The government itself owns just 31 percent of the hospital beds it supports and has been gradually decen-

tralizing the control of publicly owned facilities. This trend towards decentralization and privatization in Brazilian health reform will be important to the policy conclusions discussed below.

Private health care in Brazil has grown rapidly, and about 26 percent of Brazilians are covered by such plans. The plans vary a great deal in terms of price and quality, but usually exclude expensive, catastrophic conditions, leaving them to be covered by the public system. The private plans are subject to almost no regulation.

While the government is theoretically responsible for health care in Brazil, public expenditures in this area represent only about 4.2 percent of GNP each year, a low figure relative to some of Brazil's neighbors. In comparison, the United States allocates about 12.7 percent of its GNP to health care, France about 9 percent, El Salvador about 6 percent, and Paraguay 3 percent. This low level of spending has led to many questions about the quality of publicly provided health care in Brazil. Less than $80 per capita was spent on health care in Brazil last year, whereas the corresponding per capita expenditure in Argentina was closer to $300, and in the United States, $2,300. In all, Brazil spent almost $16 billion on health care in 1995, of which $2.7 billion was used to pay staff and $2.9 billion was used to cover old loans.

Alvarez (1998) states that "the private sector in Brazil currently does a reasonably good job of providing for the 25 percent of the population that its plans and services reach. However, the public sector SUS is doing a poor job of servicing the other 75 percent." Considerable anecdotal evidence and some empirical evidence illustrate just how ineffective the SUS may be. For example, the world record for hospitalizations occurred in Campo Grande do Sul in the State of Paraná in Brazil, where over 60 percent of the population was hospitalized in one year. An audit of the system revealed that about 24.1 percent of all diagnoses were inaccurate.

Detailed data from the 1998 household survey (Pesquisa Nacional por Amostra de Domicílios—PNAD), which are described in the following section, make it possible to look specifically at individual perceptions of health care quality. In particular, the PNAD questions respondents as to whether they sought health care during the previous two weeks, and if so, whether they were able to obtain it. If a respondent was unable to get the health care necessary, survey questions determine whether this occurred because there were no doctors at the health care facility, equipment was malfunctioning, or for other reasons. Table 3.1 divides the sample of heads-of-households in

Table 3.1. Perceptions of Health Care Quality by Type of Coverage

Region	Health care type 0 = SUS 1= Private	% who received health care sought	Reason for failing to receive health care (for those who sought care during previous two weeks)					
			No vacancy	No attending doctor	No attending expert	Malfunc-tioning equipment	Had to wait too long	Other
North	0	93.6	31.1	53.3	4.4	4.4	4.4	4.4
	1	97.7	0	50	0	0	25	25
NE	0	93.8	39.5	29.7	10.3	4.9	5.4	10.3
	1	98.2	15.4	7.7	15.4	0	7.7	53.9
CW	0	94.1	50	22.1	8.8	1.5	7.4	10.3
	1	96.8	50	40	10	0	0	0
SE	0	95.4	43.1	28.8	13.1	5	3.8	6.3
	1	98.8	52.9	23.5	0	0	0	23.5
South	0	94.7	61.6	17.2	3	0	4	7.1
	1	99.1	33.3	16.7	16.7	0	0	33.3
All	0	94.6	45.6	28.5	9.3	3.6	4.9	8
	1	98.5	36	24	8	0	4	28

Source: 1998 PNAD.

the PNAD into those who use the private health care system and those who rely on the public system, then examines their answers to each of these questions. The results indicate the clear superiority of the private health system to the SUS system. For the country as a whole, the private system had shorter waiting times than the public system. More respondents who were unable to obtain necessary treatment cited "no vacancy" in hospitals and "no attending doctors" for the public system than for the private one. Other attributes of the quality of the health systems, as perceived by patients, are similarly unfavorable towards the SUS. At the regional level the picture is similar, although less reliable due to the smaller number of observations.

A World Health Organization report in June 2001 ranked Brazil's health situation as 125th out of 191 countries worldwide, comparable to that of lesser-developed countries such as Nepal, Cambodia and Vietnam. Brazil's low ranking is attributed primarily to the inequitable manner in which the health care system is financed. WHO generally concludes that Brazilians have to pay too much for health care, and that the poor suffer the most. There are

huge income disparities in Brazil: 35 million of the 120 million Brazilians without private health insurance are below the poverty line.

According to a 1998 report by the World Bank Operations Evaluation Department, Brazil's health system might appear to be efficient (it substantially separates financing from the provision of services), but it is inching toward crisis. The public system is underfinanced and therefore exhibits severe regional inequalities, rationing, and declining quality.

The report further states: "The hyperinflation of the late 1980s and early 1990s and the irregular flow of resources to health have contributed to the evolution of a fee structure for medical treatment that has not kept pace with costs, and payment can be sporadic. Doctors frequently must work at several sites to make ends meet. Stories of long lines for hospital services, mistakes in emergency care, strikes and walkouts by medical professionals, arbitrary triage, and other crises are reported daily in the press."

This system covers about 70 percent of inpatient and outpatient care. However, while local governments are given responsibilities under the decentralization of the public health care system, they do not necessarily have the resources or incentives to deliver cost-effective services. Expenditures do not target the poor, and institutions are extremely fragmented and expensive. Brazil has one of the lowest ratios of nurses to doctors in the developing world (0.33:1 in 1996) and an average of only 13 doctors per 10,000 population. Moreover, the SUS has brought with it a decrease in physician pay. According to the National School of Public Health at Brazil's Oswaldo Cruz Foundation, doctors at the old National Institute would have received a much higher salary there than they do now, even though they still perform the same functions and maintain the same caseload and hours.

In summary, in Brazil and in most Latin American countries, the public sector still has a major responsibility for providing health services.[2] The Brazilian government uses general revenues to pay for the health care of middle- and low-income population groups, while upper-income groups tend to use their own resources to pay for private health care. Those with low income and education, as well as members of certain racial groups, tend to be discriminated against by the system. Their access to public health care is more limited, waiting times in lines are longer, as are travel distances, and these groups lack the financial resources to use the private system without

[2] See Parker and Wong (1997) and Wong (2000) for an analysis of the Mexican health system.

inordinate sacrifices of other consumption. Given this two-tiered system of service provision, use of the system would be expected to vary widely across the population, an expectation supported by literature indicating that utilization of health services differs vastly by income groups in Brazil (Alves, 2000).

Data

The 1998 PNAD is an annual household survey on socioeconomic conditions in Brazil carried out by the Instituto Brasileiro de Geografia e Estatística (IBGE). The 1998 PNAD has a special supplement dealing with health conditions, making it particularly suitable for this analysis. The survey covers 344,975 individuals and 98,166 households. In the present analysis, however, the number of observations is smaller due to missing values for some important variables, particularly income per person in the household. The use of this survey data presents a number of advantages. First, it is one of the only surveys to collect data on the health of the population in a consistent fashion; and second, data collection on health is accompanied by a full set of socioeconomic data on the individual and the household. Use of the PNAD also presents some disadvantages, however. The collection of such data relies on the training of data enumerators to record health status. Thus, when looking at reported illnesses, accuracy might not be as great as it would be if medical professionals were examining the individuals and reporting their illnesses. In addition, data for northern Brazil are restricted to the urban sector and therefore are not representative of the entire region. The full data set, however, thoroughly covers the remaining regions of the country, and the exclusion of the rural north does not harm the survey's representativeness, since this population is very small. The present analysis uses the household data set instead of the larger individual sample data. Decisions on health care utilization are major decisions made at the household level, and using individual data would likely introduce correlations between household members.

Table 3.2 shows some major characteristics of the data set. The variable *With_HP* measures the proportion of household heads that have a private health insurance plan for which they pay themselves or their employer pays. This amounts to 17 percent of total household heads. The remaining

Table 3.2. Data Summary, Household Heads
(N = 98,166)

Variable	Mean	Variable	Mean
Sex	0.727	Age	44.524
Educ	6.612	Age60	0.176
Income	338.591	f_kind2	0.284
Black	0.068	PrivHP	0.069
White	0.526	whopay	0.101
Mixed	0.400	Arthrit	0.148
Yellow	0.004	Cancer	0.004
Urban	0.831	Cardiac	0.074
Metro	0.413	Cirrhose	0.003
dcwest	0.109	Backache	0.312
Dseast	0.342	Depress	0.078
Dsouth	0.178	Diabets	0.037
Dnorth	0.068	Hipert	0.197
Dneast	0.302	Kidney	0.047
With_HP	0.170	Respir	0.042
Migrator	0.009	Tendon	0.032
Value	79.474	Tuberc	0.002
Attend	0.131	Healthy	0.491
Characteristics of private health plan (N = 6,639)			
plcons	0.980	plexam	0.956
pllist	0.925	plinter	0.929
plreemb	0.310	platend	0.805
plother	0.811	plmedic	0.045
pldent	0.219	odonto	0.031
paymore	0.178		

Source: 1998 PNAD and authors' calculations.

households use the public health system. The variable *PrivHP* indicates whether a household pays for its own health plan, and *whopay* indicates that a household has a private plan that is paid for by an employer. The average household income in the sample is R$ 962.50 monthly. The average payment of health insurance for the 6.9 percent who pay their own medical insurance is R$ 150.31, and their average household income is higher than the sample average (income reported in Table 3.2 is the average household per capita income). The average per capita income for the households paying for a private health plan is twice the average per capita income of the household sample.

The mean of the *Metro* variable gives the proportion of the household heads living in Brazilian metropolitan areas. *Age60* is a variable that specifies whether the head of the household is over 60 years of age, while *Child_14* indicates the proportion of persons below age 14 living in the household. The race-differentiation variables are *White*, *Mixed*, *Black* and *Asian*. The proportion of blacks is quite small, although it should be noted that a large portion of the people who include themselves in the mixed racial group would generally be perceived as black. The proportion of households not reporting illness is 49.1 percent. The remaining 50.9 percent reported the occurrence of at least one illness, and the proportion of people seeking medical treatment in the last two weeks is 13.1 percent. Table 3.1 reports the results of users' evaluations of the quality of the health care sought by those households. Illnesses are self-reported by the head of the household.

The characteristics of the private health insurance plans are presented at the bottom of Table 3.2. The health plan attributes are defined by a set of dummy variables. The variable *plcons* takes a value of one when the health plan allows for the prior appointment of a doctor consultation and zero otherwise. Table 3.2 shows that 98.04 percent of the health insurance plans cover pre-appointment for visits to doctors. *Pllist* is one when the health insurance policy presents a list of authorized doctors, hospitals and laboratories that can be used by the policyholder and zero otherwise; 92.47 percent of the private health insurance policies present a list of authorized doctors, hospitals and laboratories. The value for the variable *plreemb* indicates that 30.99 percent of the health plans permit reimbursement of medical expenses when the individual is attended by doctors or health centers not affiliated with the health plan. *Plother* indicates that 81.15 percent of policyholders may be attended by doctors, hospitals and laboratories in cities other than the one in which they reside, and *pldent* indicates that only 21.94 percent of health insurance plans cover dental treatment. This attribute is clearly not a widespread characteristic of private health insurance, and plans with this attribute are more expensive than those without. *Paymore* is a variable capturing the fact that some health insurance policies impose a ceiling on what they pay for health care expenses, and any expense above this limit has to be paid by the policyholder.[3] The variable *plexam* indicates

[3] This is the only attribute among the 11 described here for which the dummy variable assumes the value of one when it indicates a detrimental characteristic of the health plan.

that 95.60 percent of health plans allow the policyholder to take complimentary lab exams during treatment. Among private policyholders, 92.83 percent are covered for hospitalization, as indicated by the variable *plinter*. *Platend* indicates that 80.54 percent of policyholders are attended by medical services under contract with their heath insurance company. Very few health plans cover the acquisition of medicines and drugs. *Plmedic* indicates that this attribute is very special and covers only 4.85 percent of private health insurance holders. Among health insurance holders, only 3.07 percent of plans cover orthodontic treatment. This aspect of the health insurance is represented by the variable *odonto*.

This detailed list of private health plan attributes will prove valuable in the following analysis in that it makes it possible to impute prices for a "standardized" private health plan (i.e., a simplified plan without any of the "bells and whistles" described above) for every member of the sample, regardless of whether they actually bought a private health plan (of any type). Calculation of these imputed prices will be necessary for determining the welfare consequences for each individual of facing an increase in the shadow price of public health care.

Research Methodology

In order to characterize indirect exclusion from private health care in Brazil, a two-pronged methodological approach is adopted. First, the detailed survey information in the 1998 PNAD data set is employed to determine which groups in Brazilian society have access to private health insurance and which rely on public health care. Being relegated to public health care is not a direct form of exclusion, but rather one based on relative prices for private and public health care that may be different for individuals from different segments of society. Moreover, differences in employment patterns for individuals from different socioeconomic groups will influence their access to employer-provided private health insurance.[4]

[4] This difficult issue—health insurance as an attribute of a job for which an individual may or may not face a corresponding reduction in pay—is avoided by considering only those individuals who either buy private insurance directly (i.e., those who do not receive it through an employer) or use the SUS.

Public vs. Private Health Care and Socio-demographic Groups

Certain groups delineated by race, education and location in Brazil are expected to be systematically more reliant on the public health care system. The question of which groups fall into this category is determined with a simple probit regression (Greene, 2000) as follows:

$$P(y_1 = 1) = \Phi(X_{i,j} \_ \beta) \quad (1)$$

where

y_i = form of health care coverage for individual i (1 = private, 0 = public) and X_i = socioeconomic attributes of individual *i,* including race (black, Asian, mixed, white), age, education as defined by years of schooling, household income, and regional indicators (percentage of persons in the household under 14 years and over 60 years of age, and household income, as defined by the total sum of wages and other types of income of individuals living in the household).

The results of this regression reported in Table 3.3 correspond to general perceptions about Brazilian health care. Those who rely more on the public system are less educated, female, come from the black and mixed racial groups and from the northern and center-western regions of Brazil, have lower incomes, and are elderly. The presence of people over 60 years of age in the household is highly significant, while the presence of people below 14 years of age does not make a difference in the household's health

Table 3.3. Probit Regression PrivHP
(N = 82,900, Log likelihood = –16,753.661)

Variable	Estimate	Standard error	Variable	Estimate	Standard error
Male	0.0051	0.0173	Southeast	0.3056	0.0345
Education	0.1173	0.0020	Northeast	0.0823	0.0356
Income	0.0003	0.00001	South	0.0411	0.0374
Black	–0.4785	0.0855	Age	0.0233	0.0008
White	–0.1862	0.0785	Age 60	–0.2696	0.0290
Mixed	–0.4284	0.0799	Child_14	–0.0195	0.0206
Center-West	–0.0668	0.0409	Constant	–3.3594	0.0964

Source: 1998 PNAD and authors' calculations.

insurance decision. Given the results described in Table 3.1 regarding differences in the quality of health care across providers, this alone could be considered evidence of exclusion of these groups.

Table 3.4 provides additional evidence along these lines. Specifically, it presents the results of a number of probit regressions in which a dummy variable indicating that an individual has suffered from a particular disease (e.g., depression, arthritis, cancer, diabetes, respiratory ailment, hypertension, cardiac disease, tuberculosis, cirrhosis, tendonitis, and kidney disease) is regressed on a set of individual attributes, including the form of health care provision (i.e., SUS vs. private) that the individual uses. The idea here is that an individual's health stock, which determines how likely he or she is to suffer from any of these ailments, is in part determined by the effective quantity of health care used. This is directly a function of the quality of that care, as well as its (shadow) price. An individual who relies on the public system might therefore receive lower quality care, or less care in general (if waiting times for treatment are longer), leading to a lower health stock and a higher likelihood of disease because the quality of public health care is low, and it is difficult to access effective health care units. These results should be interpreted with extreme caution, however, as the form of health care provision may be simultaneously determined with the individual's health stock; e.g., individuals who know they are likely to develop cancer might purchase private health insurance in order to guarantee a higher quality of care. The presence of an endogenous variable in a probit regression can potentially lead to inconsistent estimates of all the model's parameters.

The results, however, are very much consistent with a priori expectations about health and about the Brazilian health system. Women are generally less likely to suffer from almost all diseases except cirrhosis; that is, women seem to be healthier than men. Higher income follows good health, meaning that poor people are more likely to suffer from some of the illnesses defined in the PNAD survey. One important point to note, however, is that people with private health plans, either self paid or employer paid, are more likely to suffer from some of the illnesses, while healthier people are less likely to go to a private paid health plan. This might simply be the result of getting better diagnoses from a private health care provider than from a public clinic. Certain illnesses seem more likely to strike the northeastern region, while people from the southern and southeastern regions are more likely to report cardiac disease, cancer, depression and respiratory diseases.

Table 3.4. Probit Regressions, Determinants of Diseases

Variable	Arthritis	Cancer	Cardiac	Cirrhosis	Backache	Depression
Sex	–0.3636	–0.0801	–0.2590	0.3099	–0.1864	–0.5768
	(29.867)*	(–2.074)**	(–18.250)*	(5.621)*	(–18.116)*	(–42.920)*
Age	0.0323	0.0168	0.0270	0.0096	0.0215	0.0129
	(52.069)*	(7.935)*	(36.155)*	(4.333)*	(45.530)*	(19.772)*
Educ	–0.0461	0.0051	–0.0124	–0.0275	–0.0368	–0.0050
	(–27.706)*	(0.978)	(–6.447)*	(–4.357)*	(–29.233)*	(–2.807)*
Income	–0.00005	–7.73e–06	–0.00002	–0.00003	–0.00006	–0.00006
	(–4.307)*	(–0.263)	(–1.998)**	(–0.591)	(–7.246)*	(–5.194)*
Black	0.2166	3.714	0.3320	0.0048	0.1259	–0.1970
	(2.308)**	(22.035)*	(3.214)*	(–0.014)	(1.796)	(1.789)
White	0.2207	3.1977	0.1884	0.0223	0.1608	0.2730
	(2.406)**	(25.526)*	(1.866)	(–0.069)	(2.357)**	(2.536)**
Mixed	0.2545	3.8188	0.1967	0.1196	0.1610	0.2719
	(2.764)*	(25.907)*	(1.938)**	(0.367)	(2.349)**	(2.515)**
Age60	–0.1085	0.0302	–0.0364	–0.2681	–0.2775	–0.1887
	(–5.474)*	(0.466)	(–1.560)	(–3.568)*	(–16.240)*	(–8.087)*
f_kind2	–0.0760*	–0.1978	–0.1109	–0.1288	–0.0064	–0.1018
	(–4.529)	(–2.934)*	(–5.304)*	(–2.320)**	(–0.545)	(–5.565)*
dcwest	–0.3241	0.2203	0.0331	–0.0873	–0.1318	–0.0085
	(–12.958)*	(1.935)**	(1.035)	(–1.053)	(–6.310)*	(–0.284)
dseast	–0.5745	0.2412	–0.0355	–0.01793	–0.2450	–0.0581
	(–26.363)*	(2.337)**	(–1.268)	(–2.450)**	(–13.413)*	(–2.230)**
dneast	–0.3596	0.0751	–0.1752	–0.2243	–0.0930	–0.0772
	(–17.034)*	(0.714)	(–6.222)*	(–3.123)*	(–5.168)*	(–2.988)*
dsouth	–0.3832	0.3750	0.0631	–0.1592	–0.1983	0.0590
	(–16.065)*	(3.530)*	(2.087)**	(–1.911)**	(–9.918)*	(2.094)**
With_HP	–0.0434	–0.0383	0.0430	0.0266	0.0135	0.0002
	(–2.428)**	(–0.708)	(2.177)**	(0.409)	(1.011)	(0.008)
N	95,565	95,565	95,565	95,565	95,565	95,565
Log likelihood	–32,729.57	–2,274.64	–21,995.40	–1,875.49	–55,542.89	–24,502.90

Source: 1998 PNAD and authors' calculations.

Note: Z- statistics are in parentheses.

* Significant at 1 percent level.

** Significant at 5 percent level.

Table 3.4. Probit Regressions, Determinants of Diseases (cont.)

Variable	Diabetes	Hyper-tension	Kidney disease	Respiratory	Tendonitis	Tuberculosis	Healthy
Sex	–0.1779	–0.2975	–0.0095	–0.1998	–0.3358	0.0935	0.3300
	(–10.124)*	(–26.486)*	(–0.563)	(–12.069)*	(–18.811)*	(1.484)	(31.543)*
Age	0.0290	0.0337	0.0130	0.0038	0.01105	0.0079	–0.0325
	(30.030)*	(60.244)*	(16.570)*	(4.890)*	(12.620)*	(2.847)*	(–70.269)*
Educ	–0.0023	0.0150	–0.0333	–0.0121	0.0060	–0.0248	0.0296
	(–0.973)	(–10.564)*	(–15.026)*	(–5.638)*	(2.678)*	(–2.877)*	(24.692)*
Income	0.00003	–0.00002	–0.00009	–0.000005	0.00004	–0.0002	0.00004
	(2.289)**	(–1.885)	(– 4.560)*	(–0.360)	(3.738)*	(–1.589)	(4.684)*
Black	0.1286	0.3107	–0.0305	0.2334	0.1738	–0.1437	–0.2140
	(1.150)	(4.204)*	(–0.266)	(1.833)	(1.264)	(–0.423)	(–3.266)*
White	0.0138	0.0863	0.0148	0.2103	0.1964	–0.3173	–0.1921
	(0.127)	(1.181)	(0.132)	(1.689)	(1.468)	(–0.957)	(–3.024)*
Mixed	0.0162	0.1273	0.0134	0.2115	0.2212	–0.2295	–0.1989
	(–0.148)	(1.732)	(0.120)	(1.691)	(1.645)	(–0.689)	(–3.112)*
Age60	–0.1718	–0.1925	–0.1600	0.0279	–0.1186	–0.2536	0.1571
	(–5.952)*	(–10.469)*	(–5.194)*	(9.076)*	(–3.850)*	(–2.585)*	(8.896)*
Child_14	–0.1546	–0.1099	0.01118	–0.0241	–0.0773	–0.0220	0.0382
	(–5.498)*	(–7.545)*	(0.594)	(–1.157)	(–3.214)*	(–0.316)	(3.452)*
dcwest	–0.0700	0.0909	0.0314	0.0566	–0.1783	–0.1003	0.1243
	(–1.637)	(3.653)*	(1.057)	(1.600)	(–4.644)*	(–0.727)	(6.007)*
dseast	0.0481	0.0794	–0.3460	–0.0146	–0.2185	0.0542	0.2505
	(1.334)	(3.646)*	(–12.628)*	(–0.466)	(–6.657)*	(0.486)	(13.775)*
dneast	–0.0585	–0.0076	–0.4078	–0.1817	–0.2026	0.0503	0.1925
	(–1.607)	(–0.348)	(–14.966)*	(–5.711)*	(–6.229)*	(0.461)	(10.669)*
dsouth	–0.0091	0.0557	–0.2549	0.1740	0.0488	0.0834	0.1482
	(–0.231)	(2.347)**	(–8.443)*	(5.222)*	(1.409)	(0.683)	(7.479)*
With_HP	0.1178	0.0961	–0.0512	–0.0048	0.1397	–0.2064	–0.0677
	(5.065)*	(6.489)*	(–2.153)**	(–0.214)	(6.128)*	(–1.869)	(–5.327)*
N	95,565	95,565	95,565	95,565	95,565	95,565	95,565
Log likelihood	–13,340.13	–40,920.07	–17,382.01	–16,213.18	–12,853.73	–1,107.16	–58,161.94

Source: 1998 PNAD and authors' calculations.
Note: Z- statistics are in parentheses.
* Significant at 1 percent level.
** Significant at 5 percent level.

A Model of Individual Health Care Choice

While describing which elements of Brazilian society are more likely to rely on publicly provided health care, the preceding analysis does not provide any way of measuring the welfare consequences of this indirect form of exclusion. To do that requires developing a more elaborate model that takes into account the fact that individuals optimally choose what form of health insurance to obtain in the face of market prices and a budget constraint. The second part of the empirical analysis develops such a model. The binding constraints on the scope of the conclusions that can be taken away from this model come from the lack of data describing individuals' full income endowments and actual expenditure patterns on health and non-health commodities. Instead, the model takes a stylized view, describing the individual's choice of health coverage as a choice between alternative types of insurance in a static context. To the extent that people change health insurance status during the course of their life, this may bias answers.

In particular, it is assumed that individual *i* chooses his or her form of health care provision in order to maximize a utility function of the form:

$$U(C_i, H_i) = C_i^{a_i} H_i^{1-a_i} \tag{2}$$

subject to a simple budget constraint:

$$C_i + P_i^H H_i 1_i \tag{3}$$

where C_i represents *i*'s consumption of a composite numeraire commodity, H_i represents the consumption of *effective* health care services (the price for these services, P_i^H is allowed to differ by individual, and to reflect the quality of the nominal health care consumed), and 1_i represents the individual's income. P_i^H will also differ according to the form of health care provision chosen; i.e., P_i^S for SUS health care and P_i^P for privately provided health care. The chief source of difficulty in this analysis is that P_i^S is not observed (all SUS health care is nominally free), but is rather only a shadow price of SUS health care consumption.

Utility maximization subject to this budget constraint yields the following indirect utility function:

$$V(P_i^H, I_i) = (\alpha_i I_i)^{\alpha_1} \left(\frac{(1-\alpha_i) I_i}{P_i^H} \right)^{1-\alpha_1} \quad (4)$$

which differs by whether the individual chooses SUS health care ($V(P_i^S, I_i)$) or private health care ($V(P_i^P, I_i)$). Taking the optimal allocation of income between composite consumption and health care as given, individual *i*'s choice between the two forms of health care provision can be modeled as a comparison of these two indirect utility functions. In particular, individual *i* will choose SUS health care as long as:

$$V(P_i^s, I_i) \geq V(P_i^P, I_i) \quad (5)$$

Because of the simple functional forms adopted for the present purpose, condition (5) boils down to whether $P_i^S \geq P_i^P$. The price of an effective unit of public health care is not an observed magnitude; nominally, public health care is free to everyone in Brazil. It has a price, however, in the form of time in and disutility of crowded waiting rooms and other factors, as discussed above. This price would be expected to differ across individuals according to their opportunity cost of time, preferences for cleanliness and quality, and disutility of congestion, i.e., differences for which control may be possible with a set of observable individual attributes (X_i).

The available data make it possible to recover each individual's shadow price for an effective unit of public health care by using inequality (5). Once this is done, all the tools necessary are available for considering the welfare impacts of an increase in the congestion costs associated with receiving health services from the SUS. In particular, assuming that the individual chooses the health care option that maximizes his or her indirect utility (with the individual's perception of the quality difference between public and private provision factored into that price), private health care will be chosen if $P_i^P > P_i^S$.

Estimation of the "Shadow" Price of Public Health Services

P_i^P is observed in available data. P_i^S, on the other hand, is not observed and has to be estimated. In particular, in PNAD data the price is observed of the

private health insurance paid for everyone who opted for that form of coverage. Private health insurance premiums are imputed for the rest of the sample using the Heckman selection model:

$$P_i^p = Z_i\gamma + u_1 \tag{6}$$

where Z_i are dummy variables describing the characteristics of the health plan and γ is a vector of coefficients associated with the matrix of attributes.

P_i^P is not observed for all households, but rather only for those that purchased private health insurance. It is not known, however, which households were more likely to purchase private health plans, even if they had not actually acquired one. The probability can be described by the probit regression in equation (1). Equation (1) thus describes the sample selection component of a Heckman selection model (Greene, 2000). Private health insurance premiums are imputed for the whole sample by estimating equation (6), accounting for selection into private health provision with a Heckman selection correction as specified in equation (1). The results of this model are presented in Table 3.5. After estimating the parameters of equation (6), it is possible to forecast the private premium for a standardized policy for all individuals. In particular, the standardization adopted sets all of the attributes of the health care policy to their simplest values—i.e., to give the price of a policy without any "bells or whistles." This creates a level playing ground for comparison of the individual's decision between public and private coverage (i.e., what the welfare effect of an increase in the price of public health care would be if everyone had the same simple private option to choose as an alternative). Using this procedure, it is possible to estimate a shadow price, P_i^S, which establishes how much each household would pay for the use of the public health system.

The natural logarithm of P_i^S is paramaterized as a linear function of individual attributes (X_i) and an unobservable determinant (ε_i), which is assumed to be identically and independently normally distributed with a variance of σ^2 and a zero mean. The choice of private health coverage is then determined by the following condition being satisfied:

$$\ln P_i^p LE\ X_i \_\beta + \varepsilon_i \tag{7}$$

Table 3.5. Heckman Procedure to Impute the Price of Private Health Services
(N = 82,800, Log likelihood = –27,520.82)

Variable	Estimate	Standard error	Variable	Estimate	Standard error
Male	0.1914	0.0244	pllist	0.0280	0.0386
Education	–0.0382	0.0492	plreemb	0.1252	0.0216
Income	0.00001	0.00001	plother	0.2927	0.0268
Black	–0.0087	0.0557	plcons	–0.7845	0.0625
White	–0.1418	0.0276	plexam	0.1088	0.0559
Asian	0.1438	0.9555	plinter	0.5942	0.0413
Center-West	–0.1110	0.0564	pldent	–0.3128	0.0235
Southeast	–0.2026	0.0484	plmedic	0.0395	0.0440
Northeast	–0.0128	0.0492	odonto	–0.0045	0.0564
South	–0.1704	0.0514	paymore	–0.3670	0.0257
Age	–0.0001	0.0013	depen	0.1983	0.0627
Age60	0.1520	0.0391	fam_dep	0.1447	0.0620
Child_14	0.0492	0.0278	platend	–0.0104	0.0243
Constant	6.6104	0.1695			
Selection equation					
Male	–0.0098	0.0163	Southeast	0.2822	0.0319
Education	0.1161	0.0018	Northeast	0.0626	0.0328
Income	0.0003	0.00001	South	0.0653	0.0345
Black	–0.0654	0.0350	Age	0.0215	0.0008
White	0.2188	0.0174	Age60	–0.1998	0.0274
Asian	0.4010	0.0769	Child_14	0.0790	0.0199
Center-West	–0.0710	0.0377	Constant	–3.5883	0.0494

Source: 1998 PNAD and authors' calculations.

which will be the case if:

$$\ln P_i^P - X_i \_\beta \, LE \, \varepsilon_i \tag{8}$$

which occurs with probability $1 - \Phi[(\ln P_i^P - X_i\beta)/\sigma]$. Similarly, the probability that individual i chooses public health coverage is given by $\Phi[(\ln P_i^P - X_i\beta)/\sigma]$. The likelihood of observing all of the health coverage choices of the individuals in the data set (y_i), given their observable attributes (X_i) and private health care price (P_i^P), can therefore be written as:

$$L(\bar{y},\bar{X},\bar{P}^{P};\beta,\sigma)=\prod_{y_i=0}\left[\Phi\left(\frac{\ln P_i^P - X_i\_\beta}{\sigma}\right)\right]\prod_{y_i=1}\left[1-\Phi\left(\frac{\ln P_i^P - X_i\_\beta}{\sigma}\right)\right] \quad (9)$$

This likelihood function is maximized over the parameter vector, β, using data describing the decisions and attributes of a 10 percent subsample of household heads in the PNAD. The use of only household heads eliminates the correlation in insurance type between members of a household that exists in the full data set, and the 10 percent subsample is chosen for computational tractability. Eliminating data with missing observations for some variables, this yields a sample size of N = 8,267. Coefficient estimates and standard errors are reported in Table 3.6. The parameter σ is not identified in this discrete choice model; in particular, the parameters β are only identified up to a scaling parameter introduced by σ. This regression is therefore performed for a range of plausible values for the conditional standard deviation of $ln\ P_i^S$: 0.5, 1, 2, and 3; higher and lower values of the standard deviation lead to numerical problems. Results for each of these standard deviations are reported in Table 3.6.

Parameter estimates generally have the expected sign and tend to be statistically significant. Those who would be expected to have greater disutility from factors such as congestion (i.e., from having a greater opportunity cost of time) face a higher imputed price for SUS health care. This is true of older and more educated individuals, although once individuals are over the age of 60 (i.e., when they begin to retire), their imputed SUS health care price falls. Individuals with higher incomes face a higher price, also because of a greater opportunity cost of time, and urban individuals face a greater cost than rural individuals, possibly because congestion problems are worse in cities. People in the south, southeast and center-west regions of Brazil face higher prices than those in the north and northeast, and blacks and those in the mixed racial category face lower prices than whites, while Asians face higher prices. While the magnitudes of the various effects may differ, the signs of each of these effects are consistent across possible values of σ.

As a measure of model fit, the predicted health coverage decisions of this model can be compared with the decisions observed in the data. The model does well, correctly predicting the choices of 87 percent of all individuals irrespective of the assumed value of σ. When the model fails to predict correctly, it tends to be in the case of incorrectly forecasting the choices

Table 3.6. Determinants of ln P_i^s
(N = 8,267)

	Analysis of estimate sensitivity to assumed value of σ							
	σ = 0.5		σ = 1		σ = 2		σ = 3	
Variable	**Estimate**	**t-stat**	**Estimate**	**t-stat**	**Estimate**	**t-stat**	**Estimate**	**t-stat**
Constant	4.7355	88.009	3.5521	29.979	1.1756	4.751	–1.2034	–3.199
Education	.03127	15.562	.08772	19.236	.20130	20.788	.31508	21.282
Employee	–.13516	–6.311	–.25389	–5.253	–.49186	–4.811	–.73000	–4.678
Self-employed	–.16128	–7.432	–.29376	–6.032	–.56029	–5.452	–.82723	–5.277
Domestic worker	–.27520	–5.137	–.53145	–4.622	–1.0444	–4.407	–1.5574	–4.332
Family size	–.02573	–3.961	–.04506	–3.147	–.08387	–2.802	–.12272	–2.696
Male	–.02050	–1.020	–.14301	–3.200	–.38863	–4.140	–.63430	–4.435
Age	.01781	19.854	.02704	13.754	.04561	11.138	.06421	10.322
Age > 60	–.01627	–.528	–.06543	–.952	–.16326	–1.130	–.26098	–1.185
Income per person	.00015	44.930	.00023	28.825	.00040	22.482	.00058	20.600
Black	–.31872	–8.112	–.48615	–5.704	–.82241	–4.651	–1.1591	–4.320
Mixed	–.17262	–8.505	–.26142	–5.844	–.43949	–4.703	–.61775	–4.348
Asian	.41099	4.092	.51005	2.096	.70864	1.320	.90733	1.091
Metro	.12928	7.682	.24534	6.507	.47933	6.036	.71375	5.896
Child_14	–.02329	–1.033	–.04119	–.817	–.07710	–.727	–.11303	–.699
Center-West	.15400	5.278	.23378	3.646	.39308	4.751	.55232	2.713
Northeast	.12446	4.490	.03421	.553	.39308	2.936	–.32815	–1.655
Southeast	.07870	3.627	.14061	2.868	.26577	2.559	.39135	2.467
North	.07897	2.074	.04931	.5856	–.01069	–.060	–.07078	–.264
Log likelihood	–3496.51		–2980.96		–2770.32		–2706.81	

Source: 1998 PNAD and authors' calculations.

made by those individuals who opt for private health coverage, i.e., high-income, more educated and older (younger than 60 years) people.

