Title
Essays on the Economics of Crime

Author
Chalfin, Aaron James

Publication Date
2013

Peer reviewed|Thesis/dissertation

Essays on the Economics of Crime

By

Aaron James Chalfin

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Public Policy

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Steven Raphael, Chair

Professor Rucker Johnson

Professor Justin McCrary

Professor Jesse Rothstein

Spring, 2013

Essays on the Economics of Crime

Abstract

Essays on the Economics of Crime

by

Aaron James Chalfin

Doctor of Philosophy in Public Policy

University of California, Berkeley

Professor Steven Raphael, Chair

This dissertation considers the role that of various inputs in informing the market for crimes. Chapter 1 considers the "national" effect of immigration. Using panel data on U.S. cities and an instrument that leverages temporal variation in rainfall in different regions of Mexico and persistence in regional Mexico-U.S. migration networks, my findings indicate that Mexican immigration is associated with no appreciable change in the rate of either violent or property crimes in U.S. cities.

Chapter 2 leverages a natural experiment created by recent legislation in Arizona to estimate the impact on crime of an extremely large and discrete decline in the states foreign-born Mexican population. I show that Arizonas foreign-born Mexican population decreased by as much as 20 percent in the wake of the states 2008 implementation of the Legal Arizona Workers Act (LAWA), a broad-based E-Verify law requiring employers to verify the immigration status of new employees, coupled with severe sanctions for employer noncompliance. In order to isolate the causal effect of the passage and implementation of LAWA on crime, I leverage a synthetic differences-in-differences estimator, using a new method of counterfactual estimation proposed by Abadie, Diamond and Hainmuller (2010). In contrast to previous literature, I find significant and large effects of Mexican immigration on Arizonas property crime rate. Results are driven, in large part, by the fact that LAWA resulted in especially disproportionate declines among Mexican migrants who are young and male and, as such, the effects are predominantly compositional.

The final chapter, coauthored with Justin McCrary, considers the responsiveness of crime to police manpower. Using a new panel data set on crime in medium to large U.S. cities over 1960- 2010, we show that (1) year-over-year changes in police per capita are largely idiosyncratic to demographic factors, the local economy, city budgets, measures of social disorganization, and recent changes in crime rates, (2) year-over-year changes in police per capita are mismeasured, leading many estimates in the literature to be too small by a factor of 5, and (3) after correcting for measurement error bias and controlling for population growth, a regression of within-state differences in year-over-year changes in city crimes on within-state differences in year-over-year changes in police yields economically large point estimates. Our estimates imply that each dollar spent on police is associated with approximately $1.60 in reduced victimization costs, suggesting that U.S. cities employ too few police.

Acknowledgements

I would like to thank my principal advisors: Steve Raphael, Rucker Johnson, Justin McCrary and Jesse Rothstein. I would also like to acknowledge helpful comments from the following individuals who have provided feedback on individual chapters: Orley Ashenfelter, Emily Bruce, Shawn Bushway, David Card, Raj Chetty, Bob Cooter, Rachael Croson, John Dinardo, John Eck, Aaron Edlin, Benjamin Hansen, Hans Johnson, Louis Kaplow, Mark Kleiman, Tomislav Kovandzic, Prasad Krishnamurthy, Ron Lee, Thomas Lemieux, Morris Levy, John MacDonald, Jeff Miron, Denis Nekipelov, Emily Owens, Alex Piquero, Jim Powell, Kevin Quinn, Daniel Richman, Seth Sanders, David Sklansky, Geno Smolensky, Gary Solon, Kathy Speier, Hosung Sohn, Sarah Tahamont, Eric Talley, John Zedlewski and Frank Zimring.

In his seminal 1968 work, Gary Becker proposed that an individual's propensity to commit crimes is principally a function of three factors: (1) the individual's probability of capture for a given crime, (2) the severity of the sanction if captured and (3) the opportunity cost of the individual's time. Since 1968, a vast literature has arisen within economics that uses aggregate data to test each of the three predictions of Becker's economic model of crime. With regard to the probability of capture, a large literature considers the responsiveness of crime to police, to various police practices and modes of deployment. With regard to the severity of the sanction, a correspondingly large literature considers the responsiveness of crime to prison along both the intensive and extensive margin. With regard to the opportunity cost of crime, the literature has focused extensively on the responsiveness of crime to wages, unemployment rates and employment conditions. All three literatures find, to varying degrees, support for Becker's theory.

This dissertation considers the role that various inputs play in informing the supply of crimes in the United States. The final chapter, Chapter 3 which is co-authored with Justin McCrary, is tightly linked to the extant literature that has developed in response to Becker's economic theory of crime. This chapter considers the responsiveness of crime to police manpower in U.S. cities. The idea is that an increase in the number of police raise an individual's probability of capture, thus dicincentivizing and deterring crime (as well as lowering crime via the incapacitation effect of prison). On this question the literature has equivocated over time with more recent papers tending to find evidence of modest elasticities of crime with respect to police. These elasticities tend to be larger in magnitude for violent crimes than for property crimes, though this has remained an item of considerable controversy. Since violent crimes are considerably more costly to society than property crimes, there remains a troubling lack of consensus on the cost-beneficiality of police and, as such, on the appropriateness of public investment in police as opposed to prisons as the primary outlet for crime control dollars. This is a first order public policy question.

Our contribution to this literature is to generate more precise estimates of the elasticities of individual crime types with respect to police manpower. Using a new panel data set on crime in medium to large U.S. cities over 1960- 2010, Justin and I show that (1) year-over-year changes in police per capita are largely idiosyncratic to demographic factors, the local economy, city budgets, measures of social disorganization, and recent changes in crime rates, (2) year-over-year changes in police per capita are mismeasured, leading many estimates in the literature to be too small by a factor of 5, and (3) after correcting for measurement error bias and controlling for population growth, a regression of within-state differences in year-over-year changes in city crimes on within-state differences in year-over-year changes in police yields economically large point estimates. Our estimates are generally similar in magnitude to, but are estimated with a great deal more precision than, those from the quasi-experimental literature. Our estimates imply that each dollar spent on police is associated with approximately $1.60 in reduced victimization costs, suggesting that U.S. cities employ too few police. The estimates confirm a controversial finding from the previous literature that police reduce violent crime more so than property crime.

Chapters 1 and 2 of this dissertation consider the effect of immigration on crime in U.S. cities. Immigration can be seen as a shock to the crime market which may be viewed as being exogenous to the economic model of crime. In particular, while immigrants may face the same incentives as natives, they may be selected quite differently. That is, immigrants may be heavily selected with regard to their responsiveness to police, prisons and the opportunity cost of their labor. To the extent that this is true, immigration may have a direct effect on the crime rate that operates only indirectly with respect to the economic model of crime.

In Chapter 1, I consider the “national effect of immigration on crime. The chapter identifies a causal effect of Mexican immigration on crime using an instrument that leverages temporal variation in rainfall in different regions in Mexico as well as persistence in regional Mexico-U.S. migration networks. The intuition behind the instrument is that deviations in Mexican weather patterns isolate quasi-random variation in the assignment of Mexican immigrants to U.S. cities. My findings indicate that Mexican immigration is associated with no appreciable change in the rates of either violent or property crimes in U.S. cities. Notably, this is a precisely estimated null effect as I can reject that a one percentage point increase in the rainfall-induced share of Mexican migrants leads to greater than a one percent increase in the crime rate.

Chapter 2 leverages a natural experiment created by recent legislation in Arizona to estimate the impact on crime of an extremely large and discrete decline in the states foreign-born Mexican population. I show that Arizona’s foreign-born Mexican population decreased by as much as 20 percent in the wake of the states 2008 implementation of the Legal Arizona Workers Act (LAWA), a broad-based E-Verify law requiring employers to verify the immigration status of new employees, coupled with severe sanctions for employer noncompliance. In order to isolate the causal effect of the passage and implementation of LAWA on crime, I leverage a synthetic differences-in-differences estimator, using a new method of counterfactual estimation proposed by Abadie, Diamond and Hainmuller (2010). In contrast to previous literature, I find significant and large effects of Mexican immigration on Arizonas property crime rate. Results are driven, in large part, by the fact that LAWA resulted in especially disproportionate declines among Mexican migrants who are young and male and, as such, the effects are predominantly compositional.

Each of the three chapters contributes to an already large and growing literature within economics that seeks to understand the causes and consequences of criminal activity in urban areas. While much remains unexplained, the literature suggests that police manpower is an important predictor of crime. The role of immigration is less clear. The majority of the literature suggests that immigration is an unimportant factor in the aggregate crime production function. However, even if the effects are largely compositional, recent evidence from Arizona motivates caution in interpreting the aggregate evidence.

Chapter 1:
What is the Contribution of Mexican Immigration to U.S. Crime Rates? Evidence from Rainfall Shocks in Mexico

May 14, 2013

Abstract

This paper identifies a causal effect of Mexican immigration on crime using an instrument that leverages temporal variation in rainfall in different regions in Mexico as well as persistence in regional Mexico-U.S. migration networks. The intuition behind the instrument is that deviations in Mexican weather patterns isolate quasi-random variation in the assignment of Mexican immigrants to U.S. cities. My findings indicate that Mexican immigration is associated with no appreciable change in the rates of either violent or property crimes in U.S. cities. Notably, this is a precisely estimated null effect as I can reject that a one percentage point increase in the rainfall-induced share of Mexican migrants leads to greater than a one percent increase in the crime rate.

I. Introduction

Since 1980, the share of the U.S. population that is foreign born has doubled, rising from just over 6 percent in 1980 to over 12 percent in 2010. Compounding this demographic shift, the share of the foreign born population that is of Mexican origin also doubled, leading to a quadrupling of the fraction of U.S. residents who are immigrants from Mexico.[1] Over the same time period, crime rates in cities across the United States have declined considerably, in many cases, reaching historic lows. While the aggregate time series suggests that increases in immigration from Mexico have had a protective effect on crime, public opinion has generally reached the opposite conclusion, with a majority of U.S. natives indicating a belief that immigration is associated with increases in criminal activity (Espenshade and Calhoun 1993). Meanwhile, the consensus in the academic literature is that immigrants to the United States are, at worst, no more likely to participate in criminal activity than U.S. natives and, at best, may be far less likely to participate in crime (Butcher and Piehl 1998a; Butcher and Piehl 1998b; Reid, Adelman, Weiss and Jaret 2005; Moehling and Piehl 2007; Butcher and Piehl 2009; Stowell, Messner, McKeever and Raffalovich 2009; Wadsworth 2010).[2]

While recent empirical work suggests an answer to the conundrum, the literature remains unsatisfying in several ways. First, while historical research has successfully disaggregated the effect of early 20th century immigration by nationality, there is little research that addresses the criminal participation of recent Mexican immigrants. Since these are the immigrants who have become such a salient issue in recent policy debates, addressing the degree to which Mexican immigration is (or is not) associated with crime would appear to be an issue that is of first order importance. Second, while prior research has employed a variety of plausible identication strategies, chief among them the use of ethnic enclaves as an instrumental variable, concerns regarding the internal validity of this strategy motivates further investigation. Finally, the majority of the least squares literature identifies an effect of immigration on crime using long differences, generally employing decennial Census data. While this strategy plausibly addresses the problem of measurement errors in immigration data, such analyses lack the granularity of research designs that employ annual data and are subject to concerns regarding internal migration of U.S. natives in response to immigration or other changes in local conditions (Borjas 2003; 2006).

This research adds to the literature on immigration and crime in several important ways. First, by utilizing annual rather than decennial data on the share of immigrants and the crime rate, I am able to estimate the relationship between the two variables at a substantially more granular level than has been done in past research. Second, by limiting my analysis to Mexican immigration, I am able to isolate the specific migration

[1]As recently as 1970, the share of Mexican immigrants in the United States was only 1.5 percent (Hanson and McIntosh (2010).

[2]See Buonanno, Bianchi and Pinotti 2011 for similar research in a sample of Italian municipalities.

flows (young, low-skilled immigrants of Mexican origin) that have become such a salient issue in the policy debate surrounding immigration reform. Finally, with regard to the crime literature, I introduce a novel source of identifying variation in constructing an instrumental variable for the cross-city stock of immigrants in the United States. Specifically, I follow the general approach of Pugatch and Yang (2011) and construct an instrument that combines data on the permanent (long run) component of Mexican state-U.S. city migration relations with data on time-varying rainfall shocks in different Mexican states. The intuition behind the instrument is that deviations in Mexican weather patterns isolate quasi-random variation in the assignment of Mexican immigrants to U.S. cities. Indeed, I find strong evidence that Mexican immigration to the United States is responsive to Mexican rainfall. My findings indicate that, on net, Mexican immigration is associated with no appreciable change in the rates of either violent of property crimes in U.S. cities. Notably, this is a precisely estimated null effect as I can reject that a one percentage point increase in the rainfall-induced share of Mexican migrants leads to greater than a 1 percent increase in violent crimes or a 1.5 percent increase in property crimes. Finally, though I do find evidence that an increase in the share of Mexican migrants leads to a modest increase in per capita robberies, the result is sensitive to the inclusion of Los Angeles, underscoring the enormous heterogeneity in the treatment effect as well as the difficulty in identifying a "national effect" of Mexican immigration.

The remainder of the paper is organized as follows. Section II provides a discussion of identification problems in this literature as well as a brief literature review. Section III provides a discussion of mechanisms underlying the decision to migrate. Section IV presents the econometric framework used to estimate an average causal response of crime to immigration and includes a discussion of the identifying assumptions of the model. Section V describes the data and sample. Section VI presents the empirical results and includes a discussion that links the results to those estimated in the prior literature. Section VII concludes.

II. Conceptual Background

A. Empirical Challenges

Findings in the extant literature arise from two strains of research that attempt to identify the criminal participation of the foreign-born. The first examines the demographic characteristics of institutionalized populations and finds that recent immigrants are substantially underrepresented among those individuals who reside in an institutionalized setting at the time of the decennial census. In particular, Butcher and Piehl (1998a) find that the foreign-born are approximately five times less likely to be institutionalized than natives, further demonstrating that this figure is unlikely to be driven substantially by selective deportation. The advantage of research designs that compare the institutionalization rates of foreign-born to the native-born is that the descriptive nature of the exercise does not require a convincing source of identifying variation.[3] However, for several reasons, this line of research may not provide an internally-valid and policy relevant estimate of the contribution of immigration to cross-city crime rates. First, since it is not possible to disaggregate the incarcerated from the otherwise institutionalized using recent data, the validity of the resulting estimates requires an assumption that immigrants and natives have the same relative propensities to be incarcerated conditional upon institutionalization.[4] Second, the institutionalized population, by definition, includes only those individuals who were apprehended, arrested and subsequently incarcerated for a crime, a potentially highly selected sample of foreign-born offenders.[5] Finally, to the extent that immigration changes the calculus of offending among U.S. natives, an examination of the institutionalization rates of the foreign-born fails to capture general equilibrium effects associated with immigration. Thus, while the approach to studying the relationship between immigration and crime using individual-level microdata provides an important benchmark of the criminal involvement of the foreign born, this research is not a substitute for an empirical estimate of the effect of immigration on crime derived from aggregate data.

A second strain of research exploits cross-city variation in the stocks and flows of the foreign born and reports associations between changes in the size of a city's immigrant population and its crime rate. This research design offers a key advantage in that the researcher is able to observe associations between immigration and crime that are not contingent on an assumption of equal apprehension or adjudication probabilities among immigrants and natives.[6] However, to achieve identification, the design necessarily relies on the exogeneity of immigrant location decisions. To the extent that immigrants endogenously select destination cities either according to city-specific crime rates or according to other unobserved city and time-varying amenities that are themselves correlated with crime, the treatment effect uncovered by this research will be biased. To see this, consider the following two-way fixed effects model that we might be interested in estimating:[7]

$$CRIME_{it} = \alpha + \beta IMM^{*}_{it} + \phi_i + \rho_t + \varepsilon_{it} \tag{1}$$

In (1), IMM^{*}_{it} is the true measure of the immigrant share and ϕ_i and ρ_t are city and

[3] Moreover, it is important to note that such analyses plausibly capture an effect which is due to solely to the criminality of immigrants, rather than an effect of immigration that is a mixture of immigrant crimes and crimes committed by natives.

[4] The U.S. Census last differentiated between the incarcerated population and the population that is otherwise institutionalized in 1980.

[5] That said, empirical evidence supports the idea that immigrants may be more likely to be institutionalized conditional upon arrest. In particular, immigrants are more likely to face pre-trial detainment which, in turn, increases the likelihood of a conviction (Hagan and Palloni 1999).

[6] These area studies are also able to capture the "general equilibrium" effects of immigration insofar as these designs capture changes in the behavior of natives that arise as a result of immigration. The cost is that the treatment effect that is captured by such designs may not isolate the criminality of immigrants themselves.

[7] This is essentially the model estimated by Stowell, Messner, McKeever and Raffalovich (2009). Wadsworth (2010) pursues a similar approach, differencing (1) to remove the fixed effects.

year fixed effects, respectively. The assumption needed to identify this equation is that, conditional on the covariates: cov(IMM_{it}, ε_{it}) = 0. This condition is violated if IMM^*_{it} is correlated with V_{it}, some time-varying city characteristic that also predicts crime and is omitted from (1) or if $CRIME_{it}$ and IMM^*_{it} are simultaneously determined. To the extent that immigrants migrate on the basis of crime rates or on unmodeled variables that are related to crime rates, this condition will be violated. Likewise, (1) will return an inconsistent estimate of the effect of immigration on crime if a city's immigrant share is measured with error. Under the classical measurement error model in which the measurement errors are idiosyncratic and additive, errors in the measurement of the Mexican population share will lead to attenuation in β. However, correlated measurement errors can lead to biases in either direction. The dual problems of simultaneity/and omitted variables and measurement error in immigration data motivates the need for an instrumental variable that is as good as random and, as such, is plausibly uncorrelated both with both omitted variables and measurement errors.

B. The "Network" Instrument

The standard solution to this problem in the immigration literature is to instrument for recent flows of country-specific immigration with country-specific immigrant flows that are predicted by the national flow of migrants to the United States and the location decisions of past migrants, an instrument pioneered by Altonji and Card (1991) in their seminal treatment of the cross-city effect of immigration on the wages and employment of natives. The approach relies on the empirical observation that immigrants tend to cluster in cities where prior immigrants from their country of origin have previously settled. Thus the network instrument achieves identification by attempting to isolate exogenous variation in factors that *pull* immigrants to particular locations.[8] Formally, the network instrument is written as follows:

$$Z_{it} = \sum_{c=1}^{n} MIG_{ct} \times P_{ic} \tag{2}$$

In (2), MIG_{ct} is the number of immigrants from country c who are living in the United States in year t and P_{ic} is a matrix of source region-U.S. destination weights that return the conditional probability of migration from each source region c to each U.S. city i. The network instrument Z_{it} is the interaction of these two terms, summed over the n source regions. Card (2001) and Card and Lewis (2007) have used this instrument to estimate a causal effect of immigration on the employment outcomes of U.S. natives while Saiz (2003) has used the instrument to estimate the effect of immigration on various aspects of urban housing markets. With respect to crime using data from the 1980s, Butcher and Piehl (1998b) estimate the effect of immigration using the

[8] In a recent working paper, Chalfin and Levy (2012), argue that the "network" instrument" can be decomposed into a component that is explained by the size of foreign birth cohorts and a component that is captured by the conditional probability of migration for each (lagged) birth. The authors argue that the former term captures plausibly exogenous variation while the latter term, in part, captures "pull" variation to U.S. destinations.

network instrument in a panel of forty-three U.S. metropolitan areas and find that immigration is not associated with any type of crime, violent or property. This basic finding is echoed in least squares estimates of U.S. city panel data in Stowell, Messner, McKeever and Raffalovich (2009) and Wadsworth (2010) and, with the exception of robbery, in a recent study of immigrants in Italy (Buonanno, Bianchi and Pinotti 2011). An exception is a recent working paper by Spenkuch (2011) which uses the network instrument at the county-level and finds large effects of immigration on crime, a finding which is particularly large for Mexican immigrants.

To the extent that the lagged values of the stock of the foreign-born population do not directly affect contemporary crime rates, the network instrument presumably satisfies the exclusion restriction needed to achieve identification and returns an unbiased estimate of the effect of a specific exogenous flow of migrants on crime. Unfortunately, there are several mechanisms through which the prior location decisions of migrants might influence current crime rates, other than via their "pull" effect on subsequent migrants. First, to the extent that there is serial correlation in unobserved city-specific factors that are correlated with crime, the instrument might isolate not only exogenous variation in migration to that city but also migration that is drawn by persistent city-specific amenities. For example, if migrants are drawn to a particular city due to certain characteristics in 1960, to the extent that these characteristics persist, today's migrants may be pulled to a city for similar reasons. Second, as noted by Card (2001) and Pugatch and Yang (2011), the exclusion restriction will be violated if there are persistent city-specific shocks that differentially affect traditional gateway cities relative to non-gateways. For example, if differentially higher crime growth (or slower crime reductions) in gateway cities was a meaningful determinant of immigrant flows, then the network instrument would lead to an estimate of the effect of immigration on crime that is positively biased.[9] Instruments that rely on exogenous variation in factors that *pull* immigrants to a given city are inevitably problematic in that they rely on the presumably endogenous location decisions of prior immigrants or a lack of persistence in the characteristics of cities over time. In their recent study of the effect of immigration on the employment rates and wages of U.S. natives, Pugatch and Yang (2010) recognize this fact and propose that a cleaner source of identifying variation may be found in factors that induce migration from source countries. They argue that *push* factors (those factors that differentially induce migration from different source regions) are less likely to be systematically related to economic (or other) variables in the United States.

Another way to describe how the network instrument can fail to isolate exogenous variation in immigration flows is to consider that the network instrument can be decomposed into two components: (1) the available supply of Mexicans who are eligible to migrate to the United States ans (2) the conditional probability of migration in a given year. To see this, consider that in a given year t there is some number of Mexicans (N)

[9] It is also possible that increases in the stock of immigrants within a city lead to emigration of U.S. natives. While Card (2001) finds no evidence that this is the case, it may be the case that the composition of natives changes in the long run, in response to immigration.

who are available to migrate to the United States. N is itself a function of the number of lagged Mexican births (where the length of the lag will correspond with the ages of likely migrants) and the number of deaths among each cohort in N. The number of Mexicans who actually migrate to the United States in a given year is $N \times p_t$ where p_t is the conditional probability of migration to the United States in year t. Whereas N is a function of conditions in Mexico many years ago, p_t is a function of contemporary conditions in both Mexico and traditional migrant destinations in the United States. It is in this way that p_t creates a potential problem for the network instrument. For example, if a particular city is experiencing positive wage growth over a given time period, this wage growth might increase the conditional probability of migration, thus building in a negative bias to the network instrument.[10] Recognizing this, we would like to find a proxy for p_t which is not a function of conditions in the U.S. gateway cities.

A natural candidate to isolate "push" variation in immigration flows employs variation in weather. Weather variation, specifically rainfall, has been used as an instrument for internal migration in Indonesia (Levine and Yang 2011; Kleemans and Magruder 2011). With respect to Mexico, rainfall has been used to predict migration by Munshi (2003) and Pugatch and Yang (2011). As an instrument for the Mexican share of the U.S. labor force in a given state, Pugatch and Yang use deviations in rainfall from the long run mean in the Mexican states from which migrants to that U.S. state have historically originated. The intuition is that rainfall affects economic conditions in Mexico, which, in turn, alters propensities for affected Mexicans to migrate to the U.S. To the extent that there is persistence in Mexican state-U.S. state migration channels, migration to a given U.S. state can be thought of being induced quasi-randomly by rainfall in a particular Mexican sending state. In order to link migration from a given Mexican state to a given U.S. state, Pugatch and Yang construct measures of regional migration patterns that developed over time in response to the construction of early 20th century railroads. The authors note that a number of studies (such as Cardoso, 1980; Massey et al, 2002; and Woodruff and Zenteno, 2007) have documented the emergence of migration patterns between Mexican and U.S. regions connected by railroads at the beginning of the 20th century, as U.S. employers would travel by rail to Mexico and return with recruited laborers. Next, using data on migrants passing through three different border crossings collected by Forrester (1925), the authors construct a set of weights reflecting the probability that a migrant from a given Mexican state migrates to a given U.S. region.

Building on the approach of Pugatch and Yang, I construct a push instrument for immigration that links weather shocks in Mexico to long-term migration patterns between Mexico and the United States. However, in a key departure from their approach, I exploit microdata on migrants collected by the Mexican Migration Project (MMP) at Princeton University to develop estimates of the permanent component of migration from a given Mexican state to each of forty-six large U.S. metropolitan statistical areas. These data offer two important advantages over the cross-sectional data from

[10] The bias is negative to the extent that positive wages growth is, other things equal, associated with a reduction in crime.

border crossings employed by Pugatch and Yang. First, I am able to observe actual long-run migration patterns from each Mexican state to each U.S. city, rather than relying on a single cross-section of migrants entering through border crossings in the early 1920s. This is particularly important because the measure of the permanent component of long-run migration trends that I observe is determined over a longer period of time and includes not only legal but also illegal immigrants. Perhaps more importantly, I am able to estimate a model at the MSA rather than the state level. This is particularly salient to the study of crime since crime is determined primarily by local contextual factors (Bailey 1984). Using an instrument that combines annual data on rainfall with these long-run Mexican state-U.S. metropolitan area migration patterns, I develop a causal estimate of the contribution of Mexican immigration to crime rates in approximately fifty of the largest U.S. metropolitan areas. Since the instrument is activated only by rainfall shocks, it is as if, in each year, different numbers of Mexican immigrants were assigned at random to each U.S. city.

