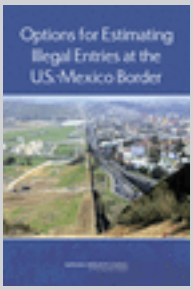
This PDF is available from The National Academies Press at http://www.nap.edu/catalog.php?record_id=13498



Estimating Illegal Entries at the U.S.-Mexico Border

ISBN
978-0-309-26422-8

143 pages
6 x 9
PAPERBACK (2013)

Alicia Carriquiry and Malay Majmundar, Editors; Panel on Survey Options for Estimating the Flow of Unauthorized Crossings at the U.S.-Mexico Border; Committee on National Statistics; Division of Behavioral and Social Sciences and Education; National Research Council

Add book to cart | Find similar titles | Share this PDF

Visit the National Academies Press online and register for...

- Instant access to free PDF downloads of titles from the
  - NATIONAL ACADEMY OF SCIENCES
  - NATIONAL ACADEMY OF ENGINEERING
  - INSTITUTE OF MEDICINE
  - NATIONAL RESEARCH COUNCIL
- 10% off print titles
- Custom notification of new releases in your field of interest
- Special offers and discounts

Distribution, posting, or copying of this PDF is strictly prohibited without written permission of the National Academies Press.
Unless otherwise indicated, all materials in this PDF are copyrighted by the National Academy of Sciences.
Request reprint permission for this book

Copyright © National Academy of Sciences. All rights reserved.

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

Options for Estimating Illegal Entries at the U.S.–Mexico Border

Panel on Survey Options for Estimating the Flow of Unauthorized Crossings at the U.S.–Mexican Border

Alicia Carriquiry and Malay Majmundar, *Editors*

Committee on National Statistics

Division of Behavioral and Social Sciences and Education

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

Copyright © National Academy of Sciences. All rights reserved.

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

Support of the work of the Committee on National Statistics is provided by a consortium of federal agencies through a grant from the National Science Foundation (No. SES-1024012). The project that is the subject of this report was supported by an allocation from the U.S. Department of Homeland Security to the National Science Foundation under this grant. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number-13: 978-0-309-26422-8
International Standard Book Number-10: 0-309-26422-7

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Copyright 2013 by the National Academy of Sciences. All rights reserved.

Printed in the United States of America

Suggested citation: National Research Council. (2013). *Options for Estimating Illegal Entries at the U.S.–Mexico Border*. Panel on Survey Options for Estimating the Flow of Unauthorized Crossings at the U.S.–Mexican Border, A. Carriquiry and M. Majmundar, Eds. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

Copyright © National Academy of Sciences. All rights reserved.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

PANEL ON SURVEY OPTIONS FOR ESTIMATING THE FLOW OF UNAUTHORIZED CROSSINGS AT THE U.S.–MEXICAN BORDER

ALICIA CARRIQUIRY (*Chair*), Department of Statistics, Iowa State University
DAVID L. BANKS, Department of Statistical Sciences, Duke University
PETER BROWNELL, RAND Corporation, Santa Monica, CA
STEPHEN E. FIENBERG, Department of Statistics, Carnegie Mellon University
MARK S. HANDCOCK, Department of Statistics, University of California, Los Angeles
GORDON HANSON, Department of Economics, University of California, San Diego
VIRGINIA LESSER, Department of Statistics, Oregon State University
PIA ORRENIUS, Federal Reserve Bank of Dallas
JEFFREY S. PASSEL, Pew Hispanic Center, Washington, DC
FERNANDO RIOSMENA, Institute of Behavioral Science and Geography Department, University of Colorado
SILVIA ELENA GIORGULI SAUCEDO, Center for Demographic, Urban, and Environmental Studies, El Colegio de Mexico

MALAY MAJMUNDAR, *Study Director*
THOMAS J. PLEWES, *Senior Program Officer*
MICHAEL J. SIRI, *Program Associate*

Copyright © National Academy of Sciences. All rights reserved.

COMMITTEE ON NATIONAL STATISTICS
2012-2013

LAWRENCE BROWN (*Chair*), Department of Statistics, The Wharton School, University of Pennsylvania
JOHN M. ABOWD, School of Industrial and Labor Relations, Cornell University
DAVID CARD, Department of Economics, University of California, Berkeley
ALICIA CARRIQUIRY, Department of Statistics, Iowa State University
CONSTANTINE GATSONIS, Center for Statistical Sciences, Brown University
JAMES S. HOUSE, Survey Research Center, Institute for Social Research, University of Michigan
MICHAEL HOUT, Survey Research Center, University of California, Berkeley
SALLIE KELLER, University of Waterloo, Ontario, Canada
LISA LYNCH, Heller School for Social Policy and Management, Brandeis University
SALLY MORTON, Graduate School of Public Health, University of Pittsburgh
RUTH PETERSON, Criminal Justice Research Center, The Ohio State University
EDWARD H. SHORTLIFFE, Columbia University and Arizona State University
HAL STERN, Donald Bren School of Information and Computer Sciences, University of California, Irvine
JOHN THOMPSON, National Opinion Research Center at the University of Chicago
ROGER TOURANGEAU, Statistical Group, Westat, Rockville, MD

CONSTANCE F. CITRO, *Director*

Copyright © National Academy of Sciences. All rights reserved.

Contents

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

Acknowledgments

In 2011, at the request of the U.S. Department of Homeland Security (DHS), the Committee on National Statistics at the National Research Council (NRC) appointed the Panel on Survey Options for Estimating the Flow of Unauthorized Crossings at the U.S.–Mexican Border to provide guidance on the use of surveys and other methodologies to estimate the number of unauthorized crossings at the U.S.–Mexico border. This report represents the final product of the panel.

The panel held its first in-person meeting (in conjunction with a public workshop) in November 2011. It held a second in-person meeting in January 2012. Several panel members participated in a field trip to the Tucson and San Diego sectors of the U.S.–Mexico border prior to the January meeting and had an opportunity to meet with U.S. Border Patrol (USBP) agents, observe the processing of apprehended migrants in the Tucson sector, tour the Nogales and the Chula Vista sections of the border fence, and see the remote sensing facility in the San Diego sector. These panel members were encouraged to ask questions about all aspects of border enforcement and received valuable information from USBP agents. The panel held its final in-person meeting in March 2012. In May 2012, several panel members traveled to Tijuana, Mexico, where they were hosted by researchers at El Colegio de la Frontera Norte (COLEF) and learned about data collection for the Survey of Migration at the Northern Border (EMIF-N).

This report would not have been possible without the contributions of many people. Special thanks go to the members of the panel, who dedicated time, thought, and energy to the report. The panel worked very well together and, due to its professional diversity, was well-positioned to address

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

the complex problem of illegal migration at the southwestern border of the United States. Special thanks also go to Michael Hoefer, Director of the Office of Immigration Statistics at DHS, who played a key role in developing this study and who served as the Department's liaison with the panel during the course of its work.

The panel learned much from discussions on surveys and other data initiatives at the November 2011 workshop, where presentations were made by Mark Borkowski, U.S. Customs and Border Protection, DHS; Alfredo Bustos, Instituto Nacional de Estadística y Geografía (INEGI); Wayne Cornelius, University of California, San Diego; Elizabeth Grieco, U.S. Census Bureau; Douglas Massey, Princeton University; Elsa Pérez Paredes, INEGI; Jeffrey Passel, Pew Hispanic Center; Michael Rendall, University of Maryland; Melissa Scopilliti, U.S. Census Bureau; and Duncan Thomas, Duke University. The work of the panel was also informed by the trips made by several members to the southwest U.S. border and to Tijuana, Mexico—arranged, respectively, by Luke Lopez of U.S. Customs and Border Protection and by Marie-Laure Coubes and Rene Zenteno from COLEF. Finally, the panel is grateful to Michael Hoefer for providing useful information about the general structure and content of DHS administrative data[1] and for his assistance in advancing the panel's data request within DHS.

Several members of the staff of the NRC made significant contributions to the report. Malay Majmundar served as study director for the panel and was instrumental to the success of the panel's work. He made sure that all the i's were dotted and t's were crossed in the report's substantive and technical discussions, kept the panel engaged in the project and abreast of developments, and wrote significant portions of initial drafts of the report. Michael Siri provided key administrative support to the panel and efficiently organized meetings and field trips. Thanks are also due to Kirsten Sampson Snyder for helping guide the report through review, Robert Katt for skillful editing, and Yvonne Wise for managing the production process. Tom Plewes provided valuable guidance and oversight during the course of the study, and his many years of experience and knowledge of the NRC study process are gratefully acknowledged. Connie Citro was helpful as usual and provided many valuable comments and suggestions.

This final report has been reviewed in draft form by individuals chosen

[1]During a briefing with DHS immediately prior to the public release of this report, the panel received additional clarifying information regarding the structure and content of DHS administrative data. Specifically, it learned that although the data are not integrated across the constituent agencies of DHS for "analytical purposes," they are so for "enforcement" purposes. Although this did not change the panel's conclusions or the thrust of its recommendations, text was added to the report in the Summary, Chapter 5, and Recommendation 5.1 to reflect this distinction.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

for their diverse perspectives and technical expertise, in accordance with procedures approved by the Report Review Committee of the NRC. The purpose of this independent review is to provide candid and critical comments that assist the institution in making its reports as sound as possible and to ensure that the reports meet institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

The panel thanks the following individuals for their review of the report: Frank D. Bean, Center for Research on Immigration, Population and Public Policy, University of California, Irvine; Víctor Alfredo Bustos y de la Tijera, Office of the Deputy Director General, National Institute of Statistics and Geography, Aguascalientes City, Mexico; Rebecca L. Clark, Demographic and Behavioral Sciences Branch, National Institute of Child Health and Human Development; Krista J. Gile, Department of Mathematics and Statistics, University of Massachusetts; Joel L. Horowitz, Department of Economics, Northwestern University; Roderick J. Little, Department of Biostatistics, School of Public Health, University of Michigan; Douglas S. Massey, Department of Sociology, Princeton University; Hal S. Stern, Donald Bren School of Information and Computer Sciences, University of California, Irvine; and Victoria Velkoff, Assistant Division Chief, Population Estimates and Projections, Population Division, U.S. Census Bureau.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations, nor did they see the final draft of the report before its release. The review of the report was overseen by John Rolph, University of Southern California, and Charles Manski, Northwestern University. Appointed by the NRC, they were responsible for making certain that the independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of the report rests entirely with the authoring committee and the National Research Council.

Alicia Carriquiry, *Chair*
Panel on Survey Options for Estimating the Flow of
Unauthorized Crossings at the U.S.–Mexican Border

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

Summary

The U.S. Department of Homeland Security (DHS) is responsible for securing and managing the nation's borders. Over the past decade, DHS has dramatically stepped up its enforcement efforts at the U.S.–Mexico border, increasing the number of U.S. Border Patrol (USBP) agents, expanding the deployment of technological assets, and implementing a variety of "consequence programs" intended to deter illegal immigration. During this same period, there has also been a sharp decline in the number of unauthorized migrants apprehended at the border.

Trends in total apprehensions do not, however, by themselves speak to the effectiveness of DHS's investments in immigration enforcement. In particular, to evaluate whether heightened enforcement efforts have contributed to reducing the flow of undocumented migrants, it is critical to estimate the number of border-crossing attempts during the same period for which apprehensions data are available. With these issues in mind, DHS charged the National Research Council (NRC) with providing guidance on the use of surveys and other methodologies to estimate the number of unauthorized crossings at the U.S.–Mexico border, preferably by geographic region and on a quarterly basis. The NRC appointed the Panel on Survey Options for Estimating the Flow of Unauthorized Crossings at the U.S.–Mexican Border to carry out this task. A better understanding of the magnitude, timing, and location of these flows will help DHS to better evaluate the effectiveness of its enforcement efforts, more efficiently allocate its resources along the border, and provide a more complete report to the public on the state of illegal immigration. This study focuses on Mexican migrants since Mexican

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

nationals account for the vast majority (around 90 percent) of attempted unauthorized crossings across the U.S.–Mexico border.

A discussion of the use of surveys and statistical methods to measure unauthorized border flows needs to be set in the context of border enforcement and the migration process, both current and past. The migration process is highly complex, and it is influenced by a variety of factors such as the economic, social, and environmental conditions in sending and destination areas; immigration policies; and interior and border enforcement. Furthermore, migrants and their smugglers have adapted, and can continue to adapt, to changes in resources and strategies on the U.S. side of the border, and enforcement efforts in one geographic area can have spillover effects into another. This situation argues for a broader conception of the border, not segmented into ports of entry and areas between the ports of entry. The migration process is also dynamic and evolving, and survey designs and modeling approaches that may be well suited for capturing certain aspects of unauthorized migration flows today may not be able to do so with the same reliability in the future. Thus, flexibility of design and continuous evaluation of how DHS is implementing border metrics will be of the utmost relevance.

Within this context, one can assess the usefulness of different types of data to capture information on migration flows. There are a number of major surveys in the United States and Mexico that collect some information about migration and border crossing. On the U.S. side, the American Community Survey (ACS) and Current Population Survey (CPS) each target U.S. households. On the Mexican side, the "long questionnaire" of the Mexican Census of Housing and Population (administered to a 10 percent sample of the population in Mexico), the National Survey of Occupation and Employment (ENOE), the National Survey of Population Dynamics (ENADID), and the longitudinal Mexican Family Life Survey (MxFLS) target households at a national level. None of these surveys was specifically designed to study migration. The Mexican Migration Project (MMP) and Mexican Migration Field Research Program (MMFRP), in contrast, focus on studying migration, although the surveys are not based on probability sampling. The Survey of Migration at the Northern Border (EMIF-N), which has a probability sample conceptual basis, targets migrants passing through northern border cities of Mexico.

The panel began by enumerating the major attributes for evaluating surveys for this purpose: the nature of the target population and related issues of sample size and survey design, the frequency with which surveys are conducted and the speed with which data are made publicly available, and the types of questions that are asked about migration. Since international migration is a relatively rare event, it is important for a general survey to have a sufficiently large sample in order to obtain reliable information on

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

migration. In particular, the sample size requirements for traditional national household surveys would be considerable, especially if one wanted precise flow estimates by the nine geographic sectors into which USBP divides the southwest border. The ever-changing nature of migration flows would also create design challenges for geographically specific estimates. Although sample sizes can be smaller for probability surveys that focus on migrants (such as EMIF-N), changing patterns of migration would still create design and analysis challenges. Moreover, although EMIF-N shows great potential for DHS purposes, its adaptive design makes assessing its accuracy problematic.

Estimating annual flows in a timely fashion using survey data is a great challenge, and doing so on a quarterly and border sector/subregion basis is an even greater challenge. For border flow estimates to have practical value, survey data would need to be collected, analyzed, and released in a timely fashion. Most of the existing surveys do not meet these criteria, and only one—ENOE—is currently capable of providing quarterly estimates on a timely basis (although EMIF-N might also have the potential to provide such timely data). There are a number of key bits of information on migration and border crossing that ENOE does not collect (e.g., documentation status at crossing and border crossing location); questions on those items could, in principle, be added to the ENOE survey instrument. Given the dynamic nature of the migration process, however, questions about the migration process that are salient today may be less so in the near future. Thus, the questions added would have to be limited.

Such improvements could prove useful to researchers and others, and they would be welcome by the panel (as would, for example, improvements in the timeliness of EMIF-N). But from the perspective of estimating flows on an annual or quarterly basis, such modifications would still take place against the backdrop of larger limitations and complexities relating to sample size and survey design. The report also notes the administrative and implementation challenges arising from the fact that ENOE falls under the jurisdiction of the government of Mexico. These challenges would be no less salient if the U.S. government attempted to put in place a new dedicated migration survey in Mexico. Although surveys, especially those that focus specifically on migration, can provide a wealth of information about the migration process, they are not sufficient by themselves in meeting the needs of DHS for obtaining estimates of unauthorized migration flows across the U.S.–Mexico border on an annual or quarterly basis. The panel believes that, although DHS could benefit from engaging with entities in Mexico that collect survey data relevant to the analysis of unauthorized migration, it should not invest substantial resources in making changes to existing surveys or in implementing a new survey for the purpose that is the subject of this study.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Apprehensions data collected by USBP in its enforcement database are also insufficient by themselves to estimate flows because those data do not contain information on unauthorized migrants who elude capture or on migrants who are deterred from trying to cross again after they were apprehended once or more. Data on the re-apprehensions of individuals do provide information that could, in theory, be used to make inferences about the size of the unauthorized population entering the United States successfully. The accuracy of those inferences, however, depends on modeling assumptions that are often difficult to validate.

Despite the limitations of DHS administrative data, their combination and integration with survey data could produce useful insights about migrant flows and the effectiveness of border enforcement. The panel was charged with reviewing administrative data collected by DHS, and in order to better inform our analysis and discussion, we formally requested that DHS provide us with apprehensions data from the USBP enforcement database. The panel made its request with the understanding that any data given to us would need to be made publicly available, in accordance with the institutional rules governing NRC studies. We indicated that DHS could provide it to us in a format that would protect any information that it deemed operationally sensitive. However, DHS would not provide these data without an exemption from public disclosure requirements. It was the judgment of the panel that the quality of its published analysis and the timeliness of its deliberations would have been unduly impaired by the classification restrictions that would have accompanied such an exemption. Therefore, the panel did not pursue its request, and DHS did not provide the panel with access to its administrative data.

The panel had sufficient information about the content, populations covered, and other features of DHS administrative data to support its conclusions about the possible contributions of the data to estimating unauthorized crossings at the U.S.–Mexico border. However, the panel believes that researcher access to DHS data would improve the quality of research in this area and contribute to a better understanding of activity at the border. Researcher access would also permit independent scholars to evaluate the quality, completeness, and reliability of the DHS data at the micro level, which has never been done and which would permit a much richer understanding of the role that apprehensions data could play in informing model-based approaches to estimating flows. Moreover, knowledge of and experience with the use of model-based approaches for estimating flows have been limited to date, and the complexities and uncertainties are considerable. To develop, apply, and continually refine specific modeling approaches, DHS will need to engage with the broader scientific community in a sustained and long-term fashion. This will be possible only if the data in its enforcement databases are made widely available.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

DHS might be reluctant to release apprehension and other administrative data to the public on the grounds that they are sensitive from a law enforcement point of view. The panel believes that most of these concerns can be addressed by implementing masking methods for problematic fields in the records, creating broad geographic identifiers that link them to USBP sectors rather than individual USBP stations, and releasing data with sufficient delay (e.g., a full year) to diminish their sensitivity for operational use and deployment.

Given the importance of looking at the border as an integrated whole, the panel also believes that DHS should integrate apprehensions data for analytical purposes from the Office of Field Operations (OFO), which is responsible for enforcement at official ports of entry; from USBP, which is responsible for enforcement between ports of entry; and from Immigrations and Customs Enforcement (ICE), which is responsible for immigration violations in the interior of the country, away from the physical border itself. Conversations with representatives from DHS suggest that that the linkages between the apprehensions records controlled by USBP, OFO, and ICE in the ENFORCE database are limited to uses that relate specifically to enforcement; linkages across the data sources for broader analytical purposes would require approval from each of the three agencies, and the full database has not been widely used for analysis. Sharing of data across the different components of DHS for analytical purposes would seem a reasonable starting point.

Given the gaps and limitations in survey and administrative data, the panel believes that a necessary approach to estimating the flow of undocumented migrants consists in developing models that can combine survey, administrative, and other types of disaggregated data. These modeling approaches could include conventional statistical regression and other models and incorporate spatiotemporal aspects of the data, but they might also include less-standard simulation-based approaches such as agent-based models.

Any modeling approach and the assumptions underlying it will need to be continually validated against historical trends and data. Although each of these approaches has its limitations, much could also be learned by comparing estimates from these multiple methods. However, without access to DHS data, the panel is unable to provide precise guidance on the modeling approaches that would be most useful for estimating migration flows. Furthermore, although the panel was aware that DHS has been considering specific modeling approaches (e.g., capture-recapture methods using apprehension data), it was not granted access to the relevant technical reports. Moving forward, DHS would greatly benefit from making the administrative data in its enforcement databases (which could be subject

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

to various procedures to protect potentially sensitive information) publicly available to the research community.

The major conclusions and recommendations offered by the panel (and numbered according to the chapter in which they are developed) are as follows:

- **Conclusion 2.1: To understand migration flows in any one sector, it is important to view the entire border as a system; localized increases in border enforcement may simply change where migrants cross without reducing the overall flow in the long run.**

- **Conclusion 2.2: The migration process is complex and dynamic. Undocumented migration is the outcome of many interrelated factors that can vary widely across people, space, and time; migrant characteristics and the geography of sending and destination areas are changing constantly. This complexity and dynamism should be incorporated into the analytical approaches and study designs used to estimate flows of unauthorized migrants.**

- **Recommendation 4.1: For the purpose of estimating unauthorized migration flows across the U.S.–Mexico border on an annual or quarterly basis, DHS should not invest substantial resources in making major changes to existing surveys or in implementing a new survey.**

- **Conclusion 4.1: Existing surveys are subject to a variety of limitations having to do with target populations and associated issues of sample size and survey design, the frequency with which surveys are conducted and the speed with which data are made publicly available, and the types of questions that are asked about migration. Therefore, although survey data are critical for understanding patterns and general trends in unauthorized migration, they will not be sufficient by themselves to meet the needs of DHS for estimating unauthorized migration flows across the U.S.–Mexico border.**

- **Conclusion 4.2: Implementing a new household survey that meets the needs of DHS would require an investment at least comparable to that associated with the American Community Survey in the United States; any such survey would also have to fall within the purview of the Government of Mexico. A survey that uses a time-location design and focuses directly on migrant populations (e.g., EMIF-N) would be more promising, but such a non-traditional**

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

design would necessitate careful adherence to the sampling protocol and, in particular, would require that concerns about coverage error be addressed. Mexican-side implementation would also be an issue. Substantial modifications of existing general household or specialty migration surveys to meet the needs of DHS would encounter similar challenges. These challenges are only magnified by the complex and dynamic nature of the underlying migration process.

- **Recommendation 5.1: DHS should integrate apprehensions data from USBP, OFO, and ICE for analytical purposes.**

- **Conclusion 5.1: Administrative data from DHS are alone insufficient to estimate the flow of unauthorized migrants across the U.S.–Mexico border. However, they could be combined with survey data to produce useful insights about migrant flows and the effectiveness of border enforcement. The use of modeling approaches in conjunction with disaggregated survey and administrative data is necessary for estimating these flows.**

- **Recommendation 5-2: DHS should sponsor and conduct research on modeling approaches for estimating the flows of unauthorized migrants across the U.S.–Mexico border.**

- **Conclusion 5.2: DHS would greatly benefit from making the administrative data from its immigration enforcement databases publicly available for research use, as that would allow DHS to engage with the broader scientific community to develop, apply, and continually refine specific modeling approaches. DHS could develop ways of constructing masked and/or aggregate files for public release in order to protect sensitive information.**

- **Conclusion 6.1: Modeling approaches, and the assumptions underlying them, must keep track of changing mechanisms of migration and be continually validated against historical trends and data. Since all modeling approaches have their limitations, there is also much that could be learned by comparing estimates from multiple methods.**

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

1

Introduction

The U.S. Department of Homeland Security (DHS) is charged with securing and managing the nation's borders. The Office of Field Operations within Customs and Border Protection is responsible for enforcing immigration laws at official ports of entry, while the U.S. Border Patrol (USBP)—which is also part of Customs and Border Protection—enforces immigration laws between ports of entry. Immigration violations in the interior of the country, away from the border, fall under the purview of Immigrations and Customs Enforcement.

According to published DHS statistics on "deportable aliens located," over the last decade USBP apprehensions in the southwest sectors (Figure 1-1) have accounted for between 86 and 91 percent of total apprehensions of unauthorized immigrants (U.S. Department of Homeland Security, 2011b). Apprehensions in the southwest sectors have fallen by approximately 64 percent between 2001 and 2010, as shown in Figure 1-2. This decline has coincided with stepped-up immigration enforcement efforts by DHS. The number of USBP agents nearly doubled in that period, from approximately 10,000 in 2004 to more than 20,500 in 2010, and DHS has deployed thousands of technology assets (including mobile surveillance units, thermal imaging systems, non-intrusive inspection equipment, and aerial surveillance) along the U.S.–Mexico border (U.S. Department of Homeland Security, 2011a).

DHS has also implemented a variety of "consequence delivery" programs in which unauthorized immigrants are not simply "voluntarily returned" without any attendant civil or criminal consequences but instead are placed in administrative and/or legal proceedings prior to being formally

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

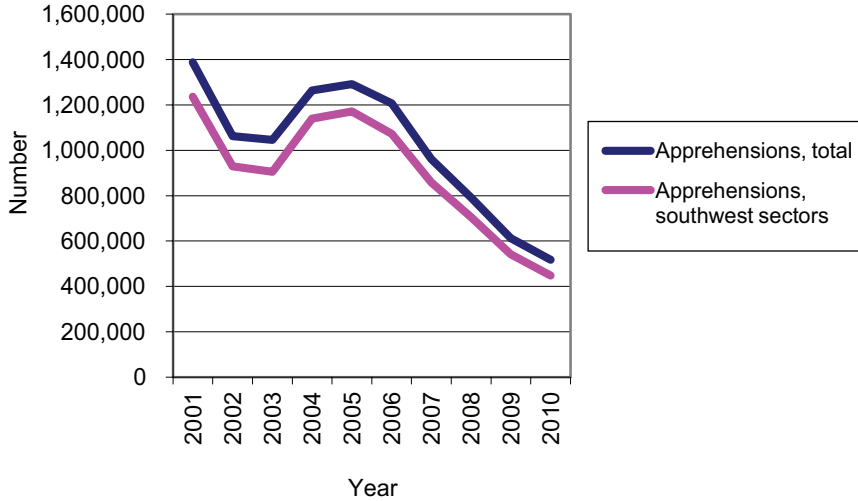


FIGURE 1-1 Apprehensions of unauthorized immigrants, 2001-2010.
SOURCE: Data from U.S. Department of Homeland Security, 2011b:Table 35.

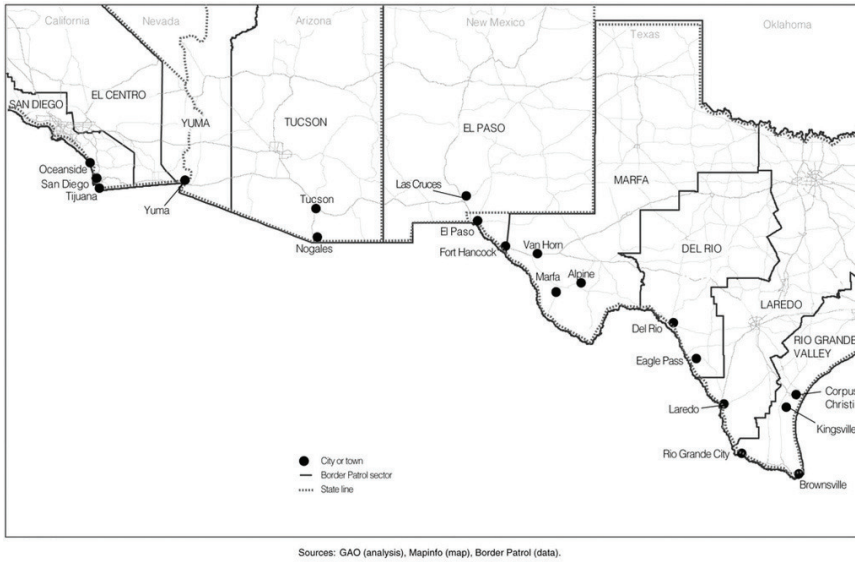


FIGURE 1-2 U.S. Border Patrol sectors along the U.S.–Mexico border.
NOTE: The sectors are San Diego and El Centro in California; Yuma and Tucson in Arizona; and El Paso, Marfa, Del Rio, Laredo, and Rio Grande Valley in Texas.
SOURCE: U.S. Government Accountability Office, 2011b:2.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

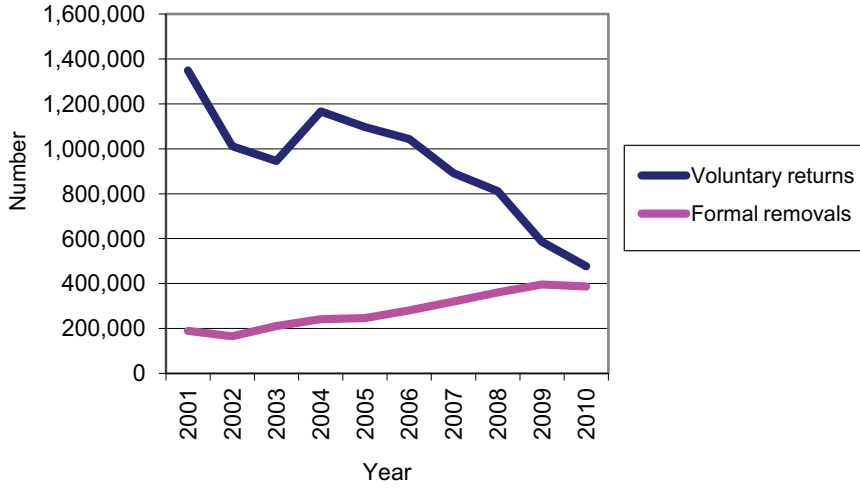


FIGURE 1-3 Returns and removals of unauthorized immigrants, 2001-2010.
SOURCE: Data from U.S. Department of Homeland Security, 2011b:Table 36.

removed. DHS has also expanded programs, such as Secure Communities, that screen for removable immigrants who come into contact with the criminal justice system.[1] As a result of such initiatives, the number of formal removals has roughly doubled between 2001 and 2010 (Figure 1-3).[2,3]

During the same period, the U.S. economy shed millions of jobs, and the unemployment rate more than doubled from 4.6 percent in 2007 to 9.6 percent in 2010. The decline in employment prospects in the United States may have acted as a deterrent to immigration. Thus, the number of apprehensions may have declined even if no change occurred in DHS's effectiveness and DHS continued to apprehend illegal crossers at the same rate as before.

To properly evaluate the effectiveness of DHS's investments in immigration enforcement, one needs an appropriate measure (or set of measures) of the total flow of unauthorized immigrants at and between the ports of entry.

[1]See National Research Council (2011:Chapter 4) for a detailed description of how the immigration enforcement system operates.

[2]The total of returns and removals in Figure 1-3 exceeds the total number of apprehensions in Figure 1-1. See National Research Council (2011:48-51) for a discussion of some of the limitations of published DHS apprehensions data.

[3]In fiscal 2010, approximately 75 percent of unauthorized immigrants who were voluntarily returned were Mexican nationals, while 1 percent were from Central America. In that same year, approximately 73 percent of unauthorized immigrants who were formally removed were Mexican nationals, while about 20 percent were from Central America (U.S. Department of Homeland Security, 2011b).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Trends in total apprehensions are, by themselves, inadequate because different outcomes may be given similar interpretations (Morral et al., 2011). For example, increases in border apprehensions may be suggestive of more effective enforcement as long as the underlying flow of unauthorized immigrants is declining, constant, or increasing by a smaller amount than the rise in apprehensions. However, declines in border apprehensions might also be suggestive of more effective enforcement as long as the magnitude of the decline in apprehensions is larger than the magnitude of the decline in the underlying flow of unauthorized immigrants.[4] In both instances, there is an increase in the ratio of apprehensions to unauthorized crossings.[5]

DHS has charged the National Research Council (NRC) with providing guidance on the use of survey options and other methodologies to estimate the number of unauthorized crossings at the U.S.–Mexico border, preferably by geographic region and on a quarterly basis (see Box 1-1). The NRC appointed the Panel on Survey Options for Estimating the Flow of Unauthorized Crossings at the U.S.–Mexican Border to carry out this task. A better understanding of the magnitude, timing, and location of these flows will help DHS to better evaluate the effectiveness of its enforcement efforts, allocate its resources along the border more efficiently, and provide a more complete report to the public on the state of illegal immigration.

Effective enforcement by DHS has at least three aspects: first, stopping individuals who are in the process of attempting to enter the United States from succeeding; second, doing this so well that potential entrants stop trying and return to their homes; and third, discouraging potential migrants from even leaving their homes in order to attempt illegal entry into the United States. The focus of most DHS measurement attempts, most surveys, and this report is on the first two of these outcomes. Full assessment of enforcement success, however, would require measuring the third outcome—discouragement at the origin. The measurement of this third outcome is beyond the scope of this report. Broad surveys of the Mexican population and complex analyses of the characteristics of potential migrants would be required to assess such discouragement.

Chapter 2 of the report provides context for the use of surveys and statistical methods for estimating flows by providing an overview of recent estimates on the stocks and flows of undocumented immigrants; a brief history of unauthorized immigration and policy responses to it, including a short summary of studies looking at the effectiveness of border enforcement; and a description of the border crossing process for unauthorized

[4]If enforcement efforts have a deterrent effect and cause the underlying flow of unauthorized immigrants to decline, enforcement can be regarded as effective even if the magnitude of the decline in apprehensions is smaller than the magnitude of the decline in-migration.

[5]However, as suggested in the previous note, the ratio of apprehensions to border crossings may not be completely informative about the effectiveness of enforcement efforts if those efforts also have a deterrent effect on unauthorized immigration.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

BOX 1-1
Statement of Task

An ad hoc panel of experts will evaluate survey options to estimate the number of foreign nationals who attempt illegal entry across the U.S.–Mexican land border each year and/or the probability of apprehension of illegal entrants. The panel will evaluate currently available and potential survey sources for estimating the number of foreign nationals who attempt illegal entry and/or the probability of apprehension of illegal entrants. These sources will include surveys currently conducted in the United States and in Mexico by both government and nongovernment entities. As part of its evaluation, the panel may consider options for initiating new survey-based data collection, modifying existing surveys, and combining data from different sources. The panel will also review the administrative data sources that the Department of Homeland Security is currently evaluating for estimating the illegal flow in order to inform the survey options. The administrative sources include, but are not limited to, apprehension and re-apprehension rates, agency estimates, results of consequence programs, interviews with apprehended foreign nationals, and results of sensor activity. In addition to survey and administrative data, the panel may also consider various modeling methodologies. The panel will provide guidance on survey implementation and cost estimates under various options. The panel will also evaluate the possibilities of collecting information through surveys that could enable estimation of annual flow and/or the probability of apprehension of illegal entrants on a quarterly basis and for particular regions of the U.S.–Mexican border. The panel will issue a report with findings and recommendations.

migrants. Chapter 3 outlines the features of the major surveys in the United States and Mexico that collect information about migration and border crossing, and Chapter 4 lays out their usefulness and limitations for estimating flows. Chapter 5 then assesses the usefulness of DHS administrative data for measuring the flow of unauthorized migrants into the United States. Finally, Chapter 6 discusses aspects of model-based approaches that would combine information from a variety of sources, including surveys and administrative data.

The panel also notes that although the Office of Immigration Statistics in DHS is not a "statistical agency" per se, it is nevertheless considered a "statistical program" by the Office of Management and Budget (U.S. Office of Management and Budget, 2011) and, as such, would benefit from adhering to the broader principles of good statistical practice as carried out in the federal statistical system (National Research Council, 2009). The panel's conclusions and recommendations, particularly with regard to the availability of administrative data in DHS's enforcement database, are informed by this presumption.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

2

The Process of Unauthorized Crossing at the U.S.–Mexico Border

A discussion of the use of surveys and statistical methods to measure unauthorized border flows needs to be set in the context of border enforcement and the migration process, both current and past. We focus on Mexican migrants since Mexican nationals account for the vast majority of attempted unauthorized border crossings at the U.S.–Mexico border. According to the U.S. Border Patrol (USBP), 90 percent of the people apprehended along the southwest U.S. border in fiscal year 2010 were from Mexico (Sapp, 2011).[1]

This chapter begins by providing an overview of the population of unauthorized migrants and recent estimates of the stock and flows of undocumented immigrants. It then presents a brief history of unauthorized immigration and policy responses to it, including a short summary of studies looking at the effectiveness of border enforcement. Then, the chapter describes the contemporary border crossing process for unauthorized migrants aiming to enter without inspection. Finally, the chapter discusses what is known about the apparent extent of the involvement of organized crime and drug cartels in migrant smuggling.