Results and Policy Analysis

Welfare Effects of a Change in the Price of Public Health Care

The initial goal of this chapter was to determine which groups would suffer most as a result of the increasing congestion of the public health care infrastructure likely accompany current socio-demographic trends in Brazil. To

measure the welfare cost of increased waiting time for public health provision that might result from more elderly Brazilians relying on the SUS without a corresponding increase in supply, it is necessary only to consider the effect on different individuals in the sample if the price of public health care were to increase (e.g., by 50 percent), taking into account the optimizing insurance decision each person makes in response to this increase. Many individuals who had chosen public health care, for example, might stick with that choice and bear the brunt of the price increase, while others might find it optimal to pay more and switch to private health care. Those who had chosen private health coverage prior to the price increase would experience no change in price or disposable income. The decisions of each individual in the data set are simulated, backing-out the overall change in the price of receiving health care he or she faces under each of the possible values that σ might take. The last item considered is the difference in the natural logarithms of the prices ultimately faced by each individual before and after the price change. This measure provides a proportional measure of the compensating variation in income needed to maintain the same level of utility:

$$\ln P_i^{H'} - \ln P_i^{H} = \frac{1}{1-\alpha_1}(\ln I_i' - \ln I_i) \tag{10}$$

where P_i^H and I_i represent the price of health care provision and the accompanying required level of income needed to reach the original level of utility, after the increase in the price of SUS health coverage. Note that the compensating variation in income cannot be calculated directly, because it is not possible to determine α_i for each individual. This results from the fact that an individual's *full* income endowment (i.e., including the value of available time, etc.) is not observed, but only the individual's monetary income, which is not expended at all if SUS health care is employed. This means that it is impossible to ultimately determine whether the difference in log prices is attributable to a compensating variation in income, or to heterogeneity in preferences. For the discussion below, the former is assumed.

In order to quickly summarize the welfare implications of an increase in the price of SUS health care like that which would accompany increasing congestion of that system, this proportional measure is regressed on a vector of socio-demographic attributes in order to determine which groups in Brazilian society will suffer the most. The difference in magnitude of the

effect across groups could not be determined from the simple probit analysis described above, because that analysis did not describe how different types of behaviors would change in response to a price change. In all, the model (assuming $\sigma = 1$) predicts that 6.6 percent of all individuals consuming public health care prior to the price change would switch from public to private health coverage in response to this simulated price increase.[5] Accounting for optimizing responses is therefore important.

The results of this regression, performed separately for each potential value of σ, appear in Table 3.7. Those in the north (i.e., the excluded region) fare worse than those in the rest of Brazil, especially the center-west and southeast. Blacks and those in the mixed racial group fare worse than whites, while Asians generally do better (owing to their greater predisposition to have been using private health care before the price increase). Older individuals do better (as they are also more likely to have been using private health care in the first place) until they reach the age of 60, at which point they generally rely more on public health care and do much worse. Men generally fare worse than women, while those with more education and higher levels of income do better in the face of rising SUS prices, again reflecting predispositions towards using private health coverage. The directions of these effects are consistent across possible values of σ.

Welfare Effects of a Private Health Care Subsidy

The apparatus developed above also makes it possible to consider the implications of counterfactual policies designed to offset increasing congestion in the provision of public health care, where a simple reduced-form analysis cannot. In particular, the implication can be considered for individuals' optimizing choices of a private health care subsidy, designed to expand the individual's budget constraint *only* if the income is used for the purchase of private health care. The welfare implications of such a policy could then (with better data describing the *full* income endowment) be compared to the implications of a simple income subsidy that could be used for any sort of consumption, indicating the value of a relatively paternalistic policy. In the absence of such data, the welfare consequences of a simple 50 percent price subsidization of private health care (i.e., the government pays 50 cents

[5] Note that this number falls dramatically as σ rises towards three.

Table 3.7. Socio-Demographic Effects on a Proportional Measure of Compensating Income Variation from a 50 Percent Increase in P_i^S
(N = 8,267)

	Analysis of estimate sensitivity to assumed value of σ							
	σ = 0.5		σ = 1		σ = 2		σ = 3	
Variable	**Estimate**	**t-stat**	**Estimate**	**t-stat**	**Estimate**	**t-stat**	**Estimate**	**t-stat**
Constant	.50761	87.310	.50440	94.572	.49925	98.390	.49769	99.341
Education	–.01482	–57.227	–.01084	–45.660	–.00910	–40.283	–.00844	–37.857
Employee	.02847	10.956	.01503	6.304	.00782	3.448	.00665	2.972
Self-employed	.02839	10.924	.01577	6.615	.00999	4.405	.00877	3.914
Domestic worker	.04455	9.764	.01758	4.200	.00902	2.265	.00690	1.754
Family size	.00248	4.319	.00085	1.605	.00069	1.370	.00071	1.434
Male	.02330	9.924	.016805	7.802	.01449	7.071	.01408	6.959
Age	–.00223	–23.455	–.16133	–18.468	–.00135	–16.231	–.00130	–15.848
Age > 60	.01999	5.460	.01386	4.126	.00988	3.090	.00964	3.054
Income per person	–.00003	–26.192	–.00004	–39.400	–.00004	–46.012	–.00004	–47.922
Black	.03493	9.483	.01591	4.707	.00953	2.9626	.00813	2.560
Mixed	.02110	10.182	.01007	5.296	.00657	3.6310	.00591	3.308
Asian	–.05146	–2.872	–.06987	–4.251	–.08442	–5.397	–.09729	–6.300
Metro	–.02115	–11.050	–.00978	–5.567	–.00597	–3.572	–.00577	–3.498
Child_14	–.00215	–.9265	–.00073	–.3447	–.00087	–.4269	–.00000	–1.095
Center-West	–.01535	–4.425	–.01078	–3.387	–.00936	–3.092	–.00956	–3.197
Northeast	–.00317	–1.105	–.00755	–2.872	–.00984	–3.934	–.01044	–4.228
Southeast	–.01725	–6.469	–.01060	–4.332	–.00992	–4.264	–.01058	–4.601
North	–.87364	–.2118	–.00085	–.2235	–.00312	–.8665	–.00463	–1.303
R^2	.561		.508		.497		.493	

Source: 1998 PNAD and authors' calculations.

on every one real spent by the individual on private health care) are considered. In order to describe how the resulting welfare gains differ across socio-demographic groups (again, a *proportional* measure of the compensating variation in income, assuming homogenous preferences, after the optimizing provider of health care is chosen[6]), these gains are regressed on a vector

[6] The model (assuming σ = 1) predicts that approximately 13.3 percent of all individuals consuming public health care prior to the price subsidy would switch to private health coverage. With such a large increase in the demand for private health care, the government might want to undertake policies to facilitate entry by new private health care providers in addition to the subsidy, so as to avoid new congestion costs.

of socio-demographic attributes. The results of this regression are described in Table 3.8 for each of the four values of σ considered. The directions of each of the marginal effects are consistent across alternative values of σ. Negative numbers describe reductions in income that return individuals to their original levels of utility (i.e., indicating a benefit). People in the center-west and southeast regions seem to benefit most from this price subsidy, while those in the north and northeast benefit the least. Whites and Asians benefit more than blacks and those in the mixed racial group, and those with higher levels of education benefit more than those with less. Similarly, wealthier and older individuals (under the age of 60) benefit more from the subsidy. Generally, these relative benefits reflect a greater predisposition towards (or propensity to switch to) private health care provision.

The natural question is how might such a private health care subsidization policy be targeted to benefit those individuals who would suffer most under an increase in the price of public care. To the extent that such subsidies, when applied broadly, seem to benefit high-income and highly educated people in the more developed parts of Brazil, they simply represent a transfer of rents, since those people suffer less than the poor from the increasing price of public health care. One possible alternative would be to implement the subsidy as part of an income tax collection regime, where participation criteria could easily be established so as to make the subsidy available only to low-income residents. Problems of fraud in the reporting of private health care expenditures, however, might make this approach difficult. Instead, it might be preferable to focus on the results of the first counterfactual simulation and target subsidy funds geographically so as to reach those individuals who lose most under the simulated increases in public prices. Such is true, for example, of black and mixed race residents,[7] particularly those with low levels of education and in rural areas. One possible solution that would certainly benefit the most disadvantaged groups would be to target subsidy funds towards lowering the cost of private health care in rural areas, possibly by establishing new health care facilities where none were previously present. Also important to remember is the relatively large magnitude of the negative effect on members of the black and mixed racial groups of increasing public health care prices, even in urban areas.

[7] Reaching these particular racial groups might be difficult, unless the subsidies took the form of monies to establish new private health care facilities in racially segregated neighborhoods.

Table 3.8. Socio-Demographic Effects on a Proportional Measure of Compensating Income Variation from a 50 Percent Reduction in P_i^P
(N = 8,267)

	Analysis of estimate sensitivity to assumed value of σ							
	σ = 0.5		σ = 1		σ = 2		σ = 3	
Variable	**Estimate**	**t-stat**	**Estimate**	**t-stat**	**Estimate**	**t-stat**	**Estimate**	**t-stat**
Constant	.13868	16.174	.18410	20.486	.17196	19.657	.16534	19.164
Education	–.03238	–84.801	–.02291	–57.254	–.01762	–45.238	–.01595	–41.506
Employee	.06971	18.193	.03836	9.551	.01943	4.971	.01444	3.746
Self-employed	.07370	19.231	.03905	9.722	.02208	5.647	.01834	4.755
Domestic worker	.14139	21.013	.05355	7.593	.02206	3.213	.01577	2.329
Family size	.00664	7.847	.00310	3.496	.00131	1.520	.00135	1.588
Male	.05934	17.137	.03629	10.000	.02683	7.595	.02404	6.899
Age	–.00480	–34.164	–.00345	–23.448	–.00261	–18.204	–.00238	–16.819
Age > 60	.03380	6.259	.03111	5.497	.02172	3.943	.01834	3.375
Income per person	–.00003	–18.538	–.00006	–36.940	–.00007	–44.557	–.00007	–46.916
Black	.09321	17.157	.04312	7.573	.02180	3.933	.01729	3.162
Mixed	.05404	17.683	.02645	8.259	.01495	4.794	.01191	3.873
Asian	–.06196	–2.344	–.13335	–4.814	–.16781	–6.223	–.18145	–6.823
Metro	–.05698	–20.181	–.02856	–9.652	–.01427	–4.952	–.01170	–4.119
Child_14	.00098	.287	–.00344	–.9583	–.00215	–.614	–.00272	–.787
Center-West	–.03182	–6.217	–.02395	–4.466	–.01790	–3.427	–.01676	–3.254
Northeast	.01186	2.808	–.01063	–2.400	–.01631	–3.781	–.01798	–4.228
Southeast	–.04273	–10.865	–.02299	–5.577	–.01739	–4.334	–.01789	–4.521
North	.00590	.971	–.00187	–.293	–.00399	–.642	–.00652	–1.065
R^2	.724		.587		.526		.511	

Source: 1998 PNAD and authors' calculations.

Finally, dealing with the impact of rising public health care costs on the elderly (a major concern given current socio-demographic trends) by subsidizing the consumption of private health care seems especially futile, since those over the age of 60 are predicted to benefit less than most other groups from this policy. This arises from the model's prediction that members of this group are not as likely to switch to private health coverage even with the change in relative prices. Indeed, in order to limit the adverse effects on this group of rising congestion in public health care consumption without affecting huge rent transfers to those who are less adversely im-

pacted, the government will likely have to take steps to directly increase the supply of public health care provision.

Conclusions

The goal of this analysis was to determine which groups in Brazilian society were most excluded from private health care. Private health care is generally considered to be of a higher quality than its public counterpart; this perception is generally supported by the PNAD survey data. Such exclusion is not of a direct form, as would be racial exclusion from a club, but is rather based on individuals' facing different relative prices for public and private health care owing to differences in their observable attributes and preferences for health care consumption.

The initial analysis of PNAD survey data documents what is generally perceived to be the case—that poor, rural, black and mixed-race Brazilians tend to rely more on public health care. This alone would not necessarily represent a source of social inequity, except that the price of this form of health care is expected to increase in the coming decades owing to the increasing congestion of an already-overburdened system, and these groups are expected to suffer disproportionately because of their inability to switch to private health coverage. In order to determine how these price increases would be distributed over different socioeconomic groups, a more elaborate model of optimal individual decision-making is needed; i.e., a model that makes it possible to determine how individuals would behave under current and counterfactual relative-price scenarios.

Operating under constraints of data availability, it is assumed that each individual was required to consume a single unit of some form of health care coverage (i.e., public or private), and that differences in the quality of care across forms would be internalized in the price confronting the individual. Differences in price might also arise from observable individual attributes (i.e., a direct form of discrimination), or from an individual's preferences for health care consumption (e.g., individuals with strong preferences for health care consumption might face an even higher price for an effective unit of public care than a similar individual who had weak preferences for health care consumption), but available data do not make it possible to identify these effects.

From a simple and stylized model of utility maximization, it was possible to recover estimates of the price of public health coverage and to use those estimates to infer which socio-demographic groups would suffer most from an increase in the congestion of the public health care system. The conclusions of this analysis conform to the general perceptions regarding race, education, and income groups, and suggest that rural individuals are more at risk than those living in metropolitan areas. Moreover, they suggest that using private health care subsidies as a solution to increased congestion of the SUS will chiefly result in rent transfers to the least affected groups. This result is significant in light of the Brazilian government's propensity toward privatization of public health care delivery in response to similar problems over the last 10 years. Instead, it is clear that the SUS system itself must be expanded to meet the growing demand.

With even more detailed data on the attributes of the alternative forms of health care provision (e.g., quality), it might also be possible to build a more realistic hedonic model in which individuals with weak preferences for health care would choose to consume the type that exhibits low levels of amenities and a low price, while those who derive a great deal of utility from the consumption of health care might choose a "deluxe" form of health care provision. This could be important in predicting how different individuals would respond to an increase in the congestion of the public system, which would increase waiting times for treatment (i.e., a specific trait of the health care commodity). For example, certain socio-demographic groups might exhibit a strong distaste for waiting time, and they would thus tend to bear more of the burden of increasing congestion of the SUS. While other survey data provide some indication of waiting time incurred in the receipt of health care services, these data exhibit many missing observations, and it is unclear whether they will be appropriate for such an analysis.

Even with the limitations and simplifications described above, the current model suggests which groups are most likely to suffer from the increasing congestion of the public health care infrastructure that is likely to accompany current demographic trends in Brazil. From an equity perspective, these groups are generally those about whom concern is greatest, suggesting that some policy (i.e., subsidizing private health care or expanding public infrastructure) must be undertaken. The results of the counterfactual simulations suggest that the latter would be a more cost-effective contribution to delivering health care assistance to those who need it most.

This page intentionally left blank

CHAPTER FOUR

Legal Status and Social Exclusion: Nicaraguans in Urban Costa Rica

Edward Funkhouser
Juan Pablo Pérez Sáinz
Carlos Sojo

Costa Rican society was transformed during the 1990s as a result of the immigration of between 350,000 and 450,000 Nicaraguans.[1] Although the United States was the principal destination of Nicaraguan emigrants during that decade, today the leading destination is Costa Rica, which in the face of declining economic conditions in Nicaragua represents a more viable alternative, both economically and psychologically, to emigration to the United States.[2] Nicaraguan immigrants now make up over 10 percent of the population of Costa Rica.

Unlike previous waves of Nicaraguan immigrants into Costa Rica, which were usually temporary or linked to marital ties in Costa Rica, the current wave increasingly is permanent and family based. Recent arrivals to Costa Rica are more likely to come from rural areas and areas outside of Managua, are younger at the time of arrival, and have less education than previous emigrants. After years of not recognizing illegal immigrants, in 1999 the Costa Rican government extended a one-year amnesty to 160,000 Nicaraguans who could show residence in Costa Rica prior to November 1998.

The main focus of this chapter is to describe the conditions of Nicaraguan immigrants in comparison with Costa Ricans, and to examine the extent to which they might be socially excluded from Costa Rican society. The

[1] For descriptions of migration patterns from Nicaragua to Costa Rica, see Funkhouser (1999), Morales and Castro (1999), and FLACSO-Costa Rica (1999).

[2] According to the World Bank, per capita income in Nicaragua fell from approximately $800 in the early 1980s to approximately $340 in the early 1990s, and then rebounded to about $430 in 1999. This compares to $2,740 in 1999 in Costa Rica.

chapter finds that most Nicaraguan immigrants, like most immigrants in general, have integrated into Costa Rican society, in contrast to many other excluded groups. It also finds that while non-citizen or illegal status alone limits the rights of Nicaraguans, this status can be changed over time, especially in the second generation. Finally, housing policy in metropolitan San José makes geographic concentration of Nicaraguans likely as the migration continues.

Social Exclusion and Vulnerability

Applying the idea of exclusion to recent migrants, while not unique to this chapter, involves examining a group that, in effect, has chosen to be excluded.[3] While many Nicaraguan immigrants to Costa Rica experience different legal treatment and earn lower incomes than natives, these conditions nonetheless represent an improvement over what they experienced or expected to experience in their home country.[4] And while current migrants have different legal status, future exclusion may depend on other distinctive characteristics that will persist after naturalization or into the second generation of migrants.

The definition of social exclusion used in this chapter is based on the idea of denial of equal access to opportunities. Whatever the causes or mechanisms of social exclusion may be, such denial of access results in consequences such as geographic exclusion, economic exclusion, and denial of citizenship. The confluence of different types of exclusion, moreover, distinguishes so-

[3] For an overview of this concept, see de Haan (1999) or Loury (1999).

[4] The better economic outcomes are a result of the capacity of the Costa Rican labor market to absorb an increase in labor supply. According to the data in the annual national household survey, employment increased by approximately 200,000, or about 20 percent, between 1993 and 1999. Of this increase, about 70,000 jobs involved people with six or fewer years of education; 60,000 jobs employed those with seven to 11 years of education; and another 70,000 employed people with 12 or more years of education. Between 1997 and 1999 (the years in which Nicaraguans can be distinguished), the increase in employment of Nicaraguans accounted for all of the 20,000 increased employment for those with six or fewer years of education, 5,000 of the 13,000 increased jobs for those with seven to 11 years of education, and only 2,000 of the 37,000 jobs for those with 12 or more years of education.

cial exclusion from other explanations for worse outcomes experienced by particular groups.

In many cases, though, the dichotomous view of exclusion—either a group is excluded in all dimensions, or it is not—is not always appropriate. This is particularly relevant for groups that have chosen to be excluded. While the labor market integration of these groups may not reflect current exclusion, the type of integration may make the group vulnerable to future exclusion. Vulnerability is also a difficult concept to define. A working definition, though, might involve a situation in which a group that maintains its current behavior into the future would experience exclusion at some point because of denial of access to opportunities. For example, employment in cyclical industries displays outcomes during periods of growth that would not appear to be the result of exclusion; during economic downturns, though, outcomes may suggest exclusion, as unemployed workers are denied access to other industries and occupations. It would also be the case for second-generation immigrants who are denied access to the labor market that their parents had entered.

Legal Status of Nicaraguans in Costa Rica

Although Nicaraguans are more likely to be of mixed race and to have darker skin than Costa Ricans, the more identifiable differences between Nicaraguans and Costa Ricans are social characteristics, such as accent, slang and dress. All of these can be changed even within a generation. As a result, one of the most important aspects of Nicaraguan integration into Costa Rican society is legal status.

Naturalization in Costa Rica may be requested for children born in Costa Rica, Central Americans who have lived officially in Costa Rica for five years, persons who have been married to Costa Ricans for more than two years and have resided in Costa Rica for two years, and persons who have lived in Costa Rica for more than 20 years.[5] There were approximately 100,000 such Nicaraguans in Costa Rica with citizenship prior to the migration wave of the 1990s. The 1999 amnesty provided legal residence for

[5] Lizano and Maklouf Weiss (1999).

the 160,000 illegal immigrants who applied before July 31, 1999 and could show residence prior to November 1998.[6] Those who arrived after 1998 are overwhelmingly illegal. Immigration has continued following the amnesty, with 8 to 9 percent of migrants arriving for the first time between 1999 and 2001.

While the constitution states that foreigners have the same obligations and rights as Costa Ricans and that the laws can limit these rights only with fairness and reasonableness, in practice there are several differences in the legal treatment of foreigners in Costa Rica. First, there are several restrictions on political participation, including intervention in political affairs, occupation of directorial positions in unions, and the specification of a maximum number of positions on the boards of directors of several other types of organizations. On the boards of directors of *barrio* organizations, for example, Nicaraguans can occupy only two positions at most. Second, while restrictions on the labor market activity of foreigners have been relaxed—including the elimination of a labor card, maximum length of labor contracts, and restrictions on employment in private security—employers are expected to select a Costa Rican over a foreigner when the two applicants are otherwise equal. Third, the possibilities for transfer of ownership of properties accessed through land invasions are limited by legal status. Fourth, affiliation with the social insurance system, which is available to legal but not illegal residents, is a prerequisite for many types of health care.[7] Finally, while the law does not establish restrictions per se on Nicaraguans in the educational system, there are problems with graduation at each level, since diplomas are not issued for those who cannot

[6] For a more detailed discussion of the amnesty and those who applied, see Morales (1999). There are three main reasons why eligible Nicaraguans might not have applied for the amnesty. First, it may be difficult for some Nicaraguans to show residency prior to November 1998. Second, it may have been unclear that the period will be renewed at the end of one year. By reporting themselves to the Costa Rican government, some Nicaraguans may feel that they increase their chances for difficulties at the end of the year. A third reason may be a lack of urgency for short-term immigrants. Cost did not appear to be a major deterrent, since the application requirements were 3,355 cordobas (approximately $12), a passport size photo, proof of identity, and lack of criminal record during the previous 10 years.

[7] Although neither emergency medical care nor care in neighborhood health clinics (EBAIS) require proof of legal residence, care in other health services, such as an appointment with a doctor, requires participation in the social insurance system.

show legal residence. This can effectively restrict entry into the next level of schooling.[8]

The primary reason for entering Costa Rica illegally is economic. Legal immigration represents significant expenditures, particularly for workers who plan to remain in Costa Rica for only a short time. Obtaining a passport from the Nicaraguan government costs $100, and obtaining a tourist visa in Costa Rica costs $25. (Nicaraguans who enter Costa Rica as tourists can subsequently regularize their status.) Compared to these high initial expenses, renewing a residency visa costs $3, while renewing an amnesty visa costs $10. These institutional features, as well as the large number of Nicaraguans who applied for amnesty in 1999 when the costs of legalization were reduced, suggest that, within a model of migration choice, legal status is a measure of the permanence of migration.

Formation of Nicaraguan Enclaves

A third important aspect of the integration process of Nicaraguans is residential location. The formation of Nicaraguan enclaves in the metropolitan area of San José has followed two dominant processes during the recent migration—land invasions in marginal urban areas and rooming houses in more established areas.[9] During the presidency of Oscar Arias in the 1980s, and prior to the Nicaraguan influx, the construction of organized developments of block housing on the outskirts of urban areas promoted urban expansion. Since then, expansion has been less organized, with illegal inva-

[8] One of the two types of scholarships offered by the Ministry of Education to poor households is not available to foreign households. The two sources of funds for the institution that gives these scholarships, IMAS, are its own budget and funds received through the Ley de Asignaciones Familiares. Funds from the latter are available only to Costa Rican households.

[9] Earlier waves of migrants to the metropolitan area tended not to be as geographically concentrated. The two other parts of the country with significant clusters of Nicaraguans are the northern and Atlantic regions, both traditional destinations of agricultural workers from Nicaragua. Because the former is the destination of workers for large banana plantations, the groupings of Nicaraguans tend to be large but without much infrastructure. Nicaraguan migrants to the northern region tend to work on smaller farms. Workers on a particular farm or grouping of farms live near each other. While some land invasions in urban areas similar to those in the Central Valley have occurred in Limón, the spatial distribution of Nicaraguans is mostly related to proximity to the work site.

sions followed by the construction of makeshift housing on unused lands, often publicly owned property. In the most recent areas of such development, there are few public services, roads are unpaved, sanitation is a problem, and distance to schools or work is an issue. Although these types of invasions occurred prior to the current wave of Nicaraguan immigrants, in recent years such invasions have increased in number.

The general strategy of these invasions is to stake a claim to a piece of land by occupying it in the hope that the claim becomes recognized over time. Because there is no legal claim to the property, however, temporary abandonment could lead to loss of the property. In most cases, there is a *barrio* organization that charges a monthly fee and promotes access to services from the government. The main incentive for paying these quotas is the belief that they will improve the chances of attaining legal ownership of the property. At present, Nicaraguans who are in Costa Rica illegally are prohibited from obtaining a title for these properties.[10]

A second type of residential area in which Nicaraguans have become concentrated involves houses that have been converted to rooming houses or rental units. These buildings consist of small sleeping rooms, generally three square meters, a shared bathroom and a common room. The two areas in the metropolitan region with large numbers of Nicaraguans living in rooming houses are fairly run-down neighborhoods. One is in the very center of the city, next to the red-light district, in an area that has experienced an increase in crime. The other is located in a marginal area on the downward slope of a canyon.

Data

Two types of data are used to examine group effects and the mechanisms of social exclusion of Nicaraguans in Costa Rica. The first is data from the

[10] The characteristics of such neighborhoods have been documented by Acuna and Olivares (1999). The expanded list of such concentrations in San José (provided by Edith Olivares) includes Los Diques, Los Guido, Rincon Grande, Tejarcillos, La Milpa, La Carpio, Marco Aurelio, Pochote, Paso la Vaca, San Juan de Dios, and Tirases. It should be noted that while many Nicaraguans live in these recently invaded areas, the areas include a significant number of Costa Rican households as well. Typically, the Costa Rican families are organized along the main road or roads and the Nicaraguan families along the minor roads.

National Multipurpose Household Survey (*Encuesta Nacional de Hogares de Propósitos Multiples*). The types of questions asked on this survey are limited and not directly related to the study of social exclusion, but they provide a large, representative sample. The second set of data comes from a smaller targeted survey in the metropolitan area of San José, conducted by the Facultad Latinoamericana de Ciencias Sociales (FLACSO) in Costa Rica specifically for this study. This survey is smaller in scope (398 households, with 192 Costa Rican heads and 206 Nicaraguan heads), but it includes a survey instrument with questions related to social exclusion. To ensure that the surveys cover a comparable population, the government household survey is restricted to metropolitan San José.

National Household Survey

The Costa Rican government's Statistics and Census Institute has carried out the National Household Survey on a regular basis in July of each year since 1976. Approximately 10,000 households are asked basic demographic, labor market, and household information. Since 1997, the survey has also included questions about place of birth and citizenship status. Immigrants are classified as either naturalized or citizens of another country, with Nicaraguans identified separately. The survey tends to undercount the number of Nicaraguans, though the number reported is increasing over time, with 75,490 Nicaraguans represented in 1997 and 133,548 in 1999 for the country as a whole.[11] Those Nicaraguans not included are likely to be worse off than those included, reflecting either a more precarious position or less permanent migration.

FLACSO Survey

The FLASCO survey included 398 households and was carried out in the metropolitan San José area in January and February 2001. While the survey sample is smaller than that of the National Household Survey, the number of Nicaraguan households in the metropolitan area is larger, and the survey

[11] The sample sizes for the country as a whole were 994 in 1997, 1,092 in 1998, and 1,633 in 1999. For San José, the sample includes 211 Nicaraguans representing 28,837 persons in 1997, 224 Nicaraguans representing 29,344 in 1998, and 345 representing 45,540 in 1999.

instrument is designed to provide some insight into the mechanisms through which social exclusion might operate.[12]

The FLACSO survey used the information from the primary sampling units (approximately 60 households) of the government's household survey to determine the *barrios* included in the FLACSO sample. The first step was to identify the *barrios* corresponding to each sampling unit (*segmento*) in the 2000 survey. Sampling units were then assigned to three strata: high presence of Nicaraguans (over 20 percent in the *segmento*), medium presence of Nicaraguans (5 to 20 percent Nicaraguans), and without presence of Nicaraguans (under 5 percent Nicaraguan, but of similar socioeconomic status).[13] Two *barrios* were chosen by randomly selecting two sampling units from each of the three strata.[14]

The *barrios* included in the final sample were La Union and La Carpio (high presence of Nicaraguans), Paso Ancho and La Aurora (medium presence), and Guarari and Barrio Mexico (low presence).[15] Households were randomly selected within each *barrio* with over-sampling of Nicaraguans in the *barrios* with a medium and high presence of Nicaraguans.

---

[12] The information in the FLACSO survey was supplemented with three interviews in each of the six communities: one with the director of the main primary school, one with the leader of the *barrio* association, and one with the director of the health center. The primary purpose of these interviews was to document differential patterns of activity or outcomes that lead to future inclusion or exclusion as the Nicaraguan community extends its time in Costa Rica.

[13] Only a few *segmentos* had a population under 5 percent Nicaraguan. The criterion for selecting *barrios* of similar socio-economic status was poverty rates above the national average.

[14] To allow over-sampling of Nicaraguans in the medium and high presence *barrios*, a field census was undertaken within each selected *barrio* in which only the head of household was interviewed. Of the original six *barrios* selected, one from the high strata was found following the field census to not have the appropriate number of Nicaraguans (because the sampling unit in the household survey was not representative of the *barrio* as a whole). A second *barrio*, originally selected from the medium strata, was not large enough to over-sample Nicaraguans. These two *barrios* were replaced in the sample.

[15] After the field census, the differences in the Nicaraguan proportion (measured by nationality of principal earner in the household) of the medium presence *barrios* was higher than originally calculated from the *segmentos* in the household survey. In the case of La Union, difficulty in entering the rooming houses led to a lower target number of households in that *barrio* and a higher number in La Carpio.

An overview of the characteristics of the *barrio* and the nature of Nicaraguan insertion into it is presented in Table 4.1. The descriptions make clear that there are two types of geographic insertion by Nicaraguans into the metropolitan area. The first involves moving into makeshift houses on previously occupied vacant lands, as described above. In La Carpio, the Nicaraguan households were part of the original land takeover, though they were located in more marginal locations within the *barrio*. In La Aurora, Nicaraguan households occupied the interior of a coffee plantation next to a more established Costa Rican community.[16] Rooming houses were found in the *barrios* of Paso Ancho and Bajos de La Union.

Summary Information

Table 4.2 summarizes information provided by persons aged 20 to 65 living in urban areas of metropolitan San José. The 1999 government survey results are reported in columns (1)–(3), and the FLACSO survey results are reported in columns (4)–(6). The first two columns for each data set show means for Costa Ricans and Nicaraguans. The third column (columns 3 and 6) includes those Nicaraguans who live in neighborhoods with a high presence of Nicaraguans.

The general patterns for Nicaraguans relative to Costa Ricans are similar in the two surveys. Nicaraguans are younger, less educated, more likely to be working,[17] work more hours, are more likely to be working in the informal sector, earn less, are less likely to have social insurance, and are more likely to be living in poverty. For most variables, the magnitudes of the descriptive statistics are similar in the two surveys.[18] There are two notable differences

[16] Primarily to limit the survey time to 45–60 minutes, the emphasis was on the principal earner in the household, rather than all household members. Basic information was asked of the respondent about all household members who share food in the household. For Nicaraguans, this included some information about migration. The household head was also asked additional labor market questions and detailed questions about social capital and perceptions. Principal earners who were Nicaraguan were also asked more detailed questions related to migration and activity in Nicaragua.

[17] The relatively large number of Costa Rican principal earners without labor income over the age of 65 in the FLACSO survey suggests that in some cases, the distinction between principal earner and head of household was misunderstood.

[18] This is a bit surprising for incomes, since the Costa Rican *barrio*s in the FLACSO survey were chosen according to poverty rate (above the national average), while those in the household survey were not.

Table 4.1. The Six Areas of the FLACSO Survey

Location	Community infrastructure	Condition of housing	Living conditions of Nicaraguans
La Carpio High presence of Nicaraguan population. First settlement established in April 1993. Located in central canton of San José, district of La Uruca.	Main road is paved and has sidewalks. The other streets are dirt, without sidewalks and in poor shape. The community has one public school and a health clinic. There are no green or recreation zones. The space that children use to play soccer and baseball has been designated for the construction of a new school.	No family has been served by housing or social programs. As a result, housing is predominately *ranchos* made with discarded materials, though there are some houses of cement. Many houses are on one lot, with a single electric meter or water faucet.	The overwhelming majority of Nicaraguan households are located in the southern section of the *asentamiento*, where conditions are most vulnerable because of the terrain, the proximity to a river, and the poor condition of the houses themselves.
Bajos de La Union High presence of Nicaraguan population in old *barrio* of San José. Located next to a bridge that crosses to La Uruca, the community is based around a narrow road that descends a steep hill without exit.	The streets are paved, but in poor shape. There are sidewalks. All basic services are available. There are problems with water and sewerage pipes because they were built with materials that have deteriorated and the water that is consumed becomes contaminated. In some cases, the wastewater runs into the river. There are no green or recreation zones.	Most houses were built many years ago and are in a process of deterioration. Many of the houses are thin but long.	Nicaraguans live primarily in *cuarterias* (rooming houses) with a series of small (3m x 3m) rooms. In some cases, new rooms have been constructed on an improvised second floor. There is a shared bathroom and electricity in nearly all rooming houses. Some occupants have installed equipment to cook.
Paso Ancho Medium presence of Nicaraguans. Traditional *barrio* in San José, located in the south of the central canton.	The streets are paved, but in poor shape. All basic services are available. There is a school, soccer field, and a nearby supermarket.	Housing types vary, ranging from wood houses in a state of deterioration to others made of cement and in better condition.	Most Nicaraguan families live in rooming houses concentrated in two sectors of the barrio. Some were constructed beneath the level of the road, so there is little light and, in some cases, the stair access is dirt. Some Nicaraguan families also rent houses.