C. The Decision to Migrate

To be sure, Mexicans may migrate to the United States for any number of social or economic reasons, the sum of which are far too complex to capture in a stylized model of migration.[11] However, since this research uses an empirical approach that relies on the exogeneity of weather shocks, the migration mechanism that I will most plausibly capture and the resulting local average treatment effect that I will be able to estimate presumably arises from weather-induced changes in economic opportunities in Mexico. Borjas (1999) lays out a simple economic model of migration that couches the decision to migrate as one that depends on economic opportunities in a particular source country versus a candidate host country. In this simple model, residents in both the source country and the host country face a given earnings distribution which depends on the permanent component of the country's wage as well as an individual-specific wage shock:

$$log(w_0) = \delta_0 + \mu_0 \tag{3}$$

$$log(w_1) = \delta_1 + \mu_1 \tag{4}$$

In (3), w_0 is the realized wage in the source country which depends on δ_0 , the mean wage in the source country and a disturbance term μ_0 , which is normally distributed with mean zero and standard deviation σ_0^2. If the entire population of the source

[11] Economic theories of migration give rise to ambiguous predictions regarding the selection of migrants along dimensions related to criminal propensities. Economic theory typically assumes that individuals migrate from Mexico (a relatively poor country) to the United States (a relatively wealthy country) in search of higher earnings. To the degree that these earnings can be either licit or illicit, theory cannot generate obvious predictions about how migrants differ according to their criminal propensities. Moreover, given that migrants are selected according to their expected earnings in the U.S., if a subset of these migrants experience an unexpected lack of viable employment options, it is possible that these individuals may be especially willing to turn to criminal activity to compensate for their poor draw in the distribution of earnings in the U.S. On the other hand, if migrants are selected according to their earnings potential in the U.S., to the degree that earnings potential is positively correlated with characteristics that are negatively associated with participation in crime, selection may work in the opposite direction.

country were to migrate to the host country (which is assumed here to be wealthier than the source country), it would face a higher earnings distribution given by w_1 in (4).[12] Using equations (3) and (4), an index function can be defined to capture the binary choice faced by a potential source country migrant. This index function is constructed by forming an index I that is given by the ratio of the cost of migrating to the cost of remaining in the source country and substituting:

$$
\begin{aligned}
I &= log\left[\frac{w_1}{w_0 + C}\right] && (5)\\
&= log\left[\frac{\delta_1 + \mu_1}{\delta_0 + \mu_0 + C}\right] && (6)\\
&\simeq (\delta_1 + \mu_1) - (\delta_0 + \mu_0 + C) && (7)\\
&\simeq (\delta_1 - \delta_0) + (\mu_1 - \mu_0) - C && (8)
\end{aligned}
$$

The index function depends on three inputs: the wage distribution in the source country, the wage distribution in the destination country and the cost of migration. If the value of the index function is greater than zero, the individual chooses to migrate. In (8), C represents the nominal costs of migration (K) divided by $\delta_0 + \mu_0$, the realized wage in the source country. This is done to reflect the fact that the cost of migration is best expressed relative to an individual's income. Thus, a more complete representation of (8) is given by:

$$I \simeq (\delta_1 - \delta_0) + (\mu_1 - \mu_0) - \frac{K}{\delta_0 + \mu_0} \tag{9}$$

The individual chooses to migrate if $I > 0$. In other words, he migrates if he expects his wage to be higher in the host country than in the source country, net of migration costs. It is straightforward to see that, according to (9), a decline in the mean wage in the source country (δ_0) leads to an increase in the probability of migration. Likewise, a decrease in μ_0, the wage shock in the source country is also associated with an increase in the probability of migration. Indexing the parameters in (9) by individual and time subscripts, a natural extension of the model is that δ_0 is the permanent component of an individual's wage in the source country and that μ_0 is the transitory shock. The effect of rainfall on migration operates through this transitory wage shock and, as such, equation (9) has a special link to the first stage regression that I estimate between Mexican rainfall, weighted by permanent migration flows and annual city-specific immigration. Specifically, the coefficient on rainfall in the first stage regression is an aggregate estimate of $\frac{\partial I}{\partial \mu_0}$,the responsiveness of the latent migration probability to the rainfall shock, aggregated over all individuals in the dataset. An examination of (9) reveals that μ_0 appears twice on the right-hand side of the equation, leading to the

[12]As Borjas notes, while the earnings distribution that is faced by migrants from the source country in the host country is higher than they would have received, it will probably still be a lower distribution of wages that is faced by natives from the host country due to a difference in human capital acquisition between the source and host countries.

following estimate of the rainfall elasticity:

$$\frac{\partial I}{\partial \mu_0} = \frac{K}{(\delta_0 + \mu_0)^2} - 1 \tag{10}$$

Thus when $\frac{K}{(\delta_0 - \mu_0)^2} - 1 > 0$, an increase in rainfall leads to increased migration and when $\frac{K}{(\delta_0 - \mu_0)^2} - 1 < 0$, an increase in rainfall leads to decreased migration. Solving for μ_0, we see that, for $\delta_0 > 0$, migration is increasing in rainfall when:

$$|\mu_0| > \sqrt{K} - \delta_0 \tag{11}$$

Thus, when the absolute value of the rainfall-induced wage shock (μ_0) is greater than the gap between the migration cost and the permanent wage in the source country, migration is increasing in rainfall. Moreover, rainfall is predicted to lead to decreased migration when K is high and δ_0 is low. This is a sensible prediction as the chief reason why migration will increase with rainfall is that a large wage shock is needed to fund K. Hence when K increases, for a given μ_0, there will be less migration. Likewise, when δ_0 is low, an increase in rainfall is more likely to lead to higher migration as the rainfall is needed to fund K. Moreover, the model predicts that migration will increase most rapidly when the rainfall shock is large in absolute value. In other words, migration should be most responsive to extreme rainfall. More broadly, the framework presented in (11) reflects that a negative income shock can have a theoretically ambiguous effect on the probability that an individual migrates. On the one hand, a negative income shock makes migration more attractive as his expected wage differential between the two countries has now grown. On the other hand, migrants face real and binding constraints on the resources necessary to fund a migration episode. In the case of rainfall, to the extent that low rainfall depresses the local economy, potential migrants may face serious credit constraints that serve to reduce migration to the U.S. On the other hand, when times are tough, migrants face enhanced incentives to migrate, as is predicted by the simple model. Ultimately, whether reduced rainfall which leads to negative economic shocks, reduces or increases migration is an empirical proposition, one which I will test in my first stage regression.[13]

III. Identification Strategy

A. Econometric Framework

Using the Current Population Survey, 1986-2004, I begin with a sample of forty-six metropolitan areas with a 1980 population that exceeded 500,000 individuals, and I generate an estimate of the proportion of each city's population that is comprised of

[13]It is worth noting that while this simple framework assumes that migrants travel only from the source country to the host country, reverse migration is also possible. Thus when economic prospects improve in the source country or when those prospects are less variable, migrants in the host country may be more likely to return home.

individuals of Mexican origin in a given year (IMM_{it}).[14] By construction, IMM_{it} can be disaggregated into the number of Mexicans who migrate to the United States from each of thirty-two Mexican states:

$$IMM_{it} = \sum_{m=1}^{32} IMM_{mit} \tag{12}$$

Thus, in (12), the total number of Mexicans living in city i in year t is simply the sum of Mexicans in that city in that year who migrated from each of thirty-two Mexican states. Since IMM_{mit} is likely endogenous, it must be estimated using a source of plausibly exogenous variation. As Pugatch and Yang note, with data available on the source region of each Mexican migrant to the U.S. in each year, an instrument could be developed by regressing the number of Mexican migrants on a particular measure of rainfall for each Mexican state-U.S. city pair in the data and aggregating. Unfortunately, the sample sizes of available datasets do not permit such a granular analysis. As an alternative, following the general approach of Pugatch and Yang, I formulate IMM_{it} as a function of the total number of Mexican migrants from each Mexican state in each year (IMM_{mt}) and a set of Mexican state- U.S. city migration weights (P_{im}). However, in a key divergence from their approach, here the weights reflect an empirical measure of the permanent component of Mexican state-U.S. city specific migration flows, as opposed to a cross-sectional measure of Mexican state-U.S. state migration relations that were determined as long ago as 1924 according to the placement of railroad tracks. Equation (13) captures this relationship, with the inclusion of a time-and city- varying disturbance term that captures idiosyncratic shocks that are unrelated to the migration weights.

$$IMM_{it} = \sum_{m=1}^{32} (P_{im} \times IMM_{mt}) + \varepsilon_{it} \tag{13}$$

The weights (P_{im}) are estimated using the mean probability that a migrant from Mexican state mm migrates to each U.S. city using data from 1921-1985. [15] Next, I reformulate (6) to reflect the fact that migration from each Mexican state (IMM_{mt}) is instrumented for using rainfall shocks. In order to scale the instrumental variable in a way that generates an interpretable first stage regression coefficient, I multiply the Mexican state-U.S. city migration weights by $MIG_{mt=1980}$, an estimate of the total number of U.S.-bound Mexican migrants from each state in 1980 and divide this quantity by the population of each U.S. city in 1980. This procedure yields the following instrumental variable:

$$Z_{it} = MIG_{mt=1980} \times \frac{\sum_{m=1}^{32} P_{im} \times RAIN_{mt}}{POP_{ct=1980}} \tag{14}$$

[14]Following the approach of Butcher and Piehl (1998a) who studied crime at the MSA level, I choose the years 1986-2004 because coding of metropolitan statistical areas was largely consistent over this time period. The reason why I restrict the analysis to the MSAs with populations above 500,000 is because migration data from Mexican states to smaller MSAs is extremely limited.

[15]I choose 1985 as an end date to ensure that all of the migration relations contained in P_{im} are pre-determined with respect to the study sample.

In (14), for each of the thirty-two Mexican states, the time-invariant vector of migration weights to each city (P_{im}) is first multiplied by a column vector of the estimated number of U.S.-bound migrants from each Mexican state. The resulting term, $P_{im} \times MIG_{mt=1980}$ is the time-invariant estimate of the number of annual migrants from each Mexican state to each U.S. city. Next, this term is multiplied by the rainfall variable which varies by Mexican state and year. Hence, the term within the summation sign is an 46xT matrix which reflects the predicted number of migrants to each of the 46 cities in the dataset from 1986-2004 for a given Mexican state. Summing each of the terms in this matrix over the thirty-two Mexican states yields a predicted number of migrants for each city-year arising from rainfall in Mexico. Finally, the term is scaled by the size of the 1980 population in each MSA so that the instrument is expressed as a predicted flow of immigrants to a city in a given year.

Pugatch and Yang formulate $RAIN_{mt}$ in a number of ways but primarily as a z-score reflecting standardized deviations in rainfall from each state's long-run mean. In this research, I utilize both the z-score as well as a set of indicator variables that capture extreme deviations in rainfall in Mexican states. As is predicted y the theoretical model in Section II.C, to the extent that migrants face fixed costs associated with migration, it is likely that extreme deviations will be more salient predictors of migration. The indicator variables are defined such that $RAIN_{mt}$ is equal to one if rainfall is one standard deviation greater than the mean annual rainfall in each Mexican state from 1941-1985 and, alternatively that $RAIN_{mt}$ is equal to one if rainfall is one standard deviation lower than its state-specific long run mean. These versions of the instrument allow me to capture changes in migration that do not vary linearly in the z-score but are instead based on unexpectedly large rainfall shocks (that are either positive or negative). Finally, before specifying the first stage regression, it is necessary to consider potential temporal variation in the relationship between rainfall shocks and migration. That is, since migrants may not respond to rainfall shocks immediately, it is especially important to capture the relationship between the instrument and migration as flexibly as possible. Hence, I begin by specifying the first stage regression using a series of lags of the instrumental variable, beginning with a contemporaneous measure and adding one, two, and then three lags in additional specifications.[16] Equation (15) is a representation of the first stage regression where r takes on values between zero (to capture the contemporaneous relationship) and three.

$$IMM_{it} = \alpha + \beta_{FS}\left[MIG_{mt=1980} \times \frac{\sum_{m=1}^{32} P_{im} \times RAIN_{mt-r}}{POP_{it=1980}}\right] + \delta_i + \psi_t + \pi_{it} + \varepsilon_{it} \quad (15)$$

Referring to (15) δ_i represents a vector of U.S. city fixed effects. These terms de-mean IMM_{it} so that the instrument predicts deviations in the percentage of a city's Mexican population from its long-run mean. By de-meaning, I am netting out time-invariant city-specific characteristics that may explain the stock of Mexicans in each city. Like-

[16] I have utilized up to five lags of the instrument in models that are not reported in the paper. The first stage models with up to three lags of the instrument yield the greatest predictive power.

wise, ψ_t represents year fixed effects which control for annual migration shocks at the national-level. I also add a vector of linear city-specific time trends π_{it} to capture (either positive or negative) linear migration trends from Mexico to each city that that are independent of rainfall. Hence, the coefficients on the vector of lagged instruments are identified under fairly stringent identifying assumptions. That is, in order to satisfy the first stage, the instrument must predict deviations from the long-run mean of the Mexican proportion of a city's population that are not explained by annual national immigration trends or linear trends in the immigration series.[17] The corresponding outcome model yields the relationship between the outcome variable, the (log of) crimes per capita (Y_{it}) and rainfall-induced Mexican migration:

$$logY_{it} = \eta + \theta_{IV} I\hat{M}M_{it} + \delta_i + \psi_t + \pi_{it} + \varepsilon_{it} \tag{16}$$

In (16), $I\hat{M}M_{it}$ is the city's predicted Mexican share. The coefficient on this term, θ, represents the impact of a one percentage point increase in a city's Mexican share on the percentage change in the crime rate. Specifying the outcome equation in this way allows for a clear interpretation of θ, the parameter of interest. Since the dependent variable is scaled by the population, under the null hypothesis that immigration does not increase crime, increases in a city's Mexican share should not affect the crime rate. Accordingly a rejection of the null hypothesis that $\theta = 0$ is taken as evidence in favor of an effect of immigration on crime.

B. Identifying Assumptions

Conditional upon instrument relevance (which I discuss in Section VI), this research design identifies a causal effect under the following conditions:

1. The instrument (persistent migration relations weighted by rainfall) affects the per capita crime rate in a given network-linked U.S. city only through its effect on migration.
2. There are no individuals who migrate to the United States *only if* rainfall in their state in not extreme.

The first condition is the standard requirement for the exclusion restriction in an instrumental variables framework.[18] The second condition (that there are no "defiers of the instrument) is a standard restriction (monotonicity) under which a local average treatment effect is identified. [19] In order for the exclusion restriction to be met, rainfall must be conditionally random – that is, rainfall must succeed in assigning different numbers of Mexican immigrants to each U.S. city in a manner that is independent of any and all other variables, whether they are observed or unobserved. Despite the apparent randomness of rainfall, there are several ways in which the exclusion

[17]The coefficient on the instrument is the effect of the estimated rainfall shock on deviations from the long-run trend of a city's Mexican population. Where the instrument equals zero, the model predicts that the city's migration changes exactly according to a linear (or, in some cases, a quadratic) time trend.

[18]Formally, we are assuming that $cov(Z,\varepsilon)$=0.

[19]see Imbens and Angrist 1994 for a detailed discussion.

restriction could potentially fail in this context. First, rainfall shocks in Mexico could be correlated with a time-varying feature of a given city that affects crime through an alternate channel. For example, rainfall in Mexico might be correlated with rainfall shocks in linked U.S. cities, or, alternatively, with Mexican trade with the United States.[20] Fortunately, in their analysis, Pugatch and Yang roundly reject that this is the case.[21] A related possibility is that exports of narcotics from Mexico to the United States might, in fact, be a function of rainfall in Mexico. Thus, to the extent that crime in U.S. cities is a function of the supply (or the price) of drugs, crime could be related to rainfall through an alternative channel aside from immigration. While I am unable to directly test this, I note here that as long as the rainfall-induced supply shock to narcotics markets affects all cities equally in a given year, such an effect is picked up by the inclusion of year fixed effects. In other words, it need not be the case that Z is completely random - only that it is as good as random, conditional on the covariates in the model.[22] A second concern underlying this research design involves the potential selection of migrants from each Mexican state. While this concern does not involve the conditional randomness of rainfall and, as such, does not threaten the consistency of 2SLS, it nevertheless has implications for how 2SLS coefficients are interpreted and, accordingly, I discuss this consideration here. Specifically, since my analysis compares the change in the immigrant stock in each city to the change in its crime rate, under a homogenous treatment effect, an assumption of the analysis is that the average criminal propensities of immigrants from each Mexican state are equal. To the extent that Mexican states differ in the underlying criminality of the individuals who migrate to the U.S. as a result of rainfall, the resulting estimates may differ a great deal from city to city. In particular, we might be concerned that migrants from certain Mexican states migrate to a U.S. city explicitly in order to participate in that city's crime market. While I am unable to reject that this is the case, by using the permanent component of migration, I am isolating variation in Mexican migration that is the result of long-standing migration networks. In other words, while an association between rainfall in Mexico and marijuana exports could potentially affect the *timing* of migration, the instrument captures only migrants who leave Mexico for historically-linked U.S. destinations. As such, the criminally-involved migrant from Baja California who settles in Philadelphia (which is not a linked U.S. destination) to pursue a career in an underground market will not contribute to the average causal response that I estimate.

Finally, it is worth noting that the exclusion restriction is likely not violated even if there are errors in the measure of the immigrant stock I obtain from the Current Population Survey, a concern highlighted by Butcher and Piehl (1998b). Given that

[20] As Pugatch and Yang (2011) note, this might be the case if higher rainfall in a U.S. state's historical migrant origin areas in Mexico led to higher demands for U.S. goods (p. 24).

[21] The authors include U.S. weather patterns as well as U.S. state-level exports to Mexico as additional regressors and fail to reject the null hypotheses that these regressors are jointly equal to zero.

[22] I further note that the bias introduced by a "near exogenous" instrument is most serious if the instrument is also weak. The F-statistic on the instrumental variable used throughout the analysis exceeds 80, thus easing this concern.

this variable is almost certainly measured with error, at first blush, this would appear to be a first-order concern. However, while classical measurement errors in the immigrant share will result in attenuated OLS coefficients, since the immigrant stock is, in this research, the endogenous covariate that I am projecting on to the instrument, classical measurement errors in this variable will only decrease the precision of resulting estimates the estimates will still be consistent under the assumption that the measurement errors are uncorrelated with rainfall. As such, the rainfall instrument plausibly fixes two problems associated with least squares estimation - the problem of endogeneity and as well as problems arising due to the presence of measurement errors.

IV. Data

This research draws primarily on four different datasets to construct a city-by-year level analysis file. I begin with data on a city's Mexican population that is drawn from the March supplements of the Current Population Survey (CPS). In order to ensure appropriate cross-city comparisons, I use data on MSAs with a 1980 population that exceeds 500,000 individuals.[23] Because a variable that captures immigration status was added to the CPS only in 1994, in order to extend the series, I follow Pugatch and Yang (2011) and use a variable indicating Mexican nationality to capture the percentage of each city's population that is Mexican in a given year. While this approach does not allow me to isolate the percentage of a city's population that is comprised of Mexican immigrants, to the extent that a first stage relationship exists between rainfall in Mexican states and changes in the Mexican population of U.S. cities linked historically to those Mexican states, it is reasonable to expect that the relationship is being driven by a subset of individuals who are immigrants. That said, if the local average treatment effect being estimated captures a modest number of U.S.-born Mexicans, the coefficient vector on the instruments will simply estimate the reduced form effect of rainfall in Mexico on a U.S. city's total Mexican population. To the extent that Mexican immigration drives changes in the number of U.S.-born Mexicans either mechanically or through network effects, this is an important consideration.

Data on rainfall in Mexican states were obtained from the Mexican Migration Project environmental file. [24] The file contains data collected from local weather stations on monthly rainfall, for each Mexican state, from 1941-2005. Because the growing season in Mexico is year-long, I generate annual rainfall for each state in each year and standardize the data by subtracting each data point from its state-specific mean and dividing by its state-specific standard deviation to obtain a z-score.

Data used to construct P_{im}, the matrix of Mexican state-U.S. city specific time- invariant migration weights were generated from the Mexican Migration Project's migrant

[23]500,000 is chosen both to ensure comparability between cities and also because the number of U.S. bound migrants from each Mexican state that I am able to observe in these cities becomes very small.

[24]The Mexican Migration Project is the product of a collaboration between researchers at Princeton University and the University of Guadalajara in Mexico. The MMP is co-directed by Jorge Durand and Douglas S. Massey.

level file. The file contains survey data on a sample of over 7,000 individuals, each of whom migrated to the United States at least once in their lifetime. The migrants are a subset of individuals who were sampled at random within each community sampled in the dataset. Communities were chosen in order to provide variation in the characteristics of sending regions. While communities were not surveyed explicitly because they send large numbers of migrants to the United States, communities nevertheless needed to send at least a few migrants in order to be surveyed.[25] Each community was sampled once and individuals who reported having migrated to the United States were asked to retrospectively recall each of their prior migration experiences.[26] Among male household heads, 23 percent reported having migrated to the United States within three years of the time of survey with 89 percent reporting an undocumented migration spell (Hanson 2002).[27] Using data on the U.S. destination for the migrants first migration episode, I remove from this file all migrants whose first migration experience occurred after 1985 and construct a matrix of weights that represent the average propensity of a migrant from a given Mexican state to migrate to each U.S. MSA in the dataset.[28] Thus, the weights were constructed from the migration experiences of 3,981 Mexican migrants. Table 1A provides descriptive detail on the weights, showing the top three U.S. destination areas for migrants from each Mexican state. The percentage of

[25] The survey sample covers the following Mexican states: Aguascalientes, Baja California Norte, Chihuahua, Colima, Durango, Guanajuato, Guerrero, Hidalgo, Jalisco, Mexico, Michoacan, Morelos, Nayarit, Nuevo Leon, Oaxaca, Puebla, San Luis Potosi, Sinaloa, Tlaxcala, Veracruz, Yucatan, and Zacatecas. Within each state, communities are classified as either *ranchos* (fewer than 2,500 inhabitants), *pubelos* (between 2,500 and 10,000 residents), *mid-sized cities* (10,000 to 100,000 residents) or *large metropolitan areas* (100,000 or more residents). In pueblos and ranchos, MMP investigators conduct a complete census of dwellings and randomly select households to survey from among the entire community. In mid-sized cities and large metropolitan areas, MMP investigators selected established neighborhoods.

[26] As is always the case, when retrospective survey data are used, there is a concern that recall bias will compromise the resulting estimates. Since I use the MMP survey data to document the first stage relationship between immigration that is predicted by rainfall and actual immigration, to the extent that recall bias leads to errors in the matrix of Mexican state-U.S. city weights, the resulting first stage estimates will be weaker than those derived from error-free survey data. However, two points are worth noting. First, errors will only accrue to the extent that an individual recounts a fictitious trip — that is, a trip to a destination to which that individual did not ever travel. In the event that individuals simply switch the ordering of trips, the weights will continue to reflect legitimate migration relations. Having constructed the weights using an individual's first trip to the United States, a trip which should be easier to recall than a second or third trip, I expect such bias to be minimized. More importantly, random errors in the migration weights will affect the reduced form and the first stage estimates equally. As such, errors in the weights will serve only to increase the standard errors in the second stage estimates without introducing bias.

[27] Hanson further notes that the MMP surveys only households in which at least one member has remained in Mexico. As such, households that have entirely moved to the United States are not counted. Moreover, the migrants who are surveyed are a selected subset of migrants who have returned to Mexico, at least temporarily. For a detailed discussion of the MMP's migrant level file, see Hanson (2002).

[28] In principle, I could have used the migrant's last migration episode. However, it is likely that the first migration experience is more likely to reflect network ties between the source and destination communities. In practice, the magnitude of the elements of the matrix are almost completely invariant to the choice of migration episode.

migrants who settled in each area is given in parentheses next to the name of the metropolitan area. For example, the top two U.S. destinations for migrants from Baja California del Notre, located along the border with San Diego, CA are San Diego and Los Angeles. Likewise, the top three U.S. destinations for migrants from Nuevo Leon, a state in eastern Mexico are Houston, Dallas and McAllen, TX. While there is a fair amount of spread in the number of U.S. destinations in the dataset, the leading cities are predictably Los Angeles, Chicago, Houston, Dallas and San Diego.[29]

Finally, data on crimes reported to police were obtained from the Federal Bureau of Investigation's Uniform Crime Reports (UCR), the standard source of data on crimes at the agency level that is employed in aggregate-level crime research. Since 1934, the UCR has, either directly or through a designated state reporting agency, collected monthly data on index crimes reported to local law enforcement agencies. The index crimes collected consistently since 1960 are: murder (criminal homicide), forcible rape, robbery, aggravated assault, burglary, larceny and motor vehicle theft.[30] In order to maintain consistency with the level of aggregation of the migration data from the MMP, I aggregate the agency-level UCR data to the MSA level.