[1]This number could be overestimated, as nationals from other countries have an incentive to falsely report that they are Mexican in order to be returned to the Mexican border. This would allow them to attempt re-entry more easily than if they were deported to their home country.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

CURRENT UNAUTHORIZED POPULATION: STOCK AND FLOW ESTIMATES

Estimates of the number of unauthorized immigrants living in the United States put this population at around 11 million in 2010 (Passel and Cohn, 2011), which is about 30 percent of the foreign-born population and over 5 percent of the U.S. workforce.[2] While approximately 6.1 million are from Mexico, representing 55 percent of the total, other countries contributing substantially to the undocumented population include Brazil, China, Ecuador, El Salvador, Guatemala, Honduras, India, Korea, and the Philippines (Hoefer et al., 2012; Passel et al., 2012). The number of unauthorized residents originating from Mexico is more than 10 times larger than the number of unauthorized residents from any other country. Between 1999 and 2009, Mexican men aged 15 to 50 years old accounted for three-fourths of all apprehensions at the U.S.–Mexico border, and Mexican women aged 15 to 50 years old accounted for an additional one-seventh of apprehensions. The remaining apprehensions are distributed among juvenile Mexican nationals, senior Mexican nationals, and migrants from other countries. Reconciling the rates of apprehensions by nationality and gender with the estimates of the stocks of illegal residents in the United States requires analyses that are beyond the scope of this report.

Data on illegal immigrant flows are even harder to come by than estimates of the stock. Still, changes in the stock of illegal immigrants should roughly capture net flows, once attrition and deaths are subtracted and an upward adjustment is made for the household survey undercount.[3] For instance, the unauthorized immigrant population increased in net terms by about 500,000 annually during the early 2000s. After allowing for undercount, mortality, and return migration, the inflow of unauthorized immigrants was estimated at around 850,000 per year during the period 2000 to 2005 (Passel and Cohn, 2010).

Unauthorized immigration slowed considerably during the U.S. housing bust and subsequent recession as unemployment rates soared (Passel and Cohn, 2009a). Return migration may also have increased. As a result, the unauthorized immigrant population decreased slightly between 2007 and 2009 and has since stabilized (Hoefer et al., 2012; Passel and Cohn, 2011). Falling immigration and rising returns were most evident in the Mexican case. According to the Mexican Census, 1.4 million Mexicans and their children left the United States between 2005 and 2010, about the same as the number who are estimated to have entered the country (Passel et al.,

[2]Of the 11 million unauthorized immigrants, 8 million are estimated to be in the labor force (Passel and Cohn, 2011).

[3]The limitations of U.S. household surveys in measuring the unauthorized population are discussed in Chapter 4.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

2012). Net migration from Mexico was therefore zero or perhaps even slightly negative during this period, compared with net illegal migration of 2.3 million from Mexico between 1995 and 2000. Other studies suggest the rising returns occurred at the tail end of the crisis (cf. Giorguli and Gutiérrez, in press; Rendall et al., 2011). As a result of declining inflows and rising returns, the unauthorized Mexican-born population fell to 6.1 million in 2011, down from 7 million at its peak in 2007 (Passel et al., 2012).

Characteristics of the Unauthorized Population

Not all unauthorized immigrants "enter without inspection" at the border. People who overstay or otherwise violate the terms of their visas, such as by working while on a tourist visa, are also considered unauthorized. Among the unauthorized population, between 25 and 40 percent are thought to be visa overstayers, while the rest entered without inspection (Passel, 2005). The efforts of the U.S. Department of Homeland Security (DHS) to identify and report on overstays have been hindered by the absence of a comprehensive biometric entry and exit system for identifying overstayers (U.S. Government Accountability Office, 2011a).

These population estimates represent unauthorized immigrants who are U.S. residents, i.e., people living in the United States on a longer-term basis (more than 1 year). However, much, and probably most, of the undocumented flow across the U.S. border consists of Mexicans who are coming to the United States for short-term employment, including seasonal work in agriculture and construction. This distinction between longer-terms residents ("settlers") and short-term residents ("sojourners") is important in considering data from different sources and policies concerning immigration and admission.

Even among the settler population, many return to their home countries. But in recent years, a larger share of the unauthorized population consists of long-term residents who appear to be permanently settled in the United States (Taylor et al., 2011). Estimates suggest that about one in seven unauthorized immigrants arrived within the past 5 years, and most (almost 60 percent) have lived here for more than a decade (Hoefer et al., 2012). This pattern represents a dramatic change from a decade earlier, when about one-third of unauthorized immigrants had been in the United States for less than 5 years and a minority (44 percent) had been in the country for more than a decade (Taylor et al., 2011). Some research suggests that tougher border enforcement has played a role in reducing circular migration and inducing immigrants to permanently settle in the United States (Angelucci, 2012; Massey and Pren, 2012; Reyes, 2004). Although the typical unauthorized resident is a male aged 18-39, there are

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

substantial numbers of undocumented women and children, and U.S. births to undocumented immigrant women have risen. Nearly half (46 percent) of unauthorized adults living in the United States have a minor child, and about 80 percent of these children are U.S.-born (Taylor et al., 2011).

Unauthorized immigrants are highly geographically clustered in urban areas but have become more dispersed during the past two decades. Almost 40 percent of unauthorized immigrants reside in just two states, California and Texas. However, while California's undocumented population remained nearly unchanged during the 2000s, Texas's population grew. States in the Southeast and Mountain West, such as Georgia, North Carolina, and Nevada, experienced large increases in their unauthorized immigrant populations during the late 1990s and the 2000s (Orrenius and Zavodny, 2012). The changing geography of destinations stems primarily from the deflection of flows away from California. According to census data, two-thirds of Mexicans who entered the United States from 1985 to 1990 went to California. By 1995-2000, the share had dropped to one-third, and it has remained at that level during 2000-2005 and 2005-2010 (Durand et al., 2005; Massey and Capoferro, 2008).

Unauthorized immigrant men have very high rates of labor force participation, which suggests that they enter the United States to work. In 2008, 94 percent of working-age undocumented immigrant men were in the labor force, either working or searching for a job (Passel and Cohn, 2009b). This compares with 83 percent labor force participation for similarly aged U.S.-born men. The undocumented population's relative youth and lack of access to government transfer programs explains some of the exceptional attachment to the labor force. However, undocumented working-age women are actually less likely to be in the labor force than either U.S. natives or legal immigrants, probably because a greater proportion of them have young children (Passel and Cohn, 2009b).

Unauthorized immigrants tend to have low levels of education and hold low-wage jobs. Almost half of adult unauthorized immigrants have not completed high school, and they make up 22 percent of all non-high school graduates in the United States. Undocumented workers are disproportionately employed in construction; food service; building, grounds keeping, and maintenance; and farming (Passel and Cohn, 2009b).

Unauthorized Immigration in the Pre–World War II Years

The United States had an open immigration policy for much of its early history (with the clear exception of immigration from Africa, which was mostly confined to the transatlantic slave trade until it was banned in 1808). As the young nation depended on immigrants to fill jobs, claim the land, and populate the frontier, immigration restrictions did not begin to

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

come into effect until the late 19th century. Once implemented, they were of two kinds. One set of laws excluded certain types of individuals, such as criminals and prostitutes (1875), people with mental disorders and others deemed likely to become a public charge (1882), and individuals with pre-arranged labor contracts (1885).[4] Other laws excluded entire nationalities, such as the Chinese Exclusion Act of 1882, which banned the entry of all Chinese immigrants, and the so-called Gentlemen's Agreement of 1907, which ended immigration from Japan (Orrenius and Zavodny, 2010).

Mexican migration across the southwest U.S. border occurred relatively late when compared with inflows of immigrants from most of Europe and East Asia, and it was for the most part unrestricted by law or regulation. Despite the fact that the United States in the mid-19th century acquired a considerable amount of territory previously belonging to Mexico, only around 50,000–100,000 Mexican nationals became U.S. citizens as a result of this territorial redistribution (Henderson, 2011).

Immigration from Mexico remained low for almost five decades, picking up only in the early 20th century. Permanent admissions from Mexico never surpassed 10,000 per decade during the 19th century and were generally below 5,000. The lowest number of admissions per decade, 734, was in the 1890s (U.S. Department of Homeland Security, 2011b:Table 2; see also Orrenius and Zavodny, 2010). With flows from China and Japan barred in the late 19th and early 20th centuries (a ban that would last until the 1940s), Mexican workers were recruited to build railroads, cultivate fields, and work in mines across the U.S. Southwest and parts of the Midwest (Cardoso, 1980; Durand and Arias, 2005; Foerster, 1925; Gamio, 1930). Mexican immigrant inflows reached 31,188 by the first decade of the 1900s (U.S. Department of Homeland Security, 2011b:Table 2). This practice continued and expanded into the 1910s, fueled by the difficulty of transatlantic movement during World War I and imposition of literacy requirements in 1917, which explicitly exempted Mexican workers. The violence of the Mexican Revolution also uprooted thousands of people fleeing the areas with most turmoil in Central and Northern Mexico between 1911 and 1917. The number of Mexican immigrants in the 1910s rose considerably to 185,334 for the decade (U.S. Department of Homeland Security, 2011b).

Mexican immigration rose further in the 1920s to almost one-half million (U.S. Department of Homeland Security, 2011b:Table 2), due in part to increasing restrictions on European immigration. The establishment of the first numerical limits on immigration using national origins quotas in 1921 capped annual admissions by country to 3 percent of the population present in the United States in 1910. The Immigration Act of 1924 lowered

[4]Prospective immigrants circumvented this law by denying that they had a pre-arranged contract.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

the quotas to 2 percent of the 1890 population by national origin. USBP was established in the same year and spent its early days stopping European and Asian immigrants seeking to circumvent immigration restrictions and liquor smugglers who violated Prohibition, rather than intercepting illegal immigrants from Mexico (Zolberg, 2006).

Once the economy collapsed during the Great Depression, which followed the crash of the stock market in October 1929, legal immigration from all over the world fell considerably (from 4.3 million admissions in the 1920s to slightly less than 700,000 in the 1930s [U.S. Department of Homeland Security, 2011b:Table 2]). While immigration from Mexico also slowed drastically during this period, to 32,709, hundreds of thousands more returned home, some voluntarily and some as a result of being "pushed out" or deported (Hoffman, 1974; Zolberg, 2006). As a result, the number of Mexicans living in the United States dropped from 641,000 in the 1930 census to 377,000 in 1940.

The Bracero Program and Unauthorized Immigration in the Post–World War II Years

Economic growth resumed during World War II, and labor shortages arose as the war effort intensified. In 1942, the U.S. and Mexican governments crafted the Bracero Program, which would bring in about 200,000 Mexican workers annually over the next 22 years, although flows varied greatly over time. The Bracero Program evolved considerably during its 22 years of existence, going from a binational accord managed by the U.S. Department of Labor and the Mexican Ministry of Foreign Affairs to a unilateral effort managed by the U.S. Immigration and Naturalization Service (Calavita, 1992). It began with 4,200 entries in 1942, rising to 62,000 in 1944, falling to around 20,000 in 1946, and then rebounding to 100,000 by 1949. The flows fluctuated around 200,000 per year from 1951 to 1953, but a large expansion occurred in the late 1950s with annual flows between 400,000 and 450,000 per year before declining in the early 1960s and then being terminated entirely in 1965 (Calavita, 1992; Massey et al., 2002).

Unauthorized migration became more commonplace during the era of the Bracero Program because of robust U.S. labor demand and the temptation of employers and immigrants to circumvent the administrative hassles of complying with the program's rules in the United States (Calavita, 1992) and Mexico (Fitzgerald, 2009:Chapter 2). During this period, there was no law barring the employment of unauthorized workers; on the contrary, the 1952 immigration law included the so-called Texas proviso, which specifi-

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

cally permitted the employment of illegal immigrants.[5] It would continue to be legal to hire undocumented labor until passage of the Immigration Reform and Control Act (IRCA) 34 years later.

The Bracero Program was intended to admit Mexican workers for short-term employment; it did not allow the migrants to settle in the United States. This is evident from the population data on Mexicans in the United States. Notwithstanding the more than 4 million admissions during the 1942-1964 period, the number of Mexican immigrants living in the United States increased by only about 200,000 over 20 years, from 377,000 in the 1940 census to 576,000 in 1960. With the end of the Bracero Program in 1964, the United States implemented a new immigrant preference system favoring family reunification and, to a lesser extent, certain forms of skilled labor. The 1965 law also imposed a Western Hemispheric cap on permanent resident visas that took effect in 1968. As far as Mexico was concerned, the changes virtually eliminated temporary Mexican worker visas and reduced the number of available permanent visas (Massey et al., 2009). The demand for Mexican labor, however, did not change. Within a short period, employers substituted unauthorized immigrant labor for Bracero workers (Massey and Pren, 2012). At the same time, the nature of Mexican labor migration began to change, with an increasing proportion of workers employed year-round and full-time in non-agricultural jobs (Cornelius, 1992; Riosmena, 2004). Perhaps predictably, undocumented migration increased steadily during the 1970s, ushering in the beginning of what some scholars have called the "Undocumented Era" (Massey et al., 2002).

The early historical experience demonstrates the evolution of the immigration system, driven by two aspects that continue to apply in the present-day context, especially with regard to illegal immigration. First, economic conditions and the state of labor demand (which some refer to as "business interests") play a central role in immigration law and the extent to which it is implemented and enforced. Second, barring extraordinary events such as the Great Depression, tighter regulations with regard to a particular group's migration or method of entry (e.g., Bracero permits or unskilled worker visas) often result in a compensatory reaction by another group's migration or method of entry (e.g., unauthorized entries). Although the immigration system is more likely to "self-correct" under conditions of

[5]Texas growers were deliberately left out of the Bracero agreement initially, at the request of Mexican authorities who cited Texas employers' abuses of Mexican workers during the prior wave of Mexican immigration. In response, these growers recruited and hired undocumented labor and lobbied against restrictions on this practice. They would become part of the program in its later stages when the Mexican government softened its stance and as the United States continued bringing in Braceros without the direct cooperation of the Mexican government (Calavita, 1992).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

relatively weak border enforcement (as the southwest U.S. border can be characterized before the mid-1990s), even increased enforcement has not fundamentally changed this dynamic, as we discuss below.

IRCA and Its Consequences

The unauthorized population was estimated at around 2.1 million in 1980 (Warren and Passel, 1987).[6] In response to the growth in undocumented immigration, Congress passed IRCA in 1986. This law represented a landmark change in U.S. policy on unauthorized immigration in several ways. First, it sought to stem future unauthorized immigration by making it illegal to knowingly hire undocumented workers, requiring employers to verify workers' eligibility for employment, increasing funding for border enforcement, and establishing the H-2A and H-2B programs for temporary agricultural and non-agricultural workers (Orrenius and Zavodny, 2012).

IRCA also included provisions to legalize most of the unauthorized population already in the country. A Legally Authorized Workers program granted legal permanent residence to immigrants who could document continuous presence in the country since January 1, 1982.[7] A Special Agricultural Workers program granted legal status to circular and seasonal migrant workers as long as they had worked at least 90 days during the past year. Together, these programs regularized 2.7 million undocumented immigrants, three-fourths of whom were from Mexico (Durand et al., 1999:523; Orrenius and Zavodny, 2003:439; Phillips and Massey, 1999:233).

Despite mass legalization and employer sanctions under IRCA, the number of unauthorized immigrants had increased to 3.4 million by 1992 (Warren, 1994, cited in Espenshade, 1995a:201).[8] By 2000, this number had grown to 8.4 million (Passel et al., 2004). Although this growth may be partly exaggerated by different types of data flaws, more present in earlier statistics,[9] it is clear even after taking these factors into account that the

[6]In the early 1980s, there was considerable uncertainty about the size of the unauthorized immigrant population, and some contemporary estimates were as high as 10 to 12 million (Edmonston et al., 1990).

[7]Approved applicants first received temporary legal status. After 18 months and the successful completion of English language and civics classes, applicants then received permanent legal status (green cards).

[8]While these two programs decreased the undocumented population considerably, some migrants did not meet the requirements and others arrived after the application deadline. In addition, it took a good part of the 1990s for all IRCA-related regularizations to go through the Immigration and Naturalization Service backlog (Rytina, 2002).

[9]Most importantly, a higher undercount of the immigrant population (and the unauthorized in particular) in the 1990 Census relative to the 2000 Census (cf. Robinson et al., 1993, 2002) may have misallocated some of this growth to the 1990s, when in fact it actually may have occurred prior to 1990.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

unauthorized population increased considerably throughout the 1990s and continued to do so until the onset of the 2007-2009 housing collapse and financial crisis. Based on a methodology similar to that of Passel and colleagues (2004), Passel and Cohn (2011) estimate that this number increased to 12 million by 2007 before declining to 11 million by 2010 (Hoefer et al., 2012; Passel and Cohn, 2011).

The Rise of Border Enforcement

IRCA's various provisions, all designed to stem unauthorized immigration, appear to have had little effect on the flow of unauthorized migrants across the U.S.–Mexico border (Orrenius and Zavodny, 2003). USBP apprehensions, sometimes used as a rough gauge of changes in illicit border crossings (keeping in mind that apprehensions are also a function of U.S. enforcement efforts), declined for three consecutive years after IRCA, but then resumed an upward trend. By the early 1990s, USBP apprehensions were back to pre-IRCA levels (see Figure 2-1).

While successful entries increased steadily between the early 1990s and 2008, apprehensions at the border have fluctuated considerably.[10] Around 1970, the number of apprehensions exhibited an upward trend, followed by sizable fluctuations in the 1980s and 1990s, and reached a peak in fiscal year 2000 at almost 1.8 million. Since 2000, the number of apprehensions has plummeted, currently standing at around half a million—a level not recorded since 1972 (U.S. Department of Homeland Security, 2011b:Table 33).

The number of apprehensions clearly depends on the number of attempted crossings, and thus researchers have focused on the changes, if any, in the probability of apprehension. Attempts are influenced by socioeconomic conditions in sending areas and potential destinations; by security conditions on the southern side of the U.S.–Mexico border; and by U.S. immigration policy, including legal immigration opportunities and border and interior enforcement. Potential migrants can consider the immigration system as a whole and assess their best migration options. Greater availability of visas may reduce the probability of migrating illegally, assuming all other factors remain the same. In a similar vein, permissive policies at the ports of entry reduce incentives to cross between ports of entry. Among migrants choosing to cross illegally, more enforcement in one area of the border can increase crossings in another.

[10]Given that the more reliable estimates pertain to stocks estimated indirectly, undocumented flows have surely fluctuated more than scholars are able to measure. Even allowing for this, flows seem to have fluctuated less than apprehensions. Compare Figure 2-1 with Massey and Singer (1995) and Passel and Cohn (2011).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

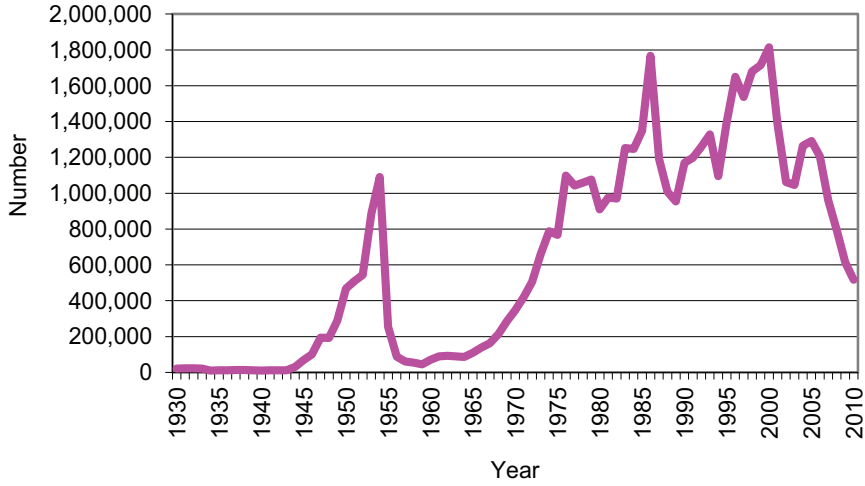


FIGURE 2-1 Apprehensions of unauthorized immigrants, 1930-2010.
SOURCE: Data from U.S. Department of Homeland Security, 2011b:Table 33.

The probability of apprehension is mostly a function of the level and spatial distribution of border enforcement effort (including not only line-watch personnel and physical infrastructure but also "consequence programs," described below) and the capacity of migrants and smugglers to anticipate and react to enforcement efforts. In the immigration system, more border enforcement encourages greater use of smugglers and innovative techniques, such as decoys; border fences may lead to digging tunnels; and so on.

Immigration policy in general, and border enforcement policies and practices more specifically, should affect the number of apprehensions through the effectiveness of border enforcement (i.e., the probability of catching people attempting to cross illegally [Espenshade, 1995b; Massey and Singer, 1995]) and by deterring people from attempting to cross (Angelucci, 2012; Cornelius and Salehyan, 2007; Hanson and Spilimbergo, 1999; Massey and Riosmena, 2010; Orrenius and Zavodny, 2005). We discuss these two factors in a separate section, "Effectiveness of Border Enforcement."

Continuing undocumented inflows in the early 1990s prompted USBP to change strategies and adopt a "prevention through deterrence" strategy, which "called for reducing unauthorized migration by placing agents and resources directly on the border along population centers in order to deter would-be migrants from entering the country" (Nuñez-Neto and Viña, 2006:1; U.S. General Accounting Office, 1999). In 1993, the El Paso sec-

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

tor launched Operation Hold the Line, which deterred crossings in the urban corridor by constructing barriers/fences and stationing agents close together.

The new strategy led to a drastic decline in apprehensions in the El Paso sector, and similar operations were subsequently launched in San Diego in 1994 (Operation Gatekeeper), Nogales, Arizona, in 1997 (Operation Safeguard), and South Texas in 1997 (Operation Rio Grande). Operation Gatekeeper was by far the most important of these initiatives, given that the great majority of unauthorized migrants at that time were crossing into California through Tijuana-San Diego.

The site-specific crackdowns resulted in two responses by border crossers, neither of which suggests that overall illegal immigration was deterred. First, operations such as Gatekeeper deflected attempted crossings away from areas with increased USBP activity toward less-patrolled territory. Migrants initially attempted to cross in these areas within the same sector (Nuñez-Neto and Viña, 2006:8-12). Eventually, migrants moved to other sectors, such as Arizona and, to a lesser extent, South Texas (Massey et al., 2002; Spener, 2009:46). Second, within highly trafficked corridors, migrants began crossing through remote country, away from built-up, settled areas. More perilous crossings led to a higher number of deaths among unauthorized border crossers (Cornelius, 2001; Eschbach et al., 1999; U.S. Government Accountability Office, 2006).

Between 1990 and 2000, the USBP budget increased 205 percent in real terms and the number of USBP agents assigned to southwest border sectors rose 165 percent, from 3,226 in 1990 to 8,525 in 2000.[11] Border enforcement rose further in the 2000s; the budget increased an additional 157 percent in real terms between 2000 and 2011[12] and the number of agents on the southwest U.S. border more than doubled to 18,506 in 2011.[13] Since 1990, USBP has also implemented the use of more advanced technology, including double fences and watch towers, ground sensors, remote video surveillance, and aerial and marine surveillance. According to USBP, as of February 2012, 651 miles of the 1,969 miles along the U.S.–Mexico border are fenced, 352 miles with a primary pedestrian fence and 299 miles with a vehicle fence.[14] There is also a secondary fence built along specific parts of urban corridors.

[11]Agents data through 2000 are from http://trac.syr.edu/immigration/reports/143/include/rep143table2.html (May 14, 2012) and are for September of each year.

[12]Border Patrol budget data by fiscal year from http://www.cbp.gov/linkhandler/cgov/border_security/border_patrol/usbp_statistics/budget_stats.ctt/budget_stats.pdf (May 14, 2012), converted to real dollars using CPI-U.

[13]See http://www.cbp.gov/linkhandler/cgov/border_security/border_patrol/usbp_statistics/staffing_92_10.ctt/staffing_92_11.pdf (May 14, 2012).

[14]See http://www.cbp.gov/xp/cgov/border_security/ti/ti_news/sbi_fence/ (April 10, 2012).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

The spillovers in the immigration-enforcement system are readily apparent by comparing apprehensions in sectors with and without increased enforcement operations. Figure 2-2 is an indexed chart of apprehensions by sector for the period 1992 to 2011. Around the time of Operations Hold the Line and Gatekeeper, the decline in activity in the El Paso and San Diego sectors is notable. It is followed by a rise in apprehensions in El Centro, the Arizona sectors (Tucson and Yuma) and Del Rio, Texas. The share of apprehensions in San Diego and El Paso fell from 71 percent in 1992 to 16 percent in 2000.

Consequence Policies

Notwithstanding this unprecedented border buildup, illegal immigration continued, reaching new highs during the U.S. economic boom and Mexican peso crisis of the mid- to late 1990s. Although apprehensions also increased, the undocumented flow continued, and the unauthorized population expanded rapidly (Orrenius, 2001). Migrants may not have been deterred by this increased enforcement in part due to the USBP's "catch and release" policy, whereby most apprehended migrants who were from Mexico signed voluntary departure contracts and boarded a bus back to the border, after which they would simply try to cross again within a day or two (Kossoudji, 1992). USBP also had difficulty identifying repeat offenders and smugglers because, when apprehended, migrants would provide different names (Koslowski, 2002).

Whereas the first phase of enforcement, still underway, focused on apprehending unauthorized crossers, the next phase attempted to address how to best prevent re-entry attempts. Efforts to deter repeat attempts centered on better recordkeeping (for example, universal finger printing since the late 1990s and crosschecking with the Federal Bureau of Investigation [FBI] IAFIS database since 2004) and harsher punishments for repeat crossers. Harsher and speedier punishment for illegal entry came to be referred to as "consequence policies" (Fisher, 2011). The Illegal Immigration Reform and Immigrant Responsibility Act of 1996 (IIRIRA) laid the groundwork for consequence policies by instituting expedited removal, interior repatriation, and 3- and 10-year admission bars for previously unauthorized immigrants seeking to be admitted legally to the United States.[15] Expedited removal is the process by which a non-U.S. citizen present in the country for less than 14 days and located within 100 miles of the border can be physically

[15]The law also set harsher punishment for smugglers. In addition, IIRIRA increased funding for border and interior enforcement and launched Basic Pilot (an employment verification program that was the precursor to E-Verify) (Orrenius and Zavodny, 2012).

Copyright © National Academy of Sciences. All rights reserved.

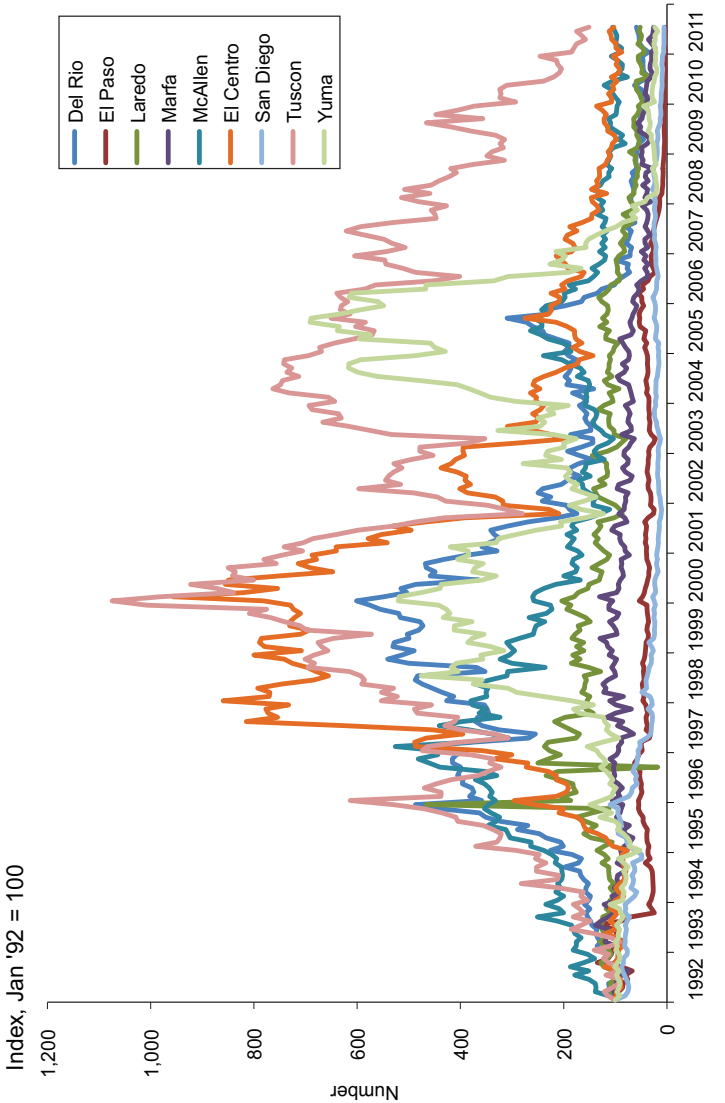


FIGURE 2-2 Border Patrol apprehensions by sector.
SOURCE: Unpublished data from U.S. Border Patrol, U.S. Department of Homeland Security, with seasonal adjustment by the Federal Reserve Bank of Dallas.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

removed from the United States with limited legal recourse.[16] Customs and Border Protection (CBP) officers, including USBP, issue expedited removal orders in a short proceeding, which also carries a 5-year re-entry bar. Expedited removal of non-Mexicans requires detention space, and expansion of expedited removal in 2005 to include all USBP sectors could only come about once authorities secured sufficient jail beds to house the undocumented aliens (National Research Council, 2011).

Over the past decade, USBP implemented the consequence policies laid out in IIRIRA in addition to its own initiatives. In 2003, USBP in Arizona launched a program of lateral repatriation: the practice of returning apprehended migrants to points along the border that are distant from the USBP station where the apprehension took place. The stated purpose of lateral repatriation was to prevent deaths in the Arizona desert, but it also results in the separation of migrants from their smugglers, complicating re-entry attempts.

USBP also launched Operation Streamline in the Del Rio sector in Texas in 2005, and by 2009 six of the nine southwestern USBP sectors had implemented versions of it (Lydgate, 2010). Under Operation Streamline, USBP and the U.S. Department of Justice cooperate to subject as many migrants as possible to federal criminal prosecution (see National Research Council [2011:Chapter 4] for further details). Although most such offenses are misdemeanors and result in very short jail terms, the punishment marks a dramatic shift away from the catch and release tactics of the 1980s and 1990s, in which migrants were quickly returned to Mexico.[17]

MOTIVATIONS FOR MIGRATION

Much of the research on the decision to illegally migrate to the United States considers the Mexican case (see Massey et al., 1998:Chapter 3). As noted above, Mexico–U.S. migration has a long history of responding largely to changes in labor demand in the United States. The combination of job opportunities and higher U.S. wages has always been a powerful incentive for Mexican migration. When jobs—particularly in the construction sector—dried up in the recession that began in 2008, unauthorized immigration plummeted (Papademetriou and Terrazas, 2009). In contrast, legal permanent immigration was largely unresponsive to the economic

[16]IIRIRA originally mandated expedited removal for arriving aliens at ports of entry but gave the U.S. Department of Justice the option of applying it more broadly, which it later did (Siskin et al., 2006:11).

[17]However, returning non-Mexican migrants to their home countries has always been a lengthy process, during which the migrants have typically been detained.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

downturn, given its emphasis on family reunification and the large backlog in applications (Papademetriou and Terrazas, 2009).

Economic opportunity in sending areas, or the lack thereof, also motivates migration (Stark and Bloom, 1985). Families create safeguards against different kinds of shocks (e.g., employment loss or inflation caused by economic crises, crop failure caused by droughts) by diversifying their sources of income through the allocation of family labor abroad. Migration serves as a mechanism for risk management if market conditions in destinations are negatively or weakly correlated with those in the sending community; if local economic conditions deteriorate, the household can then rely on migrant remittances for support. This is particularly important when origin communities lack established financial institutions, such as banks and insurance companies, or when families do not have access to loans or insurance even if the appropriate institutions exist.

The absence of formal credit markets leads to another motivation for migration: target savings. Low wages and limited lending make it very difficult for families to accumulate capital. Many migrants are therefore target earners, engaging in temporary migration for the express purpose of accumulating a set amount of money. Research has found evidence consistent with these motivations in the Mexican case (Hamilton and Villarreal, 2011; Lindstrom, 1996; Lindstrom and Lauster, 2001; Massey and Espinosa, 1997; Massey and Parrado, 1998). Such target earners would probably not settle in the United States but rather are likely to return to Mexico once they have met their savings target.

Typically it is not a lack of development or absolute poverty per se that motivates migration. Mexican migrants are not a random sample of the Mexican population; they self-select and tend to be young, male, and from the middle of the education distribution (Chiquiar and Hanson, 2005; Feliciano, 2005; Ibarraran and Lubotsky, 2007; Moraga, 2010). An important factor preventing the poorest, least-educated Mexicans from international migration is the cost of illegal border crossing. The highly educated Mexicans, meanwhile, typically do not have the need to migrate or the willingness to cross the border illegally, given the risks. Economic and institutional changes, such as urbanization, technological change, the entry of women into the labor force, or the disruption of local livelihoods associated with the development process, can displace workers and also provide reasons to migrate (Sassen, 1988).

Because the development process is uneven and social mobility is low in many sending areas, migration may also be motivated by relative deprivation—i.e., migration for the purpose of improving one's social rank in the sending community (Jones, 1998). Relative deprivation can be caused by the migration and capital accumulation of other community members.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

This is one mechanism by which the momentum created by initial migration may lead to additional movement of people (Massey, 1990).

New areas of emigration in Mexico have emerged as a result of these economic transformations (Riosmena and Massey, 2012). Mexican outmigration has traditionally been concentrated in localities in the central-western states of Guanajuato, Jalisco, Michoacán, San Luis Potosí, and Zacatecas. Since the first massive wave of Mexican immigration in the early 20th century, these states have accounted for a majority of U.S. migrants (Durand et al., 2001). However, areas of origin have become increasingly dispersed over time. Over the past three decades, new sending regions have slowly emerged (Durand and Massey, 2003), particularly in areas south and east of Mexico City, reducing the importance of the central-western region of Mexico to less than half of the flow for the first time since the initiation of Mexican migration (Durand and Massey, 2003).

Migrant networks and, more generally, connections between sending and destination areas created by the migration process are another strong facilitator of migration. While economically motivated at its core, Mexico–U.S. migration became much more complex once Mexican communities were established in the United States. Models of network migration have been able to better explain migration trends and settlement patterns than just wage and income gaps between the places of origin and destination (Massey and Espinosa, 1997; Massey and Riosmena, 2010; Massey and Zenteno, 1999). Given that the demand for immigrant labor continued after the end of the Bracero Program, and given that major changes in immigration law in the late 1960s eliminated provisions for legal temporary and permanent labor migration,[18] social networks have played an even larger role in facilitating undocumented migration from Mexico (Massey and Riosmena, 2010).

Secular social and economic changes in both the United States and in sending areas, in addition to changes in immigration policy that affect the supply of certain types of visas, have had an effect on individuals' decisions to engage in undocumented migration. These changes need to be taken into consideration when quantifying or forecasting illegal immigration and evaluating the effectiveness of border enforcement. The immigration system is dynamic. Sending areas change, and migrant motivations evolve. For example, the drastic decline in Mexican fertility over the past 40 years

[18]Because most Braceros circulated back and forth between Mexico and the United States and only a minority stayed and became permanent residents, family reunification was not a choice for most Mexican migrants prior to the passing of IRCA. Settlement occurred more steadily during the Undocumented Era, both due to secular changes in the character of Mexican migration to the United States (Cornelius, 1992; Riosmena, 2004) and because of the unintended consequences of immigration policy (Angelucci, 2012; Reyes, 2004).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

is taken as evidence that a resurgence of mass unauthorized migration is unlikely (Passel and Cohn, 2012).