La Aurora Medium presence of Nicaraguans. Located in the central canton of Heredia, district of San Francisco. The *barrio* was constructed by INVU in the 1970s. In 1993, a land takeover by about 100 families occurred on an adjacent coffee plantation.	The main community has paved roads and all basic services. The houses do not have electricity, water, or bathrooms. The residents constructed a trench for wastewater. Water is obtained from a common pipe. The access road is dirt. There is overcrowding in housing.	In the main community, the houses are made of cement, with two stories, four bedrooms. Most houses are in good condition. In the *asentamiento*, the *ranchos* have been made of discarded materials, with a large number having dirt floors.	Nicaraguan families are concentrated in the *asentamiento*, and 70 of the 80 families have a Nicaraguan principal earner. Only six of 216 households in the main community are Nicaraguan.
Guarari No presence of Nicaraguans. Located in the central canton of Heredia, district of San Francisco. The original settlement was established in 1986–87, led by a group called COPAN.	The area was constructed as part of one of the housing programs under the Arias administration. It has all basic services. There is a primary school and a health clinic, as well as green zones.	Most houses are made of cement, and there is one area with makeshift housing (*ranchos*).	
Barrio Mexico No presence of Nicaraguans. Located in the central canton of San José, near the red light district.	The streets are paved and in good condition. There are all basic services. The community has a primary school, a nearby secondary school, and a large number of businesses.	Most houses are three decades old.	

Table 4.2. Summary Data

	1999 National Household Survey Nicaraguans			2001 FLACSO Survey Nicaraguans		
	Costa Ricans (1)	All (2)	In segmentos >20% Nic. (3)	Costa Ricans (4)	All (5)	In barrios >20% Nic. (6)
Persons ages 20–65:						
Age	37.602	31.083	28.802	37.078	33.698	33.779
	(.209)	(.782)	(1.236)	(.656)	(.539)	(.769)
Female	.528	.509	.440	.540	.546	.562
	(.009)	(.033)	(.052)	(.031)	(.025)	(.035)
Years of education	9.578	7.344	6.530	8.483	6.304	5.985
	(.074)	(.283)	(.451)	(.239)	(.157)	(.218)
Male working	.864	.893	.863	.808	.919	.926
	(.009)	(.032)	(.048)	(.036)	(.021)	(.029)
Female working	.510	.655	.600	.586	.629	.614
	(.012)	(.046)	(.079)	(.041)	(.034)	(.047)
Working sample:						
Hours of work	47.692	51.046	49.912	49.355	50.895	51.169
	(.359)	(1.254)	(2.012)	(1.192)	(.963)	(1.222)
Informal sector	.300	.540	.529	.355	.556	.564
	(.010)	(.035)	(.056)	(.042)	(.033)	(.046)
Total income	132,908	73,405	58,855	122,168	77,400	74,521
	(2,807)	(9,319)	(14,614)	(7,259)	(2,206)	(2,737)
Any social insurance (head)	.833	.610	.469	.832	.481	.524
	(.011)	(.043)	(.067)	(.038)	(.038)	(.055)
Household variables:						
Household size	3.947	4.658	5.242	3.970	4.994	5.299
	(.050)	(.214)	(.330)	(.167)	(.205)	(.308)
Children under 12	.948	1.519	1.788	1.022	1.704	1.734
	(.030)	(.130)	(.201)	(.113)	(.122)	(.176)
Number working	1.643	2.089	2.212	1.436	1.687	1.712
	(.027)	(.117)	(.181)	(.080)	(.070)	(.093)
Household income	218,675	146,001	124,335	179,812	122,035	118,634
	(5,453)	(22,512)	(33,196)	(12,782)	(5,792)	(6,587)
Poverty	.147	.274	.394	.395	.514	.449
	(.010)	(.042)	(.062)	(.062)	(.050)	(.068)

Notes: Numbers in parentheses are robust standard errors. Informal sector includes all self-employed workers, domestic workers, apprentices, family workers, and wage earners and salaried workers in firms with fewer than five employees that are not in professional or managerial occupations.

in the characteristics of the Costa Rican households in the FLACSO survey and those in the government survey: in the FLACSO survey households are larger, but the number of members working is lower, which results in lower household incomes. It should also be noted that differences for the *barrios* of greatest concentration of Nicaraguans (in the final column) are not as pronounced in the FLACSO survey as in the household survey.

Nicaraguans in the FLACSO Survey

To better understand the coverage of the FLACSO survey, Table 4.3 gives more detailed summary information for Nicaraguans in each of the four *barrios* with large numbers of Nicaraguans. The first two columns include the two *barrios* with a high presence of Nicaraguans: La Union and La Carpio. The final two columns include the two *barrios* with a medium presence of Nicaraguans: Paso Ancho and La Aurora.

While the demographic characteristics of Nicaraguans ages 20–65 do not differ much by *barrio*, the household and migration characteristics are quite different. *Barrios* consisting of makeshift housing, or *ranchos*, are larger and have more children than households in more established *barrios*. Not surprisingly, households in those *barrios* are more likely to have deficient materials—42 percent in La Carpio and 84 percent in La Aurora. Residents living in La Carpio are less likely to be from the most recent wave of immigration, while those in Aurora are more likely to include recent immigrants. Nearly one-third of the residents in La Aurora are not in Costa Rica legally, reflecting arrival after 1998 and ineligibility for amnesty.

There are also differences across *barrios* in place of origin and attachment to Nicaragua. The differences are by type of *barrio*, rather than by the level of presence of Nicaraguans. Approximately three-fourths of the residents age 20–65 in the more established *barrios* of La Union and Paso Ancho are from the southern part of Nicaragua. Only one-third to two-fifths of the residents of the makeshift-housing *barrios*, La Carpio and La Aurora, are from the southern part. In La Aurora, the *barrio* including the largest proportion of most recent immigrants, over one-half of the residents age 20–65 are from the northern part of Nicaragua.[19]

[19] Over 40 percent of those in La Aurora said they were planning to return to Nicaragua.

Table 4.3. Characteristics of Nicaraguans by Barrio

	La Union	La Carpio	Paso Ancho	La Aurora
Total persons	4.219	5.662	4.292	4.740
	(.462)	(.310)	(.377)	(.369)
Children <=12	1.250	1.915	1.375	1.860
	(.287)	(.193)	(.234)	(.230)
Working members	1.750	1.718	1.688	1.500
	(.174)	(.116)	(.142)	(.130)
Poverty	.250	.380	.375	.540
	(.085)	(.057)	(.069)	(.060)
Housing:				
Deficient materials	.000	.423	.063	.840
		(.043)	(.053)	(.052)
Deficient services	.719	.620	.292	.840
	(.078)	(.053)	(.064)	(.063)
Overcrowding	.844	.718	.833	.700
	(.075)	(.051)	(.061)	(.060)
Persons ages 20–65:				
Age	33.017	33.866	35.758	31.825
	(1.346)	(.847)	(1.061)	(1.019)
Female	.576	.557	.547	.505
	(.065)	(.065)	(.051)	(.049)
Years of education	6.288	5.918	7.462	6.088
	(.392)	(.249)	(.312)	(.298)
Informal sector	.565	.543	.603	.527
	(.074)	(.056)	(.066)	(.068)
Monthly income	74,345	74,978	86,384	76,438
	(6,167)	(4,361)	(5,396)	(5,087)
Not legal	.153	.128	.074	.301
	(.046)	(.029)	(.037)	(.305)
Not citizen	.864	.718	.600	.893
	(.054)	(.034)	(.043)	(.041)
Year of arrival:				
Prior to 1980	.017	.048	.087	.010
	(.026)	(.017)	(.021)	(.020)
1980–90	.293	.241	.315	.146
	(.056)	(.036)	(.045)	(.042)
1991–95	.345	.545	.315	.388
	(.064)	(.040)	(.051)	(.048)
1996–2001	.345	.166	.283	.456
	(.058)	(.037)	(.046)	(.044)
Region in Nicaragua:				
North	.220	.286	.118	.553
	(.057)	(.036)	(.045)	(.043)
Managua	.017	.286	.172	.107
	(.048)	(.030)	(.038)	(.036)
South	.763	.429	.710	.340
	(.061)	(.039)	(.049)	(.046)
Principal earner:				
Planning to return	.250	.200	.311	.408
	(.080)	(.058)	(.067)	(.065)
Relatives in U.S.	.594	1.558	.739	1.000
	(.430)	(.312)	(.359)	(.358)
Relatives in Costa Rica	2.031	2.590	2.130	1.673
	(.643)	(.466)	(.536)	(.520)
Sends remittances	.625	.377	.500	.429
	(.088)	(.466)	(.073)	(.071)
Avg. monthly remittance ($)	37.406	16.459	40.196	26.286
	(8.265)	(5.987)	(6.894)	(6.679)

Notes: Relatives in Costa Rica are those who do not live in household. Monthly remittances are averaged over all principal earners including those who do not send remittances.

Remittance rates are highest from principal earners in the more established *barrios*: 63 percent in La Union and 50 percent in Paso Ancho, compared to 38 percent in La Carpio and 43 percent in La Aurora. Remittance levels averaged over all principal earners, including those who do not send any remittances, are also higher in the first two *barrios*, reflecting the higher remittance rate and the higher average remittance among those who do send money.[20]

Empirical Approach

The analysis begins by estimating differences in outcomes for Nicaraguans, or group effects:

$$Y_i = \alpha + X_i\beta + C_i\gamma + \delta G_i + \varepsilon_i \quad (1)$$

where Y_i is outcome Y of individual i, X_i is a set of individual characteristics, C_i is a set of community characteristics, G_i is the group effect, and ε_i is the unobservable random component of trading or business opportunities for a minority group.

The demonstration that outcomes of a group presumed to be excluded are worse than those of other groups does not show the existence of social exclusion or the mechanisms through which social exclusion impacts the excluded group. To the extent possible, it is also desirable to examine the consistency of observed patterns of social exclusion with the proposed explanations for social exclusion. These explanations include discrimination, community or class norms, contacts and networks, and epidemic effects. In addition, explanations for differences in outcomes not based on social exclusion include preferences and the possibility that low income alone leads to worse outcomes in other areas.

Table 4.5 attempts to control for some of these factors, using the data from the FLACSO survey by including proxy variables corresponding to the different mechanisms for exclusion, E, into equation (1), resulting in:

$$Y_i = \alpha + X_i\beta + C_i\gamma + \delta G_i + \lambda E_i + \varepsilon_i \quad (2)$$

[20] For more information on remittance patterns, see Funkhouser (1999), Morales and Castro (1999) and Pritchard (1999).

The main issue in equation (2) is the choice of good variables in the vector *E*. In particular, there is a possibility that the variables intended to measure social exclusion are endogenous or that the different outcomes are jointly determined. This chapter restricts itself to a reduced-form approach, using the variables available in the FLACSO survey that most closely correspond to the associated explanation for social exclusion.[21]

Measurement of Group Effects

Table 4.4 presents initial evidence on group effects by reporting the coefficient on the dummy variable for Nicaraguan status in equation (1) for a variety of outcomes. Panel A presents the results for individual outcomes for males; panel B presents the results for individual outcomes for females. Individual outcome variables are poverty, affiliation with the social insurance system, employment, and, for the working sample, hours and informal sector status logarithm of monthly income. Also included are two measures of child integration into the school system: attendance and grade/age.[22] Panel C includes household level variables, including medical attention for household members, household income, poverty, and several variables related to quality of housing.

In panels A and B, columns (1)–(5) report results from the pooled sample from the 1997, 1998 and 1999 National Household Survey and columns (7)–(13) report results from FLACSO. The first column in each group (columns 1 and 7) reports mean differences without controls, and each subsequent column adds a set of controls for basic demographic information (columns 2 and 8),[23] nationality composition of household (columns 3 and 9),[24] job controls for income and hours regressions (columns

[21] While the original idea was joint estimation of the determinants of outcomes using instrumental variables, the initial results did not suggest that more would be learned from that approach than from reduced-form results.

[22] Grade for age is measured as a binary variable indicating whether age minus grade completed minus six is greater than one.

[23] All specifications with the National Household Survey include year dummy variables. Demographic controls include age, age squared, education and education squared.

[24] Composition controls include Nicaraguan head of household married to Costa Rican spouse, Costa Rican head of household married to Nicaraguan spouse, male Nicaraguan head of household (only), and female Nicaraguan head of household (only). The reported coefficients correspond to Nicaraguan head of household and spouse.

Table 4.4. Coefficients on Nicaraguan Status, Urban Metropolitan Area, National Household Survey and FLACSO Survey

Panel A: Males

	National Household Survey						FLACSO Survey						
	Without control (1)	With control (2)	HH type control (3)	Job control (4)	Neighb. FE (5)	N (6)	Without control (7)	With control (8)	HH type control (9)	Job control (10)	Migr. control (11)	Neighb. FE (12)	N (13)
Labor market, adults ages 20–65													
Poverty	.032	–.008	.017		–.025	3,795	.079	–.001	.082		.017	.052	353
	(.023)	(.023)	(.025)		(.029)		(.092)	(.092)	(.094)		(.159)	(.101)	
In labor force	.077	.059	.047		.050	4,807							
	(.020)	(.019)	(.022)		(.024)								
Working	.051	.051	.029		.047	4,807	.114	.132	.064		.056	.060	352
	(.023)	(.023)	(.026)		(.029)		(.060)	(.059)	(.059)		(.100)	(.056)	
Working sample:													
Ln hours week	.083	.077	.071	.060	.055	4,088	.088	.015	–.019	–.014	.008	.042	298
	(.027)	(.027)	(.031)	(.031)	(.035)		(.049)	(.049)	(.050)	(.061)	(.100)	(.063)	
% informal sector	.102	.055	.024		.047	4,086	.022	.028	.025		.008	.065	216
	(.032)	(.031)	(.035)		(.040)		(.093)	(.095)	(.100)		(.166)	(.108)	
Ln monthly	–.354	–.112	–.134	–.160	–.152	3,506	–.283	–.156	–.155	–.211	–.153	–.254	226
income	(.054)	(.045)	(.052)	(.050)	(.056)		(.106)	(.100)	(.104)	(.128)	(.215)	(.128)	
Institutional participation:													
Any social	–.315	–.247	–.224	–.208	–.183	4,796	–.280	–.245	–.237		–.528	–.649	186
insurance	(.028)	(.028)	(.032)	(.030)	(.034)		(.071)	(.093)	(.099)		(.197)	(.172)	
Children 6–14:													
Attend school	–.130	–.132	–.132		–.175	1,749	–.062	–.067	–.082		–.257	–.237	90
	(.030)	(.029)	(.031)		(.039)		(.067)	(.061)	(.065)		(.204)	(.159)	
Grade/Age	–.375	–.602	–.674		–.314	1,551	–.306	.298	.348		.095	.287	141
	(.170)	(.151)	(.155)		(.202)		(.125)	(.109)	(.108)		(.345)	(.200)	

Table 4.4. Coefficients on Nicaraguan Status, Urban Metropolitan Area, National Household Survey and FLACSO Survey (cont.)

Panel B: Females

	National Household Survey						FLACSO Survey						
	Without control (1)	With control (2)	HH type control (3)	Job control (4)	Neighb. FE (5)	N (6)	Without control (7)	With control (8)	HH type control (9)	Job control (10)	Migr. control (11)	Neighb. FE (12)	N (13)
Labor market, adults ages 20–65													
Poverty	.050	–.009	.019		–.013	4,212	.172	.090	.177		.075	.065	430
	(.024)	(.023)	(.026)		(.028)		(.093)	(.092)	(.098)		(.168)	(.107)	
In labor force	.259	.272	.365		.340	5,354							
	(.030)	(.029)	(.031)		(.033)								
Working	.222	.245	.341		.318	5,354	.044	.114	.188		.226	.232	412
	(.030)	(.029)	(.031)		(.034)		(.071)	(.071)	(.074)		(.127)	(.078)	
Working sample:													
Ln hours week	.155	.196	.252	.307	.321	2,673	–.082	–.066	–.069	.017	.071	.117	236
	(.047)	(.048)	(.050)	(.049)	(.056)		(.072)	(.075)	(.084)	(.094)	(.145)	(.127)	
% informal sector	.407	.287	.272		.254	2,673	.378	.323	.330		.393	.415	210
	(.034)	(.031)	(.033)		(.037)		(.092)	(.084)	(.090)		(.148)	(.110)	
Ln monthly income	–.457	–.128	–.091	.073	.085	2,321	–.357	–.206	–.138	.345	.387	.122	229
	(.066)	(.055)	(.058)	(.054)	(.062)		(.155)	(.149)	(.163)	(.171)	(.261)	(.176)	
Institutional participation:													
Any social insurance	–.340	–.305	–.346	–.337	–.330	5,343	–.452	–.458	–.533		–.550	–.439	130
	(.021)	(.021)	(.023)	(.026)	(.029)		(.081)	(.090)	(.080)		(.343)	(.291)	
Children ages 6–14:													
Attend school	–.072	–.064	–.089		–.032	1,619	–.140	–.147	–.156		–.370	–.384	102
	(.025)	(.024)	(.026)		(.032)		(.049)	(.048)	(.054)		(.242)	(.213)	
Grade/Age	–.638	–.538	–.400		–.058	1,432	–.256	–.257	–.317		–.473	–.451	183
	(.144)	(.129)	(.138)		(.173)		(.113)	(.101)	(.105)		(.334)	(.191)	

Table 4.4. (cont.)

Panel C: Households, FLACSO Survey

	Without controls (1)	Head demog. controls (2)	HH controls (3)	Head migr. controls (4)	Neighb. FE (5)	N
Go to doctor	–.090	–.101	–.123	–.137	–.237	342
	(.071)	(.072)	(.081)	(.142)	(.098)	
Poverty	.058	.004	.055	.068	.071	349
	(.062)	(.060)	(.068)	(.118)	(.084)	
Log household income	–.087	.086	–.235	–.188	–.149	341
	(.108)	(.100)	(.104)	(.180)	(.118)	
Housing:						
Inadequate material	.236	.222	.269	.299	.165	349
	(.039)	(.040)	(.046)	(.080)	(.067)	
Inadequate services	.418	.383	.436	.541	.418	349
	(.042)	(.041)	(.048)	(.080)	(.072)	
Overcrowded	.439	.379	.385	.509	.467	349
	(.064)	(.064)	(.067)	(.117)	(.079)	

Notes for Panels A and B: Entries are coefficient on dummy variable for Nicaraguan. All regressions using National Household Survey include year dummy variables. Demographic controls: adult regressions with controls include age, age squared, education, and education squared. Controls in children regressions include age dummy variable.

Household composition controls include nationality composition of head and spouse. Job controls include dummy variables for one-digit sector of economic activity, self-employed status, domestic, and public employee status. Migration controls include years since migration, years since migration squared, and four legal status dummy variables. Sample for affiliation with social insurance is heads only.

Notes to Panel C: Entries are coefficient on dummy variable for Nicaraguan. Demographic controls include age, age squared, years of education and years of education of head. Household controls include nationality composition of head and spouse, number of household members, number of children, marital status and gender of head. Migration controls include years since migration, years since migration squared, and four legal status dummy variables.

4 and 10),[25] and sampling group fixed effects (columns 5 and 12). The FLACSO survey further includes a column with a set of controls for migration that includes legal status and year of arrival information (column 11). The final column for each data set reports the number of observations in the sample (columns 6 and 13).

Because comparable household information was not available in the data sets provided from the National Household Survey, panel C includes only information from the FLACSO survey. The organization of the table is similar to that of panels A and B, starting with mean differences without controls, then adding controls across the columns.

Labor Market Insertion

Patterns in labor market insertion show Nicaraguans to have labor force participation rates and employment rates comparable to or higher than those of Costa Ricans. This is not surprising, given the primarily economic motive for migration, and it does not suggest restricted access to entering the labor market. The most pronounced differences are for females, with employment rates over 20 percent higher than those of Costa Rican females, compared to differences of 5 to 7 percent for males. These differences are similar, though higher for males, when detailed controls including neighborhood fixed effects are included.

The type of insertion is measured crudely with the informal sector and income measures. For males, the proportion of Nicaraguans in the informal sector is higher than among Costa Ricans, though most of the difference is explained with controls. For females, the proportion working in the informal sector is 40 percent higher and remains over 25 percentage points higher when detailed controls are included. The differences in income between Nicaraguans and Costa Ricans are large: 28–35 log points for males and 36–46 log points for females. For males, about half of the difference is explained with basic demographic controls, as shown in columns (2) and (8) of Table 4.4. The inclusion of household composition and job controls leads to little change in the difference. For females, two-thirds to three-fourths

[25] Job controls include dummy variables for one-digit industry, self-employed status, domestic status, and public employee status.

of the difference is explained by basic demographic and household composition controls. After controlling for job characteristics, Nicaraguan females earn *more* than their Costa Rican counterparts. The inclusion of neighborhood fixed effects has little effect on the difference in log monthly earnings for either males or females. In fact, the slight increase in the differential for males when neighborhood effects are included in the FLACSO survey suggests that living in worse neighborhoods affects Costa Ricans' earnings even more than those of Nicaraguans.

Institutional Participation

The four measures of institutional participation are individual access to social insurance, school attendance of children ages 6–14, legal status (shown in panels A and B), and use of a doctor in the last month (shown in panel C). In each, Nicaraguan outcomes are substantially and significantly worse than those of Costa Ricans in the household survey. Rates of access to social insurance are 32–35 percent lower for Nicaraguans, falling to 16–20 percent for males with controls. School attendance rates are 6–14 percentage points lower, with little change when controls are introduced, except that the introduction of neighborhood fixed effects increases the deficit in school attendance of Nicaraguans relative to Costa Ricans.[26]

The FLACSO survey asks whether anyone in the household has been to a doctor in the last month. Nicaraguans are 15 percent less likely to have done so. Controls for age and education of principal earner and household size reduce the coefficient only slightly to 13 percent.

Housing and Household Labor Market Insertion

The most substantial differences besides legal status in the societal integration of Nicaraguans is in quality of housing (shown in panel C). For each of three general measures, binary outcome variables were constructed similar to those used to calculate satisfaction of basic human needs. The first measure is quality of housing materials, which is classified as inadequate if the

[26] For a discussion on Nicaraguans in the Costa Rican school system, see Mora Corrales and Mora (1998).

floor, roof or walls are made of inadequate materials. The second measure is access to electricity, water and sanitation; the housing unit is classified as inadequate if any of those services is not available. The third measure is overcrowding, which is defined as more than two persons per bedroom in the house.

Nicaraguan households are much more likely to live in housing with inadequate materials or lack of access to basic services. Few Costa Rican households, on the other hand, have low levels of housing inadequacy (7–8 percent in each measure). In the FLACSO survey, 24 percent of Nicaraguan households do not have adequate materials and over 40 percent do not have access to basic services. The overcrowding measure is also substantially higher for Nicaraguans—by 44 percent when no controls are included. Three-fourths of Nicaraguan households suffer from overcrowding, compared to 30 percent of Costa Rican households.

The summary measures show 56 percent more Nicaraguan households have inadequate housing in at least one dimension than Costa Rican households and 21 percent more have inadequate housing in all dimensions. These levels are not explained by household head characteristics, household characteristics, head migration characteristics, or neighborhood fixed effects. Only in the case of adding neighborhood fixed effects in the determination of deficient materials is there any change in the Nicaraguan group effect, and the change is relatively small (7 percentage points).

Correlation in Outcomes

While not surprising and indicating nothing about causality, there is a confluence of negative outcomes with a significant relationship between income and other outcomes, including housing quality, use of a doctor, and affiliation with the social insurance system. Using the FLACSO data and the sample from Table 4.4, the correlations are calculated (but not reported) between outcomes using both the levels without controls and the residuals from the estimation of the specification with household controls.

The two-way correlations range in magnitude from 0.1 for the relationship between the log of household income and use of a doctor to 0.21 for the relationship between the log of household income and both access to basic services (negative) and affiliation with the social insurance system. In addition, there is an even stronger correlation between housing quality out-

comes (inadequate materials, access to basic services, and overcrowding) and a significant relationship between affiliation with the social insurance system and other outcomes.

What is a bit surprising is that many of these relationships are much weaker, and not statistically significant, when calculated for Nicaraguan households. In particular, the correlation between housing quality outcomes and variables other than poverty are weaker and statistically insignificant. In addition, the relationship between the residuals from the specification with household controls also yields much weaker and statistically insignificant relationships.

An Initial Portrait of Nicaraguan Integration into Costa Rica

The integration of Nicaraguans into Costa Rica reflects the economic motive for migration. Labor market integration is high and outcomes are better than what would have occurred in Nicaragua. The pattern of labor market integration reflects the lower skills of Nicaraguans as much as restricted access. Though there are indications of occupational segregation, especially among Nicaraguan females, and incomes of Nicaraguans are lower than those of Costa Ricans even after controlling for other factors, the high absorption of low-skilled workers by the Costa Rican economy over the late 1990s may have led to vulnerability in the labor market, rather than exclusion.

In other areas, Nicaraguans show many patterns documented elsewhere for excluded groups. Integration into institutions is lower and quality of housing is lower for the large group of Nicaraguans living in *ranchos* in La Carpio and La Aurora. Differences in other outcomes, especially labor market outcomes, do not show significant differences by level of concentration of Nicaraguans in the neighborhood or *barrio*.

These initial findings lead us to emphasize the migrant nature of Nicaraguans in Costa Rica. In particular, the concept of exclusion may not be appropriate to describe the current situation of Nicaraguans as a group. Rather, the differential participation of Nicaraguans in Costa Rican society, accepted by current migrants as part of a package that includes economic gains, makes the group vulnerable. As Nicaraguans as a group spend more time in Costa Rica, these differences may decline, and migrants who are not successful may return to Nicaragua, perhaps enhancing the possibility for inclusion of those Nicaraguans who remain. On the other hand, the flow of

migrants may yet increase, especially if there is continued new migration and differential outcomes of the children of Nicaraguans in Costa Rica. This could enhance the possibility of exclusion.

Mechanisms of Social Exclusion

The preceding discussion suggests that, as the situation currently stands, Nicaraguans are a vulnerable rather than an excluded group in Costa Rica. Among the factors that will determine whether Nicaraguans as a group become an included or excluded group in the future are the potential for integration of those Nicaraguans who are currently residing and remain in Costa Rica, and the integration of new Nicaraguan migrants. In the view most favorable to inclusion, the only factors leading to worse outcomes of Nicaraguans are time in Costa Rica and legal status, which will mechanistically improve over time for current Nicaraguans as a group, especially once the second generation of Costa Rican citizens reaches adulthood. The alternative view is that there are other mechanisms of social exclusion—including the persistent effects of initial legal status, use of networks, geographic epidemic effects, transferability of human capital, and discrimination—that cause Nicaraguans as a group to be treated differently than Costa Ricans, and that these differences will persist over time. This section explores these possibilities and the effects that they might have on Nicaraguans.

Where possible, variables are included to proxy for the mechanisms of social exclusion in equation (2). For some mechanisms, especially social capital and discrimination, the evidence is more indirect, using the questions in the FLACSO survey related to perceptions, use of networks, and participation in organizations. In these cases, the evidence addresses the possibility that these mechanisms are important, with limited evidence on the magnitude of the impact.

Table 4.5 presents many of the coefficients included to proxy for the mechanisms of exclusion discussed below. Each row includes the results of a regression for the determinants of a separate outcome variable, including logarithm of household income, use of a doctor, inadequate housing materials, inadequate services, overcrowding, household poverty, and the working status, affiliation with the social insurance system and income of the principal household earner. Other controls include those in Table 4.4.

Table 4.5. Inclusion of Proxy Variables for Mechanisms of Social Exclusion: FLACSO Sample of Principal Earners

	Costa Ricans			Nicaraguans								
	Use of networks (1)	Have problems (2)	High presence (3)	Use of networks (4)	Have problems (5)	High presence (6)	Natur–alized (7)	Applied amnesty (8)	Elig., did not apply (9)	Time in C. R. (10)	Time in C.R. sq./100 (11)	N (12)
Ln (household income)	.120	–.352	–.043	–.009	.024	.055	–.098	.116	–.288	–.007	–.025	333
	(.143)	(.140)	(.137)	(.097)	(.115)	(.167)	(.202)	(.152)	(.228)	(.015)	(.051)	
Use doctor	.215	–.011	–.021	.119	–.080	.024	.407	.216	–.147	.012	–.090	332
	(.112)	(.109)	(.099)	(.104)	(.092)	(.126)	(.148)	(.134)	(.223)	(.013)	(.038)	
Inadequate housing mat.	.041	.037	.172	–.085	.129	–.324	–.014	.023	.137	.016	–.092	341
	(.032)	(.032)	(.063)	(.079)	(.080)	(.100)	(.153)	(.137)	(.250)	(.012)	(.035)	
Inadequate services	–.033	.035	.137	–.057	.053	–.123	–.495	–.253	–.170	.011	–.084	341
	(.041)	(.058)	(.062)	(.076)	(.073)	(.077)	(.115)	(.092)	(.225)	(.022)	(.035)	
Overcrowding	.021	–.057	–.006	.003	.058	.046	.026	–.022	.284	–.010	.045	341
	(.102)	(.107)	(.093)	(.074)	(.072)	(.115)	(.133)	(.117)	(.135)	(.012)	(.037)	
Household poverty	–.152	.210	–.089	.024	.051	.007	.071	–.050	.347	.013	–.054	341
	(.079)	(.085)	(.079)	(.077)	(.080)	(.108)	(.182)	(.165)	(.240)	(.013)	(.045)	
Working:												
Males	.014	.063	.105	–.044	–.032	–.114	–.046	–.056	.041	.003	.002	194
	(.026)	(.033)	(.042)	(.028)	(.034)	(.051)	(.042)	(.037)	(.075)	(.004)	(.011)	
Females	.016	.042	–.105	–.016	.017	.037	.132	.155	.271	–.001	–.019	147
	(.094)	(.089)	(.092)	(.051)	(.055)	(.114)	(.240)	(.231)	(.274)	(.014)	(.069)	
Affiliated social insurance:												
Males	–.115	–.098	.135	–.120	–.146	–.098	.317	.232	–.187	.009	–.034	185
	(.108)	(.098)	(.116)	(.104)	(.112)	(.149)	(.184)	(.156)	(.196)	(.017)	(.052)	
Females	–.155	–.086	–.154	–.183	.097	.329	–.043	.024	.211	.032	–.171	129
	(.143)	(.126)	(.126)	(.133)	(.129)	(.179)	(.351)	(.326)	(.423)	(.030)	(.134)	
Head income:												
Male	–.077	–.535	.004	–.068	–.066	.001	.038	–.054	–.407	.005	–.039	156
	(.168)	(.168)	(.232)	(.093)	(.103)	(.230)	(.198)	(.153)	(.289)	(.019)	(.073)	
Female	.055	.015	.166	.004	–.161	–.310	–.263	–.084	.733	–.007	–.094	104
	(.159)	(.175)	(.182)	(.210)	(.215)	(.269)	(.358)	(.284)	(.509)	(.035)	(.142)	

Notes: Numbers in parentheses are robust standard errors. Each row reports the results of a separate regression in which the use of networks and problem variables are interacted with Costa Rican and Nicaraguan status. Controls include age, age squared, years of education, years of education squared, and four dummy variables for household nationality composition. Household level variables (first four rows) include household size and number of children. Individual income regressions include dummy variables for industry, self-employed status, domestic status, and public worker status.

Legal Status

To examine the effects of legal status, three dummy variables are included: naturalized, applied for amnesty, and eligible but not applying for amnesty. The omitted group is those who were not eligible for the amnesty. The reported coefficients are shown in columns (7)–(9) of Table 4.5. While the precision of the estimates is low and there is a relationship between legal status and length of time in Costa Rica (shown in columns 10–11), legal status is a significant determinant of some outcomes. Those who are naturalized or applied for amnesty are (somewhat) significantly more likely to use a doctor and to be affiliated with the social insurance system, indicating greater integration into institutions than other immigrants. There is a large and significant difference in access to basic services by legal status, but no difference in the two other housing outcomes. In other outcomes, there are mixed results, most of which are not significant. For example, those who chose not to apply for the amnesty have worse outcomes in many dimensions, though females in this category are more likely to work and to have higher incomes. When controls for legal status are included, time in Costa Rica does not have a significant impact on outcomes other than the use of a doctor and access to basic household services.

Use of Networks

In the FLACSO survey, there are two questions related to networks: 1) How did you obtain your current job? and 2) Have you used networks to obtain a job, enter a health clinic, obtain housing, obtain credit, or enroll a child in school? The evidence does not suggest that only Costa Ricans have access to networks: there is very little difference between the proportions of Nicaraguans and Costa Ricans, respectively, who have used networks of friends and family. If anything, Nicaraguans are more likely to obtain a job and obtain housing, though the results are not statistically significant. A fairly high proportion of principal earners, however, report the use of networks. More than one-third report using contacts in the labor market, while more than one-fourth report using contacts in health centers and housing. Thus, while the quantity of network use may be similar between Nicaraguans and Costa Ricans, the general use of networks may lead to the quality of networks having an impact on the outcomes of Nicaraguans and Costa Ricans. Columns (1) and

(4) in Table 4.5 report the coefficients on the use of networks. For Costa Ricans, networks have a statistically significant correlation with outcomes only for use of a doctor and poverty. For Nicaraguans, use of networks is not a significant determinant of any outcome variable, and the sign of the coefficients in many labor market variables (household income, poverty, affiliation with social insurance, and individual income) suggests that the networks that Nicaraguans use may not be improving these outcomes.

Geographic Concentration

Much of the recent economics literature on group effects has focused on the role that geographic concentration plays in the development of worse outcomes.[27] Given the housing policy of the Costa Rican government and the effect it has had on the location decisions of Nicaraguan migrants, it is important to document the extent of geographic concentration and the relationship, if any, between concentration and outcomes. The examination of neighborhood effects begins by measuring the extent of geographic segregation. The potential effect of geographic concentration is then examined by trying to explain neighborhood effects from equation (1) with variables related to the presence of Nicaraguans.