Forty-six cities are used in the analysis. For these cities, the 1986-2004 CPS datafile is comprised of 3,067,064 individuals of whom 6.8 percent are identified as individuals of Mexican origin. In the U.S., the Mexican population is 52 percent male with an especially high number of males represented among the prime working ages. From 1986-2004, the percentage of the U.S. population that is Mexican origin nearly doubled, increasing from 4.7 percent in 1986 to 9.2 percent in 2004. Over the same time period, on a per capita basis, reported violent crimes and property crime fell by more than 25 percent and 28 percent respectively. Figures 1A and 1B plot the number of reported violent and property crimes in the United States over the 1960-2008 period. While both violent and property crime rates rose during the period from 1986-1990, since 1990 crime has fallen monotonically.

Table 2 presents summary statistics for the Mexican share and each of the crime rates for individual cities, both in levels and in logs. In particular, for each variable employed in the analysis, I present a mean, a minimum value, a maximum value and three types of standard deviations - the overall standard deviation as well as the between (cross-sectional) and within standard deviations. The average city in the sample is 10 percent Mexican, a number that ranges from less than 1 percent to 88 percent over the entire study period. Notably, nearly all of the variation in a city's Mexican share is cross-sectional, while within-city variation is relatively small. This reflects the fact that while some cities (e.g., Los Angeles and Houston) are persistent destinations for Mexican immigrants while other cities (e.g., New York and Boston) are not. With

[29] See Table 1B for additional details.

[30] The UCR employs an algorithm known as the "hierarchy rule" to determine how crimes involving multiple criminal acts are counted. In order to avoid double counting, the UCR classifies a given criminal transaction according to the most serious statutory violation that is involved. For example, a murder-robbery is classified as a murder.

regard to crime, several features of the data are worth noting. First, approximately six in seven crimes reported to police are property crimes with an average large MSA documenting 6,300 property crimes and 1,000 violent crimes per 100,000 residents. Second, the most serious crimes (murder and rape) account for less than 1 percent of all crimes reported to police while less serious offenses such as larceny account for nearly half of all crimes. With regard to the decomposition of variance, the picture for crime is more mixed than it is for the Mexican share. I note that the between variation is dominant for the violent crimes (murder, rape, robbery and aggravated assault) while the between and within variation are more equally apportioned for the property crimes (burglary, larceny and motor vehicle theft).

V. Results

A. First Stage Estimates

In the first stage, I estimate the effect of several different incarnations of the rainfall instrument on deviations from the long-run mean of the proportion of the Mexican population in U.S. cities. There are three primary conceptualizations of the instrument that I explore. First, I specify the rainfall variable as a z-score such that each Mexican state's rainfall in a given year is expressed in terms of standard deviations from it's mean over the 1941-1985 period. This variable, which uses deviations in rainfall to proxy for transitory shocks to each Mexican state's economy, captures the (migration-weighted) linear effect of rainfall where low values of the instrument reflect drought and high values of the instrument reflect an abundance of precipitation. Next, I re-specify the rainfall variable as an indicator variable that is equal to unity if the rainfall z-score in a given Mexican state-year was greater than 1. This variable allows me to test for the possibility that extreme positive deviations in rainfall that cause migration. Finally, in order to capture an effect of droughts, I specify an instrument that captures extreme low (< -1 SD) deviations in rainfall. A positive relationship between the low rainfall dummy and migration might be the case if individuals face fixed costs associated with migration and migrate to escape (extreme) economic hardship.

Prior to running the first stage regressions, it is important to examine variation in each of the instruments to get a sense of the type of variation that is being captured. The average value of the z-score instrument is -0.00012 (SD = 0.0028). It is sensible that the instrument is centered around zero as the mean rainfall z-score is zero. An analysis of variance reveals an intraclass correlation coefficient of 0.95, indicating that nearly all of the variation in the instrument is comprised of within-city variation. This too is sensible as the instrument is activated by rainfall and assumes that Mexican state-U.S. city migration relations are constant over time. Thus each city, over many years, receives a rainfall-induced migration shock that is of roughly equal magnitude though, within cities, there is much temporal variation. Turning to the dummy variable version of the instruments, the story is subtly different. Since the instrument is now equal to zero if none of the Mexican sending states experienced a (positive or negative) rainfall shock in a given year and a positive number indicating the strength of migration ties otherwise, the variables are distributed quite differently. An examination of the extreme value instruments reveals that 24.4 percent of city-years experienced at least one high rainfall shock from a Mexican sending state and an equivalent percentage of city-years experienced at least one low rainfall shock. Roughly 10 percent of city-years experienced at least one of each type of rainfall shock.

In order to carefully examine the pathways through which rainfall influences migration, I begin by specifying very parsimonious first stage models, including one lag of the instrument at a time. Table 3a presents regression results for models using the z-score instrument. In Table 3a, columns 1-4 report coefficients on the contemporaneous instrument and each of one, two and three lags. Column 5 includes a specification containing all four lags in a single model. To derive a national-level estimate, all models are estimated using weighted least squares where 1980 MSA population is used to weight the observations. In addition, all models include city and year fixed effects as well as city-specific linear time trends. Standard errors, which are clustered at the city-level, are reported in parentheses below the estimated coefficients. In each column, I report the F-statistic on all excluded instruments in the model with the corresponding critical value for the weak instruments test suggested by Stock and Yogo (2005) below. A visual inspection of the coefficients in Table 1 reveals weak predictive power of rainfall. The coefficients change sign depending upon the lag employed and an F-test on the instruments reveals that, despite a tendency for the coefficients to be positive, they are only weakly significant and typically do not exceed Stock-Yogo threshold. In order to pin down the precise mechanism through which rainfall induces migration, I re-specify the instrument using dummy variables that capture extreme (+/-1 SD) deviations in rainfall. Those first stage estimates are reported in Table 3b. Table 3b, which is laid out the same way as Table 3a, reveals a robust, positive relationship between extreme deviations in rainfall and migration. The pattern is sensible and it explains why the linear instrument was not successful in predicting migration. That is, since large deviations (both positive and negative) are positively associated with migration, the impact of the extreme positive values tends to negate the impact of the extreme negative values, leading net effect of rainfall to be not substantially different from zero. The strength of the relationship is strongest at two lags. However, at each lag, coefficients on the excluded instruments are positive and nearly always significant and the corresponding F-test generally meets or exceeds the Stock-Yogo critical values each time. Using the contemporaneous instruments as well as three additional lags, yields an F-statistic of 80.2 on the excluded instruments.[31] This value of the F-statistic easily exceeds the corresponding critical value for one endogenous covariate and eight excluded instruments as is recommended by Stock and Yogo (2002).[32]

[31] In a model with one endogenous regressor, eight excluded instruments and a desired maximal bias of 0.10, the threshold for the F-statistic is 33.8. The critical value for a maximal bias (relative to OLS is 20.3.

[32] I also apply the Kleibergen-Paap rk Wald test for underidentification a test for whether the matrix of instruments and endogenous regressors is of full column rank. The test is valid for data that is not i.i.d.

In order to demonstrate that the instrument is a valid predictor of cross-city migration, I subject the first stage model to a series of additional robustness checks. First, I estimate (7) using quadratic in addition to linear city-specific time trends. This ensures that the instrument predicts within city changes in migration above and beyond a more flexibly specified time trend. Using my preferred first stage specification, the F-statistic on the excluded instruments remains sufficiently high at 80.4. Second, to ensure that the results of the analysis are not driven by the way that each city is weighted, I estimate (7) without city population weights. Without the weights, in my preferred specification, the F-statistic on excluded instruments is 56.3. Third, I re-estimate the first stage equation excluding Los Angeles and report an F-statistic on the excluded instruments of 43.9. Fifth, I condition on a vector of time-varying controls that capture variation in a city's social and economic conditions as well as demographic composition.[33] Finally, I re-estimate (7) using leads, rather than lags of the instruments. If the instrument were spuriously correlated with migration flows, we might expect that leads of the instrument were correlated with migration just as contemporaneous and lagged versions of the instrument are. Since leads of rainfall cannot have a causal effect on migration, I interpret evidence of an association between leads of the instrument and migration as evidence of a spuriously measured relationship. In order to check that the causal pathway through which rainfall instruments that I employ in my preferred specification. Whereas the F-statistic on lags of the excluded instruments was 80.2, the F-statistic on leads of the excluded instruments is well below the Stock-Yogo critical value and none of the coefficients are significant at conventional levels.[34] Finally, I present results from a series of tests of overidentifying restrictions which unilaterally fail to reject the null hypothesis of exogeneity. In particular, because the number of instruments exceeds the number of endogenous regressors, my IV equation is overidentified allowing me to test the exogeneity of my instruments under the assumption of a constant local average treatment effect. Since I cluster my standard errors at the city level, I utilize Hansen's J-test which produces a test statistic that is robust to arbitrary dependence in the within-city errors. In Table 4, I present results from Hansen's J-test of overidentifying restrictions for each of the models that is tested in the paper. Each row contains models in which the dependent variable is the log of a different UCR crime rate. Along the columns, for each choice of regression weights, I run the J-test for all instruments. A cursory glance at Table 4 reveals that I fail to reject the null hypothesis that the instruments are exogenous for all crime models. The results of these tests provide support for (though do not automatically validate) my identifying assumption that rainfall in Mexican states is uncorrelated with U.S. crime rates except through migration. Using the final set of first

When the data are i.i.d., the test is equivalent to the test of Donald and Cragg (1993). The test statistic is significant at $p < 0.001$ allowing me to reject the hypothesis that the first stage model is underidentified.

[33]The full set of variables includes the employment-to-population ratio, the poverty rate, the 12th grade high school dropout rate, the per capita number of sworn police officers and a series of variables capturing a city's demographic composition.

[34]The F-statistic on excluded instruments using one and two leads is 5.9.

state regressions reported in Table 3b, I proceed to estimating my second stage models.

B. Least Squares and 2SLS Estimates

In the outcome equation, I regress both the level and the log of each of seven index crimes on the predicted change in a city's share of Mexican migrants. Prior to presenting 2SLS results, I present results from a series of least squares regressions of the crime rate (also measured in levels and logs) on the share of Mexican migrants. These estimates are presented in Table 5. In Table 5, the first two columns correspond to models in which the crime rate is measured in logs while the second set of columns corresponds to models in which the crime rate is measured in levels. Within each panel, each row corresponds to a different index crime with the first two rows (violent crimes and property crimes, respectively) corresponding to the two crime aggregates. Finally, each model is specified both with and without MSA population weights.[35] As with the first stage models, all regressions are estimated using city and year fixed effects and city-specific linear time trends, with standard errors clustered at the city level.

Beginning with the log crime models, I note that regression coefficients have been multiplied by 100 for ease of interpretation. Thus, referring to the weighted least squares estimates, we see that a one percentage point increase in the Mexican share is associated with a 0.1 percent decrease in the rate of violent crimes and a 0.3 percent decrease in the rate of property crimes. Both of these estimates are small both in an economic sense and relative to their standard errors. In fact, the least squares models are estimated with extraordinary precision all around as I am typically able to reject increases or decreases in the crime rate on the order of 0.5 percent. The precision of the models is due to the fact that since city fixed effects and linear time trends explain such a large share of the within-city variation in the crime rate (with corresponding R^2 values exceeding 0.98), the explanatory power of the models is extremely high and the corresponding sampling variance is small. The magnitude of the coefficients and standard errors on violent and property crimes is broadly consistent with results reported by Butcher and Piehl (1998b) for a panel of cities and cross-sectional results reported by Reid et al (2005) in which OLS results for a 2000 cross section of U.S. cities were analyzed. For example, Butcher and Piehl (1998b) conditioning of fixed effects, report a violent crime coefficient of -0.25 percent (S.E. = 1.15 percent). As their sample is less than half the size of the sample I employ, it is sensible that the standard errors I obtain are smaller.

Referring to the disaggregated crimes, several patterns in the data are worth noting. First, Mexican immigrants are associated with higher rates of per capita rapes and burglaries and lower rates of per capita larcenies, though notably the degree to which coefficients are significant depends a great deal on whether or not the analysis employs MSA population weights. This instability of the coefficients suggests a great deal of heteorgeneity amongst receiving cities. Second, in addition to estimating the models in

[35]The population weights use the MSA's 1980 population. In order to get a sense of the degree to which there is heterogeneity in the results, I also weight by the share and the size of an MSA's Mexican population in 1980. The results are largely invariant to this weighting scheme.

logs, I also provide estimates of the association between Mexican immigrants and crime in levels. Here, a one percentage point increase in the Mexican share is associated with 4.8 additional burglaries per 100,000. Taken as a whole, the mixture of positive and negative coefficients for different crime types and the sensitivity of the estimated effects to the inclusion of population weights presents little consistent evidence against a null effect. The results, once again, are broadly consistent with prior cross-city research that finds little evidence of an association between immigration and crime.[36]

Table 6 presents 2SLS estimates of the relationship between predicted Mexican immigration and crime. Because the Mexican proportion of the population is estimated, point estimates in these models are less precisely estimated than the least squares coefficients presented in Table 5. However, it is worth noting that they remain extraordinarily precise (depending on the crime, standard errors are typically under 1-2 percent.[37] Consistent with the results from OLS models, IV results presented in Table 6 provide little evidence against the null hypothesis that Mexican immigration is not associated with crime. Neither the violent nor the property crime models reveal a significant relationship. Moreover, in those models, I can reject the possibility that a one percentage point increase in the Mexican share is associated with more than a 1 percent increase or more than a 1.5 percent increase in the rate of violent crimes and property crimes, respectively. Referring to the individual crime categories, while coefficients for murder, rape, assault, burglary and larceny do not meet conventional thresholds for significance in any of the models, there is some evidence in favor of a positive effect of Mexican immigration on robbery and motor vehicle theft. While the robbery result is not significant in unweighted regressions, with population weights, a one percentage point increase in the Mexican share is associated with a 2.7 percent increase in robberies or an increase of 15.4 robberies per 100,000 residents. The motor vehicle theft coefficient is significant only in the unweighted models and, even then, is only significant when the crime rate is measured in levels. To compare my results more explicitly to those in the extant literature, I note that while Butcher and Piehl (1998b) do not report 2SLS, they indicate that those coefficients are very similar to those obtained using OLS. My results are quite similar to OLS models reported in their 1998 paper.

C. Heterogeneity and Robustness

In Table 7, I present a series of robustness checks designed to test the sensitivity of the estimates reported in Table 6 to minor specification changes. Column (1) of Table 7 replacates column (2) of Table 6 in which the log of the crime rate is regressed on the instrumented Mexican population share weighting by city population. Each subsequent column presents the equivalent estimates with a particular change in specification.

Given the notable differences between weighted and unweighted regression models,

[36]The only crime type for which results are significant in both weighted and unweighted models is rape.

[37]To the degree that I measure the Mexican population in each city with random error, the resulting IV estimates will be measured more imprecisely. However, the resulting IV estimate will remain consistent.

a natural extension of the paper would involve an exploration of the degree to which there is heterogeneity in the effect of immigration on crime across different types of cities. While the relatively small number of cities in my sample limits my power to test for heterogeneity in the estimates along a vector of initial city characteristics, I can nevertheless provide several important tests of the robustness of the reported results. I begin by testing whether the treatment effects reported in Table 6 are driven by one or two "important" cities. In particular, we might be concerned that the null effects reported in Table 6 are an artifact of a null effect in one or two influential cities, rather than a pattern that is consistent across all cities. In columns (2) and (3) of Table 7, I drop Los Angeles and Chicago, the two largest destination cities in the sample and repeat the main analyses presented in Table 6 of the paper. The coefficients reported in columns (2) and (3) differ in magnitude to an extent from the coefficients reported in column (1) which uses the full sample. However, given sampling variability, they are, on the whole, similar to those in Table 6. However, the robbery result is a notable exception. While the positive coefficient in the robbery model survives the exclusion of Chicago, it does not survive the exclusion of Los Angeles, indicating that the positive coefficient on robbery appears to derive from local conditions that are specific to Los Angeles. This finding underscores the difficulty of estimating a "national effect" of immigration and serves as a reminder that immigration may have very different effects depending on a variety of contextual factors.[38]

Next, I consider whether the results reported in Table 6 are robust to the inclusion of a standard set of covariates. Each of the regressions reported in Table 6 conditions on state and year fixed effects as well as city-specific linear time trends. I omit covariates in the main regressions because the standard covariates included in crime regressions may not be exogenous with respect to Mexican immigration. Nevertheless, it is instructive to test whether the estimated coefficients are sensitive to the inclusion of city-specific variables which may be correlated both with immigration and crime. Column (4) of Table 7 reports the results of a series of 2SLS regressions of the crime rate on instrumented Mexican immigration conditional on fixed effects, time trends and a number of key covariates, each of which is theoretically linked with crime at the city level. The employment-to-population ratio and the poverty rate, computed using the Current Population Survey March supplements, capture fluctuations in each MSA's local economy . While not directly linked to crime, The 12th grade high school dropout rate in the MSA's largest cities (those with a population exceeding 50,000) is a proxy for changes in local conditions which may correlate with higher levels of crime.[39] The

[38]In addition, in order to test for whether the inclusion of time trends results in an "overfitted" model, I re-estimate the basic specification presented in column (1) of Table 7 on the 1995-2004 subsample of the data. The estimated coefficients for the 1995-2004 subsample are as follows: murder (0.76), rape (1.26), robbery (-0.74), assault (-1.81), burglary (-1.06), larceny (-2.82) and motor vehicle theft (-0.23). None of the coefficients are significant at conventional levels.

[39]The dropout rate proxy in year t is one minus the number of 12th graders in the city in year t relative to the number of 11th graders in the city in year t 1. The data come from the National Center for Educational Statistics.

per capita number of sworn police officers with arrest powers comes from the Uniform Crime Reports police employees file and is included to capture the effect of changes in the level of law enforcement on the crime rate. Next, I include a series of demographic variables that capture changes in each city's population structure that are potentially correlated with the crime-rate. These variables include the proportion of each city's population who are unmarried, the proportion who are white, black and asian, and the proportion of the population in a granular set of age groups. The proportion of the population aged 16 through 24 are entered separately in the regression, as these are the peak ages of participation in criminal activity.

The primary takeaway from column (4) of Table 7 is that the coefficients reported in Table 6 are robust to the inclusion of these controls. For each of the nine crime rates, the estimates in column (4) are more positive than those in column (1) that do not condition on the controls but the substantive findings remain unchanged.[40] With the exception of robbery, the coefficients do not differ from zero and in no case do the estimates in column (4) differ significantly from those in column (1).

In column (5) of Table 7, I test whether the results that I obtain can be explained by rainfall-driven changes in the age and gender composition of Mexicans. In particular, if rainfall serves to induce young males to migrate to a greater extent than older females, rainfall in Mexico might result in crime in U.S. cities due to positive selection on characteristics that are associated with crime. In order to check for the importance of such a mechanism, I re-specify the 2SLS regression models presented in Table 6 adding as an explanatory variable the change in the proportion of Mexicans in a given city who are males between the ages of 15-45. To the extent that changes in offending are purely driven by changes in the demographic composition of Mexican immigrants, this variable will capture the effect of these changes. In general, we should expect the coefficients to decrease in size with the inclusion of this control. When this variable is added to each of the crime models, while the resulting coefficients are, in general, slightly smaller than those reported in Table 6, they are extremely similar in magnitude indicating that the results are not driven, to an appreciable degree, by changes in demographic composition. This finding is sensible as, conditional on year and city fixed effects and time trends, the remaining variation in the demographic composition of Mexican migrants is quite low.[41]

Finally, I consider the possibility that immigration may affect crime rates via a temporal lag. This might be the case, for example, if arriving migrants commence criminal involvement only after having failed to successfully integrate into local labor markets. In Table 8, I consider test whether instrumented immigration in year t affects crime in

[40] Due to missing data, the sample size for the regressions reported in column (4) of Table 7 is slightly lower (n=736) than in Table 6.

[41] In addition, in an auxiliary analysis, I examine the degree to which the instruments predict the change in the Mexican immigrant share for each of ten age-gender groups. I find evidence the eight instruments have predictive power for all ten age-gender groups, providing intuition that the instrument has not isolated an unusual local average treatment effect.

years *t+1* and *t+2*. Coefficients arising from these IV regressions which are population weighted and use log crime are presented in columns (2) and (3) of Table 8 alongside the basic estimates from Table 6 in column (1). Referring to the models that operate using one lag, the estimates are broadly similar to the models that assume a contemporaneous relationship. With regard to the crime aggregates, the estimates remain imprecisely estimated, with standard errors of roughly identical magnitude to those presented in column (1). An exception is the effect of Mexican immigration on the murder rate which rises by approximately 3.6 percent in response to a one percentage point increase in the immigrant share in the previous year. Unlike the significant result for robbery, the result is robust to the exclusion of Los Angeles and Chicago, indicating a plausibly national affect. The effect of immigration on murder is likewise large, though imprecise, in twice lagged models. These results leave open the possibility that Mexican immigration may be associated with serious criminal violence in the years following an immigration shock. Failure to detect significant effects in any of the other crime categories is consistent with an explanation that the effect may be associated either with gang violence or an increase in intimate partner homicides, rather than offending with an acquisitive motive.

D. Comparisons wth the Network Instrument

Given that the identification strategy employed in this research departs from the traditional "network" instrument by relying on rainfall variation, it is instructive to briefly consider how the estimated effects differ when the rainfall instrument is used as opposed to the network instrument. While the network instrument was pioneered for use with long-differenced data, it can nevertheless be constructed using annual data. Recall that the network instrument is calculated by multiplying the number of migrants who enter the United States from a given source regionin a given year by a set of weights that capture the (pre-determined) conditional probability that a migrant sojournes to each destination. Summing over all source regions, the network instrument yields a predicted number of migrants that each destination region receives in each year. This incarnation of the network instrument was adapted by Altonji and Card (1991) to predict the number of immigrants entering each city in the United States between two given Census years. Using annual data, I construct a version of the network instrument that is consistent with the Altonji and Card version of the instrument by multiplying an estimate of the stock of Mexicans in the United States in each year by the proportion of Mexican migrants who settled in each U.S. city prior to 1986. Thus the instrument predicts Mexican immigration solely using the stock of Mexicans living in the United States and the pre-determined network weights, without the benefit of rainfall.

Table 9 reproduces the basic 2SLS results using the rainfall instrument that are reported in Table 6 alongside estimates of the effect of Mexican immigration on crime using the network instrument for Mexico. On the whole, the results differ very little when the network instrument is employed with a notable exception: robbery. While the rainfall instrument identifies a positive effect of Mexican immigration on robbery (when Los Angeles is included in the data), the network instrument identifies no such positive effect. Differences between the rainfall and network instruments have implications for the direction of selection bias of immigrants with respect to robbery which, the direction of which is theoretically ambiguous. In particular, it appears that immigrants may have been more likely to migrate to the United States at times when either robbery rates (or an unmodeled factor that is correlated with robbery rates) in network-linked U.S. cities were falling . Under these conditions, the network instrument would fail to detect a positive causal association between Mexican immigration and robbery.[42]

VI. Conclusion

In this paper, I estimate the effect of Mexican immigration on the rate of crimes reported to the police in U.S. metropolitan areas. When I instrument for migration using rainfall shocks in network-linked Mexican states, the evidence suggests that Mexican immigration tends to be associated with neither higher nor lower levels of overall crime. Notably, this zero is precisely estimated as I can reject that a one percentage point increase in a city's Mexican immigrant share leads to greater than a 1 percent change in rates of either violent or property crimes. At the same time, I find evidence that Mexican immigration is associated with a modest increase in robberies, though the result appears to be driven by Los Angeles. The results are robust to the inclusion of a series of standard control variables as well to controls for changes in the age and gender composition of Mexican immigrants. The results also vary a great deal across different cities which is apparent in the sensitivity of the estimated coefficients to the inclusion of population weights suggesting that a "national effect of Mexican immigration does not exist in a meaningful sense. These results are broadly consistent with the extant literature which has tended to report either null or weakly negative effects of immigration on crime. While the results may appear surprising, they are sensible as changes in the demographic composition of cities tend to only weakly predict changes in crime, conditional on fixed effects. The results are also largely robust to allowing the effect of immigration to operate through a lag. When the effect of immigration on crime is assumed to occur via and one- or two-year lag the results remain indistinct from zero with the exception of murder which appears to rise in the year after an influx of migrants.

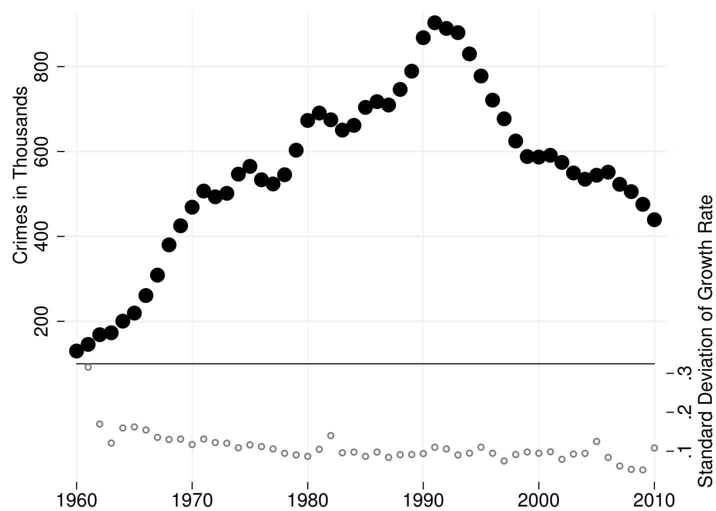
To my knowledge, this is the first paper to exploit plausibly exogenous "push variation in a source country to estimate the impact of immigration on crime. While my findings largely mirror those in the extant literature, to the extent that rainfall mimics a random assignment mechanism in allocating immigrants to U.S. cities, this research helps to resolve any remaining skepticism regarding the identification strategies employed to generate past findings. Moreover, this research isolates the effect of Mexican immigration on crime, thus addressing a key source of contention in contemporary policy debates regarding appropriate immigration policy. For several reasons, estimates in this paper likely represent an upper bound on the criminality of immigrants. First, to the extent that recent Mexican immigrants tend to possess observable characteristics

[42]It is also possible that each instrument captures a different local average treatment effect.