EFFECTIVENESS OF BORDER ENFORCEMENT

The steady growth of the unauthorized population in the United States during the past several decades might suggest at first glance that border enforcement has been ineffective. Conversely, the fact that the number of apprehensions and stock of undocumented migrants has decreased in more recent years might suggest that the stepped-up enforcement efforts at the border have been effective. As discussed briefly in the previous section, the incentives to engage in unauthorized migration have been changing due to economic and social changes in destination and sending areas and due to the expansion of migrant networks. An appropriate evaluation of immigration enforcement must consider a counterfactual scenario in which other migration factors are held constant as border enforcement is increased. This would indicate whether and to what extent border enforcement is a deterrent.[19]

Studies of migration tend to find evidence of small but significant deterrent effects of border enforcement. Gathmann (2008) found that the border buildup between 1986 and 2004 raised smugglers' fees by 17 percent and increased the time costs of crossing by 2 to 5 additional days. Orrenius (1999) found that a 20 percent increase in the smuggling fee caused a 13 to 21 percent decline in the probability of migrating. Spener (2009) notes that enforcement operations drive smuggling prices up both by requiring migrants to buy a more complex portfolio of smuggling services than in the past and by pushing traffic to remote areas, making it harder to cross the border undetected. A number of studies have found that more border enforcement negatively affects the probability of undocumented migration after accounting for the role of other forces influencing this decision, although the size of deterrence effects in most of these studies is generally modest (Amuedo-Dorantes and Bansak, 2011; Angelucci, 2012; Cornelius and Salehyan, 2007; Hanson and Spilimbergo, 1999; Massey and Riosmena, 2010; Orrenius and Zavodny, 2005). Angelucci (2012) finds that the elasticity of illegal inflows to border enforcement is between –0.4 and –0.8 and that sensitivity to enforcement has increased over time. Amuedo-Dorantes and Bansak (2011) find that an increase of one-half million linewatch hours—the average yearly increase along the U.S.–Mexico border between 1990 and 2003—reduced intentions to re-migrate among a sample of male return migrants by about 14 percent. The deterrence effects can also be short-lived: Dávila and colleagues (2002) found that, although

[19]In other words, although undocumented migration increased, it could have been even higher in the absence of increased enforcement.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

increased linewatch hours reduced apprehensions initially, the effect dissipated as migrants adapted their behavior.[20]

By using estimates or reports of unauthorized migration among people who reportedly managed to cross into the United States, these studies have looked at the net effect of enforcement on successful crossings, not on attempted ones. These studies thus assessed the combined effectiveness of border enforcement in deterring attempts and in catching migrants attempting to cross without inspection. Assuming, however, that these studies are mostly indicating the role of border enforcement on deterring an initial attempt to cross the border, the probability of apprehension would still need to be analyzed separately.

The probability of apprehension in the 1980s and early 1990s ranged between 30 percent and 40 percent (Massey and Singer, 1995). The Mexican Migration Project (MMP), which is described in Chapter 3, also publishes updated estimates of apprehension probabilities.[21] While the probabilities estimated by the MMP are somewhat lower than those presented by Massey and Singer—around 20-25 percent for the 1980s, dipping to 15-20 percent in the early 1990s—the MMP estimates indicate a sizable subsequent rise in the probability of apprehension to levels between 30 and 40 percent during the 2000s. It is difficult to determine if the increase is related to the changing composition of communities in the MMP data or to real changes at the border. If one assumes that the trend is a result of the latter,[22] then the probability of apprehension during the last decade was around one in three, a relatively low number.

It is also worth noting that increased enforcement during the past decade in the U.S. interior has worsened the labor market and living conditions of undocumented immigrants in the United States (Orrenius and Zavodny, 2009) and may have a deterrent effect on attempted crossings (Wein et al., 2009). However, scholarship on the subject has mostly been devoted to understanding how interior enforcement affects immigration to particular localities. For instance, Parrado (2012) finds that the 287(g) program[23] has only been effective in deterring immigration at a few large

[20]Enforcement can still affect the selection of people who cross, regardless of whether it affects the total number of people who cross successfully (Lozano and Lopez, 2010; Orrenius and Zavodny, 2005).

[21]See http://mmp.opr.princeton.edu/results/008apprehension-en.aspx (April 13, 2012).

[22]The step decline in the probability of apprehension between 2010 and 2011 is most likely due to a drop in the number of MMP communities available for these analyses.

[23]The Section 287(g) program was established in 1996 but implemented primarily after 2005. Under this program, state and local law enforcement agents receive DHS training and supervision to check the immigration status of arrestees in jails and prisons and to apprehend suspected unauthorized immigrants through traffic stops or other community interventions (National Research Council, 2011:42-43).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

immigrant gateways: it has not been effective in stemming immigration into most cities where the program has been implemented.

THE PROCESS OF ATTEMPTING TO CROSS THE BORDER WITHOUT INSPECTION

Even before the buildup of border enforcement, migrants would travel to the border region and attempt to cross, generally in groups and typically assisted by a guide or smuggler, known as a *coyote* or *pollero*. Smugglers typically require partial payment up front; full payment is due once the migrant is delivered to awaiting friends or relatives in the United States. Fees for smuggling services often allow for multiple attempts to cross the border, in the event that a migrant is apprehended by USBP one or more times. The share of migrants using smugglers has increased from about 80 percent in 1990 to roughly 90 percent today.[24] Smuggler prices have risen in inflation-adjusted terms from about $600-$1,000 in 1990 to $2,500 in 2010 (Roberts et al., 2010). Prices also vary by region and mode: anecdotal accounts indicate that crossing into the United States by boat in the San Diego sector is about two to three times as expensive as crossing by land, which suggests that the probability of apprehension during unauthorized maritime crossings is low.

Figure 2-3 shows a stylized version of the undocumented border crossing process, along with the sources of survey and administrative data (which are described and evaluated in Chapters 3, 4, and 5) that can potentially and partially measure these flows. The process starts in the home community, where people decide to travel to the border region and attempt an unauthorized crossing. In some cases, individuals may use border-crossing cards or visas to cross and then work in the United States (Chávez, 2011), although this is less common for Mexicans (particularly those who do not reside in the border region)[25] than for people of other nationalities (Massey and Riosmena, 2010:304-305).

As noted above, a majority of migrants use the services of smugglers, especially those attempting to cross for the first time (López Castro, 1997; Orrenius, 1999; Roberts et al., 2010:4; Spener, 2009:81). Once a smuggler is hired, migrants attempt to cross through nonfenced, more remote areas

[24]Data on smuggling usage and price are from the Mexican Migration Project website at http://mmp.opr.princeton.edu/results/results-en.aspx (March 19, 2012).

[25]Using the same regional classification of Mexican states provided in Riosmena and Massey (2012), the panel's own calculations using 2006 and 2009 ENADID data, described in Chapter 3, suggest that migrants coming from the Mexican border region have considerably lower shares of the undocumented (35 percent and 29 percent) compared to migrants from the central-western (68 percent and 67 percent), central (86 percent and 79 percent), and southeast (85 percent and 79 percent) regions.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

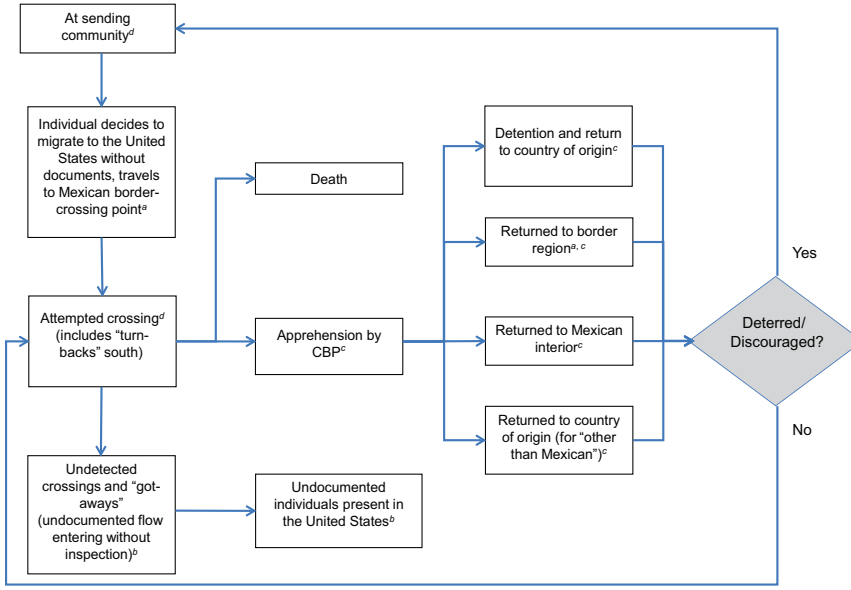


FIGURE 2-3 The undocumented border crossing process: Survey and administrative data sources.
NOTES: Data sets potentially (partially) measuring each component:
[a]EMIF-N, (MMP/MMFRP).
[b]ACS/CPS (indirectly), ENADID, ENOE, Mexican Census, MxFLS.
[c]CBP's ENFORCE database.
[d]MMP/MMFRP.

ACS:	American Community Survey
CPS:	Current Population Survey
EMIF-N:	Survey of Migration at the Northern Border of Mexico
ENADID:	National Survey of Population Dynamics
ENOE:	National Survey of Occupation and Employment
MMFRP:	Mexican Migration Field Research Program
MMP:	Mexican Migration Project
MxFLS:	Mexican Family Life Survey

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

or by climbing over the double fence in more transited areas (López Castro, 1997). If migrants are spotted while still close to the line, they generally choose to return to the Mexican side in order to attempt another crossing. This is what USBP calls a "turn-back south." USBP reportedly keeps records of the number of turn-backs south detected during a shift, but these have not been publicly released, nor does the panel know of any other data source that reports numbers of turn-backs south. In any case, such attempts are not individually identifiable and linkable to prior or subsequent attempts and apprehensions.

An undocumented crossing can result in three different outcomes (other than a turn-back south). First, some undocumented crossers succeed in eluding USBP after being detected; these are termed "got-aways" and become part of the unauthorized population, along with individuals who manage to cross undetected by USBP. If DHS could estimate the number of attempted crossings, they would then be able to indirectly estimate the number of migrants who cross undetected as the difference between the total number of attempts and the sum of turn-backs south, getaways, and apprehensions. The ratio of apprehensions to the total number of attempts is a proxy for the effectiveness of USBP in actually catching the migrants they encounter in the field, a measure DHS officials call the "interdiction rate."[26]

Death is another possible outcome of an illicit border crossing, typically brought about by exposure to extreme hot or cold temperatures and dehydration. Deaths at the border have increased (Cornelius, 2001; Eschbach et al., 1999) as more migrants have been crossing through desolate parts of the Arizona desert and remote parts of Texas, typically walking for several days in order to circumvent USBP checkpoints on the highways.

Finally, a crossing attempt can be stopped by USBP (or other federal law enforcement body), at which point the migrant is fingerprinted and photographed and his identity is run through an FBI database to check for prior criminal convictions. After the apprehending agent reviews the individual's migration history and checks for any prior apprehensions, the officer applies a "consequence" and the migrant is eventually returned to his homeland. Possible outcomes under consequence policies include voluntary return, a formal removal order, or criminal charges. The choice of the consequence depends on the migration and criminal history of the migrant, as well as on applicable USBP policies in effect in the vicinity of the apprehension.[27] Depending on the consequence applied to the migrants, they

[26]We did not have access to these measures or to data that would allow us to calculate these rates, and we know no published source that shows estimates of them.

[27]This description is based on the discussion of enforcement pipelines in National Research Council (2011:Chapter 4).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

can be returned to the closest border point, a distant border point (through lateral repatriation), or the Mexican interior. After release, they may be deterred from crossing again or may attempt re-entry, in which case they may succeed (i.e., get away), be apprehended, or die. Since around 2000, USBP has kept better records of individuals processed at its stations, identifying them through biometric measures such as fingerprints. These records are stored in the ENFORCE database, which we describe in Chapter 5.

The Role of Drug Cartels

Human and drug smuggling have typically been separate businesses along the U.S.–Mexico border. With severe penalties for drug smuggling and relatively light penalties for entering without inspection, migrants and their *coyotes* have had strong incentives to avoid carrying drugs for traffickers (Spener, 2010). Existing evidence also suggests that most human smugglers on the U.S.–Mexico border are not affiliated with organized crime (Fuentes and García, 2009; Izcara Palacios, 2012; Spener, 2004; United Nations Office on Drugs and Crime, 2010). Mexicans tend to rely on smugglers who they know personally or who they know of through their social networks. It is relatively easy for anyone with some border crossing knowledge and experience to guide another person across the border and charge money for this service. The border area is also geographically large and varied, and it would be difficult for criminal syndicates to guard the myriad crossing places in order to enforce control over all independent operators.[28]

While organized crime may not be dominating human smuggling on the U.S.–Mexico border, it appears to be playing an increasingly important role over time. According to federal authorities, the role of organized crime in human smuggling has increased (National Gang Intelligence Center, 2011). As Mexican drug cartels become more powerful, display an increasing capacity for violence, and are subject to heightened pressures from U.S. and Mexican enforcement, migrant smugglers may find it increasingly necessary to negotiate agreements with local "plaza bosses" who control specific segments of the border. The growing role of organized crime could also have something to do with the rising numbers of non-Mexican border crossers (United Nations Office on Drugs and Crime, 2010). Non-Mexicans have limited access to migrant networks and so may be more likely than Mexi-

[28]During their visit to the southwest U.S. border, panel members were given the impression by USBP that drug cartels currently play a prominent role, if not a dominant role, in regulating the flow of unauthorized migrants across the border. This is at odds with what is in the research literature. In this regard, the literature may be somewhat dated—although USBP's perceptions of the importance of cartels may also be exaggerated.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

cans to contract with a criminal syndicate that offers smuggling services. Non-Mexicans may require assistance traveling illegally through the interior of Mexico, another service more likely to be provided by transnational cartels than by local small-scale smugglers in Mexican sending villages or at the U.S.–Mexico border.

CONCLUSION

The decision to engage in unauthorized migration is highly complex, as more formal models by Chang and colleagues (2012), Guzman and colleagues (2008), MITRE Corporation (2008), and Wein and colleagues (2009) portray. It is influenced by economic, social, and environmental conditions in sending and destination areas; immigration policies; and interior and border enforcement. Furthermore, the border itself can be seen as a "system" of its own, in which enforcement efforts and policy changes can have spillover effects and generate adaptive responses by unauthorized migrants and others. As can be seen in Figure 2-3, none of the existing survey and administrative data sources captures the entire migration process.

All of these factors have evolved over the past 50 years, as have the numbers and profiles of migrants. Crossing routes, modes, and the roles of smugglers (who are separate from, but may be increasingly connected to, organized crime and drug cartels) have also changed considerably in recent times. Survey designs and modeling approaches that may be well suited for capturing certain aspects of unauthorized migration flows today may not be able to do so with the same reliability and validity in the future. The methods and assumptions used to estimate flows should be accordingly flexible.

- **Conclusion 2.1: To understand migration flows in any one sector, it is important to view the entire border as a system; localized increases in border enforcement may simply change where migrants cross without reducing the overall flow in the long run.**

- **Conclusion 2.2: The migration process is complex and dynamic. Undocumented migration is the outcome of many interrelated factors that can vary widely across people, space, and time; migrant characteristics and the geography of sending and destination areas are changing constantly. This complexity and dynamism should be incorporated into the analytical approaches and study designs used to estimate flows of unauthorized migrants.**

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

3

Migration-Relevant Surveys in the United States and Mexico: Background

Surveys are studies conducted to obtain information on characteristics of a population. Well-designed surveys based on probability sampling can provide valuable information about a large population using just a small representative sample of subjects. Because sample surveys are less expensive than conducting a full census of the population of interest, they are quite attractive as an option for obtaining population measurements. A number of sample surveys that include questions on migration are conducted in both the United States and Mexico; these surveys are discussed in this chapter.

The theory of sampling begins with a target population of observations to be studied: for example, households or individuals about which inferences will be made. The sampling frame is a list of sampling units from which the sample to be actually observed is drawn. Probability sampling assumes that each unit in the population has a known, nonzero chance of selection into the sample, and a chance method is used to select the units in the sample. Estimates for the entire population of interest and measures of the uncertainty (or variance) in these estimates can be computed from the sample data and the sampling design.

The ideal of probability sampling is to construct a perfect frame for the target population and then select a sample from that frame in accordance with the probability sampling design. The survey then observes the true value for each study variable on each unit selected into the sample (the data collection) and processes the data without introducing errors. In practice, most surveys fall short of these ideal conditions. For example, in many surveys, not all of the selected sample units respond. In addition, coverage error exists since the sampling frame may not have included part

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

of the target population. Statistical inferences are affected by nonsampling errors, such as coverage, measurement, and non-response errors. Survey researchers often use weighting procedures to account for the features of the sampling design and make adjustments for non-response or coverage error.

For the purposes of this report, it is important to determine how well the existing surveys are designed, in order to assess their usefulness. The surveys described are conducted in the United States, in Mexico, or both, and the panel evaluated them to determine their effectiveness in estimating both the number of foreign nationals who attempt illegal entry across the U.S.–Mexico land border and the probability of apprehension of illegal entrants. Our evaluation also addresses the possibility of obtaining these estimates on a quarterly basis for particular regions of the U.S.–Mexico border.

This chapter outlines the features of the major surveys in the United States and Mexico (summarized in Table 3-1) that collect information about migration and border crossing; details about some of the specific questions asked by the surveys are provided in Appendix A. Chapter 4 discusses the usefulness and limitations of these surveys in estimating the number of foreign nationals who attempt illegal entry across the U.S.–Mexico land border.

AMERICAN COMMUNITY SURVEY

The American Community Survey (ACS) conducted by the U.S. Census Bureau is intended to collect data comparable with the Census 2000 long-form sample data, but to do so every year rather than every 10 years (Grieco and Rytina, 2011). Information from the ACS is used to administer federal and state programs and distribute federal funds. The survey asks about age, gender, race, family and relationships, income and benefits, health insurance, education, veteran status, disabilities, where people work and how they get there, where people live, and how much people pay for certain essentials. Fully implemented since 2005, the ACS collects annual data in twelve monthly samples. The target population of the ACS is the entire resident population of the United States and Puerto Rico. The sampling frame reflects this target population by identifying all addresses of households for the 2005 ACS and all addresses of both households and group quarters for the ACS since 2006. Data collection uses three modes that take place over a 3-month period: mail, telephone, and personal visit. The target population for the ACS only includes people who are deemed to be residents of the United States; short-term migrants, such as many undocumented Mexican workers, would not be part of the ACS universe.

The 2010 Public-Use Sample, which is a 1 percent sample of the U.S. population, included 1,204,000 households and 3,062,000 respondents. Of these, 145,000 households were immigrant households (35,000 Mexi-

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

can and 8,400 Central American) and 348,000 respondents were foreign-born (92,900 Mexican and 22,200 Central American). Each year's ACS is weighted to the current population estimates for that year with a weighting methodology that controls for small areas, race/Hispanic origin, age, gender, and marital status/households (see U.S. Census Bureau, 2009, for details). The annual ACS aggregates the 12 monthly samples to yield annual data. The U.S. Census Bureau does not release information about the dates interviews were conducted and, therefore, there is no specific "reference date" within the year for the survey.[1] Immigration-related questions asked by the ACS include nativity/citizenship (but not legal status), year of arrival in the United States,[2] country of birth, year of naturalization, residence 12 months before interview, and language spoken at home.

The U.S. Department of Homeland Security (DHS) uses the ACS to estimate the size of the unauthorized immigrant population residing in the United States (see Box 3-1). DHS has published annual estimates for 2006-2011 using the ACS data for 2005-2010 (Hoefer et al., 2011).

CURRENT POPULATION SURVEY

The CPS is the primary source of labor force statistics for the U.S. population and is also the source of high-profile economic statistics such as the national unemployment rate. The CPS, which is sponsored by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics, is administered monthly by the U.S. Census Bureau (Grieco and Rytina, 2011). This survey frequently includes specialized supplements, and the Annual Social and Economic Supplement (ASEC), conducted in February-April each year, has an expanded sample (a double sample of Hispanic households and an oversample of other minority households and households with children). Although monthly data are generally limited to labor force items and education, a broad range of social, economic, and demographic data are collected in the CPS ASEC (formerly known as the "March supplement").

The target population of the CPS is the civilian non-institutional population living in the United States. The sampling frame is a list of housing addresses obtained from the most recent decennial census and updated with new housing units built after the census. The CPS ASEC also includes military personnel living in off-base housing with civilian adult household members. In 2011, the average monthly sample (from January to November) in the public-use sample included 54,000 households and

[1]The ACS is weighted using population estimates for July 1 of the survey year.

[2]The question about year of arrival can be subject to various interpretations for circular migrants, and estimates of migration based on year of arrival tend to be higher than those based on residence 1 year ago.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

TABLE 3-1 Overview of Migration-Relevant Surveys in the United States and Mexico

	ACS	CPS (ASEC)	Mexican Census (10% sample –long form)	ENOE
Probabilistic sample	Yes	Yes	Yes	Yes
Target population	U.S. households	U.S. households	Mexican households	Mexican households
Frequency of survey	Annual	Annual	Every 10 years	Quarterly
Mexico and U.S. samples	No	No	No	No
Total sample size	~1,204,000 households	~97,000 households per month	~2.9 million dwellings	120,260 dwellings per quarter
INCLUDES QUESTIONS ABOUT:				
Documentation status at crossing	No	No	No	No
Crossing locations	No	No	No	No
Number of attempts	No	No	No	No
Smuggler use	No	No	No	No
Reasons for migrating	No	No	No	Yes
Intention to migrate	No	No	No	Yes[a]

[a]Asked of those not currently working.

[b]Weights are estimated; due to the nature of the survey design, their accuracy is difficult to quantify.

[c]EMIF-N is conducted continuously, and, beginning with 2012, data are released on a quarterly basis. From 2013 on, the intent is to reduce the delay in reporting each quarter to 2 months.

[d]Northbound sample.

[e]Asked on U.S. side only.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

ENADID	EMIF-N	MxFLS	MMP	MMFRP
Yes	Yes[b]	Yes	No	No
Mexican households	Migrants passing through Mexican border cities	Mexican households	Sending communities in Mexico	Sending communities in Mexico
1992, 1997, 2006, 2009	Quarterly[c]	2002, 2005, 2009	Annual	Annual
No	No	Yes	Yes	Yes
~40,000-100,000 dwellings	~14,000 respondents[d]	~35,000-40,000 respondents	~600-1,000 households	~700-1,000 respondents
Yes	Yes	Yes	Yes	Yes
No	Yes	Yes	Yes	Yes
No	Yes	Yes[e]	Yes	Yes
No	Yes	Yes	Yes	Yes
No	Yes	Yes	No	Yes
No	Yes	Yes	No	Yes

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

BOX 3-1
Residual Estimates of the Unauthorized Immigrant Population

The residual methodology, which subtracts estimates of the legally resident foreign-born population from estimates of the total foreign-born population, is used to estimate the size of the unauthorized immigrant population residing in the United States. Warren and Passel (1987) applied residual methods using information from the 1980 Census. Edmonston and colleagues (1990) applied similar methods to data from the Current Population Survey (CPS)[a] in 1983 and 1986 and to data from the 1990 Census. With the regular collection of data on the foreign-born population in the CPS, beginning in 1994, more frequent estimates based on the residual method appeared (e.g., Passel and Clark, 1998; U.S. Immigration and Naturalization Service, 2001; Warren, 1994 [cited in U.S. Immigration and Naturalization Service, 2001]). An early residual-type estimate developed by Bean and colleagues (1997) for the U.S. Commission on Immigration Reform combined data from the 1996 CPS March supplement with data from Mexico to estimate the number of unauthorized Mexicans in the United States (with a correction for CPS omissions). Annual estimates of the unauthorized immigrant population have been published by DHS for 2006-2011 using the ACS (Hoefer et al., 2012) and by the Pew Hispanic Center for 2000-2010 using the CPS (Passel and Cohn, 2011).

Most residual estimates of the unauthorized immigrant population take the same broad approach:

1. An estimate of the number of legal foreign-born residents in the United States is developed using administrative data on legal admissions across a number of years. The data are combined using demographic techniques to allow for mortality and emigration after admission to the United States.
2. A survey (ACS, CPS, or Census) is used to generate tabulations of the number of immigrants found in the survey.
3. An initial estimate of the number of unauthorized immigrants appearing in the survey is derived by subtracting the results of step (1) from the results of step (2).
4. A final estimate of the total number of unauthorized immigrants in the country is derived by adjusting the results of step (3) for undercount in the survey.

The various estimates using this approach differ in terms of their specificity (e.g., countries of origin, states, and periods of entry) and in the assumptions made about various parameters of the estimation process. The DHS and Pew

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

estimates both assume that all immigrants who entered the United States before 1980 are legally present, so the construction of a legal foreign-born population estimate uses only administrative data on admissions for 1980 and later.

There are three main areas where the existing residual estimates rely on data and assumptions for which strong empirical support is lacking and which can affect the overall magnitude of the estimates. First, the United States lacks data on departures from the country, so emigration of legal residents must be estimated using a variety of methods. Second, the DHS and Pew estimates differ in how they handle legal temporary migrants with longer-term visas who appear in the ACS and CPS (such as foreign students and intracompany transfers). Third, firm estimates of coverage of unauthorized and legal immigrants in surveys such as the ACS and CPS are also lacking, so both estimates rely on information drawn from local surveys and on assumptions based on various external sources.

To place the sensitivity of the residual estimates in context, the average annual growth in the unauthorized population between 2000 and 2005 was more than 400,000 according to DHS estimates (Hoefer et al., 2011) and more than 500,000 according to the Pew estimates (Passel and Cohn, 2011). But both sets of estimates show peaks in 2007 followed by declines, so that the net annual growth between 2005 and 2010 was less than 100,000. Annual estimates of emigration of post-1980 legal immigrants in recent years have exceeded 200,000, and the estimated cumulative emigration since 1980 amounts to 3.6 million (Hoefer et al., 2011). So, relatively small variations in this hard-to-measure component could have a significant impact on the resulting estimate of unauthorized immigrants. The number of legal temporary migrants living in the United States in 2010 according to DHS estimates is 1.8 million (Hoefer et al., 2011), yet the Pew estimates find evidence in the CPS for fewer than 1 million.

Both the DHS and Pew estimates allow for undercount of legal and unauthorized immigrants amounting to about 1.5 million. Yet, these assumptions rely principally on a relatively small study done in Los Angeles after the 2000 Census (Marcelli and Ong, 2002). Clearly, better information on coverage of immigrants in the ACS and CPS would improve the precision of the residual estimates and their face validity. The importance of understanding survey coverage of immigrants has long been recognized, and the U.S. Government Accountability Office (GAO) has recommended that the U.S. Census Bureau devote time and resources to the topic (U.S. General Accounting Office, 1998). Little official government work has been done to fill this vital information gap, and the GAO recommendation is considered "open."

---

[a]Described below.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

133,900 respondents; 6,400 households were immigrant (1,600 Mexican and 440 Central American), and 14,600 respondents were foreign-born (4,000 Mexican and 1,100 Central American). The 2011 CPS ASEC sample included about 97,000 households and 205,000 respondents.

Each year's CPS is weighted to the year's population estimates generated in the previous year. The weighting methodology has final controls for states by broad age group (and race,[3] depending on sample size); national data by age, gender, and race; and national data by age, gender, and Hispanic origin. The CPS ASEC has an additional weighting step to ensure that both spouses (or both unmarried partners) have equal weights.

Immigration-related questions asked by the CPS include nativity/citizenship (but not legal status), period of arrival in the United States (released in public-use data at 2-3 year intervals),[4] country of birth, residence 12 months before interview (ASEC only), and country of birth of parents.[5] Due to the small sample size of the CPS, there are limitations for measuring year-to-year changes and date of arrival disaggregation.

The Pew Hispanic Center has used the CPS to estimate the size of the unauthorized immigrant population residing in the United States for 2000-2011.[6] It does so using a residual methodology (see Box 3-1). Using additional data (such as survey information on date of arrival) and assumptions, inflows and outflows of unauthorized immigrants can also be estimated from the ACS- or CPS-based series of annual population estimates (Passel and Cohn 2009a, 2010); such estimates are extremely sensitive to underlying assumptions, as discussed in Box 3-1.

MEXICAN CENSUS OF HOUSING AND POPULATION (LONG AND SHORT QUESTIONNAIRES)

The Mexican Census of Housing and Population is funded by the Government of Mexico, with data collected, processed, and made publicly available by the National Institute of Statistics and Geography (known in Spanish as Instituto Nacional de Estadística y Geografía [INEGI]). This census is conducted every 10 years, at the beginning of each decade. The

[3]With categories for blacks and all others.

[4]As with the ACS, the question about year of arrival can be subject to various interpretations for circular migrants, and—even more so than in the ACS—CPS estimates of migration based on year of arrival are higher than estimates based on residence 1 year ago.

[5]The CPS does not have data on language spoken at home. The CPS has asked monthly questions about place of birth, parental place of birth, U.S. citizenship status, and year of entry into the United States since 1994 (Grieco and Rytina, 2011).

[6]The Pew Hispanic Center uses the CPS for historical reasons, and it plans to switch over to the ACS in the near future.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

target population is all individuals living in Mexico, while the sampling frame is a list of dwellings located in Mexico.

In both 2000 and 2010, two different questionnaires were administered. The short form is administered to the entire Mexican population and collects basic social-demographic information for all household members, as well as characteristics of the dwelling.[7] It includes information about the place of residence 5 years prior to the interview (i.e., for 2010, the location of residence in 2005 was asked),[8] which permits estimation of the number and basic social-demographic characteristics of the population that returned from the United States in the past 5 years.

A long questionnaire was administered to a random sample of about 10 percent of the total Mexican population in the 2000 and 2010 census. In 2010, more than 2.9 million dwellings were interviewed. Expansion weights[9] were estimated for each dwelling based on the sampling weight and on an adjustment factor accounting for differential response rates. The sample provides estimates reflecting the population at the following levels: national, state, state with four community sizes predefined, municipality, and localities with more than 50,000 inhabitants (Instituto Nacional de Estadística y Geografía, 2011b, 2011c).

The long questionnaire includes the questions listed in the short questionnaire, plus a special section on the international migration experiences of household members during the prior 5 years, including those no longer in the household. In addition to the question on place of residence 5 years ago, which allows estimation of return migration between 2005 and 2010, the long questionnaire includes a set of questions on international migration. These questions refer to the migration of any person who is currently living in the dwelling or who lived in the dwelling between 2005 and 2010. Information is collected on the number of people, gender, age at migration, date of last migration, state of residence at the time of the migration, country of destination, country of current residence, date of return (for those who returned), and whether the person is currently living in the same household (Instituto Nacional de Estadística y Geografía, 2011a:36-42). The questionnaire also asks whether the household is currently receiving remittances. The 2010 census sample includes 83,757 households with at

[7]The definition of "household" changed between the 2000 and the 2010 Census. In 2000, there could be more than one household per dwelling. In 2010, there was no question about the number of households in a dwelling. This change in the definition needs to be taken into account when comparing the international migration information from the two censuses.

[8]This question is asked for all of the dwellers ages 5 years and older.

[9]Expansion weights are sampling weights. The word "expansion" is used because it describes the expansion of that sampling unit to the population from which it was sampled.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

least one return migrant,[10] 31,536 households with at least one circular migrant,[11] and 89,601 households with at least one out-migrant.[12] In total, 186,456 households reported a migration experience, and 149,000 households reported that they were receiving remittances.

Using the questions in the long questionnaire, the following analyses are possible:

- Estimates of the flow of out-migration, circular migration, and return migration in the past 5 years nationally and by state of residence at the time of the migration (see, e.g., Giorguli and Gutiérrez, in press);
- Gender and age composition of the above flows;
- Estimates of the return rates and the social-demographic characteristics of return migrants (and of circular migrants if currently living in the same household); and
- Geographical profile of the intensity of migration by municipality (based on the proportion of households where at least one member has migrated and/or received remittances).

NATIONAL SURVEY OF OCCUPATION AND EMPLOYMENT (ENOE)

Mexico's National Survey of Occupation and Employment (known in Spanish as Encuesta Nacional de Ocupación y Empleo [ENOE]) is funded by the Government of Mexico. INEGI is responsible for collecting and processing the ENOE data and making them publicly available on its website. The main objective of this national probability survey is to capture short-term changes in the occupation and labor force situation of the working age population in Mexico. It also includes detailed information on employment characteristics, access to social benefits, hours worked per week, income, social-demographic characteristics, and residential status. The first round of ENOE was conducted in 2005, although antecedents of the survey date back to 1972.

The target population is the working age population of Mexico (individuals 12 years and older). The sampling frame is INEGI's National Households Frame 2002, which is based on cartographic and demo-

[10] "Return migrant" refers to a person born in Mexico who lived in the United States before 2005 and returned any time between 2005 and 2010.

[11] "Circular" migrant refers to a person born in Mexico who moved to the United States after 2005 and returned before 2010. There is some overlap between circular migrants and return migrants (Passel et al., 2012).

[12] "Out-migrant" refers to a person born in Mexico who left for the United States between 2005 and 2010 and has not returned.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

graphic information from the General Population and Housing Census 2000. ENOE collects information from approximately 120,000 dwellings every quarter and is based on a probabilistic, two-staged, stratified cluster sampling design. The sampling frame is stratified by socioeconomic status and is designed to reflect national, state, city, and community-size levels.[13] Households within dwellings are interviewed five times within a period of 15 months. Each quarter, 20 percent of the dwellings leave the sample after completing the fifth round and are replaced by new entries, which are randomly selected with unequal probabilities of selection from the stratified sampling frame. Expansion factors include a sampling weight, a weight to reflect differential non-response rates, and calibration weights to the official projections of the population for a given year.

ENOE asks those not currently working whether they have tried to look for a job in another country or whether they are preparing to cross the border. ENOE also has information on whether individuals receive remittances, though without specifying the amount and regularity. Other migration information is captured in the household roster, which asks about the residential status of all individuals living in the dwelling at the time of the prior round of interviews. If a person is no longer living in the household, the respondent is asked about the reasons why he/she moved out (work, study, health problems, family reasons, among others) and his/her current place of residence (the same state, another state, or outside of the country[14]). This information can be complemented with the basic social-demographic variables obtained in the prior interview. The household roster also captures information on new arrivals. Aside from the social-demographic profile of the new members of the household, it asks the reasons for moving in (similar to the responses for those moving out) and which state or country the respondent was leaving before arriving to the household. In the 2010 ENOE, an average of 239 households per round reported out-migrants and 185 households reported returned migrants moving in from outside of Mexico. An average of 367 households per round had either a new arrival or somebody leaving for another country. (As noted above, each round includes more than 120,000 dwellings.) Researchers state that ENOE can capture short-term changes[15] in Mexico–U.S. migration trends, and it has been used to estimate out-migration and in-migration flows from 2006 to the present (Bustos, 2011; Zenteno, 2011).

[13]The four size categories for communities are less than 2,500 inhabitants, 2,500 to 14,999 inhabitants, 15,000 to 99,999 inhabitants, and 100,000 or more inhabitants.

[14]ENOE does not specify the country of origin or destination of international migrants. Nevertheless, the vast majority of all international movements—above 95 percent—are known to occur between Mexico and the United States (Galindo and Ramos, 2009:114).

[15]That is, quarterly migration or any movement between the first and last round (15 months).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

NATIONAL SURVEY OF POPULATION DYNAMICS (ENADID)

Mexico's National Survey of Population Dynamics (known in Spanish as Encuesta Nacional de la Dinámica Demográfica [ENADID]) is funded by the National Council on Population (known in Spanish as Consejo Nacional de Población [CONAPO]) and INEGI. The purpose of ENADID is to provide information about the different components of population dynamics (e.g., fertility, mortality, internal and international migration). INEGI is responsible for questionnaire design, operational planning, survey execution, and data design, while CONAPO (along with a group of academic experts) participates in a conceptual review and validates the results before publication. ENADID was first fielded in 1992, with subsequent cross-sections in 1997, 2006, and 2009. In its origins, ENADID was one of the few national probability-based surveys that had instruments to measure international out- and in-migration.[16]

The target population of ENADID is the population permanently residing in private homes in Mexico. The sampling frame is INEGI's National Households Frame 2002, based on cartographic and demographic information from the General Population and Housing Census 2000. Data from each of ENADID's cross-sections come from samples aimed at reflecting characteristics of the whole country, each of the 31 states and the Federal District, and urban and rural areas (Instituto Nacional de Estadística y Geografía, undated-a, undated-b; Instituto Nacional de Salud Pública, 2008). Each cross-section of the ENADID aims to represent all household members at the time of the survey (including both Mexican- and foreign-born individuals) and all individuals who were members of the household roughly 5 years prior to the survey and who left the household to move to another country during this period, whether they returned to the sampled household or not.[17] Sample sizes for the ENADID surveys have ranged between 40,000 and 100,000 households.

ENADID includes questions to measure international out- and in-migration. In the case of in-migration, all of the ENADID cross-sections use the relatively conventional retrospective question asking if people living in the household were living in another state or country 5 years prior to

[16]Before 2000, the Mexican Census only measured in-migration; ENOE's predecessors only occasionally included migration questions; the Survey of Migration at the Northern Border (see next section) had only started in 1993 and was not continuous (and it also had a different scope and use than the ENADID in terms of understanding international migration flows); and more detailed, non-representative surveys like the Mexican Migration Project (MMP) were only starting and had (more) limited coverage and sample sizes.

[17]That is, they could remain in the United States or live elsewhere in Mexico, but only those moving abroad during the prior 5 years would be recorded in these questions.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

the survey.[18] For the measurement of out-migration, the main respondent of the survey is asked if any person who currently lives or previously lived in the household left for the United States about 5 years prior to the survey (the period of reference varies slightly by cross-section), the times of out-migration and (if applicable) return, and the basic social-demographic profile of the emigrants. In 2006 and 2009, the ENADID questionnaire also included items aimed at measuring the documentation or visa held by people at the time of out-migration, including residence permits (other than a green card), work permits, green card, tourist visa, student visa, U.S. citizenship, other, and no documentation.

