The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title:	Research into Immigration and Crime: Advancing the Understanding of Immigration, Crime, and Crime Reporting at the Local Level with a Synthetic Population
Author(s):	Christopher Inkpen, Jonathan Holt, John Bollenbacher, Heather Deforge, Lilly Yu, Pranav Athimuthu, Nicole Jasperson, Renata Zablocka, Nick Kruskamp
Document Number:	310356
Date Received:	March 2025
Award Number:	2020-R2-CX-0027

Research into Immigration and Crime: Advancing the Understanding of Immigration, Crime, and Crime Reporting at the Local Level with a Synthetic Population

Final Report

Prepared for

U.S. Department of Justice
Office of Justice Programs
National Institute of Justice

Prepared by

Christopher Inkpen
Jonathan Holt
John Bollenbacher
Heather Deforge
Lilly Yu
Pranav Athimuthu
Nicole Jasperson
Renata Zablocka
Nick Kruskamp

RTI International
3040 E. Cornwallis Road, PO Box 12194
Research Triangle Park, NC 27709
www.rti.org

RTI Project Number 0217714

December 2024

Research into Immigration and Crime: Advancing the Understanding of Immigration, Crime, and Crime Reporting at the Local Level with a Synthetic Population

Final Report

Prepared for

U.S. Department of Justice
Office of Justice Programs
National Institute of Justice

Prepared by

Christopher Inkpen
Jonathan Holt
John Bollenbacher
Heather Deforge
Lilly Yu
Pranav Athimuthu
Nicole Jasperson
Renata Zablocka
Nick Kruskamp, PhD

RTI International
3040 E. Cornwallis Road, PO Box 12194
Research Triangle Park, NC 27709
www.rti.org

December 2024

Contents

Figures

Tables

Number **Page**

Executive Summary

Research into Immigration and Crime: Advancing the Understanding of Immigration, Crime, and Crime Reporting at the Local Level with a Synthetic Population

Introduction

This report, funded by the National Institute of Justice (award #: 2020-R2-CX-0027) and prepared by RTI International, describes the results of a National Institute of Justice-funded research study that uses advanced analytical methods and novel datasets to explore the complex relationship between immigration, crime, and crime reporting at the neighborhood level. The study, which employs crime and crime reporting data from ten jurisdictions across the United States paired with a synthetic population that estimates the unauthorized immigrant population, aims to provide an in-depth analysis at the Census tract level. Analyses focus on unauthorized immigration and its correlation with drug, property, and violent crime rates, while accounting for crime reporting in traditional and emerging immigrant destinations along with sites with low foreign populations.

Background

Despite persistent political discourse linking immigration and increases in crime, most academic research contradicts this notion, showing either a negative or null relationship between immigration and crime. At the individual level, first-generation immigrants tend to have lower arrest rates than native-born citizens. Yet this trend diminishes with subsequent generations, as the children of first-generation immigrants (i.e., second generation immigrants) are often arrested at similar rates to children of native-born citizens. However, few studies assess the relationship between documentation status and offending. Macro-level analyses that focus on crime and immigration in specific areas reveal that areas with higher immigrant populations often experience lower crime rates or that the prevalence of immigrants in an area is not associated with an increase in arrests. Yet these studies frequently omit distinctions in documentation status, as these data are often unavailable. Further, many macro-level analyses are conducted at the county, state, or city level, which may obscure relationships observed at local levels.

This study also attempts to control for the nuances of crime reporting among immigrant populations. Immigrant neighborhoods, especially those in emerging destinations, show lower rates of crime reporting. Trust in the police and fear of deportation are potentially significant factors influencing the likelihood of underreporting crimes. highlighting the importance of community-police relations. Strong police-community relations are crucial for public safety as they foster trust and cooperation, leading to more accurate crime reporting, effective law enforcement, and safer spaces.

Methodology and Data Sources

This study uses a variety of data sources to analyze the relationship between the presence of unauthorized immigrants in a Census tract and corresponding crime rates.

Synthetic Population Development

Traditional methods for estimating the unauthorized immigrant population in the United States rely on either demographic accounting or model-based survey imputation techniques. Defined simply, demographic accounting involves subtracting the estimated number of legally present immigrants from the total foreign-born population recorded in U.S. Census Bureau surveys. These techniques may also employ logical edits to large survey datasets that use characteristics like age, education, and place of birth to infer unauthorized status. Model-based survey imputation techniques combine data from different survey datasets to estimate unauthorized status in nationally representative surveys, using statistical techniques to merge information from surveys with immigration queries and those with extensive geographic detail. This study builds on the work of model-based imputation methods by developing models that predict unauthorized status in a survey dataset and applying them to a synthetic population of the United States, based on U.S. Census Bureau survey datasets. This approach allows for granular estimates of unauthorized immigrant populations at the Census tract level. By combining data from the Survey of Income and Program Participation (SIPP) and the American Community Survey (ACS), the study developed a robust model to predict unauthorized status and produced Census tract-level estimates of the unauthorized immigrant population for 2019. Validation efforts for these estimates included comparisons with county- and state-level estimates from sources like the Migration Policy Institute and Pew Research Center along with scaling local estimates to meet state-level figures.

Crime and Crime Reporting Data Sources

The study utilized Records Management System (RMS) and Calls for Service (CFS) data from multiple police jurisdictions in 2019. These data were standardized and cleaned to ensure consistency across different regions and different case management systems.

Analytical Approach and Findings

Ordinary least squares (OLS) linear regression models were employed to analyze the relationship between unauthorized immigration and drug, property, and violent crime rates while controlling for crime reporting in the Census tract. The study also incorporated controls for authorized immigrant populations and other known correlates of crime, such as measures of housing, income, unemployment rates, and population density. Results suggest that across the different jurisdiction types there is no significant association between the estimated unauthorized immigrant population and drug, property, or violent arrest rates when accounting for known correlates of crime and correcting for robustness tests. In addition, increases in the authorized immigrant population are found to be associated with lower crime rates for drug, property, and violent crime types in emerging immigrant destination jurisdictions and associated with lower drug crime rates in traditional immigrant destinations.

Supplementary Analyses

A supplementary analysis in this report focuses on the U visa program in one jurisdiction, which provides a pathway to legal status for immigrant victims of certain crimes, such as domestic violence, abduction, felonious assault, etc. The study found significant gaps between the estimated number of U visa-eligible crimes and the actual number of certifications, suggesting potential underutilization of the program.

Conclusion

The findings of this study have significant policy implications. Policymakers should be cautious about linking increases in crime rates to unauthorized immigrant populations without robust evidence. Further, the finding that the presence of authorized immigrants is associated with lower crime rates is consistent with the existing body of research but suggests the need to delve into how authorized and unauthorized immigrants qualitatively differ and why the association with crime rates may vary by documentation status.

Where possible, future studies should adopt nuanced approaches to examine the relationship between immigration and crime, exploring distinctions in arrest rates and crime reporting by country of origin, ethnic group, documentation status, and time in the United States. These studies can provide valuable information for law enforcement seeking to enhance police-community relations. In addition, the negative association between authorized immigration and crime rates confirms prior studies and suggests areas for further investigation that focus on understanding the mechanisms through which immigration status may impact offending.

1. Introduction

The relationship between immigration and crime is a contentious topic in public discussion, often fueled by heated political discourse. Yet most academic research on the topic in the United States typically supports the notion that there is a negative or null relationship between immigration and crime (Lee & Martinez, 2009); however, overviews of the field indicate that important methodological issues can limit our ability to draw inferences about this relationship. Ousey and Kubrin (2018) conducted a meta-analysis to examine the relationship between immigration and crime; their findings generally indicate that increased immigration is associated with reductions in violent crime rates, while highlighting the importance of considering geographic and temporal contexts in understanding this relationship. In addition, they identify issues with different measures of crime reporting and the limited ability to assess the potential impact of documentation status.

This report provides details on the development and implementation of a research project that uses novel methodologies and large datasets to provide a comprehensive assessment of the correlational relationship between unauthorized immigration, crime, and crime reporting at the Census tract level. The goal of this project was to produce a variety of tests on the association between the estimated population of unauthorized immigrants within a given Census tract and different crime rates within that tract. This report starts with a synthesis of the current literature on the association between immigration and crime. Chapter 2 provides a technical description of the methodology used to produce estimates of the unauthorized immigrant population at the Census tract level. Chapter 3 proceeds by detailing recruitment efforts for jurisdictions that participated in the project, along with a description of the Records Management System (RMS) incident records and Calls for Service (CFS) data for the jurisdiction. Chapter 4 begins by describing the analytical dataset along with the methodological approach for analyzing the association between the unauthorized immigration population, crime rates, and crime reporting before detailing findings. This chapter also includes a brief discussion of how the estimates of the unauthorized immigrant population can be used for other research purposes, here focusing on the estimates and how they are associated with U visa applications. Last, Chapter 5 discusses limitations of these analyses and draws conclusions and recommendations from the preceding analytical results.

The Immigration-Crime Nexus

The immigration-crime nexus has long been a topic of contention and a popular political tool, with historical perceptions often linking high immigration rates to increased crime (Miller, 2018; Taft, 1933). This long-posited and often politically useful association has evolved over time, with empirical analyses consistently challenging these notions. Martinez & Lee (2000) highlighted that the supposed association between immigration and crime was not born out by empirical evidence, even in the early 1900s. These findings have been further supported by studies that suggest immigration may actually suppress crime rates (Lee & Martinez, 2009; Martinez et al., 2010; Ousey & Kubrin, 2009; Wadsworth, 2010). These conclusions contrast public opinion, which often links increased immigration with crime (Hagan et al., 2008) or conflates an exaggerated perception of the size of the unauthorized immigrant population with perceived rises in crime rates (Wang, 2012). This section proceeds by describing the literature on the individual relationship between immigrant status and crime, as well as macro-level analyses that assess the relationship between increases in immigration and crime rates.

Immigration and Crime at the Individual Level

Studies that examine individual differences in offending by immigrant generation along with race and ethnicity generally find lower rates of offending for first-generation immigrants (i.e., people who are foreign-born) compared to native-born residents. This phenomenon is also associated with time spent in the United States, as Butcher & Piehl (1998) show that offending increases the longer someone lives in in the host country. This relationship persists when looking within ethnic groups, as Jennings and colleagues (2013) found that Hispanic first-generation immigrants offend less over the life course compared to their second- and third-plus generation Hispanic counterparts. Bersani (2014) used longitudinal survey data to illustrate that first-generation immigrants offend less during the transition from youth to adulthood, which is often a critical period in offending trajectories. Indeed, although first-generation immigrants are less likely to offend compared to their higher-generation counterparts, the second generation (i.e., children of the foreign-born) approach parity with the third-generation or higher in terms of offending prevalence (B. E. Bersani, 2013). This same pattern tracks within ethnicity when exploring across regional or national boundaries; one study shows that both first-generation immigrants of Mexican and other Hispanic origins offend at lower rates compared to the third-plus generation for both non-Hispanic white, non-Hispanic Black, and Hispanic individuals (Inkpen, 2024). Similarly, the second generation of both origin groups offends at similar rates to non-Hispanic white and Black members of the third-plus generation, providing support for the notion that subsequent generations “catch up” to the children of native-born residents, despite the first generation’s generally lower rates of offending. Further, although research on documentation status and offending is limited, Bersani and colleagues (2018) identified that unauthorized immigrants who are arrested and detained reported longer times to first arrest and lower post-arrest returns to criminal behavior when compared to detained authorized immigrants and detained native-born individuals.

Immigration and Crime at the Macro Level

Where individual-level analyses show lower rates of offending by immigrant generation, studies that examine the presence of immigration (i.e., increases in the foreign-born population) often show that the presence of more immigrants are associated with lower crime rates. When examining a large sample of metro areas in the United States between 1980 and 2010, Ousey and Kubrin (2009) found reductions in crime rates when observing increases in the foreign-born population. Similarly, Adelman and colleagues (2017) highlighted a negative relationship between violent and property crimes rates and the percentage of foreign-born residents in 200 metro areas in the United States. Subsequent analyses that used estimates of the unauthorized immigrant population as a percentage of the metro area population found a negative relationship between property crime and the presence of unauthorized immigrants, as well as a null relationship with violent crime, suggesting that unauthorized immigration is not on its face associated with increases in crime in American cities. Analyses that explore this relationship at a more granular level have found similar patterns. Lyons and colleagues (2013) found that the general negative relationship between immigration and crime at the Census tract level is bolstered when accounting for city-level measures of immigrant political opportunity. Akins and colleagues (2009) found that when controlling for traditional correlates of crime, recent immigration was not tied to increases in homicides in one Texas city. Similar findings persisted in Southern California, although Kubrin and colleagues(2018) found that disaggregating immigrant generation by ethnic origin provides a more accurate picture of the relationship at the neighborhood level. Although studies at the Census tract-level have traditionally identified a negative or null relationship, recent research using thousands of Census tracts suggest that the relationship follows these patterns but that the actual trend is curvilinear. Thus, this wide body of research at the macro level provides further support for the notion that the relationship between immigration and crime is negative or null.

Immigration and Crime Reporting

A commonly held assumption is that trust in the police is integral to promoting crime reporting, with those who have higher levels of trust in the institution being more likely to report crimes when they occur (Kääriäinen & Sirén, 2010). However, research suggests that levels of trust in the police have limited impact on crime reporting and that crime reporting is more associated with the nature of the crime and the victim’s age and sex, with women often more likely to report crimes (Timukaite & Buil-Gil, 2024). Similarly, rational choice theory suggests that individuals report crimes based on their perceptions of the costs and benefits of reporting it (Bowles, Reyes, & Garoupa, 2009). This calculus may be shifted for immigrants, especially when considering documentation status and the perceived need to balance the risk of deportation with benefits gained from reporting the crime. Research examining trust in the police, fear of deportation, and crime reporting for foreign-born Latina women in the United States suggests that fear of deportation is associated with lower levels of perceptions of procedural justice. Trust in the police was a more impactful measure when predicting likelihood of reporting a crime (Messing et al., 2015). This suggests that documentation status could impact crime reporting through lowered perceptions of procedural justice in the criminal-legal system. In fact, one study exploring the shift of resources in police priorities in Dallas, TX to focus immigration enforcement on individuals convicted of serious crimes (Priority Enforcement Program) showed a resulting increase in the likelihood of Latino or Hispanic individuals to report crimes (Jácome, 2022). This points to a potential moderating effect between immigration status and crime reporting, with perceptions of police involvement in deportation impacting the willingness of foreign-born people to report crimes.