(e.g., lower rates of human capital and lower wages) that are typically associated with higher criminal propensities, it is plausible to conclude that, if there is an economically meaningful effect of immigration on crime, it should be observable among Mexican migrants. Second, as the effect that I identify is a reduced form estimate of the effect of immigration on crime, is it possible that a portion of the observed effect is driven by increases in crime among natives, rather than among immigrants. This might be true, for instance, if Mexican immigrants are attractive crime victims or if Mexican immigrants destabilize employment markets for U.S. natives. Further research along these lines is needed. In particular, it is important to understand the apparently contradictory findings from the literature that examines the demographic characteristics of U.S. prisoners and the cross-city literature. While the former finds ample evidence that immigrants (including Hispanic immigrants) are less likely to be incarcerated than natives, the cross-city literature generally find little evidence of any effect of immigration on crime. While this paper does not resolve this debate, it adds a critical data point to the cross-city literature.

FIGURE 1. AGGREGATE TRENDS IN VIOLENT AND PROPERTY CRIME: EVIDENCE FROM THE UNIFORM CRIME REPORTS

A. Violent Crime: Murder, Rape, Robbery, Aggravated Assault



B. Property Crime: Burglary, Larceny, Motor Vehicle Theft

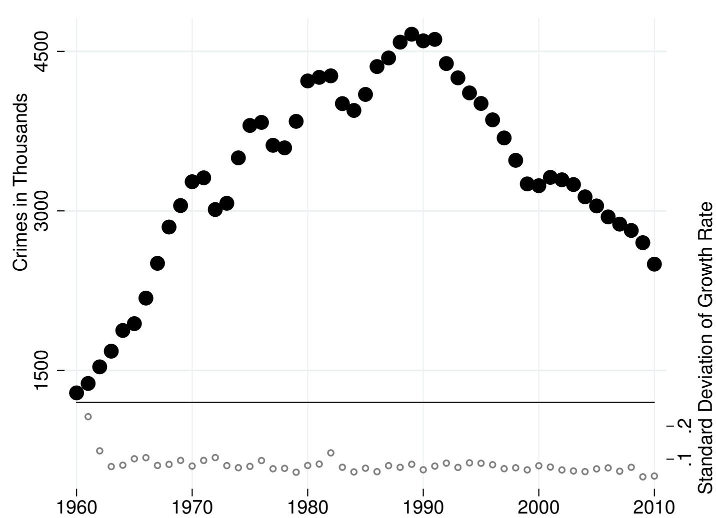


TABLE 1A. U.S. DESTINATIONS OF MEXICAN IMMIGRANTS

Mexican State	Destination #1	Destination #2	Destination #3
Aguascaliente	Los Angeles (20%)	Reno (6%)	Tulsa (6%)
Baja California del Notre	San Diego (60%)	Los Angeles (22%)	
Baja California del Sur			
Campeche			
Coahuila de Zaragoz			
Colima	Los Angeles (41%)	Fresno (9%)	
Chiapas			
Chihuahua	El Paso (16%)	Los Angeles (9%)	Dallas/Phoenix (9%)
Districto Federal	Los Angeles (20%)	Chicago (11%)	Orange County (CA) (8%)
Durango	Chicago (23%)	Los Angeles (19%)	Dallas (7%)
Guanajuanto	Los Angeles (15%)	Chicago (11%)	Houston (7%)
Guerrero	Chicago (29%)	Los Angeles (15%)	Phoenix (12%)
Hidalgo	Las Vegas (12%)	Dallas (9%)	Houston (7%)
Jalisco	Los Angeles (26%)	San Diego (6%)	San Jose (4%)
Mexico (Estado)	Chicago (32%)	Stockton (10%)	los Angeles (7%)
Michoacan	Los Angeles (20%)	Fresno (8%)	Chicago 96%)
Morelos	Los Angeles (29%)	Minneapolis (18%)	Chicago (10%)
Navarit	Los Angeles (29%)	San Jose (10%)	Orange County (CA) (7%)
Nuevo Leon	Houston (16%)	McAllen (15%)	Dallas (11%)
Oaxaca	Los Angeles (51%)	San Diego (9%)	
Puebla	New York (56%)	Los Angeles (23%)	
Querataro			
Quintana Roo			
San Luis Potosi	Houston (16%)	San Diego (16%)	Dallas (6%)
Sinaloa	Los Angeles (48%)	San Diego (10%)	Riverside (8%)
Sonora			
Tamaulipas			
Tabasco			
Tiaxcala	Los Angeles (9%)		
Veracruz	Los Angeles (14%)	Chicago (13%)	San Jose (8%)
Yucatan	Portland (31%)	San Francisco (29%)	Los Angeles (11%)
Zacatecas	Los Angeles (28%)	Fresno (5%)	Merced (5%)

Note: The table reports the three largest U.S. matropolitan area destinations for migrants from each Mexican state, among migrants in the Mexican Migration Project's Migrant File, 1924-1985.

TABLE 1B. MEXICAN STATE SOURCES OF U.S.-BOUND IMMIGRANTS
SELECTED U.S. METROPOLITAN AREAS

U.S. Metropolitan Area	Source #1	Source #2	Source #3
Atlanta	Jalisco (23%)	Nuevo Leon (12%)	Veracruz (11%)
Austin-San Marcos	San Luis Potosi (33%)	Veracruz (26%)	Guerrero (21%)
Chicago	Durango (30%)	Jalisco (25%)	Guanajuanto (19%)
Dallas	Guanajuanto (28%)	Durango (26%)	Jalisco (11%)
Denver	Yucatan (58%)	Chihuahua (14%)	Districto Federal (7%)
El Paso	Chihuahua (64%)	Zacatecas (9%)	Veracruz (5%)
Fresno	Jalisco (44%)	Michoacan (15%)	Guanajuanto (14%)
Houston	San Luis Potosi (50%)	Guanajuanto (15%)	Michoacan (7%)
Las Vegas	Jalisco (43%)	Nayarit (14%)	Districto Federal (13%)
Los Angeles-Long Beach	Jalisco (23%)	Michoacan (10%)	Guanajuanto (9%)
Merced	Nayarit (43%)	Jalisco (23%)	Michoacan (18%)
Minneapolis-St. Paul	Morelos (100%)		
New York	Puebla (56%)	Morelos (22%)	Tlaxcala (5%)
Oakland	Jalisco (58%)	Michoacan (36%)	Districto Federal (2%)
Orange County (CA)	Jalisco (25%)	Guerrero (20%)	Guanajuanto (13%)
Philadelphia	Guanajuanto (91%)	Districto Federal (4%)	
Phoenix	Chihuahua (30%)	Guanajuanto (16%)	Durango (12%)
Portland	Yucatan (91%)		
Riverside-San Bernardino	Michoacan (22%)	Jalisco (20%)	Yucatan (9%)
San Diego	Baja California del Norte (61%)	San Luis Potosi (16%)	Jalisco (7%)
San Francisco	Yucatan (54%)	Jalisco (13%)	Nayarit (10%)

Note: The table reports the three most prevalent source regions among Mexican immigrants to selected U.S. metroplitan areas.
These data are based upon the experiences of migrants surveyed in the Mexican Migration Project's Migrant File, 1924-1985.

TABLE 2. SUMMARY STATISTICS

		Logs				Levels			
		Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.
Mexican share	O	1.11	1.79	1.00	4.48	9.8	15.1	0.0	88.4
	B		1.67				14.9		
	W		0.71				2.9		
Violent crimes	O	6.82	0.51	5.18	8.27	1,047.8	569.0	178.5	3,913.0
	B		0.46				508.6		
	W		0.22				265.4		
Property crimes	O	8.71	0.30	7.66	9.72	6,316.6	1,888.1	2,111.4	16,576.4
	B		0.24				1,526.4		
	W		0.18				1,132.3		
Murder	O	2.38	0.65	0.60	4.39	13.7	11.6	1.8	80.8
	B		0.58				10.6		
	W		0.30				5.1		
Rape	O	3.88	0.42	2.17	5.15	52.7	22.8	8.7	172.8
	B		0.35				18.3		
	W		0.23				13.7		
Robbery	O	5.83	0.61	4.30	7.36	409.5	267.8	73.8	1,578.0
	B		0.55				231.8		
	W		0.28				138.0		
Assault	O	6.22	0.56	4.04	7.67	583.6	332.3	56.6	2,137.9
	B		0.51				299.6		
	W		0.25				149.9		
Burglary	O	7.16	0.40	5.78	8.29	1,392.9	571.1	323.7	3,994.3
	B		0.27				385.8		
	W		0.16				424.7		
Larceny	O	8.23	0.31	7.21	9.20	3,949.9	1,194.5	1,358.2	9,905.6
	B		0.27				1,037.0		
	W		0.16				610.9		
Motor vehicle theft	O	6.74	0.55	5.09	7.93	973.8	503.1	163.1	2,788.3
	B		0.46				408.6		
	W		0.31				299.1		

Note: The table reports summary statistics for the Mexican population share and each of the crime variables both in logs and in levels. For each variable, we report the overall mean, the standard deviation decomposed into overall ("O"), between ("B"), and within ("W") variation, as well as the minimum and maximum values.

TABLE 3A. FIRST STAGE REGRESSION MODELS [Z-SCORE INSTRUMENT]

	(1)	(2)	(3)	(4)	(5)
Rainfall instrument	52.4				-36.5
	(50.7)				(48.6)
$(t-1)$		118.0			61.4
		(30.9)			(27.5)
$(t-2)$			131.7		53.4
			(20.4)		(16.1)
$(t-3)$				85.1	51.4
				(39.6)	(32.0)
N	893	846	799	752	752
R-squared	0.984	0.985	0.986	0.986	0.986
F-statistic on excluded instrument	1.1	14.6	41.5	4.6	4.6
Stock-Yogo critical value (10% maximal bias)	16.4	16.4	16.4	16.4	24.6

Note: The table reports coefficients and standard errors for a series of least squares regressions of the city's Mexican immigrant share on several different lags of the rainfall instrument, in z-scores. Depending on the number of lags, models utilize between 752 and 893 observations covering 46 metropolitan statistical areas from 1986-2004. The table reports WLS regressions using 1980 MSA population weights. All models contain city and year fixed effects as well as city-specific linear time trends. The F-statistic that is reported is the joint hypothesis test that the coefficients on all excluded instruments are equal to zero. The Stock-Yogo critical value is the critical value associated with one endogenous regressor and the appropriate number of excluded instruments. Standard errors (in parentheses) are clustered at the city level.

TABLE 3B. FIRST STAGE REGRESSION MODELS [EXTREME VALUES INSTRUMENT]

	(1)	(2)	(3)	(4)	(5)
High rainfall	507.1				36.9
	(260.3)				(441.7)
$(t-1)$		876.9			300.0
		(184.1)			(159.6)
$(t-2)$			831.5		840.9
			(86.6)		(196.3)
$(t-3)$				679.8	492.9
				(219.8)	(356.2)
Low rainfall	504.0				-180.7
	(114.9)				(183.1)
$(t-1)$		662.8			631.5
		(151.6)			(212.6)
$(t-2)$			320.0		352.1
			(184.7)		(290.4)
$(t-3)$				251.8	-36.3
				(331.4)	(362.1)
N	893	846	799	752	752
R-squared	0.984	0.985	0.985	0.986	0.986
F-statistic on excluded instrument	14.7	19.1	53.2	22.3	80.2
Stock-Yogo critical value	19.9	19.9	19.9	19.9	33.8
(10% maximal bias)					

Note: The table reports coefficients and standard errors for a series of least squares regressions of the city's Mexican immigrant share on several different lags of the rainfall instrument, in extreme values. Depending on the number of lags, models utilize between 752 and 893 observations covering 46 metropolitan statistical areas from 1986-2004. The table reports WLS regressions using 1980 MSA population weights. All models contain city and year fixed effects as well as city-specific linear time trends. The F-statistic that is reported is the joint hypothesis test that the coefficients on all excluded instruments are equal to zero. The Stock-Yogo critical value is the critical value associated with one endogenous regressor and the appropriate number of excluded instruments. Standard errors (in parentheses) are clustered at the city level.

TABLE 4. HANSEN'S J-TEST OF OVERIDENTIFYING RESTRICTIONS

	Logs		Levels	
	Unweighted	Weighted	Unweighted	Weighted
Violent crimes	0.22	0.19	0.35	0.52
Property crimes	0.43	0.45	0.35	0.37
Murder	0.15	0.78	0.51	0.96
Rape	0.20	0.55	0.45	0.72
Robbery	0.07	0.10	0.35	0.34
Assault	0.54	0.34	0.51	0.41
Burglary	0.31	0.48	0.32	0.41
Larceny	0.58	0.48	0.48	0.41
Auto theft	0.23	0.61	0.21	0.49

Note: The table reports the p-values from Hansen's heteroskedasticity and cluster robust J-test of overidentifying restrictions. All models utilize 729 observations covering 46 metropolitan statistical areas from 1986-2004. The first set of estimates report the p-value of the test using the log of the number of crimes. The second set of estimates report the p-value of the test using number of crimes in levels. Unweighted regressions and WLS regressions using 1980 MSA population weights are reported. All models contain city and year fixed effects as well as city-specific linear time trends.

TABLE 5. LEAST SQUARES MODELS OF THE EFFECT OF MEXICAN IMMIGRATION ON CRIME

	Logs		Levels	
	Unweighted	Weighted	Unweighted	Weighted
Violent crimes	-0.05	-0.10	0.7	-0.2
	(0.23)	(0.24)	(2.0)	(2.7)
Property crimes	-0.42	-0.29	-22.7	-11.5
	(0.20)	(0.23)	(11.1)	(13.1)
Murder	0.15	0.08	0.0	0.0
	(0.30)	(0.57)	(0.5)	(0.1)
Rape	0.73	0.62	0.4	0.4
	(0.21)	(0.25)	(0.1)	(0.1)
Robbery	-0.46	-0.05	-1.5	-0.4
	(0.22)	(0.26)	(0.9)	(1.5)
Assault	0.10	-0.26	1.8	-0.3
	(0.31)	(0.33)	(1.5)	(1.8)
Burglary	0.08	0.28	-0.2	4.8
	(0.21)	(0.27)	(2.7)	(3.3)
Larceny	-0.80	-0.61	-27.7	-19.0
	(0.22)	(0.21)	(7.5)	(8.5)
Auto theft	0.02	-0.03	5.2	2.7
	(0.25)	(0.30)	(2.5)	(3.4)

Note: The table reports coefficients and standard errors for a series of least squares regressions of the number of crimes on the Mexican population share, conditional on the MSA population. Each model utilizes 752 observations covering 46 metropolitan statistical areas from 1989-2004. The first set of estimates report the effect of a one percentage point increase in the Mexican population share on the log of the number of crimes. The second set of estimates report the effect of a one percentage point increase in the Mexican population share on the number of crimes in levels. Unweighted regressions and WLS regressions using 1980 MSA population weights are reported. All models contain city and year fixed effects as well as city-specific linear time trends. Standard errors (in parentheses) are clustered at the city level.

TABLE 6. TWO STAGE LEAST SQUARES MODELS OF THE EFFECT OF MEXICAN IMMIGRATION ON CRIME

	Logs		Levels	
	Unweighted	Weighted	Unweighted	Weighted
Violent crimes	-1.14	0.31	-8.8	7.6
	(1.31)	(0.99)	(9.9)	(12.9)
Property crimes	-0.88	-0.05	-31.5	23.6
	(1.76)	(1.54)	(90.5)	(81.1)
Murder	0.49	0.57	0.1	0.1
	(1.22)	(1.33)	(0.1)	(0.2)
Rape	-0.52	-0.96	-0.4	-0.3
	(0.56)	(0.63)	(0.3)	(0.3)
Robbery	1.03	2.73	5.6	15.4
	(1.08)	(1.06)	(3.9)	(7.4)
Assault	-2.88	-0.68	-13.7	-3.0
	(1.67)	(1.35)	(7.2)	(8.8)
Burglary	-1.07	1.23	-4.9	31.6
	(1.75)	(1.86)	(19.3)	(20.4)
Larceny	-2.35	-0.95	-71.1	-35.0
	(1.92)	(1.64)	(58.6)	(51.9)
Auto theft	2.07	0.12	44.6	27.0
	(1.50)	(2.19)	(20.2)	(29.7)

Note: The table reports coefficients and standard errors for a series of 2SLS regressions of the crime rate on the Mexican population share. Mexican population share is instrumented using predicted rainfall-induced immigration. Each model utilizes 752 observations covering 46 metropolitan statistical areas from 1986-2004. The first set of estimates report the effect of a one percentage point increase in the Mexican population share on the log crime rate. The second set of estimates report the effect of a one percentage point increase in the Mexican population share on the number of crimes in levels. Unweighted regressions and WLS regressions using 1980 MSA population weights are reported. All models contain city and year fixed effects as well as city-specific linear time trends. Standard errors (in parentheses) are clustered at the city level.

TABLE 7. TWO STAGE LEAST SQUARES MODELS OF THE EFFECT OF MEXICAN IMMIGRATION ON CRIME: ROBUSTNESS CHECKS

	(1)	(2)	(3)	(4)	(5)
Violent crimes	0.31	-0.98	0.31	1.18	0.20
	(0.99)	(1.47)	(0.99)	(0.91)	(1.00)
Property crimes	-0.05	-0.90	-0.20	0.93	-0.11
	(1.54)	(2.09)	(1.53)	(1.43)	(1.54)
Murder	0.57	1.11	-0.17	0.77	0.80
	(1.33)	(1.37)	(1.52)	(1.86)	(1.37)
Rape	-0.96	-0.47	-0.96	0.16	-0.87
	(0.63)	(0.81)	(0.63)	(0.97)	(0.63)
Robbery	2.73	1.24	2.52	3.17	2.81
	(1.06)	(1.27)	(1.08)	(1.18)	(1.08)
Assault	-0.68	-1.59	-1.23	1.07	-1.00
	(1.35)	(1.89)	(1.30)	(1.49)	(1.68)
Burglary	1.23	-1.29	0.88	2.09	1.06
	(1.86)	(1.97)	(1.89)	(1.85)	(1.34)
Larceny	-0.95	-3.22	-1.03	0.26	-0.98
	(1.64)	(2.07)	(1.60)	(1.85)	(1.64)
Auto theft	0.12	0.72	0.06	0.47	0.07
	(2.19)	(2.48)	(2.18)	(2.65)	(2.21)
Los Angeles included	yes	no	yes	yes	yes
Chicago included	yes	yes	no	yes	yes
Economic/social controls	no	no	no	yes	no
Control for prime-age males	no	no	no	no	yes

Note: The table reports coefficients and standard errors for a series of 2SLS regressions of the log crime rate on the instrumented Mexican population share. Column (1) reproduces 2SLS estimates reported in Table 6 that use the full sample and condition on city and year fixed effects and city-specific linear time trends. In columns (2) and (3), Los Angeles and Chicago are excluded from data, respectively. Column (4) conditions on a vector of control variables: the employment-to-population ratio, the porverty rate, the 12th grade dropout rate, the number of sworn police officers per capita and a series of demographic controls. Column (5) controls for the change in the proportion of the Mexican population that is comprised of prime age males. All models are weighted by 1980 MSA population. Standard errors are clustered at the MSA level.

TABLE 8. TWO STAGE LEAST SQUARES MODELS OF THE EFFECT OF MEXICAN IMMIGRATION ON CRIME ROBUSTNESS TO FUNCTIONAL FORM

Year of crime measurement relative to immigration:	Year *t*	Year *t+1*	Year *t+2*
Violent Crimes	0.31	0.72	-0.48
	(0.99)	(1.13)	(1.30)
Property crimes	-0.05	0.49	-0.13
	(1.54)	(1.64)	(1.23)
Murder	0.57	3.60	4.57
	(1.33)	(1.61)	(3.66)
Rape	-0.96	-0.77	-1.91
	(0.63)	(0.85)	(1.17)
Robbery	2.73	2.72	0.64
	(1.06)	(1.28)	(1.12)
Assault	-0.68	0.01	-0.42
	(1.35)	(1.56)	(1.82)
Burglary	1.23	1.28	-0.46
	(1.86)	(1.80)	(1.32)
Larceny	-0.95	-0.38	-0.48
	(1.64)	(1.74)	(1.34)
Auto theft	0.12	1.16	-0.42
	(2.19)	(2.08)	(1.65)

Note: The table reports coefficients and standard errors for a series of 2SLS regressions of the log crime rate on the Mexican population share. Mexican population share is instrumented using predicted rainfall-induced immigration. Column (1) reproduces estimates presented in column (2) of Table 6. Columns (2) and (3) assume that the effect of immigration on crime operates with one- and two-year lags, respectively. Regressions are weighted using 1980 MSA population weights. All models contain city and year fixed effects as well as city-specific linear time trends. Standard errors (in parentheses) are clustered at the city level.

TABLE 9. TWO STAGE LEAST SQUARES MODELS
OF THE EFFECT OF MEXICAN IMMIGRATION ON CRIME
COMPARING THE RAINFALL AND THE NETWORK INSTRUMENT

Instrument	Violent crimes	Property crimes	Murder	Rape	Robbery	Assault	Burglary	Larceny	Auto theft
Rainfall instrument	0.31 (0.99)	-0.05 (1.54)	0.57 (1.33)	-0.96 (0.63)	2.73 (1.06)	-0.68 (1.35)	1.23 (1.86)	-0.95 (1.64)	0.12 (2.19)
Network instrument	0.40 (1.27)	-0.31 (1.45)	0.44 (0.89)	0.48 (0.62)	-0.10 (1.29)	0.69 (1.80)	0.69 (1.41)	-0.82 (1.66)	-0.40 (1.08)

Note: The table reports coefficients and standard errors for a series of 2SLS regressions of the log crime rate on the instrumented Mexican population share. Two different instruments are employed. The top row reproduces the estimates from column (1) of Table 7 and reports output from models in which the rainfall instrument is used to predict the Mexican population share. The bottom row of the table reports output from models in which the "network" instrument for Mexico is used to predict the Mexican population share. Regressions are weighted using 1980 MSA population weights. All models contain city and year fixed effects as well as city-specific linear time trends. Standard errors (in parentheses) are clustered at the city level.

Chapter 2: New Evidence on Mexican Immigration and U.S. Crime Rates: A Synthetic Control Study of the Legal Arizona Workers Act

Abstract

In this study, I leverage a natural experiment created by recent legislation in Arizona to estimate the impact on crime of an extremely large and discrete decline in the state's foreign-born Mexican population. I show that Arizona's foreign-born Mexican population decreased by as much as 20 percent in the wake of the state's 2008 implementation of the Legal Arizona Workers Act (LAWA), a broad-based E-Verify law requiring employers to verify the immigration status of new employees, coupled with severe sanctions for employer noncompliance. By contrast, the law appears to have had no effect on the state's share of other foreign-born individuals or U.S.-born Hispanics. In order to isolate the causal effect of the passage and implementation of LAWA on crime, I leverage a synthetic "differences-in-differences" estimator, using a new method of counterfactual estimation proposed by Abadie, Diamond and Hainmuller (2010). To provide a direct estimate of the effect of Arizona's Mexican immigrant share on its crime rate, I extend the synthetic differences-in-differences framework to construct implied synthetic instrumental variables estimates, using LAWA as an instrument for a state's Mexican population share. In contrast to previous literature, I find significant and large effects of Mexican immigration on Arizona's property crime rate. Results are driven, in large part, by the fact that LAWA resulted in especially disproportionate declines among Mexican migrants who are young and male and, as such, the effects are predominantly compositional. The remainder of the paper considers how to interpret these estimates. In particular, I present a theoretical model of immigrant offending and characterize analytically the conditions under which an empirical estimate of the immigrant share on the reported crime rate will be a conservative estimate.