ENADID has been used to study international migration, mostly to describe the general social-demographic profile of migrants (Canales, 1999; Durand and Massey, 2003; Durand et al., 2001; Riosmena and Massey, 2012). It has also been used to complement analyses using convenience samples, such as the MMP (McKenzie and Rapoport, 2007) and as a criterion to evaluate the validity of the MMP (Massey and Capoferro, 2004; Zenteno and Massey, 1999), the Survey of Migration at the Northern Border of Mexico (Rendall et al., 2009), and ENOE (for measuring return migration) (Rendall et al., 2011). Estimates of international out- and in-migration using ENADID (along with many other data sources) are provided in Galindo and Ramos (2008).

SURVEY OF MIGRATION AT THE NORTHERN BORDER OF MEXICO

The Survey of Migration at the Northern Border of Mexico is known by the acronym EMIF-Norte (or simply EMIF-N) for its Spanish name, Encuesta sobre Migración en la Frontera Norte de México. EMIF-N is a multistage probability sample of flows across the U.S.–Mexico border and has been used to estimate migration flows (Rendall et al., 2009). It is sponsored by the Mexican government's Ministry of Labor and Social Welfare, CONAPO, National Migration Institute (Instituto Nacional de Migración, INM), and Ministry of Foreign Relations. Funding has also been provided by the World Bank (Secretaría de Gobernación et al., 2010, 2011). The survey is carried out by El Colegio de la Frontera Norte (COLEF) and has been fielded annually since 1993. Microdata for the period from 1995 to the first quarter of 2012 are presently available on the COLEF website (http://www.colef.mx/emif) for public download. Data and documentation are also available on the CONAPO website (http://www.conapo.gob.mx/es/CONAPO/Encuestas).

[18]For a discussion of the advantages and limitations of these questions, see Rogers and colleagues (2003).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

The EMIF-N survey design is based on methods originally developed by biologists for sampling migratory populations (Bustamante, 1998; Santibáñez, 1999) and is similar to the United Kingdom's International Passenger Survey (Economic and Social Data Service, 2008; Office for National Statistics, 2007). Like the International Passenger Survey, EMIF-N samples passengers at airports, bus depots, and train stations, but it also samples passengers at ports of entry, international bridges, and Mexican Customs inspection points (Bustamante et al., 1998; Consejo Nacional de Población, undated; Santibáñez 1997, 1999; Secretaría de Gobernación et al., 2010, 2011; Secretaría de Trabajo y Previsión Social, 1997, 1998, 1999).

EMIF-N is made up of four subsamples of migrants entering or leaving border cities: first, migrants entering from the interior (i.e., southern Mexico), who are identified by the survey as intending to cross the U.S.–Mexico border; second, apprehended migrants returned to the Mexican side of the border by U.S. authorities; third, migrants leaving the U.S. border region for the Mexican interior; and fourth, migrants returning to the interior from northern Mexican border cities (not of interest to the current study). Within each of the specific localities, multi-stage probability sampling is used and incorporates geographic and temporal stages of selection (Secretaría de Gobernación et al., 2010, 2011).

The target population is all individuals 15 and older: (1) who are not residents in a border city or the United States, not born in the United States, and arriving in a Mexican border region;[19] and (2) whose travel is due to a job or job search; change of residence; being in transit to the United States; or other reasons such as study, tourism, or visiting family or friends; and (3) who have neither a fixed return date nor employment in their place of origin. The sampling frame is a list of all zones (such as airports and bus stations) and points (such as airport gates and entries to bus terminals from unloading areas) identified by EMIF-N investigators in preselected border cities.

In the data collected in 2008, the size of the subsample of migrants from the interior (i.e., southern Mexico) was 13,792, of which 8,075 reported intentions of crossing the border within the next 7 days. The size of the subsample of migrants returning from the United States was 7,729, and the size of the subsample of detained migrants returned by DHS was 6,989. Response dispositions for each of the four subsamples are not clearly documented, although information on refusers is collected. EMIF-N's weights currently presume that refusers are nonmigrant travelers; the weights are not adjusted for survey non-response. Differential non-response is therefore a potential source of bias in EMIF-N.

[19]With the exception of the subsample of migrants returning voluntarily from the United States, which includes migrants reporting that they reside in the United States.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

EMIF-N has an adaptive time-location sampling design. The set of locations in the sampling frame evolves in response to the EMIF-N team's perception of changes in the geographical and temporal distributions of migratory flows, with the intent of covering the total flow through the border region. Fieldwork prior to the first (1993-1994) wave of EMIF-N indicated that 23 border localities (cities or towns) "constituted practically the [entire] universe of crossing locations for the labor flows to and from the United States" (Secretaría de Trabajo y Previsión Social, 1998:20).[20] The first wave surveyed in 18 of these localities and found that 8 cities accounted for 94 percent of the migrant flows between the United States and Mexico (Secretaría de Trabajo y Previsión Social, 1997, 1998). Although the sampling frame is continually updated, these eight cities (Cuidad Juárez, Matamoros, Mexicali, Nogales, Nuevo Laredo, Piedras Negras, Reynosa, and Tijuana) have been surveyed in every year of EMIF-N's subsamples of northbound and southbound migrants.

Qualitative fieldwork is carried out on an ongoing basis using informants in localities (i.e., cities or towns in Mexico along the U.S.–Mexico border) who are knowledgeable about trends and geographic shifts in migrant flows. Based on this research, the sampling frame of localities (and sampling zones and points within localities) is updated annually to adjust for shifts in flows and to maintain high levels of coverage.[21] Documentation of the localities sampled in each wave are available on the COLEF website for EMIF-N (http://www.colef.mx/emif).

Within localities included in the sampling frame, EMIF-N survey workers sample at zones (such as bus and train stations and airports) and at points within these zones through which migrants must pass (such as airport gangway doors or the points at which passengers disembark from buses). The probability of sampling a particular point and zone is based on estimates of the share of migrant flows passing through that zone during a particular time period (usually one of three 8-hour shifts each 24 hours). Since at least 2009, these estimates of the share of flows of migrants through sampling zones are based on periods during which EMIF researchers observe each sampling zone simultaneously for 24 hours a day for 7 consecutive days, enumerating all people passing through the zone, and administering to as many as possible the filter used to distinguish migrants

[20]Translated from Spanish by the panel.

[21]In the 2012 wave, northbound migrants were surveyed in the eight cities listed above and four additional cities (Agua Prieta, Altar, Cananea, and Ciudad Acuña). The subsample of apprehended migrants returned by the U.S. immigration authorities has also been consistently surveyed in the same eight core cities, with the exception of Piedras Negras, which was dropped from the two most recent waves. The 2012 apprehended subsample also included migrants surveyed in Ciudad Acuña and San Luis Río Colorado.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

eligible for the survey from others, such as local travelers.[22] The information thus gathered is used on an annual basis to update the sampling frame with regard to the share of migrant flows passing through each zone during each time period. The sampling frame is further updated on a quarterly basis based on less exhaustive ongoing fieldwork identifying shifts in migrant flows between years (Secretaría de Gobernación et al., 2011).

Full updates (based on 24-hour observations on 7 consecutive days) of the sampling frame have been carried out in 2009, 2011, and 2012. In each of these years, the update observations were carried out in the first quarter (during February or March) in order to measure flows during seasonal high levels of migration from Mexico to the United States. In 2013, EMIF-N investigators plan to extend these intensive week-long enumerations of flows to 20 localities, which will include all currently sampled localities plus those included in the first wave and dropped in the second wave. This plan will allow for updated information on the share of flows passing through the localities currently in the sampling frame.

The EMIF-N investigators have made an operational decision to focus resources on sampling as exhaustively as possible within localities in the sampling frame, but to exclude from the frame those localities that account for very small shares of migrant flows. The EMIF-N weights do not attempt to inflate estimates to account for the small share of migrants passing through localities (and within localities' points and zones) not in the frame. Strictly speaking, estimates from EMIF-N using the weights produced by the EMIF-N team are estimates of the flows passing through sampling zones used in the survey, although, as noted, much effort is expended to make sure that sampling localities, zones, and points are selected and updated to adjust for any shifts in migrant flows and to maintain a high level of coverage for the population that migrates through a city. The design implicitly assumes that the flows through the localities not in the sampling frame at a given time are zero. Thus, the weights for the localities at times they are not in the sampling frame are assumed to be zero.

The sampling localities and sites for the EMIF-N subsample of apprehended migrants returned to Mexico by the U.S. authorities are selected based on data that the EMIF-N researchers receive from the Mexican immigration authorities (INM). For sampling frame updating, data are used on apprehended migrants returned during the same week as EMIF-N's enumeration activities. In the past, the INM annual data on apprehended migrants returned have reported fewer migrants than data reported by the U.S. immigration authorities. INM data count only returned migrants

[22]These estimates are based partially on direct observation (the periodic process whereby EMIF-N researchers observe each sampling zone continuously for a week) and partially on records of the operating hours and managers' records for the zones in the sample.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

received by INM from U.S. authorities under the framework of binational agreements. Differences between U.S. and Mexican numbers can be attributed to some combination of error in the statistical systems of one or both countries and "deficiencies and omissions in the application of the relevant binational procedures and protocols" (Sandoval et al., 2011:5).[23] At present, EMIF-N researchers adjust their survey weights for migrants from localities based on INM counts.[24] Lack of information sharing and coordination between DHS and INM can cause INM to undercount the flow of returned apprehended migrants and, therefore, downwardly bias EMIF-N's estimates of such migrant flows.[25]

All subsamples gather information on the demographic characteristics and educational attainment of the respondent. For respondents intending to cross, there are survey items asking about the reason for the trip, the possession of valid documents, the location (city) of planned crossing, the reason for selecting this crossing site, the intended U.S. destination, whether the respondent has already arranged employment in the United States, the intended sector of employment, how long the migrant expects to stay in the United States, whether a smuggler has or will be hired, where the smuggler was/will be hired, and how much the migrant expects to pay the smuggler. There are also survey items regarding the number of previous trips to the United States and details of the most recent trip, including questions regarding previous use of a smuggler to cross, Mexican city of crossing, U.S. city in which respondent stayed the longest, whether the respondent had family or friends in that U.S. city, whether family or friends provided assistance, and employment and earnings in the United States on the last trip.

Similar questions are asked of migrants returning from the United States. The subsample of migrants returned by U.S. immigration authorities corresponds to DHS administrative data on apprehended aliens but provides a richer detail of respondents' characteristics, history, and crossing experience. A question on how many times they had been captured by the U.S. Border Patrol and returned to Mexico allows the calculation of a probability of apprehension based on a repeated trials model (discussed in more detail in Chapter 5). EMIF-N has been used to estimate elements of unauthorized migration (Escobar-Latapí, 1999; Reyes et al., 2002), exam-

[23]Translated from Spanish by the panel.

[24]If the INM counts for returnees for a given locality are not within the 95 percent confidence interval of the EMIF-N estimate (based on the original weights), the EMIF-N weights are post-stratified to the INM counts.

[25]In addition to the adjustment of weights, EMIF-N researchers told panel members of a situation in 2010 in which DHS began returning migrants to San Luis Rio Colorado, a locality not previously used for such returns. Because the EMIF-N team only learned of this shift after it began and did not have operations in place in San Luis Rio Colorado, 3 months passed before the EMIF-N was able to begin collecting data at the new site.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

ine routes of unauthorized migration (Anguiano Téllez, 2007), study legal status as a factor in migrant remittances (Amuedo-Dorantes and Pozo, 2005), and look at differences in wages by temporary/permanent status among unauthorized immigrants (Brownell, 2010).

MEXICAN FAMILY LIFE SURVEY

The Mexican Family Life Survey (MxFLS) is a longitudinal survey of individuals, households, families, and communities in Mexico. The survey is jointly administered by Luis Rubalcava (Centro de Investigaciones Demograficas y Economicas, Mexico), Graciela Teurel (Universidad Iberoamericana, Mexico), and Duncan Thomas (Duke University), with funding from the Ford Foundation and other public and private sources in Mexico and the United States. The main objective of the survey is to gather information on the dynamics of socioeconomic indicators, demographic and health indicators, and population migration over a period of at least 10 years (see, e.g., Genoni et al., 2011). The MxFLS is the first Mexican survey with a longitudinal structure that attempts to track participants regardless of migration decisions. The baseline sample of participants selected in 2002 was a probability sample selected from the national population in Mexico at that time.

The MxFLS consists—thus far—of three waves. The first wave (MxFLS-1) or baseline was administered in 2002. The target population was all households in Mexico in 2002. The sample for MxFLS-1 was selected by INEGI and was a multi-stage, stratified probability sample of the Mexican population in 2002. The baseline sample consisted of approximately 35,000 individuals living in approximately 8,440 households.

The second wave of the survey (MxFLS-2), which was conducted in 2005 and 2006, was designed to follow all baseline participants and their children born after 2002 ("panel respondents"). In addition, all other people living with panel respondents at the time of the re-interview were included in the survey. The re-interview rate was about 90 percent. After adding new children and co-residents, the sample in MxFLS-2 included approximately 37,000 respondents. A salient aspect of the MxFLS is that it attempts to follow panel respondents regardless of their migration decisions. Accordingly, survey administrators tried to contact those panel respondents thought to have migrated to the United States between waves 1 and 2. Of those, 91 percent (or 774 people) were located and re-interviewed in the United States.

The third wave of the survey (MxFLS-3) went into the field in 2009 and was still ongoing as of spring 2012. The intent is to re-interview as many of the original panel respondents as possible. To date, the re-interview rate is approximately 85 percent; after including children born after the first

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

re-interview (MxFLS-2), the current number of respondents in MxFLS-3 is approximately 38,000. Attempts to track down all panel respondents who have not yet been located continue. MxFLS-3 participants include individuals who never left Mexico, those who migrated to the United States before MxFLS-2 and are still in the United States, those who migrated to the United States after MxFLS-2, and those who migrated to the United States sometime between MxFLS-1 and MxFLS-3 but have since decided to return to Mexico.

A unique aspect of the MxFLS is that it gathers the same type of social-demographic information (including health, reproductive decisions, education, and other data) on Mexicans who have stayed in Mexico and those who have decided to migrate. To date, about 1,179 panel respondents have been re-interviewed in the United States. It is estimated that they constitute about 80 percent of the panel respondents who migrated and still reside in the United States. Approximately half of the panel respondents who were in the United States at the time of MxFLS-2 had returned to Mexico by MxFLS-3.

Every panel respondent is asked several questions that attempt to capture migration history at the personal level. Respondents are asked about their place of birth and about their place of residence at age 12. After age 12, respondents are asked to provide more detailed information about moves within Mexico and between Mexico and other countries. For each move that lasts more than 1 year, respondents are asked for the migration date, destination (locality, municipality, state, and country), reason for leaving, people moving with the respondent, and source of funds. If the respondent moved to the United States, the type of documentation (none, visa, green card, citizenship) carried by the respondent is also recorded. For the 2-year period preceding a survey wave, respondents are asked to report the same information, but only for moves that last a month or longer. Finally, respondents are also asked prospective questions about their intentions to move. If they plan on moving sometime in the future, the likely destination, reason for the move, and the existing networks in their destination are recorded.

MEXICAN MIGRATION PROJECT

The MMP is a binational research effort that is codirected by Jorge Durand (Professor of Social Anthropology at the University of Guadalajara) and Douglas S. Massey (Professor of Sociology and Public Affairs at Princeton University).[26] Since its creation in 1982, the MMP has focused

[26]The description of MMP that follows is based largely on information from the public website http://mmp.opr.princeton.edu/research/design-en.aspx.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

on gathering social, demographic, and economic information on Mexican households and their U.S. migration experience. The MMP is supported by grants from the National Institute of Child Health and Human Development and the William and Flora Hewlett Foundation. MMP data are made available to the public through a website housed at Princeton University.

Each year during December and January (when seasonal and other migrants typically return home), the MMP randomly samples 200 households in several communities located throughout Mexico (for a total of 600 to 1,000 households interviewed per year). MMP communities are not randomly selected but are chosen to reflect a broad range of sizes, locations, urbanicity, and migration prevalence. Once communities have been chosen, frames are constructed by the enumeration of all households in smaller communities or all households in a neighborhood in larger metropolitan areas. Households are randomly selected within communities from among eligible households.

After gathering social, demographic, and economic information on the household and its members, interviewers collect basic information on each migrant's first and last trip to the United States. From household heads who have migrated in the past, they compile a year-by-year retrospective history of U.S. migration and administer a detailed series of questions about the last trip northward, focusing on employment, earnings, and use of U.S. social services.

Following completion of the Mexican surveys, U.S.-based samples are gathered in some cases. Using tips from community contacts where available, interviewers travel to destination areas in the United States the following summer to administer identical questionnaires to migrants from the same communities sampled in Mexico; the exact number of migrants sampled (typically 10-20) depends on how many, if any, can be identified. These are permanent migrants who have settled north of the border and no longer return home. U.S.-based sampling is intended to generate a binational sample that is representative of permanent and return migrants. However, while most of the MMP communities sampled before 2000 have a U.S. counterpart, the majority sampled in 2000 and later do not. Because the U.S. sample is very small and limited to certain years, the MMP disproportionately represents return migrants and underrepresents permanent emigrants.

The data include community-based weights that reflect the community population. In the Mexican sample, the weight is calculated as the inverse of the sampling fraction where the number of households interviewed is divided by the number of eligible households in the predefined survey area from which the 200 surveyed households were drawn. The U.S. weight is also the inverse of the sampling fraction; the population size is estimated by comparing the number of adult children among surveyed household heads

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

who settled in the United States with the number who stayed in Mexico. Once the data are weighted, they are representative only of the population in the predetermined sampled communities, not of Mexico or of any well-defined geographic area in Mexico.[27]

The MMP contains detailed information on U.S. migration, including timing and location of illegal border crossings; number of failed attempts; smuggler usage and cost; past trips; U.S. destination, occupation, wages, and social ties; and duration of trip. It includes data on nonmigrant households as well, which is useful in modeling the migration decision. MMP data have been thoroughly documented and are widely used in the migration literature (see, e.g., Durand and Massey, 2004), but they are almost never used to measure migration flows.

THE MEXICAN MIGRATION FIELD RESEARCH PROGRAM

The Mexican Migration Field Research Program (MMFRP) was established in 2004 at the Center for Comparative Immigration Studies (CCIS) at the University of California, San Diego, which remains a cosponsor. It has been funded through multi-year grants from the Ford Foundation, the University of California Office of the President, and smaller extramural grants. By following migrants in rural Mexico and the United States over time, the MMFRP seeks to document and explain changes in their migration and settlement behavior.

The MMFRP follows three small, rural Mexican communities and their U.S. satellite communities over time.[28] One community is interviewed per year and is then re-interviewed at two- or three-year intervals. Since 2004, the study has rotated between the three cities: San Miguel Tlacotepec, Oaxaca; Tlacuitapa, Jalisco; and Tunkas, Yucatan.[29] The three small communities of fewer than 3,000 people were chosen to be "broadly representative of high-emigration communities in west-central and southern Mexico" (Cornelius et al., 2009:x). Similar to the MMP, the MMFRP administers household surveys to residents in the small towns at times when seasonal migrants are most likely to be present, typically in the months of January or February. The U.S. part of the survey is administered using a "snowball" technique, where participants in Mexico give contact information for their friends and family in the United States so that they may also answer

[27]See *The Mexican Migration Project Weights* on the MMP website: http://mmp.opr.princeton.edu/databases/studydesign-en.aspx.

[28]One community, Las Animas, Zacatecas, was interviewed just once in 2005.

[29]Tlacuitapa was interviewed in 2005, 2007, and 2010; Tunkas in 2006, 2009, and 2012; and Tlacotepec in 2007-2008 and 2011.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

the questionnaire. Most satellite communities are in Southern California, although Oklahoma City is a top destination for migrants from Tlacuitapa.

The size of the combined Mexico- and U.S.-based sample is typically based on 700 to 1,000 interviews per year. The survey usually targets household heads, but it also interviews other adults (ages 15-65) in the household. Since a census is essentially taken within the community, there is no sampling error and hence there are no sample weights. Refusal rates tend to be less than 10 percent.

The focus of the interviews and, hence, the survey questionnaire changes every year. This allows the MMFRP to address the most recent developments, such as the impact of the U.S. recession on migration or the deterrent effect of border enforcement. Although the annual themes differ, the surveys consistently include information similar to the MMP on social, demographic, and economic variables. In addition, it asks about intended and actual migration, networks, remittances, perceptions of border enforcement and other immigration laws, and legal status. For the undocumented, it asks about how the border crossing was made, coyote usage and price, and any abuse or mistreatment, if applicable (for first and last trip, as in the MMP). The MMFRP also asks migrants why they return to Mexico or why they stay in the United States. Other information varies by survey year. At one time or another, questions have included welfare program usage in Mexico, race relations and ethnic identity, politics and civic participation, religion, future plans, and more.

The MMFRP data and documentation have recently been made available on the CCIS website. Researchers who participate in the MMFRP program in a given year write up their results as chapters in a book that focuses on that year's survey theme. (See, e.g., Cornelius et al., 2010, which examines how the U.S. economic recession that began in 2007 has affected flows of Mexican immigrants to and from the United States.[30])

[30]A list of MMFRP and other publications of CCIS is available at http://ccis.ucsd.edu/publications/books/.

Copyright © National Academy of Sciences. All rights reserved.

4

Migration-Relevant Surveys in the United States and Mexico: Usefulness and Limitations

This chapter discusses the usefulness and limitations of the surveys described in Chapter 3 for estimating the number of foreign nationals who attempt illegal entry across the U.S.–Mexico land border. The major criteria for evaluation are the nature of the target population and related issues of sample size and survey design, the frequency with which surveys are conducted and the speed with which data are made publicly available, and the types of questions that are asked about migration.

TARGET POPULATION

A probability survey is critical to drawing inferences to a population much larger than the number of individuals actually observed (i.e., questioned). The American Community Survey (ACS), Current Population Survey (CPS), Mexican Census (10 percent sample for the long form), Mexican National Survey of Occupation and Employment (ENOE), National Survey of Population Dynamics (ENADID), Mexican Family Life Survey (MxFLS), and Survey of Migration at the Northern Border of Mexico (EMIF-N) were designed as probability surveys. However, for the purposes of this study, each suffers from various limitations.

The ACS and CPS are U.S.-based surveys of U.S. residents rather than border crossers, and they cannot be used to directly estimate flows of unauthorized immigrants across the U.S.–Mexico border. Instead, inferences about the flows of unauthorized immigrants have to be made based on changes in the estimated stock of unauthorized residents in the United States. The estimates produced by these methods are necessarily imprecise

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

(see Box 3-1 in Chapter 3). Given their focus on issues of residence and intention to live in the United States, the ACS and CPS may also have problems covering people who have been in the United States for a short time (less than 1 year). They also appear to omit most of the seasonal workers, who usually live in Mexico and cross into the United States to work for a few months each year (a group that may account for a significant share, or even a majority, of unauthorized border crossers). Like all surveys tied to the decennial census, the ACS and CPS suffer from undercount. The undercount rate for unauthorized immigrants appears to be larger than for the rest of the population; estimates of the total unauthorized immigrant population based on the ACS (Hoefer et al., 2012) and the CPS (Passel and Cohn, 2011) make an adjustment for undercount in the range of 10 to 15 percent. Moreover, year-to-year comparisons have been complicated by the introduction of new population weighting methods in 2007 and 2008, redesigned questionnaires in 2008 (ACS only), and the switch to the 2010 Census as the base for weighting adjustments (in 2010 for the ACS and 2012 for the CPS).

The Mexican Census, ENOE, ENADID, and the MxFLS focus on Mexican households, a target population that is more relevant to this study than the U.S. households that are targeted by the ACS and CPS. One limitation of the Mexican Census, ENOE, and ENADID is that they miss entire households that have migrated,[1] thereby potentially underestimating flows from Mexico to the United States. ENOE will also miss whole households returning to Mexico because the migration information is only based on the second through fifth interviews; data from the 2010 Mexican Census suggest that about half of returning migrants return to households that did not exist prior to the return (Passel et al., 2012). The MxFLS, which tries to follow people when they move from Mexico to the United States, is less likely to miss the migration of entire households. However, only the MxFLS baseline sample (selected in 2002) reflects the national population in Mexico, and that sample is tied to the population at that time. Subsequent survey waves do not refresh the sample with new households and, hence, do not reflect the Mexican population at the time of data collection.

A more fundamental and general concern, however, has to do with the sample size of these traditional national household surveys. International migration is a relatively rare event, and it is important for a general survey to have a sample that is sufficiently large to obtain reliable information on

[1]In the ENOE 2010, for example, an average of 3.5 percent of households were declared *hogares mudados*, or households who moved out between rounds. The reasons may be residential change in the same locality, internal migration, or international migration; we do not have precise information on the nature of the geographic mobility associated with the *hogares mudados*.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

TABLE 4-1 Surveying Mexican Households with Migration Experience: Total Sample Sizes Required

Number of Mexican Households in Survey Sample with Migration Experience	Total Number of Mexican Households That Would Need to Be Sampled by the Survey
1,000	78,431
5,000	392,157
10,000	784,314
25,000	1,960,784
50,000	3,921,569
100,000	7,843,137

migration. Specifically, sample sizes should be sufficiently large so as to accurately detect relatively small changes in flow rates (i.e., by a few percentage points) associated with changes in enforcement policies, market forces, and other factors. Some simple calculations by the panel using information from ENOE illustrate the challenges at hand, both for existing surveys as well as any new ones that may be put in place to specifically address the migration question. Table 4-1 shows the total number of households across Mexico that would need to be sampled in any given time period (be it quarterly or yearly) in order to obtain a target number of sampled households with "migration experience" (i.e., having crossed, or intending to cross, the U.S.–Mexico land border). The panel made two assumptions. The first assumption is that roughly 1.5 percent of households in Mexico each year have an individual who crosses the border. This assumption is based on a recent per person out-migration rate in ENOE of 3.78 per 1,000 (0.00378 percent) (Instituto Nacional de Estadística y Geografía, 2012), with average household size being around four people. The second assumption, based on documentation material for ENOE (Instituto Nacional de Estadística y Geografía, 2007:48), is that the survey response rate is approximately 85 percent. The number of households that would need to be interviewed is equal to the target sample size divided by the product of the response rate and the household out-migration rate.[2] It is possible, in principle, to reduce sample sizes by oversampling in traditional Mexican "sending regions" or by otherwise using stratification or clustering based on what is known about the migration process to date. However, as discussed in Chapter 2, the sampling design would have to be adaptive to changing patterns of population migration, and strategies for oversampling could all too easily become out of date.

Total sample sizes would have to be even larger if one wanted precise flow estimates by, for example, each of the nine geographic sectors into

[2]For example, 1,000/(0.015 * 0.85) = 78,431.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

which the U.S. Border Patrol divides the southwest U.S. border. Assuming that the survey had information about crossing location—which ENOE currently does not—the survey design would have to capture flows from points of origin throughout Mexico to each of the geographic areas of interest at the U.S.–Mexico border. Pilot studies would very likely be needed to inform such a complex design, and the complexity of the design would make it all the more vulnerable to being rendered obsolete by changes in migration patterns.

There are different approaches to determining the benchmark sample size (represented in the left-hand column of Table 4-1). One approach consists in first deciding the magnitude of the change in migration flows that the U.S. Department of Homeland Security (DHS) would like to detect with a given probability. Another approach is based on the uncertainty that DHS is willing to accept surrounding estimates of, for example, the number of attempts by undocumented migrants by sector and by quarter.

To illustrate the type of calculations that could be carried out to estimate a benchmark for the number of sampled Mexican migrants, suppose that p is the true proportion of Mexican households with migration experience. The width (w) of the confidence interval around that proportion will be $2z[p(1-p)]^{0.5}n^{-0.5}$, from which the number of sampled households with migration experience (n) can be calculated for any given width and for any given level of confidence. Specifically, solving for n gives $n = 4z^2p(1-p)/w^2$. Narrower intervals correspond to a more precise estimate. Supposing that $p = 0.015$, and that one wishes to detect a 5 percent change in p with 95 percent probability (so that $z = 1.96$), $w = 0.0015$ (calculated as $0.015 \times 0.05 \times 2$),[3] so the sample would require approximately 105,000 households with migration experience in any given time period, be it quarterly or yearly (see Table 4-2). According to Table 4-1, this would require a nationwide sample of about 8 million households. Such a survey would be on a scale much larger than the ACS, which in 2009 had a $197 million appropriation for a final interview sample of approximately 1.9 million housing units (U.S. Census Bureau, 2008, 2012). As indicated by Table 4-2, wider confidence intervals (which imply less precise estimates for DHS's evaluation and operational purposes) require that fewer Mexican households with migration experience be sampled by the survey. Even so, the sample sizes would not be trivial—detecting a 15 percent change in p would require a sample size of approximately 12,000 households with migration experience, which in turn (according to Table 4-1) would require a nationwide sample of approximately 915,000 households.

[3]Supposing $p = 0.015$ and wanting to be able to detect a change (plus or minus) of 5 percent, one would want to see whether p goes down to 0.01425 or up to 0.01575. Therefore, $w = 0.015775 - 0.01425 = 0.0015$.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

TABLE 4-2 The Number of Mexican Households with Migration Experience That Needs to Be Sampled in Order to Detect (with 95 percent confidence) Changes in the True Proportion of Mexican Households with Migration Experience

Change in p	$4z^2p(1\text{-}p)$	w	w^2	n
5%	.2364	0.0015	0.00000225	105,067
10%	.2364	0.003	0.000009	26,267
15%	.2364	0.0045	0.00002025	11,674

NOTE: n = the number of Mexican households with migration experience that needs to be sampled; p = the true proportion of Mexican households with migration experience; w = width of the confidence interval around p; z = 1.96.

There are also other approaches to thinking about how to compute the benchmark sample size. In the preceding example, the problem could have been formulated in terms of "tolerance intervals" rather than confidence intervals (see, e.g., Krishnamoorthy and Mathew, 2009), which would have resulted in an even higher n (as tolerance intervals are similar to prediction intervals). Or, one could propose that two consecutive surveys be carried out to estimate the observed change in the number of households with migration experience in the intervening period. Regardless of how the problem is formulated, however, designing a Mexico-based survey of this size to interview households about their entries (and intended entries) across the U.S.–Mexico border would be complex, and the costs of administering and conducting such a survey would be very high. Response rates associated with a survey sponsored by a foreign government (such as the United States) and its immigration enforcement agency would likely be far lower than those currently associated with national surveys such as ENOE; the probability of erroneous response would also be higher. As discussed in Chapter 2, moreover, the survey design would need to be adaptive to changing patterns of population migration and to enforcement changes that influence non-apprehension rates. Although the actual details might vary somewhat by changing the assumptions, the order of magnitude and the complexity would not.

In contrast to the ACS, CPS, Mexican Census, ENOE, ENADID, and MxFLS, EMIF-N has a target population that is directly relevant to the estimation of unauthorized flows and that also includes significantly larger sample sizes of migrants compared to other annually collected data sources on both sides of the U.S.–Mexico border (Rendall et al., 2009:36). However, the panel notes that there are uncertainties surrounding the weighting methodology that is meant to ensure that the collected data reflect the entire Mexican population. Unlike traditional survey sampling designs, the

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

sampling design of EMIF-N is dynamic and adaptive. Since 1993, units in the sampling frame (i.e., cities, zones, and points) have been added and removed in response to perceived changes in the geographical and temporal distributions of migratory flows. Not every locality at Mexico's northern border is in the sampling frame of localities at every EMIF-N administration (survey wave), and the weighting assumes that all flows are through the cities in the sampling frame (and through no other cities) at the time of that survey wave.

The accuracy of the sample weights depends on the quality of the adaptability of the sampling frame, and this is difficult to quantify. Specifically, localities and transportation modes (e.g., private cars) not in the sampling frame at a given point in time are presumed to have zero flows. The survey will have coverage error if significant flows are missing in the sampling frame. This coverage error can be reduced by expanding the covered localities and modes of transportation. Although EMIF-N investigators believe the coverage to be between 90 and 95 percent of the flows, the size of the coverage error needs to be quantified, or at least bounded. In addition, the weights for a time-location design like EMIF-N are estimated from quantities collected during the survey and require careful adherence to the sampling protocol. In particular, it requires an accurate count of the total number of people passing through sampling points during the application of the survey. Another concern relates to possible deviations from the random selection of people at sampling points (due, for example, to traveling groups). This can be resolved by refining the sampling design. In addition, new statistical methodology could be developed to adjust for uncertainty in the weights due to deviations from the desired design.

Unlike the ACS, CPS, Mexican Census, ENOE, ENADID, and EMIF-N, the Mexican Migration Project (MMP) cannot be regarded as a probability survey for two reasons. First, although households are randomly selected within communities, the communities themselves are not randomly selected. Second, the additional companion sample collected in the United States introduces additional selection bias, as it is unclear who volunteers their relatives for the U.S.-based survey and how they might differ from other Mexican emigrants. Therefore, inferences cannot be drawn from the MMP results to the larger population of communities. Even though MMP data on migrant characteristics are similar to those from ENADID (Massey and Zenteno, 2000), the MMP "is not a technique for aggregate statistical estimation" (Massey et al., 1987:12-13). Rather, it is best used in causal models, such as modeling the determinants of migration in a multivariate setting. Similarly, the Mexican Migration Field Research Program (MMFRP) seeks to explain changes in migration and settlement behavior;

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

its data are not meant to be representative of any larger groups (especially since survey samples and questionnaires change from year to year[4]).

TIMELINESS

Estimating annual flows in a timely fashion using survey data is a great challenge, and doing so on a quarterly and border sector/subregion basis is an even greater challenge. For border flow estimates to have practical value, survey data would need to be collected, analyzed, and released in a timely fashion.

There are two components that need to be considered when discussing the timeliness of surveys: the frequency with which estimates are reported and the turnaround time from data collection to release of the data. ENOE is conducted on a quarterly basis and can be used to look at short-term population movements occurring within the 15 months when its five interviews are conducted, thereby providing a snapshot of seasonality and yearly changes in migration patterns. EMIF-N (along with the CPS) also contains information that is granular at the monthly level or smaller intervals. EMIF-N has historically been released in 12-month waves, although the investigators at COLEF have recently released data for the first quarter of 2012. Surveys such as ENADID and the Mexican Census, in contrast, provide an accumulated picture of migration during a 5-year period. Moreover, ENADID—like the MxFLS—is characterized by an irregular periodicity that makes it difficult to use for planning purposes. And, because of the challenges in tracking participants across communities and international borders, the periods during which survey personnel are in the field have continued to increase between MxFLS-1 and MxFLS-3.

Granularity aside, survey data will be most useful when they are made publicly available as quickly as possible. Monthly microdata from the CPS are released very quickly (within less than a month); the Annual Social and Economic Supplement of the CPS (formerly known as the "March supplement") is usually released in August or September. Data from ENOE are also made available relatively promptly, within a year of collection. Until recently, EMIF-N had at least a 2-year delay for the public release of data. However, the turnaround time for the most recent EMIF-N data released has been reduced to less than 12 months (as well as being released on a quarterly rather than annual basis).

[4]In some years, surveyors interview all 15 to 65-year-olds, while in other years they only interview those with migration experience.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

MIGRATION QUESTIONS

EMIF-N, the MMP, and the MMFRP are specifically designed to focus on migration and, not surprisingly, have the richest array of information on border crossings and the migration process more generally. While ENOE, the Mexican Census, ENADID, and the MxFLS do not contain as much information about migration as EMIF-N, the MMP, and the MMFRP, they do contain a number of items that are relevant to this study. Specifically, ENADID and the MxFLS ask about the documentation status of migrants, the MxFLS asks about crossing locations and (in the U.S. sample) the number of crossing attempts, and ENOE (for those not currently working) and the MxFLS ask about intentions to cross the border. (Unauthorized crossings are not illegal in Mexico, and there is substantial experience collecting data in Mexico on documentation status and mode of crossing; see Table 3-1.)