Geographic studies of immigration and crime reporting further bolster the idea of a nuanced relationship between willingness to report crimes and one’s immigration status. One study using crime reporting and the proportion of foreign-born and noncitizens at the metro level shows an inverse relationship between the proportion of the foreign-born in the metro area and likelihood of reporting a crime (Gutierrez & Kirk, 2015). Using survey data from the area-identified National

Crime Victimization Survey, Xie and Baumer (2019) found—after controlling for racial composition and other factors associated with crime—that although traditional immigrant neighborhoods tend to exhibit similar rates of crime reporting to neighborhoods with low foreign-born populations, immigrant neighborhoods in emerging destinations display substantially lower rates of crime reporting. This study highlights the importance of understanding the context in which crime reporting can occur and the need to account where possible for the potential impact of crime underreporting in immigrant neighborhoods.

Producing Estimates of the Unauthorized Immigrant Population

The foreign-born population in the United States has grown substantially as a proportion of the population over the last half-century, rising from roughly 5% in 1970 (9.6 million people) to nearly 14% in 2022 (46.2 million individuals) (Azari et al., 2024). Roughly half of the country's foreign-born population live in the traditional immigrant destinations states of California, Texas, New York, Florida, and New Jersey; although half of the foreign-born population is from Latin America, the proportion of Asian and African immigrants has increased in recent years (Azari et al., 2024). National estimates of the unauthorized immigrant population generally place the figure at roughly 11 million people (Baker & Warren, 2024; Capps et al., 2020). However, recent estimates of the changing composition of the unauthorized population and their migration dynamics suggest that increases in the unauthorized immigrant population are due to fewer outflows rather than an increase in new arrivals (Van Hook, 2024). Reliable and robust methods of estimating the unauthorized immigrant population are key to understanding both the size and the characteristics of this group.

Methods of Estimating the Unauthorized Immigrant Population

Researchers have developed several methods for estimating the number of unauthorized immigrants in the United States, using demographic accounting, logical edits, and model-based survey imputation techniques. The demographic accounting method, also known as residual estimation, involves using official visa and immigration statistics from federal sources to estimate the unauthorized population. This is done by subtracting the number of legally present immigrants, derived from official government statistics on granted and active visa numbers, special immigration statuses (e.g., asylee status), and naturalizations from the total foreign-born population recorded in Census Bureau surveys such as the American Community Survey (ACS) (U.S. Census Bureau, 2022a). The remaining amount is presumed to represent the universe of the potential unauthorized immigrant population (Baker & Warren, 2024). Researchers also use logical rules to attribute a probable unauthorized status to survey respondents based on characteristics like age, place of birth, and educational attainment (J. Passel & Cohn, 2009; J. S. Passel, 2006; J. S. Passel et al., 2004).

In contrast, model-based imputation methods involve data fusion techniques that combine information from different survey datasets to estimate unauthorized status. For example, the Survey of Income and Program Participation (SIPP) (U.S. Census Bureau, 2022b) includes specific queries on immigration status that allow for the estimation of unauthorized immigration status but has limited geographic coverage, while the ACS provides extensive geographic detail. By merging these datasets, researchers can impute missing immigration status information across a larger sample. Put simply, this process involves training a predictive model of immigration status with the SIPP data using measures observed in both datasets, then applying it to the ACS microdata to estimate the number of unauthorized immigrants nationally and at different geographic levels (e.g., state or large metro areas) (Van Hook et al., 2015a). This approach allows for more localized and accurate estimates, which are essential for effective policy development and resource allocation (Capps et al., 2018a).

2. Estimating the Unauthorized Immigrant Population in Survey Data

Introduction

Assessing the relationship between unauthorized immigration and crime at the neighborhood level requires neighborhood-specific estimates of the unauthorized population, but granular projections of the unauthorized immigrant population are unavailable when using traditional methods of estimation. This chapter details the method for producing estimates of the unauthorized immigrant population at the Census tract level by using multiple U.S. Census Bureau survey products.

Accurate estimates of the quantity and geographic distribution of unauthorized immigrants in the United States are crucial for informed research, policymaking, and resource allocation. Traditional data sources of the unauthorized population often lack sufficient granularity and representation, inhibiting comprehensive understandings of how unauthorized immigrant populations impact labor markets, welfare systems, and local resource planning (Bachmeier et al., 2014; Clark et al., 2009). This data gap also complicates assessments of access to housing and health care, which are critical issues facing unauthorized immigrants (Arbona et al., 2010; Hall & Greenman, 2013). As described in the previous chapter, researchers have developed several methods to estimate the unauthorized immigrant population using a combination of demographic accounting, logical assignment of legal status, and nationally representative survey data. Despite this, local estimation is often unfeasible due to data limitations in the geographic coverage and sample sizes in the survey products used. This report introduces a novel methodology to improve local-level estimation of unauthorized immigrant populations by employing a survey-based imputation method complemented by synthetic population modeling. The result is a national dataset of estimates of unauthorized immigrants within each U.S. Census tract, which is scaled specifically for the jurisdictions participating in this project. This approach enhances the resolution and accuracy of demographic predictions and paves the way for more effective policy interventions and community engagement.

A Data Fusion Approach

The process of estimating unauthorized immigrant populations hinges on a data fusion approach using statistical imputation. This method integrates information from a “donor” dataset, namely SIPP, with a broader “recipient” dataset, ACS. The imputation method combines unique aspects of both surveys to infer unobserved variables, which are critical in estimating the unauthorized status of immigrants (Van Hook et al., 2015a).

Fundamentally, this process involves two stages: first, preparing SIPP data to model unauthorized status, using known indicators like educational background and employment sector to extrapolate this status onto noncitizens. SIPP, a Census Bureau-administered survey, provides essential immigration-related data, albeit with some missing variables that need imputation (Capps et al., 2018b). The second stage involves constructing a synthetic population reflecting Census tract demographics, enabling the spatial translation of these findings down to highly localized areas. Iterative Proportional Fitting (IPF) methods ensure that the synthesized demographic distributions align with real-world data. We used the model developed in the first stage to predict unauthorized status in the synthetic population, effectively providing the estimated probability of unauthorized status for each person in the tract. These probabilities were used to generate counts of the unauthorized population in each tract, county, and state. We also incorporated state-level topline estimates to ensure that counts at the tract level aggregate to external higher-level estimates derived using a similar methodology to generate state-level figures.

The synergy of statistical imputation with synthetic population modeling facilitates a highly granular view of unauthorized populations, unlocking insights previously occluded by data unavailability. This method equips policymakers and researchers with powerful tools for examining immigration patterns across varying geographic scales, providing a foundation for enhanced resource allocation and targeted policy development.

Data Sources

SIPP, administered by the U.S. Census Bureau, is a longitudinal survey that provides detailed information on the income, employment, and governmental program participation among U.S. residents, as well as a section on immigrant naturalization for foreign-born survey participants (U.S. Census Bureau, 2022b). To prepare the survey dataset for modeling, we pooled several recent SIPP waves (2008, 2014, 2018). SIPP asks foreign-born respondents about citizenship, and all noncitizens are asked about their status upon arrival to the United States. Furthermore, several panels of the SIPP—most recently 2008—asked foreign-born respondents who arrived as noncitizens whether they had since changed their status (U.S. Census Bureau, 2013). This variable—adjusted legal status since arrival—is missing for years 2014 and 2018. To account for this missingness, we employed multiple imputation (n = 10) to derive an estimate of this status measure in the 2014 and 2018 SIPP waves, using similar imputation strategies (i.e., the imputation model equation) used to impute adjustment of legal status in SIPP when deriving national estimates of the unauthorized population (Capps et al., 2018a).

Identifying Unauthorized Status in the SIPP

Following previous research (Greenman & Hall, 2013; Van Hook et al., 2015b), we inferred that immigrants arriving without possessing lawful permanent resident (LPR) status who do not adjust to LPR status represent the universe of foreign participants who may be unauthorized immigrants or legal temporary immigrants. Furthermore, we applied a series of logical filters to refine the pool of suspected unauthorized immigrants and remove likely authorized temporary immigrants from the pool of potentially unauthorized immigrants (Borjas, 2017) (see Table 2-1). For instance, recently immigrated foreign-born individuals with college degrees working in high-tech industries are assumed to have H1B visa status and are removed from the unauthorized pool (J. S. Passel & Cohn, 2018).

Table 2-1. Basis for Assigning Legal Status in SIPP—Each Respondent Assigned as "Naturalized," "Legal Noncitizen," or "Unauthorized"

Criteria	Description
Baseline criteria	▪ If naturalized à *Naturalized* ▪ If not naturalized AND (arrived as LPR OR adjusted status since arrival) à *Legal Noncitizen* ▪ If not naturalized AND (arrived as non-LPR AND did not adjust status since arrival) à *Unauthorized*
Additional criteria	If *Unauthorized* meets any of the following criteria, move to *Legal Noncitizen:* ▪ Reports holding certain types of visas (A, FM, J, H1, H1B, L1, G1, R1, O1/P1, J1 TN) ▪ Works in a specialty profession (computer, engineering, math, architecture, post-secondary education, law, health care, criminal justice, government) ▪ Arrived before 1980 (likely to have obtained legal status through the 1986 Immigration Reform and Control Act) ▪ Receives public assistance ▪ Served in the U.S. military ▪ Born in Cuba (likely able to obtain legal status through the Cuban Adjustment Act) ▪ Resides in public housing ▪ Married to an LPR ▪ Likely asylee/refugee status based on arrival year and county of origin

The dataset used for modeling unauthorized status consisted of foreign-born SIPP respondents who were either "legal noncitizen" or "unauthorized" (i.e., the "naturalized" group was removed). Thus, the predictive model of unauthorized immigrant status is framed as a binary classification to determine whether noncitizens are in the United States legally. The predictor variables in the model are education, health insurance coverage, region of birth, sex, region of residence, age, years since immigrating, income, English proficiency, and Hispanic origin (Table 2-2). Although more accurate models of unauthorized status can be developed using additional variables in SIPP, this task requires that we use measures that are observed both in the SIPP and in the ACS tract-level tables in order to use the data fusion approach described below. Some of our numeric variables were binned to categorical levels to match the levels of the ACS measures reported by the U.S. Census Bureau at the tract level.

Table 2-2. Summary Statistics of the Training Data Used to Develop a Binary Classification Model to Predict Authorized vs. Unauthorized Immigrants

Variable	Levels	Count	Fraction (%)
Response variable			
Legal authorization status	Authorized	11,058	80.53
	Unauthorized	2,674	19.47
Predictor variables			
Education	Less than high school	5,301	38.6
	High school	3,387	24.7
	Some college	2,168	15.8
	Bachelor's degree or higher	2,876	20.9
Health insurance	Insured	7,735	56.3
	Uninsured	5,997	43.7
Region of birth	Latin America	9,125	66.5
	Asia	3,019	22.0
	Europe	963	7.0
	Other	625	4.6
Sex	Male	6,732	49.0
	Female	7,000	51.0
State/region	California	3,051	22.2
	Texas	1,659	12.1
	Florida	970	7.1
	Pacific/Southwest	1,033	7.5
	Northwest	920	6.7
	Southeast	856	6.2
	Other	5,243	38.2
Age	< 18	470	3.4
	18+	13,262	96.6
Years in the U.S.	< 10	6,369	46.4
	10–19	4,160	30.3
	20–29	1,989	14.5
	30+	1,214	8.8

(continued)

Table 2-2. Summary Statistics of the Training Data Used to Develop a Binary Classification Model to Predict Authorized vs. Unauthorized Immigrants (continued)

Variable	Levels	Count	Fraction (%)
Income	No income	5,539	40.3
	< $25k	4,562	33.2
	$25k–$75k	2,842	20.7
	$75k+	789	5.7
English proficiency	Proficient	6,934	50.5
	Not proficient	6,798	49.5
Hispanic origin	Hispanic	7,800	43.2
	Not Hispanic	5,932	56.8
Total		13,732	100

Source: SIPP, Waves 2008, 2014, 2018

Modeling the Unauthorized Population in SIPP

To ensure the robustness of our modeling approach, the SIPP dataset was divided into training and test sets. Eighty percent of the data was allocated to the training set, which was used to develop and optimize the machine learning models, while the remaining 20% constituted the test set, reserved for evaluating the models' predictive performance on unseen data. This split was essential for assessing generalization capabilities, mitigating overfitting, and ensuring that our statistical imputation approach was both reliable and accurate across different applications (Hastie, Friedman, & Tibshirani, 2001).

A suite of machine learning classifiers was considered to model unauthorized immigrant status effectively. Linear models like logistic regression and naïve Bayes, along with non-linear approaches such as random forest and support vector machine (SVM), were evaluated. Each model offered unique strengths: logistic regression's interpretability, random forest's robustness to overfitting, and SVM's capability to handle complex variable relationships (Breiman, 2001). Random forest emerged as the superior model, with an area under the curve of 0.74 on cross-validation, demonstrating its efficiency in capturing non-linear interactions critical for this application.

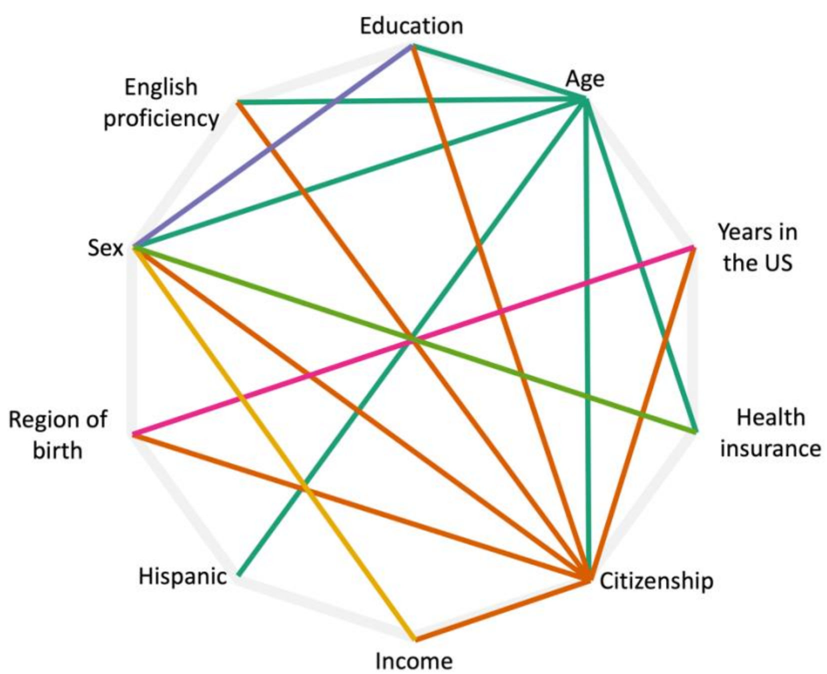
Synthetic Population Development

Synthetic populations serve as artificial datasets reflecting real-world statistical distributions without compromising individual privacy. That is, synthetic populations allow you to create individual-level datasets at a local level that match observed Census estimates but do not represent a real person. By maintaining key demographic characteristics, they enable researchers to explore scenarios that are otherwise restricted due to data limitations or privacy concerns (Beckman et al., 1996). In this study, synthetic populations allow us to create high-resolution estimates of unauthorized immigrant populations by using Census tract-level estimates from observed surveys (here, the American Community Survey) which can be used to predict an unobserved measure (i.e., unauthorized status) and allow researchers to generate local estimates of a measure that is never captured directly by the Census.