VII. Introduction

Over the past thirty years, crime rates in cities across the United States have plummeted, in many cases, reaching fifty-year lows (Zimring 2006). At the same time, the share of the foreign born among the U.S. population has increased rapidly, with the foreign-born Mexican share of the population quadrupling since 1980. A research literature in both criminology and economics suggests that, at a minimum, immigration has played no role in this historic decline in crime (Butcher and Piehl 1998b, Reid, Adelman, Weiss and Jaret 2005, Chalfin 2013a). Indeed some authors have identified immigration as having contributed importantly to the decline in crime (Butcher and Piehl 1998a; Lee, Martinez and Rosenfeld 2001; Stowell, Messner, McKeever and Raffalovich 2009; Ousey and Kubrin 2009; Martinez, Stowell and Lee 2010; Wadsworth 2010; MacDonald, Hipp and Gill 2012). While the extant literature supports the view that increases in immigration may have had a protective effect on crime, public opinion has generally reached the opposite conclusion, with a majority of U.S. natives indicating a belief that immigration is associated with increases in criminal activity (Espenshade and Calhoun 1993; Muste 2012).[43] Though recent empirical work is consistent with patterns in the aggregate time series, the literature remains unsatisfying in several ways. First, the available literature rarely disaggregates the effects of immigration on crime by nationality. In particular, there is little research that addresses the criminal participation of recent Mexican immigrants.[44] As Mexican immigrants comprise over one third of all immigrants to the United States and over half of all undocumented immigrants, assessing the effect of Mexican immigration on crime would appear to be particularly relevant. Second, prior literature has not been able to disaggregate between crimes committed by immigrants and crimes committed against immigrants. This is particularly concerning both because immigrants may be less likely than natives to report being victimized to the police and because immigrants may be especially attractive crime victims. In general, immigrant underreporting of crime will tend to make an empirical estimate of the effect of immigration on crime conservative with respect to immigrant criminality while the tendency for immigrants to be attractive crime victims will have the opposite effect. To address this issue I develop a theoretical model of immigrant offending that characterizes analytically the conditions under which an empirical estimate of the effect of immigration on crime will yield a conservative estimate of immigrant criminality.[45]

[43] Muste (2012) reviews twenty years of public opinion data on immigration. According to GSS data, in 1996, 32 percent of American natives believed that immigrants increased crime rates. In 2004, 25 percent of Americans indicated such a belief. Gallup polls indicate stronger beliefs with regard to immigrant criminality. In June 2001, 50 percent of respondents indicated a belief that immigration "made the crime problem worse." In June 2007, 58 percent of Americans indicated such a belief.

[44] Spenkuch (2012) and Chalfin (2013a) offer the first analyses that disaggregate the effect of immigration on crime by Mexican nationality. Chalfin(2013a) studies the effect of immigration on crime at the MSA level and finds no consistent effect of Mexican immigration on any type of crime while Spenkuch (2012), using county level data, finds important effects on crimes with a pecuniary motive.

[45] Discussion of reporting bias can be traced back at least as far as the 1931 report on the National Committee on Law Observation and Enforcement, also called the "Wickersham Commission" (Tonry

Finally, estimates of the effect of immigration on crime available in prior literature can only be given a causal interpretation under stringent assumptions regarding the inability of immigrants to adjust the timing and destination of migration in response to conditions in U.S. cities. To the extent that migrants select into U.S. cities on the basis of city-specific characteristics, standard regression estimates will return an inconsistent estimate of the effect of immigration on crime. The vast majority of the prior literature does little to address these concerns (Butcher and Piehl 1998b; Spenkuch 2012; Chalfin 2013a).[46] Because it is generally difficult to leverage credibly exogenous variation in destination-specific immigrant flows, there is promise in searching for a natural experiment. In the spirit of Card's seminal 1990 research on the labor market impacts of the Mariel Boatlift on Miami, in this study, I leverage a natural experiment created by recent legislation in Arizona to estimate the impact of an extremely large and discrete decline in the state's foreign-born (noncitizen) Mexican population. I show that Arizonas foreign-born Mexican population decreased by as much as 20 percent in the wake of the states 2008 implementation of the Legal Arizona Workers Act (LAWA), a broad-based E-Verify law coupled with severe sanctions for noncompliance. By contrast, the law appears to have had no effect on the state's share of other noncitizens or U.S.-born Hispanics. In order to isolate the causal effect of the passage and implementation of LAWA on crime, I employ a synthetic "differences-in-differences" estimator, using a new method of counterfactual estimation proposed by Abadie, Diamond and Hainmuller (2010). To calculate a direct estimate of the effect of Arizona's Mexican immigrant share on its crime rate, I extend the synthetic differences-in-differences framework to construct implied synthetic instrumental variables estimates, using LAWA as an instrument for the Mexican population share. In contrast to previous literature, I find significant and large effects of Mexican immigration on Arizona's crime rate. The estimates are robust to a variety of specification checks including changing the composition of the synthetic comparison group as well as using agency-level and monthly data and are supported by a series of placebo tests that examine the impact of dummy E-Verify laws in states that never received one. However, the results are driven predominantly

1997; MacDonald and Saunders 2012).

[46]Only a handful of papers in the prior literature attempt to explicitly address concerns over the endogeneity of the timing and concentrations of immigrant location decisions. Each of these papers uses an instrumental variables strategy pioneered by Altonji and Card (1991) in which the historic distribution of country-specific immigration among counties, cities or neighborhoods is used to predict the current spatial concentration of immigrants. Because the instrument relies on the presence of immigrant networks, it is known as the "network instrument." Using data from the 1980s, Butcher and Piehl (199b) present estimates using 45 U.S. MSAs and find no evidence of an effect of immigration on crime. On the other hand, Spenkuch (2012), using more recent data at the county level, finds large effects of immigration on property crime, an effect which is even larger for Mexican immigrants. A recent paper by MacDonald, Hipp and Gill (2012) presents results at the neighborhood level using data from 200-2005 in Los Angeles and finds that higher immigrant shares predict a decline in crime. While the network instrument is likely an improvement upon conventional least squares estimates of the effect of immigration on crime, several authors point out that the network instrument is potentially biased in the presence of persistent pull factors that attract or repel immigrants to U.S. destinations (Pugatch and Yang 2011; Chalfin and Levy 2013).

by the fact that LAWA resulted in especially large declines among Mexican migrants who are young and male and, as such, the effects are largely compositional. Indeed, for most crimes, the treatment effect is fully explained by age and gender composition. For motor vehicle theft, I estimate that between one third and 87 percent of the decline in crime that is associated with LAWA can be attributed to compositional changes among Arizona's foreign-born Mexican population. The remainder of this paper is organized as follows. Section II provides a brief review of theoretical linkages between immigration and crime that are found in the extant literature. Section III describes the Arizona Legal Workers Act and its "E-Verify" provisions. Section IV motivates the identification strategy and describes the modeling framework. Section V provides a description of the data employed in the study. Section VI presents results, robustness checks and considers the local average treatment effect of the legislation, Section VII lays out a general model of immigration and crime and Section VIII concludes.

VIII. Theoretical Links Between Immigration and Crime

While the majority of empirical work that examines links between immigration and crime has appeared in the past two decades, interest among U.S. policymakers in the relationship between the two variables goes back at least a century. Early sociological theories of criminal offending generally concluded that immigrants were more likely to participate in crime than natives as a result of economic and social deprivation (Sellin 1938; Shaw and McKay 1969; Reid et. al 2005). However, more recent theoretical work has highlighted the potential for immigrants to contribute to the economic and social development of urban areas in ways that are protective of crime (Portes and Mooney 2002, Ousey and Kubrin 2009). In this section, I briefly summarize theoretical arguments either in favor of a positive or a negative causal relationship between immigration and crime. For a more detailed treatment, I note that theories of immigrant criminality have been ably summarized by Reid, Adelman, Weiss and Jaret (2005), Ousey and Kubrin (2009) and MacDonald and Saunders (2012), among others. The degree to which immigration and crime are related at a macro level is nuanced and depends on the types of migrants that the United States tends to attract as well as contextual factors at work in receiving communities. Economists have tended to focus on selection among migrants (see, for example, Borjas 1999) while criminologists and sociologists have written at length about social forces which inform migrants' experiences in the United States. Generally, immigration can contribute to U.S. crime rates through one of three channels. First, immigrants to the United States may differ from natives according to characteristics that are typically observed by researchers. The most important of these characteristics are age and gender which criminologists have long and convincingly linked to participation in crime. To the extent that differences in criminal propensities among immigrants and natives are explained by observable characteristics, the differences are purely compositional and, as such, the contribution of immigration to U.S. crime rates is more or less mechanical.[47] A second way in which immigration can affect the U.S. crime rate is through selection on characteristics that are typically unobserved by researchers. These characteristics include personality traits such as intelligence, motivation and impulsiveness – traits that have been shown to predict criminal involvement but are difficult to measure in national samples. An alternative but related possibility is that migrants bring with them different tolerances for risk and, as such, respond differently than natives to traditional criminal justice policy levers such as police and prisons. To the extent that migrants differ systematically from natives along unobserved dimensions, differences in criminal involvement will persist even if the demographic composition of immigrants is similar to that of natives.[48] Economic theories of migration have posited that selection of migrants to the United States is a function of differences in the distribution of earnings between the United States and a candidate source country. In particular, economic theory assumes that individuals migrate from a relatively poor country (e.g., Mexico) to a relatively wealthy country (e.g., the United States) in search of higher earnings (Borjas 1999). Since the earnings gap between Mexico and the United States is largest at the lowest portion of the skills distribution, Mexican migration to the United States is predicted to be concentrated among those with less valuable labor market skills. Empirical support for this theory of migration has been mixed. However, even if this theory of migration is empirically valid, it has little to say about selection of migrants along dimensions related to criminal propensities. To the degree that these earnings can be either licit or illicit, economic theory cannot generate obvious predictions about how migrants differ according to their criminal propensities. Moreover, given that migrants are selected according to their expected earnings in the U.S., if a subset of these migrants experience an unexpected lack of viable employment options, it is possible that these individuals may be especially willing to turn to criminal activity to compensate for their poor draw in the distribution of earnings (Chalfin 2013a). On the other hand, if migrants are selected according to their earnings potential in the U.S., to the degree that earnings potential is positively correlated with characteristics that are negatively associated with participation in crime, selection may work in the opposite direction. A third possibility is that, independent of any underlying differences between migrants and natives, contextual variables that shape the experiences of migrants and natives alike may either incentivize or deter crime. Examples of the types of contextual variables that can inform the relationship

[47] An example of such compositional effects can be found in a historical analysis in Moehling and Piehl (2009) who examine the criminality of migrants to the United States in the early 20th century. They find that Italian immigrants were considerably more likely than U.S. natives (and other immigrants) to end up incarcerated in the United States. However, this finding is no longer true when examining age- and gender-specific arrest rates. Italian immigrants were more likely to be involved in crime because they were substantially more likely than other immigrants to be young and male.

[48] Duncan and Trejo (2013) note that while Mexican immigrants have lower levels of education, on average, than U.S. natives, they are nevertheless likely to be drawn from the upper half of the Mexican skill distribution. This fact may have implications for the degree to which Mexican immigrants are negatively selected with regard to criminal participation along unobservables.

between immigration and crime are numerous and suggest that the relationship between the two variables is complex. Theories that suggest a positive association between immigration and crime generally focus on material hardship, social disadvantage, and a lack of social cohesion (Bankston 1998; DeJong and Madamba 2001). With regard to material hardship, migrants may engage in crimes with a pecuniary motive as a means of supplementing their incomes out of necessity born out of a lack of opportunities for legitimate earnings (Freeman 1996). Likewise, the substantially lower wages faced by Mexican immigrants suggests enhanced incentives to participate in crime in order to supplement one's legitimate earnings.[49] A more dynamic version of this story posits that sustained material deprivation may lead individuals to engage in violent crimes as an expression of rage or frustration (Blau and Blau 1982; Agnew 1992). With regard to social disadvantage, researchers have posited that assimilation of immigrants into poorer or more violent destination communities might influence participation in crime, above and beyond the characteristics of the immigrants themselves (Martinez 2002, Martinez Lee and Nielsen 2004).[50] For example, a sustained lack of opportunity for advancement within the legitimate labor market may lead to the creation of immigrant subcultures organized around ethnic gangs (Short 1997; Reid, Weiss, Adelman and Jaret 2005). More fundamentally, to the extent that neighborhood poverty is positively associated with crime, Mexican immigrants, who are, on average, poor, tend to settle in high crime neighborhoods within a central city, these migrants may be exposed to a higher degree of criminality and a greater concentration of anti-social norms (Shaw and McKay 1969, Hagan and Polloni 1999). This is especially true of neighborhoods that have pre-existing ties to illegal drug markets (Ousey and Kubrin 2009). To the extent that this arrangement is associated with greater crime than an arrangement that randomly assigns Mexican immigrants to neighborhoods, assimilation can be said to drive immigrant criminality, above any beyond any effects of selection. Finally, researchers have suggested that neighborhoods that are settled by immigrants, particularly those from Mexico, are unstable and lack social cohesion. In particular, the high degree of population turnover and frequent and rapid social change in immigrant neighborhoods creates an environment that is conducive to sustained criminal activity. The literature has focused primarily on the breakdown of informal social control (Bankston 1998; Lee, Martinez and Rosenfeld 2001; Mears 2002). On the other hand, the tendency of immigrants to settle in ethnic enclaves might have a protective effect on their welfare and, as such, a dampening effect on crime (Logan el. al 2004). For example, immigrants tend to settle in communities and work for businesses that cater to other immigrants from their source country, shielding them from the effects of labor

[49] An economics literature documenting theoretical linkages between the wage and the opportunity cost of crime can be traced back to Becker (1968). Other references include Ehrlich (1973, 1976) and Grogger (1991). Recent surveys of the relationship between wages and crime can be found in Mustard (2010) and Chalfin and Raphael (2011).

[50] Martinez, Lee and Nielsen (2004) refer to this phenomenon as the "Americanization hypothesis."

market discrimination.[51] Likewise, immigrant neighborhoods may be associated with a greater degree of formal social control (Desmond and Kubrin 2010). Finally, the loss of utility that arises from an arrest and subsequent conviction may be greater for undocumented immigrants who make up a substantial portion of the newly arrived Mexican foreign-born population. This is because an arrest leads not only to a criminal sanction but also, in many cases, to deportation. As a result, immigrants may have an especially strong incentive to “fly under the radar. A final consideration worth mentioning concerns responses to immigration rather than the experiences or behavior of migrants themselves. This consideration follows from an empirical literature on the social and economic effects of immigration on the experiences of U.S. natives and adds to the number of mechanisms through which immigration might affect crime by pointing out that immigration can also change the calculus of offending among natives. For example, if immigrants depress the wages or employment opportunities of natives whose crime-wage elasticities are highest, crime might rise in response to immigration even if immigrants are not responsible for the new crimes that are committed (Grogger 1998).[52] Alternatively, immigrants might prove to be attractive crime victims and accordingly they might lower the search costs of potential offenders (Butcher and Piehl 1998). Related to this is the potential for immigration to contribute to ethnic tension that subsequently spills over into crime.

IX. Institutional Setting

In this paper, I leverage a large and discrete change in Arizona’s noncitizen Mexican population following the passage and implementation of the Legal Arizona Workers Act (LAWA) to estimate the contribution of Mexican immigrants to crime. LAWA’s primary provision is a broad-based mandate that employers verify the legal work eligibility of all new hires using a federal database known as “E-Verify.” This section provides a brief description of both E-Verify as well as LAWA and argues that the timing of LAWA’s implementation in Arizona is plausibly exogenous.

A. The Federal “E-Verify” System

Given the inherent difficulty involved in policing a porous U.S.-Mexico border, recent advances in U.S. immigration enforcement have emphasized policies that address undocumented immigration within the countrys interior. Enforcement in the interior has taken two main forms: 1) expanded cooperation between federal and local law enforcement and 2) workplace-centered measures which seek to either incentive or compel employers to deny employment access to undocumented immigrants. Federal sanctions

[51] The extent to which immigrants suffer from labor market discrimination is debatable given that survey research has found that employers report a greater willingness to hire low-skilled immigrants than their low-skilled native counterparts (Beck 1996; Wilson 1996).

[52] A very large literature addresses the labor market impacts of immigration. While the majority of the literature, particularly those studies that study the effect of immigration on local labor markets, find that immigration has little impact on the labor market prospects of U.S. natives, there are important exceptions to these findings — for example, Borjas (2003) and Pugatch and Yang (2011).

on employers who “knowingly hire unauthorized workers date to the 1986 Immigration Reform and Control Act (IRCA). Motivated by the reality that undocumented migration is, in large part, a function of employer demand for unauthorized labor and widely supported by a nontrivial share of the American public, employer-based enforcement has nonetheless proved challenging to implement.[53] To address these problems, the 1996 Illegal Immigration Reform and Immigrant Responsibility Act (IIRIRA) mandated the pilot program that eventually developed into the Internet-based “E-Verify” system in 2004.[54] The E-Verify system works as follows: Under federal law, all U.S. employers are required to fill out an I-9 tax form for all new employees. Using data provided by new hires during the Form I-9 process, employers who elect to use E-Verify will also submit a new hire’s name, date of birth and either a social security number or an alien identification number into the E-Verify system through a secure website. The information provided is then verified against Social Security Administration (SSA) and Department of Homeland Security (DHS) databases. If the data provided by the applicant do not match administrative records, a “tentative non-confirmation result induces an investigation to ascertain the source of discrepancy. If the identification data ultimately cannot be corroborated, a “final non-confirmation is issued.[55] To date, E-Verify has had a 46 percent success rate in identifying undocumented immigrants (Westat 2009). The use of E-Verify has expanded rapidly in recent years rising from 9,300 participating employers in 2006 to 243,000 participating employers as of January 2011 (Rosenblum and Hoyt 2011). Likewise, Rosenblum and Hoy document a dramatic rise in the number of employer queries — from 1.7 million in 2006 to 13.4 million in 2010.[56] While employers in any state may utilize the system for a minimal cost, much of the recent rise in utilization is due to the passage of state laws mandating its use.[57] To date, fifteen states have passed some sort of legislation that mandates the use of E-Verify while eight states have passed an E-Verify law that has broad applicability to a large proportion of the states workforce. [58] However, the first state to pass a

[53]The proliferation of false identity documents renders the Form I-9 process susceptible to fraud. Employers often claim they strive for rigor but fear running afoul of IRCAs anti-discrimination provision.

[54]As Rosenblum and Hoy (2011) note, political support for something like E-Verify can be traced as far as 1982, when the Senate passed an employer sanctions bill that would have created a “national identification card.” Likewise, in 1984, both the House of Representatives and the Senate passed sanctions bills that would have mandated a national “call-in” system which could be used to verify employment eligibility. However, both bills died in committee.

[55]Recently, DHS has also made available such features as the photo-tool that allows employers to prevent fraud by comparing the photograph on the identity card provided against a photo in the database.

[56]Despite a rapid rise in uptake, as of January 2011, fewer than 3 percent of all U.S. employers had signed up with E-Verify.

[57]say what this cost is.

[58]These states include Arizona, a traditional destination for undocumented immigrants in the United States, Utah, and a number of “new destinations in the southeastern United States: Georgia, Alabama, North Carolina, South Carolina, Mississippi, and Tennessee. A number of additional states have passed an “E-Verify” law that covers specific sectors of the state’s economy — generally public employment. Naturally, there are very few undocumented immigrants working in the public sector. Colorado became

broad-based E-Verify law that covers nearly all employers in the state was Arizona.

B. The Legal Arizona Workers Act

The Legal Arizona Workers Act (LAWA) (also sometimes referred to as the "Employer Sanctions Law") was signed into law in July 2007 and took effect on January 1, 2008. LAWA prohibits Arizona employers from knowingly or intentionally hiring an undocumented immigrant after December 31, 2007. LAWA also mandates the use of E-Verify by all employers in Arizona to establish the identity and work eligibility of all new hires. Not only is the law broad-based, it also imposes harsh sanctions on non-compliant employers. The penalty for an employer's first offense is a suspension of business license with the second offense carrying a potential penalty of revocation.[59] As of January 2011, Arizona accounted for just over 7 percent of businesses nationwide that were enrolled in E-Verify (Rosenblum and Hoyt 2011). Within Arizona, 35,988 (25.7 percent) of the state's 140,081 employers had enrolled in the system. The enrollment rate in Arizona is thus over ten times higher than that in California (2.4 percent), Texas (2.6 percent) or New Mexico (2.5 percent), three other states with large undocumented populations. As Bonn, Lofstrom and Raphael (2011) note, recent reports suggest that at least 700,000 new hires made between October 2008 and September 2009 were subject to E-Verify checks in Arizona, equaling roughly 50 percent of all new hires in the state. As such the law has quite plausibly made it considerably more difficult for unauthorized migrants to obtain gainful employment in Arizona than in other U.S. states. To the extent that LAWA decreases the share of Arizona residents who are undocumented, this may occur through two different channels (Bohn, Lofstrom and Raphael 2013). First, undocumented immigrants currently residing in Arizona may choose to leave the state either settle in another U.S. state or return to their country of origin. Second, foreign nationals planning to migrate to Arizona might choose to migrate elsewhere or to remain in their country of origin. While the legislation targets undocumented immigrants, there is also the possibility that the legislation may cause certain U.S. citizens to leave the state as well. This might occur, for instance, in families in which some members were born in the United States while others are undocumented. Section VI of the paper examines, in detail, changes in the demographic composition of Arizona's population in the wake of the passage and implementation of LAWA. In particular, I examine the impact of LAWA on the foreign-born (noncitizen) Mexican population, a population that has been shown to be both largely undocumented as well as the largest contributor to the undocumented population.[60] If LAWA provides a plausible natural experiment for a change in the foreign-born Mexican share of the state's

---

the first state to pass an E-Verify law in 2006.

[59] As Bohn, Lofstrom and Raphael (2013) note, "to date, legal action taken against employers for violating the provision of LAWA has been quite rare. As of April 2010, more than two years after implementation, only three employers have been indicted under the provisions of LAWA, and all of those in a single county (Maricopa)."

[60] As Passel and Cohn (2009) note, between 80 and 90 percent of recently arrived Mexican immigrants are undocumented and Mexican nationals comprise approximately 60 percent of the undocumented population in the U.S.

population, it should be true that Mexican nationals are the only subpopulation of immigrants whose population numbers are affected by the law. I provide evidence in Section VI that this is the case. Before I present results, however, the following section provides context for thinking about the timing of LAWA's passage as being plausibly exogenous.

C. Threats to Internal Validity

Following Bohn, Lofstrom and Rahpael's 2013 study of the effect of LAWA on Arizona's demographic composition, the identification strategy employed in this research relies on the exogeneity of LAWA's timing. In other words, the consistency of treatment effects estimated in Section VI are valid only if the timing of LAWA's passage and subsequent implementation is as good as random. Threats to validity include the possibility that LAWA was passed in response to an increase in crimes committed by immigrants or a factor that is correlated with crime such as the strength of the state's economy or trends in employment conditions. Likewise, estimates in Section VI cannot be interpreted as causal if LAWA's timing coincided with important changes in federal immigration enforcement that differentially affected Arizona. This section considers potential threats to internal validity of the differences-in-differences estimator described in the following section. As Bohn, Loftstrom and Raphael (2013) note, a number of features of Arizona's legislative environment suggest that the passage of LAWA was not a response to recent crime or employment conditions. Indeed, prior to 2007, violent and property crime rates in Arizona had been constant and falling respectively. Likewise, Arizona's unemployment rate had been falling and its employment-to-population ratio had been rising for nearly a decade prior to LAWA's passage. Instead, the legislative debate surrounding LAWA suggests that the law was a reaction to perceived long-term discontent regarding an increasing presence of undocumented immigrants in the state. As evidence for the randomness of the timing, Bohn, Loftstrom and Raphael note that legislative debate over LAWA spanned several legislative sessions and, due to several federal lawsuits challenging the constitutionality of LAWA, there was substantial uncertainty as to when the act would go into effect once passed by the state legislature.[61] Even if the timing of LAWA's passage and implementation was as good as random, estimated treatment effects can only be thought of as causal to the extent that LAWA's passage did not coincide with other changes in crime markets that differentially affected Arizona relative to other U.S. states. The remainder of this section considers specific potential confounders — namely the rollout of the "Great Recession" which differentially affected Arizonas construction-heavy economy and unrelated changes in federal immigration enforcement during the post-treatment period. LAWA was considered and initially implemented during a period of broad economic growth. However, the great recession began to roll out in late 2008 and, to the extent that it differentially affected Arizona's labor markets and thus its crime market, the great recession has the potential to confound differences-in-differences estimates of the effect of LAWA

[61] The key federal lawsuit was dismissed in December of 2007 thus clearing the way for LAWA to take effect on January 1, 2008.

on crime.[62] To address this concern, I control extensively for pre-treatment trends in Arizona's unemployment rate, its employment-to-population ratio and employment shares in construction, wholesale and retail trade, manufacturing, restaurants and other leading industries. Since a synthetic control region is selected for Arizona on the basis of pre-treatment trends in crime as well as a broad range of economic and social covariates, the analysis controls for these potential confounders as long as the synthetic control method finds a "close" match for Arizona. As I show in Section VI, this condition is satisfied. Finally, it is important to consider whether changes in federal immigration policy coincide with the timing of LAWA's implementation. Bohn, Lofstrom and Raphael report that a review of Department of Homeland Security arrest and apprehension data reveals that the proportion of border apprehensions for the Tucson border sector did not change during LAWA's implementation period. Moreover, they note that the Arizona Border Control Initiative which was responsible for an increase in the intensity of border enforcement pre-dated LAWA by several years. A remaining concern is DHS' rollout of its "Secure Communities" program, a federal initiative that induces cooperation between DHS' Office of Immigration and Customs Enforcement (ICE) and local law enforcement agencies. Under Secure Communities, local police agencies are required to send identifying information, including fingerprints, for all arrestees to federal immigration authorities so that arrestees who are illegal aliens can be identified using federal databases. If ICE identifies an arrestee as a potential immigration violator, ICE can require local law enforcement to hold the individual in jail for up to forty-eight hours so that the individual can be transferred to federal custody for the initiation of deportation proceedings (Cox and Miles 2010). While Secure Communities is currently required of all jurisdictions, during the initial rollout, local police agencies were given the choice to voluntarily opt in to the program. Arizona counties are heavily represented amongst those opting into the program with key counties such as Maricopa (activation date: January, 2009) Pima (November, 2009) and Yuma (January, 2009) activating early. As of December 2012, ICE has identified 84,976 alien arrestees in Arizona of whom 16,177 had a prior criminal conviction. Of the 84,976, 3,497 had a prior ICE removal. While relatively few of these individuals have been removed by federal authorities, it is not entirely possible to separate the effect of Secure Communities from that of LAWA in the years after 2008.[63] Therefore, in Section VI, all results are shown using 2008-2010 as the post-treatment period and using 2008 only. Happily, the results are similar whichever post-treatment period is employed.