Although each of these household surveys contains useful information, they also have various gaps and limitations in terms of the questions asked. If one wanted to make substantive improvements to those survey instruments, one of the best options would be to add questions to the high-frequency and timely ENOE study. For international migration, the survey could specify the country to which individuals migrated. Questions could also be added that ask, for example, about the documentation status of migrants crossing to the United States, the number of attempts that were made, and crossing location. Furthermore, questions about intentions to move to or look for a job in the United States could be asked of all household members, not just of those currently not working. It may not be appropriate to go much beyond this, however, since the dynamic nature of the migration process (discussed in Chapter 2) could very well render questions about the migration process that are salient today less so in the near future. It is also important to ensure the "reliability" of survey instruments, which has to do with the degree to which a survey instrument elicits similar responses from different individuals under similar conditions.[5]

Such improvements could prove useful to researchers and others, and they would be welcome by the panel—as would, for example, improvements in the timeliness of EMIF-N. But from the perspective of estimating flows on an annual or quarterly basis, such survey modifications would still take place against the backdrop, as discussed above, of larger limitations and complexities relating to sample size and survey design. The panel also notes the challenges that arise from the fact that ENOE falls under the jurisdiction of the Government of Mexico. The addition of appropriate survey

[5]Instrument reliability is typically tested in pilot studies that precede the survey but are then rarely discussed. The panel did not find documentation associated with reliability in the descriptions of the various surveys carried out in Mexico.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

questions on the Mexican side would be challenging enough (to say nothing of potential problems arising from low and erroneous response rates), and the coordinated collection of complementary data from U.S. samples (which could provide useful information on undocumented migrants who successfully crossed the border) would be even more challenging. The panel does not discourage DHS from engaging with entities in Mexico that collect survey data relevant to the analysis of unauthorized migration, and in fact we believe that engagement would be beneficial. However, the difficulties of doing so, especially on such a politically sensitive issue, should also be acknowledged.

CONCLUSION

DHS appears to have multiple goals associated with obtaining information on unauthorized migration flows across the U.S.–Mexico border. Annual estimates of flows and apprehension probabilities would allow DHS to better evaluate (and report on) the effectiveness of its enforcement efforts, and estimates obtained on a quarterly basis and by specific geographic region (e.g., U.S. Border Patrol sector) might inform operational decisions regarding, for example, the allocation of enforcement resources along the border. More generally, this information would allow DHS to provide a more complete report to the public on the state of illegal immigration. With these DHS needs and interests in mind, the panel evaluated a range of surveys according to the following criteria: the nature of the target population and related issues of sample size and survey design; the frequency with which surveys are conducted and the speed with which data are made publicly available; and the types of questions that are asked about migration. The criteria chosen by the panel were based in large part on the standards of federal statistical agencies (National Research Council, 2009).

For border flow estimates to have practical value, survey data need to be collected, analyzed, and released in a timely fashion. Since ENOE is conducted on a quarterly basis and its data are released relatively promptly, it does the best job of meeting DHS's need for timeliness. The significant reduction in the turnaround time for public release of EMIF-N data, which took place during the drafting and revision of this study, have also increased the usefulness of EMIF-N to DHS.

Although ENOE has historically fared better than EMIF-N in terms of timeliness, EMIF-N collects a much broader range of information about border crossings than does ENOE (see Table 3-1 in Chapter 3). This is unsurprising considering that EMIF-N is a specialty migration survey whereas ENOE is a general labor force survey. Even so, both ENADID and the MxFLS—neither of which are specialty migration surveys—ask questions about border crossings that ENOE does not. In order for ENOE to be use-

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

ful to DHS, basic questions about the legal status of migrants, the number of crossing attempts, crossing location, and so on would have to be added to the survey instrument. Such modifications would entail a number of administrative challenges, given that ENOE falls under the jurisdiction of the Government of Mexico.

A more important concern of the panel regarding both ENOE and EMIF-N is the nature of the target populations of those surveys. Since the focus of DHS is a specific population group (unauthorized migrants) in a particular geographic area (the U.S.–Mexico border), a survey with a relatively narrow target population—such as EMIF-N—would appear to be of greatest use to DHS. The accuracy of the sampling weights in such a survey with an EMIF-N design is difficult to quantify, however, and the panel's concerns about its coverage error are significant. These issues are especially important given the adaptive and dynamic nature of the migration process. The panel's concerns about coverage error and the accuracy and transparency[6] of sampling weights are less pronounced for national-level household surveys such as ENOE. The issue there, rather, is that existing sample sizes are generally inadequate for detecting changes in flow rates. (Several possibilities for the magnitude of change that DHS may want to detect with 95 percent confidence are presented in Table 4-2.) Sample sizes would have to be even larger to obtain estimates for geographic subregions, and the design of the survey would be all the more complex—and, therefore, all the more vulnerable to obsolescence because of changing patterns of migration and enforcement along the border.

The financial costs that DHS would incur in establishing a new household survey in Mexico—or in adding a host of border crossing-related questions to ENOE and dramatically expanding its sample size (which would be tantamount to creating a new survey)—would be very high. The challenges associated with Mexican-side implementation and coordination would also be formidable (much more so than with adding new questions to ENOE while keeping the sample size the same), and the involvement of DHS could create additional problems relating to low and erroneous response rates. The financial costs and administrative complexity associated with any such survey would be multiplied by several factors if DHS wanted to obtain estimates by geographic subregion.

A survey such as EMIF-N that uses a time-location design and focuses directly on migrant populations holds greater promise (setting aside the panel's concerns regarding coverage error and accuracy of sampling weights) for estimating unauthorized flows across the U.S.–Mexico border. Nevertheless, extensive and direct involvement by DHS—be it in instituting a new survey similar to EMIF-N or working to improve the existing

[6]It was not until several panel members made a site visit to Tijuana, Mexico, that the panel was able to understand EMIF-N sampling weights and procedures.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

EMIF-N—would still raise concerns regarding Mexican-side implementation and low and erroneous response rates, similar to those that would be raised if DHS were to be involved in a Mexican-side household survey.[7]

- **Recommendation 4.1: For the purpose of estimating unauthorized migration flows across the U.S.–Mexico border on an annual or quarterly basis, DHS should not invest substantial resources in making major changes to existing surveys or in implementing a new survey.**

- **Conclusion 4.1: Existing surveys are subject to a variety of limitations having to do with target populations and associated issues of sample size and survey design, the frequency with which surveys are conducted and the speed with which data are made publicly available, and the types of questions that are asked about migration. Therefore, although survey data are critical for understanding patterns and general trends in unauthorized migration, they will not be sufficient by themselves to meet the needs of DHS for estimating unauthorized migration flows across the U.S.–Mexico border.**

- **Conclusion 4.2: Implementing a new household survey that meets the needs of DHS would require an investment at least comparable to that associated with the American Community Survey in the United States; any such survey would also have to fall within the purview of the Government of Mexico. A survey that uses a time-location design and focuses directly on migrant populations (e.g., EMIF-N) would be more promising, but such a non-traditional design would necessitate careful adherence to the sampling protocol and, in particular, would require that concerns about coverage error be addressed. Mexican-side implementation would also be an issue. Substantial modifications of existing general household or specialty migration surveys to meet the needs of DHS would encounter similar challenges. These challenges are only magnified by the complex and dynamic nature of the underlying migration process.**

The next chapter explores the usefulness and limitations of another approach: the use of DHS administrative data as they relate to the estimation of unauthorized migration flows across the U.S.–Mexico border.

[7]Sample sizes and financial costs would also be significantly smaller for an EMIF-N-like design than they would for a nationwide Mexican household survey.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

5

Administrative Data on Undocumented Migration Across U.S. Borders

Administrative data collected by the U.S. Department of Homeland Security (DHS) represent an important source of information about the activities of undocumented migrants. As part of its operations to secure U.S. borders and the U.S. interior against illegal immigration, DHS records the number of undocumented migrants it apprehends, the disposition of these migrants, and the resources it devotes to enforcement activities. The relevant administrative data come in three primary forms: apprehensions data collected by the U.S. Border Patrol (USBP), which contain individual records of migrants apprehended by USBP between ports of entry; data on apprehensions at ports of entry by the Office of Field Operations (OFO); and data on apprehensions in the U.S. interior by Immigration and Customs Enforcement (ICE).

The panel was tasked with reviewing administrative data collected by DHS, and it formally requested access to data from the enforcement database, indicating that DHS could provide the data to the panel in a format that would protect any information that DHS deemed operationally sensitive. The panel made this request with the understanding that any data given to it would need to be made publicly available, in accordance with the institutional rules governing National Research Council studies. However, DHS would not provide these data without an exemption from public disclosure requirements. It was the judgment of the panel that the quality of its published analysis and the timeliness of its deliberations would have been unduly impaired by the classification restrictions that would have accompanied such an exemption. Therefore, the panel did not pursue its request, and DHS did not provide the panel with access to its administrative data.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Nevertheless, the Office of Immigration Statistics (OIS) at DHS did provide information on the main fields included in the apprehensions database. Non-DHS data sources such as the Survey of Migration at the Northern Border of Mexico (EMIF-N) also provide some information on the apprehension of migrants at the border. In this chapter, we use the descriptions of the DHS data sources provided to the panel by OIS, as well other non-DHS data, to assess the usefulness of administrative data for measuring the flow of unauthorized migrants into the United States.

Our conclusion is that administrative data alone do not permit reliable estimation of the inflow of unauthorized migrants across the U.S.–Mexico border. The data provide no direct information on the number individuals who elude capture and enter the United States successfully. By making assumptions about the behavior of unauthorized migrants, one can use the volume of apprehensions to estimate the magnitude of unauthorized flows (see, e.g., Espenshade, 1995b; Massey and Singer, 1995). In this chapter, the panel discusses how one can generalize an approach, previously developed and reported in the sociology literature, which uses a "repeated trials" model. We note that all such estimates are based on strong assumptions that are difficult to validate empirically. As discussed in this chapter, estimation methods based on capture-recapture techniques (described in detail in Appendix B), which offer a sophisticated approach to determining the size of a population based on the fraction of initially "sampled" (i.e., apprehended) individuals who are subsequently "re-sampled" (i.e., re-apprehended) cannot, unfortunately, solve the problem about the lack of direct information on the number individuals who elude capture and enter the United States successfully. While administrative data have limitations, they could still offer potential insights into unauthorized migration flows if they were combined with other data sources. The panel outlines some strategies for characterizing key features of unauthorized migration flows, based on combining administrative and survey data. These approaches are discussed in more in detail in Chapter 6.

SOURCES OF ADMINISTRATIVE DATA ON APPREHENSIONS

USBP seeks to apprehend all individuals who attempt to cross U.S. borders illegally. Data on these apprehensions are a major source of DHS administrative records on unauthorized migration. In the last decade, USBP resources have increased dramatically, with the number of USBP officers growing from 9,000 in 2001 to 21,000 today (see Kessler, 2011). The expansion in resources, combined with the drop-off in the number of migrants apprehended at the U.S.–Mexico border since 2007, means that USBP currently has the manpower to document virtually all individuals with whom it comes in contact. But it also creates difficulties for comparing in-

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

formation across time. In the mid-2000s, when USBP had fewer agents and the booming U.S. economy encouraged high levels of unauthorized entry, recordkeeping on apprehensions may have been incomplete. Additionally, the expansion of USBP resources has been uneven across the nine sectors of the U.S.–Mexico border (Borger et al., in press), which complicates spatial comparisons of administrative data.

Beginning in fiscal year 1999, USBP created an electronic record of each apprehension made by a USBP agent. Table 5-1 is a partial list of the variables contained in the ENFORCE database, in which USBP, OFO, and ICE apprehensions data are recorded. Individual USBP apprehension records contain demographic information on the person apprehended, including gender, date of birth, country of origin, and (if a Mexican national) state of birth. The records also include an event number—which permits linking of individuals apprehended together—and information on when and where the apprehension took place, including the nearest port of entry and the USBP station, district, and sector of the arresting officer. A fingerprint identification number (IDFINS), which since fiscal year 2000 has been based on all 10 fingers, essentially identifies an individual and enables one to "tag" migrants who are apprehended more than once. On days (or more precisely, evenings) when apprehensions run at high levels, individual USBP stations may not have the personnel to fingerprint or interview all

TABLE 5-1 Variables in ENFORCE Database

U.S. Border Patrol Data
IDFINS
Event number
Date of apprehension
Site of apprehension
Nearest port of entry
Border Patrol sector
Disposition
Time in U.S.
Arrest method
Country of citizenship
Country of residence
Country of birth
State of residence (Mexico only)
State of birth (Mexico only)
Date of birth
Gender
Marital status

SOURCE: Office of Immigration Statistics, U.S. Department of Homeland Security (personal communication).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

apprehended migrants—a situation that appears to have been more common in the early 2000s, when the ratio of apprehensions to USBP officers was much higher than today. Currently a USBP station needs the approval of its station chief to forego fingerprinting and interviewing a subject. USBP asserts that it enters information into its database on nearly all current apprehended migrants.

Additional information in USBP apprehensions data describes:

- **Arrest method.** The method for the overwhelming majority of arrests is capture by USBP at the border. Other arrest methods are capture by USBP agents in the U.S. interior, capture by other law enforcement agencies, and capture at traffic checkpoints.
- **Status at entry.** The overwhelming majority of records have this status as "Present without Authority from Mexico," which indicates the individual was attempting to cross the U.S.–Mexico border as opposed to entering from Canada or by sea.
- **Status when found.** The overwhelming majority of records indicate the individual is in transit rather than working or seeking employment.
- **Time in the United States.** The overwhelming majority are for arrests at entry.
- **Smuggler use.** Whether the individual hired a smuggler to cross the border, and if so, the price paid.

The records also describe the disposition of the individual after apprehension. Most apprehendees are returned to their countries of origin, rather than being subject to incarceration in the United States.

Apprehensions data in aggregate form (i.e., the total number of apprehensions in a given month) have been used in a large body of academic research (see Hanson [2006] for a survey). Apprehensions tend to rise when average U.S. wages increase relative to average Mexican wages or when Mexico's real exchange rate depreciates vis-à-vis the United States (Hanson and Spilimbergo, 1999). They tend to fall initially, but later recover, in response to increases in USBP enforcement activities (Bean et al., 1990; Cornelius and Saleyhan, 2007; Dávila et al., 2002; Donato et al., 1992; Espenshade, 1994; Kossoudji, 1992; Orrenius and Zavodny, 2003).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

USING THE APPREHENSIONS DATA TO EVALUATE UNAUTHORIZED MIGRANT FLOWS

Using Apprehensions to Infer Unauthorized Flows

There are complications with using apprehensions data to estimate the number of individuals crossing the border illegally. The more serious problem relates to the inherent nature of the data: while apprehensions provide data on the number of individuals captured at the border, they provide no direct information on those who elude capture, which is the population of interest for this study. Minor problems include misreporting of key variables (e.g., given USBP return policies, non-Mexicans have a strong incentive to claim Mexican nationality) and possible missing data during peak apprehension periods (owing to failure to record all arrests during these times).

Nevertheless, the data can be used to apply capture-recapture techniques to make inferences about the size of the undocumented population and the flow of individuals into this population at regular time intervals. The *re-apprehension* of individuals provides information that can, in theory, be used to make inferences about the size of the unauthorized population entering the United States successfully. Of apprehensions of Mexican men over the period from 1999 to 2009, approximately three-fifths were of individuals who were apprehended only once, one-fifth are of individuals who were apprehended twice over the period, and one-fifth are of individuals who were apprehended three or more times (Borger et el., in press). Re-apprehensions of individuals typically occur within a few days or weeks of the initial capture. Such re-apprehensions are likely part of a single crossing episode. Of apprehended Mexican men who are subsequently re-apprehended, three quarters occur within 90 days, with the vast majority of these occurring in the first 30 days.

With the appropriate data in hand, how would one describe apprehensions analytically? Consider a simple model of the apprehensions process, whose main elements are described in Table 5-2. Suppose that at time t there are $M(t)$ individuals in Mexico who are considering crossing the U.S.–Mexico border illegally (this exercise ignores other nationalities). Suppose further that a fraction $m(t)$ of these individuals choose to attempt illegal migration, where $m(t)$ may be affected by economic conditions in the United States and Mexico and by the intensity of enforcement at the border. Let the probability of apprehension by USBP at time t be $a(t)$. Upon being apprehended, an individual decides whether to cross again or to return home (the latter group having been successfully deterred from further crossing attempts). Let the probability of retrying, conditional on being apprehended, be $r(t)$. It is likely that $m(t)$, $a(t)$, and $r(t)$ will vary

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

TABLE 5-2 Modeling Apprehensions

Variable	Definition
M(t)	Population of potential migrants in Mexico
m(t)	Probability of attempting migration
a(t)	Probability of apprehension, conditional on attempting migration
r(t)	Probability of re-attempting migration, conditional on being apprehended

across individuals according to demographic characteristics (age, gender, region of birth, marital status, family structure), skill (education, occupation, work experience), and knowledge about migration (previous crossing experience, access to migrant network), among other factors. For purposes of illustration, we first ignore these sources of individual heterogeneity but then consider some ways to address them below.

In each period *t*, the number of individuals apprehended on their first attempt to cross the border is represented by the product, $a(t)m(t)M(t)$: the number of potential migrants in Mexico times the probability that an individual attempts to migrate times the probability of apprehension. Similarly, the number of individuals apprehended on their second attempt to cross the border is equal to the product, $a(t)r(t)[a(t-1)m(t-1)M(t-1)]$: the number of individuals apprehended last period (in brackets) times the probability of attempting to cross again (after an initial apprehension) times the probability of apprehension. Assuming that the IDFINS data are known for all migrants, one can separate apprehensions associated with first attempts to cross the border from those associated with repeat attempts. Ignored here are complications associated with individuals waiting for more than one period before retrying. Allowing for waiting complicates the math, but does not change the basic structure of the problem.

Using the value of first-time apprehensions in the previous period and second-time apprehensions in the current period, one can identify the value for $a(t)r(t)$, which is the joint probability of retrying to cross the border and being apprehended. But unless one assumes that $r(t) = 1$, which means that all individuals who attempt to cross the border keep trying until they succeed, the value of $a(t)$ cannot be determined; if $a(t)$ can be identified, then the magnitude of apprehensions can be used to estimate the magnitude of illegal attempts to cross the border or successful attempts to cross the border. Previous uses of apprehensions data to estimate unauthorized migration flows, such as the analyses by Espenshade (1995b) and by Massey and Singer (1995), have assumed that $r(t) = 1$. But the EMIF-N data do not support this assumption. Some individuals who are apprehended become discouraged and do not attempt to cross the border again. Furthermore, the

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

EMIF-N data suggest that the fraction of attempted crossers who become discouraged has risen over time.

By imposing additional structure on the data, one can use the apprehensions data to infer more about unauthorized migration flows. One approach would be to develop an economic model of the process governing the decision to migrate, which would allow one to characterize how the migration probability, *m(t)*, responds to changes in economic conditions, and how the apprehensions probability, *a(t)*, and the probability of retrying to cross the border, *r(t)*, respond to changes in the level of enforcement activity. The data sets from ENOE, ENADID, the MMP, and the MxFLS contain information that permit the migration decision to be modeled (Gathmann, 2008; McKenzie and Rapoport, 2010; Orrenius and Zavodny, 2005) with some limitations (as discussed in Chapter 4). Furthermore, EMIF-N and the MMP contain information that allows examination of how re-apprehensions respond to changes in enforcement. The use of survey data would further allow one to address heterogeneity across individuals in how the decision to migrate responds to changes in environmental conditions. McKenzie and Rapoport (2010), for instance, find that the propensity to migrate varies across Mexican communities according to the past migration experience of community members. Chapter 6 discusses different classes of models that combine survey, administrative, and other types of data.

A second approach in the use of apprehensions data is to make assumptions about the stochastic process governing apprehensions, generalizing the repeated-trials method developed by Espenshade (1995b). In a section below, we use EMIF-N data on repeat apprehensions to illustrate such an approach (which could be explored more fully using ENFORCE data).

Using Data on Smuggling Costs

The price of smuggling services is a potentially useful indicator of the effectiveness of border enforcement. If intensifying enforcement causes the risk of apprehension, incarceration, or physical harm for smugglers to rise, one would expect them to charge higher prices for ferrying migrants across the border. Gathmann (2008) finds in data from the MMP that higher levels of enforcement activity by USBP are associated with higher prices for smuggling services paid by migrants in surveyed communities. Because the ENFORCE database contains information on smuggling costs, it could be used to construct a measure of the effectiveness of enforcement.

In practice, there are myriad problems with using existing smuggling cost data for analytical purposes. One is that USBP collects information on smuggling costs in an inconsistent manner. Borger and colleagues (in press) report that, in the USBP apprehensions data, fewer than 20 percent of apprehendees report whether or not they use a smuggler. Even if one

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

were able to collect information on smuggling and smuggling prices from all individuals apprehended, the problem would remain that these prices correspond to just those individuals who were apprehended. One would therefore have a selected sample of individuals from whom to extract price data. Smuggling prices in such a sample may be subject to downward bias owing to the possibility that individuals purchasing inferior smuggling services at low-end prices may be more likely to be apprehended than those who paid higher prices.

A further problem is that there is likely to be differentiation in the market for smuggling services. There is likely to be variation across migrants in the duration of smuggling services being purchased (transport immediately across the border versus delivery to an interior U.S. city), the size or composition of the group being guided (small numbers of adult males versus entire families including children), experience in crossing the border, mode of transportation (by land or by sea), and the physical risks being confronted (cold in winter, heat in summer, longer routes when crossing through the Sonoran desert or mountainous regions). Borger and colleagues (in press) find that smuggling prices tend to be higher for groups that include women, children, or the elderly. With information on the characteristics of the services being purchased, one could in principle construct a hedonic price index for migrant smuggling that would adjust for product differentiation, similar to how the U.S. Bureau of Labor Statistics adjusts price indices for consumer goods to account for changes in the quality of goods over time. However, the information currently reported by USBP in its ENFORCE database is insufficient for such an exercise. The systematic collection of data on smuggling prices would expand the options available to DHS for analyzing the behavior of undocumented migration at the U.S.–Mexico border.

Frequency of Apprehension Frequencies Analysis

An alternative approach for using apprehensions data to estimate the number of individuals crossing the border successfully is to impose assumptions about the underlying stochastic process that governs attempts to cross the border. If one treats apprehensions as a count variable and the number of apprehensions for a given individual as the outcome of a draw of a random variable from a defined distribution, then one can use the observed frequency of outcomes for one apprehension, two apprehensions, three apprehensions, and so forth to estimate the "missing category" of zero apprehensions, which corresponds to the number of individuals who cross the border without being apprehended. This approach generalizes the repeated-trials model used by Espenshade (1995b). Implementing this approach requires one to impose the untested assumption that data on ob-

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

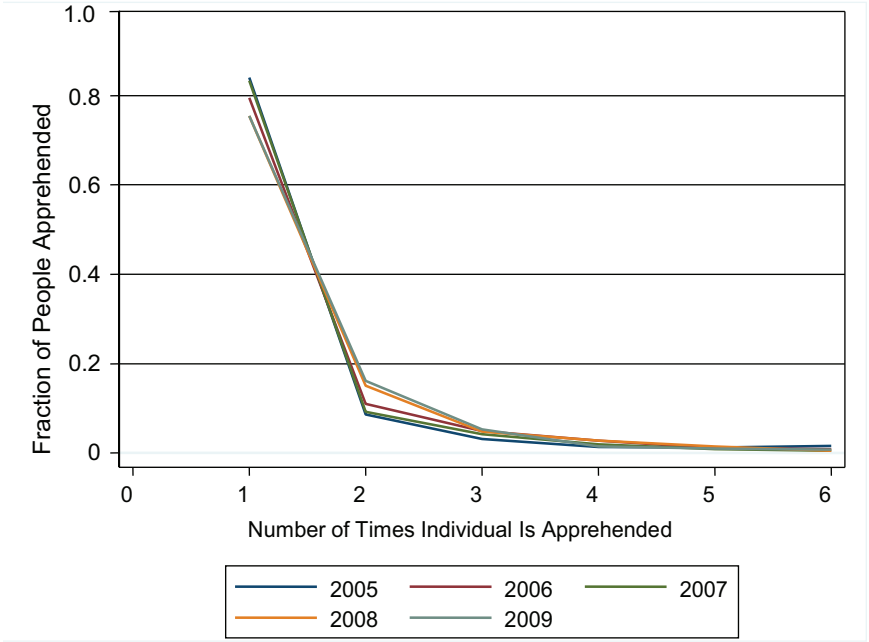


FIGURE 5-1 Frequency of apprehensions across time.
SOURCE: Data from EMIF-N.

served behavior (the number apprehended) is informative about unobserved behavior (the number avoiding apprehension).[1]

To provide an example of how one could use data on apprehensions, Figure 5-1 shows the number of times individuals surveyed in EMIF-N report being apprehended in a given series of attempts to cross the U.S.–Mexico border. The length of the time window used to define apprehensions for a single crossing episode is an important issue. In EMIF-N, the length of this window is not defined precisely. Because most subsequent apprehensions occur within a few weeks of the initial apprehension, the length of the crossing window may not be important for the results. As more of those apprehended are subject to consequence programs, however, the appropriate window for defining a migration episode may change.

[1]Related ideas have been used to estimate the number of animals or plants present in a community (Bunge and Fitzpatrick, 1993; Royle and Dorazio, 2008), the number of unique records in a filing system with duplicates (Arnold and Beaver, 1988), and statistical disclosure risk assessment (Fienberg and Makov, 1998). In addition, there is a growing literature on Good-Turing methods to estimate the probability of types being unobserved (Good, 1953). The latter were originally developed to estimate the frequencies of words in a corpus.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Each plot in Figure 5-1 shows the relative frequency of outcomes (as the fraction of all outcomes, or density) based on observations for the calendar year of the initial apprehension. For individuals first apprehended in 2009, 75.4 percent are apprehended only once, 16.1 percent are apprehended twice, 4.4 percent are apprehended three times, 1.3 percent are apprehended four times, 0.6 percent are apprehended five times, and 0.5 percent are apprehended six or more times. To estimate the number of zero apprehensions, the frequency of apprehensions is projected to the left, based on estimation of the distributional parameters governing apprehension, as discussed below.

The figures for 2005 to 2009 overlap one another to a considerable degree. There is a slight increase in the fraction apprehended two times in 2008 and 2009 compared with earlier years. This suggests that, if the same distribution governs apprehensions in each period, the fraction of individuals in the zero-apprehension category would be similar across time, although slightly lower in recent years. An important implication of such an outcome is that the probability of crossing the border successfully may have changed only modestly across time, despite the massive increase in border enforcement resources. The panel emphasizes that drawing inferences about the effectiveness of enforcement policies from apprehensions data is problematic. In the simple analysis illustrated here, the probability that those who are apprehended re-attempt to cross the border is assumed to be stable over time (an assumption also implicit in the repeated trials approach discussed earlier). If this assumption is incorrect, there may be less stability in the zero apprehensions category than Figure 5-1 appears to suggest. We also note that this "frequency of frequencies" approach cannot separate the proportion of migrants who cross successfully from those who are deterred from further attempts.

The data used for Figure 5-1 provide mild evidence of variation in outcomes for apprehensions across time (from year to year). What about variation across space (i.e., border sectors)? Figure 5-2 plots the frequencies of apprehensions from 2007 to 2009 for four USBP regions that account for a large share of apprehensions: Tijuana/San Diego, Nogales/Nogales, Ciudad Juarez/El Paso, and Nuevo Laredo/Laredo. Across sectors, there is minor variation in the frequency of apprehensions. Projecting back to the category of zero apprehensions, it would appear that the frequency of successful crossings is slightly higher for Ciudad Juarez/El Paso and lower for the other locations. The absence of notable regional variation in the zero apprehensions category suggests that, despite large cross-sector differences in the scale of enforcement activities, the probability of apprehension may be stable across regions.

It would have been preferable to perform the analyses represented in Figures 5-1 and 5-2 using ENFORCE data. ENFORCE covers the universe

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

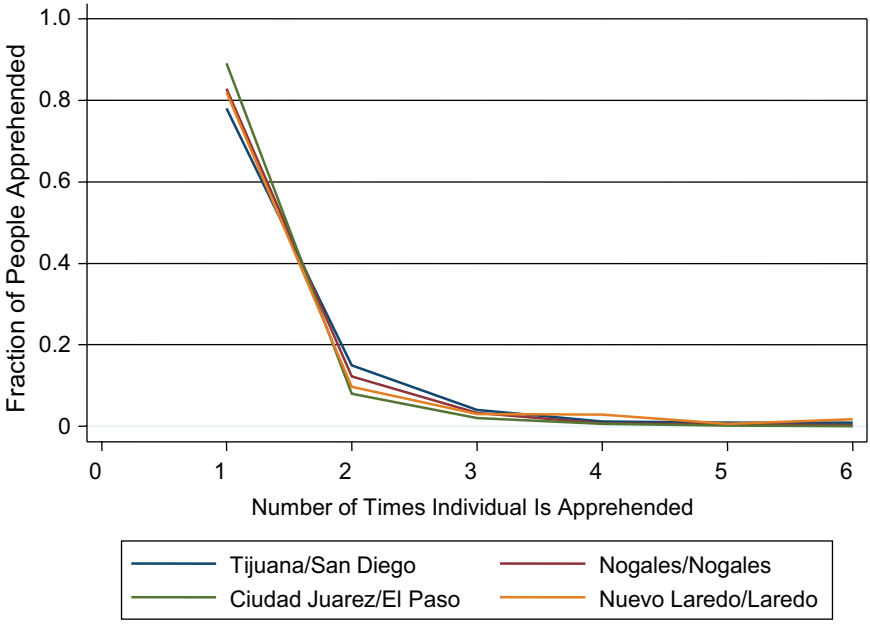


FIGURE 5-2 Frequency of apprehensions across space.
SOURCE: Data from EMIF-N.

of individuals apprehended, whereas EMIF-N only covers those individuals questioned by survey enumerators, whose choice of survey zones and points to find individuals being returned to Mexico after apprehension may introduce unknown sources of bias into the sample. Of course, both ENFORCE and EMIF-N are subject to the limitation that the population of individuals who are apprehended once but not seen again includes both those who, on their subsequent attempt, cross into the United States successfully and those who, after the initial apprehension, become discouraged and return home to Mexico. The conflation of successful crossers and discouraged crossers contaminates the analysis. For instance, if no apprehended crossers became discouraged, then the ratio of those apprehended for a second time to those apprehended just once would equal the probability of apprehension. But if some apprehended crossers become discouraged, then this ratio equals the apprehension probability multiplied by the probability that initially apprehended crossers do not become discouraged, a product that does not allow one to pin down the apprehension probability itself.

Despite the concerns noted above, there would have been additional value in using the ENFORCE data for this analysis. Because ENFORCE data

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

contain the universe of apprehensions, as well as demographic information on captured migrants, they would permit one to evaluate the stability of the zero-apprehensions category across time, space, and individuals by age, gender, and region of birth within Mexico. The panel could have used the ENFORCE data to determine whether there were systematic changes in the zero-apprehensions category as DHS boosted enforcement along the border and imposed consequence programs at specific points along the border. Such changes may indicate that apprehension probabilities are responsive to changes in border enforcement (e.g., the zero-apprehensions category expands because more individuals are being caught) or that the composition of border crossers is responsive to changes in border enforcement (e.g., the zero-apprehensions category expands because more-determined crossers account for a higher fraction of those crossing). Although the panel would not have been able to attach likelihoods to these or other explanations, knowing whether the size of the zero-apprehensions category was correlated with the intensity of border enforcement would have been helpful to the panel in considering approaches (discussed in Chapter 6) to formally modeling migration flows. In conducting these analyses, moreover, the panel would have gained at least some insight into the quality, completeness, and reliability of the administrative data.

More Detailed Frequency of Apprehension Frequencies

The frequency of frequencies approach is based on fitting statistical distributions to the counts of the number of times an individual is apprehended (within a given window).[2] Three core assumptions are sufficient to make such models meaningful:

1. The individuals are apprehended independently. That is, their propensity to be apprehended is independent of that of other individuals who attempt to cross.
2. Apprehended individuals are not deterred from subsequent attempts. In fact, the approach assumes that they will attempt to cross until they are successful.
3. If individuals cyclically migrate then their propensity to be apprehended is independent of their prior attempts. That is, their probability of apprehension is the same as if they were a new individual attempting to cross.

[2]The frequency of frequencies may be thought of as a "species problem" (see Efron and Thisted [1976] on how many words Shakespeare knew).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

BOX 5-1
Classes of Statistical Distributions

The following classes of distributions are relevant for modeling frequency of apprehension frequencies: Poisson, negative binomial, geometric, and Conway-Maxwell-Poisson. The choice of a Poisson distribution can be motivated in a number of ways. A simple one is to assume that individuals have an "apprehendability," defined as a measure of their propensity to be apprehended. Specifically, assume that apprehendability is measured by the expected number of apprehensions before a successful attempt for the individual. Further assume that the ratio of the probability of being apprehended $k - 1$ times to that of being apprehended k times is proportional to k (unconditional on being apprehended the kth time). If this apprehendability is common to all individuals, then the number of times an individual is apprehended is a Poisson distribution whose mean is their apprehendability.

The choice of a negative binomial distribution can be motivated by individual heterogeneity in apprehendability. Assume that the individual apprehendabilities vary but can be modeled as independent draws from a gamma distribution. Then the number of times an individual is apprehended is a negative binomial distribution whose mean is the mean apprehendability of the group of individuals.

A geometric distribution is a special case of negative binomial distribution. That is, it presumes a specific relationship between the mean apprehendability and the variance of the apprehendability in the population. There is an alternative motivation for choosing it. Suppose there is a common probability of apprehension per attempt, a, and the apprehension events are independent for the same individual over time. Then the number of times an individual is apprehended is represented by a geometric distribution whose mean is $(1 - a)/a$.

The Conway-Maxwell-Poisson is a variant of the Poisson distribution that allows over-dispersion (like a negative binomial distribution) as well as under-dispersion relative to a Poisson distribution. While over-dispersion may be expected, it is possible that under-dispersion in the number of apprehensions occurs, and this possibility should be represented in the array of models whose fit to the data is tested.

All three assumptions are important, and the possibility of deviating from them is high. Details on statistical distributions that incorporate these assumptions are presented in Box 5-1.

The observed number of apprehensions is truncated at zero. That is, there are no observations of individuals who are not apprehended on their first attempt. However, candidate distributions can be fit to the available data, taking this into account. The fit can be tested using a maximum likelihood estimation method, such as that provided in the **degreenet** package from the Comprehensive R Archive Network (CRAN).[3] Measures of the

[3]Available: http://statnet.org and http://CRAN.R-project.org/package=degreenet (accessed August 2012).

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

uncertainty due to sampling, including confidence intervals, can be estimated, although such estimates are not included in this discussion. We note that the assumptions of the model allow it to extrapolate beyond the data themselves (i.e., to the frequency of the number of times an individual is not apprehended). If, however, the distribution is misspecified or the assumptions are false, then this extrapolation can be subject to substantial error.

Figure 5-3 shows the total proportions for the number of apprehensions from the 2009 EMIF-N data, represented by the circles on the plot (notice that no value is shown for zero apprehensions). The colored lines on the plot represent the fits of the four types of statistical distributions described in Box 5-1. The negative binomial, geometric, and Conway-Maxwell-Poisson distributions all fit the observed counts closely—indeed, the lines overlap so that the separate colors are not visible. The Poisson

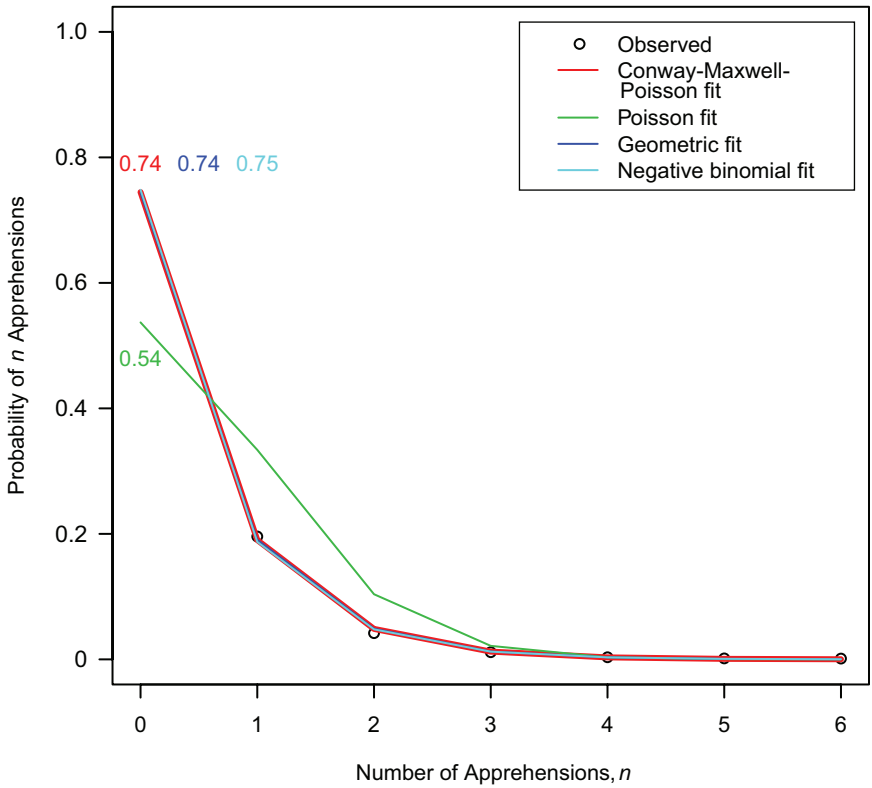


FIGURE 5-3 Fits of naïve apprehensions models for 2009.
SOURCE: Data from EMIF-N.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

distribution does not provide a good fit for one to three apprehensions. The numerical values for the probability of non-apprehension on the first crossing attempt, as estimated by each of the four distribution classes, are shown close to the vertical axis. (The values of the distributions at zero apprehensions are off the scale of this graph.) The probabilities for zero apprehensions for the three good-fitting distributions are close together (74 to 75 percent), while that of the poorer-fitting Poisson distribution is substantially lower (54 percent). Because the sample sizes are large, the nominal confidence intervals (not plotted) for the probability of non-apprehension on the first crossing are narrow, but these estimates are not adjusted for possible model misspecification. A variant of this distribution-fitting approach is the Good-Turing frequency model (Good, 1953). The simplest version of this approach estimates the probability of non-apprehension as the proportion of those apprehended who were apprehended exactly once. For the 2009 EMIF-N data, this estimate is 75.4 percent, which is in close agreement with the first three good-fitting distributions.