The construction of these synthetic populations relies on the IPF method, which adjusts sample weights iteratively to match observed marginal totals from ACS (Beckman et al., 1996; Fienberg et al., 1985). The IPF method effectively aligns synthetic dataset variables with ACS-reported marginal distributions, resulting in a dataset that reflects real-world characteristics at fine geographic scales. Importantly, synthetic populations provide a comprehensive picture of demographic distributions, offering valuable insights into potential policy impacts and resource allocation.

For this project, a synthetic population was developed for each U.S. Census tract, based on the ACS 5-year survey samples. To generate the synthetic population, we collected ACS data using the U.S. Census API (U.S. Census Bureau, 2022a). For each variable (age, education, income, etc.) we extracted tract-level counts. In addition, we obtained multiway contingency tables (i.e., two-way crosstabs) when available that reflected the marginal distributions between two or more variables. For cases where three- or four-way contingency tables were available, we collapsed them into multiple two-way tables. The complete set of intersecting tract-level relationships is detailed in Figure 2-1.

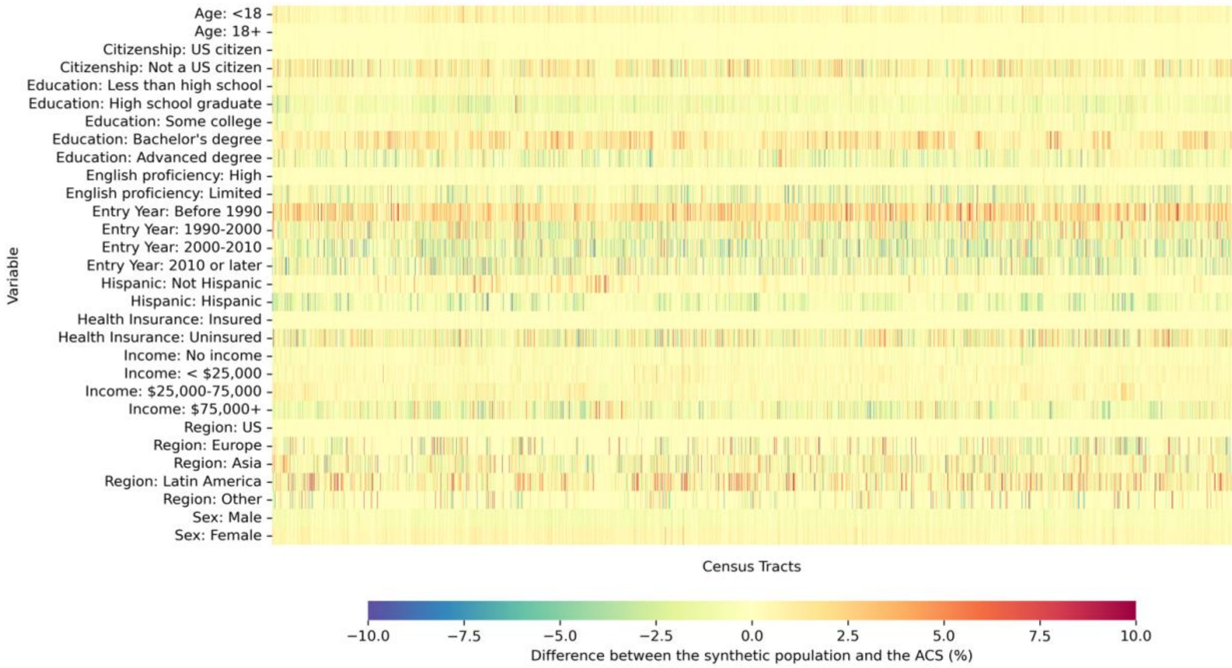
Figure 2-1. Contingency Tables (Crosstabs) Used to Construct the Synthetic Population



Note: Connecting lines indicate two-way tables downloaded from the ACS at the tract level. This collection of contingency tables was used to constrain the IPF algorithm.

Since IPF is a statistical algorithm designed to help one matrix approximate another, there is expected to be error associated with the synthetic population. To quantify the level of fidelity between the synthetic population and the true population, we compared the counts of each variable across Census tracts. The difference is reported as the percentage difference between the synthetic population and the ACS. Using the nation's largest state as a case study, we assessed the quality of the synthetic population for all Census tracts in California (Figure 2-2).

Figure 2-2. Accuracy of the Synthetic Population

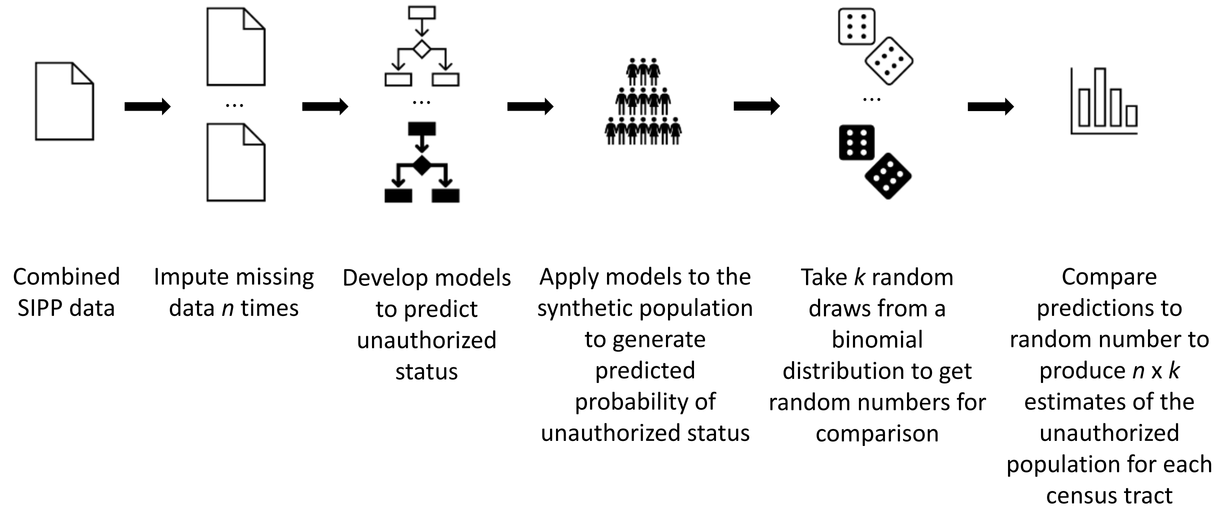


Note: Rows are the variables in the synthetic population, and columns are Census tracts. Here, we display all tracts in the state of California. Most counts of tract-variable combinations in the synthetic population fell within 10% of the ACS counts.

Imputing Unauthorized Status into Synthetic Population

Imputing unauthorized status within the synthetic population framework involves utilizing the predictive model developed from the SIPP dataset. This model leverages key demographic variables—such as education, income, and region of birth—to estimate the probability that an individual is unauthorized. A Monte Carlo simulation approach was employed, performing multiple draws to capture the variability and uncertainty inherent in the imputation process (see Figure 2-3). By simulating numerous iterations, with each individual assigned a probability-based unauthorized status, we generated a range of estimates, providing a robust measure of the unauthorized population for each U.S. Census tract.

Figure 2-3. Process for Estimating the Uncertainty Associated with Predictions of Unauthorized Status



Notes: Two sources of uncertainty were accounted for: (1) The imputation model that fills in missing data in the SIPP and (2) the classification model that predicts legal status into the synthetic population.

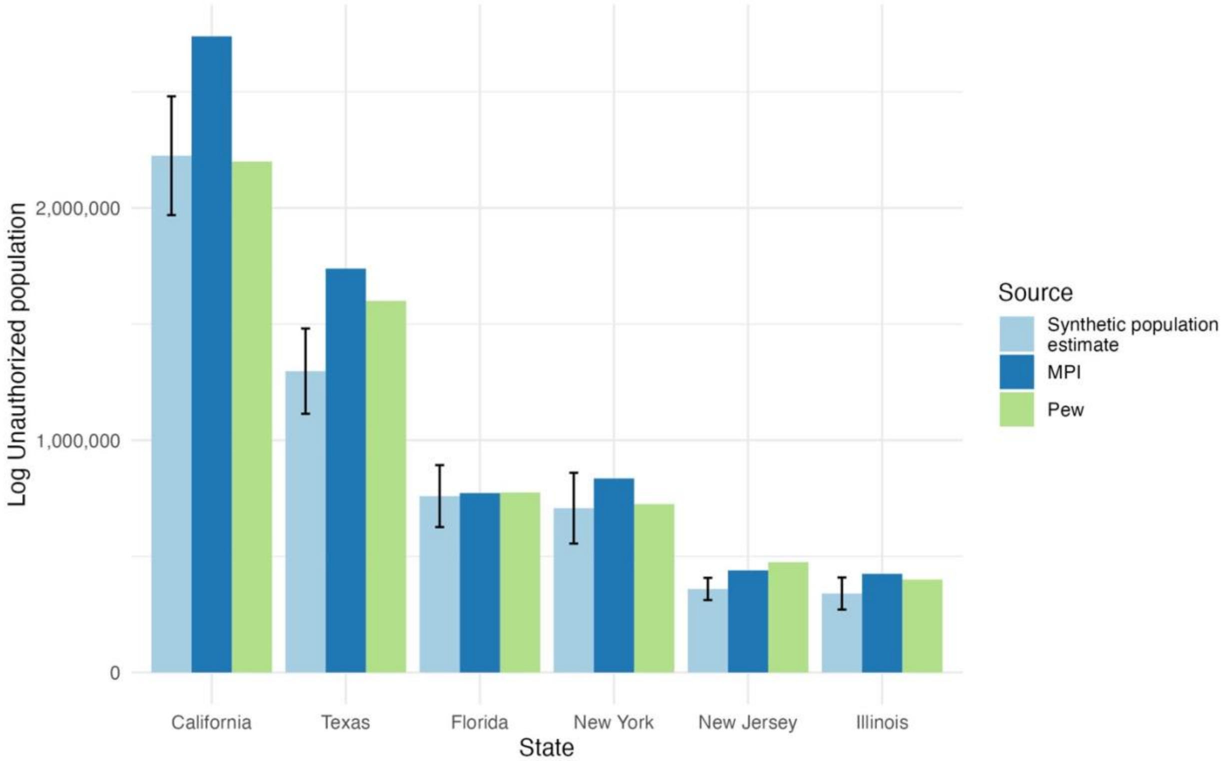
Validation Efforts

Validation is crucial to ensure the reliability and accuracy of the imputed estimates of unauthorized immigrants. This study employed a multifaceted validation strategy, comparing its tract-level and aggregated results with existing state- and county-level estimates from external well-regarded sources.

The model's state-level estimates were validated against those produced by the Migration Policy Institute (MPI) and Pew Research Center. These institutions provide comprehensive estimates of unauthorized populations, which serve as benchmarks for evaluating the proposed methodology's effectiveness. In this analysis, the generated state-level totals were aggregated from the tract-level estimates and compared to the MPI and Pew figures.

Figure 2-4 illustrates the close alignment between the imputed model's state-level estimates and those of MPI and Pew, highlighting the reliability of the approach across multiple states and further corroborating the model's robustness in capturing demographic nuances accurately.

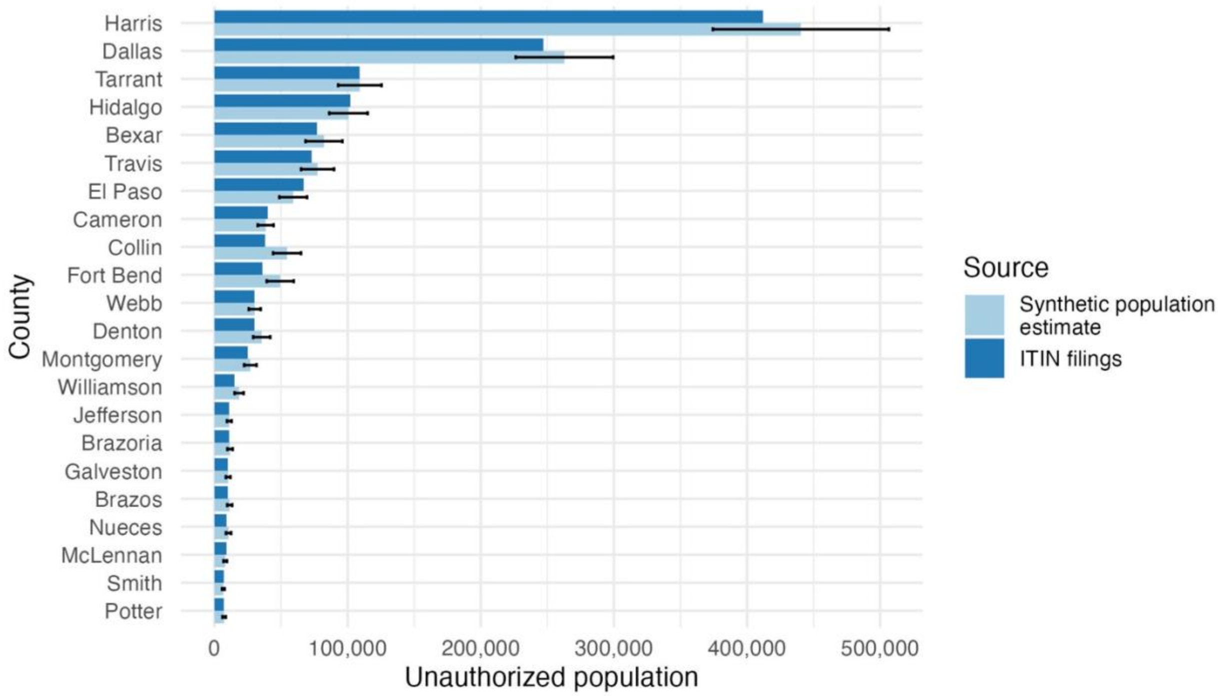
Figure 2-4. Estimated Counts of Unauthorized Immigrants in U.S. States



Further validation was conducted using Individual Taxpayer Identification Number (ITIN) filings, a proxy frequently used to approximate the size of the unauthorized immigrant population, as shown through data from Texas. ITIN filings provide an additional layer of validation, helping evaluate the model's effectiveness at the county level by comparing the predicted counts with ITIN-based estimates.