[62]There is evidence that Arizona was differentially affected by the 2008 recession as Arizona's economy is disproportionately reliant on the construction industry. However, Arizona's decline in construction employment broadly mirrors drops in other states.

[63]In a recent working paper, Chalfin, Loeffler and Treyger (2013) examine the impact of Secure Communities on crimes reported to police and find little evidence of crime declines in response to program roll-out.

X. Empirical Strategy

A. The Standard "Differences-in-Differences" Estimator

The standard approach to computing "differences-in-differences" (D-D) estimates of the effect of a state-level policy shock is to regress a state- and time-varying outcome, Y_{it} on a treatment dummy, D_{it}, a vector of time varying covariates, X_{it} and state and year fixed effects, ψ_s and ϕ_t, respectively:

$$Y_{it} = \alpha + \beta D_{it} + X_{it}'\delta + \psi_s + \phi_t + \varepsilon_{it} \tag{17}$$

In (1), β yields an estimate of the treatment effect of the policy shock.[64] Typically regression estimates are computed using weighted least squares such that the comparison group for the treated state(s) is a population-weighted average of other U.S. states and the confidence interval around β is typically computed by clustering the standard errors at the state level.[65] The identifying assumption under which β represents a causal estimate of the effect of D_{it} is that the treated state(s) and the comparison states experience parallel trends but for the treatment. Naturally the degree to which this assumption holds depends on the appropriateness of using (population-weighted) untreated U.S. states as a control group for (population-weighted) treated state(s). While the identifying assumptions of the D-D estimator are well understood, in practice, researchers rarely provide a direct test of the validity of this assumption.[66]

B. The Synthetic "Differences-in-Differences" Estimator

In order to estimate the effect of the passage and implementation of LAWA on crime at the state level, I employ a new method of counterfactual estimation developed by Abadie, Diamond and Hainmuller in an influential 2010 article published in the *Journal of the American Statistical Association*. The method, which uses a data-driven algorithm to identify a synthetic comparison group from among a pool of potential comparison states, represents the latest advance in the estimation of treatment effects for discrete aggregate-level policy interventions.[67] In the context of a state-level intervention in the United States, the methodology works by assigning an analytic weight to each U.S. state that has not implemented a given policy (e.g., an E-Verify law), where the weights are computed such that the difference in a given pre-intervention outcome (e.g., crime) between a treated state (e.g., Arizona) and its pool of potential comparison states is minimized. In this way, the methodology generates a comparison group which,

---

[64] Absorbing the covariates and fixed effects, this can be seen by considering that $E[Y_{it} | D=0] = \alpha$ and $E[Y_{it} | D=1] = \alpha + \beta$. Thus $E[Y_{it} | D=1] - E[Y_{it} | D=0]$ is β, the D-D estimate of the effect of the treatment.

[65] Bertrand, Duflo and Mullainathan (2003) show that clustering the standard errors is the only reliable means of accounting for arbitrary unit-specific serial correlation.

[66] In principle, (1) can also be estimated without population weights or using some other weighting scheme. Regardless, as long as the choice of weights is arbitrary with respect to the parallel trends assumption, regression-based D-D estimators will potentially suffer from the above problem.

[67] Abadie, Hainmuller and Diamond (2010) apply the methodology to estimate the effect of the passage of Proposition 99, a California ballot proposition designed to reduce consumption of tobacco. An older reference can be found in Abadie and Gardeazabal (2003) who study the effects of terrorism on economic development in Spain's Basque Country.

conditional on pre-treatment observables, meets the assumption of parallel trends prior to implementation of the treatment. The methodology represents an advance on designs that select comparison states based on arbitrary or ad hoc criteria and standard two-way fixed effects D-D estimators which implicitly use a population-weighted or unweighted average of the remainder of the United States as a comparison group. By using a data-driven method to generate an appropriate control group, the estimated treatment effect is robust to a common misspecification problem. Moreover, the method offers a series of placebo tests that ensure that the resulting D-D estimate is not the result of an intervention whose timing is insufficiently random. In this section, I provide a formal treatment of the "synthetic control methodology used to estimate the effect of LAWA on state-level crime rates. Formally, let the index $j = (1,2,,J)$ denote the J states in the United States.[68] The value j=1 corresponds to Arizona, and j=(2,,J) correspond to each of the other states that are candidate contributors to the control group.[69] I begin by defining Y_0 as a kx1 vector with elements equal to the seven annual index crime rates and two crime aggregates (violent and property crimes) for Arizona for the 2000-2007 pre-intervention period. Likewise, I define the kx$J$ matrix Y_1 as a stack of similar vectors for each of the other J states in the donor pool. The synthetic control method identifies a convex combination of the J states in the donor pool that best approximates the pre-intervention data vectors for the treated state. Define the Jx1 weighting vector W=(w_1, w_2, , w_J)' such that:

$$(A1) \qquad \textstyle\sum_{i=1}^{J} w_j = 1$$
$$(A2) \qquad w_j \geq 0 \;\text{ for }\; j=(1,,J)$$

Condition (A1) guarantees that the weights sum to 1 while condition (A2) constrains that the weights are weakly positive. The product Y_1W then gives a weighted average of the pre-intervention vectors for all states in the donor pool (omitting Arizona), with the difference between Arizona and this average given by Y_0 -Y_1W. The synthetic control method selects values for the weighting vector, W, that result in a synthetic comparison group that best approximates the pre-intervention violent crime trend in Arizona. Once the optimal weighting vector $W*$ is computed, both the pre-intervention path as well as the post-intervention values for the dependent variable in "synthetic Arizona can be tabulated by calculating the corresponding weighted average for each year using the donor states with positive weights. The post-intervention values for the synthetic control group serve as the counterfactual outcomes for Arizona. My principal estimate of the impact of LAWA on the crime rate uses the pre- and post-treatment values for both Arizona and its synthetic control group to calculate a simple D-D estimate. Specifically, define

[68] The discussion in this section is drawn, in part, from a 2013 working paper by Chalfin and Raphael entitled "New Evidence on the Deterrence Effect of Harsher Sanctions: Re-examining the Impact of California Proposition 8.

[69] Excluded from the donor pool of the remaining J states are Alabama, Georgia, Mississippi and South Carolina, states that have likewise passed an expansive E-Verify law after 2008.

Y^{AZ}_{PRE} as the average value of the violent crime rate for Arizona for the pre-intervention period 2000 through 2007 and Y^{AZ}_{POST} as the corresponding average for a defined post-treatment period, 2008-2010. Y^{SYNTH}_{PRE} and Y^{SYNTH}_{POST} are the corresponding quantities for Arizona's synthetic control group. Then the synthetic D-D estimate is given as follows:

$$DD = (Y^{AZ}_{POST} - Y^{SYNTH}_{POST}) - (Y^{AZ}_{PRE} - Y^{SYNTH}_{PRE}) \tag{18}$$

To formally test the significance of any observed relative change in Arizonas violent crime rate, I apply a permutation test suggested by Abadie, Hainmuller and Diamond (2010) and implemented by Bonn, Lofstrom and Raphael (2013) to the D-D estimator given in equation (2). Specifically, for each state in the donor pool that *did not* receive the treatment, I re-compute weights to generate a synthetic control group. Next, I re-compute the synthetic D-D estimates under the assumption that the other states, in fact, passed an E-Verify law on the same date as Arizona. Because, the causal effect of the placebo laws must be zero, the distribution of these "placebo" difference-in-difference estimates then provides the equivalent of a sampling distribution for the estimate DD_{AZ} (see Abadie, Diamond and Hainmuller 2010 for a detailed discussion).

C. The Synthetic Instrumental Variables Estimator

The synthetic D-D estimator described in the prior section computes estimates of the average treatment effect of LAWA on state-level crime rates. In this paper, I present evidence that the passage and implementation of LAWA appears to both reduce Arizona's noncitizen Mexican population and its crime rate substantially. While the reduced form effect of LAWA on crime is of interest, the parameter that has been of greater interest in prior literature is the effect of an increase in a state's *Mexican population share* on crime. Using the fact that LAWA induced emigration of noncitizen Mexicans from Arizona, in this section, I show that the synthetic D-D framework advanced by Abadie, Diamond and Hainmuller can be conveniently extended to compute implied IV estimates of the effect of Mexican emigration from Arizona on its crime rate. Recall that the instrumental variables estimator, $\beta^{IV} = (D'M)^{-1}\, D'Y$ where D is the instrument, M is the noncitizen Mexican population share which is potentially endogenous and Y is the crime rate. Re-writing in terms of covariances, we get:

$$\beta^{IV} = \frac{cov(D_i, Y_i)}{cov(D_i, M_i)} \tag{19}$$

Equation (3) provides intuition for how the implied synthetic IV estimator can be constructed. Dividing both the numerator and denominator in (3) by $var(D_i)$ yields the following characterization of the IV estimator:

$$\beta^{IV} = \frac{cov(D_i, Y_i)}{var(D_i)} \Big/ \frac{cov(D_i, M_i)}{var(D_i)} \tag{20}$$

Examining (4), it is straightforward to see that the numerator is the least squares coefficient obtained from a regression of Y_i on D_i and the denominator is the least squares coefficient obtained from a regression of M_i on D_i. The former is simply the reduced form estimate of the effect of LAWA (D_i) on crime (Y_i) while the latter is the first stage estimate of the effect of LAWA on M_i, the Mexican population share. Using the synthetic D-D estimator, these quantities can be written as:

$$DD^{RF} = (Y^{AZ}_{POST} - Y^{SYNTH}_{POST}) - (Y^{AZ}_{PRE} - Y^{SYNTH}_{PRE}) \tag{21}$$

$$DD^{FS} = (M^{AZ}_{POST} - M^{SYNTH}_{POST}) - (M^{AZ}_{PRE} - M^{SYNTH}_{PRE}) \tag{22}$$

Finally, the synthetic IV estimator is constructed as $\frac{DD_{RF}}{DD_{FS}}$, that is by dividing the D-D estimate in (5) by the D-D estimate in (6). When M is measured as the foreign-born Mexican share of each state's population, the IV estimator yields the predicted percentage change in the crime rate arising from a one percentage point increase in the noncitizen Mexican share. One complication is worth noting. In principle, the construction of an IV estimator from first stage and reduced form estimates requires that both equations contain the same control variables. Since the synthetic D-D estimator implicitly assigns different weights to states in both the first stage and the reduced form equation, this condition will not be met. One solution is to re-estimate both the first stage and reduced form equations using the same weights in each equation. However, in practice, this approach is difficult because the quality of the synthetic match depends heavily on matching on lagged values of the dependent variable. A second approach is to simply control for past values of the non-citizen Mexican share and observe if the results differ from the original approach. Such results are reported in Appendix A. The results are consistent with those presented in Section VI.

XI. Data

Data used in this research are drawn from two primary data systems. Crimes reported to police were obtained from the Federal Bureau of Investigation's Uniform Crime Reports (UCR), the standard source of data on crimes at the agency level that is employed in aggregate-level crime research. Since 1934, the UCR has, either directly or through a designated state reporting agency, collected monthly data on index crimes reported to local law enforcement agencies. The index crimes collected consistently since 1960 are: murder (criminal homicide), forcible rape, robbery, aggravated assault (hereafter "assault"), burglary, larceny and motor vehicle theft.[70] The majority of the analyses reported in the paper utilize monthly agency-level data that have been aggregated to the state-year. In an auxiliary analysis, I report results using the higher frequency quarterly and monthly data. Data on the foreign-born noncitizen population come from the American Community Survey (ACS), a one percent sample of U.S. households, collected annually since 2000 by the U.S. Census. The ACS asks respondents whether or not they were born in the United States and, if not, in what country were they born. For each state, I calculate the share of the population in each year that is (1) noncitizen Mexican, (2) noncitizen other than Mexican and (3) U.S.-born Hispanic. I also calculate age- and

[70] The UCR employs an algorithm known as the "hierarchy rule" to determine how crimes involving multiple criminal acts are counted. In order to avoid double counting, the UCR classifies a given criminal transaction according to the most serious statutory violation that is involved. For example, a murder-robbery is classified as a murder.

gender-specific versions of each of these three population shares.[71] Finally, I collect key control variables from the ACS. These variables include measures of a state's native demographic composition — the percentage white, the percentage black, the percentage married and percentage in the following age groups: 0-14, 15-24, 25-39, 40-54 and 55+. In addition, I control for each state's labor force participation rate, its employment-to-population ratio and its unemployment rate as well as each state's industry concentration for the following industries in which immigrants are disproportionately employed: agriculture, construction, manufacturing, restaurants and retail trade. Table 1 presents summary statistics for key independent and dependent variables for Arizona in each year between 2001-2010. Panel A presents means for selected demographic variables — the percentage of residents who are white and black, the percentage married and the percentage in each of five age groups (0-14, 15-24, 25-39, 40-54 and 55+). Panel B presents summary statistics for several measures of the immigrant population: the foreign-born Mexican population share, the foreign-born non-Mexican share, the total immigrant share and the Mexican share among all immigrants. In addition, I provide data on the share of Hispanic U.S. citizens. Panel C presents summary statistics for three key economic variables: the labor force participation rate, the employment-to-population ratio and the unemployment rate. Panel D presents the summary statistics on industry concentration, that is the proportion of state residents employed in each of five industries which employ a large share of immigrants: agriculture, construction, manufacturing, restaurants and retail trade. Finally, Panel E presents data on crimes per 100,000 individuals for each UCR index crime. Examining Table 1, it is clear that key covariates for Arizona appear to be smooth across the 2007-2008 treatment threshold. For example, referring to Panel C, between 2007 and 2008, there is hardly any change in either the labor force participation rate, the employment-to-population ratio or the unemployment rate in Arizona. Likewise, in Panel D, there is no discernable change in the Arizona's industry composition between 2007-2008. Notably however, the financial crisis generated large changes in the strength of Arizona's economy in 2009-2010, increasing the unemployment rate from 5.8 percent to 11.7 percent and decreasing the employment-to-population ratio from 58 percent to 52 percent. Arizona's economic decline is best seen in the share of construction employment which decline from 5 percent in 2008 to 4.1 percent in 2010. Recognizing that the state of Arizona's macroeconomy may have affected both its immigrant share and its crime rate, I control for each measure of the state's local economy and industry concentration in standard differences-in-differences estimates reported later in the paper. However, I note that there is little evidence of any meaningful change in Arizona's social, demographic

[71] A decision that commonly arises in immigration research concerns whether foreign-born citizens should be counted as immigrants or natives. On the one hand, the foreign-born are immigrants whether or not they subsequently obtain citizenship. Likewise, foreign-born citizens are likely a heavily selected subpopulation of the foreign-born. On the other hand, when researchers refer to "natives," they are commonly referring to individuals who have standing as natives in U.S. society. The majority of the literature on the labor market impacts of immigration count foreign-born citizens as natives and, accordingly, I maintain that convention here.

or economic covariates until 2009, making 2008 an ideal post-treatment year.

XII. Results

A. Main Results

I begin a discussion of the results with an analysis of the effect of LAWA on Arizona's foreign-born Mexican population. The raw data in Table 1 suggests that Arizona's foreign-born Mexican population share declined from 6.1 percent in 2007 to 5.5 percent in the immediate aftermath of LAWA and falling further to 4.9 percent by 2010. This remarkable decline in the foreign-born Mexican share of nearly 20 percent over a three year period is unprecedented in recent U.S. history. Importantly, Table 1 indicates no such change in the share of the foreign-born population of nationalities other than Mexican (this population share was 2.5 percent in 2007 and 2.3 percent in 2010). While the overall noncitizen share declined from 8.6 percent in 2007 to 7.3 percent in 2010, this was fueled entirely by a 3 percentage point drop in the share of the foreign-born who are of Mexican origin. Figure 1 provides a more formal analysis of these trends by presenting graphically synthetic differences-in-differences estimates of the effect of LAWA on three key population shares. Panel A considers the foreign-born Mexican share, Panel B considers the foreign-born non-Mexican share and Panel C considers the share of the population that is comprised of Hispanic U.S. citizens. Each of the panels presents two figures. The figure on the lefthand side of the page compares the relevant population share in Arizona (using a solid line) to the same population share in Arizona's synthetic comparison region (using the dashed line). This comparison region is formed using a weighted average of states that did not pass an E-Verify law, where the weights are computed to minimize the gap between the comparison states and Arizona prior to the implementation of the treatment. The figure on the righthand side of the page compares the estimated treatment effect of Arizona's LAWA (using the blue line) to placebo treatment effects in the other states in the donor pool, each of which is plotted using a gray line. In particular, the blue line plots the difference in the relevant population share over time between Arizona and synthetic Arizona while each gray line plots the difference in the relevant population share over time between each state in the donor pool and its synthetic comparison group. Because none of the other states in the donor pool passed an E-Verify law, the distribution of the gray lines is equivalent to the sampling distribution of the estimated treatment effect. To the extent that the post-treatment difference in Arizona is especially large or especially small relative to the untreated states, the estimated difference is unlikely to have occurred due to chance. Panel A considers the effect of LAWA on Arizona's Mexican population share. From 2000-2007, Arizona and synthetic Arizona track each other extremely closely, indicating that the synthetic matching algorithm has performed well. After 2007, Arizona's noncitizen Mexican share falls dramatically relative to its synthetic comparison region. By 2008, the estimated difference was 0.5 percentage points. By 2010, the difference is approximately 1 percentage point (which represents a

17 percent reduction in the share). Accordingly the synthetic differences-in-differences estimates confirm the trend that can be seen in the raw data. The figure on the righthand side of Panel A considers whether the estimated effect is likely to be due to chance. While the difference between Arizona and its synthetic comparison group is in the middle of the distribution of the distribution prior to LAWA, after LAWA, the synthetic D-D estimate for Arizona is larger in magnitude than for any state in the donor pool. This is already true by 2008 and, by 2010, the gap is even larger. The results provide intuition that the drop in Arizona's foreign-born Mexican share is unlikely to have been due to chance. The figures presented in Panel A assume that the drop in the foreign-born Mexican share can be attributed to LAWA. However, this interpretation could reasonably be called into question if the share of groups that should be far less or entirely unaffected by LAWA also change across the treatment threshold. Accordingly, panels B and C present identical figures for the foreign-born non-Mexican share and the share of Hispanic citizens, respectively. Referring to Panel B, there is little evidence in favor of a decline in the noncitizen population that are nationals of a country other than Mexico. If anything, the share of this group initially rises in Arizona relative to the synthetic comparison group. Referring to the placebo figure, the treatment effect on this group is very close to zero and is entirely consistent with sampling variability. Panel C considers the effect of LAWA on the U.S. citizen Hispanic share. Here too there is little evidence of any effect of LAWA. The Hispanic citizen share increases in both Arizona and synthetic Arizona across the treatment threshold with any differences being consistent with sampling variability. Having presented graphically the first stage relationship between LAWA and the change in the foreign-born Mexican population share, I next present a series of graphs that capture the reduced form effect of LAWA on seven different UCR index crimes and two crime aggregates: violent crimes (murder, rape, robbery and aggravated assault) and property crimes (burglary, larceny and motor vehicle theft). These graphs are presented in Figure 2. Each panel in Figure 2 presents synthetic D-D estimates for a given crime type along with the associated placebo test. Panel A presents estimates for the violent crime aggregate. Referring to figure on the left-hand side, we see that Arizona and synthetic Arizona have very similar violent crime rates prior to the introduction of LAWA. After LAWA's passage, violent crime falls by approximately 15 percent in Arizona relative to its synthetic control region. Referring to the placebo test, this difference appears to be larger than the average among the placebo states. However, approximately five other placebo states have larger drops in their violent crimes rates and accordingly it is difficult to conclude that Arizona's reduction is significant. Disaggregating violent crimes by crime type, we see very little evidence of a lasting effect on murder and rape and only a small decline in robberies relative to the synthetic control region. A slightly larger decline is seen for aggravated assault which decreased by approximately 12 percent. There is stronger evidence that property crimes decreased in Arizona in the aftermath of LAWA. Referring to Panel F of Figure 2, we see that property crimes declined by approximately 20 percent after LAWA's passage, an effect which looks particularly large relative to

the sampling distribution formed by the placebo states. Disaggregating by crime type, this result appears to be largely driven by motor vehicle theft which declined by 20-40 percent, depending on whether 2008 or 2010 is used as the relevant post-treatment year. Before turning to a more detailed discussion of the D-D estimates of the effect of LAWA, I pause to briefly consider the composition of Arizona's synthetic comparison group for each of the outcome variables presented in Figures 1 and 2. The weights used to construct the synthetic D-D estimates are presented in Table 2. For each outcome, a variety of states contribute to the comparison group with the District of Columbia, New Mexico and Hawaii receiving positive weights for the greatest number of crimes. Table 3 presents the information in Figures 1 and 2 in tabular form. In particular, the table presents the reduced form (Panel A) and first stage (Panel B) D-D estimates given in equations (5) and (6) along with implied IV estimates (Panel C) which are computed by dividing each reduced form estimate by the corresponding first stage estimate. For each dependent variable, the table computes the mean difference between Arizona and its synthetic control region in the pre-2008 intervention period. Next, for two post-treatment periods (2008 and 2008-2010), I compute the mean post-treatment difference between Arizona and its synthetic control group. Subtracting the mean pre-treatment difference from a given post-treatment difference yields the D-D estimate of the average treatment effect. Finally, the table reports the p-value from the one-tailed test of the null hypothesis that Arizona's D-D estimate is non-negative against the alternative that the D-D estimate is negative. As suggested by Abadie, Diamond and Hainmuller (2010), the p-value is computed by dividing Arizona's rank in the distribution of D-D estimates by the total number of D-D estimates among Arizona and the placebo states. For example, when Arizona's D-D estimate for a given variable is the most negative among all of the states studied, the p-value on that D-D estimate would be 1/47 or 0.021. I begin discussion of Table 3 by considering first stage estimates of the effect of LAWA on the foreign-born Mexican share presented in Panel B. Prior to LAWA, the mean difference between Arizona and its synthetic control region is zero indicating that the matching algorithm in the synthetic control procedure performed exceptionally well. In 2008, Arizona's foreign-born Mexican share was 0.46 percentage points lower than that in its synthetic control region and, by, 2010, its Mexican share was 0.96 percentage points lower. Since this is the largest difference among the sampling distribution, the associated one-tailed p-value is 0.021 (1/46). Next, I consider reduced form estimates of the effect of LAWA on the log crime rate given in Panel A. For each crime type, mean pre-treatment differences are small indicating that the algorithm identified a good match for Arizona. The largest of these differences was 2.3 percent (for the property crime aggregate) with most differences falling well below 1 percent. Two sets of D-D estimates are given — those calculated using 2008 as the post-treatment period and those using 2008-2010 as the post-treatment period. While the latter uses more information, 2009 and 2010 are potentially compromised by DHS' early rollout of its Secure Communities program in a number of densely-populated Arizona counties. Hence 2008 may give a cleaner estimate of the average treatment effect of LAWA. Using 2008 as the post-treatment period, we see that the largest D-D estimates are found for rape (-0.15), motor vehicle theft (-0.15), aggravated assault (-0.15) and murder (-0.07). Estimates for rape, aggravated assault, motor vehicle theft and the property crime aggregate are significant at the 10 percent level. Notably, all of the estimated treatment effects are negative. Referring to the 2008-2010 post-treatment period, once again all of the estimated treatment effects are negative. Largest effects are found for motor vehicle theft (-0.39), aggravated assault (-0.14) and burglary (-0.12). However, only estimates on motor vehicle theft and the property crime aggregate are significant at conventional levels. The D-D estimate for rape falls by more than 50 percent relative to the estimates that use 2008 only as the post-treatment period. Interestingly, the magnitudes of the D-D estimates using the 2008-2010 treatment period are largely similar to those computed only 2008. To the extent that Secure Communities confounds the estimates by resulting in the removal of criminal aliens, the expected bias would go in the negative direction. Therefore, a conservative reading of the evidence suggests that per capita property crimes fell by approximately 12 percent in the aftermath of LAWA, with the effect largely driven by a 15 percent reduction in per capita motor vehicle thefts. Because measuring the effect of LAWA is largely interesting insofar as it yields an estimate of the contribution of Mexican immigration to crime, Panel C computes implied IV estimates of the effect of a change in the foreign-born Mexican share on crime. For each crime type, these estimates are computed by dividing the reduced form estimate in Panel A by the first stage estimate in Panel B. Since both the reduced form and the first stage estimates are negative indicating that LAWA reduced both crime and the foreign-born Mexican share, the implied IV estimates are positive implying that the effect of Mexican immigration on crime is positive. Estimates for all crime types are large and, using 2008 as the post-treatment period, range from 0.04 for larceny to 0.32 for motor vehicle theft. Since only motor vehicle theft and the property crime aggregate are significant for both post-treatment periods, I focus most heavily on these results. The implication is that a one percentage point increase in Arizona's foreign-born Mexican share led to a 23 percent increase in property crimes. These effects are very large, far larger than those that are found in the literature. Accordingly, in the remainder of the paper, I report a series of robustness checks designed to generate further confidence in the research design, I seek to characterize the local average treatment effect of LAWA and, in Section VII, I further consider conditions under which these estimates are likely to yield an accurate portrayal of immigrant criminality in Arizona.