CONCLUSION

At least three agencies within DHS collect administrative data on apprehensions of unauthorized immigrants. USBP collects data on apprehensions between ports of entry, OFO collects data on apprehensions at ports of entry, and ICE collects data on apprehensions in the interior of the United States. Because fingerprints on those apprehended are collected in all three data sources, DHS's ENFORCE database can integrate data across the three sources at the individual level. However, conversations with representatives from DHS suggest that the linkages between the apprehensions records controlled by USBP, OFO, and ICE in the ENFORCE database are limited to uses that relate specifically to enforcement. Linkages across the data sources for broader analytical purposes would require approval from each of the three agencies, and the full database has not been widely used for analysis.

If one wants to analyze apprehensions at the border, integrating USBP and OFO apprehensions records is essential.[4] To understand how U.S. enforcement, either at the border or in the interior, affects attempts at unauthorized entry, integration of the ICE and USBP databases is necessary. As discussed in Chapter 2, increasing enforcement in one border sector may

[4]See also National Research Council (2011:50-51) for a discussion of how the immigration enforcement data published in the widely used DHS *Yearbook of Immigration Statistics* do not completely reflect the immigration enforcement activities undertaken by all relevant DHS agencies.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

redirect attempts at entry to other sectors, implying that analysis of enforcement requires looking at the border as an integrated whole.

That said, administrative data from DHS are alone insufficient to estimate the flow of unauthorized migrants across the border. These data report the number of individuals who are captured by USBP at entry but contain no information on those who elude capture. Data on re-apprehensions of individuals provide insight into the migration process but do not carry information about the "got aways." Because those apprehended may either try to cross the border again or return to their home in interior Mexico, knowing the fraction of those re-apprehended does not allow one to identify the probability of apprehension. Without knowledge of the apprehension probability, one cannot use the level of apprehensions to make inferences about the flow of undocumented migrants across the border. It is unlikely that having access to the DHS administrative data would have changed the panel's conclusion.

Despite their limitations, the administrative data still have many uses for understanding unauthorized migration. By combining administrative data with survey data, one can produce a model of individuals' migration decisions in Mexico that is informative about how attempts at illegal entry respond to changes in the economic environment. Such an approach would combine a behavioral model of the decision to migrate, analyzed using survey data, with a model of the stochastic process governing apprehensions, analyzed using administrative data. Although this approach can produce estimates of the flow of unauthorized migrants across the border, it incorporates assumptions about migrant behavior and the statistical properties of apprehensions that may not be open to empirical validation. Still, models of migration that combine administrative and survey data would provide DHS with additional tools for analyzing the effectiveness of border enforcement and expected future workloads for USBP agents. Chapter 6 examines in more detail a variety of modeling techniques and approaches.

Because DHS administrative data have not been made public, they have never been evaluated by independent scholars for their quality, completeness, or reliability—an omission that is significant in light of the role that apprehensions data could play in informing model-based approaches to estimating flows. As the discussion in Chapter 6 will also make clear, knowledge of and experience with the use of model-based approaches for estimating flows are so far limited, and the attendant complexities and uncertainties are considerable. In order to develop, apply, and continually refine specific modeling approaches, DHS will need to engage with the broader scientific community in a sustained and long-term fashion. This will be possible only if the administrative data discussed in this chapter are made widely available.

Currently, however, DHS shares its administrative data on apprehen-

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

sions of unauthorized immigrants only with those with whom it contracts to perform confidential analyses. Such analysis comes at high monetary cost to DHS and is rarely subject to peer review in the manner typical of academic research. There is a large community of scholars actively studying illegal immigration. Providing this community with access to the ENFORCE database would likely produce dozens of new academic studies that would be available to DHS at no charge. Because most of these studies would be subject to peer review by academic journals, they would arguably be of higher quality than the consulting reports that DHS currently acquires. The wide dissemination of data (along with the integration of data from surveys with data from administrative records) is also a recommended practice for federal statistical agencies, which have developed a number of procedures for providing research data access while protecting the confidentiality of the information (National Research Council, 2009).

A further benefit of putting individual-level apprehensions data in the public domain is that this could potentially improve the quality of data collection by EMIF-N and any future surveys that target individuals who have been apprehended by USBP. DHS's administrative data presumably represent the full universe of apprehended migrants. As noted in Chapter 3, these data would be valuable in efforts, such as EMIF-N, to survey and estimate flows of such repatriated migrants. In particular, detailed data on the number and basic demographic characteristics of migrants repatriated, by time, date, and port of entry of return, would provide an independent measurement of this return flow that would be extremely helpful in both designing the survey's sampling frame and correctly weighting estimates.

Some in DHS have expressed concern that restricting the release of administrative data is necessary because information contained in the files is law enforcement sensitive. However, the number of apprehended individuals subject to criminal prosecution for terrorism or the trafficking of drugs, arms, or people appears to be very small. Records on these individuals could be excised from the USBP apprehensions database before their release to the public, without affecting the value of these data for analytical purposes. Even though the smugglers of illegal aliens already appear to have relatively accurate information on the rates of apprehension and successful entry into the United States, important operational information could nevertheless be safeguarded through broad geographic identifiers that link, for example, to USBP sectors rather than individual USBP stations. Others have argued that releasing administrative data risks violating the privacy of individuals who are apprehended. However, it is simple to transform individual identifiers in the USBP apprehensions database in a manner that would make the risk to privacy very low. For many research purposes, individual-level data would also be unnecessary. It would be sufficient to have aggregate data on the frequency of apprehensions for individuals bro-

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

ken down by age, gender, country of birth, sector of apprehension along the border, and time period (e.g., month and year). The panel believes that USBP would be able to release a much more complete individual-level file by implementing masking methods for problematic fields in the records and by releasing data with sufficient delay, for example a full year, to diminish their sensitivity for operational use and deployment.

It should be noted that there are other mechanisms in addition to properly constructed public use files for providing researcher access to data while protecting privacy. These mechanisms include: establishing one or more secure enclaves for researchers to access microdata, similar to the U.S. Census Bureau's network of Research Data Centers or its National Science Foundation–Census Research Network; developing remote, monitored online data access services such as the system maintained by the National Center for Health Statistics (NCHS); and providing licenses to individual researchers for using confidential data at their institutions, as is done by the National Center for Education Statistics (NCES). However, these systems, unlike the construction of one or more public use files, require a level of staff and resources that would likely be difficult for DHS to establish and maintain over a sustained period. Given the basic nature of the information included in the DHS enforcement administrative databases and the population in question (i.e., unauthorized border crossers), DHS's confidentiality and privacy concerns may also be different from those of NCHS, NCES, and the Census Bureau.

- **Recommendation 5.1: DHS should integrate apprehensions data from USBP, OFO, and ICE for analytical purposes.**

- **Conclusion 5.1: Administrative data from DHS are alone insufficient to estimate the flow of unauthorized migrants across the U.S.–Mexico border. However, they could be combined with survey data to produce useful insights about migrant flows and the effectiveness of border enforcement. The use of modeling approaches in conjunction with disaggregated survey and administrative data is necessary for estimating these flows.**

- **Recommendation 5.2: DHS should sponsor and conduct research on modeling approaches for estimating the flows of unauthorized migrants across the U.S.–Mexico border.**

- **Conclusion 5.2: DHS would greatly benefit from making the administrative data from its immigration enforcement databases**

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

publicly available for research use, as that would allow DHS to engage with the broader scientific community to develop, apply, and continually refine specific modeling approaches. DHS could develop ways of constructing masked and/or aggregate files for public release in order to protect sensitive information.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

6

Model-Based Approaches to Estimating Migration Flows

INTRODUCTION

The migration process for undocumented people is complex and dynamic, as described in Chapter 2. Undocumented migrants and their agents adapt quickly to changes in resources and strategies on the U.S. side of the border. Furthermore, the migration of undocumented people responds not only to enforcement efforts by the United States but also to labor market factors in the United States and Mexico, local laws and regulations on both sides of the border, and "competing" traffic across the border—including the highly profitable drug traffic going north and arms traffic heading south.

To estimate the number of illegal crossings at the U.S.–Mexico border, enforcement agencies require information that is not only precise but also timely (e.g., available on a quarterly basis and soon after the end of a quarter). The fact that the migration process is highly dynamic makes this difficult. For example, neither the emergence of drug violence along the border nor the severe economic recession in the United States were anticipated as recently as 5 years ago. The need for accurate estimates at the border at a geographically detailed level introduces additional challenges. Any effective information system will have to be agile, adapting to altered flows and externalities as the process evolves.

No single data source is able to provide direct estimates of the number of illegal attempts to cross the U.S.–Mexico border (see the discussion in Chapters 4 and 5). While several U.S. and Mexican surveys (described in detail in Chapter 3) address specific aspects of the migration process, they tend to do so only in limited ways (as discussed in Chapter 4). A new

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

survey, or substantial modification of a current one, would be very costly. Moreover, its design would have to be sufficiently flexible to reflect the dynamic nature of the migration process.

Similarly, administrative data collected by the U.S. Department of Homeland Security (DHS) along the U.S.–Mexico border (which were collected for purposes other than the estimation of migration flows) are likely to provide only a partial picture of the activities of undocumented migrants and cannot be used in isolation to draw inferences about migration flows (see Chapter 5). The difficulty in estimating flow from current data sources persists even if statistical modeling techniques, such as capture-recapture methodology and other sampling strategies, are used to estimate these hard-to-count populations.

Based on the panel's conversations with U.S. Border Patrol (USBP) agents during site visits to Arizona and California, it is clear that USBP already attempts to combine information from different sources to forecast border-crossing activity, albeit in informal ways. In addition to whatever the surveys may indicate, agents make use of their own administrative data, their previous experience, and other sources of information that include, for example, occupancy rates of hotels on the Mexican side of the border, sign-cutting (i.e., observing and tracking footprints and other physical signs of migrant passage), and remote sensing data.

Building upon what USBP already does in practice, this chapter discusses more formal ways for combining varied sources of information to estimate unauthorized migration flows with geographic and annual/quarterly specificity. These methods include conventional approaches, such as probability models, regression models, and spatiotemporal processes, and more recent methods such as agent-based modeling.

To fit a model, one wants to have a training sample for which both the explanatory variables (such as economic pressure, enforcement effort, point of origin) are known, and also the true values of the response variable (such as the flow of illegal immigrants at a specific portion of the border). Such a training sample is difficult to obtain in this situation and will never be fully achieved. Nevertheless, a model for illegal flow will include many components for which data exist. For example, each border station records the number of people in different demographic segments who are interdicted that month, and surveys are available that indicate how many people in a particular town chose to seek work in the United States. A mathematical model for illegal immigration that is founded on good social science theory can be fit to the available data, and it offers reasonable hope of correctly tracking the unmeasured data. This hope can be approximately validated, or disconfirmed, if the model's broad predictions for, say, the total number of illegal Mexican immigrants are not consistent with estimates obtained

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

from other sources (e.g., the cost of day labor or the number of illegal aliens found during random traffic stops).

TELEPHONE CARDS: A THOUGHT EXPERIMENT ON QUANTIFYING A DIFFICULT-TO-MEASURE POPULATION

Two glaring gaps in the information required to estimate the effectiveness of the resources that have been deployed at the U.S.–Mexico border during the past decade are the proportion of undocumented crossers who succeed in their first or later attempts and the proportion of apprehended migrants who are deterred from further crossing attempts. In the course of its deliberations, the panel discussed a number of different ideas concerning creative sampling methods for estimating different components of undocumented immigration. The panel describes here one such idea for quantifying one type of deterrence effect (i.e., the fraction of apprehended migrants who choose not to attempt to cross the border again). This simple thought experiment involves providing telephone cards to undocumented immigrants who are apprehended in the United States and are then returned to the Mexican side of the border.

Typically, individuals who intend to cross the border without documentation arrive in the border area and make arrangements for illegal crossing with assistance from a smuggler. Most of those who are apprehended during their first attempt and are returned to the Mexican side of the border will try to cross again within the next few days. If the second attempt is also unsuccessful, they tend to keep trying, usually over a period of several days, until finally they either succeed or give up. USBP could, in principle, provide a phone card from a Mexican telephone company to a randomly selected subset of apprehended migrants who are about to be returned to Mexico. The phone cards, which would come preloaded with a certain usage value, could be used to call from either side of the border, but only after the caller is identified as the person who actually received the card. The toll-free number to activate the card would differ depending on whether the individual was in the United States or Mexico at the time of activation. The fraction of individuals activating the card in Mexico would provide an estimate of the fraction of apprehended individuals who are deterred from crossing again, and the fraction activating the card in the United States would provide an estimate of the fraction of individuals that cross successfully on the next attempt.

Several practical problems would need to be resolved in order to implement the phone card experiment. First, phone cards would have to be sufficiently attractive so that migrants actually use them, but not so valuable that they render the program too expensive (or induce criminal elements to prey on returning migrants, or create a black market in phone cards). This

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

issue might be addressed by experimenting with different phone card values and by varying the fraction of migrants who receive phone cards across time and border locations. Second, it would be necessary to ensure that the user of the phone card is the intended card recipient and not someone else to whom the card was given or sold. This issue might be addressed by having the migrant answer simple questions at the time of phone card activation based on demographic information collected at the time of apprehension. Third, it is necessary to ascertain whether the location from which the migrant activates the card is the migrant's final destination (be it in Mexico or the United States). This issue might be addressed by having the card's earliest possible activation date be 1 to 2 weeks after the apprehension, so that cards activated in Mexico would predominantly be activated by discouraged crossers, while cards activated in the United States would predominantly be activated by successful crossers. Finally, one would expect that successful undocumented migrants would be more reluctant to activate their cards in the United States than would unsuccessful undocumented migrants still in Mexico, so the undercount is likely to be different for successful undocumented crossers than for unsuccessful crossers who are deterred from further attempts. The risks of non-use and differential use cannot be ignored. Steps would need to be taken to address the concerns of undocumented migrants, such as allowing callers not to self-identify and providing assurances that the identity and locations of the callers are not required and would not be traced (beyond the country of origin).

This thought experiment highlights the kinds of data that would be needed if apprehensions data were to be used to estimate stocks or flows of unauthorized immigrants. However, even though the phone card experiment might be useful in estimating the number of crossers who are successful after having been apprehended once or more than once, it still would not provide any information about individuals who cross successfully in their first attempt. One approach to counting such elusive populations is based on network sampling or link-tracing sampling (see Box 6-1). However, these methods require careful implementation and additional assumptions to be usable, and they are not yet sufficiently developed to be clearly helpful to DHS in filling the critical data gaps.

SURVEY AND ADMINISTRATIVE DATA—INFORMED MODELING

Statistical models can provide plausible descriptions of immigration behavior, and some aspects of their fit can be validated against available survey and administrative data. A model can be applied to historical data to see whether its predictions agree with results from previous surveys. Even though the available surveys do not directly address all questions of interest to DHS (see Chapters 3 and 4), if a statistical model agrees

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

with the findings of the surveys on those aspects of flow that the surveys do capture, then one can reasonably expect that the model has predictive power for estimating other relevant aspects of flows. Similarly, if a model produces results that are not supported by previous data, then one of three conclusions is plausible: the model does not fit the data well, the migration process has changed significantly over time, or both these conditions apply. The model must be flexible, and one should expect that it will be necessary to extend it when new factors come into play, leading to a new round of model retrofitting and validation.

Beyond timeliness and the possibility of greater accuracy, modeling has additional advantages. A good model allows policy makers to explore "what if" scenarios by changing model inputs. In particular, DHS can explore the impact of different allocations of enforcement resource among border stations or the impact of new enforcement policies. More importantly, the process of building a good model can create a stronger understanding of the social process underlying immigration behavior. Finally, a good model should produce accurate estimates of prediction uncertainty. Predictions from a model that are not paired with estimates of their prediction error have limited value.

This chapter reviews several approaches to survey-informed modeling and, to the extent possible, offers some comparative guidance in the context of estimating unauthorized migration flows. There are at least three standard strategies for doing survey-informed estimation: build a probability model, fit a regression model, or employ a spatiotemporal model. Newer approaches such as agent-based modeling are based on simulation but still rely on survey and administrative data for parameter settings and model components. The estimates that result from the use of models should usually be fairly accurate if the data are representative and reliable, the model is valid, and the immigration process does not change. In this case, disaggregated data on migrants and nonmigrants that allow for the explicit modeling of the migration decision can be used to verify and validate findings from aggregate data, such as apprehensions.

Probability Models

Massey and Singer (1995) developed a simple probability model for the number of unsuccessful attempts at illegal immigration before a successful crossing. Their basic model was a geometric distribution for the number of attempts before the first success, where it was assumed that attempts were independent trials with constant probability of success. The estimated probability of success was obtained from interview data on the number of people who were successful on the first try, the second try, and so forth. Using survey data on the number of crossing attempts by migrants, collected

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

BOX 6-1
Network Sampling

Link-trace sampling and its variant, respondent-driven sampling (Salganik and Heckathorn, 2004; Thompson and Seber, 1996), have been used for sampling elusive, hard-to-reach populations, such as unregulated workers, the homeless, drug users, and sex workers. These populations are typically characterized not only by the absence of a serviceable sampling frame but also by the presence of social relations among members of the population. These social connections provide a means of reaching individuals in the population via the people they know. Such approaches are often informally referred to as "snowball sampling." Link-tracing sampling strategies such as snowball sampling and respondent-driven sampling are often used to leverage those social relations beyond the small population subgroup available to researchers. The initially selected group is referred to as the "seed" sample. Information about the social links of individuals in the seed sample to other members of the population is used to identify and contact members outside the original population subgroup available to researchers. In most applications of link-tracing sampling, the seeds are a convenience sample, so the probability of inclusion of any member is unknown. Therefore, a serious drawback of this type of sampling is that probability-based inferential methods are problematic. However, it is possible for the seeds in link-tracing designs to be selected randomly, even in applications to hard-to-reach populations—for example, by using a spatial sampling frame (Felix-Medina and Thompson, 2004).[a]

Respondent-driven sampling is a variation of link-tracing sampling in which the respondents themselves choose and contact people they know and invite them to participate in the survey (Heckathorn, 1997, 2007; Salganik and Heckathorn, 2004; Volz and Heckathorn, 2008). While it is possible to reduce the dependence of the final sample on the seeds, recent simulation studies (Handcock and Gile, 2010) suggest that substantial biases can remain. A common feature of networked populations is that they exhibit homophily by attributes; that is, the social ties are more likely to occur between people who have similar attributes. In the case of undocumented migrants, homophily might occur when the initial group of seeds

from family members in four Mexican states, they modified the initial geographic distribution into a Poisson regression model in which the response was the number of attempts before the first successful trip and the mean of the Poisson included covariates related to economic factors, gender, and other variables that might affect the success rate. They also (coarsely) corroborated their model's predictions against data from Mexican surveys and against the data on the number of people legalized under the Immigration Reform and Control Act of 1986 (IRCA). Their estimates suggest that, as of 1995, the U.S.–Mexico border was becoming increasingly porous and the probability of apprehension on any given attempt was about one-third and

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

includes men from the same geographic area in Mexico. While link-tracing designs are often effective at acquiring a sample, the degree to which data so collected can be considered a probability sample is unclear. To allow valid inference to the population, the designs need to be implemented carefully and the mechanism of selection of successive waves of the sample must be well understood. Recent research (Gile, 2011; Gile and Handcock, 2011) discusses model-based estimation methods introduced by convenience samples of seeds. However, methodological development lags the data collection efforts.

There have been some applications of link-tracing sampling to unauthorized border crossings. Two such efforts have been discussed in Chapters 3 and 4: the Mexican Migration Project and the Mexican Migration Field Research Program. Neither of the two surveys results in samples that can be used for quantifying migration flows. Respondent-driven sampling does not appear to have been systematically used in the context of Mexican migration. Morral and colleagues (2011) discuss the potential use of respondent-driven sampling to estimate the stock of undocumented migrants, the probability of eluding capture at the border, and other quantities associated with the migration process.

One approach to network sampling is to start with seeds who are recent immigrants from Mexico and use them to recruit other recent immigrants. Key questions asked of them would include how long they have been in the United States since their most recent crossing and how many recent immigrants they know. The recruitment "coupons" for this respondent-driven sampling would not need to be physical (e.g., an identifying number would be sufficient). Methods have been developed to estimate population size from respondent-driven sampling data (Handcock et al., 2011; Salganik et al., 2011). Given the size of the flows, it is likely that network scale-up methods, which are a form of post-stratification, would be the most effective means to estimate the population size. However, significant methodological development and empirical testing are required before these methods can be recommended.

[a]See Thompson (2002) for an informative general discussion of link-tracing designs.

falling. They concluded that about 98 percent of individuals who attempted to cross the border illegally were ultimately successful.

Massey and Espinosa (1997) extended this methodology in various ways—for example, by proposing a model for estimating the probability of a first trip and of recurrent trips. The extended model includes not only macroeconomic variables but also individual and household characteristics, migration experience of other members of the household, and macroeconomic variables from the community/country of origin. Additional variables, such as those related to the political and legal contexts in Mexico and the United States (the Bracero Program, the IRCA period, and so on)

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

might also have been considered. Nonetheless, the situation at the border has changed markedly since 1997, and the panel has no confidence that these older models, which antedate the drug corridors, modern enforcement technology, and other innovations, can provide good guidance for the current era. Since the older models are unlikely to have the correct form, it would probably be necessary to rebuild them rather than just refit them with new data.[1]

While the policy environment can be updated in a rebuilt model, another shortcoming of much survey-based regression type modeling is the endogeneity of many of the migration determinants. In the presence of endogenous covariates and dual causality, the ability to simulate counterfactuals is compromised. More recent studies along the lines of Massey and Espinosa (1997) address this problem by using instrumental variables estimation (Angelucci, 2012; Gathmann, 2008; Orrenius, 1999).

Wein and colleagues (2009) and Liu and Wein (2008) extended the probability modeling to include compartmental modeling. They describe a system of four submodels, each of which is tuned with historical data and which interact to produce a probabilistic model for immigration flows. The four submodels are as follows:

- a multinomial logit model, as in Ben-Akiva and Lerman (1985), which gives a probabilistic description of the choices made by undocumented border crossers, such as the location of the attempt;
- an enforcement model, which describes the probabilities of interdiction as a function of enforcement effort and resources;
- a repatriation model, which describes how an apprehended alien is returned to Mexico; and
- an economic model that accounts for how supply and demand affect the wages of unskilled immigrant workers.

These submodels can become arbitrarily more sophisticated, incorporating elements of game theory, queuing theory, and portfolio analysis. The analyst must solve systems of non-linear equations or differential equations. Chang and colleagues (2012) developed a practical computational tool for implementing this model.

The main difficulty with this approach, even in its most mathematically advanced form, is that it is difficult to tune the model from historical data since the various submodels rely on information that was collected using different designs, at different data scales, in different time frames, and with

[1]Since the changing relationship between the U.S. and Mexican economies will also have a non-linear effect on the incentive to migrate, the older models are likely to have moved outside the range in which approximate linearity would allow simple retrofitting to work.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

different degrees of precision. This makes it almost impossible to calculate the errors associated with model predictions. This difficulty is not unique to this multipart model, and Chang and colleagues (2012) consider it to be among the most practical strategies for assessing cross-border flow. In principle, however, this type of complex probability modeling could be useful to address several aspects of illegal migration flows, with the exception of the probability of a successful first attempt.

Regression Models

Multiple linear regression modeling is a standard tool in demography and econometrics. Regression has been applied to various problems in immigration studies, perhaps most pertinently by Lewer and Van den Berg (2007), but it is more commonly used to estimate the economic impact of undocumented laborers. When circumstances admit locally linear approximation to complex phenomena, regression models can be quite effective, even in nonlinear applications. They are easy to implement, and results are relatively robust to departures from model assumptions. Multiple regression is transparent: the coefficients are often directly interpretable, and standard statistical inference enables tests of those coefficients, sensitivity analysis, and the calculation of confidence intervals. In the context of illegal immigration, the response variable might be the total number of successful illegal crossing attempts in a month, or it might be a multivariate response, such as a vector of illegal crossings at each of a number of different locations.

To estimate immigration flows, an economics perspective would start by modeling them in terms of the difference in earnings realizable in the United States and the prospective migrant's current home; one such approach has already been briefly discussed in Chapter 5. Migration costs are then subtracted from the potential gains, where costs include actual travel expenses, foregone earnings, and the disutility of being away from home (often ameliorated by migrant networks, which can also be readily modeled with the right data). The model might allow for benefits and costs to vary by age, gender, and education level. For example, young men with long work horizons, greater facility for learning English, and facing less risk in an illicit border crossing would have greater migration incentives than older people, women, and high-education workers (the latter have relatively high earnings in Mexico).

There is a large literature on models for immigration forecasting. Howe and Jackson (2005) recently surveyed this area, describing and comparing methodologies that have been adopted in the United States, Canada, and various European countries. But they emphasized that "[t]he poverty of explanatory models in the current practice of immigration projection contrasts sharply with the abundance of theories proposed and discussed

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

by experts in a variety of social science and policy disciplines" (Howe and Jackson, 2005:19). This apparent gap suggests that there is potential for research on modeling approaches that formalize and combine the various theories on illegal immigration (e.g., those based on social networks, relative strength of dual economies, social capital effects, and policy analysis) into formal regression models whose accuracy can be assessed using historical data.

Nonetheless, much has already been done to apply regression methods, broadly defined, to immigration flows. McKenzie and Rapoport (2010), for example, used ordinary least squares regression on data from the 1997 National Survey of Population Dynamics (ENADID) to study how the education level of illegal immigrants depends upon demographic and economic variables. Their work responds in part to work by Orrenius and Zavodny (2005), who used a Cox proportional hazards regression model on data from the Mexican Migration Project to quantify the effect of economic conditions, border enforcement, and migrant networks on the education level of unauthorized border crossers. Massey and Espinosa (1997) undertook a broader examination of 41 covariates that, from one or more theoretical perspectives, might be linked to illegal immigration; their work was based on 25 samples drawn from border states in western Mexico. Even before that, Taylor (1987) built explicit economic models that used regression to estimate income gains that provided incentive for illegal immigration. There is a great deal more literature on the topic, some of which is discussed in Chapter 2. But the question that remains to be answered is whether these regression tools can produce estimates of illegal migration flows (or their components) with sufficient accuracy and timeliness to meet DHS's needs.

On the positive side, there are some grounds for optimism. DHS wants to estimate flows in the recent past, which tends to be easier than the forecasting problem that has driven much of the literature, especially that summarized in Howe and Jackson (2005). Also, the DHS has access to administrative data, which can help inform the social, economic, and political theories that have driven previous modeling efforts. On the negative side, much of the previous literature was developed using data from before 2005, and it is clear that the illegal immigration process has changed in important ways.

Nonetheless, using historical survey data and administrative records, a multiple regression model of some kind (including multivariate regression, principal components regression, Cox proportional hazards regression, and so on) could be fit to describe how a specific component of the total flow might depend on the explanatory variables. For example, if the component of interest were counts of males aged 17-30, then that variable would be used as the response, and a model would be fit that included such information as the expected difference in income opportunity between the United

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

States and Mexico, the level of interdiction effort, and perhaps such covariates as the cost of being smuggled, the size of the Hispanic population in the United States, and indicator variables for seasonality, which affects migrant farm labor and the home construction industry. The flow $Y_{i(t)}$ can be estimated from a national survey such as ENADID or the Mexican National Survey of Occupation and Employment (ENOE), albeit with the difficulties and limitations that have already been discussed in Chapter 4. The difference in expected income by age, education, and other individual attributes can be obtained from economic records, and the amount, distribution, and type of border enforcement are available from the administrative records. This kind of modeling can, in principle, be implemented for each demographic component. An estimate of the total flow is the sum of the estimates computed for each segment.

The simplest nonlinear model for immigration, the gravity model, assumes that the magnitude of the population flow between two locations is proportional to the product of the sizes of the populations in each of those locations and inversely proportional to some monotonically increasing function of the distance between the two locations (e.g., the square of the distance, as in the Newtonian model for gravitation in astronomy).[2] In the context of this study, the measure of distance would be supplemented by variables that reflect the amount of border security, the costs and dangers of traveling, and so forth. Also, the product of the sizes of the populations might be replaced by the "attractiveness" of the United States, probably measured in terms of economic advantage.

Regression models are popular statistical tools because they are relatively easy to fit and lend themselves to straightforward interpretations. Relative to other approaches, they also tend to make the least consequential assumptions. However, the regression modeling approach will not do a good job of tracking changing mechanisms of migration, and it is unlikely to provide fine geographic detail given the kinds of survey data currently available. Over time, as the migration process evolves and the measures of the inputs become outdated, the validity of the model will drift, and its predictive accuracy will surely decline.

Spatiotemporal Processes

A third standard approach is to model the correlation in flow data across time and space. A simple time series is a starting point, but it requires a great deal of faith in the model and its stability. For estimating immigra-

[2]See Sen and Pruthi (1983) for a discussion of the use of regression in fitting a gravity model to migration flows and Sen and Smith (1995) for a discussion of gravity models in general.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

tion flows, the natural time series formulation would disaggregate the total flow into more homogeneous flows, for example:

- men between the ages of 17 and 40, who are entering the United States to work as migrant farm laborers;
- pregnant women who are entering the United States to ensure that their child has U.S. citizenship;
- people who are joining family members already established in the United States (either legally or illegally); and
- other cases, such as drug smuggling, employees in the building trade, and so on.

A time series model would be fit to each such flow, probably with additional covariates to capture the dynamics of the process, such as the impact of the recent drop in the construction sector or changes in the level of enforcement activity at the border.

A time series model can easily capture the seasonality of farm labor demand and construction starts, but it will probably not do a good job of capturing changing dynamics of flows. Some aspects, such as changing levels of law enforcement, can be handled through transfer functions, since the times at which such interventions occur is known. But other aspects, such as the changing impact of criminal cartels on the immigration pipeline, will not. The time series literature is rich, and there are certainly strategies for handling some of these kinds of interventions, but those methods quickly become complex. Prado and West (2010) offer a recent survey of the area, with special emphasis on dynamic time series models, which seem likely to be the type of time series model most relevant to estimating immigration flows when there may be feedback effects (e.g., if increased flow triggers increased border security).

A common time series model is the autoregressive process. The simplest such model treats, say, the total flow $Y(t)$ at time t as a regression on the past, so that mean function $\mu(t)$ satisfies

$$\ln \mu(t) = \theta_0 + \Sigma_j \theta \ln \mu(t - j) + \varepsilon(t)$$

where $\varepsilon(t)$ is "white noise," t is the period of interest, and j is the number of prior time periods. If the time period is a month, then j might take on values from 1 to 12 when the model postulates that the flow in month t depends on the flows during the previous year. In the case of immigration flow, one expects that the coefficients θ_1 and θ_{12} are both positive, since it seems likely that recent secular trends are well forecasted by flow rate in the previous month, whereas seasonal effects are captured by flow rate during the same month in the preceding year. The coefficients for other months

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

may be well approximated by 0. In other words, the flow in, for example, May 2012 can be expected to be positively correlated with the flow in April 2012 (because of local conditions), and also with the flow in May of 2011 (because migration tends to be seasonal). This is a common formulation for time series models for employment and travel data. The flow in the previous month captures recent events and trends, such as changes in the U.S. economy; the flow in the same month of the previous year captures the annual cycles associated with home construction and migrant farm labor. However, it would surely be better to disaggregate the total flow by demographic characteristics, and then model those component flows separately.

A more sophisticated instantiation of this strategy is to build a spatiotemporal model for the flows. Such models extend time series analyses to include spatial correlation structure. Simple versions might allow association among flows in Mexican states for decisions about whether to immigrate, while more complicated versions might be able to capture discouragement at particular border-crossing locations, redirection of flow due to fences or smuggling cartels, and so on. In particular, this last aspect of redirection offers the possibility of modeling the "squeezed balloon" aspect of cross-border traffic, in which increased interdiction at one region simply relocates the flow to a less monitored region.

Specifically, an example of a standard spatiotemporal model is a Conditional Autoregressive (CAR) model. In this case it might model the flow at a particular time t and location s as a Poisson random variable with mean μ_{st}. The mean is then modeled as $\ln \mu_s t = \beta_0 + \beta_{1x1}(s, t) + \ldots + \beta_p x_p (s, t) + \varepsilon(s, t)$, where the log function is motivated as the natural link and the linear regression relates covariates to the response through regression, as previously discussed. Those covariates would typically include time series terms, such as the flow in the previous month or in the same month of the preceding year. The spatial structure can be incorporated in two ways: as covariates (e.g., the population size in regions of Mexico), or through correlation among the error terms $\varepsilon(s, t)$. The correlation structure is the most likely avenue for handling the "squeezed balloon" effect. (See Banerjee and colleagues (2004) for an extensive treatment of modern spatial modeling.)

As is typically the case, more sophisticated modeling approaches realize their promise of improved accuracy only when data are available at increasingly higher levels of resolution (both in space and time). Some detailed spatial information about migration experiences at the municipal level (intensity of out-migration) can be obtained from Mexican Census data. If it is possible to cross the spatial information on migration with other economic, social, and enforcement information (such as the intensity of operations of organized crime along the U.S.–Mexico border), then it could be possible to produce estimates of the number of attempts to cross by sector and by time period.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

To the panel's knowledge, appropriately tailored and tested spatiotemporal models have not been previously used for immigration flow processes. From that perspective, this modeling strategy may be better viewed as prospective research, rather than a reasonable plan for near-term implementation. A first step would require exploring the availability of the data needed to fit spatiotemporal models.

Simulation Models

The preceding discussion suggested that standard statistical methods, even with reasonably enhanced survey data, may not be adequate to completely satisfy the needs of DHS. More recently developed strategies include simulation models, of which agent-based modeling is an example. Although the basic idea behind agent-based modeling dates back to the early 1900s, the approach is computationally intensive and therefore did not become widely used until the 1990s. Today, agent-based modeling is used in many disciplines, including economics (e.g., Holland and Miller, 1991), military preparedness, battlefield management, and epidemiological planning (Caplat et al., 2008), to name a few. The approach has also been implemented to address the question of migration of human populations (Edwards, 2008).

Agent-based modeling represents a new strategy for modeling immigration flows using data from surveys and official records. Such models endow a set of artificial agents with "rules" and then observe the emergent behavior as the agents interact with each other and their environment. The notion of "agent" can be quite general. In weather forecasting, for example, agents can be cubic kilometers of atmosphere, which exchange temperature, pressure, and humidity according to the laws of physics. In traffic-flow modeling, the agents are automobiles, with probabilistic rules that prefer certain spacings, speeds, origins, and destinations, which are chosen by the programmer to mimic the known activity of the community under study.

The best-known example of agent-based modeling is the "Sugarscape" created by Epstein and Axtell (1996). In that program, agents are allocated at random on a flat plane where a nutrient, "sugar," grows at a fixed rate. In order to survive, the agents must consume the sugar, which they do at a faster rate than the sugar grows back. Thus, when the sugar is depleted, the agent must move to a new location where the sugar has not been consumed. The first layer of rules prescribes how agents move and generates migration patterns similar to those in hunter-gatherer societies. A second layer of rules creates two genders, which reproduce when sufficient resources are available; this leads to behaviors that reflect the differential equations seen in population dynamics. Higher-order layers enable barter economics and division of labor. The main point of Sugarscape is that seemingly complex

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

social behaviors can be generated by a handful of simple and transparent rules (Epstein and Axtell, 1996).