Measures of Segregation, Isolation and Exposure

The National Household Survey is used to examine geographic concentration at the level of the metropolitan area. The primary sampling unit, or *segmento*, is used as a rough measure of neighborhoods. The *segmentos* of the surveys for 1997 to 1999 are divided into those that are 0–5 percent Nicaraguan, 5–10 percent Nicaraguan, 10–15 percent Nicaraguan, 15–25 percent Nicaraguan, and over 25 percent Nicaraguan. There is significant representation of Nicaraguans in each category for each year, with a greater proportion in the lower percentage groups in 1997 compared to 1999. In 1997, only 29 percent of

[27] Most relevant for this chapter are Cutler and Glaeser (1997), Cutler, Glaeser and Vigdor (1999), Borjas (1998), Glaeser, Laibson and Sacerdote (2000), and Bertand, Luttmer and Mullinathan (2000). Several other authors, including Lazear (1999), Portes (1987) and Portes and Jensen (1989), have discussed the possibility that the formation of enclaves may improve residents' economic opportunities.

Nicaraguans lived in areas with over 15 percent Nicaraguans; by 1999, this proportion had grown to 58 percent, in part due to the growth in the number of Nicaraguans included in the survey. In each year, 80 to 85 percent of Costa Ricans lived in areas with less than 5 percent Nicaraguans and an additional 10 percent lived in areas with 5 to 10 percent Nicaraguans.

Despite this initial evidence of geographic concentration, Nicaraguans are not isolated from Costa Ricans. Costa Ricans are a majority even in the areas where most of the Nicaraguans live. In only one *segmento* in one (1999) of the three survey years were Nicaraguans a majority, and in only nine *segmentos* (four in 1997, two in 1998 and three in 1999) did Nicaraguans account for more than 25 percent of residents.

To formalize these ideas, measures of segregation, isolation and exposure were calculated using data from the metropolitan area. The Duncan index—which measures the proportion of Nicaraguans that would have to move from disproportionately Nicaraguan survey areas to other areas to have a proportional population in each survey area—ranges from 0.62 in 1997 to 0.60 in 1999, and shows that Nicaraguans are geographically segregated.[28] The measure of isolation—that is, the proportion of Nicaraguans in the survey areas where Nicaraguans live, scaled between 0 and—is low, but does increase from 0.12 in 1997 to 0.23 in 1999. This is not surprising, given the large number of Costa Ricans living in survey areas in which Nicaraguans live. Similarly, the measure of exposure—the proportion of other groups in survey areas where Nicaraguans live, again scaled between 0 and 1—also does not show lack of exposure to other groups. The figure was 0.88 in 1997, falling to 0.77 in 1999.[29]

Including Neighborhood Variables for Nicaraguan Presence in Outcome Regressions

Neighborhood fixed effects are included in the regressions in Table 4.4. To summarize those findings, the coefficient on Nicaraguans' status changed in

[28] Values between 0.3 and 0.6 are considered to show some segregation; over 0.6 is high segregation.

[29] For comparison, similar measures are calculated for other divisions of the population by characteristics that might explain Nicaraguan segregation: poor versus non-poor, and primary education or less versus those with more than primary education. For each of these divisions, the Duncan index is much lower (0.29–0.49), isolation is similar (0.12–0.20), and exposure is also similar (0.80–0.87).

a way that indicated that Nicaraguans experience worse outcomes on the basis of neighborhood only for the outcome variables of affiliation with the social insurance system and grade/age of children ages 6–14 in the household data, and only for housing quality variables in the FLACSO data. Despite this weak evidence that neighborhood effects are important for Nicaraguans, the possible importance of Nicaraguan concentration, rather than all neighborhood characteristics, is explored by including variables related to the concentration of Nicaraguans in the regressions for outcomes.[30] Dummy variables are included in Table 4.5 for living in neighborhoods with a high presence of Nicaraguans for Costa Ricans (column 3) and Nicaraguans (column 6).

These results are also weak. Those living in *barrio*s with a high presence of Nicaraguans have worse outcomes only in the case of housing materials (both Nicaraguans and Costa Ricans) and working status (Nicaraguan males). The data also suggest that Nicaraguan females are more likely to be affiliated with the social insurance system in these *barrios*, a finding that runs counter to the existence of negative externalities.

Occupational Concentration and Transferability of Human Capital

Table 4.6 shows the occupational concentration of Nicaraguan workers in metropolitan San José and their work status prior to migration. The first column presents the five most represented three-digit occupations for males and females calculated from the 1999 National Household Survey. The second column shows the same information from the FLACSO survey. The row labeled "These occupations" includes the totals for all of the listed occupations. In each, 40–50 percent of males are employed as vendors, carpenter's assistants (or other construction), *talabarteros* (leather workers), and security guards. Females are even more concentrated, with 44–51 percent employed as domestics and 71–81 percent employed in the three-digit classification associated with vendors, sewing, domestics, restaurants and cleaning occupations.

[30] The qualitative results are similar when the Nicaraguan concentration variables are included in the household data. Because the other variables in Table 4.6 are not available in the household survey, only the results from the FLACSO data in the table are reported.

Table 4.6. Transferability of Skills

Panels A and B: Occupation Distribution, Urban Metropolitan Area

			FLACSO Survey Principal earner only: Status in Nicaragua	
	1999 Household Survey	Occupation in C.R.	Working	Same 3-digit
Panel A: Males				
Sellers	8.13	4.12	100.00	0.00
Carpenter asst.	14.63	28.35	66.67	30.30
Other construction		6.70	88.89	22.22
Skilled manuf.		7.22	75.00	0.00
Talaberteros	5.69			
Security	9.76	5.67	71.43	0.00
These occupations	38.21	47.94	73.33	19.67
All occupations			78.63	19.49
Panel B: Females				
Sellers	6.67	6.21	100.00	25.00
Sewing	4.76	4.83	83.33	50.00
Domestic	51.43	44.83	68.97	17.24
Restaurant/Hotel	12.38	5.52	100.00	37.50
Cleaning	5.71	9.66	80.00	0.00
These occupations	80.94	71.05	78.85	23.08
All occupations			71.43	19.23

Panel C: Coefficients on Years of Education (log monthly income)

	1999 Household Survey	2001 FLACSO Survey
Males	.096	.061
	(.005)	(.012)
***Nicaraguan**	–.036	–.052
	(.021)	(.019)
Females	.135	.073
	(.007)	(.016)
***Nicaraguan**	–.028	–.054
	(.029)	(.026)

Note: Principal earner sample sizes for individual occupations are small.

The final two columns report the work status and occupation in Nicaragua for principal earners only. The bottom row labeled "All occupations" for males and females, respectively, includes the sample of all workers in Costa Rica. Prior to migration, 79 percent of males and 71 percent of females employed in Costa Rica were working in Nicaragua. Of these, only 19 percent were working in the same three-digit occupations. The sample sizes for some three-digit occupations are small when calculated for principal earners only, but it is worth noting that of the most represented occupations for males in Costa Rica, only those employed in construction have a high percentage that were employed prior to migration in the same occupation. Among females, only 17 percent of those employed as domestics in Costa Rica were employed in the same occupation in Nicaragua.

To further examine the possibility that human capital is not completely transferable, or that the quality of Nicaraguan education is lower than that in Costa Rica, the coefficient on years of education from log monthly earnings regressions is calculated separately for Nicaraguans and Costa Ricans in regressions similar to those presented in Table 4.4. The National Household Survey shows a return to education for Nicaraguan males of about two-thirds that for Costa Rican males and insignificantly different for females. The FLACSO survey shows much lower returns to education for Costa Ricans and very little relationship between education and earnings for Nicaraguans.

These findings provide fairly strong support for the idea that skills are not transferable or that the quality of education in Nicaragua is much lower than that in Costa Rica. Combined with the patterns on occupation, patterns of skill transferability—and the intergenerational transmission of skill—represent an important factor in explaining the future condition of Nicaraguans in Costa Rica.

Social Capital

The FLACSO survey included several questions related to the existence and use of organizations. In general, a significantly higher proportion of Costa Ricans than Nicaraguans report the existence of a type of organization in the *barrio*; Costa Ricans have a slightly higher but still low and insignificant rate of participation in that type of organization; and Nicaraguans think that participation is more important. The most prevalent and important

types of organizations identified are religious ones, followed by neighborhood groups.

The low rates of participation in organizations and the similarity in participation rates of Costa Ricans and Nicaraguans suggest that social capital, as manifested in participation in organizations, plays a minor role in explaining differences between Costa Ricans and Nicaraguans. The disparity between the two groups in identifying the existence of organizations, combined with the stated importance of organizations by Nicaraguans, does suggest that Nicaraguans may be uninformed.

Discrimination

The large and unexplained differences between Nicaraguans and Costa Ricans in many outcomes are consistent with, but do not show, discrimination against Nicaraguans. Although reliable direct measures of discrimination against Nicaraguans are unavailable, the FLACSO survey does provide some evidence on whether discrimination could be a determinant of the relative outcomes.

In the FLACSO survey, there were three sets of questions related to perceptions of discrimination. The first set asked whether there is discrimination in Costa Rica related to the labor market, health centers, housing, credit and school enrollment. The striking results show that over half of Costa Ricans (58 percent) think that Nicaraguans are discriminated against in the labor market and between one-third and two-thirds of Costa Ricans think there is discrimination against Nicaraguans in health centers, credit and housing. Among Nicaraguans, the proportion that believes they are discriminated against is even higher. Eighty-three percent believe there is discrimination in the labor market, 55 percent believe there is discrimination in health centers, and 67 percent believe there is discrimination in housing. The high proportion of all people who believe that there is discrimination against Nicaraguans, and the differential perception between Nicaraguans and Costa Ricans, suggests that discrimination may play a role in the outcomes of Nicaraguans. It is interesting to note that less than one-fourth of both Costa Ricans and Nicaraguans think that there is discrimination in school enrollment.

The second set of questions related to the perceptions of Costa Ricans and Nicaraguans of the attributes of the majority of persons of each nation-

ality. The principal earners were asked if Costa Ricans were good workers, honest, trustworthy, and racist, respectively. The questions were repeated for Nicaraguans. It is important to note that all interviewers were Costa Rican and that this may have had an effect on the answers in Nicaraguan households. When asked if most Costa Ricans are hardworking, honest or trustworthy, Nicaraguans responded more favorably than the Costa Ricans themselves. However, 68 percent of Nicaraguans said most Costa Ricans are racist, while less than one-fourth of Costa Ricans consider themselves to be racist.

Nearly all Nicaraguans consider themselves hardworking, honest and trustworthy, and do not consider themselves to be racist. While nearly all Costa Ricans consider Nicaraguans to be hardworking, they are less favorable in describing the honesty and trustworthiness of Nicaraguans, and 41 percent of Costa Ricans consider Nicaraguans to be racist. These patterns are consistent with a view, expressed independently by several observers, that as a group, Costa Ricans consider that Nicaraguans' main contribution to Costa Rican society to be in the labor market, and especially in low-skill occupations.[31]

The third set of questions asked both Costa Ricans and Nicaraguans if they avoid having Nicaraguan friends, and asked Nicaraguans if they avoid Costa Rican friends. The responses do not show strong avoidance on either side; approximately 20 percent of Costa Ricans reported avoiding having Nicaraguan friends, while 6 percent of Nicaraguans reported avoiding having Costa Rican friends.

Taken at face value, the results of these questions provide some evidence that discrimination may be an issue and is likely to be a determinant of some of the differential outcomes of Nicaraguans. Given the responses indicating that discrimination is an issue, it is somewhat surprising that there are not greater differences in outcomes for Nicaraguans. One possible explanation is that although discrimination exists, its effect is small. As in other

[31] A survey by Sandoval et al. (1999) also reported a high percentage of Costa Ricans who think Nicaraguans are hard workers. They also find a high proportion that think Nicaraguans help the economy of Costa Rica, and a high proportion that think Nicaraguan customs are different than those of Costa Ricans. The survey found Costa Ricans of a mixed view as to whether Nicaraguans are equal to them, whether they would like to have Nicaraguan relatives, whether Nicaraguans only bring problems, and whether Nicaraguans should be permitted to enter Costa Rica. Less than half of Costa Ricans supported the amnesty.

studies of discrimination, determining the magnitude of any effect is difficult without adequate measures of the phenomenon. A second explanation is that the survey question does not adequately measure the extent of discrimination.

Summary

The proxy measures for the mechanisms through which social exclusion might affect outcomes of Nicaraguans are imperfect, and the findings should be taken as suggestive rather than definitive. Nevertheless, there is evidence that uniquely Nicaraguan characteristics are likely to be affecting outcomes in a way that is consistent with some explanations for social exclusion. The results do not distinguish well among the explanations that are consistent with the evidence, but they do rule out some explanations.

The explanations that are consistent with the evidence include the following. First, labor market integration appears to be related to the occupational concentration that results from the lack of transferability of human capital, the use of suboptimal networks, or discrimination. Second, housing outcomes are affected by time in Costa Rica and legal status. Third, there is fairly strong evidence that Nicaraguan integration into Costa Rican institutions is affected by factors particular to Nicaraguans, especially legal status.

On the other hand, there is not much evidence of the effects of social capital or Nicaraguans' use of networks; these findings, though, could be the result of imperfect measures for those mechanisms.

Perhaps the most surprising finding is that there is little relationship between geographic concentration of Nicaraguans and many of the outcomes under consideration. This result has significant policy implications, given that geographic concentration of Nicaraguans is high, that it is increasing over time, and that the current housing policy of the Costa Rican government is likely to lead to increased concentration as Nicaraguan migration continues. There are alternate means of exploring this finding, such as looking at *rancho* communities rather than the proportion of Nicaraguans in a given neighborhood. Nonetheless, Nicaraguan presence continues to have strong effects on housing outcomes, but weaker results for other outcomes. The results suggest that the effects of geographic location on outcomes, where there are effects, operate through characteristics other than nationality, such as poverty.

Of the mechanisms for exclusion that have some support, such as transferability of skills, legal status and discrimination, only discrimination is the result of denial of access. Nicaraguans who migrate have chosen to do so, even though their skills are not transferable; and Nicaraguans who go to Costa Rica illegally have chosen to do so for economic rather than institutional reasons.

Conclusions and Policy Recommendations

This study provides a benchmark against which the future inclusion or exclusion of Nicaraguans in Costa Rica may be measured. While many outcomes of Nicaraguans are worse than those of otherwise similar Costa Ricans, the data do not show many of the common features associated with social exclusion. First, the negative relationship between group effects and geographic concentration, which dominates much of the discussion of group effects of African-Americans in the United States and social exclusion in Europe, is not strong in these data. Second, integration into institutions is correlated with legal status. Third, the worse outcomes of Nicaraguans in the main public services of education and health appear to be more related to factors other than denial of access or discrimination. Fourth, the general paternalistic role of the Costa Rican state towards the poor has been extended, to a certain extent, towards Nicaraguans. And fifth, the strong performance of the Costa Rican economy during the 1990s has generated demand for labor, especially unskilled labor.

Despite this mixed evidence on current levels of exclusion, Nicaraguans as a group are vulnerable. First, they are very concentrated in terms of both geography and occupation. Second, their housing conditions, as measured by quality of housing, access to services and overcrowding, are significantly worse than those of Costa Ricans. Third, integration into the political process is limited by legal status, and the type of integration is strongly correlated with geographic concentration. Finally, the effects of future migration on Nicaraguans as a group are likely to create a stronger relationship between geographic concentration and outcomes, and to test the capacity of the Costa Rican labor market to absorb low-skill labor. In particular, current housing policy makes it likely that future Nicaraguan immigrants will be concentrated in *rancho*-style *asentamientos*.

The main policy recommendations from these findings are divided according to the three aspects of Nicaraguan integration into Costa Rican society. First, the characteristics of Nicaraguans, especially low income levels, make it more likely that they will have worse outcomes in other dimensions. Policies in these areas are related to improving the quality of such factors as access to school, rather than improving Nicaraguan status per se. While these policies redistribute resources from native Costa Ricans to Nicaraguans, they are likely to produce net benefits for Costa Rica as a whole, given the permanent nature of the current migration for many Nicaraguans.

Second, Nicaraguan status itself affects outcomes by increasing the costs of residing legally in Costa Rica. It also reduces the benefits of investment in institutional integration if migration is not permanent, and through the experience of discrimination. Because the main costs of migrating are the cost of the Nicaraguan passport, which is outside the control of the Costa Rican government, the only possibility for increasing the probability of legal residence is to eliminate the passport requirement, as in the amnesty that was offered in 1999. When legal status proxies for permanence of migration, the outcomes for which such policies will have an impact are limited.

Third, while the finding that there is little relationship between outcomes and geographic concentration is weak, there is a possibility that this relationship will strengthen as the number of Nicaraguans in Costa Rica increases. This possibility is greater if one considers the finding that there does appear to be some evidence of discrimination against Nicaraguans. Policies in this area should recognize the likelihood of continued urbanization in metropolitan San José and the benefits of planned urbanization over unplanned takeovers. As in the case of anti-poverty policies, policies designed to rationalize urban development in poorer areas involve some redistribution to Nicaraguan households and in all likelihood would have net benefits to all those residing in metropolitan San José. But unlike any additions to current anti-poverty programs, housing policies would not be aimed specifically at Nicaraguan immigrants.

CHAPTER FIVE

Geographic Isolation and Labor Markets in Rural El Salvador

Ana Regina Vides de Andrade
Anabella Lardé de Palomo
Lissette Calderón Martínez

El Salvador is characterized by the vulnerability of large segments of society, the absence of opportunities, and the social exclusion of many productive people, especially in the rural areas that are home to 43 percent of the population. An estimated 61.5 percent of the rural population live in poverty, only 35 percent have access to running water, and only 56 percent have access to electricity. On average, people living in rural areas complete only 3.2 years of schooling. Rural emigrants cite insufficient job opportunities and widespread violence as reasons for fleeing to the city, or for leaving the country altogether (DYGESTIC, 1999, 2000).

In terms of economic participation, 37.3 percent of the Salvadoran labor force lives in rural areas (DYGESTIC, 2000). From 1992 to 1998, the labor force participation rates in rural areas increased by 3 percent for men (from 43 to 46 percent) and by 5 percent for women (from 23 to 28 percent) (DYGESTIC, 1993, 1999). A large number of these rural workers are clustered in non-remunerated family employment (13 percent), retail trade (15 percent), home production (31 percent), and subsistence agriculture (11 percent)—all economic activities where labor productivity is extremely low. Rural labor incomes are, on average, only 49 percent of urban levels; in fact, 90 percent of rural workers make less than the urban minimum wage. These workers are a great concern for policymakers, and several questions are typically posed regarding them. Who are these workers? Is their geographical location a determinant of poor labor market outcomes? Are they socially excluded? Even if they had the necessary human capital, would they have equal access to better conditions, job networks and good jobs?

Social exclusion in El Salvador, however, especially in the rural areas, is still not well understood. Little is known about the extent of social exclusion, or the consequences that social exclusion may have for labor markets. Moreover, the mechanisms that lead to such social exclusion have not been extensively studied. This lack of knowledge has in turn prevented the formulation of policies to promote a more inclusive society, produce more egalitarian labor income distribution, and improve the efficiency and equity of human resources.

This chapter attempts to address the gap in current research by examining one aspect of social exclusion, namely, the geographic isolation of people living in El Salvador's rural areas and its impact on three labor market outcomes: labor force participation decisions, sector of employment, and labor income. It is hypothesized that living in geographic isolation has a negative impact on rural workers' labor outcomes and that geographic isolation—through a combination of security hazards and increased transaction and working costs—depresses labor force participation rates, increases the likelihood of working in low-productivity jobs, and results in lower labor income levels.

Data

This chapter uses micro-level data from the 1999 Rural Household Survey carried out by the Salvadoran Foundation for Economic and Social Development (FUSADES). The survey was designed as a stratified random sample representative of the rural population at a 10 percent significance level. The Rural Household Survey design is similar to the Living Standards Measurement Surveys, which include a broad array of socioeconomic information on individuals and households. The data cover education, family composition, distances (measured in time and kilometers from place of residence to closest post office, workplace, primary school, health care facility, etc.), health, demographics, migration, remittances, agricultural activities, employment, sector of employment, income, and perception of access to financial and health care markets.[1]

Prior to the 1999 Rural Household Survey, FUSADES carried out similar surveys for 1995 and 1997. These surveys, however, lacked key variables

[1] A copy of the 1999 questionnaire can be found on Ohio State University Website at: http://www-agecon.ag.ohio-state.edu/programs/ruralfinance/Basis.htm

needed to analyze geographic exclusion. This chapter is therefore a cross-sectional study using the yearly data from 1999, as that year's survey contained the necessary elements to examine the link between geographic exclusion and labor market outcomes.

Data Description

This study focuses on individuals between 16 and 65 years of age. Table 5.1 provides descriptive statistics for the main variables of interest for workers and non-workers. On average, working individuals are older, have less edu-

Table 5.1. Worker and Non-Worker Basic Statistics by Gender

	Whole sample		Men		Women	
	Non-workers	Workers	Non-workers	Workers	Non-workers	Workers
Variable	**Mean**	**Mean**	**Mean**	**Mean**	**Mean**	**Mean**
Age	32.79	34.77*	30.09	34.89*	33.42	34.60
Marital status (married = 1)	0.27	0.31	0.12	0.31*	0.31	0.30
Household members	6.99	2.81	6.72	6.97	7.05	6.87
Children in household	2.49	2.66	1.86	2.61*	2.64	2.71
Literacy	0.75	0.78	0.83	0.80	0.73	0.75
Schooling	4.70	4.36**	6.63	4.50*	4.24	4.16
Avg. schooling of household	2.97	2.73*	3.47	2.69*	2.86	2.78
Annual remittances (col.)	3,959	2,399*	5,051	2,151*	3,703	2,750**
No. of household migrants	1.05	0.79*	1.02	0.79	1.05	0.78*
Electricity	0.72	0.69	0.79	0.68*	0.71	0.70
Potable water	0.47	0.46	0.44	0.46	0.48	0.46
Belongs to an organization	0.31	0.34	0.35	0.34	0.30	0.35**
Organizations belonged to	0.38	0.46*	0.46	0.47	1.27	0.46*
Time the family has lived in community (years)	60.45	62.35**	60.87	61.93	60.35	62.94*
Distance to paved road (km)	4.67	5.07	3.97	5.15	4.83	4.96
Distance to post office (km)	5.19	5.51	5.27	5.54	5.18	5.47
Distance to market (km)	7.63	7.81	7.14	7.94	7.74	7.62
Time to paved road (minutes)	28.63	33.11*	24.28	33.33*	29.67	32.78
Time to post office (minutes)	32.99	35.83**	33.22	35.93	32.94	35.68
Time to market (minutes)	40.55	42.66	39.68	43.12	40.75	42.02

* Significant at the 5 percent level.
** Significant at the 10 percent level.
Source: FUSADES (1999).

cation, have fewer migrants in the household, receive lower remittances, and participate more in community organizations than non-workers. In addition, the relatives of workers live in the same community longer, and in terms of geographic location workers live farther from post offices, markets and paved roads than non-working individuals.

Regardless of gender, on average individuals live four to eight kilometers (32 to 35 minutes away) from the closest paved road, market and post office and commute 18 minutes to work. Also, 21 percent travel an hour or more to reach the closest paved road. Whether geographic isolation affects the decision to work, the sector of employment and labor productivity is explored in this study.

Working women receive lower remittances, have fewer migrants in the household, and belong to more organizations than non-working women, and their relatives remain in the same community longer than those of non-working women. Working men spend more time getting to the closest paved road than non-working men, are more likely to be married, and have more children living in the same household. The differences between working and non-working men—i.e., age, schooling, marital status, and number of children in household—suggest the importance of estimating separate labor market equations by gender.

Table 5.2 examines three sectors of employment by gender: salaried, home production, and a mixed sector. The salaried and home production sectors were divided into agricultural and non-agricultural activities. The mixed sector aggregates a large group of individuals who, during a particular year, work in several sectors. There are substantial gender differences in the sector of employment. Women are heavily represented in the home production and non-agriculture salaried sectors. Men are mostly concentrated in the mixed sector, which accounts for men's higher participation rates.

Total labor income and hours worked by sector and gender are shown in Table 5.3. Working men earn higher total income than women in all sectors but the mixed sector. At the same time, men work longer hours than women in all sectors but the salaried non-agriculture sector. Self-employed men working in agricultural activities earn 47 percent more labor income and work twice as many hours as women in the same sector. On the other hand, men working in salaried non-agriculture activities (i.e., *maquila*) earn 45 percent more than women while working the same number of hours.

Table 5.2. Workers by Sector and Employment Status by Gender
(In percent)

Sector and status	Men	Women	Both
Participation in labor force	90	60*	75
Sector			
I. Salaried			
Agriculture salaried	13	10	12
Non-agriculture salaried	18	21**	19
II. Own production			
Agriculture own production	20	32*	25
Non-agriculture own production	4	17*	10
III. Mixed sectors	45	19*	34

Note: This table uses a sample of individuals 15–65 years of age; workers are individuals working at least 26 hours during the whole year.
* Significant at the 5 percent level.
** Significant at the 10 percent level.
Source: FUSADES (1999).

Table 5.3. Labor Income and Hours Worked by Sector and Gender

	Hours worked		Total labor income	
Sector and status	Men	Women	Men	Women
Sector				
I. Salaried				
Agriculture salaried	1,715	774*	¢ 8,481	¢ 3,720*
Non-agriculture salaried	2,218	2,306	¢ 20,497	¢ 14,675*
II. Own production				
Agriculture own production	1,334	628*	¢ 5,136	¢ 1,726*
Non-agriculture own production	2,380	1,879*	¢ 22,145	¢ 14,661**
III. Mixed sectors	2,052	1,660*	¢ 9,166	¢ 10,133
Total	1,907	1,409*	¢ 10,796	¢ 8,486*

Note: This table uses a sample of individuals 15–65 years of age. Total hours and labor income are given per year. Workers are individuals working at least 26 hours during the whole year.
* Significant at the 5 percent level.
** Significant at the 10 percent level.
Source: FUSADES (1999).

Women who work in agricultural activities, whether salaried or in home production, work fewer hours than men.

Table 5.4 shows total labor income by gender and schooling level. Overall, more educated workers earn more than less educated workers. Men and women with some high school education (10 school years and more) earn 45 and 29 percent more, respectively, than workers with four to six years of schooling. Table 5.4 also shows labor force participation rates by schooling level. Individuals with less education have higher participation rates than individuals with higher education levels. Conversely, women with no education have lower labor participation rates than women with some primary schooling. The total labor income differentials between genders are especially large for workers with no education, and between those with four to six years of schooling. However, the total labor income differential between men and women with seven and more years of schooling is not significant.

To assess the characteristics of people living in greater or lesser isolation, all individuals were first sorted and ranked based on the time (in minutes) they needed to reach the closest paved road, so that five groups could be created. The fifth quintile refers to individuals with the longest travel times to the nearest paved road, and therefore living in the greatest isolation. The first quintile consequently refers to the least isolated population. Table 5.5 shows selected gender characteristics for the first and fifth quintile. Regardless of gender, individuals living in less isolated places have more education (+years of schooling) than individuals living farther from paved roads.

Table 5.4. Labor Incomes by Education Level

	Total labor income				Labor force participation rates	
Schooling level	Whole sample	Men	Women	T-test	Men	Women
None	6,519	7,621	5,209	2.44	0.92	0.59
1–3	8,732	9,032	8,354	0.45	0.94	0.64
4–6	10,982	12,319	8,665	2.54	0.95	0.62
7–9	10,670	11,323	9,755	0.90	0.86	0.60
>10	15,910	15,888	15,944	–0.02	0.76	0.57

Source: FUSADES (1999).

Table 5.5. Characteristics by Geographic Exclusion Level

	Men		Female	
Characteristics	**– Isolated**	**+ Isolated**	**– Isolated**	**+ Isolated**
Education	5.38	3.56*	5.21	3.52 *
Age	34.41	34.62	34.27	32.29
Marital status (Married = 1)	0.29	0.36	0.33	0.37
Household members	6.71	6.60	6.81	6.58
No. of children living in household	2.22	2.63	2.40	2.72
Land ownership	0.63	0.72	0.65	0.65
Remittances (colons per year)	¢3,869.64	¢2,575.27	¢3,534.23	¢3,861.84
Parents living/born in same place	0.73	0.87*	0.78	0.85
Land inherited	0.22	0.32**	0.25	0.40*
Land given by govt.	0.09	0.17*	0.07	0.09
Working (Yes =1)	0.85	0.97*	0.61	0.68
Working as salaried in agric.	0.06	0.12	0.04	0.06
Working as salaried in non-agric.	0.20	0.07*	0.18	0.06*
Working as agric. own prod.	0.14	0.24*	0.17	0.36*
Working as non-agric. own prod.	0.09	0.04	0.12	0.05**
Mixed sectors	0.36	0.49*	0.09	0.15
Hours worked	1,748.81	1,869.56	987.36	675.79*
Income (average per hour)	5.21	4.17	4.13	2.48*
Time to closest paved road	2.56	115.56*	2.48	114.95*
Time living in same community	32.45	40.91*	31.32	38.70*
Total labor income	12,702.48	9,269.31	6,971.18	2,544.04*
Total sample	187.00	90.00	208.00	95.00

Note: This table uses the time to closest paved road as a measure of geographic exclusion.
* Significant at the 5 percent level.
** Significant at the 10 percent level.
Source: FUSADES (1999).

Men living in geographic isolation have less schooling, reside where their parents and relatives have lived or were born, and have inherited land or obtained it through a government program. In addition, men living in isolated areas are more likely to be self-employed in agricultural activities or work in more than one sector during a given year; their labor income is 60 percent lower than that of men who are less geographically excluded. Women living in greater isolation are also more likely to have inherited their land, work fewer hours, and work in home agriculture production; they make only 36 percent of what women living closer to a paved road earn.

Men living in isolation have resided for longer periods in the same community where their parents and relatives lived or were born. This may suggest that the decision by people to remain in one location may be affected by their parents' social capital accumulation. Regardless of gender, remittances and degree of geographic isolation were not correlated.

Theoretical Framework

The large body of literature on social exclusion is mostly descriptive or theoretical in nature, discussing concepts and social exclusion indicators and elaborating on what exclusion is or should be about. Gore and Figueiredo (1997) recommend measuring social exclusion with both multiple social and material deprivation indicators (using the categories of Townsend, 1993) and by including groups' perceptions of their position in society in order to determine if they have chosen to exclude themselves from society.

For this chapter, social exclusion is defined as the process by which "individuals or groups are partially or totally excluded from their participation in the society where they belong" (European Foundation, 1995; Trouillot, 2000), a process that results in a "denial of equal access to opportunities imposed by certain groups of the society upon others" (IDB, 2000). This exclusion process is accumulative and multi-dimensional (Figueroa, 2000, and Gacitua and Davis, 2000), and the concept assumes that at the individual level social exclusion is involuntary (Gore and Figueiredo, 1997).

One area where there is a significant body of empirical research is on geographical segregation ("ghetto neighborhoods") and racial and ethnic tension.[2] The literature on this subject analyzes three distinct issues. The first set of studies measures the extent of geographic segregation faced by particular groups by counting the number of persons who reside in particular geographic areas and calculating various segregation indices from these counts.[3] These studies focus on residential segregation and are concerned with the impact of social exclusion on large numbers of individuals/workers living in isolated geographical areas. Such an impact may foster cultural

[2] See Borjas (1997 and 1995a), Cutler, Glaeser and Vigdor (1997), Case and Katz (1991), Crane (1991), Evans, Oates and Schwab (1992), Jencks and Meyer (1990), and Manski (1993).

[3] See Bean and Tienda (1987), Massey and Denton (1989), and McKinney and Schnare (1989).

attitudes, social contacts and economic opportunities that affect individuals throughout their lives.

A second set of studies examines the implications of residential segregation on labor market outcomes.[4] Borjas (1995a) found that earnings of children are strongly affected not only by parental earnings, as traditional models of intergenerational income suggest, but also by the earnings of the ethnic group of the parents' generation, or what he calls "ethnic capital." Borjas suggests that the ethnic neighborhood is one of the mechanisms through which ethnic externality works, linking residential segregation and human capital accumulation. Ethnic capital is measured as the mean earnings of the ethnic group in the parents' generation. This variable is then utilized as a proxy for the socioeconomic background of the neighborhood where the children were raised.

A third strand of research debates the identification problem, where the same unobserved factors that lead to a particular location choice might also lead to other socioeconomic outcomes. These studies conclude that measures of residential segregation and their impact on labor market outcomes may only reflect spurious correlation.[5] To address the endogeneity issue of neighborhood choice, Borjas (1997) examines how individuals choose the neighborhood where they wish to reside. His findings suggest that choice of neighborhood depends on both the household's skills and economic resources and on aggregate characteristics of ethnic groups.

Another group of studies relevant to this chapter focuses on the spatial separation or mismatch between poor workers and job location. Kain (1968, 1992) studied the effects of job decentralization (factories moved from inner cities where many minority groups live to suburban areas) and housing discrimination on both the spatial distribution of black workers and on their ability to find work (see also Holzer, 1991). He suggested that minorities' poor labor outcomes were due to the spatial mismatch between workers and jobs locations. Since Kain's seminal work in 1968, a vast number of studies have debated all aspects of spatial mismatch.[6] Most conclude that

[4] See Cutler and Glaeser (1997), Elliot et al. (1996), Borjas (1995a), Crane (1991), and Case and Katz (1991).

[5] See Evans, Oates and Schwab (1992), Jencks and Meyer (1990), and Manski (1993).

[6] See Bell (1974), Madden (1980), Madden and White (1980), Reid (1985), Vroman and Greenfield (1980), McLafferty and Preston (1992), Ihlandfeldt and Sjoquist (1990 and 1998), Ihlandfeldt and Young (1996), Sánchez (1999), Sawicki (2000), and Thompson (1997).

the separation of neighborhoods from entry-level jobs has a negative impact on the success of minorities in the labor market. Moreover, spatial separation of poor workers from entry-level jobs creates not just physical separation but also spatial isolation from job networks and information. Spatial separation increases when public and private transportation for poor individuals is weak, lessening their mobility.

Methodology

This chapter limits its scope to the geographic separation of workers with respect to local markets. The labor market outcomes of individuals living close to urban jobs and closest to paved roads are compared to the outcomes of individuals living further away. Two measures of geographic exclusion are used to study the effect of geographic isolation: distance to the closest paved road (kilometers) and a location index. The content, advantages and limitations of these two measures of geographic isolation are explained below.