Figure 2-5 demonstrates reasonable congruence between the estimated unauthorized counts and ITIN filings across select Texas counties, reinforcing the model's accuracy at refined geographic scales. The comparative analysis validates the model's capacity to reflect actual conditions, showcasing its utility as a reliable tool for policy planning and demographic studies.

Figure 2-5. Estimated Counts of Unauthorized Citizens in Counties in Texas, Compared to an Approximation Based on ITIN Records



Note: Error bars reflect two standard deviations from the mean estimate.

Conclusion

This chapter outlines the methodology and validation efforts for estimating the number of unauthorized immigrants in the United States at the Census tract level. It describes the adaption of traditional model-based methods of imputation and classification models for filling in missing data and predicting legal status in the SIPP. In addition, we describe the process of generating a project-specific synthetic population prior to developing estimates of the unauthorized immigrant population within that dataset. The validation process includes comparing state-level estimates with those from reputable sources such as the MPI and Pew Research Center, and analyzing ITIN filings at the county level, particularly in Texas. The results show strong alignment with external benchmarks, affirming the model's accuracy and reliability for policy planning and demographic studies. This first step effectively provides Census tract-level estimates of the unauthorized immigrant population. This report now turns to a discussion of the police jurisdiction data necessary to identify crime rates at the Census tract level for a selection of jurisdictions.

3. Police Jurisdictions and Records Management System/Calls For Service Data

To explore the relationship between unauthorized immigration, crime, and crime reporting at the Census tract level requires detailed crime incident and crime reporting data that can be correctly allocated to the corresponding Census tract. Unfortunately, national coverage of crime data from either the Uniform Crime Report (UCR) (Lejins, 1965) or the National Incident-Based Reporting System (NIBRS) (Maxfield, 1999) rarely provide data at geographies lower than the state or, in some cases, the county. Testing the relationship at this more granular level of geographic resolution requires requesting and cleaning incident-level details and using geographic information to allocate incidents to their Census tract. This chapter describes the process of jurisdiction recruitment along with data management and cleaning of the incident-based crime records stored in agencies' Records Management System (RMS) and the computer-aided dispatch (CAD) Calls For Service (CFS) data.

Cohort Selection

Examining the relationship of immigration and crime at a national level is unfeasible for a number of reasons, chief among them that national crime data do not exist at the level of granularity necessary to produce useful, geographically specific estimates of the relationship. As such, RTI sought to recruit a diverse sample of police jurisdictions that included variety on geography as well as population size and type of immigrant destination. RTI partnered with a purposive sample of police jurisdictions designed to gather crime data from traditional and emerging immigrant destinations and places without an established immigration pattern (i.e., a comparatively low immigrant population). The research team privileged site selection based on (1) previous participation in the Police Data Initiative's Open Data Portal (Police Data Initiative, n.d.), (2) current collaboration of NIBRS or other incident-based records of crime, and (3) availability of CFS data. The research team created a memo to send to potential partner jurisdictions that included the goals of the project, expected outputs, research design, and requirements for each jurisdiction. The research team recruited more than 40 partner jurisdictions with the goal of selecting 10. This would ensure a balanced sample for subsequent analyses with a sufficient number of Census tracts to detect any relationship between unauthorized immigration and crime. Table 3-4 (seen on page 24) describes the 10 jurisdictions by population size, type of immigrant destination, and percentage of foreign-born residents, although data have been presented in ranges to preserve the anonymity of the sites.

Data Types and Measures

To explore the relationship of crime rates and crime reporting, we requested both RMS crime incident data and CFS data. This section describes the data types and measures included in these datasets as well as the cleaning and standardization process necessary to analyze often disparate data types in a uniform manner.

Incident-Based Crime Reports

Each jurisdiction provided the research team with both CAD/CFS data of crime reports or general emergency service calls, as well as RMS data. Each RMS dataset contained rows of individual crime or arrest incidents to which police responded. Some jurisdictions also record medical or fire emergency incidents in their RMS. Each column in the RMS data contains information about each incident, including the date and time the incident occurred, a report or event identification number, the type of event that occurred, the disposition (the outcome of the event, e.g., arrest, report written, transport to hospital), and event location information. Event location is typically provided as either an address or latitude/longitude coordinates and is often aggregated to the nearest 100 block level. For this analysis, we requested the private (i.e., non-jittered) latitude/longitude coordinates as accurate aggregation to the Census tract was integral for subsequent Census tract-level analyses.

CFS Data

Each jurisdiction provided numerous years of CAD or CFS data containing all recorded police activity in a jurisdiction, including all incoming 911 calls or other community-initiated police contact, police-initiated events, police administrative actions or events, traffic incidents, medical incidents, and fire incidents. Each CAD/CFS dataset contained rows of community-initiated police contacts (e.g., 911 calls or texts, flagging down an officer), as well as police-initiated activities (police administrative activity, traffic stops, proactive policing activities, police foot patrol stops, etc.), and fire- or medical-related 911 calls. The columns in the CAD/CFS data contained information about each event, including an event number, the date and time of the call/event, the type of event, and—in some cases—the disposition of the call, although not all jurisdictions record this in their CAD/CFS. In addition, location information in the form of an address or latitude/longitude was included, although in certain cases (described below) this field was missing and allocation to the Census tract was not viable.

Cleaning and Standardization

Each jurisdiction's RMS and CAD/CFS data were cleaned so as to contain all of the same variables and values across jurisdictions, allowing for cross-jurisdiction analyses. Some jurisdictions, for example, provided dispositions for their CAD/RMS data while others did not, requiring that analyses on call disposition be dropped or limited to a subsample of jurisdictions. Other jurisdictions collect much more detailed CAD information, but the data frames contained variables that were not needed or usable for the purposes of this study. Such variables were removed from all jurisdictions to create new data frames for each study site that contained all of the variables needed for subsequent analyses. These variables included event or report numbers (individual event identifiers), event dates and times, event descriptions/event types, event disposition, and event location (either an address or latitude/longitude coordinates) although identifying information was removed from records and geographic and date information were only used to aggregate incidents or calls to the relevant Census tract or year.

Due to the custom nature of CAD systems and data, each jurisdiction contained a wide variety of event types or event descriptions in their CAD and RMS data. To make cross-jurisdiction comparisons possible, the research team undertook a standardization process for the data. Based on previous efforts to develop a standardization schema for 911 call data (see generally (Langton et al., 2023; Ratcliffe, 2021; Wu et al., 2022), RTI developed a standardization schema that included primary and secondary categories for each event type. Each event was then hand-coded into the appropriate primary and secondary categories, shown in Table 3-1. The primary category describes the kind of event that took place. For example, the “administrative” category contains all the police administrative actions recorded in the CAD. Other primary categories include “investigative,” “community,” “medical/mental health,” “property crime,” “violent crime,” and “other crimes.” Each primary category contains several secondary categories that specify events such as “noise complaints,” “disturbance/disorder,” “neighbor/civil dispute,” “missing persons/runaway,” “lost/found property,” “mental health complaint,” “health emergency,” “vandalism/graffiti,” “disorderly conduct,” “theft/larceny,” “auto theft,” “burglary,” “assault,” “shots fired/non-fatal shooting,” and “robbery.” This standardization allowed the authors to create datasets that made cross-jurisdiction comparison of event types possible.

Table 3-1. Primary and Secondary Event Categories

Primary Category	Secondary Category	Description/Examples
Administrative	Court/warrant service	Time spent in court/serving warrants/other courts and warrants related activities
	Other agency assist	Assistance to other police agency
	Other miscellaneous activity	Uncategorized police activity in the CAD; recorded activities of officers
	Special detail	Police-initiated/-directed activities (e.g., directed patrol, hot spots policing, public events detail, school detail)
Community	Abandoned vehicle	Vehicles abandoned/calls for towing
	Alarm	Alarms that contact police/calls about alarms
	Lost/found property	Lost or found property reported to police
	Missing person/runaway	Calls about missing persons or runaways
	Noise complaint	Calls about noise complaints/nuisances
	Other disturbance/ disorder	Uncategorized miscellaneous calls about disturbances/disorder (e.g., loitering, littering, broken windows)
	Public assistance	Police assist the public (e.g., motorist assist, take reports, lockout assist)

(continued)

Table 3-1. Primary and Secondary Event Categories (continued)

Primary Category	Secondary Category	Description/Examples
Investigative	911 hangup	Calls that came into 911 but were disconnected, made in error, hung up, all silent, etc.
	Follow-up	Investigative follow-ups, interviews, etc.
	Suspicious activity	Calls about suspicious activity (suspicious person/vehicle/package)
	Welfare check	Checks for personal welfare, property and building checks, open window/door checks
Medical/Mental	Death	All deaths except crime-related deaths
	Health emergency	Medical emergencies
	Mental distress	Mental crisis emergencies
	Suicidal	Suicidal persons emergencies
Other Misc. Crimes (Non-Violent)	Disorderly conduct (criminal)	Calls about or logged officer responses to disorderly conduct where criminal charges are applicable
	Domestic/family (non-violent)	Calls about domestic incidents or disputes, non-violent
	Drug offense	Calls or logged events having to do with drugs/drug crime/drug paraphernalia/etc.
	Fraud/financial	Fraud or financial related crimes
	Kidnapping/ abduction	Calls about kidnappings/abductions
	Other miscellaneous crimes (non-violent)	Other uncategorized, minor, non-violent crimes (littering, curfew violations, etc.)
	Other sex offense	Other uncategorized sexual crimes (non-rape; indecent exposure, possession of pornographic materials, etc.)
	Threats/ intimidation	Harassment, threats, bribery, intimidation, harassing phone calls, etc.
	Trespass	Forbidding a person(s) from entering/continuing to enter a property/grounds
	Vandalism/ graffiti	Defacing, coloring, writing, otherwise damaging public or private property
	Weapon offense	Offenses regarding unlawful possession/use of weapons
Property Crime	Arson	Calls about arson incidents
	Auto theft	Motor vehicle thefts
	Burglary	Calls about burglaries/breaking and entering
	Theft/larceny	Any thefts/larceny/shoplifting (not including robbery)

(continued)

Table 3-1. Primary and Secondary Event Categories (continued)

Primary Category	Secondary Category	Description/Examples
Traffic	Collision	Traffic crashes
	Direction/point control	Officers logging time spent directing traffic in events where needed (signals out, accident, road work, etc.,)
	DUI	Driving under the influence, drugs or alcohol
	Enforcement	Traffic enforcement/traffic stops
	Roadway hazards	Hazards in the roadway (branches, wires, poles, trees, flooding, abandoned vehicle, debris, etc.)
Violent Crime	Assault	Any assault incidents
	Domestic/ family (violent)	Calls about domestic incidents, violent
	Homicide	Murder/attempted/suspected murder
	Non-fatal shooting	Shots fired incidents where injuries are unknown or unconfirmed
	Rape	Reports of rape
	Robbery	Forcible taking of property with violence or threat of violence (carjacking included)

RMS Measures

Each site's RMS data included rows of incidents to which police or fire/medical personnel responded. The types of events recorded in the RMS varied according to each police department. Each column in the RMS data contained information about each incident, including the date and time the incident occurred. Some sites provided the date and time the event was reported, a report or event identification number, the type of event that occurred, the disposition (outcome of event, i.e., arrest, report written, transport to hospital), and event location information (as an address or latitude/longitude coordinates, blurred to the nearest block/street). RMS event types were often text-entry variables specific to the agency or used Uniform Crime Reporting coding to name event types. Because of these differences between agency RMS systems, the event types were hand-coded into the standardization schema in Table 3-1 to allow for cross-jurisdiction comparisons. Standardization also helped data cleaning, as some jurisdictions use different types of phrases to indicate the same kind of event. For example, one jurisdiction had over 10 different ways to indicate or describe a burglary. Standardization grouped these into a single "burglary" category and did the same for all categories in all jurisdictions, which would eventually be coded to "property crimes" as a broad category for analysis.

RMS events differ from CFS events in that RMS events are confirmed events that emergency personnel responded to, whereas CFS include not only emergency events, but also calls about non-emergency events, non-police matters, records of police-logged activity (such as out of patrol car, special detail, special patrols, etc.), and events that are unconfirmed or incorrect. For example, if someone called 911 about a person behaving aggressively in public, it could be entered into the CAD system as “suspicious/aggressive person,” because that is what the caller identified as the issue. If, upon arrival, officers discovered this person was in crisis and arranged appropriate medical care, the event would be recorded in the RMS as a mental health emergency. This may not necessarily be how all CAD/RMS systems work, as they can be customized to each department’s needs, but this example identifies a primary difference in CAD and RMS datasets. Some departments provide identification numbers in the CAD and RMS data that allow for event matching (i.e., someone would be able to see the above “aggressive persons” call in the CFS data being coded as a mental health event in the RMS data). However, many jurisdictions do not provide the identification numbers necessary to connect their CAD and RMS events. As such, analyses are often unable to account for specific incident reporting. Additionally, many CAD events, such as police administrative activity recorded in the CAD, do not connect to any RMS events. All RMS incidents, when possible, were coded to primary and secondary incident types to be used in the production of type-specific crime rates for each Census tract.