B. Robustness Checks

Before characterizing the estimated treatment effects presented above, I explore several robustness checks designed to test the sensitivity of the synthetic D-D estimates to decisions made during the research process and as an implicit check on the identification strategy. I also present estimates of the effect of LAWA using more granular monthly and quarterly crime data as well as using crime data available at the city rather than the state level. I begin a discussion of robustness checks by re-specifying the synthetic D-D models presented in Table 3 excluding Arizona's border states from the donor pool. These estimates address the potential for bias caused by spillovers from Arizona to its border states. If Arizona's border states receive an increase in Mexican immigration from Arizona as a result of LAWA, the estimated effect of LAWA will be too large. Thus, in the presence of LAWA-induced spillovers, the suitability of border states as a comparison region for Arizona can be called into question. These estimates are presented in Table 4. Since Arizona's border states (California, New Mexico, Colorado, Utah and Nevada) contribute importantly to the donor pool in only a few instances, many of the estimates in Table 4 are equivalent to those in Table 3. Among the estimates that change marginally are those for murder using 2008 as the post-treatment period and robbery using the 2008-2010 post-treatment period. These are both larger in magnitude. For motor vehicle theft, the results are identical for the 2008-2010 post-treatment period and only slightly different for the 2008 post-treatment period. The same is true for the property crime aggregate. The results therefore lend credence to estimates presented in Table 3 and suggest that cross-border spillovers do not have a measurable effect on the estimated treatment effect. Next, I re-specify the D-D models including 2007 as a post-treatment year. The intuition behind this re-specification of the model is that while LAWA was implemented on January 1, 2008, the law was passed in July, 2007 and, as such, there might have been an anticipatory effect of the law. Table 5 accounts for potential anticipatory effects. Here, the results are broadly similar to those estimated using 2008-2010 as the post-treatment period though the implied IV coefficient for property crimes (0.16 using 2007-2008 as the post-treatment period or 0.17 use 2007-2010 as the post-treatment period) is somewhat smaller than estimates that exclude 2007. Next, I leverage monthly crime data available from the FBI to re-compute synthetic D-D estimates that estimate the treatment effect of LAWA with more temporal granularity. This is potentially important as the law takes effect in the first quarter of 2008 and, as such, the treatment effect should potentially be seen immediately. Figure 3 plots synthetic D-D estimates using quarterly crime data. In each panel of the graph, the x-axis plots quarterly crime data where quarters are numbered relative to 2008Q1 which is denoted "0." Overall, the quarterly plots closely resemble those that employ annual data in both trend and magnitude of the D-D effect. For motor vehicle theft and the property crime aggregate, differences between Arizona and its synthetic treatment region occur immediately beginning in the first quarter of 2008 and increase in magnitude over time. Finally, I present regression-based D-D estimates of the effect of LAWA using monthly police agency-level data for U.S. cities with 50,000 or greater population. These are the most granular data that can be used to estimate the effect of LAWA.[72] With 11 Arizona cities meeting the population threshold of 50,000, the treatment effect is estimated using 396 treated city-months (11 cities $\times$ 36 treated months for each city). For each crime type, I regress the log per capita crime rate on

[72]The cost is larger standard errors due to the fact that there is more sampling variability in monthly crimes.

a treatment dummy and a full set of city and month fixed effects.[73] Results are broadly consistent with synthetic control estimates computed using state-by-year variation. Table 6 presents the results of this exercise. In Table 6, Panel A presents the basic D-D results while Panel B reports D-D estimates controlling for two placebo dummies one which captures Arizona cities in 2007 and a second dummy that captures Arizona in 2005-2006 (2-3 years prior to LAWA's implementation). Estimates in Panel B are useful both because they provide a test of anticipatory effects (for 2007) and for the existence of a pre-treatment trend that would tend to invalidate the identification strategy. Referring to Panel A, LAWA is estimated to have reduced monthly assaults in Arizona cities by 10 percent and motor vehicle thefts by 43 percent. These are highly consistent with the state-level D-D estimates which were 14 percent and 39 percent, respectively. Overall, Arizona cities witnessed a decline of 7 percent and 27 percent for violent and property crimes respectively in the aftermath of LAWA. Referring to Panel B, there is little evidence that crime reductions began to accrue in Arizona cities prior to LAWA's implementation. If anything, there is evidence that murder and rape had risen in the years preceding LAWA with reversals in trends after the law's passage. With regard to property crimes, there is some evidence of a pre-LAWA decline. However, even conditioning on this decline, the estimated effect of LAWA remains negative and significant.

C. LATE

This research has uncovered large crime declines in the aftermath of LAWA that survive a battery of robustness checks. Accordingly, the remainder of the paper considers how to characterize these effects. I begin by assessing whether the large effects of LAWA on crime are likely to be driven by a local average treatment effect that is uneven in its impact on the demographic composition of Arizona's Mexican immigrant population. In particular, because participation in crime among any population is so highly concentrated among young males, I test to see whether LAWA was especially likely to induce young males among the foreign-born Mexican population to leave Arizona. Figure 4 provides insight into the local average treatment effect that is induced by LAWA. Each panel of Figure 4 presents synthetic D-D plots for a different age-gender sub-group of the foreign-born Mexican population. As with Figure 1, the left-hand figure in each panel compares Arizona to its synthetic counterpart while the right-hand figure shows where Arizona's D-D estimate falls in the sampling distribution of placebo estimates. Panel A considers the effect of LAWA on the share of Arizona's population that is comprised of foreign-born Mexican male children ($<$ age 14). In turn, Panel B considers the effect of LAWA on Arizona's 15-24 year old foreign-born Mexican population. There is, in fact, evidence that this subgroup leaves the state in large numbers, declining by over one third after LAWA's passage, an effect that is large relative to the sampling distribution. In 2007, just prior to LAWA's implementation, Arizona's foreign-born Mexican 15-24 male population share was 0.66 percent. By 2010, that same share was just 0.36 percent indicating that this subgroup declined by

[73]Monthly population is linearly interpolated using the ACS population measures for each year.

approximately 46 percent compared to an overall decline of 19 percent. Overall, the decline in the 15-24 year old male population accounts for approximately 27 percent of the overall decline in the Mexican population share despite the fact that this subgroup is only approximately 10 percent of the foreign-born Mexican population. Put differently, while young males comprised 11 percent of the foreign-born Mexican population prior to LAWA, they comprised just 7 percent of this population by 2010. Given that the decline is concentrated most heavily in the age-gender subpopulation that is responsible for a disproportionate share of crime, further attention is warranted to sort out the degree to which results are compositional as opposed to behavioral. The post-LAWA decline in the young male population among foreign-born Mexicans can be used to estimate the proportion of the decline in crime that was brought about by LAWA that can be attributed to a change in the demography of Arizona's Mexican immigrant population. I begin by counting the proportion of arrestees who are 15-24 year old males. In 2009, 31.4 percent of the 13.7 million U.S. arrestees for Part I. crimes were males in this age group. Likewise, according to the U.S. Census, 15-24 year old males comprised 7.2 percent of the U.S. population, making individuals in this group 4.3 times more likely to be arrested than other U.S. residents. Using this information, I can back out the expected decline in the crime rate when the 15-24 year old male share of the foreign-born Mexican population declines. Let M_1 be the initial share of young males among the foreign-born population and M_2 be the young male share post-LAWA. Furthermore let X be the crime rate among Mexican immigrants who are not young males and let a be ratio of the share of young male offenders among young males to the share of other offenders among all others. Then the predicted percent decline in the crime rate (ΔV_p) for a decline in the share of young males from M_1 to M_2 can be computed as:

$$\Delta V_p = \frac{(M_2 - M_1)(a-1)X}{M_1 a X + (1 - M_1)X} \tag{23}$$

The denominator of (7) is simply the initial crime rate while the numerator gives the decrease in crime owing to a decline in the young male population from M_1 to M_2. Simplifying (7) yields an expression that no longer contains X:

$$\Delta V_p = \frac{-(M_1 - M_2)(a-1)}{1 + M_1 a - M_1} \tag{24}$$

Setting $a = 4.3$ (the degree to which young males are overrepresented among arrestees) and using $M_1 = 11$ percent and $M_2 = 7$ (the empirical decline in Arizona's young male share among Mexican immigrants) yields a predicted drop in crime of 9.8 percent. This computation can be extended so as to be crime-specific. Such calculations are presented in Table 7 which, for each crime type, computes a along with the predicted decrease in crime (ΔV_p) given that $M_1 = 0.11$ and $M_2 = 0.07$. The predicted crime drop is compared with synthetic D-D estimates (ΔV_a) presented in Panel A of Table 3. Young males are most overrepresented among arrestees for robbery (a=7.5), burglary (6.4), murder (6.1) and motor vehicle theft (6.0). Given the values of a given in the table, a decline in the young male share of the Mexican immigrant population would be predicted to mechanically reduce robbery by 15 percent, burglary by 14 percent and murder and motor vehicle theft by 13 percent each. Comparing these predicted crime declines with those estimated from the data yields important insight. Relative to the prediction, in the aftermath of LAWA, murder, robbery, burglary and larceny implying that, conditional upon age and gender, Mexican immigrants have a protective effect on crime. For motor vehicle theft, between 33 percent and 87 percent of the decline is explained by compositional effects depending on whether the 2008 post-treatment period estimate or the 2008-2010 post-treatment period estimate is preferred. The implications of this exercise are therefore nontrivial as the crime decline that is associated with LAWA is almost entirely compositional.

XIII. A Simple Model of Immigration and Crime

The results presented in Section VI suggest that implementation of LAWA was associated with significant declines in rape (15 percent), aggravated assault (15 percent) and motor vehicle theft (15 percent). Dividing these reduced form estimates by first stage estimates of the effect of LAWA on the noncitizen Mexican population share implies that a one percentage point decline in Arizona's noncitizen Mexican share led to declines in each these crimes of over 30 percent. These effects are qualitatively large even in the presence of important compositional effects. Accordingly, the remainder of this paper further considers how to interpret these results. Aggregate data analyses of the effect of the immigrant share on the rate of reported crimes are compromised by several important concerns that have received mention in but have not been addressed by the prior literature. Two such conditions are of first order importance. First, there will be slippage between the estimated effect of immigration on crime and the actual contribution of immigrant offenders to crime when immigrants report crimes to police at a lower rate than do natives. Under this scenario, when the immigrant share increases, other things equal, the rate of reported crimes will decrease mechanically since immigrants are less likely to report crimes to the police than are natives. Likewise, when the immigrant share decreases (as it did in the aftermath of LAWA), the rate of reported crimes will increase mechanically. As a result, other things equal, the estimated effect of LAWA on reported crime will be biased upwards resulting in bias that is conservative with respect to a negative treatment effect in a regression of the crime rate on the immigrant share. A second concern has to do with the degree to which immigrants are associated with higher levels of crime only insofar as they are attractive crime victims to native offenders (Butcher and Piehl 1998b, Chalfin 2013a). Under this scenario, when the immigrant share increases, other things equal, the crime rate will rise but only because the supply of attractive crime victims has risen. Likewise, when the immigrant share falls, so too does the supply of attractive crime victims. The increase in crime that is attributed to immigrants in a simple empirical model would, under such conditions, in fact, be caused by native offenders. Accordingly, given that the immigrant share falls, the estimated treatment effect will be too large with respect to immigrant criminality, rendering the estimate anti-conservative. The degree to which an empirical assessment of the impact of LAWA yields a conservative estimate of the treatment effect depends on the magnitude of differential crime reporting rates by immigrants versus natives as well as the relative attractiveness of immigrants as crime victims. To date, the literature has not provided a formal treatment of the conditions under which estimated effects will be conservative. To do so, I motivate a simple model of immigrant and native offending and solve analytically for the conditions under which the above is true.

A. Model Setup

Consider a society consisting of N individuals in which the share of Mexican immigrants (hereafter "immigrants") is S_m and the share of natives is S_n. If the two designations are mutually exclusive and collectively exhaustive, the share of natives can be written as $1-(S_m)$. Thus, in this society, the number of immigrants is NS_m and the number of natives is $N(1-S_m)$. Further assume that immigrants and natives are assumed to be either offenders or non-offenders with the proportion of offenders among each group given by C_n and C_m, respectively.[74] While immigrants and natives are allowed to differ in their criminal behavior, I assume that all immigrants and all natives, respectively are identical to one another. In this model, the total supply of offenders is given by $N[(S_mC_m+(1-S_m)C_n]$ and the total supply of potential victims is N. The behavioral dynamics of the model are governed by four conditional probabilities which describe interactions between four types of citizens: immigrant potential victims, immigrant offenders, native potential victims and native offenders. Specifically, I allow the probability that an offender victimizes a given potential victim to depend on the nationality of both the victim and offender. As such, the model allows for scenarios in which crime is largely endogamous (e.g., immigrant-on-immigrant or native-on-native crime) or in which immigrants are especially attractive crime victims for native offenders. Let:

P_{mm}: conditional probability that an immigrant offender victimizes an immigrant potential victim

P_{mn}: conditional probability that an immigrant offender victimizes a native potential victim

P_{nm}: conditional probability that a native offender victimizes an immigrant potential victim

P_{nn}: conditional probability that a native offender victimizes a native potential victim

By conditional probability, I refer to the probability that a victimization occurs conditional upon a meeting between an offender and a potential victim of a given type. For example, if P_{mn} is 0.25, there is a 25 percent chance that a meeting between an immigrant offender and a native potential victim results in a crime against the native victim. For simplicity, these probabilities are assumed to be exogenous in the sense that they are invariant to the immigrant share, the offender share or an other variable in

[74]I assume here that offenders are also potential victims. However, this assumption is not pivotal. The intuition of the model will be no different under the assumption that the offender and non-offender groups are mutually exclusive.

the model.[75] If each of the N individuals is allowed to encounter one other individual, the actual number of victimizations of immigrants and natives will depend on the immigrant population share, the offender share among each group and the conditional victimization probabilities. Let V^a_{mm}, V^a_{mn}, V^a_{nm} and V^a_{nn} denote the actual number of immigrant-on-immigrant (M-M), immigrant-on-native (M-N), native-on-immigrant (N-M) and native-on-native (N-N) victimizations, respectively. Then the actual number of each type of victimization is given by:

$$V^a_{mm} = NS_mC_mS_mP_{mm} \tag{25}$$

$$V^a_{mn} = NS_mC_m(1-S_m)P_{mn} \tag{26}$$

$$V^a_{nm} = N(1-S_m)C_nS_mP_{nm} \tag{27}$$

$$V^a_{nn} = N(1-S_m)C_n(1-S_m)P_{nn} \tag{28}$$

Using (9) as an example, the number of M-M victimizations is equal to the product of the total number of immigrants in the population (NS_m), the probability that an immigrant is a potential offender (C_m), the probability that an encountered potential victim is an immigrant (S_m) and the conditional probability that an immigrant offender chooses to victimize an immigrant potential victim given that this encounter occurs (P_{mm}). Likewise, in (10), the number of M-N victimizations is given by the product of the total number of immigrant offenders (NS_mC_m), the probability of a native victim ($1-S_m$) and the conditional probability of a victimization (P_{mn}). Gathering terms, the total number of crimes committed in this society will be:

$$\begin{aligned} V^a = N\{&[S_m^2C_mP_{mm}] + [S_m(1-S_m)C_mP_{mn}] + \\ &[S_m(1-S_m)C_nP_{nm}] + [(1-S_m)^2C_nP_{nn}]\} \end{aligned} \tag{29}$$

Next, because a key question concerns whether natives are especially likely to victimize immigrants, I take an average of P_{mm}, P_{mn} and P_{nn}, so that there are two conditional probabilities that are of relevance: the conditional probability that a native victimizes an immigrant (P_{nm}) and the conditional probability of any other sort of victimization (P_o). The conditional probability of a victimization type other than N-M is equal to the conditional probability an N-M victimization multiplied by a constant, r which is the relative conditional victimization probability. Here, r is hypothesized to be less than 1 allowing for the possibility that N-M victimizations are more likely than other victimizations. Were $r > 1$, other things equal, an empirical estimate of immigration on crime is conservative. Finally, I introduce the possibility that natives and immigrants report crimes to the police with different probabilities using R_m and R_n as the reporting probabilities for immigrants and natives, respectively. Setting $R_m = k \times R_n$, the constant k yields the relative crime reporting rate of immigrants. Here, k is hypothesized to be less than 1. Similarly, the ratio of native to immigrant offenders $(\frac{C_n}{C_m}) = c$. Using these reporting probabilities, and fixing the population, $N = 1$, the

[75] A violation of this assumption would occur if Mexican immigration were to change P_{nn} by altering natives' opportunities in the legitimate labor market. However, the majority of the extant literature finds little evidence of such an effect.

number of crimes reported to the police (V_r) is given by:

$$V^r = \{crk[S_m^2 C_n P_{nm} R_n] + cr[S_m(1-S_m) C_n P_{nm} R_n] + \\ k[S_m(1-S_m) C_n P_{nm} R_n] + r[(1-S_m)^2 C_n P_{nm} R_n]\} \tag{30}$$

In the next section, I provide a descriptive analysis of the intuition contained in (14).

B. Comparative Statics

Equation (14) is an expression for the per capita reported crime rate in our simplified society. Differentiating this expression with respect to a given variable yields an analytic expression for the effect of a small increase in this variable on the crime rate. The partial derivative of the victimization rate with respect to the immigrant share gives an expression for the contribution of the immigrant population share to crime:

$$\frac{\partial V^r}{\partial S_m} = C_n P_o R_n \times S_m(2ck - c + kr - 2kr) \tag{31}$$

When $\frac{\partial V^r}{\partial S_m} > 0$, an increase in immigration increases per capita crimes that are reported to the police. This first order condition can be recast in order to solve for the relative crime reporting rate of immigrants vis-a-vis natives (k) that is required to render an empirical estimate of the effect of immigration on crime conservative. In particular, by setting $\frac{\partial V^r}{\partial S_m} = 0$ and solving for k^*, we can characterize analytically how high the relative crime reporting rate must be in order for the reported crime rate to fall when the immigrant share increases given that natives potentially victimize immigrants at an especially high rate. This expression is given in (16):

$$k^* = \frac{-(2S_m + c - 2cS_m - 2)}{r + 2cS_m - 2rS_m} \tag{32}$$

Equation (16) presents the "breakeven" relative crime reporting rate as a function of the relative victimization probability (r). In addition to r, this relationship depends on S_m as well as the relative offender share (c).[76] When c = 1 so that the offender shares among immigrants and natives are equal as would be true in a neutral case, (16) reduces to:

$$k^* = \frac{1}{2S_m(1-r) + r} \tag{33}$$

Examining (17) yields some useful insight. First, when r = 1 (meaning that immigrants and natives report crimes to the police with equal probability), k^* = 1. The intuition is that with equal offender shares and equal reporting probabilities, whenever $k^* < 1$, an empirical estimate will be conservative. This is sensible as when r = 1, any tendency for immigrants to report at a lower rate will introduce conservative bias.[77] When $r > 1$, the

[76]As it turns out, C_m and C_n do not affect the breakeven value of the relative reporting rate as long as the two parameters are equal. This can be seen by differentiating k^* with respect to C_m and C_n respectively. The first derivative with respect to C_m equals -1 times the first derivative with respect to C_n. As a result, when both variables are increased by the same amount, k^* remains unchanged.

[77]For $k > 1$, the estimate will be anti-conservative. However, there is very little reason to believe that immigrants report crimes to police at a higher rate than natives.

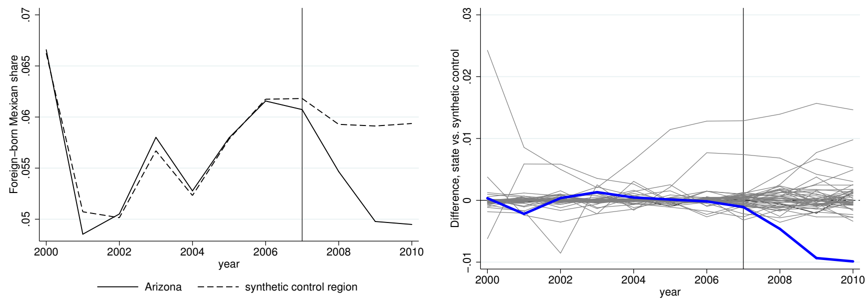
breakeven reporting rate exceeds 1 as long as $S_m > 0.5$. For example, with $r = 2$ (i.e., natives are twice as likely to victimize immigrants relative to all other offender-victim encounters) and $S_m = 0.75$ (i.e., three quarters of the population are immigrants), $k^* = 2$, indicating that as long as the immigrant reporting rate is less than twice the native reporting rate, the empirical estimate will be conservative. Since we typically think of k as being less than 1, an empirical estimate is very likely to be conservative when the immigrant share is large. Alternatively, when $S_m < 0.5$, a value of r that is greater than 1 leads to a k^* that is less than 1. For example, when $r = 2$ and $S_m = 0.25$, $k^* = 0.67$. As such, under this scenario, an empirical estimate will be conservative as long as the immigrant reporting rate is less than 67 percent of the native reporting rate. In order to formally characterize the conditions under which an empirical estimate of the effect of immigration on crime is conservative, I continue to assume that the share of offenders is equal among immigrants and natives. Given equal offender shares, if $\frac{\partial V^r}{\partial S_m} < 0$, then the potentially lower reporting rate among immigrants outweighs the potentially higher victimization rate among immigrants and the estimate is conservative. Under these conditions, the only parameter of importance is S_m. The dynamics of the model are summarized visually in Figure 5. This figure plots the breakeven relative crime reporting rate (k^*) as a function of the relative victimization probability (r) for several different values of S_m. Regardless of S_m, when $r = 1$ (i.e., there is no tendency for immigrants to be especially good victims for native offenders), the breakeven reporting rate is 1. This means that as long as $R_m \leq R_n$, an empirical estimate of $\frac{\partial V^r}{\partial S_m}$ will be conservative. However, when r=2, the breakeven reporting rate falls to approximately 56 percent at $S_m = 0.1$ or 61 percent at $S_m = 0.2$. Here, the estimate is conservative as long as the immigrant crime reporting rate is less than 0.56 (or 0.61) times the native reporting rate. Notably as the immigrant share rises, so too does the likelihood that the estimate is conservative. This is because as the immigrant share rises, the native share and, along with it, the number of native offenders falls. Overall, given that the relative victimization probability is less than 3, an empirical estimate of immigrant offending will be conservative so long as immigrants report crimes at a rate that is less than half that of natives. Empirical estimates of the likelihood that an individual reports a crime to police by immigration status are difficult to come by. However, survey research by Davis and Henderson (2003) suggests an answer. The authors surveyed residents of Jackson Heights, a neighborhood with a high concentration of immigrants in Queens, NY. Whereas 93 percent of African-Americans indicated they would report a break-in to the police, 76 percent and 87 percent of Ecuadorian and Colombian immigrants indicated that they would do so. The numbers are similar for muggings (African-Americans: 88 percent, Ecuadorian immigrants: 70 percent, Colombian immigrants: 78 percent). The numbers suggest that immigrants are about 85 percent as likely as natives to report serious crimes. If so, an empirical estimate would be conservative as long as the relative victimization rate were less than 1.3 (natives are 30 percent more likely to victimize immigrants than any other victimization mix). However, several cautions are warranted in using these estimates. First, the estimates reported by Davis and Henderson are based on a small sample of individuals. Second, New York City is a sanctuary city and accordingly the New York Police Department is not authorized to ask residents about their immigration status. Finally, these numbers are self-reported and are not based on the actual reporting behavior of these individuals.

XIV. Conclusion

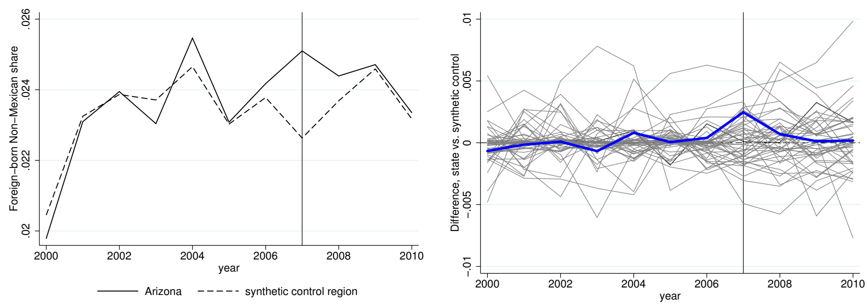
This research leverages a natural experiment in Arizona to estimate the contribution of Mexican immigrants to the state's crime rate. In the aftermath of the passage and implementation of the Legal Arizona Workers Act, there was a large and discrete decline in Arizona's foreign-born Mexican population share relative to other states. On the other hand, the law's passage seems to have had no effect on either the foreign-born non-Mexican share or the share of U.S.-born Hispanics. After 2008, Arizona's crime rate (particularly its property crime rate) declined by approximately 10 percent implying that the decline in the foreign-born Mexican share induced by LAWA resulted in a decline in property crimes of more than 20 percent. These effects are robust to the exclusion of Arizona's border states from the control group, a series of placebo tests, analysis of quarterly as opposed to annual crime data and analysis of crime data at the agency level. The large and significant decline in crime in the aftermath of LAWA is unusual given a literature that has consistently found null or even negative effects of immigration on crime. Further analysis of LAWA's local average treatment effect provides an answer to this conundrum as young males (aged 15-24) were especially likely to leave the state after the law's passage. Given this subpopulation's disproportionate involvement in criminal activity, back-of-the-envelope calculations suggest that between one third and 85 percent of the estimated treatment effect can be accounted for by compositional changes in Arizona's Mexican population along the dimensions of age and gender. As a result, this research remains broadly consistent with prior research that suggests that adjusting for age and gender, immigrants are not more likely than natives to be arrested or incarcerated. Interestingly, results in this paper are similar to those of Moehling and Piehl (2009) who arrived at similar conclusions for Italian immigrants (who were also disproportionately likely to be young and male) at the turn of the 20th century. Finally, this research provides the first attempt to model analytically conditions under which the estimated effect of LAWA on crime will be a conservative estimate of immigrant criminality. This result of this modeling exercise is that an estimate is likely to be conservative when immigrants report crimes to police at a rate that is no higher than 75 percent that of natives and when natives are no more than twice as likely to victimize an immigrant as compared to all other victimizations. While data needed to empirically calibrate this model are sparse, the model suggests that under modest assumptions on the reporting rate and the relative victimization rate, there are not strong reasons to believe that an empirical estimate of the effect of immigration on crime will be very anti-conservative.