The methodology consists of two steps. First, factors are identified that are correlated with the choice of where to live. This step explores the link between choice of location and intergenerational income effects (parents' land ownership, place of residence, and socioeconomic status), remittances, landownership, household income/resources, and human capital variables. The identification of these factors provides the instrumental variables estimates used to explore the relation between location and labor market outcomes. As a second step, adjusting for self-selection due to individuals' location choice, instrumental variable estimation is used where the predicted value of the geographic isolation measure is included in each of the three labor market outcome estimations: labor force participation, sector of allocation, and labor income per gender.

Measures of Geographic Isolation

The Rural Household Surveys contain information on individuals' place of residence, topographic characteristics of the land, and exact residential coordinates (latitude and longitude);[7] in addition, it contains detailed information

[7] The topographic characteristics and coordinates data were still being processed at the time of this study and could not be used.

on the time and distance it takes individuals to travel from their place of residence to the closest postal office, primary school, secondary school, health care facility, paved road, market, and bus station, among other destinations.

For the purposes of this chapter, "travel time to closest paved road" is used as one measure of geographic isolation. Access to paved roads is believed to decrease transaction costs and increase labor mobility, linking individuals not only to more productive jobs, but also to job networks, health care facilities, and markets where rural products can be sold. The main drawback of using the simple measure of geographic separation is the implicit assumption that access to all paved roads has the same impact on individuals' labor outcomes. That is, travel time to paved roads ignores the market differences of an alternative destination; it also ignores differences in terms of job availability, wages, financial and commercial institutions, and economic intensity. For instance, two individuals "A" and "B," spending 30 minutes to reach the closest paved road, will be assigned the same degree of geographic isolation, without considering that individual A's closest paved road leads him or her to a tiny community with fewer than 100 inhabitants, while B's closest paved road leads to San Salvador, the largest city in the country.

To overcome this limitation, the "location index" developed by Lardé de Palomo and Argüello de Morera (2000) is used as a measure of job access. The index includes two components. The first captures the number of urban jobs available in the closest urban population of 50,000 inhabitants or more, adjusted by the distance each individual would have to commute to get there. The second component measures the number of jobs available in free trade zones located within a 30-kilometer radius of the household, adjusted by the distance that must be traveled to get to the zone. The location index adds both urban jobs and free trade zone jobs. The location index ranges from 0 to 1. The closer to 1, the higher access to urban jobs or free trade zone jobs.

This does mean that access to markets with 50,000 inhabitants or more is the only appropriate benchmark for measuring lack of access to employment or geographic isolation. El Salvador has three developed urban markets: San Salvador, San Miguel and Santa Ana. It could easily be argued that a measure of geographic isolation could use any or all of these markets as a reference. A similar argument can be made when using other regional or international markets such as Guatemala, Costa Rica, Mexico or even the United States. These considerations, however, lie beyond the scope of this chapter.

Factors Determining Geographic Location

Of particular interest is the link between choice of location and intergenerational income effect (parents' landownership, place of residency, parents' employment status) as proxies of parent's socioeconomic status, remittances, land ownership, size of the community, other household income/resources, and human capital variables. These estimates provide information on the factors behind individuals' geographical decisions and will additionally identify some of the instrumental variables needed to explore the relations between location and labor market outcomes.

People with higher skills are expected to live in less isolated locations. The correlation between income and location choice, though, can be either positive or negative. On the one hand, individuals with higher income may live in less isolated areas due to their taste and capacity to afford the expenses of living there. On the other hand, individuals with higher incomes may wish to remain in isolated communities, where perhaps they have lived most of their lives. It is possible that the travel and opportunity costs of living in isolated communities may be lower for individuals with higher incomes.

Also included in the geographic location equation is a binary variable controlling for individuals living in the same place where parents or relatives lived or were born; it is hypothesized that people may choose to live and remain in the same place where parents or relatives were born or lived most of their lives in order to capitalize on their parents' social capital accumulation and sense of belonging. The significance rather than the sign of these dummy variables is of interest. If significant, the coefficient will show that individuals make location decisions, even in remote areas, in order to capitalize on parents' social capital gains. On the other hand, it is assumed that parents' or relatives' geographic location and the time they have remained in that location are not directly correlated with current individual labor outcomes. It is assumed instead that current individual labor outcomes fully depend on human capital accumulation and on whether the person currently lives in geographic exclusion from society.[8]

[8] Previous studies argue that intergenerational factors such as parents' education do have a strong correlation with offspring human capital accumulation and therefore on their current labor productivity. The lack of this variable in the data on hand made it impossible to test this hypothesis.

Another dummy variable is added to control for the land provided through government programs. Under the terms of the peace agreements following the end of the civil war in 1990, the government gave away land. In these cases, location choices and the degree of isolation were not up to the individual, but determined by the government, making it necessary to control for such cases (less than 5 percent in the sample). For these cases, a positive coefficient will show that the government gave away land located in remote areas and that the people who were given it were unable to sell or rent it and consequently had no choice but to live on it.

A linear regression equation is estimated for each of the two measures of geographic exclusion, first for the whole sample and then by gender.

Geographic Location and Work Participation

The working decision, and therefore the labor supply of men and women, appears to be jointly determined at the household level. The proper way to model family decision-making is not entirely clear, and economists have adopted various approaches (Killingsworth, 1983). One assumes that the marriage partners have a collective set of preferences and behave as a single unit. Another approach assumes that each partner has an individual utility subject to a family budget (Manser and Brown, 1980). Modeling and estimating the joint decision is complex; a multinomial logit model (to estimate each partner's choice probability) and simultaneous equations are some of the techniques being used. For simplicity, however, this study analyzes the probability of labor force participation for men and women separately. A probit model is used to examine the consequences of geographic location for individuals' labor participation decisions.

The probit model is $P(Y = 1 \mid X) = F(\beta X)$ where F(.) is the standard normal cumulative distribution function (cdf). The decision of whether to work in the market depends on the market wage, W, and the shadow price or reservation wage, Z. Market wage is used here if $W > Z$, the person works either in home production or as an employee. Thus, the dependent variable takes on the value of one (1) if the person participates in the labor market and zero (0) if the person does not.

The shadow price or reservation wage (Z) depends on productivity in activities other than labor market work. For both men and women, the reservation wage depends on personal characteristics such as marital status,

age and number of young children, and on the "need" for income, which is measured by remittances, land ownership, family income (excluding the individual's own labor income), and geographical isolation. Whether to work would depend on those factors and the reservation wage. According to the theory, a woman's decision is based on the value of her time at home or her previous earnings versus the offered wage or potential earnings. A man's work decision, however, is based on his value of time if he remains unemployed—waiting for a better match—or his previous earnings versus the offered wage or potential earnings. Despite the gender differences in the factors affecting the labor force participation decision, both men and women can be viewed as setting minimum standards for their decision to work. As such, when the earnings from that job surpass a critical value the job will be accepted; otherwise it will be rejected. This means that the samples of working men and women are self-selected.

In the probit participation functions, age is entered as a series of dummy variables for each age group (in six-year cohorts) to take into account any non-linearity in the effect of age on participation. It is not clear a priori what the signs should be on the coefficients for age variables. Younger individuals may be more likely to be out of the work force and in school, which would decrease the probability of this group's participation. When enrollment in higher education is high, labor force participation of young people will also be lower. However, given the small percentage of individuals with higher education, especially in rural areas of El Salvador, this variable may have little effect.

Heads of household are more likely to work to support the family. Married women are expected to have a lower probability of participation than unmarried women, while the opposite relationship is expected for men. Remittances and family income, originating from an individual's spouse or other relatives living in the household or out of the country, are also included in the probit equation to measure the income effect on the participation decision, and they are expected to have a negative impact on participation. Family remittances may also have a positive impact on participation by increasing individuals' access to more-developed financial and labor markets (rural, urban and international). When available, time living in the same location is also entered to control for migration effects.

According to the spatial mismatch literature, residential location is an important determinant of individuals' labor participation. The impact

of geographic isolation on labor force participation rates can go in either direction. Some empirical studies on spatial mismatch show that living at a distance from jobs increases worker mobility costs and detaches individuals from job networks. Together, both of these factors decrease individual labor force participation rates.

On the other hand, individuals in rural areas, even those living in great geographic isolation, can still work on their own land for commercial or subsistence purposes. In times of crisis, when non-agricultural or salaried employment becomes unavailable, labor force participation rates may nonetheless increase through work in agricultural activities. It is hypothesized that geographic isolation imposes higher opportunity costs and greater hardships on women than men, which may lower women's labor force participation rates further than men's. Security reasons, plus women's cultural role in the society as child-bearer and family caretaker, may be the principal mechanisms whereby geographic isolation limits women's labor force participation more than men's.

In addition, El Salvador is one of the most violent countries in Latin America; therefore the risk of walking alone on dark, muddy paths (rather than paved roads), and of being physically attacked are additional transaction costs faced more by rural women than by men. Such security risks, as well as time and transportation costs, impose higher working opportunity costs on women than on men. Moreover, geographic isolation may also increase the value of women's household production activities by increasing their value as protector of children and caregiver, and further reducing the probability of their participation in the labor market.

Geographic isolation, through a combination of security hazards and women's cultural role in society, is thus an important factor in women's exclusion from the rural labor market. On the other hand, higher working costs and the minimization of risk hazard costs may lead women living in isolated communities to increase their participation decision by working in agricultural activities at home or close to home. Geographic isolation may thus increase rather than decrease labor force participation rates.

Geographic Isolation and Sector Allocation

A multinomial logit model is used to test the impact of geographic isolation on the sectors of employment. Rural Salvadoran workers are divided into

those in salaried sectors, both agricultural and non-agricultural, and those in home production, both agricultural and non-agricultural.[9] To classify all rural workers into mutually exclusive sectors, it was also necessary to include a fifth sector, the "mixed" sector. This sector includes all individuals who worked in two or more sectors during the previous year. The underlying assumption is that rural Salvadoran labor markets are fully described by these five sectors.

Specification of the Sector Allocation Model

This model first assumes that an individual chooses to work in a specific sector, and then his or her labor income is observed. Therefore, the sector assignment estimation cannot include any variables determined by sector membership such as experience in the labor market, payments in kind, and whether a job is temporary or permanent. The sector assignment estimation does, however, need to include some variables that are related to productivity and could influence employers' preferences for workers. The following equations are estimated for the whole sample, and then for males and females:

$$\log (p_1/p_2) = \alpha_{12} + \beta_{12}\,GEO + \gamma_{12}\,ED1 + \delta_{12}\,ED2 + \zeta_{12}\,ED3 + \eta_{12}\,ED4 + \omega_{12}\,REG2 + \varsigma_{12}\,REG3 + \tau_{12}\,REG4 + \theta_{12}\,AGE16\text{–}20 + \Theta_{12}\,AGE21\text{–}25 + \phi_{12}\,AGE26\text{–}30 + \psi_{12}\,AGE31\text{–}35 + \chi_{12}\,AGE36\text{–}40 + \varepsilon_{12}\,AGE41\text{–}45 + \nu_{12}\,MAR + e_{12}$$

$$\log (p_3/p_2) = \alpha_{32} + \beta_{32}\,GEO + \gamma_{32}\,ED1 + \delta_{32}\,ED2 + \zeta_{32}\,ED3 + \eta_{32}\,ED4 + \omega_{32}REG2 + \varsigma_{32}\,REG3 + \tau_{32}\,REG4 + \theta_{32}\,AGE16\text{–}20 + \Theta_{32}\,AGE21\text{–}25 + \phi_{32}\,AGE26\text{–}30 + \psi_{32}\,AGE31\text{–}35 + \chi_{32}\,AGE36\text{–}40 + \varepsilon_{32}\,AGE41\text{–}45 + \nu_{32}\,MAR + e_{32}$$

$$\log (p_4/p_2) = \alpha_{42} + \beta_{42}\,GEO + \gamma_{42}\,ED1 + \delta_{42}\,ED2 + \zeta_{42}\,ED3 + \eta_{42}\,ED4 + \omega_{42}\,REG2 + \varsigma_{42}\,REG3 + \tau_{42}\,REG4 + \theta_{42}\,AGE16\text{–}20 + \Theta_{42}\,AGE21\text{–}25 + \phi_{42}\,AGE26\text{–}30 + \psi_{42}\,AGE31\text{–}35 + \chi_{42}\,AGE36\text{–}40 + \varepsilon_{42}\,AGE41\text{–}45 + \nu_{42}\,MAR + e_{42}$$

[9] Ideally, the salaried sector should be divided into private and public sectors. Previous studies (Briones and Andrade-Eekhoff, 2000) have shown the importance of separating them because the public sector has different skill requirements and wage-setting mechanisms than the private sector, especially for women. However, the structure of the Rural Household Survey did not separate these two sectors.

$$\log (p_5/p_2) = \alpha_{52} + \beta_{52}\,GEO + \gamma_{52}\,ED1 + \delta_{52}\,ED2 + \zeta_{52}\,ED3 + \eta_{52}\,ED4 + \omega_{52}\,REG2 + \varsigma_{52}\,REG3 + \tau_{52}\,REG4 + \theta_{52}\,AGE16\text{–}20 + \Theta_{52}\,AGE21\text{–}25 + \phi_{52}\,AGE26\text{–}30 + \psi_{52}\,AGE31\text{–}35 + \chi_{52}\,AGE36\text{–}40 + \varepsilon_{52}\,AGE41\text{–}45 + \nu_{52}\,MAR + e_{52}$$

where:

1	=	Agriculture salaried sector
2	=	Agriculture home production
3	=	Non-agriculture salaried sector
4	=	Non-agriculture home production
5	=	Mixed sector
p_i	=	Probability of i^{th} individual of working in the k^{th} sector
GEO	=	Predicted geographic isolation measure
ED1	=	1 if one to three years of schooling, 0 otherwise
ED2	=	1 if four to six years of schooling, 0 otherwise
ED3	=	1 if seven to nine years of schooling, 0 otherwise
ED4	=	1 if more than 10 years of schooling, 0 otherwise
REG2	=	1 if western region, 0 otherwise
REG3	=	1 if central region, 0 otherwise
REG4	=	1 if eastern region, 0 otherwise
AGE1620	=	1 if ages 16 to 20
AGE2125	=	1 if ages 21 to 25
AGE2630	=	1 if ages 26 to 30
AGE3135	=	1 if ages 31 to 35
AGE3640	=	1 if ages 36 to 40
AGE4145	=	1 if ages 41 to 45
MAR	=	1 if married (or living together), 0 otherwise.

For each year, four sector assignment equations are estimated, one set of four equations for each geographic exclusion measure: time and index location. The reference groups include individuals with zero years of formal education, living in the metropolitan area, and between 46 and 65 years old. In the estimation of sector assignment, the agriculture home production sector is used as the reference sector; that is, all coefficients on the agriculture home production sector are normalized to zero. The outcome of the sector assignment estimation indicates the worker's propensity to be in the non-agriculture home production sector or located in the salaried agricul-

ture or non-agriculture sectors rather than in the agriculture home production sector.

Educational attainment is measured using four dichotomous variables. It is assumed that the more education a worker has, the more likely he or she is to find work in the non-agricultural salaried sector; similarly, workers with less education are more likely to be working in the agriculture sector, especially in the salaried agriculture sector. Therefore, with regard to the propensity of being in the non-agriculture salaried sector rather than in the agriculture home production sector, the coefficients with respect to education of workers with high education (EDU3 and EDU4) should be positive: $\zeta_{32} > 0$ and $\eta_{32} > 0$. In contrast, the higher the worker's education, the lower the probability of being in the salaried agriculture salaried sector. Hence, the coefficients with respect to high education should be negative: $\eta_{12} < 0$, $\zeta_{12} < 0$.

In order to control for the cohort effect, age is measured using six dichotomous variables. It is difficult to predict the sign on these coefficients. Age serves as a proxy for the date of entry into the labor market and may also represent workers' labor market experience. Higher market competition faced by salaried enterprises increases their demand for younger workers, who may be considered much easier to train, and makes those firms more willing to sacrifice an experienced labor force. Therefore, new entrants into the labor market may have better job opportunities in the salaried sector than elsewhere. From this partial analysis, it is predicted that $\theta_{32} > 0$ and $\theta_{12} > 0$.

At the same time, the effect of greater labor market experience can make some of these age coefficients negative. More experience may lead to more know-how, more information about the functioning of the financial markets, more contacts, and other advantages. Therefore, it can be predicted that some of the coefficients, especially for prime-age workers or older workers, may be negative, showing a higher propensity for working in the home production sector than elsewhere.

In general, married people look for job stability and may have a stronger preference for jobs that offer better working conditions. Firms in the salaried non-agriculture sector are more likely to offer job stability and security, and more likely to comply with legal requirements regarding benefits. If these firms also have greater tastes for married workers, who may be seen as more stable and responsible than other workers, the coefficients on

these dummy variables are expected to be positive: $\nu_{32} > 0$ and $\nu_{12} > 0$. However, these coefficients may turn out to be exactly the opposite for women. If firms in the salaried sector also have a greater preference for single women who do not have children to take care of but have more time for their jobs, or if married women may prefer jobs with more flexible time schedules, the married coefficients may then be negative: $\nu_{32} < 0$ and $\nu_{12} < 0$.

To control for differences in sector allocation across regions, three dummy variables are included. The omitted region is the metropolitan area, which includes San Salvador. As in many developing countries, economic activity is highly concentrated in urban areas, especially in the capital. Assuming that the home production sector is more common in places where employment opportunities are scarce in the salaried sectors and where economic activity is depressed, workers living in the metropolitan area (the capital and immediate surroundings) may have a greater opportunity to obtain employment in the salaried non-agriculture sector than elsewhere. Also, the effect of the civil war, mostly concentrated in the eastern part of the country, reinforces the fact that enterprises may be located far from the east and closer to the capital and immediate areas.

In this study, two measures of geographic exclusion are used. The predicted value of geographic isolation is again included in order to measure the impact of residential location on sector allocation. Assuming that non-agriculture salaried jobs are located in more urbanized areas, it was hypothesized that workers living closer to urbanized jobs, i.e., living in less isolation, will have more access to and work in the non-agriculture salaried sector. Workers living in more isolated places are expected to work in home production activities, especially in the agriculture sector.

Geographic Isolation and Labor Income

The chosen earnings equation is a variant of Mincer's (1974) standard human capital equation. The exact functional form is (ignoring the male, female and sector-specific subscripts):

$$\ln W = \beta_0 + \beta_1 EDU + \beta_2 EXP + \beta_3 EXP2 + \beta_4 GEO + \xi_1$$

where

$\ln W$ = the natural logarithm of hourly wage

EDU = years of schooling
EXP = years of potential labor market experience
GEO = geographic isolation measures
ξ_1 = error term

Education increases individuals' labor productivity and labor income; the coefficient on education is thus expected to have a positive sign. Similarly, based on the concavity of the experience/earnings profile, it is expected that the estimated coefficient on experience (β_2) will have a positive sign and the quadratic term (β_3) a negative sign. Age is included as a proxy of experience.

To measure the impact of residential location on labor income, the predicted value of geographic exclusion is included in the earning equation. Living in isolation from the rest of society may impact labor productivity via increasing transactions costs (increasing the human effort to reach places), limiting access to markets for selling their products or buying raw material with better prices and selection, limiting access to key information on new technology, etc. If more geographically isolated individuals are those who work in agricultural activities, especially as self-employed workers, and the labor productivity of these workers is lower than others working in salaried activities or who work in more than two sectors during the year, then the workers living farther from urban markets would be expected to have lower labor income.

Results

Geographic Exclusion

The two aims of estimating the geographic isolation regression were to identify the factors behind individuals' geographical decision and to define the instrumental variables that would later be needed to explore the relation between location and labor market outcomes. These instrumental variables should be uncorrelated with current labor market outcomes but correlated with choice of location. Table 5.6 shows the results of the geographic isolation regressions where isolation is measured by the location index (from 0 to 1; closer to 1 shows greater access to urban jobs and industrial parks) and

Table 5.6. Geographic Isolation Regression Results by Gender

	Time to paved road (minutes)			Location index (0–1)		
	Whole sample	Women	Men	Whole sample	Women	Men
Variable	**Parameter estimate**	**Parameter estimate**	**Parameter estimate**	**Parameter estimate**	**Parameter estimate**	**Parameter estimate**
Intercept	20.3773***	26.0016***	14.5508**	0.1711***	0.1796***	0.1685***
	(5.39)	(7.98)	(7.12)	(0.01)	(0.02)	(0.01)
Income	0.0002	0.0000	0.0003	0.0000	0.0000	–0.0000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Remittances	–0.2	–0.1	–0.3 *	–0.000*	–0.0000	–0.0000
(thousands of colons)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Land ownership	–3.5406	–3.9068	–2.5801	–0.0081*	–0.0111	–0.0065
	(2.15)	(3.08)	(3.00)	(0.00)	(0.01)	(0.01)
Schooling	–1.2843***	–1.4256***	–1.2116 ***	0.0009*	0.0016**	0.0003
	(0.24)	(0.35)	(0.33)	(0.00)	(0.00)	(0.00)
Age	–0.1939**	–0.3067**	–0.1031	0.0001	0.0003	–0.0000
	(0.06)	(0.10)	(0.08)	(0.00)	(0.00)	(0.00)
No. of children	1.3793***	0.9544	1.7198**	–0.0007	–0.0002	–0.0009
	(0.42)	(0.60)	(0.59)	(0.00)	(0.00)	(0.00)
Inherited land	3.5257	4.6173	2.0199	0.0005	–0.0004	0.0025
	(2.15)	(3.04)	(3.03)	(0.00)	(0.01)	(0.01)
Land given by	1.4291	–1.2105	3.8021	–0.0006	0.0033	–0.0041
government	(2.76)	(3.99)	(3.81)	(0.01)	(0.01)	(0.01)
Families same	–0.0035***	–0.0036***	–0.0036***	0.0000	0.0000	0.0000
community	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
No. of organizations	5.5127***	3.3626**	7.3456***	–0.0056**	–0.0031	–0.0076**
	(1.10)	(1.63)	(1.49)	(0.00)	(0.00)	(0.00)
Live same place	7.6027***	6.2376**	9.1958**	–0.0038	–0.0053	–0.0025
as parents	(2.21)	(3.17)	(3.07)	(0.00)	(0.01)	(0.01)
Parents landowners	1.2444	1.6809	0.9410	0.0046	0.0090	–0.0001
	(1.81)	(2.57)	(2.53)	(0.00)	(0.01)	(0.00)
West	4.3249	5.4028	3.0025	–0.0938***	–0.1072***	–0.0829***
	(4.11)	(6.12)	(5.53)	(0.01)	(0.01)	(0.01)
Central	13.4485***	15.9061**	10.8815**	–0.1256***	–0.1461***	–0.1082***
	(4.02)	(6.02)	(5.41)	(0.01)	(0.01)	(0.01)
East	0.9072	1.9507	–0.2144	–0.0427***	–0.0484***	–0.0383***
	(1.38)	(2.06)	(1.87)	(0.00)	(0.00)	(0.00)
Time living in	0.1991***	0.1588**	0.2367***	–0.0002**	–0.0001	–0.0003**
same location	(0.04)	(0.06)	(0.06)	(0.00)	(0.00)	(0.00)
Gender (Women = 1)	–1.3551			0.0046		
	(1.60)			(0.00)		
Log likelihood	–8902.12	–4666.85	–4214.45	1235.53	441.61	821.19
N	1,661	841	820	1,672	847	825

Notes: Standard errors in parentheses. Missing values in variables: schooling, families in same community, time living in same location. The following levels of significance apply: * p<0.10, ** p<0.05, *** p<0.01, two-tailed test.

Source: FUSADES (1999).

by time (in minutes) required to reach the closest paved road. The adjusted r-squared of all regressions ranged from 6 to 16 percent; that is, the control variables explain very little of individuals' decisions on where to live. The results in Table 5.6 lead to the following conclusions:

1. People with higher levels of education are located in less isolated areas. This is supported by the positive schooling coefficient in the location index regression and the negative schooling coefficient in the time to closest paved roads regression. The former suggests that individuals with more skills live closer to urban markets where access to urban employment and industrial parks is higher. The latter suggests that less education correlates with longer travel times to paved roads. Using age as a proxy for labor market experience, it can be concluded that more experienced (or older) individuals live closer to paved roads.
2. Men receiving remittances from relatives outside the household may have better access to urban jobs, and people receiving or reporting remittances may be those who live in less isolated places. The effect of remittances on geographic isolation, however, is small; when significant, an increase of 10,000 colons in remittances (equivalent to $1,142) reduces the time to the closest paved road by three minutes. In the sample, only 24 percent of men living in isolated areas received some remittances, and only 10 percent received 10,000 or more colons per year.
3. Variables for land, either inherited or received through a government program, were not significantly correlated with measures of residential location.
4. The variable "living in same place where parents or relatives lived or were born" is positively correlated with geographic location. This result may indicate that individuals in fact decide to live and stay even in isolated areas to capitalize on parents' social capital gains.
5. The number of families living in the same community correlates with geographic separation. This result was expected and indicates that larger communities with more families have better access to paved roads, and therefore residents are less isolated from the rest of the society. Individuals living in isolation reside in small

communities where only a few families live close by, which in fact confirms their geographic isolation.

6. The number of group memberships (number of organizations in which the household participates) correlates positively with geographic isolation. Networking and affiliation with different groups seems to be important when living in isolated communities; therefore, membership in social, religious, sports, political and other types of organizations may be an important form of social capital for individuals living spatially excluded from the rest of society.
7. As expected, in all regions but the metropolitan area (region of reference), individuals live in greater isolation. People from the central region take longer to get to paved roads than those living in the other regions.
8. Geographic isolation is not gender related. The coefficient controlling for gender differential was not significant; hence, women are not more isolated than men.

From these results, the variables that correlate with residential location (measured through time to the closest paved road) that are uncorrelated with labor market outcomes are "living in same place as parents" and "families in same community." The correlations between these two variables and labor outcomes such as "working status" and "labor market income" are 0.057 and –0.013, respectively. These two variables were regressed with the residuals of labor income and the coefficient became not significant ($p>0.05$). Two variables were then used as exclusion restrictions, explaining the location decision and not included in the labor outcome equations. Other variables, such as schooling, remittances, number of organizations and experience correlate with both individuals' residential location and their labor market participation decision and earnings level.

Labor Market Participation

Table 5.7 reports the results of the probit regression for the work participation functions. Moving from left to right, the first three columns show the probit results when using the predicted time to the closest paved road value as a measure of geographic exclusion for the whole sample, for men, and for

Table 5.7 Labor Force Participation Results

Dependent variable: working/non-working

	Time to paved road (minutes)			Location index (0–1)		
	Whole sample	Women	Men	Whole sample	Women	Men
Variable	**Parameter estimate**	**Parameter estimate**	**Parameter estimate**	**Parameter estimate**	**Parameter estimate**	**Parameter estimate**
Age: 16–20	–0.1153	0.5053*	–0.3067*	–0.1279	0.4959*	–0.3222*
	(0.14)	(0.30)	(0.17)	(0.14)	(0.30)	(0.17)
Age: 21–25	0.2787*	1.0880 ***	0.0803	0.2697*	1.0857***	0.0723
	(0.15)	(0.33)	(0.17)	(0.14)	(0.33)	(0.17)
Age: 26–30	0.4081**	0.8913**	0.3576*	0.4139**	0.8746**	0.3639*
	(0.16)	(0.35)	(0.19)	(0.16)	(0.35)	(0.19)
Age: 31–35	0.2442	0.9712**	0.1426	0.2442	1.0000**	0.1369
	(0.17)	(0.39)	(0.20)	(0.17)	(0.39)	(0.20)
Age: 36–40	0.2431	0.4022	0.3024	0.2542	0.5035	0.3254*
	(0.16)	(0.35)	(0.19)	(0.16)	(0.36)	(0.19)
Age: 41–45	0.2959*	0.4042	0.3151*	0.2717*	0.3149	0.2982*
	(0.15)	(0.38)	(0.17)	(0.15)	(0.37)	(0.17)
No. of family members	–0.0051	–0.0382	0.0049	–0.0070	–0.0450	0.0039
	(0.02)	(0.04)	(0.02)	(0.02)	(0.04)	(0.02)
Schooling	–0.0174	–0.0425**	–0.0050	–0.0143	–0.0387**	–0.0018
	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
No. of children	0.0284	0.1348**	–0.0091	0.0340	0.1465**	–0.0007
	(0.03)	(0.06)	(0.04)	(0.03)	(0.06)	(0.04)
Married	–0.0706	0.2772	–0.2066*	–0.0554	0.2856	–0.1845*
	(0.10)	(0.23)	(0.11)	(0.10)	(0.23)	(0.11)
Landowner	–0.1312	–0.0662	–0.1551	–0.1276	–0.1819	–0.0925
	(0.09)	(0.17)	(0.10)	(0.09)	(0.17)	(0.10)
Remittances	–0.0000**	–0.0000*	–0.0000	–0.0000**	–0.0000**	–0.0000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Head of household	0.6670***	0.9967***	0.4226**	0.6708 ***	1.0032***	0.4447**
	(0.13)	(0.26)	(0.19)	(0.13)	(0.27)	(0.19)
Family living abroad	–0.0204	–0.0132	–0.0238	–0.0201	–0.0296	–0.0158
	(0.03)	(0.06)	(0.04)	(0.03)	(0.06)	(0.04)
Health problems last 12 months	0.1344*	0.2025	0.1055	0.1414*	0.1981	0.1121
	(0.08)	(0.14)	(0.10)	(0.08)	(0.14)	(0.10)
Predicted geographic exclusion	0.0003	0.0042	–0.0031	–0.7637	–5.1288**	2.0499
	(0.00)	(0.01)	(0.00)	(1.23)	(2.20)	(1.45)
Gender (Women = 1)	–0.9327***			–0.9459***		
	(0.09)			(0.09)		
Intercept	1.1601***	0.5029	0.4939**	1.1939***	1.0146**	0.1880
	(0.22)	(0.41)	(0.25)	(0.21)	(0.40)	(0.23)
Log likelihood	8902.1243	4214.4501	4666.8512	1235.5334	821.1889	441.6140
N	1,661	820	841	1,672	825	847

Notes: Corrected standard errors in parentheses. The following levels of significance apply: * p<0.10, ** p<0.05, *** p<0.01.

Source: FUSADES (1999).

women, respectively.[10] The last three columns show the probit results when using the predicted location index as a measure of geographic exclusion for the whole sample, for men, and for women, respectively. The results indicate the following:

1. The age profile of participation has an inverted U-shape for both men and women. Men younger than 36 years old are likely to be working. Women's participation decision is more age sensitive than that of men.
2. Married women are less likely to work. For men, marital status did not impact their decision whether to work.
3. Being a household head, regardless of gender, is positively correlated with the probability of participation.
4. Remittance amount decreases men's likelihood of working but has no impact on women's working decision.
5. A greater number of children living in a household increases men's likelihood of working.
6. Higher schooling levels decrease men's probability of participation. This would suggest that the reservation wages for men with higher education are higher than the actual wage offered in the market, and therefore these men tend to participate less. Education is not a determinant of women's labor participation decisions.
7. When significant, geographic isolation impacts only men's working decision. When only the men's sample is used, the probit results suggest that living apart from urban centers (lower location index) increases men's likelihood of working. On the other hand, geographic isolation seems to depress women's labor participation, but both measures (predicted time to a paved road and predicted location index) are not statistically significant. However, the effect of geographic location on men's working decision is small. For instance, a man living less than two minutes away from the closest paved road has a working probability of 89 percent, while a man living 50 minutes away has a working probability that is 4 percent higher (93 percent).

[10] The authors would like to thank Alejandro Gaviria for facilitating the procedure designed by Deon Filmer that estimated the correct standard errors when performing an instrumental estimation using probit and continuous equations.

Sector Allocation

Tables 5.8 and 5.9 provide the maximum likelihood logit estimates for the sector allocation decisions of male and female workers. In the two estimations, the predicted value of time to the closest paved road was used to control for individuals' geographic isolation.

Coefficients for the agriculture home production sector are normalized to zero. A likelihood ratio test of each logit shows that the coefficients, taken as a group, are significantly different from zero at the one percent level of significance; this indicates that male and female Salvadoran workers are assigned non-randomly to these five sectors: agriculture and non-agriculture, salaried and home production, and the mixed sector. The pseudo R-squared was 18 and 17 percent in men and women's multinomial logit, respectively.

Tables 5.10 and 5.11 report the expected sector probability for a worker at different levels of education, age, region, and degree of geographic isolation. Generally, higher education increases the probability of obtaining a job in the salaried sector, especially in the non-agriculture sector, regardless of gender. The likelihood of working in the salaried non-agriculture sector for a woman with some high school education is 40 percent, while the probability for a similar worker with no schooling is only 10 percent.[11] Lower or no education in women increases their likelihood of being allocated into the home production agriculture sector. Women with less than a third grade education have more than a 40 percent chance of working in home agriculture production (most likely for personal consumption); men with less than a third grade education have a greater probability of working in the mixed sector. Regardless of gender, workers under 30 years of age are more likely to work in the salaried sector in both agriculture and non-agriculture. Older workers are more likely to work in home production activities. One possible explanation is that, after a period of employment in the salaried sector, experienced workers decide to become self-employed.[12]

[11] The results of the age effect estimated within each educational level were basically the same: individuals with more education regardless of age were more likely to work in the non-agriculture salaried sector than elsewhere.

[12] The behavior of these workers may be due to inappropriate pension programs. With a pension that is smaller than the wage level, older people may decide to enter the home production labor market in an attempt to make up for lost income. The home production sector becomes a source of extra income that does not require workers to report their activity and thereby jeopardize their legal status as retirees.