CFS Measures

Each site’s CAD/CFS data contained rows of individual calls to 911 or events recorded in the CAD system. CAD data contain all recorded police activity in a jurisdiction, including all incoming 911 calls or other community-initiated police contact, police-initiated events, police administrative actions or events, traffic incidents, medical incidents, and fire incidents. Each CAD/CFS dataset contained rows of community-initiated police contacts (911 calls, text to 911, flagging down an officer), as well as police-initiated activities (police administrative activity, police traffic stops, police foot stops, etc.), and fire- or medical-related 911 calls. The columns in the CAD/CFS data contained information about each event, including an event number, the date and time of the call/event, the type of event, the disposition (not all jurisdictions record this in their CAD/CFS), and location information in the form of an address or latitude/longitude (blurred to the nearest block/street level or tract level). As described above, some jurisdictions include event numbers in their CAD system that allow for linking to RMS data, but many do not provide this information. Other sites provide the date and time of the call, as well as the date and time of officers’ dispatch and arrival. Given that date and time were used to allocate to the year for these analyses, information on the specific timing of each incident, while available, was not used in the final analytical dataset. Jurisdictions also record officer activity in the CAD, and CAD events log when officers leave their vehicles, are on or off patrol, and are engaging in specialized specific activities, hot spots initiatives, or any other activity a jurisdiction chooses to record.

Event naming conventions also vary depending on the jurisdiction or the CAD system. Some jurisdictions in this project use NIBRS or UCR codes systems and others use custom text-entry or code schemas specific to their jurisdictions. To allow for cross-jurisdiction comparison, the CAD call categories were standardized by hand-coding them into the schema detailed in Table 3-1. Hand-coding was necessary, as no two jurisdictions share all of the same CAD event nature types, and spelling or phrasing for the same kind of events can differ even within the same jurisdiction (e.g., “armed robbery” vs. “robbery w/weapon” vs. “rob w/weap:unkn” vs. “code 1293”). This is common when dealing with text-based records but requires standardization to allow for cross-jurisdiction analyses. Currently, there is no generally accepted or widely used standardization schema for CAD call types. The schema this project employs is based on previous research involving CAD standardization or call type grouping.

Importantly, the CAD data differ from the RMS in that CAD is where all police activity is recorded whereas the RMS data only contain events that police or other emergency personnel respond to and clear. The CAD contains all calls that come into the police from community members, including calls that are not police matters, calls that end up referred elsewhere, calls where the caller is inaccurate about what happened, and all the other situations or events that people choose to call the police about. The CAD also contains information that police enter and record about their activity: investigative activities, patrol logs, and officer-initiated stops are all recorded in the CAD. Thus, the CAD records are often larger data files than the RMS datasets and provide a larger picture of both resident-initiated and police-initiated. In this analysis, and as described in subsequent sections, we employ the CAD data as a measure of crime reporting by crime type.

Geo-allocation of Incidents to Census Tracts

Location data for each event in CFS and RMS data were provided as either an address (in same cases blurred to the nearest 100-block level) or as specific latitude/longitude coordinates. The analyses for the current study were conducted at the Census tract level and required counts of specific call and incident types at that level. The first step in this process involved using jurisdiction shapefiles that outline the police jurisdictional boundaries of each department and overlaying that onto a Census tract map of the jurisdiction. The area of each Census tract was calculated; using the R language for statistical computing (R Core Team, 2023), the area of each tract within the jurisdiction boundaries was calculated. From there, the percentage of each tract within the jurisdictional boundaries was calculated, and the population values for each tract could then be allocated. When necessary, the WorldPop population density grid (2018) for each Census tract was used to assess the proportion of the tract-level population that lived within the jurisdiction using population density measures (vs. area-based measures) to allocate population totals.

Handling Missing Data

Though all jurisdictions provided location data of some kind for their CFS and RMS events, some data were not accurate or could not be geocoded to a Census tract, and other events did not have any provided location information. Missingness analyses were conducted for all jurisdictions in the CFS and RMS datasets. Missingness was tabulated by event at the primary and secondary category level for each jurisdiction, as well as by month/year. Jurisdictions with large amounts of missing or inaccurate location data were contacted and the authors worked with them to obtain new data; in cases where jurisdictions did not respond to attempts at contact, some rows of data were excluded from analyses due to location missingness. Table 3-2 shows the percentage of each primary event category that was missing location data from each jurisdiction's CFS data. Table 3-3 displays the same for each jurisdiction's RMS data. Notably, Sites 4 and 5 demonstrate a high level of missingness for geolocation data for CFS measures, with calls to report violent and property crimes showing missingness for location data between 30-43%. Further, Sites 4 and 10 show high levels of missingness for RMS data with violent and property crime location missingness ranging from 26-54%.

Table 3-2. Percentage of CFS Incidents Missing Location Data by Primary Event Category and Jurisdiction

Primary Category	Site 1 (%)	Site 2 (%)	Site 3 (%)	Site 4 (%)	Site 5 (%)	Site 6 (%)	Site 7 (%)	Site 8 (%)	Site 9 (%)	Site 10 (%)
Administrative	6.5	14.2	0.1	99.1	44.5	7.5	11.1	4.8	10.6	0.6
Community	0.1	0.5	0.02	87.1	41.8	0.2	2.3	3.6	3.5	2.5
Investigative	1.4	2.6	0.02	97	57.5	2.8	2.7	3.6	13.5	0.2
Medical/mental	0.01	0.7	0.0002	87.9	41	0.4	1.7	2.5	1.8	0.2
Other miscellaneous crimes	0.2	1.2	0.006	60.6	41.4	5.9	2.1	2.4	7.9	0.1
Property crimes	0.1	0.6	0.006	29.8	43.2	3.2	2.4	1.8	4.5	0.1
Traffic	0.4	4.1	0.01	99.9	38.2	2.7	2.7	6.2	0.9	1.3
Violent crimes	0.04	0.5	0	30	43.4	0.8	2.8	4.3	2.1	0.03

Table 3-3. Percentage of RMS Incidents Missing Location Data by Primary Event Category and Jurisdiction

Primary Category	Site 1 (%)	Site 2 (%)	Site 3 (%)	Site 4 (%)	Site 5 (%)	Site 6 (%)	Site 7 (%)	Site 8 (%)	Site 9 (%)	Site 10 (%)
Administrative	0	2.7	0.2	99.1	0	3.5	0	0	0	26.5
Community	0	2.1	0.1	87.1	0.1	1.2	3.1	0	1.4	32.3
Investigative	0	2	0.01	97	0	1.7	0	0	0	30
Medical/mental	0	1	0	87.9	0	0.6	0	0	0	22.9
Other miscellaneous crimes	0	1.9	0	60.6	0.4	3.2	3.8	0	1.9	27.3
Property crimes	0	2.5	0	29.8	0.1	2.5	1.9	0	1.3	54.4
Traffic	0	2.1	0.1	99.9	0.3	2.7	0.9	0	4.9	28.2

Violent crimes	0	0.8	0	30	0.2	0.7	0.9	0	2.4	26.3

Jurisdiction-Level Descriptive Detail

Each jurisdiction provided several years of CFS and RMS data. Table 3-4 shows a list of jurisdictions, their approximate populations, years of data they provided, region of the U.S. the jurisdiction is in, total numbers of CFS and RMS events provided, and their average number of CFS/RMS events per year. The 10 participating jurisdictions were from various regions of the United States and ranged in size from smaller metro areas to larger more urban and population-dense places. The populations of these jurisdictions ranged from about 75,000 people to one million+.[1] On average, police in these jurisdictions are responding to anywhere from 86,000 to over one million CFS per year. Most jurisdictions provide at least 4 years of data, with a few providing 5 years of CFS and RMS data. As described above, the average number of CFS calls is usually substantially larger than the RMS data as CFS incidents include all police activity or calls for emergency service whereas the RMS data include details on crime incidents. Table 3-4 summarizes the descriptives for the CFS and RMS data for the 10 participating jurisdictions.

Table 3-4. Study Jurisdiction Sites Demographic and Descriptive Information

Site	Population (appx.)	Years of Data	Region	Total CFS Incidents	Average CFS Events per Year	Total RMS Incidents	Average RMS Events per Year
1	~300,000	2017–2020	Southeast	1,139,691	284,922.8	122,880	30,720
2	~1,000,000+	2017–2021	South	5,316,966	1,063,393	1,869,997	373,999.4
3	~75,000	2017–2020	Midwest	1,072,123	268,030.8	74,632	18,658
4	~300,000	2017–2020	Midwest	456,457	114,116.8	456,457	114,116.8
5	~1,000,000+	2017–2021	West	5,404,752	1,080,950	1,084,987	216,997.4
6	~450,000	2017–2020	Southeast	1,328,826	332,206.5	315,284	78,821
7	~190,000	2017–2021	Northeast	607,270	121,454	62,891	12,578.2
8	~75,000	2017–2020	Southeast	346,351	86,587.8	78,109	19,527.3
9	~750,000	2017–2020	Northwest	2,838,847	709,711.8	95,360	23,840
10	~540,000	2018–2021	Southwest	1,252,428	313,107	409,022	102,255.5

[1] Actual population values upward of one million are listed as one million+ to avoid identifying the participating jurisdiction.

4. Analyses of Immigration and Crime

This chapter describes the analytical approach to analyzing the relationship between unauthorized immigration, crime rates, and crime reporting at the Census tract level. In addition, we include a supplementary analysis in one jurisdiction that examines the relationship between unauthorized immigration, crime rates, and the U visa program, which provides a pathway for immigrant authorization for immigrant victims of qualifying crimes. This research builds on existing studies exploring the relationship between immigration and crime by applying the previously described method for estimating the unauthorized immigrant population at the Census tract level to gauge the association with different charge-specific crime rates and crime reporting. This analysis draws on crime and crime reporting data from 2019 for the 10 jurisdictions described in the previous chapter. As discussed, these jurisdictions provide a unique picture of the metro areas in the United States given that they vary by geographic region, size, and immigrant destination type. As such, this analysis provides a comprehensive picture of the association between the estimated unauthorized immigrant population, crime, and crime reporting at the Census tract level, while also controlling for known correlates of crime.

The specific goals for this study include the following research questions (RQs):

- What is the association between unauthorized immigration and crime rates?
- Does this relationship vary by crime type or immigrant destination types?

Data

As described in previous chapters, this analysis uses Census tract-level measures for crime, crime reporting, and the percentage of the tract population that is estimated to be comprised of unauthorized immigrants. We obtained datasets for crime rates, crime reporting, unauthorized immigrant populations, and known correlates of crime. These measures include several known correlates of crime using the ACS (U.S. Census Bureau, 2022a) and other relevant measures associated with different crime rates (Land et al., 1990). Table 4-1 provides the specification of the dependent variable, independent variable, and all control measures with the hypothesized direction of the association between the measure and crime rates. As seen in Table 4-1, the dependent variable is the tract-specific crime rate for drug, property, and violent crime. Drug crime rates consist of all drug offenses in the RMS divided by the tract population, whereas property crime rates (arson, burglary, auto theft, theft/larceny) and violent crime rates (assault, homicide, rape, robbery) are a subset of all property and violent offenses recorded in each jurisdiction’s RMS.

Table 4-1. Description of Variables Included in Regression Analysis

Variable	Definition	Data Source	Hypothesized Correlation
Crime rate (dependent variable) ▪ **Drug** ▪ **Property** ▪ **Violent**	Crimes per 100k population per year ▪ Drug: All reported drug incidents in the RMS ▪ Property: Arson, burglary, auto theft, theft/larceny events ▪ Violent: Assault, homicide, rape, robbery	RMS	N/A
Unauthorized immigrants (independent variable)	Estimated percentage of the tract population who are unauthorized immigrants	Synthpop Estimation, ACS- and SIPP-based	None
Authorized immigrants	Percentage of the population who are foreign-born minus estimated percentage who are unauthorized immigrants	ACS	Negative
Home ownership	Percentage of housing that is owner-occupied	ACS	Negative
Median income	Median income	ACS	Negative
Vacant housing	Percentage of houses or housing considered vacant	ACS	Positive
Below poverty	Percentage of people in tract living below federal poverty line	ACS	Positive
Divorced men	Percentage of men who are divorced	ACS	Positive
Unemployed rate	Percentage of adults unemployed	ACS	Positive
Aged 15–29	Percentage of tract population between 15–29 years of age	ACS	Positive
Population density	People per square mile	ACS	Positive
Reporting index—Drug crimes	See Methods section	CFS, RMS	Negative
Reporting index—Property crimes	See Methods section	CFS, RMS	Negative
Reporting index—Violent crimes	See Methods section	CFS, RMS	Negative
Reporting index—Drug crimes	See Methods section	CFS, RMS	Negative

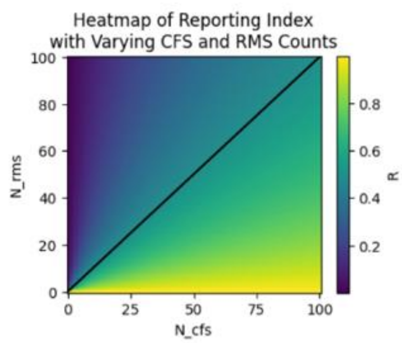
As described in previous chapters, this study operationalized the unauthorized immigrant population by including the estimated percentage of the tract population predicted to comprise unauthorized immigrants. Similarly, the foreign-born population percentage was used to calculate the expected percentage of the population comprising authorized immigrants.

Drawing on prior research exploring the structural correlates of crime (Land et al., 1990), we included a number of controls available in the ACS. Specifically, we accounted for structural characteristics (e.g., home ownership as a percentage of homes, vacant housing), economic characteristics (median income in dollars, the percentage of the tract population living below the poverty line, the tract-level unemployment rate), and demographic characteristics (percentage of divorced men in the tract, percentage of the tract population 15–29 years of age), and the tract-level population density.

Crime Reporting Index

In addition to accounting for correlates for crime, this study used the CFS data to include a reporting bias index that indicates over- or underreporting of crimes in a specific Census tract, for a specific crimes category. This study used the CFS records for specific crime types and RMS records of the same crime incident types within a Census tract. The reporting index ranged from 0 to 1, with an index close to 1 suggesting significant overreporting (i.e., more CFS incidents of a specific type compared to the same RMS incident type), whereas a reporting index close to 0 suggests underreporting, and 0.5 indicates a like for like level of calls and incidents by crime type. This is operationalized as the number of CFS incidents, N_{cfs}, and the number of RMS incidents, N_{rms}, specified below:

$$R = \frac{\left(\frac{N_{cfs}}{N_{rms}}\right)}{\left(\frac{N_{cfs}}{N_{rms}}\right) + 1}$$



Analytical Approach

To test the relationship between per capita crime rates and unauthorized immigration population percentage, we employed an ordinary least squares (OLS) linear regression model, including several known social determinates of crime (SoDC) as control variables in the model (Buonanno, 2003) (see Table 4-1). For each jurisdiction and crime type, we construct a dataset of crime rates, SoDC, and unauthorized and authorized immigrant percentages in each Census tract in the year 2019 and fit the OLS model to this dataset using the python *statsmodels* package. As part of these analyses, we tested feature reduction algorithms that identify variables that can be omitted from the model. These algorithms (here the L1 and L2 regularization methods) penalize measures that are not contributing to the explanation of variance in the model, but this analysis found no changes in the significant results, so regularization was not applied in the final models.