A significant difficulty with agent-based models is that their statistical properties are understudied. In general, it is not clear how one should make principled uncertainty statements about such models, nor how one can assess goodness-of-fit. On the other hand, these models enjoy a high level of face validity: if the rule sets are reasonable, then the model may seem more plausible than a model that encodes human behavior in complex mathematics. Also, agent-based models are easy to assess; if the emergent behavior is unreasonable, then the model is inadequate.

In the context of modeling immigration flows with an agent-based model, administrative data and surveys offer important opportunities for model tuning and falsification. For example, consider rules of the following kind:

- An agent decides whether to attempt illegal immigration according to a coin toss, where the probability of heads is a function of the agent's age, income, marital status, the distance from the U.S. border, and other relevant covariates.
- If the coin toss leads the agent to attempt to immigrate, then the agent tries a certain number of times, until discouragement, where the number of attempts is a probabilistic function of the agent's covariates.
- If the agent succeeds, then the agent will attempt to engage in various kinds of activity in the United States, such as migrant labor, home construction, joining a family member, and so on.

Obviously, these rules are simplistic and offered only as illustration. The important point is that one can tune these rules, in principle, according to data in the administrative records and surveys. If, in a given year, the age mix of those interdicted at the border does not match the mix generated by the agent-based model, then this indicates that the model is incorrectly specified. More directly, the data enable the modeler to fit the functions that determine how the covariates affect the coin toss, or how easily an agent with certain characteristics will be discouraged.

One can address the problem of making inference from agent-based models in at least two ways. One way is to do sensitivity analyses and see how the outputs vary across reasonable ranges of inputs. This is particularly useful given that certain important information (e.g., the probability of successfully crossing the border in the first, second, or later attempts) is not available. A second way is to build an emulator, which creates a mathematically simpler model that approximates the agent-based model. Using methods introduced by O'Hagan (2001) and developed by Gramacy

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

and Lee (2008) and Higdon and colleagues (2008), one can use Bayesian inferences to set credible regions on model outputs.

The advantages of agent-based models for inferring immigration flow are that the method is relatively easy to program, relatively easy to validate, and allows decision makers to flexibly explore "what if" scenarios. The disadvantages are that the methods for formal statistical inference are still under development and that building and fitting such a model requires expertise that DHS has yet to acquire. As discussed in Chapter 5, DHS would be able to cheaply and effectively "outsource" this analysis to the scholarly community if it were to make the administrative data from its enforcement database more widely available.

In the context of immigration modeling, the Secure Border Initiative (MITRE Corporation, 2008) attempted to produce a simulation model for cross-border traffic that is essentially an agent-based model. That model has been criticized for making ad hoc assumptions, and to the best of our knowledge it has not been retrospectively validated against historical data (Chang et al., 2012). Nonetheless, if DHS decides to pursue an agent-based model as a strategy for producing flow estimates, the Secure Border Initiative model is a natural starting point.

CONCLUSION

Existing surveys and administrative data sources do not suffice to estimate some important aspects of the migration process; two fundamental data gaps include the proportion of undocumented migrants who cross the border undetected and the proportion of migrants who were successfully deterred after one or more apprehensions. The use of modeling approaches informed by survey data and administrative data is therefore necessary for estimating the flows of unauthorized migrants across the U.S.–Mexico border. Any modeling approach, and the assumptions underlying it, will need to keep track of mechanisms of change and be continually validated against historical trends and data. Since all modeling approaches will have their limitations, there is also much that could be learned by comparing estimates from multiple methods.

Without access to DHS administrative data, the panel was unable to assess the strengths and weaknesses of each modeling approach in the context of estimating the components of illegal migration flows along the U.S.–Mexico border. If the panel had had access to these data, it might have been able to make some basic comparisons between the different approaches and gain some insight into the accuracy of the information obtained from surveys. As a specific example, consider the analysis carried out using EMIF-N data in Chapter 5. The panel found that several probability models appeared to fit the re-apprehension estimates from EMIF-N quite

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

well. If the apprehensions data from DHS had also been available, the panel might have been able, at least to some extent, to validate (or fail to validate) EMIF-N. The panel would also have been able to evaluate the impact of violating standard model assumptions (e.g., the assumption of a constant population size) on the performance of capture-recapture approaches. More generally, using administrative data collected over several time periods, the panel might have been able to fit models using earlier information and evaluate them by comparing their predictions to observed data from later periods. This out-of-sample validation approach would have allowed the panel to compare the predictive ability of different models and explore the importance of the various assumptions underpinning those models.

Although the panel was aware that DHS has been considering specific modeling approaches (e.g., capture-recapture methods using apprehensions data), it could not get access to the relevant technical reports commissioned by DHS. Because the broader scientific community has not hitherto been engaged with DHS in developing, applying, and continually refining specific modeling approaches, the evidentiary base to which the panel could refer was also limited. For all of these reasons, much of the discussion in this chapter was general in nature. As was discussed in Chapter 5, DHS would benefit from making the administrative data in its enforcement databases publicly available to the research community, even if it were necessary to protect potentially sensitive information through data masking, aggregation, and other such procedures.

- **Conclusion 6.1: Modeling approaches, and the assumptions underlying them, must keep track of changing mechanisms of migration and be continually validated against historical trends and data. Since all modeling approaches have their limitations, there is also much that could be learned by comparing estimates from multiple methods.**

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

References

Agresti, A. (2002). *Categorical Data Analysis* (2nd ed.). New York: John Wiley & Sons.

Amuedo-Dorantes, C., and Bansak, C. (2011). *The Effectiveness of Border Enforcement in Deterring Repetitive Illegal Crossing Attempts*. Unpublished manuscript. Available: http://www.shsu.edu/~eco_www/resources/documents/Amuedo-DorantesandBansakApril172007.pdf [August 2012].

Amuedo-Dorantes, C., and Pozo, S. (2005). On the use of differing money transmission methods by Mexican immigrants. *International Migration Review,* 39(3):554-576.

Angelucci, M. (2012). U.S. border enforcement and the net flow of Mexican illegal migration. *Economic Development and Cultural Change,* 60(2):311-357.

Anguiano Téllez, M.E. (2007). Detour through the Sonora-Arizona Desert: New routes of Mexican emigration to the United States. *The Journal of Latino-Latin American Studies,* 2(4):4-19.

Arnold, B.C., and Beaver, R.J. (1988). Estimation of the number of classes in a population. *Biometrical Journal,* 30(4):413-424.

Baillargeon, S., and Rivest, L. (2007). Rcapture: Loglinear models for capture-recapture in R. *Journal of Statistical Software,* 19(5):1-31.

Banerjee, S., Carlin, B., and Gelfand, A. (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton, FL: Chapman & Hall/CRC.

Bean, F., Espenshade, T., White, M., and Dymowski, R. (1990). Post-IRCA changes in the volume and composition of undocumented migration to the United States. Pp. 111-158 in *Undocumented Migration to the United States,* F. Bean, B. Edmonston, and J. Passel, eds. Santa Monica, CA: RAND.

Bean, F., Corona, R., Tuirán, R., and Woodrow-Lafield, K. (1997). *The Quantification of Migration between Mexico and the United States*. Team Report to Mexico/United States Binational Study on Migration. Available: http://www.utexas.edu/lbj/uscir/binpapers/v1-1bean.pdf [August 2012].

Ben-Akiva, M., and Lerman, S. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, MA: MIT Press.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. (1975). *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press.

Bohning, D. (2008). Editorial—recent developments in capture-recapture methods and their applications. *Biometrical Journal*, 50(6):954-956.

Borger, S., Cornwell, D., Hanson, G., and Roberts, B. (in press). An analysis of migrant smuggling costs along the Southwest border. *Contemporary Economic Policy*.

Brownell, P.B. (2010). Wages differences between temporary and permanent immigrants. *International Migration Review,* 44(3):593-614.

Bunge, J., and Fitzpatrick, M. (1993). Estimating the number of species: A review. *Journal of the American Statistical Association,* 88(421):364-373.

Bustamante, J.A. (1998). Some thoughts on perceptions and policies Mexico–United States labor migration flows: Some theoretical and methodological innovations and research findings. Pp. 810-867 in *Mexico–United States Binational Migration Study*. Mexico City: Mexican Ministry of Foreign Affairs and Washington, DC: U.S. Commission on Immigration Reform.

Bustamante, J.A., Jasso, G., Taylor, J.E., and Legarreta, P.T. (1998). Data sources. Pp. 709-745 in *Mexico–United States Binational Migration Study*. Mexico City: Mexican Ministry of Foreign Affairs and Washington DC: U.S. Commission on Immigration Reform.

Bustos, A. (2011, November 14-15). *Recent Mexican Data Sources for the Study of International Migration*. Presentation to the Workshop on the Feasibility of Estimating the Flow of Unauthorized Crossings, Committee on National Statistics, National Research Council, Washington, DC.

Calavita, K. (1992). *Inside the State: The Bracero Program, Immigration, and the I.N.S.* New York: Routledge.

Canales, A. (1999). Periodicidad, estacionalidad, duración y retorno. Los distintos tiempos en la migración México–Estados Unidos. *Papeles De Poblacion,* 22:11-41.

Caplat, P., Anand, M., and Bauch, C. (2008). Symmetric competition causes population oscillations in an individual-based model of forest dynamics. *Ecological Modeling,* 211:491-500.

Cardoso, L.A. (1980). *Mexican Emigration to the United States, 1897–1931: Socio-Economic Patterns*. Tucson, AZ: University of Arizona Press.

Chang, J.C., Kohout, R., Blacksten, H.R., and Benard, S. (2006). *CBP Apprehensions at the Border.* Arlington, VA: Homeland Security Institute.

Chang, J.C., Reilly, A., and Judson, D. (2012). *A Unified Model of the Illegal Migration System*. Paper presented at the 2012 Meetings of the American Economic Association, Chicago, IL. Available: http://www.aeaweb.org/aea/2012conference/program/meetingpapers.php [August 2012].

Chávez, S. (2011). Navigating the U.S.–Mexico border: The crossing strategies of undocumented workers in Tijuana, Mexico. *Ethnic and Racial Studies*, 34:1320-1337.

Chiquiar, D., and Hanson, G.H. (2005). International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States. *Journal of Political Economy,* 113(2):239-281.

Consejo Nacional de Población. (undated). *Encuesta sobre Migración en la Frontera Norte de México [Survey of Migration at the Northern Border of Mexico (EMIF-N)]*. Tijuana, Mexico: El Colegio de la Frontera Norte. Available: http://www.conapo.gob.mx/es/CONAPO/Encuestas [October 2012].

Cormack, R.M. (1989). Log-linear models for capture-recapture. *Biometrics*, 45(2):395-413.

Cornelius, W.A. (1992). From sojourners to settlers: The changing profile of Mexican migration to the United States. Pp. 155-195 in *U.S.–Mexico Relations: Labor Market Interdependence,* J.A. Bustamante, C.W. Reynolds, and R.A. Hinojosa Ojeda, eds. Stanford, CA: Stanford University Press.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Cornelius, W.A. (2001). Death at the border: Efficacy and unintended consequences of U.S. immigration control policy. *Population and Development Review,* 27(4):661-685.

Cornelius, W.A., and Salehyan, I. (2007). Does border enforcement deter unauthorized immigration? The case of Mexican migration to the United States of America. *Regulation and Governance,* 1(2):139-153.

Cornelius, W.A., Fitzgerald, D., and Borger, S. (2009). *Four Generations of Nortenos: New Research from the Cradle of Mexican Migration.* San Diego: Center for Comparative Immigration Studies at the University of California.

Cornelius, W.A., Fitzgerald, D., Lewin Fischer, P., and Muse-Orlinoff, L., eds. (2010). *Mexican Migration and the U.S. Economic Crisis: A Transnational Perspective*. San Diego: Center for Comparative Immigration Studies at the University of California.

Darroch, J.N., Fienberg, S.E., Glonek, G.F.V., and Junker, B.W. (1993). A three-sample multiple-recapture approach to census population estimation with heterogeneous catchability. *Journal of the American Statistical Association*, 88(423):1137-1148.

Dávila, A., Pagán, J., and Soydemir, G. (2002). The short-term and long-term deterrence effects of INS border and interior enforcement on undocumented immigration. *Journal of Economic Behavior and Organization*, 49(4):459-472.

Donato, K.M., Durand, J., and Massey, D.S. (1992). Stemming the tide? Assessing the deterrent effects of the Immigration Reform and Control Act. *Demography,* 29(2):139-157.

Durand, J., and Arias, P. (2005). *La vida en el norte: Historia e iconografía de la migración México–Estados Unidos*. San Luis Potosí, Mexico: El Colegio de San Luis.

Durand, J., and Massey, D.S. (2003). *Clandestinos: Migración México–Estados Unidos en los Albores del Siglo XXI*. Zacatecas, Mexico: Universidad Autónoma de Zacatecas.

Durand, J., and Massey, D.S. (2004). *Crossing the Border: Research from the Mexican Migration Project*. New York: Russell Sage Foundation.

Durand, J., Massey, D.S., and Parrado, E.A. (1999). The new era of Mexican migration to the United States. *Journal of American History,* 86:518-536.

Durand, J., Massey, D.S., and. Zenteno, R.M. (2001). Mexican immigration to the United States: Continuities and changes. *Latin American Research Review,* 36(1):107-127.

Durand, J., Massey, D.S., and Capoferro, C. (2005). The new geography of Mexican immigration. Pp. 1-22 in *New Destinations: Mexican Immigration in the United States,* V. Zúñig and R. Hernández-León, eds. New York: Russell Sage Foundation.

Economic and Social Data Service. (2008). *International Passenger Survey, 2008*. Available: http://www.esds.ac.uk/findingData/snDescription.asp?sn=5993 [June 2012].

Edmonston, B., Passel, J.S., and Bean, F.D. (1990). Perceptions and estimates of undocumented migration to the United States. Pp. 11-33 in *Undocumented Migration to the United States: IRCA and the Experience of the 1980s*, F.D. Bean, B. Edmonston, and J.S. Passel, eds. Washington, DC: Urban Institute Press.

Edwards, S. (2008). Computational tools in predicting and assessing forced migration. *Journal of Refugee Studies,* 21:347-359.

Efron, B., and Thisted, R. (1976). Estimating the number of unseen species: How many words did Shakespeare know? *Biometrika*, 63(3):435-447.

Epstein, J., and Axtell, R. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. Cambridge, MA: MIT Press.

Eschbach, K., Hagan, J., Rodriguez, N., Hernandez-Leon, R., and Bailey, S. (1999). Death at the border. *International Migration Review,* 33(2):430-454.

Escobar-Latapí, A. (1999). Low-skill emigration from Mexico to the United States. Current situation, prospects and government policy. *International Migration,* 37(1):153-182.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Espenshade, T.J. (1990). Undocumented migration to the United States: Evidence from a repeated trials model. Pp. 159-181 in *Undocumented Migration to the United States: IRCA and the Experience of the 1980s*, F.D. Bean, B. Edmonston, and J.S. Passel, eds. Washington, DC: Urban Institute Press.

Espenshade, T.J. (1994). Does the threat of border apprehension deter undocumented U.S. immigration? *Population and Development Review*, 20(4):871-891.

Espenshade, T.J. (1995a). Unauthorized immigration to the United States. *Annual Review of Sociology,* 21:195-216.

Espenshade, T.J. (1995b). Using INS border apprehension data to measure the flow of undocumented migrants crossing the U.S.–Mexico frontier. *International Migration Review*, 29(2):545-565.

Feliciano, C. (2005). Educational selectivity in U.S. immigration: How do immigrants compare to those left behind? *Demography,* 42(1):131-152.

Felix-Medina, M.H., and Thompson, S.K. (2004). Combining link-tracing sampling and cluster sampling to estimate the size of hidden populations. *Journal of Official Statistics,* 20:19-38.

Fienberg, S.E. (1972). The multiple recapture census for closed populations and incomplete 2k contingency tables. *Biometrika*, 59(3):591-603.

Fienberg, S.E., and Makov, U.E. (1998). Confidentiality, uniqueness, and disclosure avoidance for categorical data. *Journal of Official Statistics*, 14:385-397.

Fienberg, S.E., Johnson, M.S., and Junker, M.S. (1999). Classical multilevel and Bayesian approaches to population size estimation using multiple lists. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 162(3):383-405.

Fisher, M.J. (2011, February 15). *Testimony of Michael J. Fisher, Chief, United States Border Patrol, U.S. Customs and Border Protection, Department of Homeland Security before the House Committee on Homeland Security, Subcommittee on Border and Maritime Security, "Securing our Borders—Operational Control and the Path Forward."* Available: http://www.dhs.gov/ynews/testimony/testimony_1297796735363.shtm [March 2012].

Fitzgerald, D. (2009). *A Nation of Emigrants: How Mexico Manages Its Migration*. Berkeley and Los Angeles: University of California Press.

Foerster, R.F. (1925). *Racial Problems Involved in Immigration from Latin America and West Indies to U.S.* Washington, DC: U.S. Government Printing Office.

Fuentes, J., and García, O. (2009). Coyotaje: The structure and functioning of the people-smuggling industry. Pp. 79-102 in *Four Generations of Norteños: New Research from the Cradle of Mexican Migration*, W.A. Cornelius, D. Fitzgerald, and S. Borger, eds. La Jolla: Center for Comparative Immigration Studies at the University of California.

Galindo, C., and Ramos, L.F. (2008). Un nuevo enfoque para estimar la migración internacional de México. *La situación demográfica de México,* 45-72.

Galindo, C., and Ramos, L.F. (2009). Niveles y tendencias de la migración internacional en México. Pp. 93-126 in *El estado de la migración en México: Las políticas públicas ante los retos de la migración mexicana a Estados Unidos,* P. Leite and S. Giorguli, eds. México, DF: Consejo Nacional de Poblacion. Available: http://www.conapo.gob.mx/es/CONAPO/Las_politicas_publicas_ante_los_retos_de_la_migracion_mexicana_a_Estados_Unidos_ [August 2012].

Gamio, M. (1930). *Mexican Immigration to the United States: A Study of Human Migration and Adjustment*. Chicago, IL: University of Chicago Press.

Gathmann, C. (2008). Effects of enforcement on illegal markets: Evidence from migrant smuggling on the southwestern border. *Journal of Public Economics*, 92(10-11):1926-1941.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Genoni, M., Rubalcava, L., Teruel, G., and Thomas, D. (2011). *Mexicans in America*. Paper presented at the 2012 Population Association of America Annual Meeting, San Francisco, CA. Available: http://paa2012.princeton.edu/download.aspx?submissionId=120862 [July 2012].

Giorguli, S., and Gutiérrez, E. (in press). Migración entre México y Estados Unidos y desarrollo local en la coyuntura económica y política actual. *Hommes et Societés*.

Gile, K.J. (2011). Improved inference for respondent-driven sampling data with application to HIV prevalence estimation. *Journal of the American Statistical Association,* 106:135-146.

Gile, K.J., and Handcock, M.S. (2011). *Network Model-Assisted Inference from Respondent-Driven Sampling Data*. Available: http://arxiv.org/abs/1108.0298 [October 2012].

Good, I.J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika*, 40(3/4):237-264.

Goudie, I.B.J., and Goudie, M. (2007). Who captures the marks for the Petersen estimator? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170(3):825-839.

Goudie, I.B., Jupp, J.P.E., and Ashbridge, J. (2007). Plant-capture estimation of the size of a homogeneous population. *Biometrika*, 94(1):243-248.

Gramacy, R., and Lee, H. (2008). Bayesian treed Gaussian process models with an application to computer modeling. *Journal of the American Statistical Association,* 103:119-130.

Grieco, E., and Rytina, N.F. (2011). U.S. data sources on the foreign born and immigration. *International Migration Review*, 45(4):1001-1016.

Groves, R., Fowler, F., Couper, M., Lepkowski, J., Singer, E., and Tourangeau, R. (2004). *Survey Methodology.* New York: John Wiley & Sons.

Guzman, M., Haslag, J., and Orrenius, P. (2008). On the determinants of optimal border enforcement. *Economic Theory,* 34(2):261-296.

Hamilton, E.R., and Villarreal, A. (2011). Development and the urban and rural geography of Mexican emigration to the United States. *Social Forces,* 90(2):661-683.

Handcock, M.S., and Gile, K.J. (2010). Modeling networks from sampled data. *Annals of Applied Statistics,* 272(2):383-426.

Handcock, M.S., Gile, K.J., and Mar, C.M. (2011). *Estimating Hidden Population Size Using Respondent-Driven Sampling Data.* Available: http://arxiv.org/abs/1209.6241 [October 2012].

Hanson, G.H. (2006). Illegal migration from Mexico to the United States. *Journal of Economic Literature*, 44:869-924.

Hanson, G.H., and Spilimbergo, A. (1999). Illegal immigration, border enforcement and relative wages: Evidence from apprehensions at the U.S.–Mexico border. *American Economic Review,* 89:1337-1357.

Heckathorn, D.D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems,* 44:174-199.

Heckathorn, D.D. (2007). Extensions of respondent-driven sampling: Analyzing continuous variables and controlling for differential recruitment. *Sociological Methodology,* 37:151-207.

Henderson, T.J. (2011). *Beyond Borders: A History of Mexican Migration to the United States*. Malden, MA: Wiley-Blackwell.

Higdon, D., Gattiker, J., Williams, B., and Rightley, M. (2008). Computer model calibration using high-dimensional output. *Journal of the American Statistical Association*, 103:570-583.

Hoefer, M., Rytina, N., and Baker, B.C. (2011). *Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2010*. Washington, DC: U.S. Department of Homeland Security.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Hoefer, M., Rytina, N., and Baker, B.C. (2012). *Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2011*. Washington, DC: Office of Immigration Statistics, U.S. Department of Homeland Security. Available: http://www.dhs.gov/xlibrary/assets/statistics/publications/ois_ill_pe_2011.pdf [April 4, 2012].

Hoffman, A. (1974). *Unwanted Mexican Americans in the Great Depression: Repatriation Pressures, 1929–1939*. Tucson: University of Arizona Press.

Holland, J.H., and Miller, J.H. (1991). Artificial adaptive agents in economic theory. *The American Economic Review,* 81:355-370.

Hopper, K., Shinn, M., Laska, E., Meisner, M., and Wanderling, J. (2008). Estimating numbers of unsheltered homeless people through plant-capture and postcount survey methods. *American Journal of Public Health*, 98(8):1438-1442.

Howe, N., and Jackson, R. (2005). *Projecting Immigration: A Survey of the Current State of Practice and Theory*. Washington, DC: Center for Strategic and International Studies.

Ibarraran, P., and Lubotsky, D. (2007). Mexican immigration and self-selection: New evidence from the 2000 Mexican census. Pp. 159-192 in *Mexican Immigration to the United States*, G.J. Borjas, ed. Chicago, IL: University of Chicago Press.

Instituto Nacional de Estadística y Geografía. (2007). *Cómo se Hace la ENOE: Métodos y procedimientos*. Available: http://www.inegi.org.mx/est/contenidos/espanol/metodologias/enoe/ENOE_como_se_hace_la_ENOE1.pdf [August 2012].

Instituto Nacional de Estadística y Geografía. (2011a). *Censo de Población y Vivienda 2010: Marco conceptual* (México). Available: http://www.censo2010.org.mx [June 2012].

Instituto Nacional de Estadística y Geografía. (2011b). *Censo de Población y Vivienda 2010. Diseño de la muestra censal 2010* (México). Available: http://www.censo2010.org.mx [June 2012].

Instituto Nacional de Estadística y Geografía. (2011c). *Censo de Población y Vivienda 2010. Síntesis metodológica y conceptual del Censo de Población y Vivienda 2010* (México). Available: http://www.censo2010.org.mx [June 2012].

Instituto Nacional de Estadística y Geografía. (2012). *Migración Internacional a Partir de la ENOE al Cuarto Trimestre de 2011*. Available: http://www.inegi.org.mx/inegi/contenidos/espanol/prensa/Boletines/Boletin/Comunicados/Especiales/2012/marzo/comunica.pdf [August 2012].

Instituto Nacional de Estadística y Geografía. (undated-a). *Características Metodológicas de la Encuesta Nacional de la Dinámica Demográfica (ENADID) 1992*. Available: http://www.inegi.org.mx/est/contenidos/espanol/metodologias/encuestas/hogares/metodenadid92.pdf [June 2012].

Instituto Nacional de Estadística y Geografía. (undated-b). *Características Metodológicas de la Encuesta Nacional de la Dinámica Demográfica (ENADID) 1997*. Available: http://www.inegi.org.mx/est/contenidos/espanol/metodologias/encuestas/hogares/metodenadid97.pdf [June 2012].

Instituto Nacional de Salud Pública. (2008). *Diseño Muestral*. Available: http://www.sinais.salud.gob.mx/descargas/pdf/ENADID06_DisenoMuestral.pdf [July 2012].

International Working Group for Disease Monitoring and Forecasting. (1995a). Capture-recapture and multiple-record systems estimation I: History and theoretical development. *American Journal of Epidemiology*, 142(10):1047-1058.

International Working Group for Disease Monitoring and Forecasting. (1995b). Capture-recapture and multiple-record systems estimation II: Applications in human diseases. *American Journal of Epidemiology*, 142(10):1059-1068.

Izcara Palacios, S.P. (2012). Opinión de los Polleros Tamaulipecos sobre la Política Migratoria Estadounidense. *Migraciones internacionales*, 6(3):173-204.

Jones, R.C. (1998). Remittances and inequality: A question of migration stage and geographic scale. *Economic Geography,* 74(1):8-25.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Kessler, G. (2011, May 11). Obama administration boasting about border security. *Washington Post*. Available: http://www.washingtonpost.com/blogs/fact-checker/post/obama-administration-boasting-about-border-security/2011/05/10/AFj71ZkG_blog.html [June 2012].

Koslowski, R. (2002, April 25). *Information Technology, Migration and Border Control*. Presented at the Institute for Government Studies, University of California, Berkeley.

Kossoudji, S. (1992). Playing cat and mouse at the U.S.–Mexican border. *Demography,* 29(2):159-192.

Krishnamoorthy, K., and Mathew, T. (2009). *Statistical Tolerance Regions: Theory, Applications, and Computation*. Hoboken, NJ: John Wiley & Sons.

Laska, E.M., and Meisner, M. (1993). A plant-capture method for estimating the size of a population from a single sample. *Biometrics*, 49(1):209-220.

Lewer, J., and Van den Berg, H. (2007). A gravity model of immigration. *Economics Letters*, 99:164-167.

Lindstrom, D.P. (1996). Economic opportunity in Mexico and return migration from the United States. *Demography,* 33(3):357-374.

Lindstrom, D.P., and Lauster, N. (2001). Local economic opportunity and the competing risks of internal and U.S. migration in Zacatecas, Mexico. *International Migration Review,* 35(4):1232-1256.

Liu, Y., and Wein, L. (2008). A queueing analysis to determine how many additional beds are needed for the detention and removal of illegal aliens. *Management Science,* 54:1-15.

López Castro, G. (1997). Coyotes and alien smuggling. Pp. 965-974 in *The Binational Study on Migration between Mexico and the United States, Volume 3: Research Reports and Background Material*. Mexico City and Washington, DC: Mexican Ministry of Foreign Relations and U.S. Commission on Immigration Reform.

Lozano, F.A., and Lopez, M. (2010). *Border Enforcement and Selection of Mexican Immigrants in the United States*. IZA Discussion Paper No. 4898. Available: http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=1595539## [April 16, 2012].

Lydgate, J. (2010). Assembly line justice: A review of Operation Streamline. *California Law Review,* 98:481-544.

Manrique-Vallier, D., and Fienberg, S.E. (2008). Population size estimation using individual level mixture models. *Biometrical Journal*, 50(6):1051-1063.

Marcelli, E.A., and Ong, P.M. (2002, May). *2000 Census Coverage of Foreign-Born Mexicans in Los Angeles County: Implications for Demographic Analysis*. Paper presented at the 2002 annual meeting of the Population Association of America, Atlanta, GA.

Massey, D.S. (1990). Social structure, household strategies, and the cumulative causation of migration. *Population Index,* 56(1):3-26.

Massey, D.S., and Capoferro, C. (2004). Measuring undocumented migration. *International Migration Review,* 38(3):1075-1102.

Massey, D.S., and Capoferro, C. (2008). The geographic diversification of American immigration. Pp. 25-50 in *New Faces in New Places: The Changing Geography of American Immigration*, D.S. Massey, ed. New York: Russell Sage Foundation.

Massey, D.S., and Espinosa, K. (1997). What's driving Mexico–U.S. migration? A theoretical, empirical, and policy analysis. *American Journal of Sociology*, 102:939-999.

Massey, D.S., and Parrado, E.A. (1998). International migration and business formation in Mexico. *Social Science Quarterly,* 79(1):1-20.

Massey, D.S., and Pren, K.A. (2012). Unintended consequences of U.S. immigration policy: Explaining the post-1965 surge from Latin America. *Population and Development Review,* 38(1):1-29.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Massey, D.S., and Riosmena, F. (2010). Undocumented migration from Latin America in an era of rising U.S. enforcement. *Annals of the American Academy of Political and Social Science,* 630(1):294-321.

Massey, D.S., and Singer, A. (1995). New estimates of undocumented Mexican migration and the probability of apprehension. *Demography,* 32(2):203-213.

Massey, D.S., and Zenteno, R.M. (1999). The dynamics of mass migration. *Proceedings of the National Academy of Sciences*, 96(9):5328-5335.

Massey, D.S., and Zenteno, R. (2000). A validation of the ethnosurvey: The case of Mexico–U.S. migration. *International Migration Review*, 34(3):766-793.

Massey, D.S., Alarcon, R., Durand, J., and Gonzalez, H. (1987). *Return to Aztlan: The Social Process of International Migration from Western Mexico*. Berkeley and Los Angeles: University of California Press.

Massey, D.S., Arango, J. Hugo, G., Kouaouci, A., Pellegrino, A., and Taylor, J.E. (1998). *Worlds in Motion: Understanding International Migration at the End of the Millennium*. Oxford, UK: Clarendon Press.

Massey, D.S., Durand, J., and Malone, N.J. (2002). *Beyond Smoke and Mirrors: Mexican Immigration in an Age of Economic Integration.* New York: Russell Sage Foundation.

Massey, D.S., Pren, K.M., and Durand, J. (2009). Nuevos escenarios de la migración México-Estados Unidos. Las consecuencias de la guerra antiimigrante. *Papeles de Población*, 15(61):101-128.

McKenzie, D., and Rapoport, H. (2007). Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico. *Journal of Development Economics,* 84(1):1-24.

McKenzie, D., and Rapoport, H. (2010). Self-selection patterns in Mexico–U.S. migration: The role of migration networks. *Review of Economics and Statistics*, 92(4):811-821.

MITRE Corporation. (2008). *Secure Border Initiative (SBI), Decision Support Model of Immigration Systems: Detailed Model Description*. Spiral 3, Version 0.07. McLean, VA: MITRE Corporation.

Moraga, J.F.-H. (2010). New evidence on emigrant selection. *Review of Economics and Statistics,* 93(1):72-96.

Morral, A.R., Willis, H.H., and Brownell, P. (2011). *Measuring Illegal Border Crossing Between Ports of Entry: An Assessment of Four Promising Methods*. Santa Monica, CA: RAND Corporation.

National Gang Intelligence Center. (2011). *2011 National Gang Threat Assessment: Emerging Trends*. Available: http://www.fbi.gov/stats-services/publications/2011-national-gang-threat-assessment [August 2012].

National Research Council. (2009). *Principles and Practices for a Federal Statistical Agency, Fourth Edition*. Committee on National Statistics. C.F. Citro, M.E. Martin, and M.L. Straf, eds. Washington, DC: The National Academies Press.

National Research Council. (2011). *Budgeting for Immigration Enforcement: A Path to Better Performance*. Committee on Estimating Costs of Immigration Enforcement in the Department of Justice. S. Redburn, P. Reuter, and M. Majmundar, eds. Washington, DC: The National Academies Press.

Nuñez-Neto, B., and Viña, S. (2006). *Border Security: Barriers along the U.S. International Border*. Washington, DC: Congressional Research Service, Library of Congress.

Office for National Statistics. (2007). *Travel Trends 2006*. Available: http://www.statistics.gov.uk/downloads/theme_transport/TravelTrends2006.pdf [June 2012].

O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society, Series B (Statistics in Society),* 63:425-464.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Orrenius, P.M. (1999). *The Role of Family Networks, Coyote Prices and the Rural Economy in Migration from Western Mexico: 1965–1994*. Research Department Working Paper 9910. Dallas, TX: Federal Reserve Bank of Dallas.

Orrenius, P.M. (2001). *Illegal Immigration and Enforcement along the U.S.–Mexico Border: An Overview*. Available: http://www.dallasfed.org/assets/documents/research/efr/2001/efr0101a.pdf [August 2012].

Orrenius, P.M., and Zavodny, M. (2003). Do amnesty programs reduce undocumented migration? Evidence from IRCA. *Demography*, 40(3):437-450.

Orrenius, P.M., and Zavodny, M. (2005). Self-selection among undocumented immigrants from Mexico. *Journal of Development Economics*, 78(1):215-240.

Orrenius, P.M., and Zavodny, M. (2009). The effects of tougher enforcement on the job prospects of recent Latin American immigrants. *Journal of Policy Analysis and Management*, 28(2):239-257.

Orrenius, P.M., and Zavodny, M. (2010). *Beside the Golden Door: U.S. Immigration Reform in a New Era of Globalization*. Washington, DC: AEI Press.

Orrenius, P.M., and Zavodny, M. (2012). The economic consequences of amnesty for unauthorized immigrants. *Cato Journal*, 32(1):85-106.

Papademetriou, D.G., and Terrazas, A. (2009). *Immigrants in the United States and the Current Economic Crisis*. Washington, DC: Migration Policy Institute. Available: http://www.migrationpolicy.org/pubs/lmi_recessionJan09.pdf [March 2012].

Parrado, E.A. (2012). Immigration enforcement policies, the economic recession, and the size of local Mexican immigrant populations. *The ANNALS of the American Academy of Political and Social Science*, 641(1):16-37.

Passel, J.S. (2005). *Unauthorized Migrants: Numbers and Characteristics: Background Briefing Prepared for Task Force on Immigration and America's Future*. Washington, DC: Pew Hispanic Center. Available: http://www.pewhispanic.org/files/reports/46.pdf [March 2012].

Passel, J.S., and Clark, R.L. (1998, April). *Immigrants in New York: Their Legal Status, Incomes, and Taxes*. Washington, DC: Urban Institute Press. Available: http://www.urban.org/url.cfm?ID=407432 [June 2012].

Passel, J.S., and Cohn, D. (2009a). *Mexican Immigrants: How Many Come? How Many Leave?* Washington, DC: Pew Hispanic Center.

Passel, J.S., and Cohn, D. (2009b). *A Portrait of Unauthorized Immigrants in the United States*. Washington, DC: Pew Hispanic Center. Available: http://pewhispanic.org/files/reports/107.pdf [August 2012].

Passel, J.S., and Cohn, D. (2010). *U.S. Unauthorized Immigration Flows Are Down Sharply Since Mid-Decade*. Washington, DC: Pew Hispanic Center.

Passel, J.S., and Cohn, D. (2011). *Unauthorized Immigrant Population: National and State Trends, 2010*. Washington, DC: Pew Hispanic Center. Available: http://www.pewhispanic.org/files/reports/133.pdf [August 2012].

Passel, J., and Cohn, D. (2012, April 26). Why wave of Mexican migration stopped. *CNN Opinion*. Available: http://www.cnn.com/2012/04/26/opinion/passel-cohn-mexican-immigration/index.html [August 2012].

Passel, J.S., Van Hook, J., and Bean, F.D. (2004). *Estimates of the Legal and Unauthorized Foreign-Born Population for the United States and Selected States, Based on Census 2000*. Report to the Census Bureau. Washington, DC: Urban Institute Press.

Passel, J.S., Cohn, D., and Gonzalez-Barrera, A. (2012, April 23). *Net Migration from Mexico Falls to Zero—and Perhaps Less*. Washington, DC: Pew Hispanic Center. Available: http://www.pewhispanic.org/files/2012/04/PHC-04-24-Mexican-Migration.pdf [April 2012].

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Phillips, J.A., and Massey, D.S. (1999). The new labor market: Immigrants and wages after IRCA. *Demography,* 36(2):233-246.

Pledger, S. (2000). Unified maximum likelihood estimates for closed capture-recapture models using mixtures. *Biometrics*, 56(2):434-442.