Table 5.8. Maximum Likelihood Logit Estimates of Sector Allocation for Male Workers

	Agriculture salaried	Non-agriculture salaried	Own production non-agriculture	Mixed
Schooling: 1–3 years	0.6488	–0.1293	1.1236	0.6966
	(0.63)	(0.66)	(0.87)	(0.43)
Schooling: 4–6 years	1.5175	–0.4470	1.8957	1.1190
	(1.13)	(0.99)	(1.59)	(0.76)
Schooling: 7–9 years	2.5943	–0.9805	2.4398	1.5154
	(1.79)	(1.50)	(2.48)	(1.21)
Schooling: 10+ years	2.4373	–2.6949	2.0424	1.1438
	(2.48)	(2.06)	(3.52)	(1.64)
West	1.1583	–0.2205	1.8189	0.7186
	(0.84)	(0.67)	(1.31)	(0.62)
Central	–0.6087	–0.7453	0.8940	–0.0553
	(0.85)	(0.64)	(1.30)	(0.60)
East	–0.0936	–0.3976*	–0.1360	0.1144
	(0.28)	(0.23)	(0.48)	(0.20)
Age: 16–20	–0.7040	–2.4548***	3.1521	–0.5043
	(0.75)	(0.63)	(2.06)	(0.57)
Age: 21–25	–0.7558	–1.4711**	3.3096*	–0.1334
	(0.72)	(0.58)	(1.84)	(0.53)
Age: 26–30				
Age: 31–35	1.8468**	0.8335	4.1467**	1.3949**
	(0.73)	(0.68)	(1.46)	(0.63)
Age: 36–40	–0.1231	–0.4462	1.7847	0.6342
	(0.72)	(0.62)	(1.31)	(0.47)
Age: 41–45	1.1336*	0.3352	2.5810**	1.1541**
	(0.64)	(0.66)	(1.02)	(0.49)
Married	–0.9850**	0.6325*	–0.6108	–0.2343
	(0.41)	(0.38)	(0.51)	(0.28)
Predicted time	0.0014	–0.0140**	–0.0121	–0.0001
to paved road	(0.00)	(0.00)	(0.01)	(0.00)
Constant	2.3993*	5.1430***	–7.4468**	1.7834*
	(1.44)	(1.25)	(3.31)	(1.05)
–2 Log likelihood ratio	1666.7137			
N	744			

Notes: This table uses "predicted time to paved road" as a measure of geographic exclusion. All coefficients on the own production agriculture sector have been normalized to zero. Figures in parentheses are asymptotic standard errors. Workers ages 26–30 were excluded because of missing values in some sectors.
Source: FUSADES (1999).

Table 5.9. Maximum Likelihood Logit Estimates of Sector Allocation for Female Workers

	Agriculture salaried	Non-agriculture salaried	Own production non-agriculture	Mixed
Schooling: 1–3 years	–0.2843	–0.6430	–0.6668	0.3580
	(0.75)	(0.64)	(0.58)	(0.53)
Schooling: 4–6 years	–0.0283	–0.4568	–1.1689	1.3538
	(1.26)	(1.03)	(1.05)	(0.99)
Schooling: 7–9 years	1.3515	–0.9421	–1.0062	2.6146 *
	(1.99)	(1.58)	(1.63)	(1.58)
Schooling: 10+ years	0.9433	–1.5553	–1.9052	3.7562
	(2.84)	(2.27)	(2.36)	(2.28)
West	–1.5359 *	–1.7145 **	–0.9994	0.1425
	(0.81)	(0.80)	(0.80)	(0.97)
Central	–1.9410 **	–1.0815	–0.8188	0.1950
	(0.83)	(0.79)	(0.79)	(0.98)
East	–1.2052 ***	–0.8357 **	–0.6349 **	–0.1932
	(0.30)	(0.27)	(0.27)	(0.33)
Age: 16–20	6.7254 **	3.9698 *	1.5029	–0.3089
	(3.39)	(2.26)	(1.69)	(1.63)
Age: 21–25	6.3394 **	4.2103 **	1.3552	0.8727
	(2.99)	(1.96)	(1.47)	(1.39)
Age: 26–30	5.9228 **	3.5440 **	0.4802	0.3875
	(2.55)	(1.68)	(1.30)	(1.18)
Age: 31–35	3.5804	2.2579	1.1732	1.1120
	(2.24)	(1.46)	(1.08)	(0.99)
Age: 36–40	3.9154 **	2.1725 *	0.6170	0.7886
	(1.81)	(1.18)	(0.89)	(0.81)
Age: 41–45	3.3966 **	1.3856	0.6565	0.7126
	(1.40)	(0.94)	(0.68)	(0.60)
Married	–0.5110	–1.1684 **	0.0130	–0.3677
	(0.44)	(0.41)	(0.32)	(0.29)
Predicted time	–0.0027	–0.0155 **	–0.0080 *	–0.0012
to paved road	(0.00)	(0.00)	(0.00)	(0.00)
Constant	–6.0066	–3.7382	–0.8718	–1.2828
	(4.96)	(3.39)	(2.59)	(2.55)
–2 Log likelihood ratio	1341.8753			
N	523			

Notes: This table uses "predicted time to paved road" as a measure of geographic isolation. All coefficients on own production agriculture sector have been normalized to zero. Figures in parentheses are asymptotic standard errors.

Source: FUSADES (1999).

Table 5.10. Probabilities of Working in Each Sector by Education, Age, Marital Status and Residential Location for Women

	Agriculture salaried	Non-agriculture salaried	Own production agriculture	Own production non-agriculture	Mixed
Education level					
No school	0.097	0.097	0.407	0.166	0.234
1–3 years	0.098	0.128	0.414	0.158	0.203
4–6 years	0.095	0.254	0.317	0.143	0.190
7–9 years	0.203	0.230	0.203	0.216	0.149
>Middle school	0.089	0.400	0.156	0.200	0.156
Age groups					
16–20	0.167	0.267	0.311	0.178	0.078
21–25	0.143	0.341	0.209	0.143	0.165
26–30	0.203	0.234	0.328	0.078	0.156
31–35	0.048	0.119	0.310	0.214	0.310
36–40	0.094	0.151	0.358	0.151	0.245
41–45	0.111	0.095	0.333	0.175	0.286
46–65	0.025	0.075	0.458	0.217	0.225
Married	0.077	0.071	0.423	0.199	0.231
Time to paved road					
<2 minutes	0.113	0.326	0.222	0.176	0.163
<10 minutes	0.104	0.291	0.247	0.205	0.154
10–15 minutes	0.118	0.180	0.314	0.190	0.197
15–30 minutes	0.108	0.190	0.344	0.161	0.198
30–50 minutes	0.104	0.126	0.374	0.159	0.238
>50 minutes	0.119	0.129	0.418	0.132	0.202
Location index					
Less isolated	0.137	0.214	0.271	0.195	0.183
More isolated	0.083	0.159	0.409	0.143	0.207

Notes: Expected probabilities are based on estimates coefficients reported in Table 5.9, using the formula: $P_{ij} = \exp(X_i B_j)/\text{Sum}_{k=1,\ldots,j} \exp(X_i B_k)$. More isolated = individuals whose location index is less than 0.024, N = 1,000 when using whole sample. Less isolated = individuals whose location index is more or equal to 0.024, N = 977 when using whole sample.
Source: FUSADES (1999).

Table 5.11. Probabilities of Working in Each Sector by Education, Age, Marital Status, and Residential Location for Men

	Agriculture salaried	Non-agriculture salaried	Own production agriculture	Own production non-agriculture	Mixed
Education level					
No school	0.154	0.053	0.237	0.041	0.515
1–3 years	0.119	0.094	0.188	0.056	0.544
4–6 years	0.108	0.185	0.180	0.050	0.477
7–9 years	0.121	0.293	0.155	0.034	0.397
>Middle school	0.065	0.377	0.273	0.013	0.273
Age groups					
16–20	0.129	0.147	0.282	0.024	0.418
21–25	0.092	0.252	0.210	0.034	0.412
31–35	0.197	0.225	0.056	0.056	0.465
36–40	0.069	0.138	0.155	0.034	0.603
41–45	0.143	0.095	0.111	0.063	0.587
46–65	0.101	0.083	0.238	0.069	0.509
Married	0.075	0.186	0.199	0.049	0.491
Time to paved road					
<2 minutes	0.134	0.225	0.149	0.078	0.414
<10 minutes	0.119	0.252	0.156	0.059	0.413
10–15 minutes	0.119	0.200	0.185	0.057	0.440
15–30 minutes	0.106	0.178	0.215	0.038	0.464
30–50 minutes	0.120	0.152	0.202	0.037	0.489
>50 minutes	0.121	0.092	0.244	0.029	0.515
Location index					
Less isolated	0.137	0.219	0.154	0.052	0.438
More isolated	0.099	0.128	0.249	0.034	0.490

Notes: Expected probabilities are based on estimates coefficients reported in Table 5.8, using the formula: $P_{ij} = \exp(X_i B_j)/\text{Sum}_{k=1,\ldots,j} \exp(X_i B_k)$. More isolated = individuals whose location index is less than 0.024, N = 1,000 when using the whole sample. Less isolated = individuals whose location index is more or equal to 0.024, N = 977 when using whole sample.
Source: FUSADES (1999).

Turning to geographic location, workers living in less isolation have greater chance of getting jobs in the salaried sector, especially the non-agriculture sector. Residential location determines women's sector allocation more than men's; women living separated from economic activity and in greater isolation are concentrated in home production agriculture activities. The likelihood of working in the home production agriculture sector for a woman living 50 minutes or more away from the closest paved road is 42 percent, while the probability for a man living at a similar distance is 24 percent.

The coefficient on marital status in the case of women indicates that married women are concentrated in the agriculture home production sector and have either a lower preference for or less access to jobs in the salaried sector.[13] Married men are concentrated in the mixed sector. The multinomial logits were also estimated using the predicted location index as a measure of geographic exclusion; the results were very much the same as above.[14]

Earnings Regression Results

Tables 5.12 and 5.13 show the labor income equations for all five sectors for males and females, respectively. In this set of labor income equations, the predicted location index is added as a measure of geographic exclusion. When significant, the coefficients have the expected signs. Particularly striking are the small adjusted R-squares reported by the regressions. They range from 0 to 25 percent, showing how little the usual human capital variables explain labor income for the rural sample. There must be other variables, besides skills, which perform much better at explaining the marginal product of labor. The results of the estimations may then have omitted variable biases and should be treated with caution.

For women, when using all sectors, living geographically excluded decreases labor income. When the estimated coefficients from the whole sample equation are used, a woman with three years of schooling, 20 years old, and

[13] The category "married" also includes men and women who are *acompañado/a* (i.e., in a relationship comparable to marriage, such as cohabitation).

[14] Tables 5.10 and 5.11 show the multinomial logit results using the predicted value of time to the closest paved road. The multinomial logits using the predicted location index were also estimated (but not shown), and the results were very similar.

Table 5.12. Income Labor Function for Male Workers by Sector

		Salaried		Own production		
	Whole sample	Agriculture	Non-agriculture	Agriculture	Non-agriculture	Mixed sector
Dependent variable: Ln labor income per hour						
Schooling	0.0694***	0.0005	0.0576***	0.0773*	0.0762	0.0441***
	(0.0109)	(0.0124)	(0.0129)	(0.0424)	(0.0668)	(0.0122)
Experience	0.0727***	–0.0078	0.1047***	0.0911	0.0256	0.0364**
	(0.0170)	(0.0158)	(0.0266)	(0.0640)	(0.0717)	(0.0175)
Experience squared	–0.0009***	0.0001	–0.0012***	–0.0011	–0.0003	–0.0004
	(0.0002)	(0.0002)	(0.0004)	(0.0008)	(0.0009)	(0.0002)
Predicted location index	–0.7686	–0.3176	–0.6136	–6.9794	5.5923**	–0.6888
	(1.51)	(0.8377)	(0.8398)	(7.0285)	(1.7259)	(1.3732)
Constant	–0.2475	1.7190***	–0.1359	–0.9822	0.7244	0.2255
	(0.3032)	(0.2786)	(0.4479)	(1.1366)	(1.1441)	(0.3145)
R^2	0.0717	0.0085	0.2343	0.0308	0.2451	0.0537
F-statistic	14.59***	0.11	10.53***	1.31	4.52**	5.72***
N	690	86	127	113	32	332

Notes: This table uses "predicted location index" as a measure of geographic isolation. Figures in parentheses are corrected standard errors. The following levels of significance apply: * p<0.1, **p<0.05, *** p<0.01.
Source: FUSADES (1999).

living far away (location index equal to 0.2) earns 3.78 colons per hour; a women with similar characteristics but living closer to urban markets (location index of 0.6) earns 10.37 colons per hour. Schooling correlates positively with labor income, especially for those working in the salaried non-agriculture sector. Women's schooling level or labor market experience does not seem to be correlated with labor income when working in any sector other than the salaried non-agriculture sector.

For male workers, education is a determinant of labor income when allocated in the salaried non-agriculture sector and when working in home production agricultural activities. Assuming education increases workers' labor market productivity, this result implies that the agriculture home production sector rewards men and women in different ways, increasing men's income while not being correlated with women's. A possible explanation may be that women are mainly producing for personal consumption, while men produce for commercial purposes. Future research on labor rural income should control for these two types of rural production.

Table 5.13. Income Labor Function for Female Workers by Sector

		Salaried		Own production		
	Whole sample	Agriculture	Non-agriculture	Agriculture	Non-agriculture	Mixed sector
Dependent variable: Ln labor income per hour						
Schooling	0.0639***	0.0170	0.0588***	0.0186	0.0333	0.0545*
	(0.0142)	(0.0147)	(0.0174)	(0.0465)	(0.0258)	(0.0284)
Experience	0.0295	0.0111	0.0479	0.0221	0.0809	0.0037
	(0.0253)	(0.0249)	(0.0310)	(0.0553)	(0.0538)	(0.0557)
Experience squared	–0.0002	0.0000	–0.0006	0.0001	–0.0009	0.0000
	(0.0003)	(0.0003)	(0.0004)	(0.0007)	(0.0007)	(0.0007)
Predicted location index	2.5246*	–1.1726	–0.8480	–5.0917	3.9406	4.4038
	(1.3835)	(0.9139)	(1.3031)	(6.4898)	(2.7647)	(3.5891)
Constant	0.1415	1.1366**	0.5615	–0.3649	–0.2314	0.8447
	(0.4635)	(0.4275)	(0.5024)	(1.0451)	(0.9995)	(1.0675)
R^2	0.0421	0.0417	0.098	0.0532	0.0177	0.0268
F-statistic	6.48***	2.14*	4.28**	2.16*	1.23	1.61
N	485	58	96	143	88	100

Notes: This table uses "predicted location index" as a measure of geographic isolation. Figures in parentheses are corrected standard errors. The following levels of significance apply: * p<0.1, **p<0.05, *** p<0.01.
Source: FUSADES (1999).

Experience increases men's labor income when they work in the salaried non-agriculture sector. When working in home production, geographic isolation decreases men's labor income. According to the estimated coefficients, a man with three years of schooling, 20 years old, living far away (with a location index of 0.2) earns 11.80 colons per hour; a man with similar education and experience but living closer (with a location index of 0.4) earns three times more (36 colons per hour). Women with the same schooling and experience but living in isolation (location index 0.2) earn 3.78 colons per hour, while women with the same human capital accumulation but living closer to urban jobs (location index = 0.4) earn nearly twice that amount (6.25 colons per hour).

Conclusions and Policy Implications

The main objective of this chapter has been to study the consequences of geographic exclusion on three labor market outcomes: the labor force par-

ticipation decision, sector of employment, and labor income. The following results stand out:

- Geographic isolation does not discourage rural men from working. On the contrary, living at a greater distance from urban and *maquila* jobs increases men's likelihood of working, but the size of the effect is small (as 90 percent of men in the sample work).
- Male and female Salvadoran workers are non-randomly allocated into five distinct sectors: agriculture and non-agriculture, salaried and home production, and the mixed sector. The rural labor market of El Salvador is then segmented. Each sector has a different labor income-setting mechanism.
- The degree of geographic isolation determines women's sector allocation more than men's. Women living in isolation are especially highly concentrated in home production agricultural activities where women's skills are not rewarded as highly as men's. Home production in agriculture is a sector where women's human capital accumulation does not determine income labor level, whereas it does reward men's skills. By concentrating in self-employed agriculture production, women living in geographic isolation have worse labor outcomes than men.
- Education is positively correlated with labor income, especially when individuals work in the salaried sector. However, traditional human capital variables explain only a small percentage of the labor income variation in the rural sample; consequently, other variables must better explain the marginal product of labor when a rural sample is used. The results of the present income labor estimations may have omitted variable biases.
- According to the labor income results, living geographically excluded decreases women's labor income. When working in home nonagricultural production, geographic isolation has a negative impact on men's labor income.

In light of these findings, providing individuals with general skills or with incentives to obtain those skills may not be by itself a sufficient economic and social policy to reduce poverty and to better incorporate into society those Salvadorans who are presently excluded. Increasing education may increase

men's and women's labor income, but only when they are working in the non-agriculture salaried sector. By increasing individuals' education in rural areas, it may be possible to promote a more egalitarian labor income distribution (i.e., decreasing the rural-urban labor income gap, improving the general welfare of women and their children). However, education in rural areas does not automatically translate into higher labor income; there are sectors where education does not correlate with labor productivity.

The chapter has shown that women are neither more nor less geographically excluded than men. However, women living in isolation do concentrate in home agricultural production, an economic sector where women's productivity may have less value than men's. Whether sector allocation is by choice, or whether that choice is affected by the perceived or real lack of access to employment in other sectors, remains unanswered. Men have traditionally worked in agricultural activities and now capitalize on that training and specialization; to raise women's productivity in this sector, agricultural training could be promoted specifically for women, particularly those living in geographic isolation.

In general, a combination of policies may be of value. These include expanding and improving infrastructure through measures such as extending paved roads, increasing the availability of potable water and electricity, and improving public transportation. These measures can be further accompanied by the deconcentration of economic activities from urban to rural areas. Technology, information on access to new markets, diversification, and sources of employment other than agriculture may in that way reach rural and isolated areas. Ultimately, these policies may help better incorporate geographically excluded people into society and improve the efficiency of the labor market and society at large.

This page intentionally left blank

CHAPTER SIX

Language Barriers and Schooling Inequality of the Indigenous in Mexico

Susan W. Parker
Luis Rubalcava
Graciela Teruel

At the beginning of the 20th century, 80 percent of Mexico's population was illiterate, and the average Mexican completed only three years of schooling.[1] Today, half the population has reached the first year of secondary school, and only one of every 10 Mexicans is unable to read or write.[2]

These improved social indicators and increases in human capital formation have not, however, been associated with much reduction in income inequality.[3] And the social progress that has occurred has been far from homogeneous. Whereas there have been improvements in social indicators in urban areas, such indicators in remote rural communities have continued to lag far behind those of cities. These isolated communities are home to the vast majority of indigenous groups, which in many dimensions remain socially excluded from the rest of the Mexican population.

This chapter concentrates on one type of social exclusion faced by the indigenous: access to the formal education system. Unlike the average Mexican, the indigenous have seen little improvement in their schooling performance over time. Yet, educational attainment is most likely to determine the income and poverty levels of children in the future; increasing it is thus critical for avoiding the transmission of social exclusion to future generations. Table 6.1 shows the importance of education by comparing the labor income of indigenous and non-indigenous prime-age males. Overall, average labor

[1] INEGI (1994).
[2] Barro and Lee (1996) and Scott (2000).
[3] In 1950, the country's Gini coefficient was 52, but by 1990 it had increased to 60.5 (Scott, 2000).

Table 6.1. Indigenous Status, Education and Labor Income: Adult Men Aged 16 to 55

	Indigenous	Non-indigenous	T stat-sig. dif.
National sample	967.9	1,748.4	9.8
	(1,138.4)	(1,829.9)	
By years of completed schooling			
0	570.0	1,010.3	3.7
	(750.9)	(1,118.9)	
1 to 5	695.0	1,096.2	5.4
	(586.5)	(1,010.8)	
6	830.9	1,324.5	4.2
	(666.1)	(1,298.6)	
7 to 9	1,339.9	1,520.0	1.0
	(1,720.1)	(1,396.7)	
More than 9	2,359.7	2,835.1	1.5
	(1,702.8)	(2,461.5)	

Note: Standard deviations in parentheses.

earnings for the indigenous are only about half those of non-indigenous men. These differences, however, are greatest for individuals with low levels of schooling (primary or less), whereas there are no significant differences in labor earnings between indigenous and non-indigenous men with secondary education. The table suggests that the negative effect of being indigenous on earnings is greatly reduced, or even disappears, with increasing education levels. In this way, improving education of the indigenous can significantly reduce the degree of their social exclusion, at least in economic terms.

There is little evidence either in the literature as a whole or in the specific case of Mexico regarding the potential problems and limitations that indigenous children may face in school—this despite the fact that the country's indigenous population is nearly 10 million, the largest in Latin America. This chapter analyzes to what extent being indigenous is associated with lower schooling outcomes for children living in remote rural communities. It compares different schooling indicators among children with the same socioeconomic opportunities, household demographics and community characteristics, age and gender, but with different ethnic backgrounds.

The analysis exploits unique survey information on households in Mexico's poor rural communities, where most of the country's indigenous households are located. The chapter also carries out the same analysis with a nationally representative household survey to investigate the comparability of the results.

This chapter represents one of the first studies to specifically analyze the determinants of indigenous educational outcomes. Whereas there is some previous descriptive evidence suggesting that indigenous children tend to have lower educational outcomes than non-indigenous children, the potential factors associated with low performance have not been studied. Potential explanatory factors include fewer family resources, access to lower quality schools, discrimination, and cultural and language barriers (Caso et al., 1981). If indigenous households are poorer than other households and poverty affects school decisions, then simple correlations cannot determine whether poor performance among indigenous children is due to fewer family resources or other causes such as cultural factors, language barriers or access to lower-quality schools. This chapter attempts to shed some light on the extent to which family resources versus cultural factors, and versus language barriers, are related to the lower educational outcomes of indigenous children. This also makes it possible to analyze the extent to which being indigenous is a heterogeneous concept, that is, whether some indigenous children are more socially excluded than others and why.

The results show that indigenous children do indeed fare worse than their non-indigenous classmates, even within the relatively homogenous rural and marginalized communities of the sample. Nevertheless, there is important heterogeneity within the indigenous group. In particular, monolingual indigenous children (those who speak only a native language) do much worse in school than bilingual indigenous children who speak Spanish as a second language.

Regression models of the determinants of children's schooling outcomes provide concrete evidence as to the reasons for this poor performance. The models first control only for background family characteristics and resources at the household level, followed by community effects. Instrumental variable methods are then used to explore the possible endogeneity of the language spoken within the indigenous population. While controlling for parental characteristics, household wealth and community variables reduces the overall size of the negative effect of speaking only a native language, the effect remains

significant. These results indicate that while family and community resources are clearly important, they cannot explain all of the differences in educational attainment between bilingual and monolingual indigenous children. Rather, language barriers or cultural factors represent an important aspect of problems that indigenous children face in school.[4] Indigenous children who do not learn Spanish are more likely to face greater social exclusion both in and out of school than their bilingual counterparts.

In order to better understand the extent to which these language barriers affect schooling outcomes and how they differ from other unobserved cultural factors, this chapter exploits the fact that the Mexican public education system offers bilingual education at the primary level in a system parallel to its traditional schools. The Secretary of Public Education (SEP) provides bilingual primary schools that include bilingual teachers as well as textbooks in native languages. It is thus possible to examine whether the availability of bilingual schools can reduce the language barriers described above. Based on a difference-in-difference estimation that sweeps out most of the unobserved cultural heterogeneity across indigenous groups, the results suggest that the educational disadvantage due to language barriers is reduced for children who have the option of attending a bilingual primary school. In this sense, having access to a bilingual school may reduce the social exclusion of monolingual indigenous children.

These results have important policy implications for indigenous learning in Mexico. Controlling for an important number of measures of family resources, access to schools and community characteristics, indigenous monolingual children continue to perform worse than their bilingual counterparts as well as non-indigenous children. This suggests that not knowing

[4] There are many potential factors that may affect the investment (or lack thereof) of indigenous families in the education of their children. These factors may or may not relate to market failures. Possible market failures affecting this investment include the fact that indigenous families may have less information than non-indigenous families about the returns to education, or less access to credit with which to finance their children's schooling. Alternatively, if the quality of education available to indigenous children is lower than that available to non-indigenous children, potential returns to indigenous education may be lower, and in this sense it may be "rational" or "efficient" for indigenous families to invest less. Alternatively, indigenous parents may expect their children to work in activities with low returns to schooling and so may see the overall benefits to schooling as lower than non-indigenous parents. This chapter does not attempt to distinguish between the efficiency and distributional aspects of lower educational attainment by indigenous children.

Spanish represents an important factor in their lower levels of performance. Indigenous primary schools that practice bilingual education improve the educational performance of monolingual children relative to bilingual children, evidence that corroborates the importance of language barriers.

Thus, higher poverty levels can only partially explain differences between indigenous and non-indigenous educational outcomes, suggesting that social programs aiming only at lessening the marginality of indigenous communities will not be sufficient to eliminate schooling inequality among the indigenous. The policy prescriptions would thus call for study of the best ways to promote learning by indigenous children, and in particular the learning of Spanish. To the extent that bilingual primary schools seem to improve the performance of indigenous children, expansion of these integrated educational programs may be warranted.

One caveat, however, should be noted. The effectiveness of these programs may be reduced if bilingual primary schools decrease the probability that indigenous children learn Spanish and the corresponding likelihood of assimilation. Stated another way, whereas bilingual schools may reduce the social exclusion of school-age indigenous children, the issue is whether they may eventually increase social exclusion of adults if bilingual schools reduce the probability of learning Spanish during childhood.

Background

There appear to be no previous studies that have analyzed the effect of language and being indigenous on children's schooling attainment. Perhaps the closest in nature is a study by Rosenthal, Baker and Ginsburg (1983) that examines the impact of language (English literacy) on children's achievement in math and reading in primary school in the United States. They find that household economic factors explain most of the estimated differences between the achievement of native English and Spanish-speaking children. Nevertheless, a sizable language barrier remains even after controlling for household economic variables.

There is also literature in the United States on the impact of language spoken on the earnings of immigrants. Chiswick (1991) analyzes the determinants of English language fluency among migrants to the United States and the relationship this fluency has with earnings. He shows that length of

time in the United States is closely associated with English language fluency and that language fluency is an important determinant of earnings. In a subsequent study, Chiswick and Miller (1995) attempt to treat language fluency as endogenous in the context of an analysis of the foreign-born population in Australia, analyzing whether individuals who anticipate greater earnings from fluency are more likely to become fluent. They conclude that language fluency is in fact endogenous to earnings, although the exogeneity of their identifying instruments is somewhat questionable.

Even at a purely descriptive level, there has been little previous research on the educational attainment of indigenous children in Mexico or elsewhere. One exception is Panagides (1994), who uses the Survey of Income and Expenditures (ENIGH) to look at various economic dimensions of indigenous individuals and families in Mexico, including educational attainment and earnings. Nevertheless, since this survey contains no information on whether individuals are indigenous, the indicator constructed to measure indigenousness is a community-based indicator, defined by the overall percentage of individuals speaking an indigenous language in the municipality of residence.

López (1999) has analyzed the impact of a Mexican social program called PARE (*Programa para abatir el rezago educativo*) that provided school resources such as textbooks and teacher training to primary schools in Mexican states with high rates of poverty and low educational attainment on student test scores in math and Spanish. Prior to the program, test scores were lowest for children enrolled in bilingual schools. As a result of the program, improvements in test scores occurred in all areas, with the greatest improvements occurring in bilingual schools. Even after the program, however, test scores in bilingual schools remained lower than in other primary schools in rural areas. While insightful, the study does not have the richness of the household level data used in the present study.

Indigenousness is a complex concept, involving cultural traditions, languages and practices that have developed over centuries. In the case of Mexico, there are 62 different ethnic indigenous groups, speaking more than 80 different languages and with different sets of traditions. While the definition of being indigenous would ideally be multi-dimensional, including not just language but other indicators as well, data constraints restrict the present definition to language spoken.

Fortunately, however, the available data include specific definitions of language spoken. Each individual is asked if he or she speaks an indigenous

language. Those that report they do speak an indigenous language are then asked if they also speak Spanish. In this way the important distinction can be made between indigenous children who speak only a native language and indigenous children who are bilingual.[5]

There are also separate indicators of whether the parents of an indigenous child speak only an indigenous language or are bilingual. While parental language is highly correlated with the language spoken by the child, there is some important variation, particularly between the father and mother. It is much more common for fathers than mothers to be bilingual.

This chapter exploits the richness of individual-level data to analyze the impact of being indigenous on children's schooling outcomes. Taking into account the variation that exists between mother, father and child languages spoken, it is possible to control for endogeneity of languages spoken within the indigenous population. That is, it is possible to take into account the fact that learning Spanish may be a choice.

Descriptive Analysis

The nationally representative sample of the Mexican population used for this study—the Survey of Household Socioeconomic Characteristics (ENCASEH)—found that approximately 5 percent of all Mexican children are indigenous. Of these, 70 percent report speaking Spanish. The survey also shows that a majority of the indigenous population is located in rural areas. Within urban areas, only 1.1 percent of children are indigenous, versus 11.8 percent of children in rural areas.[6]

[5] Ideally one would prefer a more objective measure of indigenous status—one not dependent on self-reporting, which is potentially susceptible to a stigma effect. If indigenous respondents do not accurately report their real status because of concerns about possible discrimination, a downward bias may result in the estimation of the differential schooling outcomes between non-indigenous and indigenous children whose school performance is the lowest.

[6] It should be noted that the ENCASEH survey only contains information on 9,910 households, so the number of indigenous respondents is quite small. There were 658 indigenous children between the ages of six and 18 in the survey, and fewer than 200 said they spoke only an indigenous language. The reported nationwide percentages of the indigenous population may therefore differ from other nationally representative surveys, such as the census. ENCASEH is used only to ensure that the results are not overly biased by the use of a sample that is not representative at the national level.

Using a non-parametric regression with a national sample of the Mexican population, Figure 6.1 illustrates the relationship between completed years of schooling for three groups of children: non-indigenous, monolingual indigenous and bilingual indigenous.[7] The figure shows little differences between the three groups below eight years old, but larger and increasing differences as ages increase. As expected, non-indigenous children show the highest achievement of the three groups, followed by bilingual children. Monolingual indigenous children lag behind in all age groups. In general, the figure indicates that indigenous children who remain monolingual achieve very low levels of education on average, while indigenous children who learn Spanish over time (bilingual) perform better, although not as well as non-indigenous children. This is true across the entire child-age distribution.

Using the same axes as Figure 6.1, Figure 6.2 shows results for a sample of marginalized rural areas, where most of the indigenous population is concentrated. Interestingly, the figure shows little difference in terms of years of completed schooling between non-indigenous children and bilingual indigenous children. There are, however, huge differences between monolingual indigenous children and the other two groups above the age of eight. By the age of 18, the average monolingual indigenous child has achieved only about 2.5 years of completed schooling versus the other two groups, which achieve more than twice that level, or on average about seven years of schooling. The results from this regression foreshadow the regression analysis, which will demonstrate the great importance of language in determining educational outcomes.

Data Description

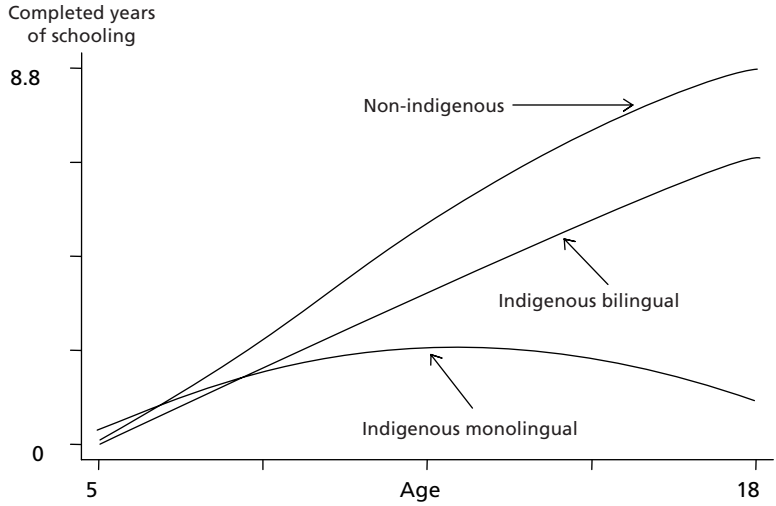
The data used for this chapter come directly from the Mexican Education, Health, and Nutrition Program (PROGRESA), a large anti-poverty program

[7] The non-parametric estimator applied carries out locally weighted, smoothed scatter plots (LOWESS). In this procedure, the regression is weighted so that the point in the middle gets the highest weight and points farther away receive less weight. This local average depends on the amount of smoothing, which in turn is affected by the choice of bandwidth h, as in

$$f_k = \frac{1}{nh}\sum_{i=1}^{n} K\left[\frac{x - X_i}{h}\right]$$

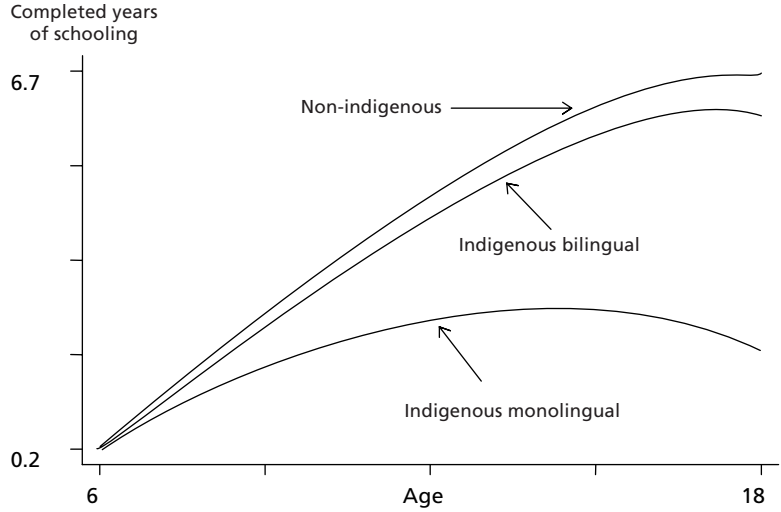
where K was chosen to be the Epanechnikov, since it has the property that it is most efficient in minimizing the mean integrated squared error.

Figure 6.1. Age and Education of Children Ages 6–18, by Language Spoken
(Nationally representative)



Source: INEGI (2000) and authors' calculations.

Figure 6.2. Age and Education of Children Ages 6–18, by Language Spoken
(Rural representative)



Source: INEGI (2000) and authors' calculations.

in rural areas that provides monetary and in-kind benefits linked to regular school attendance and health clinic visits by the family. The program has collected a large amount of socioeconomic information as a result of its beneficiary selection mechanism and evaluation procedures.