Model Metrics

This study reports the regression coefficients, p-values, and adjusted R^2 values of the regression models for each jurisdiction and jurisdiction group across the three crime categories, drug crimes, property crimes, and violent crimes. P-values are computed by the bootstrap method, to ensure robustness against heteroscedasticity and multicollinearity. We also report the Variance Inflation Factor (VIF) for the primary independent variable (crime rates) in each model.

Results

This section begins by describing the analytical cohort, focusing on the dependent variables and the primary independent variable of the percentage of the tract population estimated to comprise unauthorized immigrants. It then turns to a discussion of model fit and model results.

Descriptive Analyses

Table 4-2 introduces the descriptive statistics for the cohort of cities, presenting the values for the quantiles of each measure (25%, median, 75%). These measures show the quantile for the tract values within that jurisdiction, as opposed to the overall sample of Census tracts. For example, the median percentage of tract population estimated to comprise unauthorized immigrants in emerging jurisdiction 1 is 4%, representing the median value of tracts in this jurisdiction. The values presented in other columns follow this logic, referencing tracts only within that jurisdiction.

As seen in Table 4-2, there was a wide range of estimated tract values by immigrant destination, with median values in emerging jurisdictions ranging from 2 to 4%, compared to 1 to 2% for low foreign-born population cities and 2 to 8% for traditional immigrant destination jurisdictions. Understandably, the percentage of each tract estimated to comprise authorized immigrants was much higher, with overall median values ranging from 3 to 46%. Median drug crime rate values ranged from 130 to 500+ per 100,000 people per year in low foreign-born jurisdictions, compared to 40 to 97 and 53 to 88 per 100k in emerging and traditional jurisdictions, respectively. Property crime rates presented a much broader distribution, with median values at the tract level ranging from 385 to 1,540 per 100,000 residents. Median values for property crime rates in emerging immigrant destinations ranged from 43 to 663, showing a similar large spread, while traditional immigrant destination jurisdictions had median values of property crime rates ranging from nearly 380 to 760 per 100,000 residents. Violent RMS records were understandably lower, with low foreign-born tracts having median rates ranging from 200 to 540 incidents for 100,000 residents, compared to 50 to 213 for emerging tracts and 218 to 340 for traditional tracts. Despite this large range, tracts within immigrant designation types appeared to be fairly similar, as there were only slight outliers for median values in any specific jurisdiction when compared with its peers in that immigrant destination type.

Table 4-2. Descriptive Statistics of Independent and Dependent Variables in Each Jurisdiction

Jurisdiction	Tract Range		Unauthorized Immigrant Pct.	Authorized Immigrant Pct.	Drug Crimes per 100k Annually	Property Crimes per 100k Annually	Violent Crimes per 100k Annually
Emerging 1	30–50	Median	4.06	9.58	97.7	663	137
		25% quantile	2.06	6.47	54.8	371	64.7
		75% quantile	5.73	11.4	211	1,030	315
Emerging 2	30–50	Median	3.58	25.4	41.1	511	214
		25% quantile	2.55	16.8	29.6	348	150
		75% quantile	4.96	30.3	56.8	659	302
Emerging 3	70–100	Median	2.02	12.4	40.3	43.2	52.3
		25% quantile	1.4	9.3	24.5	33.2	39.7
		75% quantile	3.37	18.7	78.3	62.7	70.6
Low 1	50–70	Median	1.18	5.38	133	385	201
		25% quantile	0.414	3.24	73.3	278	95.5
		75% quantile	2.13	10	221	652	276
Low 2	<15	Median	1.47	3.4	514	921	404
		25% quantile	0.736	2.67	389	710	233
		75% quantile	1.77	5.1	594	995	500
Low 3	<15	Median	1.63	5.92	224	1,540	543
		25% quantile	0.795	4.47	125	558	245
		75% quantile	2.67	8.93	281	2,610	1510
Traditional 1	200–300	Median	7.1	17.8	60.3	741	261
		25% quantile	3.62	10.9	40.1	486	147
		75% quantile	11.3	23.9	92.5	1,100	418
Traditional 2	500+	Median	7.33	28.8	53.3	560	340
		25% quantile	3.46	23.7	40.2	394	197
		75% quantile	11.8	34.4	72.2	782	572

(continued)

Table 4-2. Descriptive Statistics of Independent and Dependent Variables in Each Jurisdiction (continued)

Jurisdiction	Tract Range		Unauthorized Immigrant Pct.	Authorized Immigrant Pct.	Drug Crimes per 100k Annually	Property Crimes per 100k Annually	Violent Crimes per 100k Annually
Traditional 3	50–70	Median	7.93	46.5	52.6	759	239
		25% quantile	4.86	35.2	38	514	140
		75% quantile	11.3	57.4	76.2	1,280	469
Traditional 4	100–200	Median	2.17	9.93	87.9	379	218
		25% quantile	1.24	7.63	53.5	215	145
		75% quantile	4.42	15.5	147	553	309

Analytical Results

This analysis now turns to a discussion of the linear regression results, with details on the overall relationship before moving to specific details of coefficients and model fit statistics.

Linear Relationships

Across all jurisdictions, jurisdiction groups, and crime types, the relationship between unauthorized immigration and crime rates were largely null (i.e., insignificant or having a p-value of > 0.05) (see Table 4-3). Traditional immigration jurisdiction 1 had a slightly significant ($p < 0.008$) and negative association of the unauthorized immigrant population with violent crimes. Additionally, for drug crimes in traditional immigration jurisdiction 2, there was a small positive association with a p-value of 0.005; however, after accounting for multiple tests over many jurisdictions (e.g., using a Bonferroni correction - Sedgwick, 2012) and the poor model fits (low adjusted R^2), revealed these findings to also be largely insignificant. The Bonferroni correction would suggest a significance threshold of p=0.05/36=0.0014, which neither of these results met. The near-significant result on traditional 2 was given by a model with very poor model fit (adjusted R^2=0.058).

Table 4-3. Model Results for the Association of Immigration and Crime at the Census Tract Level

Jurisdiction	Crime Category	p-value	Regression Coeff.	Adj. R^2	VIF of IV
Low immigration jurisdictions	Drug crimes	0.081	27.103	0.415	6.632
Low immigration jurisdictions	Property crimes	0.455	5.629	0.359	6.652
Low immigration jurisdictions	Violent crimes	0.403	16.611	0.29	6.632
Emerging immigration jurisdictions	Drug crimes	0.162	9.376	0.355	2.019
Emerging immigration jurisdictions	Property crimes	0.116	24.43	0.583	2.017

(continued)

Table 4-3. Model Results for the Association of Immigration and Crime at the Census Tract Level (continued)

Jurisdiction	Crime Category	p-value	Regression Coeff.	Adj. R^2	VIF of IV
Emerging immigration jurisdictions	Violent crimes	0.382	2.945	0.515	2.014
Traditional immigration jurisdictions	Drug crimes	0.068	-0.869	0.125	3.086
Traditional immigration jurisdictions	Property crimes	0.491	-0.258	0.124	3.072
Traditional immigration jurisdictions	Violent crimes	0.087	-4.076	0.122	2.896
Low immigration jurisdiction 1	Drug crimes	0.211	12.294	0.524	7.275
Low immigration jurisdiction 1	Property crimes	0.372	15.318	-0.449	7.212
Low immigration jurisdiction 1	Violent crimes	0.41	3.105	0.543	7.333
Emerging immigration jurisdiction 1	Drug crimes	0.582	1.247	0.411	3.373
Emerging immigration jurisdiction 1	Property crimes	0.26	33.706	0.515	3.609
Emerging immigration jurisdiction 1	Violent crimes	0.495	3.204	0.719	3.59
Emerging immigration jurisdiction 2	Drug crimes	0.291	-3.991	0.059	4.773
Emerging immigration jurisdiction 2	Property crimes	0.122	93.514	0.181	4.671
Emerging immigration jurisdiction 2	Violent crimes	0.458	11.978	0.202	4.29
Emerging immigration jurisdiction 3	Drug crimes	0.188	6.42	0.369	3.489
Emerging immigration jurisdiction 3	Property crimes	0.121	9.795	0.443	3.518
Emerging immigration jurisdiction 3	Violent crimes	0.432	1.848	0.478	3.49
Traditional immigration jurisdiction 1	Drug crimes	0.018	-3.254	0.24	4.172
Traditional immigration jurisdiction 1	Property crimes	0.049	-16.35	0.232	4.172
Traditional immigration jurisdiction 1	Violent crimes	0.008	-7.94	0.507	4.157
Traditional immigration jurisdiction 2	Drug crimes	**0.005**	1.314	**0.058**	3.792
Traditional immigration jurisdiction 2	Property crimes	0.417	0.431	0.123	3.817
Traditional immigration jurisdiction 2	Violent crimes	0.121	6.269	0.139	3.791
Traditional immigration jurisdiction 3	Drug crimes	0.128	6.488	0.071	6.07
Traditional immigration jurisdiction 3	Property crimes	0.278	-62.587	0.076	6.208
Traditional immigration jurisdiction 3	Violent crimes	0.368	-6.894	0.339	6.11
Traditional immigration jurisdiction 4	Drug crimes	0.266	12.91	0.245	4.391
Traditional immigration jurisdiction 4	Property crimes	0.401	-2.505	0.359	4.338
Traditional immigration jurisdiction 4	Violent crimes	0.33	11.002	0.229	4.32

Note: Regression coefficients and p-values of unauthorized immigration fraction, and model fit measures for each fit including adjusted R^2 and the VIF of unauthorized immigration fraction with respect to the control variables. Results for jurisdictions Low 2 and Low 3 are not reported, as they had too few data points to be fit on independently.

In addition to providing a test of the association between the percent of unauthorized immigrants within a Census tract and crime rates, this analysis includes a control for authorized immigrants. As described above, this is operationalized as the number of foreign-born people in a Census tract minus the estimated number of unauthorized immigrants in that tract. Table 4-4 shows the regression coefficients for full model for the percent of authorized immigrants in a Census tract, accounting for the variance explained by the unauthorized immigrant percentage in a tract. In these analyses, the direction of the regression coefficient is almost entirely negative, despite a lack of significance in many models. However, combined models of emerging immigrant destinations show that the percentage of the authorized immigrant population within a Census tract is associated with lower drug, property, and violent crimes in the tract. Further, the combined models for traditional immigrant destinations show that the percentage of authorized immigrants in a tract is associated with lower drug crime rates.

Table 4-4. Model Results for the Association of Authorized Immigration and Crime at the Census Tract Level

Jurisdiction	Crime Category	p-value	Regression Coeff.	Adj. R^2	VIF of IV
Low Immigration Jurisdictions	Drug Crimes	0.088	-8.173	0.415	5.593
Low Immigration Jurisdictions	Property Crimes	0.433	6.22	0.359	5.566
Low Immigration Jurisdictions	Violent Crimes	0.179	-13.766	0.29	5.572
Emerging Immigration Jurisdictions	Drug Crimes	<0.001	-5.46	0.355	2.527
Emerging Immigration Jurisdictions	Property Crimes	0.009	-9.806	0.583	2.515
Emerging Immigration Jurisdictions	Violent Crimes	0.003	-4.968	0.515	2.521
Traditional Immigration Jurisdictions	Drug Crimes	<0.001	-1.539	0.125	2.122
Traditional Immigration Jurisdictions	Property Crimes	0.24	1.874	0.124	1.961
Traditional Immigration Jurisdictions	Violent Crimes	0.076	-3.235	0.122	1.943
Low Immigration Jurisdiction 1	Drug Crimes	0.144	-5.925	0.524	5.485
Low Immigration Jurisdiction 1	Property Crimes	0.316	-6.365	-60.506	5.368
Low Immigration Jurisdiction 1	Violent Crimes	0.196	-7.334	0.543	5.445
Emerging Immigration Jurisdiction 1	Drug Crimes	0.284	-11.819	0.411	3.114
Emerging Immigration Jurisdiction 1	Property Crimes	0.235	-36.808	0.515	3.177
Emerging Immigration Jurisdiction 1	Violent Crimes	0.111	-18.682	0.719	3.213
Emerging Immigration Jurisdiction 2	Drug Crimes	0.099	2.342	0.059	6.598
Emerging Immigration Jurisdiction 2	Property Crimes	0.339	-11.72	0.181	6.77
Emerging Immigration Jurisdiction 2	Violent Crimes	0.568	-0.142	0.202	6.677
Emerging Immigration Jurisdiction 3	Drug Crimes	0.23	-0.898	0.369	4.235
Emerging Immigration Jurisdiction 3	Property Crimes	0.187	-1.204	0.443	4.246
Emerging Immigration Jurisdiction 3	Violent Crimes	0.045	-2.815	0.478	4.269

Table 4-4. Model Results for the Association of Authorized Immigration and Crime at the Census Tract Level (continued)

Jurisdiction	Crime Category	p-value	Regression Coeff.	Adj. R^2	VIF of IV
Traditional Immigration Jurisdiction 1	Drug Crimes	0.108	-0.921	0.24	2.339
Traditional Immigration Jurisdiction 1	Property Crimes	0.425	0.998	0.232	2.375
Traditional Immigration Jurisdiction 1	Violent Crimes	0.037	-3.479	0.507	2.33
Traditional Immigration Jurisdiction 2	Drug Crimes	0.016	-0.651	0.058	1.526
Traditional Immigration Jurisdiction 2	Property Crimes	0.095	-6.668	0.123	1.543
Traditional Immigration Jurisdiction 2	Violent Crimes	0.001	-9.999	0.139	1.526
Traditional Immigration Jurisdiction 3	Drug Crimes	0.049	-5.173	0.071	5.144
Traditional Immigration Jurisdiction 3	Property Crimes	0.369	9.349	0.076	5.093
Traditional Immigration Jurisdiction 3	Violent Crimes	0.29	-1.649	0.339	5.104
Traditional Immigration Jurisdiction 4	Drug Crimes	0.44	0.693	0.245	4.032
Traditional Immigration Jurisdiction 4	Property Crimes	0.236	5.264	0.359	4.056
Traditional Immigration Jurisdiction 4	Violent Crimes	0.418	0.859	0.229	3.981

Supplementary Analyses of Unauthorized Immigration, Crime Rates, and U Visa Certification

Beyond assessing the general relationship of unauthorized immigration and crime, this study sought to expand on the research on unauthorized immigration and crime by exploring the usage of the U visa program in one select jurisdiction. The U visa is a form of nonimmigrant visa available for immigrants who (1) become the victims of one of 28 qualifying crimes while in the United States and (2) cooperate with a law enforcement or other government agency investigation of that crime. Congress established the U visa program as part of the 2000 Violence Against Women Act in recognition that unauthorized immigrants who experience serious crimes may be hesitant to aid law enforcement and prosecutors to investigate and potentially prosecute their case due to fears about their lack of immigration status. The statute allows for up to an annual cap of 10,000 visas given out per year; applicants who are deemed eligible for the visa are placed on a waitlist if the cap is met, which can take approximately 10 years to receive in hand.