Pledger, S., Pollock, K.H., and Norris, J.L. (2003). Open capture-recapture models with heterogeneity: I. Cormack-Jolly-Seber model. *Biometrics*, 59(4):786-794.

Prado, R., and West, M. (2010). *Time Series: Modeling, Computation and Inference*. Boca Raton, FL: Chapman & Hall/CRC.

Rendall, M.S., Aguila, E., Basurto-Dávila, R., and Handcock, M.S. (2009). *Migration between Mexico and the U.S. Estimated from a Border Survey*. Paper prepared for presentation at the 2009 Annual Meeting of the Population Association of America, Detroit, and the 2009 Metropolis British Columbia and Center for Research and Analysis of Migration Workshop on the Economics of Immigration. Available: http://mbc.metropolis.net/assets/uploads/files/Rendallnew.pdf [June 2012].

Rendall, M.S., Brownell, P., and Kups, S. (2011). Declining return migration from the United States to Mexico in the late-2000s recession: A research note. *Demography,* 48:1049-1058.

Reyes, B.I. (2004). U.S. immigration policy and the duration of undocumented trips. Pp. 299-320 in *Crossing the Border: Research from the Mexican Migration Project,* J. Durand and D.S. Massey, eds. New York: Russell Sage Foundation.

Reyes, B.I, Johnson, H.P., and Van Swearingen, R. (2002). *Holding the Line? The Effect of the Recent Border Build-up on Unauthorized Immigration*. San Francisco: Public Policy Institute of California. Available: http://www.ppic.org/content/pubs/report/R_702BRR.pdf [August 2012].

Riosmena, F. (2004). Return versus settlement among undocumented Mexican migrants, 1980 to 1996. Pp. 265-280 in *Crossing the Border: Research from the Mexican Migration Project*, J. Durand and D.S. Massey, eds. New York: Russell Sage Foundation.

Riosmena, F., and Massey, D. (2012). Pathways to El Norte: Origins, destinations, and characteristics of Mexican migrants to the United States. *International Migration Review*, 46(1):3-36.

Roberts, B., Hanson, G., Cornwell, D., and Borger, S. (2010). *An Analysis of Migrant Smuggling Costs along the Southwest Border.* Washington, DC: U.S. Department of Homeland Security. Available: http://www.dhs.gov/xlibrary/assets/statistics/publications/ois-smuggling-wp.pdf [August 2012].

Robinson, J.G., Ahmed, B., Gupta, P.D., and Woodrow, K.A. (1993). Estimation of population coverage in the 1990 United States Census based on demographic analysis. *Journal of the American Statistical Association,* 88(423):1061-1071.

Robinson, J.G., West, K.K., and Adlakha, A. (2002). Coverage of the population in Census 2000: Results from demographic analysis. *Population Research and Policy Review,* 21(1):19-38.

Rogers, A., Raymer, J., and Newbold, K.B. (2003). Reconciling and translating migration data collected over time intervals of differing widths. *The Annals of Regional Science,* 37(4):581-601.

Royle, J.A., and Dorazio, R.M. (2008). *Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations, and Communities*. London, UK: Academic Press.

Royle, J.A., and Dorazio, R.M. (2010). Parameter-expanded data augmentation for Bayesian analysis of capture-recapture models. *Journal of Ornithology*, 152(Suppl. 2):521-537.

Rytina, N.F. (2002). *IRCA Legalization Effects: Lawful Permanent Residence and Naturalization Through 2001*. Available: http://www.dhs.gov/xlibrary/assets/statistics/publications/irca0114int.pdf [April 2012].

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Salganik, M.J., and Heckathorn, D. (2004). Sampling and estimation in hidden populations using respondent-driven sampling. *Sociological Methodology*, 34:193-240.

Salganik, M.J., Fazito, D., Bertoni, N., Abdo, A.H., Mello, M.B., and Bastos, F.I. (2011). Assessing network scale-up estimates for groups most at risk of HIV/AIDS: Evidence from a multiple-method study of heavy drug users in Curitiba, Brazil. *American Journal of Epidemiology,* 174:1190-1196.

Sandoval, B.S., Martínez, L.F.R., and Hernández, I.U. (2011). *Migrantes Mexicanos aprehendidos y devueltos por Estados Unidos. Estimaciones y características generals*. Instituto Nacional de Migración. Available: http://www.inm.gob.mx/static/Centro_de_Estudios/Investigacion/Avances_Investigacion/APUNTES_N2_Septiembre2011.pdf [June 2012].

Santibáñez, J. (1997). Metodología de la Encuesta sobre Migración en la Frontera Norte de México. Pp. 206-229 in *Taller de Medición de la Migración Internacional*, J. Bustamante, D. Delaunay, and J. Santibáñez, eds. Tijuana, Mexico: Colegio de la Frontera Norte. Available: http://horizon.documentation.ird.fr/exl-doc/pleins_textes/pleins_textes_7/b_fdi_03_01/010010505.pdf [August 2012].

Santibáñez, J. (1999). Migratory flows between Mexico and the United States. Pp. 141-156 in *Free Markets, Open Societies, Closed Borders? Trends in International Migration and Immigration Policy in the Americas*, M. Castro, ed. Coral Gables, FL: North-South Center Press.

Sapp, L. (2011). *Apprehensions by the U.S. Border Patrol: 2005-2010*. Available: http://www.dhs.gov/xlibrary/assets/statistics/publications/ois-apprehensions-fs-2005-2010.pdf [March 2012].

Sassen, S. (1988). *The Mobility of Labor and Capital: A Study in International Investment and Labor Flow.* Cambridge, MA: Cambridge University Press.

Seber, G.A.F. (2002). *The Estimation of Animal Abundance* (2nd ed.). New York: The Blackburn Press.

Secretaría de Gobernación, Consejo Nacional de Población, Instituto Nacional de Migración, Secretaría de Relaciones Exteriores, Secretaría del Tragajo y Previsión Social, and El Colegio de la Frontera Norte. (2010). *Encuesta sobre Migración en la Frontera Norte de México, 2008*. Available: http://www.inm.gob.mx/static/Centro_de_Estudios/Publicaciones/descargas/EMIF_NORTE_2008.pdf [August 2012].

Secretaría de Gobernación, Consejo Nacional de Población, Instituto Nacional de Migración, Secretaría de Relaciones Exteriores, Secretaría del Tragajo y Previsión Social, and El Colegio de la Frontera Norte. (2011). *Encuesta sobre Migración en la Frontera Norte de México, 2009*. Tijuana, Mexico: El Colegio de la Frotera Norte. Available: http://www.conapo.gob.mx/es/CONAPO/Encuesta_sobre_Migracion_en_la_Frontera_Norte_de_Mexico_2009_ [September 2012].

Secretaría de Trabajo y Previsión Social. (1997). *Encuesta sobre Migración en la Frontera Norte de México 1993–1994*. Mexico City: Secretaría de Trabajo y Previsión Social.

Secretaría de Trabajo y Previsión Social. (1998). *Encuesta sobre Migración en la Frontera Norte de México 1995*. Mexico City: Secretaría de Trabajo y Previsión Social.

Secretaría de Trabajo y Previsión Social. (1999). *Encuesta sobre Migración en la Frontera Norte de México 1996–1997*. Mexico City: Secretaría de Trabajo y Previsión Social.

Sen, A., and Pruthi, R. (1983). Least squares calibration of the gravity model when intrazonal flows are unknown. *Environment and Planning, Series A*, 15:1545-1550.

Sen, A., and Smith, A. (1995). *Gravity Models of Spatial Interaction Behavior*. New York: Springer-Verlag.

Singer, A., and Massey, D.S. (1998). The social process of undocumented border crossing among Mexican migrants. *International Migration Review*, 32(3):561-592.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Siskin, A., Bruno, A., Nuñez-Neto, B., Seghetti, L.M., and Wasem, R.E. (2006). *Immigration Enforcement within the United States*. Report RL33351. Washington, DC: Congressional Research Service. Available: http://assets.opencrs.com/rpts/RL33351_20060406.pdf [September 2012].

Spener, D. (2004). Mexican migrant-smuggling: A cross-border cottage industry. *Journal of International Migration and Integration,* 5:295-320.

Spener, D. (2009). *Clandestine Crossings: Migrants and Coyotes on the Texas–Mexico Border*. Ithaca, NY: Cornell University Press.

Spener, D. (2010). *Clandestine Crossings: The Stories*. Ithaca, NY: Cornell University Press.

Stark, O., and Bloom, D.E. (1985). The new economics of labor migration. *American Economic Review*, 75(2):173-178.

Sudman, S., Sirken, M.G., and Cowan, C.D. (1988). Sampling rare and elusive populations. *Science*, 240(4855):991-996.

Taylor, J.E. (1987). Undocumented Mexico–U.S. migration and the returns to households in rural Mexico. *American Journal of Agricultural Economics*, 69:626-638.

Taylor, P., Lopez, M.H., Passel, J.S., and Motel, S. (2011, December 1). *Unauthorized Immigrants: Length of Residency, Patterns of Parenthood*. Washington, DC: Pew Hispanic Center. Available: http://www.pewhispanic.org/2011/12/01/unauthorized-immigrants-length-of-residency-patterns-of-parenthood/ [June 2012].

Thompson, S.K. (2002). *Sampling* (2nd ed.). New York: John Wiley & Sons.

Thompson, S.K., and Seber, G. (1996). *Adaptive Sampling*. New York: John Wiley & Sons.

United Nations Office on Drugs and Crime. (2010). *The Globalization of Crime: A Transnational Organized Crime Threat Assessment*. Vienna, Austria: United Nations.

U.S. Census Bureau. (2008). *U.S. Census Bureau's Budget Estimates as Presented to Congress, Fiscal Year 2009*. Available: http://www.osec.doc.gov/bmi/budget/09CBJ/Census%20FY%202009%20Congressional%20Justification.pdf [August 2012].

U.S. Census Bureau. (2009). Weighting and estimation. Chapter 11 in *Design and Methodology: American Community Survey*. Washington, DC: U.S. Government Printing Office.

U.S. Census Bureau. (2012). *American Community Survey: Sample Size—Data*. Available: http://www.census.gov/acs/www/methodology/sample_size_data/ [August 2012].

U.S. Department of Homeland Security. (2011a). *Budget in Brief: Fiscal Year 2012*. Washington, DC: U.S. Department of Homeland Security.

U.S. Department of Homeland Security. (2011b). *Yearbook of Immigration Statistics: 2010*. Washington, DC: U.S. Department of Homeland Security.

U.S. General Accounting Office. (1998). *Immigration Statistics: Information Gaps, Quality Issues Limit Utility of Federal Data to Policymakers*. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office. (1999). *Status of Southwest Border Strategy Implementation*. Washington, DC: U.S. Government Printing Office.

U.S. Government Accountability Office. (2006). *Illegal Immigration: Border-Crossing Deaths Have Doubled Since 1995; Border Patrol Efforts to Prevent Deaths Have Not Been Fully Evaluated*. Washington, DC: U.S. Government Printing Office.

U.S. Government Accountability Office. (2011a). *Overstay Enforcement: Additional Mechanisms for Collecting, Assessing, and Sharing Data Could Strengthen DHS's Efforts but Would Have Costs*. GAO-11-411. Washington, DC: U.S. Government Printing Office.

U.S. Government Accountability Office. (2011b). *Preliminary Observations on Border Control Measures for the Southwest Border*. GAO-11-374T. Washington, DC: U.S. Government Printing Office.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

U.S. Immigration and Naturalization Service. (2001, December). *Illegal Alien Resident Population: Estimates of the Undocumented Immigrant Population Residing in the United States (October, 1996)*. Available: http://www.dhs.gov/files/statistics/publications/archive.shtm#content [June 2012].

U.S. Office of Management and Budget. (2011). *Statistical Programs of the United States Government, Fiscal Year 2012*. Washington, DC: U.S. Government Printing Office.

Volz, E., and Heckathorn, D.D. (2008). Probability based estimation theory for respondent driven sampling. *Journal of Official Statistics*, 24(1):79-97.

Warren, R. (1994). *Estimates of the Unauthorized Immigrant Population Residing in the United States, by Country of Origin and State of Residence: October 1992*. Washington, DC: U.S. Immigration and Naturalization Service.

Warren, R., and Passel, J. (1987). A count of the uncountable: Estimates of undocumented aliens counted in the 1980 United States Census. *Demography*, 24(3):375-393.

Wein, L.M., Liu, Y., and Motskin, A. (2009). Analyzing the homeland security of the U.S.–Mexico border. *Risk Analysis*, 29(5):699-713.

Zenteno, R.M. (2011). *Recent Trends in Mexican Migration to the U.S.: The Mexico Perspective*. Paper presented at 2011 Annual Population Association of America Meeting, Washington, DC. Available: http://www.somede.org/documentos/zenteno-2011.pdf [June 2012].

Zenteno, R.M., and Massey, D.S. (1999). Especificidad *versus* representatividad: Enfoques metodológicos en el estudio de la migración mexicana hacia Estados Unidos [Specificity versus representativeness: Methodological approaches to the study of Mexico–U.S. migration]. *Estudios Demográficos y Urbanos*, 14(40):75-116.

Zolberg, A. (2006). *A Nation by Design: Immigration Policy in the Fashioning of America*. Cambridge, MA: Harvard University Press.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

Appendix A

Survey Questions About Migration and Border Crossing

This appendix provides examples of questions about migration and border crossing that are asked in the surveys discussed in Chapters 3 and 4.[1]

American Community Survey (ACS) and Current Population Survey (CPS)

- In what country was this person born?
  [There is a drop-down list of precoded countries. If the United States, skip to country of birth of parents. If Puerto Rico or outlying area, skip to country of birth of parents and then go to year of entry.]
- Were you born a citizen of the United States? *[Yes or no.]*
  [Asked if born outside of the United States, Puerto Rico, or in an outlying area—except if both parents were born in the United States, Puerto Rico, or in an outlying area.]
- Did you become a citizen of the United States through naturalization? *[Yes or no.]*
  [For people who report being U.S. citizens but were born outside the United States.]
- Did you become a citizen of the United States through naturalization? *[Yes or no.]*[2]

[1]The panel was unable to obtain relevant migration and border crossing questions from the Mexican Family Life Survey.

[2]The CPS does not ask for year of naturalization, but the ACS does.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

- When did you come to live in the United States?
  [For anyone who reported being born outside the United States.]

SOURCES: http://www.census.gov/acs/www/methodology/questionnaire_archive/ and http://www.census.gov/cps/methodology/questions.html.

Mexican Census—Basic Questionnaire (2010)

Individual questionnaire. Questions on migration:

- For all individuals in the household: In what state of Mexico or in what country was (NAME) born?
  - In the state of residence
  - In other state
  - In the United States
  - In other country (specify)
- For all individuals 5 years and older in the household: Five years ago, in June 2005, in what state of Mexico or in what country was (NAME) living?
  - In the same state
  - In other state
  - In the United States
  - In other country (specify)

Mexican Census Sample—Long Questionnaire (2010)

Individual questionnaire. Questions on migration:

- (Same as basic questionnaire)

Special module on international migration:

- During the last 5 years, that is, from June 2005 up to today, has somebody currently living with you or who lived with you (in this dwelling) left to live in another country? *[Yes or No.]*
- How many people?
- For each person mentioned:
  - When (NAME) left the last time, was he/she living with you? *[Yes or No.]*
  - Gender of (NAME)
  - Age of (NAME) when he/she left the last time (in years)
  - In what month and year did (NAME) go to live in another country the last time?

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

  - In what state of Mexico was (NAME) living when he/she left the last time?
  - To what country did (NAME) go?
    - United States
    - Another country (specify)
  - (NAME) is currently living:
    - In the United States
    - In another country
    - In Mexico
- In what month and year did (NAME) return to Mexico?
- Is (NAME) currently living here, in this dwelling? *[Yes or No.]*

Other migration information:

- For all people in the household 12 years and older:
  - Does (NAME) receive money from?
    - Government programs
    - Pension
    - Other people who live abroad
    - Other people who live in Mexico
    - Other sources

SOURCE: http://www.censo2010.org.mx/ (Translated from Spanish by the panel.)

National Survey of Occupation and Employment (ENOE)

No migration questions included in the first interview. For the subsequent interviews (four rounds), based on the household roster from the first interview, the questionnaire asks:

- I will mention all the people that were living in this household in the prior interview:
  - Indicate if they still live here.
  - Is there somebody else living here that is part of the household?
- For those who left:
  - What was the reason why (NAME) left?
    - Work
    - Study
    - Marriage or start of a union
    - Separation or divorce
    - Health problems
    - Meet with family

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

        - Insecurity
        - Dead
        - Other
    - To what state of Mexico or country did (NAME) go?
        - Same state
        - Other state
        - Other country
- For new arrivals:
    - What was the reason why (NAME) moved in?
        - Work
        - Study
        - Marriage or start of a union
        - Separation or divorce
        - Health problems
        - Meet with family
        - Insecurity
        - Newborn
        - Missed in the prior round
        - Other
    - Where did (NAME) come from?
        - Same state
        - Other state
        - Other country

Short questionnaire (administered to all household members 12 years and older):

- For those not occupied:
    - Have you tried to look for a job in another country or have you prepared to cross the border?
    - The last time you lost your job, what was your situation?
    - Returned or was deported from the United States
- For those working:
    - During the last 3 months, have you tried to look for a job in another country or have you prepared to cross the border?

Long questionnaire (administered to a random sample of household members 12 years and older):
(Aside from the questions included in the short questionnaire)

- For those working:
    - Did you move from another city or community to get or keep your current job?

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

  - In what state or country were you living before this change of residence?
- For all members:
  - In the last 3 months have you received economic support from somebody who lives or works in another country?

SOURCE: http://www.inegi.org.mx/est/contenidos/proyectos/encuestas/hogares/regulares/enoe/default.aspx (Translated from Spanish by the panel.)

Mexican National Survey of Population Dynamics (ENADID, 2006 and 2009 cross-sections)

For each individual living in the household at the time of the survey, provide information on:

- If the person lived in the United States a year prior to the survey, and the month/year of his/her return
- If the person lived in the United States 5 years prior to the survey

For each ("regular" or habitual) household member 5 years prior to the survey (whether or not the person is living in the household at the time of the survey), information on:

- Whether the person emigrated to the United States in the 5 years prior to the survey and, for those who did:
  - Month/year of emigration
  - U.S. state of destination
  - U.S. immigration documentation that the person carried on his/her last trip (within this time period)
  - Month/year of return from the United States (if applicable)
  - If returned to Mexico, if the person is still part of the household
  - Main reason for emigrating to the United States (only in 2006)

SOURCES: 2006 ENADID questionnaire, available at http://sinais.salud.gob.mx/demograficos/enadid/cuestionarios.html (April 4, 2012) and 2009 ENADID questionnaire, available at http://www.inegi.org.mx/est/contenidos/Proyectos/encuestas/hogares/especiales/Enadid/Default.aspx (April 4, 2012). (Translated from Spanish by the panel.)

Survey of Migration at the Northern Border (EMIF-N)

- Are you thinking of crossing to the other side? Are you thinking of doing it in the next 30 days?

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

- Through which Mexican city are you going to cross? Principally, why did you choose this city to cross? Did you or will you contract with a person (coyote, guide, smuggler, etc.) to help you cross the border?
- Where did you contract with the person to help you cross the border? How much did you agree to pay?
- Do you have documents to cross? Which documents do you have? Since when have you had them?
- Do you have documents to work?
- *PAST TRIPS*: In total, how many times have you crossed to the United States to work or look for work? What was the first time? What was the last time? In what month and year was the last time that you crossed to the United States? Through which city did you cross? On that occasion, did you have some type of document to cross to the United States? Did you carry some type of document to work? Did you contract with someone to help you cross the border? Where did you contract with that person?

SOURCE: http://www.colef.net/emif/ (Translated from Spanish by the panel.)

Mexican Migration Project (MMP)

Information about each undocumented border crossings of household head or another migrant from the household:

- Year of crossing
- Crossing place (city and state in Mexico)
- Crossed with whom
- Used coyote? How much paid? Who paid?
- Number of deportations
- Successful crossing?

SOURCE: Table K in http://mmp.opr.princeton.edu/databases/pdf%20surveys/MMP%20Ethnosurvey%20Version%20V%20Applied%20from%202007%20to%202011.pdf.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Mexican Migration Field Research Project (MMFRP)[3]

2011-2012

- For first and last trips, questions about:
  - Destination country and state
  - Year of arrival
  - Length of stay (months)
  - Occupation at destination
  - Position at work
- Have you ever been to or tried to go to the United States to live or work?
- Were you caught on your last trip across the border? Were you able to cross?
- Did you pass a checkpoint?
- How did you go through the checkpoint?
- Attempting to cross, have you been a witness to or victim of violence by:
  - A coyote?
  - Immigration officer?
  - Mexican police?
  - Mexican army?
  - Other?
- Currently, how hard do you think it is to evade immigration officers?
- Currently, how dangerous do you think it is to evade immigration officers?
- Do you personally know someone who died trying to cross the border?
- Now I will show you some pictures. Of the things shown in these drawings, which are the three that most concern a person crossing the border with papers and in what order?

2010-2011

- For first and last trips, questions about:
  - Year of arrival
  - Month of arrival
  - Documents
  - Number of times stopped during crossing
  - Used a coyote
  - U.S. dollars paid to coyote

[3]Questions vary from year to year.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

  - Duration of the trip to the United States (month)
  - First occupation in the United States
  - Position at work
  - Hours worked per week on the job
  - Total earnings
    - Dollar amount
    - Frequency
- How did you gather the money to pay for the trip?
- Which part of the border did you cross or try to cross?
- When you crossed or tried to cross the border, did you have papers or did you cross without papers?
- To enter the United States on your last trip, did immigration officers or customs officers
  - Stop you?
  - Inspect your car?
  - Search your body or clothing?
- During your last trip, how many times did you try to cross the border?
- How many times were you detained?
- Were you finally able to cross?
- In what year were you not able to cross?
- Did you pass a checkpoint?
- How did you go through the checkpoint?
- In what manner did you cross the border?
- In your last crossing or attempt, where did you meet the coyote who helped you cross?
- Who recommended the coyote?
- How satisfied were you with the coyote?
- Attempting to cross, have you witnessed an act of violence by:
  - A coyote?
  - Immigration officer?
  - Bandits?
- Attempting to cross, were you a victim of violence by:
  - A coyote?
  - Immigration officer?
  - Bandits?
- In your last trip, did you have to pay a bribe to the Mexican police to reach the United States?
- Currently, how difficult is it to evade immigration?
- Currently, if you don't have papers, how dangerous is it to cross the border?
- Did you personally know someone who went to the United States and died trying to cross the border?

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

- According to what you know, what is the best month to cross without papers?
- Now I will show you some pictures. Of the things shown in these drawings, which are the three most concerning to a person crossing without papers, in the order of concern?
  - Bandits or gangs
  - Extreme temperatures or other natural hazards
  - Immigration officer
  - Imprisonment if they catch you
  - Mexican police
  - Not finding work on the other side
  - The wall

SOURCE: Mexican Migration Field Research Project. (Translated from Spanish by the panel.) Questionnaire on file with panel.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

Appendix B

Review of Capture-Recapture Ideas for Measuring the Flow of Unauthorized Crossings at the U.S.–Mexico Border

GENERAL APPROACHES TO CAPTURE-RECAPTURE SAMPLING

Capture-recapture sampling (CRC) has a history reaching back at least to the 19th century (Bohning, 2008; Goudie and Goudie, 2007). It is often used to estimate the total number of individuals in a population. In its simplest form, an initial sample is obtained from the population and the individuals in the sample are "marked" in such a way that one can subsequently observe if the individual was in the sample. A second sample is obtained independently, and the number of individuals marked in the first sample is recorded. Under simplifying assumptions about the representativeness of marked individuals in both samples, the total number of individuals in the population can be estimated (Thompson, 2002). In the case of more than one recapture sample, the names "multiple-recapture," "multiple-system methods," or "multiple list" are often used.

CRC methods have a long history in the estimation of the abundance of biological populations, such as fish, birds, and mammals. More recently, they have been used to estimate the abundance of hard-to-reach human populations such as the homeless (Hopper et al., 2008; Laska and Meisner, 1993; Sudman et al., 1988) and to adjust for census undercounts of minorities (Darroch et al., 1993). For human populations, CRC methods are referred to as "dual-system methods" or "dual-list methods."

Let N be the population size, n and m be the initial and second sample sizes, and X be the number of marked individuals in the second sample. Intuitively, if the second sample is representative of the population as a whole, then the proportion of marked individuals in it will be close to

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

the proportion in the population. Thus, the size of the population can be estimated by equating these two proportions and solving for it: $N = mX/n$. This is the so-called Petersen estimator (Seber, 2002).

The International Working Group for Disease Monitoring and Forecasting (1995a, 1995b) provides an excellent discussion of classical capture-recapture ideas. Other good discussions are given by Seber (2002) and Thompson (2002:Chapter 18). In a special issue of an academic journal focusing on recent developments in CRC, an editorial by Bohning (2008) also succinctly describes the state of CRC research.

Log-linear models are important in demography and are very useful in analyzing CRC data (Bishop et al., 1975). Such models have been proposed to allow for departures from homogeneity of the capture probabilities between individuals and/or associations between the two sampling processes (Fienberg, 1972). The capture history of an individual can be classified into four categories based on observation or non-observation in the first and second sample. This can be represented by a four-cell multinomial model. If the capture probabilities of the individuals are homogeneous within each of the samples, then the maximum likelihood estimate of N is the integer part of the Petersen estimator. If the captures and recaptures are treated as separate factors, then the number of capture histories falling into the various categories can be modeled as Poisson or multinomial counts. Different estimators can be derived under different assumptions about the population and sampling processes. More importantly, log-linear models allow for (positive or negative) dependencies between the captures to be modeled, especially if there are multiple recaptures (Bishop et al., 1975). A good application of this approach when two recaptures are made is given by Darroch and colleagues (1993). Pledger (2000) developed a unified linear-logistic framework for fitting many of these models. Baillargeon and Rivest (2007) present an R package to estimate many capture-recapture models, focusing on those that can be expressed in log-linear form.

Other approaches tend to model the heterogeneity in specific forms, typically by incorporating random effects for them. Darroch and colleagues (1993) developed Rasch-type models for CRC in the context of human censuses and supplementary demographic surveys. They also developed log-linear quasi-symmetry models. Other extensions include methods of finite mixtures to partition the population into two or more groups with relatively homogeneous capture probabilities. Examples of these are the logistic-normal generalized linear mixed model and log-linear latent class models with homogeneity within the classes (Agresti, 2002:Sections 12.3.6, 13.1.3, 13.2.6).

Fienberg and colleagues (1999) integrate many of the above approaches for multiple-recapture or multiple-list data in developing a mixed effects approach (fixed effects for the lists and random effects for the individuals). This approach allows the modeling of the dependence between lists

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

and the incorporation of covariates. They develop Bayesian inference for their specification. Manrique-Vallier and Fienberg (2008) expand on this approach, modeling individual-level heterogeneity using a Grade of Membership model wherein individuals are postulated as mixtures of latent homogeneous but extreme "ideal" types.

Many populations, including that of unauthorized crossers, are open in the sense that the population experiences change during or between the sampling (e.g., births, deaths). Many of the models reviewed above implicitly presume the population is closed (i.e., have fixed and unchanging membership). For open populations, interest typically has focused on the case where the population is closed during the period of each capture and experiences immigration and mortality between the capture periods. Cormack (1989) reviews many of the classical models for this case. Pledger and colleagues (2003) extend these to allow for individual heterogeneity in survival and capture rates using a finite mixture formulation. These models are receiving continuous development (see the review by Royle and Dorazio [2010]).

CAPTURE-RECAPTURE APPLICATIONS TO UNAUTHORIZED BORDER CROSSINGS

The most direct expression of capture-recapture ideas as applied to unauthorized border crossings is the work of Espenshade (1990, 1995b) and Singer and Massey (1998). They develop simple CRC models in the context of apprehensions ("capture") and re-apprehension ("recapture") of unauthorized crossers. Specifically, Espenshade (1995b) models as a geometric distribution the number of crossings an individual makes until a successful crossing. Under assumptions that individuals continue to attempt crossings until they succeed, that the probability of success is the same for each attempt, and other strong assumptions, he derives the equivalent of the Petersen estimator for the number of unauthorized crossers. He does not develop measures of uncertainty of this estimate, nor does he tie the work into the broader CRC literature. This approach is similar in spirit to that of the "frequency of apprehension frequencies" discussed in Chapter 5. Chang and colleagues (2006) extend these methods to treat "discouragement" due to prior apprehension and "return and rentry" due to unobserved exit and reentry into the United States. However, the panel did not have access to their paper and therefore could not review it; the only available description was by Morral and colleagues (2011).

A variant of CRC is "red teaming," in which individuals are recruited to attempt to cross so as to get an estimate of the probability of apprehension. This is referred to as plant-capture in the ecological literature (Goudie et al., 2007).

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.

Appendix C

Biographical Sketches of Panel Members

Alicia Carriquiry (*Chair*) is distinguished professor of liberal arts and sciences, professor of statistics, and director of graduate education at Iowa State University. Her research is in applications of statistics in human nutrition, bioinformatics, and traffic safety. She has published more than 70 peer-reviewed articles in areas of statistics, economics, nutrition, bioinformatics, mathematics, and animal genetics. She is associate editor of the *Annals of Applied Statistics* and Editor of *StatProb*, an electronic encyclopedia of statistics and probability. She is an elected member of the International Statistical Institute, a fellow of both the American Statistical Association and the Institute of Mathematical Statistics, and a national associate of the National Research Council. She has served as vice president of the American Statistical Association, president of the International Society for Bayesian Analysis, member of the Executive Committee of the Institute of Mathematical Statistics, and member of the Board of Trustees of the National Institute of Statistical Sciences. She received an M.Sc. in animal science from the University of Illinois and an M.Sc. in statistics and Ph.D. in statistics and animal genetics from Iowa State University.

David L. Banks is a professor of the practice of statistics at Duke University. Previously, he worked in three federal agencies: the National Institute of Science and Technology, the Bureau of Transportation Statistics, and the U.S. Food and Drug Administration. His research centers on applied Bayesian statistics, including network models, problems in transportation statistics, adversarial risk analysis, metabolomics, and agent-based models. He is a past editor of the *Journal of the American Statistical Association*

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

and currently editor of *Statistics, Politics and Policy*. He has a Ph.D. in statistics from Virginia Polytechnic Institute and State University.

Peter Brownell is an associate social scientist at RAND Corporation. Prior to joining RAND, he was a visiting research fellow at the Center for U.S.–Mexican Studies and a guest scholar at the Center for Comparative Immigration Studies, both at University of California, San Diego. His primary research interest has been on immigrants and immigration, with a particular focus on migration between Mexico and the United States. Past projects have addressed Mexican immigrants' wages in the United States, the role of U.S. policy in structuring immigrants' labor market outcomes and decisions regarding migration and settlement, the effects of the recent recession on return migration flows to Mexico, and other topics concerning Hispanic immigration to the United States. He holds a Ph.D. in sociology from the University of California, Berkeley.

Stephen E. Fienberg is Maurice Falk university professor of statistics and social science in the Department of Statistics, the Machine Learning Department, and the Heinz College at Carnegie Mellon University. His principal research interests lie in the development of statistical methodology, especially for problems involving categorical variables. His recent research has focused on approaches appropriate for disclosure limitation in multidimensional tables and their relationship with bounds for table entries; estimating the size of populations, especially in the context of census taking; and Bayesian approaches to the analysis of contingency tables. He is an elected member of the American Academy of Arts and Sciences, the National Academy of Sciences, and the Royal Society of Canada. He is a member of the Editorial Board of the *Proceedings of the National Academy of Sciences*. He has a Ph.D. in statistics from Harvard University.

Mark S. Handcock is a professor of statistics at the University of California, Los Angeles, where he is also an affiliate of the California Center for Population Research. He previously taught at the University of Washington, Pennsylvania State University, and New York University. His work focuses on the development of statistical models for the analysis of social network data, spatial processes, and longitudinal data arising in labor economics. His research involves methodological development motivated largely by questions from the social sciences and demography. Recent research has focused on survey sampling techniques and missing data methods, especially for network data. He also works in the fields of distributional comparisons, environmental statistics, spatial statistics, and inference for stochastic processes. He served as associate editor of *Annals of Applied Statistics, Journal of the American Statistical Association,* and is a fellow of the American Sta-

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

tistical Association. He holds a B.Sc. in mathematics from the University of Western Australia and a Ph.D. in statistics from the University of Chicago.

Gordon Hanson is the Pacific Economic Cooperation chair in international economic relations at the University of California, San Diego, as well as director of the Center on Emerging and Pacific Economies. He holds faculty positions in the School of International Relations and Pacific Studies and the Department of Economics. He previously was on the economics faculty at the University of Michigan (1998-2001) and the University of Texas (1992-1998). He specializes in the economics of international trade, international migration, and foreign direct investment. His current research examines the international migration of skilled labor, how border enforcement affects illegal immigration, the impact of imports from China on the U.S. labor market, and the global determinants of comparative advantage. He is a research associate at the National Bureau of Economic Research and a co-editor of the *Review of Economics and Statistics*. He has an A.B. in economics from Occidental College and a Ph.D. in economics from the Massachusetts Institute of Technology.

Virginia Lesser is a professor and chair in the Department of Statistics and director of the Survey Research Center at Oregon State University. Her research interests are in sampling, survey methodology, environmental statistics, and applied statistics. She has written on non-sampling error, the effects of item and unit non-response on non-response error, and multi-mode surveys. She is a fellow of the American Statistical Association, an elected member of the International Statistical Institute, and member of the Technical Advisory Committee for the Bureau of Labor Statistics. She has a Ph.D. in public health and biostatistics from the University of North Carolina.

Pia Orrenius is assistant vice president and senior economist at the Federal Reserve Bank of Dallas. As an officer in the regional group, she analyzes the regional economy, manages the *Texas Business Outlook Surveys,* and serves as editor of *Southwest Economy*, a publication of the Federal Reserve Bank of Dallas. Her research focuses on the causes and consequences of Mexico–U.S. migration, unauthorized immigration, and U.S. immigration policy. She is a fellow of the Tower Center for Political Studies at Southern Methodist University and a research fellow at the IZA Institute of Labor in Bonn, Germany. She is also an adjunct professor in the executive M.B.A. program at Baylor University (Dallas campus). During the 2004-2005 academic year, she was a senior economist on the Council of Economic Advisers in the Executive Office of the President, where she advised the Bush administration on labor, health, and immigration issues. She holds a Ph.D. in economics from the University of California, Los Angeles.

Copyright © National Academy of Sciences. All rights reserved.

Estimating Illegal Entries at the U.S.-Mexico Border

Jeffrey S. Passel is a senior demographer at the Urban Institute. He previously served as principal research associate at the Urban Institute's Labor, Human Services, and Population Center. His expertise is immigration to the United States and the demography of racial and ethnic groups, and he has authored numerous studies on immigrant populations in America, undocumented immigration, the economic and fiscal impact of the foreign born, and the impact of welfare reform on immigrant populations. He holds a Ph.D. in social relations from Johns Hopkins University.

Fernando Riosmena is an assistant professor in the Geography Department and the Population Program of the Institute of Behavioral Science at the University of Colorado, Boulder. His research investigates how demographic processes are associated with the spatial and social mobility, well-being, and development in Latin American societies and with immigrant communities in the United States from those societies. His main research areas are immigrant health throughout different stages of the migration process and the role of U.S. immigration policy and social, economic, and environmental conditions in sending communities on the dynamics of migration between Latin America and the United States. He holds a Ph.D. in demography from the University of Pennsylvania.

Silvia Elena Giorguli Saucedo is director of the Center for Demographic, Urban, and Environmental Studies at El Colegio de Mexico. She has also taught at El Colegio de Mexico, the Universidad Nacional Autonoma de Mexico, and the Instituto Tecnologico Autonomo de Mexico. Her research focuses on adolescents and family structure, international migration, and the impact of population change in Mexico. She holds a Ph.D. in sociology from Brown University.

Copyright © National Academy of Sciences. All rights reserved.

COMMITTEE ON NATIONAL STATISTICS

The Committee on National Statistics (CNSTAT) was established in 1972 at the National Academies to improve the statistical methods and information on which public policy decisions are based. The committee carries out studies, workshops, and other activities to foster better measures and fuller understanding of the economy, the environment, public health, crime, education, immigration, poverty, welfare, and other public policy issues. It also evaluates ongoing statistical programs and tracks the statistical policy and coordinating activities of the federal government, serving a unique role at the intersection of statistics and public policy. The committee's work is supported by a consortium of federal agencies through a National Science Foundation grant.

Copyright © National Academy of Sciences. All rights reserved.

Copyright © National Academy of Sciences. All rights reserved.