Two principal, related sources of information are used for the analysis. The first is drawn from PROGRESA's targeting mechanism, which involves carrying out a socioeconomic survey (ENCASEH) of all households in the isolated rural communities eligible to participate in the program. As of 2000, this involved about 3 million household interviews. The data collected include information on educational attainment, monetary income, durable goods and labor force participation. To facilitate analysis of the characteristics of the indigenous population, a random (representative) sampling was taken equal to about 120,000 households. This data, which include cross-sectional information for all 32 Mexican states and are the main source for this chapter, provide an overview of the indigenous population living in marginalized rural areas throughout Mexico. While not representative at the national level, the survey does provide data on an estimated 60 percent of indigenous households in Mexico.

Nevertheless, to ensure that the results are valid in making inferences about the Mexican population,[8] a nationally representative survey is also used. A separate national sample of the same ENCASEH questionnaire was carried out in 1997 and provides a convenient way to compare the results. This national sample includes 9,910 households and is representative of both urban and rural areas in Mexico.[9]

The dependent variables address two short- and long-term educational outcomes for boys and girls between the ages of 6 and 18: enrollment and years of completed schooling.[10] Household and student-level data are supple-

[8] For instance, the results could be biased by sample selection if indigenous individuals in marginalized areas are not representative of all indigenous individuals.

[9] To avoid confusion, this survey will be referred to as the national ENCASEH, whereas the ENCASEH drawn from the rural communities will be called the 32-state ENCASEH.

[10] The analysis was also performed using the schooling gap measure, defined as age-years of schooling –6: an indicator of the extent to which a child is "behind" where he or she should be in school. The lack of presence of non-linearities in the relation between age and education resulted in schooling gap estimates quite similar to those obtained with years of schooling, so the former are not reported. These results showed that indigenous children have, on average, a schooling gap 2.3 percent higher than non-indigenous children.

mented with school level information from the Secretary of Public Education (SEP), which makes it possible to link the characteristics of available schools to children's educational outcomes. These data, drawn from a census collected by the SEP at the time the household data were surveyed, provide specific information about each school, including number of pupils, education of teachers, and characteristics of the school infrastructure. In particular, at the primary level, information is used on two parallel programs offered by the SEP (regular and bilingual) to analyze the effect of bilingual education on the educational achievement of indigenous children.

According to the 32-state ENCASEH survey of 127,844 families, 29.2 percent (37,346) of heads of households reported speaking an indigenous language. Of these, 87.7 percent (32,435) also reported that they speak Spanish, suggesting that a minority of household heads speak *only* an indigenous language. Of children aged 6 to 18, 23.8 percent reported speaking an indigenous language, and of these, 81.7 percent also reported speaking Spanish. It is thus clear that this 32-state sample has a much larger concentration of indigenous families than the national level.

Methodology and Results

Disentangling Economic Conditions

The previous descriptive non-parametric results clearly showed that monolingual indigenous children lagged behind in schooling outcomes. However, this descriptive evidence cannot determine whether the poor school performance of indigenous children simply reflects cultural and language barriers or in fact represents the inferior social and economic conditions that indigenous households may face. Disentangling these effects is crucial for policymaking. If the poor school performance of indigenous children is mainly driven by the poor economic conditions in which they live, then anti-poverty programs would be largely sufficient to reduce the education gap between indigenous and non-indigenous children. However, if poor schooling outcomes are the result of other structural factors, such as a language or cultural barrier—holding poverty levels constant—then social programs aimed only at improving the marginal conditions of indigenous communities will not be sufficient.

The attempt to disentangle the effect of a language barrier from that of socioeconomic resources begins with a regression analysis. The first step is to assess the association between schooling opportunities (as measured by enrollment and years of completed schooling) and belonging to an indigenous group. The following relationship is estimated for each household child in the sample:

$$S_{ic} = B_0 + X_{1ic}B_1 + X_{2ic}B_2 + \delta_1 INDIG_{ic} + u_c + \varepsilon_{ic} \quad (1)$$

where S_{ic} stands for the education outcome variables of the child i in community c, X_1 represents his or her observed characteristics including age and sex; X_2 represents a set of household characteristics including mother's and father's education and age, and measures of household wealth and dwelling characteristics. These measures include ownership of land, access to water and electricity, whether the floor of the house is made of cement, and ownership of durable goods including a refrigerator and a stove. *INDIG* is an indicator of whether the child belongs to an indigenous group, which takes the value of one if the child speaks an indigenous language. The model also includes a community fixed effect u_c, which sweeps out any community time-invariant characteristic that may be correlated with indigenous child schooling outcomes, such as local infrastructure or cultural effects. The term ε_{ic} corresponds to an error component that reflects all remaining unobserved characteristics of the model.

The particular hypothesis of interest for this study is whether children belonging to an indigenous group display lower school productivity, all else being equal. In other words, the hypothesis being tested is whether δ is different than zero. A negative coefficient would imply a negative effect of group membership with respect to school opportunities, suggesting that indigenous children are in a disadvantaged position relative to their non-indigenous classmates.

Table 6.2 shows the determinants of years of schooling and, in particular, the effect of being indigenous. The table employs a general measure of being indigenous that includes children who are either bilingual or monolingual. The analysis will later disaggregate the two groups.

Table 6.2 also reports a number of specifications, beginning with a minimal specification in which years of completed schooling depend only on child characteristics. Subsequently added specifications include paren-

tal characteristics, household wealth indicators, and community effects. This makes it possible to analyze the extent to which the impact of being indigenous is altered by separately including these household characteristics. Column (1) shows that an indigenous child is on average half a year behind relative to non-indigenous classmates of the same age. As expected, years of completed schooling is a monotonic function of the years of age of the child. Column (2) adds the age of the parents in years. Children with older mothers tend to be more educated, but only marginally. In order to see whether the maternal age effect is contaminated with a human capital effect (older parents tend to be less educated, since they belong to older generations), parental levels of education are further controlled for in column (4). The effect of the mother's age rises marginally and the father's age effect becomes positive and significant. As expected, children whose parents have higher levels of education are more likely to have more years of completed schooling. This may reflect either parental ability in child-rearing or economic conditions of the household, as parental human capital is also a measure of permanent income. It is worth noting that the magnitude of the coefficient of being indigenous decreases, suggesting that the effect of being indigenous is highly correlated with household resources.

Column (5), along with parental characteristics, further controls for household wealth. In particular, controls are included for whether the dwelling has concrete floors, walls and ceilings, and whether the household has access to running water and electricity, owns agricultural land, and owns durable goods such as a refrigerator and stove. All wealth measures are significant and have the expected sign. However, and perhaps more importantly, comparison of columns (5) and (1) shows that the impact of being indigenous on educational outcomes is reduced by half. That is, controlling for parental and household characteristics demonstrates that at least half of observed differences between the indigenous and the non-indigenous are due to family background.

Moreover, it is very likely that indigenous households live in poorer communities with poorer infrastructure relative to the rest of the population. Therefore, failing to control for observed and unobserved community heterogeneity could cause an overestimation of the indigenous impact as well as the effect of household resources on child schooling outcomes. To correct this problem, column (6) introduces community fixed effects. It is

Table 6.2. Determinants of Years of Completed Schooling: Effect of Being Indigenous, Children Ages 6 to 18

	OLS					CFE	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Child Characteristics							
Child is indigenous	–0.571	–0.58	–0.372	–0.348	–0.246	–0.253	–0.501
	[0.010]***	[0.010]***	[0.010]***	[0.010]***	[0.011]***	[0.025]***	[0.234]**
Gender (Boy = 1)	0.002	–0.007	–0.003	–0.003	–0.004	–0.002	0.043
	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]	[0.008]	[0.054]
Age 9 to 11	2.203	2.205	2.259	2.214	2.203	2.18	2.321
	[0.008]***	[0.008]***	[0.008]***	[0.008]***	[0.008]***	[0.011]***	[0.078]***
Age 12 to 14	4.227	4.233	4.341	4.254	4.232	4.191	4.385
	[0.010]***	[0.010]***	[0.010]***	[0.010]***	[0.010]***	[0.012]***	[0.081]***
Age 15 to 18	5.324	5.354	5.505	5.374	5.347	5.326	5.577
	[0.012]***	[0.013]***	[0.012]***	[0.012]***	[0.012]***	[0.012]***	[0.083]***
Parental Characteristics							
Father's age		–0.002		0.008	0.007	0.003	0
		[0.001]**		[0.001]***	[0.001]***	[0.001]***	[0.005]
Mother's age		0.002		0.011	0.009	0.008	0.013
		[0.001]***		[0.001]***	[0.001]***	[0.001]***	[0.006]**
Father's edu 1 to 5 years			0.398	0.434	0.406	0.318	0.266
			[0.012]***	[0.012]***	[0.012]***	[0.014]***	[0.095]***
Father's edu 6 + years			0.656	0.748	0.645	0.425	0.288
			[0.014]***	[0.014]***	[0.014]***	[0.017]***	[0.111]***

Mother's edu 1 to 5 years			0.484	0.522	0.479	0.37	0.333
			[0.011]***	[0.011]***	[0.011]***	[0.013]***	[0.091]***
Mother's edu 6 + years			0.792	0.879	0.746	0.5	0.722
			[0.013]***	[0.013]***	[0.013]***	[0.017]***	[0.107]***
Assets							
Cement floor					0.306	0.258	0.292
					[0.012]***	[0.016]***	[0.088]***
Household has water and electricity					0.302	0.233	0.158
					[0.013]***	[0.016]***	[0.083]*
Household owns agric. land					0.034	0.154	0.257
					[0.009]***	[0.011]***	[0.081]***
Household has refrig. and stove					0.469	0.486	0.367
					[0.011]***	[0.015]***	[0.075]***
Observations	220,008	220,008	220,008	220,008	220,008	220,008	4,640
R-squared	0.53	0.53	0.55	0.55	0.56	0.57	0.59
Number of communities						25,905	255

Notes: Columns (1) through (6) present results using the ENCASEH 32-state sample. Column (7) estimates correspond to the rural national sample. Robust standard errors in [brackets]. Coefficients marked with (***) are significant under Schwartz (1978) a posteriori criteria, where the most likely model is chosen with a t-statistic no smaller than 3.5.

important to mention that the variation across indigenous groups derives mostly from comparisons across communities. Most rural communities in the sample tend to be either primarily indigenous or non-indigenous. Only about 10 percent of the 26,079 communities with at least one child in the sample have both indigenous and non-indigenous children.[11] Nevertheless, the fact that indigenous coefficients do not change significantly when community controls are added suggests not only that the potential sample selection bias is low, but also that family background variables capture most of the community effects.[12]

Column (7) presents the same specification as in column (6), but now a *national* sample is used, though one restricted to rural areas. This resulting sample is representative of all rural areas in Mexico.[13] The estimated effects of being indigenous are about double the size of those reported in the 32-state ENCASEH. This is perhaps not surprising, as the sample is much more heterogeneous than the sample under study, which only contains households living in highly marginalized communities. However, it is also likely that a more heterogeneous sample makes it less successful in controlling for a household's economic (unobserved) circumstances, and thus part of the estimated effect in this national sample of being indigenous reflects uncontrolled economic factors at the household level. The rest of the analysis continues with the poorer and more homogeneous 32-state ENCASEH, which better controls for household resources and thus better isolates the impact of being indigenous.

It is obvious that completed years of schooling for boys and girls between the ages of 6 and 18 is a long-term indicator, as opposed to current

[11] About 86 percent of the variation with respect to the indigenous derives from variation between communities.

[12] The household wealth coefficients change in magnitude but not in sign. Under community fixed effects, the coefficients on water and electricity and concrete dwelling characteristics are reduced, contrasting with the increase in the effect of owning agricultural land. The differential change of the wealth variables may be related to the fact that wealthier households with a concrete dwelling and available public services are located in wealthier communities, whereas agriculturally oriented households are more likely to be in rural communities with relatively less development.

[13] Urban areas are excluded from the national sample because of the very low proportion of indigenous children in urban areas. Of the 8,978 urban children ages 6 to 18 in the national ENCASEH, only 101 reported speaking an indigenous language.

school enrollment, which is short term. Table 6.3 presents specifications (5) and (6) for the second schooling outcome, current school enrollment. For comparability, the last two specifications of Table 6.2 are presented as the first two columns of Table 6.3. Columns (3) and (4) correspond to OLS and community fixed effects specifications for the probability of current enrollment, respectively.[14] In general, the results are similar to those using years of completed schooling: when household-level variables and community fixed effects are controlled for, indigenous children remain likely to do worse than their non-indigenous classmates. Column (2) shows that indigenous children are, on average, 3.2 percent less likely to be enrolled in school, even after controlling for household and community characteristics.

It should be noted that the effect of the coefficient of being indigenous on the likelihood of currently attending school reverses in sign after community fixed effects are controlled for. This is not the case for years of schooling, which suggests that unobserved community characteristics are differentially correlated to schooling attendance of indigenous children. The last two columns of Table 6.3 attempt to clarify these differences. In columns (3) and (4) the indigenous are divided into two groups, monolingual and bilingual, to show the differential impact on school enrollment.[15] The effect of being monolingual for indigenous children increases in magnitude but remains negative when including community fixed effects. This is not the case for the bilingual indigenous coefficient, which turns from being positive and significant to negative and significant. These results, which are meant only to be suggestive,[16] show that unobserved community characteristics differentially affect the schooling enrollment of monolingual and bilingual indigenous children.[17] Therefore, the regressions presented above

[14] For comparability purposes, and given the large sample size, the analysis exploits the fact that the Linear Probability Model (LPM) coefficients are consistent and estimates the probability of school enrollment using OLS models.

[15] As in previous specifications, non-indigenous is the omitted child category.

[16] As discussed further below, the language division is considered to be endogenous to the determinants of educational outcomes.

[17] A more detailed examination of the characteristics of the communities reveals that an important characteristic explaining this reversal is the percentage of indigenous children in the community and, in particular, the percentage of those speaking only a native language. This community variable is obviously endogenous to the model, however, and this specification is therefore not presented in the regressions.

Table 6.3. Long- vs. Short-term Schooling Outcomes

	Years of schooling		School enrollment			
	OLS [5]	CFE [6]	OLS [1]	CFE [2]	OLS [3]	CFE [4]
Child characteristics						
Child is indigenous	–0.246	–0.253	1.81	–3.227		
	[0.011]***	[0.025]***	[0.205]***	[0.501]***		
Only speaks indigenous language					–10.662	–17.036
					[0.454]***	[0.725]***
Speaks indigenous & Spanish					4.334	–2.664
					[0.214]***	[0.501]***
Gender (Boy = 1)	–0.004	–0.002	3.353	3.431	3.253	3.344
	[0.008]	[0.008]	[0.159]***	[0.162]***	[0.159]***	[0.162]***
Age 9 to 11	2.203	2.18	5.319	4.985	4.884	4.601
	[0.008]***	[0.011]***	[0.179]***	[0.230]***	[0.178]***	[0.230]***
Age 12 to 14	4.232	4.191	–14.386	–14.635	–15.059	–15.225
	[0.010]***	[0.012]***	[0.241]***	[0.236]***	[0.240]***	[0.237]***
Age 15 to 18	5.347	5.326	–57.617	–57.304	–58.378	–58.001
	[0.012]***	[0.012]***	[0.245]***	[0.240]***	[0.243]***	[0.241]***
Parental Characteristics						
Father's age	0.007	0.003	0.056	–0.002	0.048	–0.004
	[0.001]***	[0.001]***	[0.013]***	[0.015]	[0.013]***	[0.015]
Mother's age	0.009	0.008	0.003	–0.014	–0.004	–0.013
	[0.001]***	[0.001]***	[0.014]	[0.016]	[0.014]	[0.016]

Father's edu 1 to 5 years	0.406	0.318	4.968	4.086	4.642	4.07
	[0.012]***	[0.014]***	[0.245]***	[0.283]***	[0.244]***	[0.283]***
Father's edu 6 + years	0.645	0.425	9.45	7.196	8.941	6.94
	[0.014]***	[0.017]***	[0.289]***	[0.350]***	[0.289]***	[0.349]***
Mother's edu 1 to 5 years	0.479	0.37	6.073	4.989	5.542	4.898
	[0.011]***	[0.013]***	[0.221]***	[0.259]***	[0.221]***	[0.259]***
Mother's edu 6 + years	0.746	0.5	10.957	7.669	10.369	7.515
	[0.013]***	[0.017]***	[0.269]***	[0.332]***	[0.269]***	[0.332]***
Assets						
Cement floor	0.306	0.258	2.473	3.317	2.401	3.308
	[0.012]***	[0.016]***	[0.241]***	[0.321]***	[0.241]***	[0.321]***
Household has water and electricity	0.302	0.233	3.135	2.768	3.095	2.766
	[0.013]***	[0.016]***	[0.243]***	[0.314]***	[0.242]***	[0.314]***
Household owns agric. land	0.034	0.154	1.473	2.463	1.694	2.47
	[0.009]***	[0.011]***	[0.167]***	[0.229]***	[0.167]***	[0.228]***
Household has refrig. and stove	0.469	0.486	5.471	5.329	5.579	5.372
	[0.011]***	[0.015]***	[0.215]***	[0.292]***	[0.215]***	[0.292]***
Observations	220,008	220,008	220,712	220,712	220,716	220,716
R-squared	0.56	0.57	0.34	0.35	0.35	0.35
Number of communities		25,905		25,907		25,907

Notes: See Table 6-1. Results use the ENCASEH 32-state sample. School enrollment indicator multiplied by 100.

are, to some extent, an average effect of these two "types" of children and may mask important differences between the two groups. Furthermore, the persistent negative effect of being monolingual indigenous, as shown in columns (3) and (4) of Table 6.3, suggests that indigenous children are not only economically disadvantaged, but may also face other cultural or language barriers when attending school.

Determining Schooling Outcomes of Indigenous Children

To disentangle the possible cultural or language barrier effect from other factors, the sample is restricted to indigenous children, dividing them into two groups: those who speak Spanish (bilingual) and those who are monolingual. The purpose is to examine the extent to which there is heterogeneity among the indigenous population and the extent to which learning or not learning Spanish affects children's performance in school. The following specification is used:

$$S_{ic} = \mathrm{B}_0 + X_{1ic}\mathrm{B}_1 + X_{2ic}\mathrm{B}_2 + \delta_2 MONO_{ic} + u_c + \varepsilon_{ic} \qquad (2)$$

where S_{ic}, X_{1ic}, and X_{2ic} are as defined above, *MONO* is an indicator of whether the indigenous child is monolingual or bilingual (speaking Spanish as a second language), u_c is a community fixed effect, and ε_{ic} corresponds to all remaining unobserved characteristics.

Table 6.4 presents the findings on the impact of only speaking a native language on years of completed schooling and school enrollment of indigenous children. Columns (1) through (4) show varying specifications relating to the inclusion and exclusion of speaking only the native language, relative to the excluded category of also speaking Spanish as a second language. Column (1) shows OLS estimates of the impact of a child's language, controlling only for child characteristics, whereas column (2) includes parental and household characteristics and community fixed effects as well. Column (3) excludes the language spoken by the child and includes the languages spoken by the parents, while column (4) includes the languages spoken by both the child and the parents.

Columns (1) and (2) make clear that the child being monolingual has very large negative effects on schooling outcomes relative to other indigenous children that do speak Spanish. For example, an indigenous child who

does not speak Spanish is (on average) 1.1 years behind in terms of completed years of schooling relative to bilingual classmates with the same household and community resources (column 2); and is 14 percent less likely to be currently enrolled in school (column 4). These results make clear the large educational disadvantage that results when indigenous children do not learn Spanish. It is noteworthy that these differences are much greater than the overall differences previously observed between indigenous and non-indigenous children (Table 6.2).

Furthermore, specifications (3) and (4) show that the language barrier effect only operates through the child's ability to speak Spanish. Column (3) shows that while parental language has significant impacts on the child's educational outcomes, the (absolute) size of the coefficients is much smaller than the size of the coefficient when whether the child is monolingual is controlled for. Moreover, column (4) demonstrates that the child's monolingual effect is robust to the inclusion of controls of whether parents are also monolingual. The mother's and father's language has no significant effect on the child's human capital assessment after the child's language is controlled for. These two facts suggest that the child's "monolingual" effect is likely to reflect more a language barrier at school, rather than a parental or household (unobserved) cultural factor.

It is important to emphasize here, however, that the variable *MONO* is unlikely to be exogenous to schooling determinants. Using non-parametric analysis on the probability of being monolingual for indigenous children, Figure 6.3 shows that indigenous children who lag behind in school are those who are unlikely to learn a second language. This implies that *being bilingual* is a dynamic concept and integrally related with school attendance. As children participate in school at a given age, learning obviously occurs in terms of languages. The sample also shows that whereas 37 percent of indigenous children speak only an indigenous language at the age of six, by the age of 18, only 10 percent of indigenous children are monolingual.

Consequently, failing to control for unobserved characteristics *at the household level*, such as parental tastes or child-rearing ability in human capital formation, may lead to an overestimation of the true language barrier effect if the decision to send the child to school is correlated with the characteristics of the parents. To overcome this problem, the language spoken by the child is treated as endogenous, and the child's probability of being mono-

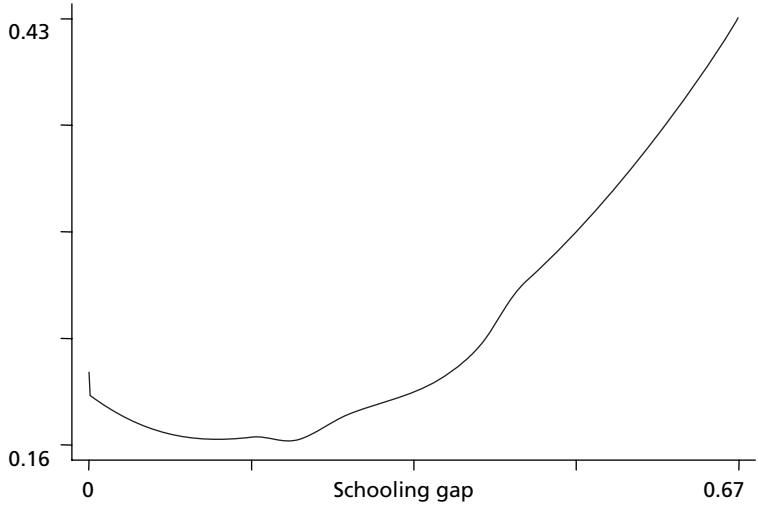
Table 6.4. Long- vs. Short-term Schooling Outcomes for Monolingual and Bilingual Indigenous Children

	Years of schooling					School enrollment	
	OLS [1]	CFE [2]	CFE [3]	CFE [4]	IV [5]	CFE [6]	IV [7]
Child characteristics							
Only speaks indigenous language	–1.455	–1.073		–1.091	–0.942	–14.255	–9.926
	[0.023]***	[0.029]***		[0.030]***	[0.055]***	[0.580]***	[1.146]***
Gender (Boy = 1)	0.193	0.189	0.214	0.188	0.181	5.992	5.511
	[0.018]***	[0.017]***	[0.017]***	[0.017]***	[0.017]***	[0.342]***	[0.339]***
Age 9 to 11	1.786	1.799	1.912	1.797	1.848	4.008	5.081
	[0.016]***	[0.024]***	[0.024]***	[0.024]***	[0.017]***	[0.485]***	[0.425]***
Age 12 to 14	3.479	3.518	3.694	3.515	3.593	–14.758	–13.685
	[0.021]***	[0.025]***	[0.025]***	[0.025]***	[0.023]***	[0.502]***	[0.551]***
Age 15 to 18	4.302	4.414	4.618	4.411	4.49	–56.786	–55.741
	[0.026]***	[0.026]***	[0.025]***	[0.026]***	[0.028]***	[0.514]***	[0.566]***
Parental characteristics							
Father's age		0.003	0.004	0.003	0.007	0.036	0.071
		[0.002]*	[0.002]**	[0.002]*	[0.001]***	[0.031]	[0.027]***
Mother's age		0.005	0.006	0.005	0.008	–0.073	–0.002
		[0.002]***	[0.002]***	[0.002]***	[0.002]***	[0.033]**	[0.030]
Father's edu 1 to 5 years		0.321	0.331	0.33	0.414	4.438	5.772
		[0.027]***	[0.029]***	[0.028]***	[0.023]***	[0.543]***	[0.478]***
Father's edu 6 + years		0.409	0.437	0.42	0.628	8.317	10.626
		[0.036]***	[0.038]***	[0.037]***	[0.029]***	[0.726]***	[0.597]***
Mother's edu 1 to 5 years		0.213	0.22	0.23	0.415	3.96	6.295
		[0.026]***	[0.028]***	[0.027]***	[0.021]***	[0.524]***	[0.438]***

Mother's edu 6 + years		0.283	0.32	0.304	0.604	5.281	9.258
		[0.038]***	[0.040]***	[0.039]***	[0.029]***	[0.768]***	[0.602]***
Father only speaks indigenous language			–0.13	0.039			
			[0.041]***	[0.041]			
Mother only speaks indigenous language			–0.115	0.069			
			[0.032]***	[0.032]**			
Assets							
Cement floor		0.306	0.323	0.305	0.371	2.827	1.852
		[0.047]***	[0.048]***	[0.047]***	[0.040]***	[0.950]***	[0.745]**
Household has water and electricity		0.194	0.202	0.194	0.119	2.843	1.033
		[0.039]***	[0.040]***	[0.039]***	[0.032]***	[0.787]***	[0.631]
Household owns agric. land		0.158	0.162	0.159	0.14	3.022	4.189
		[0.026]***	[0.026]***	[0.026]***	[0.020]***	[0.518]***	[0.397]***
Household has refrig. and stove		0.579	0.602	0.583	0.748	5.905	7.361
		[0.059]***	[0.060]***	[0.059]***	[0.056]***	[1.182]***	[0.939]***
Hypothesis testing							
IV overidentification test $X^2_{[21]}$					0.14		0.00
					(1.000)		(1.000)
R^2 First-stage regression					0.27		0.27
Hausman test for exogeneity $X^2_{[21]}$					42.04		23.97
					(0.004)		(0.244)
Observations	51,229	51,229	51,229	51,229	51,229	51,492	51,492
R-squared	0.47	0.5	0.49	0.5	0.5	0.34	0.33
Number of communities		6,432	6,432	6,432		6,433	

Notes: Results use ENCASEH 32-state sample for indigenous children only. School enrollment indicator multiplied by 100. Robust standard errors in [brackets]. Test p-values in (parenthesis). Coefficients marked with (***) are significant under Schwartz (1978), where the most likely model is chosen with a t-statistic no smaller than 3.3.

Figure 6.3. Probability of Being Monolingual Indigenous by Schooling Gap



Source: INEGI (2000) and authors' calculations.

lingual is instrumented with his or her parents' ability to speak Spanish as a second language.[18]

The results indicate that the mother's and father's ability (inability) to speak Spanish is a good instrument for a child's ability (inability) to speak Spanish. First, the mother and father's ability to speak Spanish does not significantly affect their child's schooling outcomes once a child's own ability to speak Spanish is controlled for. Second, the child's probability of speaking Spanish is highly correlated with her or his parents' language ability.[19] Finally, the Basmann (1960) IV over-identification tests show that parental language ability is a good instrument for identifying the structural model.

[18] An interesting, largely theoretical body of work has developed on motivations and incentives for learning a second language (see Lazear, 1999; Lang, 1986; Church and King, 1993). Lazear argues that incentives to learn a second language and other forms of assimilation are lower when the large majority of the group speaks the native language. However, Church and King argue that the overall benefits of language acquisition are increasing in the number of individuals who speak the language, giving rise to possible externalities that in turn imply that language acquisition decisions by the population may be inefficient.

[19] Coefficients of determination of the first-stage regressions of the probability of the child being monolingual against parental Spanish knowledge show an R^2 of 42 percent for years of completed schooling and an R^2 of 23 percent for school attendance.

Columns (5) and (7) in Table 6.4 present 2SLS estimates of the child's lack of ability to speak Spanish on completed years of schooling and on the probability of current enrollment in school, respectively. For both indicators, the effect remains negative and significant, although it decreases in magnitude. For example, 2SLS estimates on the child's probability of school enrollment show that failing to control for potential endogeneity would overestimate its effect by almost 50 percent: 2SLS estimates show that a child who does not speak Spanish is "only" 10 percent less likely to attend school, as opposed to an OLS community fixed effects (negative) probability of 14 percent (see columns (7) and (6), respectively, of Table 6.4). Moreover, Hausman specification tests reject the exogeneity of the child's being monolingual as an explanatory variable.

One final exercise examines how the language barrier may change or accumulate over time. The previous descriptive results have suggested a widening difference by age in years of completed schooling between monolingual and bilingual as well as non-indigenous children. The last specification is thus repeated in Table 6.4, using an instrumental variable approach to look at the determinants of years of completed schooling and school enrollment by age group. Table 6.5 presents the summary of these results, which show that initially (for the age group of 6 to 8), the effect of being monolingual on the years of completed schooling is relatively small for indigenous children, showing a reduction of only about 0.1 percent of one year. The effects of this language barrier, however, greatly increase with age. By the age group 15 to 18, the average difference due to language barriers is more than two years of schooling.

With respect to school enrollment, it is interesting to note that the estimated negative effect of being monolingual is large (about 10 percent) and apparent even at early ages (perhaps partially reflecting late school entry of monolingual relative to bilingual children). This negative effect continues to be strong on different age groups, with the exception of the oldest group (15 to 18), where the estimated effect of being monolingual is no longer a significant determinant of school enrollment.[20]

[20] This does not imply, of course, that there are no crude differences between school enrollment of monolingual and bilingual children at the ages of 15 to 18, but rather that these differences are no longer due to language barriers or other cultural differences.

Table 6.5. Estimated Effects of Being Monolingual on Schooling Indicators of Indigenous Children, by Age Instrumental Variable Estimations

	Age group			
	6 to 8	**9 to 11**	**12 to 14**	**15 to 18**
Years of schooling				
Child only speaks indigenous language	–0.069	–0.414	–1.602	–2.265
	[0.030]**	[0.071]***	[0.132]***	[0.213]***
Observations	12,638	12,227	12,236	14,128
R^2	0.33	0.22	0.21	0.24
School enrollment				
Child only speaks indigenous language	–10.586	–9.364	–15.181	–4.268
	[1.555]***	[1.712]***	[2.967]***	[3.434]
Observations	12,967	12,222	12,226	14,076
R^2	0.09	0.07	0.12	0.13

Notes: Results use the ENCASEH 32-state sample for indigenous children only. School enrollment indicator multiplied by 100. Robust standard errors in [brackets]. Test p-values in (parentheses). All models include child characteristics: sex and age; parental characteristics: father's age, mother's age, father and mother education groups; household assets: cement floor, water and electricity, agricultural land owning, refrigerator and stove.

In summary, failing to control for endogeneity leads to an overestimation of the language barrier effect. Nevertheless, even after controlling for endogeneity, the effect remains quite large, suggesting that language barriers result in a significant disadvantage in terms of overall human capital acquired by indigenous children, a disadvantage that accumulates as children grow older. Whereas most indigenous children do in fact learn Spanish, those who do not achieve much lower levels of education. It is these children who as adults will most likely experience the long-term consequences of extreme poverty.[21]

[21] Models carried out with gender interactions reveal a number of differences that, for reasons of space, can only briefly be discussed here. Replicating the results in Table 6.4 for girls and boys, the estimated negative effect of being monolingual on years of schooling is approximately 20 percent larger for girls than for boys. While the proportions of indigenous girls who learn Spanish are comparable with those of boys until about the age of 15, thereafter a lower proportion of girls than boys report learning Spanish. This coincides with an increasing gap between boys and girls in years of completed schooling and school enrollment during these ages. The largest gender gaps are evident in the monolingual indigenous population, with smaller gaps in the bilingual indigenous population and little or no gender gaps within the non-indigenous population. These patterns of school attainment and language acquisition are also somewhat related to marriage patterns, as by the age of 15, a significant number of indigenous girls begin to marry.

Disentangling the Effect of Language Barriers: Bilingual Education

The previous analysis has made abundantly clear that indigenous children who learn Spanish do much better in school than those who remain monolingual, which is a strong indicator of language barriers. To better understand how these barriers differ from the effect of other unobserved cultural factors, the possible role of bilingual education in improving the relative educational performance of indigenous children is now examined. Note that monolingual children may be more affected by cultural traditions that reduce their school attendance, as opposed to bilingual children who are more likely to have assimilated into more Western culture.

The language barrier is isolated from other possible cultural effects by exploiting the dual educational system provided by the SEP, which consists of *Spanish-type* and *bilingual-type* schools in rural communities. SEP currently operates bilingual primary schools and preschools in a number of indigenous communities. These schools use bilingual teachers with textbooks in the indigenous language. The goal is to "favor the acquisition, strengthening and development of the indigenous languages as well as the Spanish language by avoiding the imposition of one language over the other" (Caso et al., 1981). Proponents of indigenous education generally suggest that bilingual schools encourage students' educational attainment by encouraging them in class and by prompting teachers to pay more attention to and discriminate less against students who do not speak Spanish.

If the large educational difference between monolingual and bilingual children is mainly a language effect, it would be expected, first, that a bilingual school would reduce the negative effect of language between the two groups in terms of educational attainment; and second, that other school characteristics would have less or no impact in reducing this gap. Thus the following equation is estimated:

$$S_{ic} = B_0 + X_{1ic}B_1 + X_{2ic}B_2 + \delta_2 \mathrm{MONO}_{ic} + \delta_3 \mathrm{BILPRIM}_c + \delta_4 \mathrm{MONO}_{ic} * \mathrm{BILPRIM}_c + u_c = \varepsilon_{ic} \quad (3)$$

Model (3) is an extension of model (2). It interacts the condition that the child is monolingual (as opposed to bilingual) with an indicator variable, *BILPRIM,* that takes the value of one if the community provides a bilingual primary school.

The difference-in-difference nature of the coefficient δ_4 sweeps out the cultural unobserved heterogeneity between bilingual and monolingual children (to the extent that the impact of the unobserved heterogeneity does not vary across school programs) and thus better measures the size of the language barriers. The coefficient indicates whether the language gap in education between monolingual and bilingual children is different for children with a bilingual primary school in their community as opposed to those who do not have access to a bilingual primary school. If the language barrier hypothesis is correct, a positive coefficient would be expected. That is, assuming that bilingual schools reduce the language barrier, one should see a narrowing of the educational disadvantage between bilingual and monolingual indigenous children in communities with bilingual schools.