A key question of the U visa program is whether it fulfills its goals of encouraging immigrants to report crime and participate in criminal investigations. As a result, one of the primary requirements of the U visa application is the submission of law enforcement certification or a signed Form I-918 Supplement B from an authorized certifying agency attesting that the victim was indeed a victim of a qualifying crime and had been helpful to an investigation. Without this Supplement B, an immigrant victim cannot apply for the U visa. However, certifying agencies are also not mandated to certify U visa applications. Their participation in the program is ultimately discretionary and influenced by a range of factors, including internal policy around certification, knowledge of the program, agency capacity, and agency desire and willingness to do so. Certifiers' judgements of whether a victim experienced a qualifying crime and was cooperative and helpful to the investigation are also discretionary, subjective decisions.

This supplementary analysis examined U visa certification data from one of the law enforcement agencies participating in this study, Southeastern Police Department (SPD). The analysis produces one of the first estimates of a U visa "uptake" rate to understand the extent to which potentially qualifying immigrant victimizations ultimately make it to the agency in the form of a certification request.

U Visa Data Description

This analysis uses three primary forms of data to produce a crosswalk of both certification data charges and RMS charges to U visa qualifying offenses as determined by the U.S. Department of Homeland Security (Table 4-5). The research team then uses this crosswalk in the analyses. First, we use 1,066 U visa certification requests submitted to SPD between 2019 to 2023. Each U visa request contains dates for the incident, certification request date, certification decision date, qualifying crime if applicable, and denial reason if applicable. Background conversations with SPD regarding these requested data provided additional context for how the certifying authority maps their internal RMS crime categories to the U visa program's list of qualifying crimes and how they deem a potential applicant sufficiently helpful to the investigation. Next, the study employs RMS data from SPD in 2019, which tags events by date, charge type, initiation, disposition, and location. Lastly, the analysis uses Census tract estimates of the unauthorized immigrant population for SPD tracts as described in Chapter 2.

Table 4-5. Proposed Crosswalk of U Visa–Eligible Offenses to Observed RMS Incidents

U Visa Certification Charge Description [1]	Census Tract Crimes Categories from RMS Data [2]	Corresponding U Visa–Eligible Crime [3]
Aggravated assault, domestic violence, child abuse	Aggravated assault, domestic violence, child abuse	Aggravated Assault, Domestic Violence, Child Abuse
Blackmail, extortion	Blackmail, extortion	Blackmail
Homicide, murder	Homicide/murder/non-negligent manslaughter, justifiable homicide	Manslaughter/Murder
Kidnapping	Kidnapping/abduction	Abduction
Prostitution, trafficking, exploitation	Prostitution, trafficking, exploitation	Prostitution, Trafficking, Exploitation
Rape, forcible rape	Rape, statutory rape	Rape
Sexual battery, fondling	Incest, fondling, sodomy, sexual assault with object	Non-rape sex related (sexual assault, sexual exploitation)

[1] SPD U visa certification data.
[2] SPD RMS data.
[3] USCIS U visa–eligible crimes.

Sub-analysis Analytical Approach

Using SPD's certification data, RMS data, and estimates of the unauthorized population, all from the same Census tracts, we estimate the expected count of U visa–eligible victimizations in each Census tract and approximate gaps in certification rates by charge types. To do this, the study leverages expected counts of relevant crimes within each Census tract, by scaling the tract-wide crime rate for each crime type by that tract's estimated unauthorized population. We then estimate the number of these charges that would likely result in certification by using the observed U visa approval rate for that charge category in that Census tract based off the observed SPS U visa certification request data. We then use these counts to produce a U visa coverage rate, shown below in Equation 1. Here, to produce the coverage gap (CG) for charge type *i*, we subtract the observed count of certified requests for charge *i* (*Oi*) from the expected count of certified requests for charge *i* (*Ei*). This figure is divided by the expected count of certified requests for charge *i* (*Ei*) and multiplied by 100 to obtain a percentage.

$$CG_i = \left(\frac{E_i - O_i}{E_i}\right) * 100$$

Sub-analysis Results

Table 4-6 shows the result of estimating U visa–eligible crimes based on Census-tract-level data. The first column displays the forecasted number of eligible crimes across all Census tracts by scaling tract-specific crime rates for each crime category to the tract's unauthorized population. We then present an empirical forecast of eligibility that uses the percentage of U visas for that charge type approved in each tract across the study period to approximate the number of crimes occurring that could have been U visa–eligible. By comparing these empirical crime estimates to the actual number of U visa requests generated in 2019 in SPD, we estimate coverage gaps between the projected number of U visa–eligible crimes across tracts and the observed number of requests in the jurisdiction.

Across crime categories in 2019, potential coverage gaps range from 88% to 100%, indicating significant differences in the estimated number of U visa–eligible victimizations across tracts and the actual number of certification requests. As an example, using the expected count of aggravated assaults/domestic violence/child abuse incidents where the victim is estimated to be an unauthorized immigrant yields 1,067 potentially qualifying incidents in 2019, compared with 36 observed eligible offenses in 2019. Using the expected eligibility threshold based on charge-specific approval rates results in a coverage gap of 95%. Similarly, the example of estimated kidnapping and abduction incidents for unauthorized immigrant victims shows that using the charge-specific eligibility threshold produces a coverage gap of 95% of potentially qualifying incidents. These results suggest that there could be a potential underutilization of the U visa program if assumptions about victimization by nativity status hold.

Table 4-6. U Visa Coverage Estimates Based on Tract-Wise Crime Data and SPD Certification Data

Crime Category	Expected Crime Count [1]	Expected U Visa–Eligible Crime Count (Charge-specific Approval Rates) [2]	Charge-specific U Visa Approval Rates (2019-2023) [3]	Total Certification Requests, 2019 [4]	Potential Coverage Gap – Observed Approval Rate, 2019 [5]
Aggravated assault/domestic violence/child abuse	**1067**	**686**	**64%**	**36**	**95%**
Blackmail and extortion	**36**	**18**	**50%**	**0**	**100%**
Kidnapping and abduction	**75**	**57**	**76%**	**3**	**95%**
Murder and homicide	**50**	**17**	**34%**	**2**	**88%**
Non-rape sex offenses (e.g., incest, fondling)	**145**	**124**	**86%**	**0**	**100%**
Prostitution/trafficking/exploitation	**3**	**2**	**66%**	**0**	**100%**
Rape	**97**	**78**	**80%**	**7**	**91%**

[1] Derived from RMS offense data by dividing total crimes in each category per tract by the total tract population, multiplying by each tract's unauthorized population, and aggregating across census tracts.
[2] Derived from SPD actual U visa certification rates and multiplying these by column 1.
[3] Derived from U visa certification data by aggregating U visa requests by binned RMS offense and calculating the percent certified out of all requests.
[4] Total number of U visa certification requests in 2019, derived from U visa request data from SPD.
[5] Percentage difference between expected crime count and actual 2019 U visa requests certified, determined using observed charge-specific tract approval rates.

5. Conclusion

This chapter highlights some of the limitations of the research presented here along with consideration of ethical concerns that accompany the research. We then turn to a discussion of the primary findings and the implications for future research.

Limitations and Considerations

Because this research relies on several constituent methodological pieces, we present assumptions and limitations that are integral in understanding the scope, strengths, and weaknesses of each methodological piece.

There are a number of methodological decisions that were necessary for the analyses and require certain assumptions but may have affected their accuracy and generalizability.

SIPP limitations. To develop the estimates of the unauthorized immigrant population in the SIPP using all available SIPP data (waves 2008, 2014, and 2018), these analyses required imputing certain immigrant status questions from the 2008 wave into the 2014 and 2018 waves. This necessitated the assumption that the underlying populations included in all samples are similar and drawn from the same universe and that the unauthorized immigrant population in 2008 is similar to those identified in the 2014 and 2018 waves. In addition, the models do not explicitly make corrections for the presence of immigrants who may have potentially benefited from the Deferred Action for Childhood Arrivals (DACA) measure. Although this limitation can be corrected to some extent by scaling the tract-level estimates to topline figures of the unauthorized immigrant population using the residual method drawn from current official statistics, this issue does highlight the limitations placed on this methodology when survey questions are omitted from wave to wave.

Validation limitations. Although validation efforts were comprehensive based on potential available checks, comparing estimates to external benchmarks does not eliminate the possibility of discrepancies. This is a limitation of any research on the unauthorized immigrant population, as often the local “ground truth” is unobtainable due to ethical concerns and potential bias in survey responses. Instead, this study leveraged available methods—such as aggregating unweighted tract-level estimates for comparison against state-level and national estimates—that use other techniques for obtaining estimates of the unauthorized immigrant population or for comparing with other proxies of the unauthorized immigrant population at the county level.

Limitations in RMS/CFS data. In addition to the constraints of estimating the unauthorized immigrant population, there are inherent constraints with using observed crime and crime reporting data. First, using official statistics for crime records and crime reporting can ignore undercounts of unreported crimes. To alleviate this concern, we focused on crimes most likely to be reported as checks against non-reporting (e.g., homicide for violent crimes, arson and burglary for property charges). In addition, we privileged the use of charge categories with the least amount of missingness in geographic detail to ensure the analyses had comprehensive geographic coverage within each jurisdiction.

Causal inference limitations. It should be noted that this analysis is correlational and does not claim the ability to draw causal inference. This analysis did seek to control for known correlates of crime and aims to produce rigorous estimates of the relationship of unauthorized immigration and crime rates through robustness tests (e.g., Bonferroni correction) and adherence to recommended limits for variance inflation factors and adjusted R^2 values, but the design is not inherently causal. In addition, this analysis, like all macro-level associational analyses, assumes that crimes occur where people reside. Future analyses plan to account for this limitation by assessing the impact of neighboring unauthorized immigrant population values on crime rates.

Ethnical concerns. Producing estimates of the unauthorized immigrant population at a fine geographic resolution could raise privacy concerns and heightened risk that the data could be used with malicious intent. However, the estimates developed for these analyses included several steps that incorporated uncertainty into the development of the final counts. First, imputation is used within the SIPP and we also employ a proxy measure for legal status, although this model is likely useful for identifying characteristics often associated with undocumented status. Second, the process of generating estimates in the synthetic population requires (1) incorporating uncertainty by comparing predicted probabilities for synthetic records to a random draw from the binomial distribution and (2) using multiple draws to produce a distribution of potential counts of unauthorized immigrants within a Census tract and taking the mean value. The mean values do not represent an absolute count but instead provide a possible value drawn from a range.

Limitations of During-COVID and Post-COVID Analyses

The onset of the COVID-19 pandemic brought unique challenges to research on the relationship between unauthorized immigration and crime. Specifically, this study depends on a synthetic population developed from the American Community Survey (ACS) that primarily use the ACS 5-year survey product for broader estimates with lower margins of error but may validate with the ACS 1-year estimates for annual population estimates, although the annual estimates will have larger margins of error due to the decreased sample size. However, in addition to the many challenges presented by the COVID-19 pandemic, Census products were also impacted as the ACS was only able to collect "two-thirds of the responses it usually collects in a survey year."[2] As a result, the one-year 2020 estimates are listed as experimental in nature, and not informative for products that require one-year estimates to make generalizations. In addition, when incorporating the 2020 estimates into the ACS 5-year estimates (necessary to make post-COVID synthetic population estimates), the Census has indicated that the coefficients of variance (standard error divided by the estimate) were outsized for the majority of key measures tested for five-year products that include 2020. As such, although this study did produce synthetic population estimates for 2020 and 2022, the heighted potential for mismeasurement suggested that the synthetic populations would not offer a true representation of the unauthorized immigrant population. Furthermore, limited spatial characteristics for CFS/RMS data collection for post-COVID years prevented the research team from making generalizable

[2] https://www.census.gov/programs-surveys/acs/technical-documentation/user-notes/2022-04.html

conclusions about the post-COVID relationship between the unauthorized immigrant population and crime. These data collection issues highlight the need for further data harmonization and validation efforts based on more recent years in future research.

Discussion

The results of these analyses largely demonstrate that the relationship between unauthorized immigration and crime rates is largely insignificant (i.e., null, neither positive or negative) across various jurisdiction types and crime categories. This finding aligns with the large body of research on immigration and crime indicating that there is no *strong positive correlation* between the presence of immigrants and heightened crime rates, although this study does identify a negative relationship with crime rates and authorized immigration. Although the slight negative association observed in one traditional immigration destination for violent crimes and the small positive association seen in another traditional jurisdiction for drug crimes emphasize the complexity of these relationships, these associations lose significance when accounting for robustness tests and considerations with model fit. This suggests that unauthorized immigration does not substantially influence crime rates in a consistent manner.

These findings underscore the recommendation for future research to adopt more specific and nuanced approaches when examining the association of unauthorized immigration with crime rates. Furthermore, although the methods to develop estimates of the unauthorized immigrant population are novel and leverage accepted methodologies for national estimates, the limitations encountered in this study, such as the use of imputed data and the challenges of validating estimates, highlight the need for improved data collection methods and more comprehensive datasets. Additionally, the exploration of authorized immigration's negative association with crime rates suggests that further investigation into the varying impacts of different immigrant populations on crime could yield important insights. Future studies should also delve into the mechanisms through which reductions in crime rates that correlate with immigration operate.

Last, these results have notable policy implications. Policymakers should be wary of attributing increases in crime rates to unauthorized immigrant populations without robust evidence. Efforts should also be made to address the root causes of crime and to enhance reporting mechanisms to ensure that all crimes are accurately documented and analyzed. These findings also underline the importance of public communication. Misconceptions that link unauthorized immigration and increases in crime can lead to heightened xenophobia and discrimination. Disseminating accurate, research-based information can help mitigate these issues. Future research should build upon these analyses, seeking to bolster methodological weaknesses when present and ensuring that the discourse around unauthorized immigration and crime is grounded in empirical evidence.

References

Adelman, R., Reid, L. W., Markle, G., Weiss, S., & Jaret, C. (2017). Urban crime rates and the changing face of immigration: Evidence across four decades. *Journal of Ethnicity in Criminal Justice*, *15*(1), 52–77. https://doi.org/10.1080/15377938.2016.1261057

Akins, S., Rumbaut, R. G., & Stansfield, R. (2009). Immigration, Economic Disadvantage, and Homicide. *Homicide Studies*, *13*(3), 307–314. https://doi.org/10.1177/1088767909336814

Arbona, C., Olvera, N., Rodriguez, N., Hagan, J., Linares, A., & Wiesner, M. (2010). Acculturative Stress Among Documented and Undocumented Latino Immigrants in the United States. *Hispanic Journal of Behavioral Sciences*, *32*(3), 362–384. https://doi.org/10.1177/0739986310373210

Azari, S. S., Jenkins, V., Hahn, J., & Median, L. (2024). *The Foreign-Born Population in the United States: 2022*.

Bachmeier, J. D., Van Hook, J., & Bean, F. D. (2014). Can We Measure Immigrants' Legal Status? Lessons from Two U.S. Surveys. *International Migration Review*, *48*(2), 538–566. https://doi.org/10.1111/IMRE.12059

Baker, B., & Warren, R. (2024). *Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2018 to January 2022*.

Beckman, R. J., Baggerly, K. A., & McKay, M. D. (1996). Creating synthetic baseline populations. *Transportation Research Part A: Policy and Practice*, *30*(6), 415–429. https://doi.org/10.1016/0965-8564(96)00004-3

Bersani, B. E. (2013). A Game of Catch-Up? The Offending Experience of Second-Generation Immigrants. *Crime & Delinquency*, *60*(1), 60–84. https://doi.org/10.1177/0011128713502406

Bersani, B. E. (2014). An Examination of First and Second Generation Immigrant Offending Trajectories. *Justice Quarterly*, *31*(2), 315–343. https://doi.org/10.1080/07418825.2012.659200

Borjas, G. J. (2017). The Earnings of Undocumented Immigrants. *National Bureau of Economic Research Working Paper Series*. http://www.nber.org/papers/w23236

Bowles, R., Reyes, M. G., & Garoupa, N. (2009). Crime reporting decisions and the costs of crime. *European Journal on Criminal Policy and Research*, *15*(4), 365–377. https://doi.org/10.1007/S10610-009-9109-8/METRICS

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324/METRICS

Buonanno, P. (2003). The socioeconomic determinants of crime: A review of the literature. *Working Paper Dipartimento Di Economia Politica, Università Di Milano Bicocca;63*. https://boa.unimib.it/handle/10281/22981

Butcher, K. F., & Piehl, A. M. (1998). Recent Immigrants: Unexpected Implications for Crime and Incarceration. *ILR Review*, *51*(4), 654–679. https://doi.org/10.1177/001979399805100406

Capps, R., Bachmeier, J. D., & Van Hook, J. (2018a). Estimating the Characteristics of Unauthorized Immigrants Using U.S. Census Data: Combined Sample Multiple Imputation. *The ANNALS of the American Academy of Political and Social Science*, *677*(1), 165–179. https://doi.org/10.1177/0002716218767383

Capps, R., Bachmeier, J. D., & Van Hook, J. (2018b). Estimating the Characteristics of Unauthorized Immigrants Using U.S. Census Data: Combined Sample Multiple Imputation. *Annals of the American Academy of Political and Social Science*, *677*(1), 165–179. https://doi.org/10.1177/0002716218767383/ASSET/IMAGES/LARGE/10.1177_0002716218767383-FIG1.JPEG

Capps, R., Gelatt, J., Ruiz Soto, A., & Van Hook, J. (2020). *Unauthorized Immigrants in the United States: Stable Numbers, Changing Origins*.

Clark, R. L., Glick, J. E., & Bures, R. M. (2009). Immigrant Families Over the Life Course. *Journal of Family Issues*, *30*(6), 852–872. https://doi.org/10.1177/0192513X09332162

Fienberg, S. E., Meyer, M. M., & Wasserman, S. S. (1985). Statistical analysis of multiple sociometric relations. *Journal of the American Statistical Association*, *80*(389), 51–67. https://doi.org/10.1080/01621459.1985.10477129

Greenman, E., & Hall, M. (2013). Legal Status and Educational Transitions for Mexican and Central American Immigrant Youth. *Social Forces*, *91*(4), 1475–1498. https://doi.org/10.1093/SF/SOT040

Gutierrez, C. M., & Kirk, D. S. (2015). Silence Speaks: The Relationship between Immigration and the Underreporting of Crime. *Crime and Delinquency*, *63*(8), 926. https://doi.org/10.1177/0011128715599993

Hagan, J., Levi, R., & Dinovitzer, R. (2008). The Symbolic Violence of the Crime - Immigration Nexus: Migrant Mythologies in the Americas. *Criminology and Public Policy*, *7*. https://heinonline.org/HOL/Page?handle=hein.journals/crpp7&id=97&div=&collection=

Hall, M., & Greenman, E. (2013). Housing and neighborhood quality among undocumented Mexican and Central American immigrants. *Social Science Research*, *42*(6), 1712–1725. https://doi.org/10.1016/J.SSRESEARCH.2013.07.011

Hastie, T., Friedman, J., & Tibshirani, R. (2001). *The Elements of Statistical Learning*. https://doi.org/10.1007/978-0-387-21606-5

Inkpen, C. (2024). Differences in Time to Reported First Arrest by Race, National Origin, and Immigrant Generation: A Test of Assimilation Theories. *Crime & Delinquency*. https://doi.org/10.1177/00111287231225125

Bersani, B., Fine, A., Piquero, A., Steinberg, L., Frick, P., & Cauffman, E. (2018). Investigating the Offending Histories of Undocumented Immigrants. *Migration Letters*, *15*(2), 147–166.

Jácome, E. (2022). The effect of immigration enforcement on crime reporting: Evidence from Dallas. *Journal of Urban Economics*, *128*, 103395. https://doi.org/10.1016/J.JUE.2021.103395

Jennings, W. G., Zgoba, K. M., Piquero, A. R., & Reingle, J. M. (2013). Offending Trajectories Among Native-Born and Foreign-Born Hispanics to Late Middle Age. *Sociological Inquiry*, *83*(4), 622–647. https://doi.org/10.1111/SOIN.12017

Kääriäinen, J., & Sirén, R. (2010). Trust in the police, generalized trust and reporting crime. *European Journal of Criminology*, *8*(1), 65–81. https://doi.org/10.1177/1477370810376562

Kubrin, C. E., Hipp, J. R., & Kim, Y. A. (2018). Different than the Sum of Its Parts: Examining the Unique Impacts of Immigrant Groups on Neighborhood Crime Rates. *Journal of Quantitative Criminology*, *34*(1), 1–36. https://doi.org/10.1007/S10940-016-9320-Y/TABLES/6

Land, K. C., McCall, P. L., & Cohen, L. E. (1990). Structural Covariates of Homicide Rates: Are There Any Invariances Across Time and Social Space? *American Journal of Sociology*, *95*(4), 922–963. https://doi.org/10.1086/229381

Langton, S., Ruiter, S., & Verlaan, T. (2023). Describing the scale and composition of calls for police service: a replication and extension using open data. *Police Practice and Research*, *24*(5), 523–538. https://doi.org/10.1080/15614263.2022.2102494

Lee, M. T., & Martinez, R. (2009). Immigration reduces crime: An emerging scholarly consensus. *Sociology of Crime Law and Deviance*, *13*, 3–16. https://doi.org/10.1108/S1521-6136(2009)0000013004/FULL/XML

Lejins, P. P. (1965). Uniform Crime Reports. *Michigan Law Review*, *64*. https://heinonline.org/HOL/Page?handle=hein.journals/mlr64&id=1031&div=&collection=

Lyons, C. J., Vélez, M. B., & Santoro, W. A. (2013). Neighborhood Immigration, Violence, and City-Level Immigrant Political Opportunities. *American Sociological Review*, *78*(4), 604–632. https://doi.org/10.1177/0003122413491964

Martinez Jr., R., & Lee, M. (2000). On Immigration and Crime. *Criminal Justice*, 485–524.

Martinez, R., Stowell, J. I., & Lee, M. T. (2010). Immigration and Crime in an Era of Transformation: A Longitudinal Analysis of Homicides in San Diego Neighborhoods, 1980–2000*. *Criminology*, *48*(3), 797–829. https://doi.org/10.1111/J.1745-9125.2010.00202.X

Maxfield, M. G. (1999). The National Incident-Based Reporting System: Research and Policy Applications. *Journal of Quantitative Criminology*, *15*(2), 119–149. https://doi.org/10.1023/A:1007518620521/METRICS

Messing, J. T., Becerra, D., Ward-Lasher, A., & Androff, D. K. (2015). Latinas' Perceptions of Law Enforcement. *Affilia*, *30*(3), 328–340. https://doi.org/10.1177/0886109915576520

Miller, H. (2018). On the History of Immigration and Crime. In H. Ventura Miller & A. Peguero (Eds.), *Routledge Handbook on Immigration and Crime* (pp. 5–18). Routledge. https://doi.org/10.4324/9781317211563-2

Ousey, G. C., & Kubrin, C. E. (2009). Exploring the Connection between Immigration and Violent Crime Rates in U.S. Cities, 1980–2000. *Social Problems*, *56*(3), 447–473. https://doi.org/10.1525/SP.2009.56.3.447

Ousey, G. C., & Kubrin, C. E. (2018). Immigration and crime: Assessing a contentious issue. *Annual Review of Criminology*, *1*, 63–84. https://doi.org/10.1146/ANNUREV-CRIMINOL-032317-092026/CITE/REFWORKS

Passel, J., & Cohn, D. (2009). *A Portrait of Unauthorized Immigrants in the United States*. www.pewhispanic.org.

Passel, J. S. (2006). The size and characteristics of the unauthorized migrant population in the US. In *procon.orgJS PasselPew Hispanic Center, 2006•procon.org*.

Passel, J. S., & Cohn, D. (2018). *U.S. Unauthorized Immigrant Total Dips to Lowest Level in a Decade*.

Passel, J. S., Van Hook, J., & Bean, F. D. (2004). *Estimates of Legal and Unauthorized Foreign Born Population for the United States and Selected States, Based on Census 2000*. https://scholar.google.com/scholar?hl=en&as_sdt=0%2C34&q=Passel%2C+Bean+and+Van+Hook+%282004%29&btnG=#d=gs_cit&t=1705604831617&u=%2Fscholar%3Fq%3Dinfo%3AiIzhJrnp17UJ%3Ascholar.google.com%2F%26output%3Dcite%26scirp%3D0%26hl%3Den

Police Data Initiative. (n.d.). *Open Data Portal*. Retrieved December 29, 2024, from https://www.policedatainitiative.org/

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Ratcliffe, J. H. (2021). Policing and public health calls for service in Philadelphia. *Crime Science*, *10*(1), 1–6. https://doi.org/10.1186/S40163-021-00141-0/FIGURES/4

School of Geography and Environmental Science University of Southampton, Department of Geography and Geosciences University of Louisville, Departement de Geographie Universite de Namur, Center for International Earth Science Information Network (CIESIN), & Columbia University. (2018). *WorldPop: Population Density*. Global High Resolution Population Denominators Project. https://hub.worldpop.org/geodata/summary?id=40759

Sedgwick, P. (2012). Multiple significance tests: the Bonferroni correction. *BMJ*, *344*(7841). https://doi.org/10.1136/BMJ.E509

Taft, D. R. (1933). Does immigration increase crime? *Social Forces*, *12*(1), 69–77. https://doi.org/10.2307/2570119

Timukaite, A., & Buil-Gil, D. (2024). Trust in the police and crime reporting: Reassessing assumptions in the context of violent crime in England and Wales. *International Review of Victimology*.

https://doi.org/10.1177/02697580241296569/ASSET/IMAGES/LARGE/10.1177_02697580241296569-FIG1.JPEG
U.S. Census Bureau. (2013). *Survey of Income and Program Participation*. U.S. Census Bureau.
U.S. Census Bureau. (2022a). *American Community Survey 5-Year Data (2009-2022)*. https://www.census.gov/data/developers/data-sets/acs-5year.html
U.S. Census Bureau. (2022b). *Survey of Income and Program Participation (SIPP)*. https://www.census.gov/sipp/
Van Hook, J. (2024). Beyond Stocks and Surges: The Demographic Impact of the Unauthorized Immigrant Population in the United States. *Population and Development Review*, *50*(4). https://doi.org/10.1111/PADR.12683
Van Hook, J., Bachmeier, J. D., Coffman, D. L., & Harel, O. (2015a). Can We Spin Straw into Gold? An Evaluation of Immigrant Legal Status Imputation Approaches. *Demography*, *52*(1), 329–354. https://doi.org/10.1007/S13524-014-0358-X/METRICS
Van Hook, J., Bachmeier, J. D., Coffman, D. L., & Harel, O. (2015b). Can we spin straw into gold? An evaluation of immigrant legal status imputation approaches. *Demography*, *52*(1), 329–354. https://doi.org/10.1007/S13524-014-0358-X
Wadsworth, T. (2010). Is Immigration Responsible for the Crime Drop? An Assessment of the Influence of Immigration on Changes in Violent Crime Between 1990 and 2000*. *Social Science Quarterly*, *91*(2), 531–553. https://doi.org/10.1111/J.1540-6237.2010.00706.X
Wang, X. (2012). Undocumented Immigrants as Perceived Criminal Threat: A Test of the Minority Threat Perspective. *Criminology*, *50*(3), 743–776. https://doi.org/10.1111/J.1745-9125.2012.00278.X
Wu, X., Koper, C., & Lum, C. (2022). Measuring the Impacts of Everyday Police Proactive Activities: Tackling the Endogeneity Problem. *Journal of Quantitative Criminology*, *38*(2), 343–363. https://doi.org/10.1007/S10940-021-09496-8/TABLES/8
Xie, M., & Baumer, E. P. (2019). Neighborhood immigrant concentration and violent crime reporting to the police: A multilevel analysis of data from the National Crime Victimization Survey*. *Criminology*, *57*(2), 237–267. https://doi.org/10.1111/1745-9125.12204