To test specification (3), the 32-state ENCASEH data are merged with data from the Secretary of Public Education (SEP) covering the same data collection period to determine whether a bilingual primary school is available for children in the community where they live. Availability is initially defined using the school closest (in kilometers) to the community where the child lives. At the primary level, this is normally the school or schools located within the community, as over 80 percent of communities have at least one primary school. When there is no school in the community, the distance in kilometers is calculated to the nearest community with a school, with a maximum distance of up to five kilometers,[22] and the characteristics of a nearby school or schools are used to represent the available supply of bilingual schools for the child. Nevertheless, in the hopes of better capturing behavior in terms of the school that children actually attend, the estimation is restricted in several ways.

First, the empirical model restricts attention to communities that have a primary school. A key underlying assumption in this matching procedure is that primary-level children do not attend community schools other than their own; this would preclude matching the *true* underlying community school infrastructure to the corresponding child. It appears, however, that *conditional on the community having a primary school*, children may be unlikely to commute to other communities, given that the sample is characterized by households with few resources and located in very isolated communities.

[22] This is done through the use of Geographical Information Systems (GIS) software.

It may still be the case in the sample, however, that a community has more than one available primary school. In this case, there would be no way to verify which primary school each child attends and it would be necessary to arbitrarily construct "averages" of school characteristics. Thus, to maximize the probability of matching school infrastructure to each child, the sample is further restricted to communities where there is only one available primary school.[23]

Finally, the analysis is restricted to current schooling enrollment, the short-run schooling outcome, to avoid the potential problem of migration. That is, during their life course, some children may have attended schools other than the school in the community where they live at the time of the interview.

Table 6.6 summarizes the findings. Columns (1) and (2) suggest that results drawn from the restricted sample are consistent with earlier evidence. IV estimates on monolingual indigenous children attending primary level with only one school in their community suggest they are 11 percent less likely to be currently enrolled relative to bilingual children, as opposed to the unconstrained sample with a probability of 10 percent (see Table 6.4). Columns (3) to (5) display the results of the difference-in-difference estimation. Here, the focus is on the difference-in-difference coefficient (δ_4), which should be unaffected by the possible endogeneity of language. From column (3), it can be seen that this coefficient implies a 3.5 percentage point reduction in the language barrier gap; however, it is not significant, given the large sample size.[24]

A potential empirical problem that arises here is that of endogenous program placement (Rosenzweig and Wolpin, 1986). It is likely that the Mexican government locates bilingual schools precisely in areas with higher indigenous populations and where the indigenous tend to be less integrated, more isolated and consequently less likely to learn Spanish. If this is the case, the difference-in-difference OLS estimators may be biased downward. To

[23] Overall, 55 percent of indigenous children in the sample have access to a bilingual school. Within this total, 51.8 percent of bilingual children have access to a bilingual school, whereas the figure for monolingual indigenous children rises to 71.3 percent.

[24] Given the large sample sizes, it may be appropriate to adopt a Bayesian approach to model selection. Following Schwarz (1978), the a posteriori most likely model will be chosen if a t-statistic greater than 3.15 is judged significant in the regressions in the table.

Table 6.6. School Enrollment: Language and Bilingual School Interactions

	OLS [1]	IV [2]	OLS [3]	OLS [4]	CFE [5]
Child only speaks indigenous language	–15.132	–11.443	–18.523	–16.38	–12.981
	[0.653]***	[1.426]***	[1.333]***	[3.909]***	[4.645]***
Bilingual school			2.83	2.474	
			[0.508]***	[0.652]***	
Indigenous language * bilingual school			3.622	6.178	6.793
			[1.512]**	[1.772]***	[2.000]***
Indigenous language* teacher education level 2				–4.251	–4.16
				[1.512]***	[1.792]**
Indigenous language * teacher education level 3				0.56	–6.235
				[2.225]	[2.734]**
Indigenous language * student teacher ratio				–0.057	0.033
				[0.096]	[0.104]
F(3, 21268)				3.25	3.00
Observations	21,274	21,274	21,274	21,268	21,268
R-squared	0.38	0.38	0.38	0.38	0.38
Number of communities					3,266

Notes: Results use the ENCASEH 32-state sample for indigenous children between 6 and 12 years old in communities with only one primary school. Robust standard errors in [brackets]. Coefficients marked with (***) are significant under Schwartz (1978) with a t-statistic no smaller than 3.15. All models include child characteristics: sex and age groups; parental characteristics: father's age, mother's age, father and mother education groups; household assets: cement floor, water and electricity, agricultural land owning, refrigerator and stove. Columns (4) and (5) also include school characteristics: teachers' educational categorical variable levels, student-to-teacher ratio, and number of classrooms.

correct for possible endogenous program placement, observed and unobserved community heterogeneity is further controlled for and model (3) is re-estimated using community fixed effects. As expected, column (4) shows that once community fixed effects are controlled for, the difference-in-difference coefficient increases in magnitude but remains not significant, given the significance levels used.

This may be due, however, to the large heterogeneity in school quality of the Mexican educational system (see López, 2001). In the present sample, teachers in bilingual schools (who are fluent in a native language) demonstrate lower levels of education than their non-indigenous colleagues in other primary schools.[25] It may correctly be assumed that community effects would account for level effects (there is one school per community). Nevertheless, these school quality effects may still bias downward the double difference coefficient in the event of nonlinear spurious correlation between school performance of monolingual children and low-quality indigenous schools that are largely attended by monolinguals.

Therefore, the final specifications in Table 6.6 further control for additional school quality variables, including the education level of teachers, the student teacher ratio, and the interaction of these variables with a dummy variable that indicates whether a child is monolingual.[26] The results are as expected. The double difference estimate on the interaction term of a child being monolingual with access to a bilingual school further increases to 6.8 percentage points and becomes significant. Thus, bilingual schools significantly reduce the enrollment gap between monolingual and bilingual children. The estimated reduction is quite large, corresponding to a reduction of over 50 percent in the gap in school attendance, compared with the original estimated IV language barrier of 11 percentage points. Moreover, it is also important to note that in the final specification (column 6), none of the

[25] López (2001) shows that test scores in bilingual primary schools are lower than for any other type of primary school in Mexico.

[26] Teachers are classified into three categories. The first includes teachers with secondary school or less education; the second includes teachers with a high school education; and the third includes teachers with a university education or above, including those who attended a general university or one that specialized in teacher training. Because of the number of very small schools in the sample, a dummy variable specification is used for each group, e.g., schools with at least one teacher in the relevant group.

school quality variables interacted with a child's language is significant. An F-test of the joint significance of all of the school quality interactions (excluding access to a bilingual school) confirms that these interactions are jointly insignificant. Among school characteristics, only the bilingual educational attribute (teachers who speak the native language, textbooks in the indigenous language, etc.) reduces the enrollment gap between monolingual and bilingual children. These results are interpreted as confirming the existence of a large language barrier responsible for schooling inequality among the indigenous.

Conclusions

While it is routinely believed that the indigenous population tends to be among the poorest in terms of income or consumption measures in Mexico, there has thus far been little evidence on the educational attainment of indigenous children. Educational attainment is likely an important determinant of the level of social exclusion that indigenous children may face in the future. As a first step toward providing a diagnostic of the factors affecting educational attainment of indigenous children, this chapter has analyzed how the education level of the indigenous can be improved.

The chapter has shown that indigenous children on average fare worse in educational outcomes than non-indigenous children, even within highly marginalized rural areas of Mexico. Nonetheless, there is great heterogeneity within the indigenous population. When indigenous children learn Spanish, they achieve educational outcomes that are almost equivalent to their non-indigenous counterparts. When they do not learn Spanish, however, their educational outcomes are far inferior. In this way, the level of social exclusion faced by the indigenous is not homogenous. There are important differences between bilingual and monolingual indigenous children.

The analysis helps explain why some monolingual indigenous children perform worse than bilingual indigenous children. Instrumental variable procedures and evidence on schooling outcomes in bilingual educational programs suggest that the language barrier for children who do not speak Spanish is an important factor that greatly reduces their educational achievement. Language barriers can thus be taken as an important factor behind the continued social exclusion of the monolingual indigenous.

Bilingual education is a relatively new educational phenomenon in Mexico. The results suggest potential positive effects of bilingual primary schools in the sense that the negative effect of estimated language barriers is reduced when indigenous children have access to bilingual education; that is, bilingual schools can reduce the degree of social exclusion faced by monolingual indigenous children. Nevertheless, the impacts and results of bilingual schools constitute an important area for evaluation and should be monitored on an ongoing basis. Clearly, it is still too early to speculate on the long-term effects. In the case of the United States, for example, Duignan (2000) argues that there is strong evidence that bilingual schooling reduces the probability that children learn English and reduces assimilation rates. Therefore, one area of concern is the possible impact of bilingual education on indigenous children's learning of Spanish. Positive impacts of bilingual education could be undermined if bilingual education reduces the probability of indigenous children learning Spanish and thus increases their level of social exclusion as adults.

This page intentionally left blank

Bibliography

Acuna González, G., and E. Olivares Ferreto. 1999. Diagnóstico global. La población migrante nicaragüense en Costa Rica: Realidad y respuestas. San José, Costa Rica: Fundación Arias para la Paz y el Progreso Humano.

Albo, X., and M. Preiswerk. 1986. *Cabalgando entre dos mundos: Aymaras migrantes en la ciudad de La Paz*. La Paz: Centro de Investigación y Promoción del Campesinado (CIPCA).

Alesina, A., and E. LaFerrara. 2000. Participation in Heterogeneous Communities. *Quarterly Journal of Economics* 115(3): 847–904.

Alvarez, I. 1998. 21st Century Challenges Facing the Brazilian Health Sector: A Report on the 1998 Roundtable Held in São Paulo, Brazil. San Diego: University of California at San Diego, Institute of the Americas. Available at the Website: http://www.iamericas.org/publications

Alves, D. 2000. Parametric and Semi-Parametric Modeling of Healthcare Expenditure: A Household Data Analysis for the City of São Paulo. Texto para Discussão no. 5/00. IPE-USP, São Paulo.

Barro, R., and J. Lee. 1996. International Measures of Schooling Years and Schooling Quality. *American Economic Review* 86(2): 218–23.

Basmann, R. 1960. On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics. *Journal of the American Statistical Association* 55(292): 650–59.

Bean, F. D., and M. Tienda. 1987. *The Hispanic Population of the United States*. New York: Russell Sage Foundation.

Behrman, J. 2001. Targeting Social Protection to the Poor: What Are the Tradeoffs? University of Pennsylvania, Department of Economics. Mimeo.

Behrman, J., and A. B. Deolalikar. 1991. School Repetition, Dropouts, and the Rates of Return to Schooling: The Case of Indonesia. *Oxford Bulletin of Economics and Statistics* 53(4): 467–80.

Behrman, J., and J. C. Knowles. 1999. Household Income and Child Schooling in Vietnam. *World Bank Economic Review* 13(2): 211–56.

Behrman, J., and P. E. Todd. 1999. Randomness in the Experimental Samples of PROGRESA. International Food Policy Research Institute, Washington, DC. Website: http://www.ifpri.org/themes/progresa.htm

Behrman, J., N. Birdsall, and M. Székeley. 1999. Intergenerational Mobility in Latin America: Deeper Markets and Better Schools Make a Differ-

ence. In N. Birdsall and C. Graham (eds.), *New Markets, New Opportunities? Economic and Social Mobility in a Changing World*. Washington, DC: Brookings Institution/Carnegie Endowment for International Peace.

Behrman, J., P. Sengupta, and P. Todd. 2000. Final Report: The Impact of PROGRESA on Achievement Test Scores in the First Year. International Food Policy Research Institute, Washington, DC. Website: http://www.ifpri.org/themes/progresa.htm

Bell, D. 1974. Residential Location, Economic Performance, and Public Employment. In G. M. Von Furstenberg, B. Harrison, and A. R. Horowits (eds.), *Patterns of Racial Discrimination*. Vol. 1. Lexington: Lexington Books.

Bertrand, M., E. Luttman, and S. Mullainathan. 2000. Network Effects and Welfare Cultures. *Quarterly Journal of Economics* 115(3): 1019–55.

Besley, T., and R. Burgess. 2001. Political Economy of Government Responsiveness: Theory and Evidence from India. London School of Economics. Mimeo.

Blake, J. 1989. *Family Size and Achievement*. Berkeley: University of California Press.

Borjas, G. J. 1995a. Ethnicity, Neighborhoods, and Human-Capital Externalities. *American Economic Review* 85(3): 365–90.

———. 1995b. Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s? *Journal of Labor Economics* 13(2): 201–45.

———. 1997. To Ghetto or Not to Ghetto: Ethnicity and Residential Segregation. *Journal of Urban Economics* 44: 228–53.

Briones, C., and K. Andrade-Eekhoff. 2000. *Participación en los mercados laborales de los residentes en las áreas rurales: Limitaciones y desafíos.* Documento de Investigación BASIS no. 2. Fundación Salvadoreña para el Desarrollo Económico y Social (FUSADES), San Salvador, El Salvador.

Case, A. C. 1992. Neighborhood Influence and Technological Change. *Regional Science and Urban Economics* 22(3): 491–508.

Case, A. C., and L. F. Katz. 1991. *The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths.* NBER Working Paper 3705. National Bureau of Economic Research, Cambridge, MA.

Caso, A., S. Zavala, J. Miranda, et al. 1981. *La política indigenista en Mexico: Métodos y resultados.* Mexico City: Consejo Nacional para la Cultura y las Artes.

Chiswick, B. R. 1991. Speaking, Reading, and Earnings among Low-skilled Immigrants. *Journal of Labor Economics* 9(2): 149–70.

Chiswick, B. R., and P. W. Miller. 1995. The Endogeneity between Language and Earnings: International Analyses. *Journal of Labor Economics* 13(2): 246–88.

Chiswick, B. R., H. A. Patrinos, and M. Hurst. 2000. Indigenous Language Skills and the Labor Market in a Developing Economy: Bolivia. *Economic Development and Cultural Change* 48(2): 349–68.

Church, J., and I. King. 1993. Bilingualism and Network Externalities. *Canadian Journal of Economics* 26(2): 337–45.

Coleman, J., 1990. *Foundations of Social Theory.* Cambridge, MA: Harvard University Press.

Crane J. 1991. The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing. *American Journal of Sociology* 96(5): 1226–59.

Cutler, D. M, and E. Glaeser. 1997. Are Ghettos Good or Bad? *Quarterly Journal of Economics* 112(3): 827–72.

Cutler, D. M. and E. Meara. 1998. The Medical Costs of the Young and Old: A Forty-Year Perspective. In D. A. Wise (ed.), *Frontiers in the Economics of Aging.* Chicago: University of Chicago Press.

Cutler, D. M., E. L. Glaeser, and J. L. Vigdor. 1997. *The Rise and Decline of the American Ghetto.* NBER Working Paper 5881. National Bureau of Economic Research Cambridge, MA.

———. 1999. The Rise and Decline of the American Ghetto. *Journal of Political Economy* 107(3): 455–506.

de Farias, P. C. L. 1998. *Social Security in Brazil: Problems and Trends.* Washington, DC: George Washington University, Institute of Brazilian Issues, Minerva Program.

de Haan, A. 1999. Social Exclusion: Towards a Holistic Understanding of Deprivation. Paper prepared for the *World Development Report 2001.* World Bank, Washington, DC.

Dirección General de Estadística y Censos (DIGESTYC). 1993, 1999, 2000. *Encuesta de hogares de propósitos múltiples.* San Salvador, El Salvador: DIGESTYC.

do Valle Silva, N. 2000. Raza, pobreza y exclusión social en Brasil. In E. Gacitua, C. Sojo, and S. H. Davis (eds.), *Exclusión social y reducción de la pobreza en América Latina y el Caribe*. Washington, DC: World Bank/Facultad Latinoamericana de Ciencias Sociales.

Duignan, P. 2000. Bilingual Education: A Critique. Hoover Institution, Stanford, CA. Mimeo.

Elliot, D. S., W. Julius, D. Huizinga, et al. 1996. The Effects of Neighborhood Disadvantage on Adolescent Development. *Journal of Research in Crime and Delinquency* 33(4): 389–426.

European Foundation for the Improvement of Living and Working Conditions. 1995. *Public Welfare Services and Social Exclusion: The Development of Consumer Oriented Initiatives in the European Union*. Dublin: European Foundation for the Improvement of Living and Working Conditions.

Evans, W. N., W. E. Oates, and R. M. Schwab. 1992. Measuring Peer Group Effects: A Study of Teenage Behavior. *Journal of Political Economy* 100: 966–91.

Figueroa, A. 2000. La exclusión social como una teoría de distribución. In E. Gacitua, C. Sojo, and S. H. Davis (eds.), *Exclusión social y reducción de la pobreza en América Latina y el Caribe*. Washington, DC: World Bank/Facultad Latinoamericana de Ciencias Sociales.

Filmer, D., and L. Pritchett. 1998. Estimating Wealth Effects without Income or Expenditure Data: Educational Enrollment in India. World Bank Development Economics Research Group, Washington, DC. Mimeo.

FLACSO-Costa Rica. 1999. *El paso incesante: Problemática regional, laboral, y social de las migraciones nicaragüenses a Costa Rica*. San José, Costa Rica: United Nations Development Programme.

Fundación Salvadoreña para el Desarrollo Económico y Social (FUSADES). 1999. *Rural Household Survey*. San Salvador, El Salvador: FUSADES.

Funkhouser, E. 1999. Migration and Remittances from Nicaragua. University of California at Santa Barbara. Mimeo.

Gacitua, E., and S. H. Davis. 2000. Pobreza y exclusión social en América Latina. In E. Gacitua, C. Sojo, and S. H. Davis (eds.). *Exclusión social y reducción de la pobreza en América Latina y el Caribe*. Washington, DC: World Bank/Facultad Latinoamericana de Ciencias Sociales.

Gaviria, A., and C. Pagés. 2001. Patterns of Crime Victimization in Latin American Cities. *Journal of Development Economics*.

Glaeser, E. L., M. Kahn and J. Rappaport. 2000. *Why Do the Poor Live in Cities?* NBER Working Paper 7636. National Bureau of Economic Research, Cambridge, MA.

Glaeser, E. L., D. Laibson, and B. Sacerdote. 2000. *The Economic Approach to Social Capital.* NBER Working Paper 7728. National Bureau of Economic Research, Cambridge, MA.

Gore, C., and J. B. Figueiredo. 1997. *Social Exclusion and Anti-Poverty Policy: A Debate.* Geneva: International Labour Organization.

Gray-Molina, G. 2000. The Politics of Popular Participation in Bolivia, 1994–1999. Doctoral dissertation, Oxford University, Nuffield College, United Kingdom.

Gray-Molina, G., E. Pérez de Rada, and E. Yañez. 1999a. *La economía política de las reformas institucionales en Bolivia.* Research Network Working Paper R-350. Inter-American Development Bank Research Department, Washington, DC.

———. 1999b. *Transparency and Accountability in Bolivia: Does Voice Matter?* Research Network Working Paper R-381. Inter-American Development Bank Research Department, Washington, DC.

Gray-Molina, G., W. Jiménez, E. Pérez de Rada, and E. Yañez. 1999. Activos y recursos de los pobres: ¿Qué rol desempeña el capital social? *Trimestre Económico* 66(3/263): 365–418.

Greene, W. H. 2000. *Econometric Analysis.* Upper Saddle River: Prentice Hall.

HAM-La Paz. 1972. Levantamiento y mapa catastral de La Paz. Honorable Alcaldía Municipal de La Paz.

Harmeling, S. 1999. Health Reform in Brazil, Case Study for Module 3: Reproductive Health and Health Sector Reform. World Bank Institute, Washington, DC.

Hauser, R. M., and D. Kuo Hsiang-Hui. 1995. *Does the Gender Composition of Sibships Affect Educational Attainment?* CDE Working Paper 95-106. Center for Demography and Ecology, Madison, WI.

Hausman, J. A. 1978. Specification Tests in Econometrics. *Econometrica* 46(5): 1251–72.

Heer, D. M. 1985. Effects of Sibling Number on Child Outcome. In R. H Turner and J. F. Short, Jr. (eds.), *Annual Review of Sociology* 11: 27–47.

Hernsein, R. J., and C. Murray. 1994. *The Bell Curve.* New York: Free Press.

Holzer, H. 1991. The Spatial Mismatch Hypothesis: What Has the Evidence Shown? *Urban Studies* 28: 105–22.

Ihlanfeldt, K. R., and D. L. Sjoquist. 1990. Job Accessibility and Racial Differences in Youth Employment Rates. *American Economic Review* 80: 267–76.

———. 1998. The Spatial Mismatch Hypothesis: A Review of Recent Studies and their Implications for Welfare Reform. *Housing Policy Debate* 9(4): 849–92.

Ihlanfeldt, K. R., and M. Young. 1996. The Spatial Distribution of Black Employment between the Central City and the Suburbs. *Economic Inquiry* 64: 613–707.

Instituto Brasileiro de Geografia e Estatística (IBGE). 1998. *Pesquisa nacional por amostra de domicilios (PNAD) 1998*. Brasilia: IBGE.

Instituto Geográfico Militar (IGM). 1998. Cartografía para la ciudad de La Paz. IGM, La Paz.

Instituto Nacional de Estadística (INE). 2000a. *Proyecciones poblacionales 1990–2000*. La Paz: INE.

———. 2000b. *MECOVI-Encuesta de Hogares 1999*. La Paz, Bolivia: INE.

Instituto Nacional de Estadística Geografía e Informática de Mexico (INEGI). 1994. *Estadísticas históricas de México*. México City: INEGI.

———. 2000. Información preliminar, Censo Nacional de Población 2000. INEGI, México City.

Inter-American Development. 2000. *Development beyond Economics. Economic and Social Progress Report 2000*. Washington, DC: Inter-American Development Bank.

Jencks, C., and S. E. Meyer. 1990. The Social Consequences of Growing Up in a Poor Neighborhood: A Review. In L. E. Lynn and M. G. H. McGeary (eds.), *Inner City Poverty in the United States*. Washington, DC: National Academy Press.

Jiménez, W., and E. Yañez, 1997. Ingresos familiares en un contexto de crecimiento: Un modelo de hogares en pseudo-panel. Unidad de Análisis de Políticas Sociales (UDAPSO), La Paz.

John, A., and K. Yi. 1997. Language, Learning and Location. University of Virginia, Charlottesville, VA. Mimeo.

Kain, J. E. 1968. Housing Segregation, Negro Employment and Metropolitan Decentralization. *Quarterly Journal of Economics* 82: 175–97.

———. 1992. The Spatial Mismatch Hypothesis: Three Decades Later. *Housing Policy Debate* 3(2): 371–460.

Killingsworth, M. R. 1983. *Labor Supply*. Cambridge, UK: Cambridge University Press.

King, E., and L. A. Lillard. 1987. Education Policy and Schooling Attainment in Malaysia and the Philippines. *Economics of Education Review* 6(2): 167–81.

Kremer, M. 1997. How Much Does Sorting Increase Inequality? *Quarterly Journal of Economics* 112(1): 115–39.

Lang, K. 1986. A Language Theory of Discrimination. *Quarterly Journal of Economics* 101(2): 363–82.

Lardé de Palomo, A., and A. Argüello de Morera. 2000. *Integración a los mercados de los hogares rurales y generación de ingresos.* Documento de Investigación BASIS no. 3. Fundación Salvadoreña para el Desarrollo Económico y Social (FUSADES), San Salvador, El Salvador.

Latinobarómetro. 2000. *Latinobarómetro 1999–2000*. Santiago, Chile: Corporación Latinobarómetro.

Lazear, E. P. 1999. Culture and Language. *Journal of Political Economy* 107(6): 95–126.

Lizano Ortiz, N., and G. Maklouf Weiss. 1999. Marco jurídico de los extranjeros en Costa Rica. Paper presented at the conference "Hacia una política Migratoria Integral," Instituto Centroamericano de Estudios Sociales, Coronado, Costa Rica.

Lombardo Toledano, V. 1991. *Escritos acerca de la situación de los indigenas.* Mexico, City: Centro de Estudios Filosóficos, Politicos y Sociales Vicente Lombardo Toledano.

Long, J. S. 1997. *Regression Models for Categorical and Limited Dependent Variables.* London: Sage Publications.

López Acevedo, G. 1999. *Learning Outcomes and School Cost-Effectiveness in Mexico.* Policy Research Working Paper 2128, World Bank, Washington, DC.

———. 2001. *Professional Development and Incentives for Teacher Performance in Schools in Mexico.* LCSHD Paper Series 64. World Bank, Washington, DC.

Loury, G. 1999. Social Exclusion and Ethnic Groups: The Challenge to Economics. Paper presented at the Annual World Bank Conference on Development Economics, Washington, DC.

Madden, J. 1980. Why Women Work Closer to Home. *Urban Studies* 18: 181–94.

Madden, J., and M. J. White. 1980. Spatial Implications of Increases in the Female Labor Force: A Theoretical and Empirical Synthesis. *Lan Economics* 56: 432–46.

Manser, M., and M. Brown. 1980. Marriage and Household Decision-Making: A Bargaining Analysis. *International Economic Review* 21(1): 31–44.

Manski, C. F. 1993. Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies* 60: 531–42.

———. 1995. *Identification Problems in the Social Sciences.* Cambridge, MA and London: Harvard University Press.

Massey, D. S., and N. A. Denton. 1989. Hypersegregation in United States Metropolitan Areas: Black and Hispanic Segregation along Five Dimensions. *Demography* 26: 373–91.

McKinney, S., and A. B. Schnare. 1989. Trends in Residential Segregation by Race: 1960–1980. *Journal of Urban Economics* 26: 269–80.

McLafferty, S., and V. Preston. 1992. Spatial Mismatch and Labor Market Segmentation for African-American and Latina Women. *Economic Geography* 68: 406–31.

Mincer, J. 1974. *Schooling, Experience and Earnings.* New York: Columbia University Press.

Ministerio de Planificación. 1992. *Encuesta de hogares de propósitos múltiples.* San Salvador, El Salvador: Ministerio de Planificación

Modiano, N. 1988. Public Bilingual Education in Mexico. In C. B. Paulston (ed.), *Handbook of Bilingualism and Bilingual Education.* New York: Greenwood Press.

Mora Corrales, M., and S. C. Mora. 1998. El impacto de la migración nicaragüense en el sistema educativo de Costa Rica. Paper presented at the International Seminar on Population at the End of the Millenium, Jaco, Costa Rica.

Morales, A. 1999. *Amnistia migratoria en Costa Rica.* San José: FLACSO-Costa Rica.

Morales, A., and C. Castro. 1999. *Inmigracion laboral nicaragüense en Costa Rica.* San José: FLACSO-Costa Rica.

National School of Public Health. 1998. *Health Care Around the World: Brazil.* Rio de Janeiro: Ministry of Health, Oswaldo Cruz Foundation.

O'Faircheallaigh, C. 1998. Resource Development and Inequality in Indigenous Societies. *World Development* 26(3): 381–94.

Ordoñez, J. 2000. Los derechos fundamentales como referente del paradigma de ciudadanía civil y de la definición de la frontera de exclusión social. In E. Gacitua, C. Sojo and S. H. Davis (eds.), *Exclusión social y reducción de la pobreza en América Latina y el Caribe*. Washington, DC: World Bank/Facultad Latinoamericana de Ciencias Sociales.

Palafox, J. C., J. Prawda, and E. Velez. 1994. Primary School Quality in Mexico. *Comparative Education Review* 38(2): 167–80.

Panagides, A. 1994. Mexico. In G. Psacharopoulos and H. A. Patrinos (eds.), *Indigenous People and Poverty in Latin America: An Empirical Analysis*. Washington, DC: World Bank.

Parker, S. W., and R. Wong. 1997. Household Income and Health Care Expenditures in Mexico. *Health Policy* 40: 237–55.

Pérez de Rada, E., 1997. Discriminación salarial por género y etnia en ciudades principales de Bolivia. La Paz: Unidad de Análisis de Políticas Sociales (UDAPSO).

Pool, J. 1991. The Official Language Problem. *American Political Science Review* 85(2): 495–514.

Portes, A. 1987. The Social Origins of the Cuban Enclave Economy in Miami. *Sociological Perspectives* 30(4): 340–72.

———. 1998. Social Capital: Its Origins and Applications in Modern Sociology. *Annual Review of Sociology* 24: 1–24.

Portes, A., and L. Jensen. 1989. The Enclave and Entrants: Patterns of Ethnic Enterprise in Miami before and after Mariel. *American Sociological Review* 54: 929–49.

Portes, A., and P. Landolt. 2000. Social Capital: Promise and Pitfalls of its Role in Development. *Journal of Latin American Studies* 32(2): 529–47.

Pritchard, D. 1999. Nicaragua: Uso productivo de las remesas familiares. Document LC/MEX/L.414, United Nations, Economic Commission for Latin America, Mexico City.

Putnam, R. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.

———. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster.

Reid, C. E. 1985. The Effect of Residential Location on the Wages of Black Women and White Women. *Journal of Urban Economics* 18: 350–63.

Rivero, R., and W. Jiménez, 1999. *Diferencias salariales en el mercado de trabajo urbano en Bolivia, 1981–1997.* La Paz: Unidad de Análisis de Políticas Económicas (UDAPE).

Robichaux, D. 1994. Clase, percepción étnica y transformación regional: Unos ejemplos tlaxcaltecas. *Boletín de Antropología Americana* 30.

Rosenthal, A. S., K. Baker, and A. Ginsburg. 1983. The Effect of Language Background on Achievement Level and Learning Among Elementary School Students. *Sociology of Education* 56: 157–69.

Rosenzweig, M. R., and K. I. Wolpin. 2000. Natural "Natural Experiments" in Economics. *Journal of Economic Literature* 38: 827–74.

———. 1986. Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions. *American Economic Review* 76(3): 70–82.

Sánchez, T.W. 1999. The Connection between Public Transit and Employment: The Cases of Portland and Atlanta. *Journal of the American Planning Association* 65(3): 284–96.

Sandoval, I., M. Carrillo, O. Calderon, et al. 1999. *La poblacion costarricense ante la crisis, la migración, y las instituciones públicas.* San José, Costa Rica: Universidad Nacional, Instituto de Estudios de Poblacion.

Sawicki, D. S. 2000. Developing Transportation Alternatives for Welfare Recipients Moving to Work. *Journal of the American Planning Association* 66(3): 306–18.

Schultz, T. P. 2000. School Subsidies for the Poor: Evaluating a Mexican Strategy for Reducing Poverty. Yale University, New Haven, CT. Mimeo.

Schwarz, G. 1978. Estimating the Dimension of a Model. *Annals of Statistics* 6: 461–64.

Scott, J. 2000. Who Benefits from the State in High-Inequality, Middle-Income Countries? The Case of Mexico. Centro de Investigación y Docencias Económicas, Mexico City. Mimeo.

Secretaria de Educación Pública. 2000. *Perfil de la Educación en México.* México City: Secretaria de Educación Pública.

Sojo, C. 2000. Dinámica socio política y cultural de la exclusión social. In E. Gacitua, C. Sojo and S. H. Davis (eds.), *Exclusión social y reducción de la pobreza en América Latina y el Caribe.* Washington, DC: World Bank/Facultad Latinoamericana de Ciencias Sociales.

Terborgh, A., J. E. Rosen, R. Santiso Galvez, et al. 1995. Family Planning among Indigenous Populations in Latin America. *International Family Planning Perspectives* 21(4): 143–49.

Thompson, M. A. 1997. The Impact of Spatial Mismatch on Female Labor Force Participation. *Economic Development Quarterly* 11(2): 138–45.

Townsend, P. 1993. *The International Analysis of Poverty*. New York: Harvester Wheatsheaf.

Trouillot, M. R. 2000. Exclusión social en el Caribe. In E. Gacitua, C. Sojo and S. H. Davis (eds.), *Exclusión social y reducción de la pobreza en América Latina y el Caribe*. Washington, DC: World Bank/Facultad Latinoamericana de Ciencias Sociales.

Unidad de Análisis de Políticas Económicas (UDAPE). 2000. Estrategia boliviana de reducción de la pobreza. UDAPE, La Paz.

Vos, R., H. Lee, and J. A. Mejía. 1998. *Structural Adjustment and Poverty in Bolivia.* INDES Working Paper I-3. Inter-American Development Bank Social Development Institute, Washington, DC.

Vroman, J., and S. Greenfield. 1980. Are Blacks Making it in the Suburbs? Some New Evidence on Intrametropolitan Spatial Segmentation. *Journal of Urban Economics* 7: 155–67.

Wacquant, L. J. D., and W. J. Wilson. 1989. The Cost of Racial and Class Exclusion in the Inner City. *Annals of the American Academy of Political and Social Science* 501: 8–25.

Wilson, W. J. 1992. The Plight of the Inner-City Black Male. *Proceedings of the American Philosophical Society* 136(3): 320–25.

Wong, R. 2000. Health, Utilization of Health Care, and Aging in Mexico. Georgetown University Department of Demography, Washington, DC. Mimeo.

World Bank. 1994. *Averting the Old Age Crisis: Policies to Protect the Old and Promote Growth.* Oxford, UK: Oxford University Press.

———. 1998. The Brazilian Health System. Impact Evaluation Report. World Bank Operations Evaluation Department, Washington, DC.

———. 2000. *Entering the 21st Century: World Bank Development Report 1999/2000.* Oxford, UK: Oxford University Press.

ECONOMIC DEVELOPMENT

Social exclusion is closely linked with many economic problems in Latin America, yet seldom does it take the form of a "keep out" sign. More commonly, groups are excluded because they lack access to opportunities enjoyed by others in health care, education, housing and employment. These barriers prevent people from reaching their full productive potential—in turn constraining growth and revenues—and make them more likely to incur public health and social service costs.

Who's In and Who's Out explores various forms of social exclusion, including residential segregation in Bolivian cities, exclusion from health care in Brazil, barriers to legal status of Nicaraguan immigrants in Costa Rica, geographic isolation in El Salvador, and educational inequality among the indigenous in Mexico.

The book describes how the exclusion process is exacerbated by self-perpetuating networks of association, prohibitive prices for certain services, and misperceptions between the societal mainstream and excluded groups. In identifying the causes, mechanisms and effects of these types of social exclusion, this study marks a critical first step towards formulating policies in the region that will enable the greatest number of people to lead productive lives and access all the benefits of society.

COVER DESIGN: LEILANY GARRON

Inter-American Development Bank
1300 New York Ave., N.W.
Washington, D.C. 20577

www.iadb.org/pub

Distributed by
The Johns Hopkins University Press
Baltimore and London

ISBN 1-931003-42-4