FIGURE 1. SYNTHETIC
"DIFFERENCE-IN-DIFFERENCE" ESTIMATES OF THE EFFECT OF THE LEGAL ARIZONA WORKERS ACT [LAWA] ON THE FOREIGN BORN POPULATION SHARE

A. Foreign-Born (Noncitizen) Mexican Population



B. Foreign-Born (Noncitizen) Non-Mexican Population



C. U.S. Citizen Hispanic Population

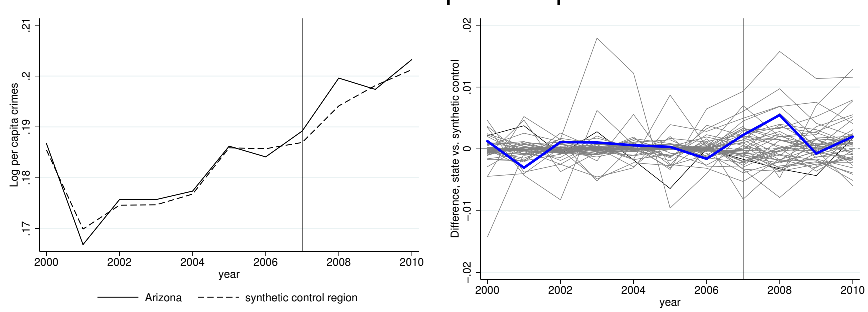
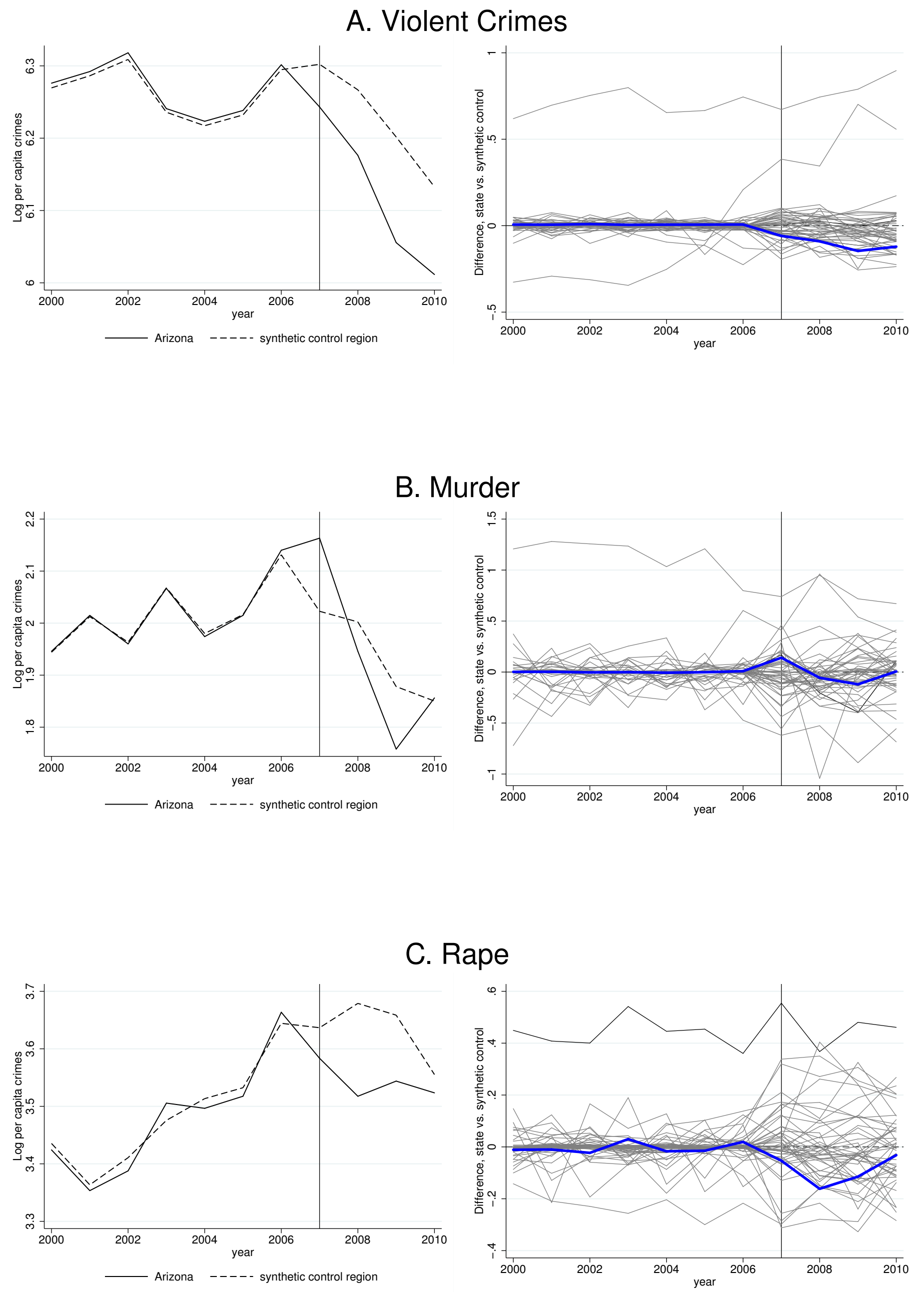
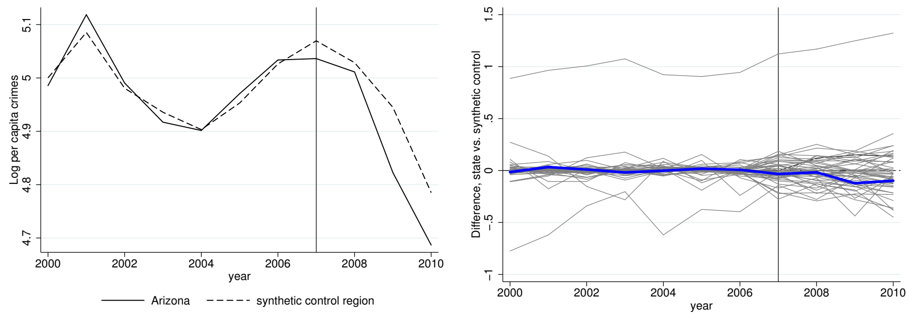


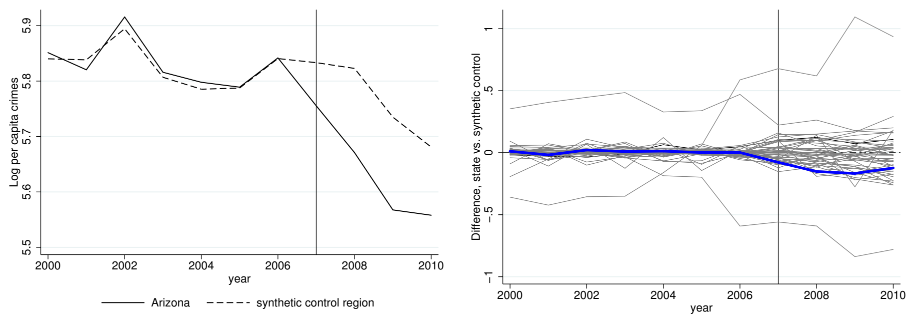
FIGURE 2. SYNTHETIC "DIFFERENCE-IN-DIFFERENCE" ESTIMATES OF THE EFFECT OF LEGAL ARIZONA WORKERS ACT [LAWA] ON LOG PER CAPITA CRIMES REPORTED TO THE POLICE



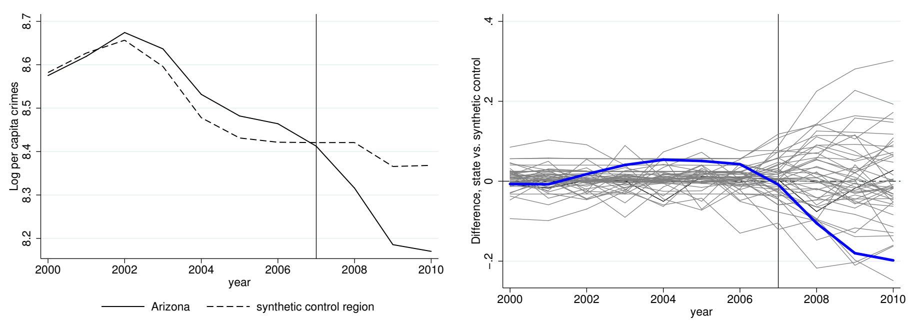
D. Robbery



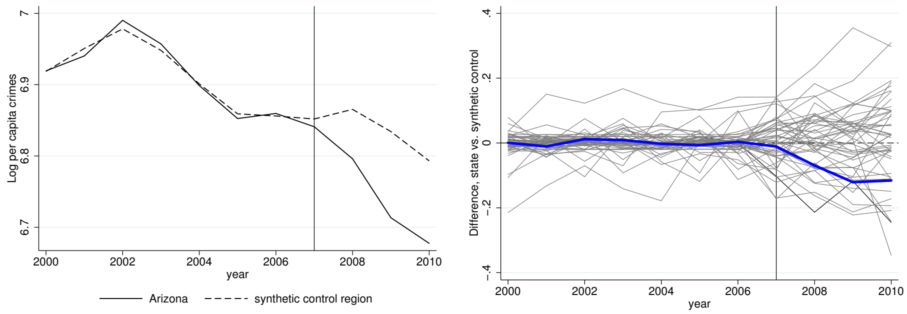
E. Assault



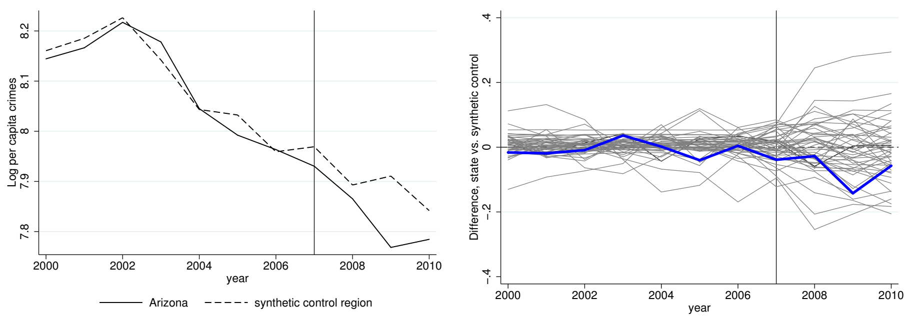
F. Property Crimes



G. Burglary



H. Larceny



I. Motor Vehicle Theft

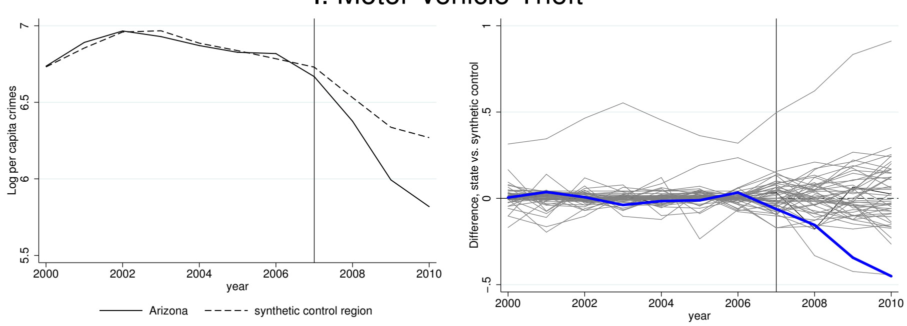
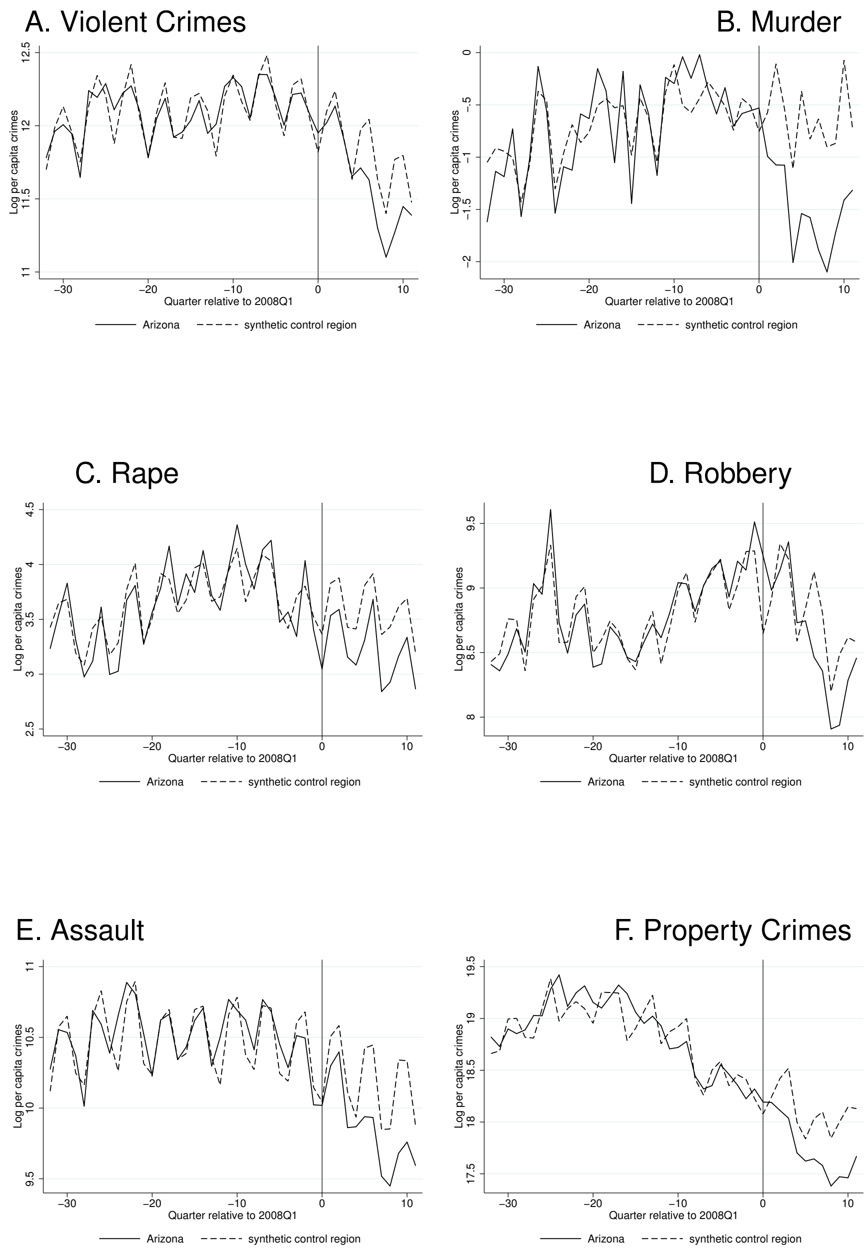
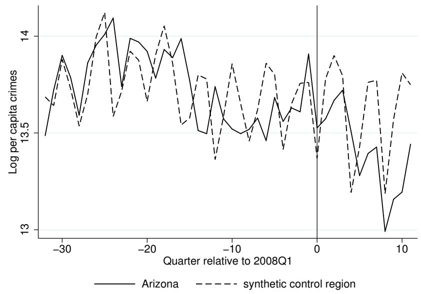


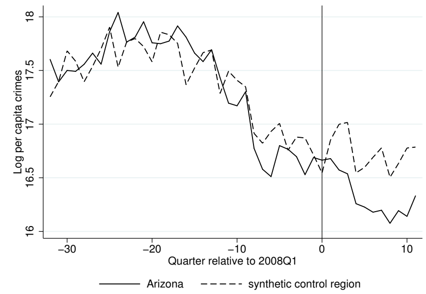
FIGURE 3. SYNTHETIC "DIFFERENCE-IN-DIFFERENCE" ESTIMATES OF THE EFFECT OF LEGAL ARIZONA WORKERS ACT [LAWA] ON LOG PER CAPITA CRIMES REPORTED TO THE POLICE
QUARTERLY DATA



G. Burglary



H. Larceny



I. Motor Vehicle Theft

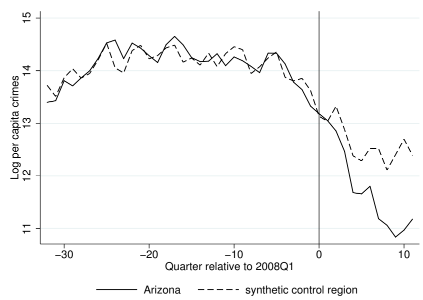
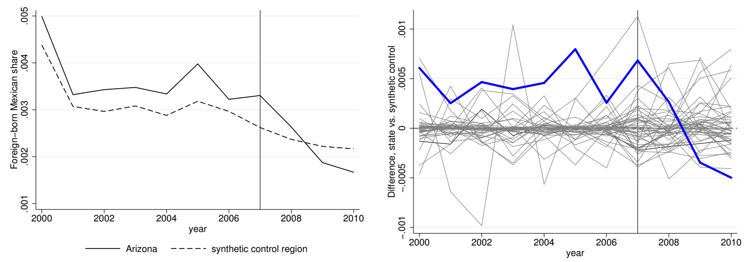
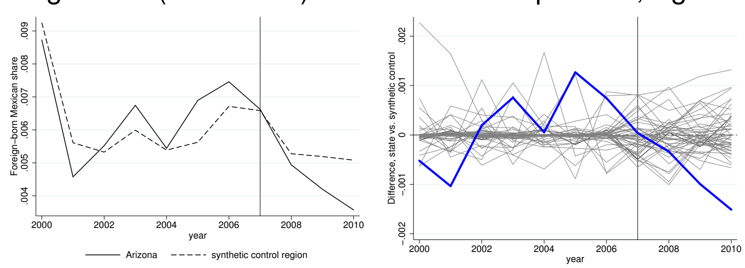


FIGURE 4. SYNTHETIC "DIFFERENCE-IN-DIFFERENCE" ESTIMATES OF THE EFFECT OF LAWA ON THE FOREIGN-BORN MEXICAN POPULATION SHARE BY AGE AND GENDER

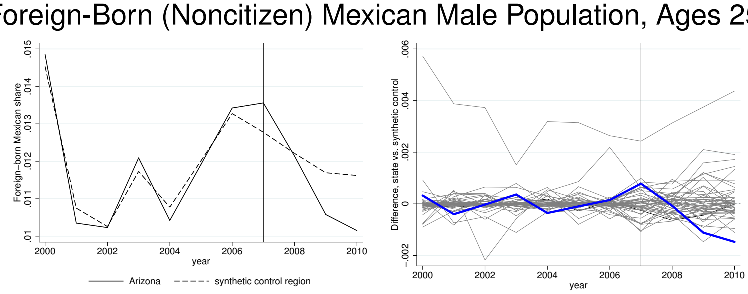
A. Foreign-Born (Noncitizen) Mexican Male Population, Ages 0-14



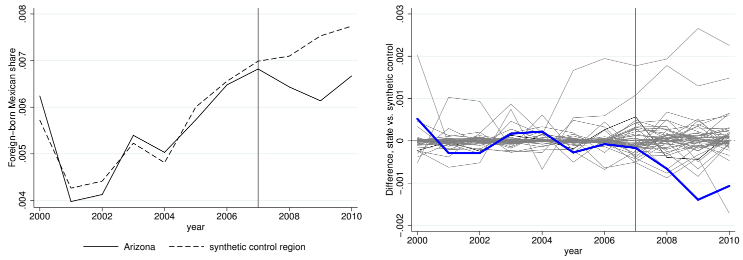
B. Foreign-Born (Noncitizen) Mexican Male Population, Ages 15-24



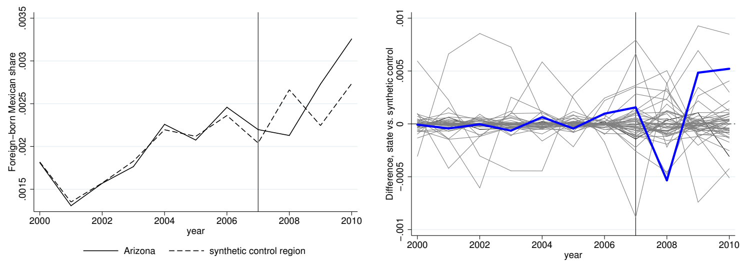
C. Foreign-Born (Noncitizen) Mexican Male Population, Ages 25-39



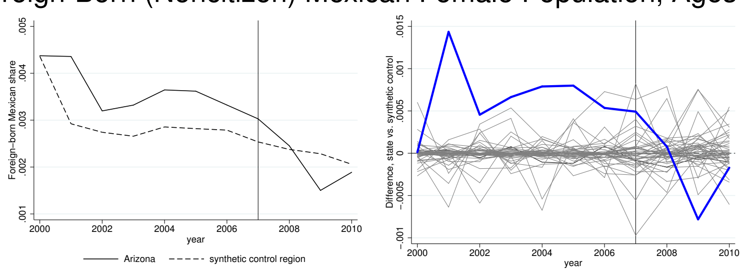
D. Foreign-Born (Noncitizen) Mexican Male Population, Ages 40-54



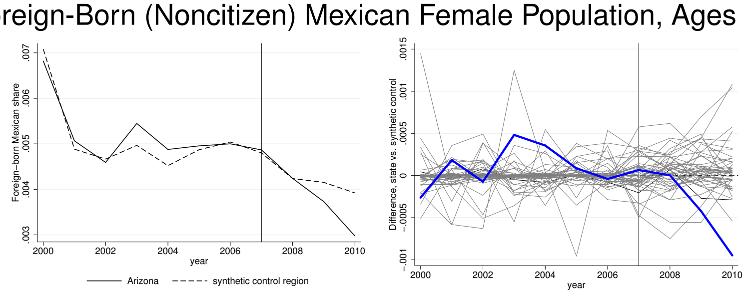
E. Foreign-Born (Noncitizen) Mexican Male Population, Ages 55+



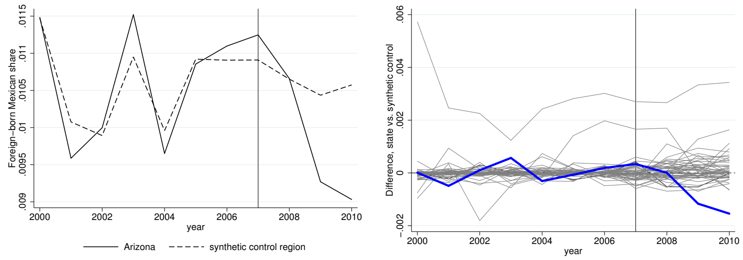
F. Foreign-Born (Noncitizen) Mexican Female Population, Ages 0-14



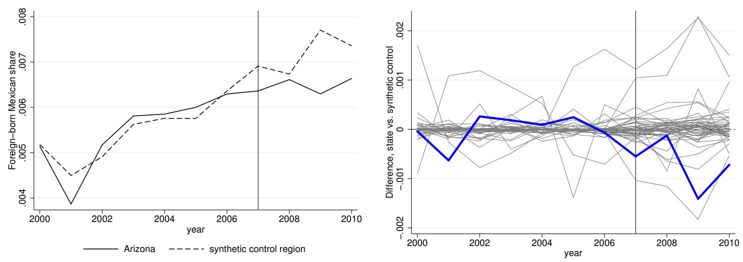
G. Foreign-Born (Noncitizen) Mexican Female Population, Ages 15-24



H. Foreign-Born (Noncitizen) Mexican Female Population, Ages 25-39



I. Foreign-Born (Noncitizen) Mexican Female Population, Ages 40-54



J. Foreign-Born (Noncitizen) Mexican Female Population, Ages 55+

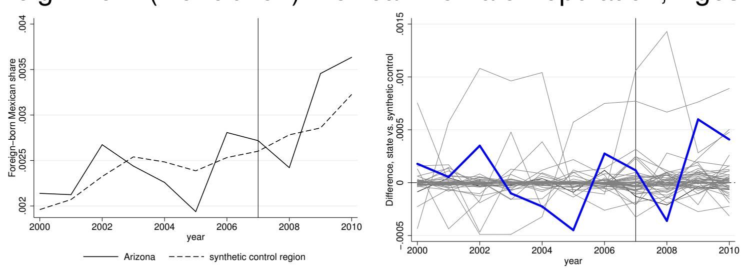
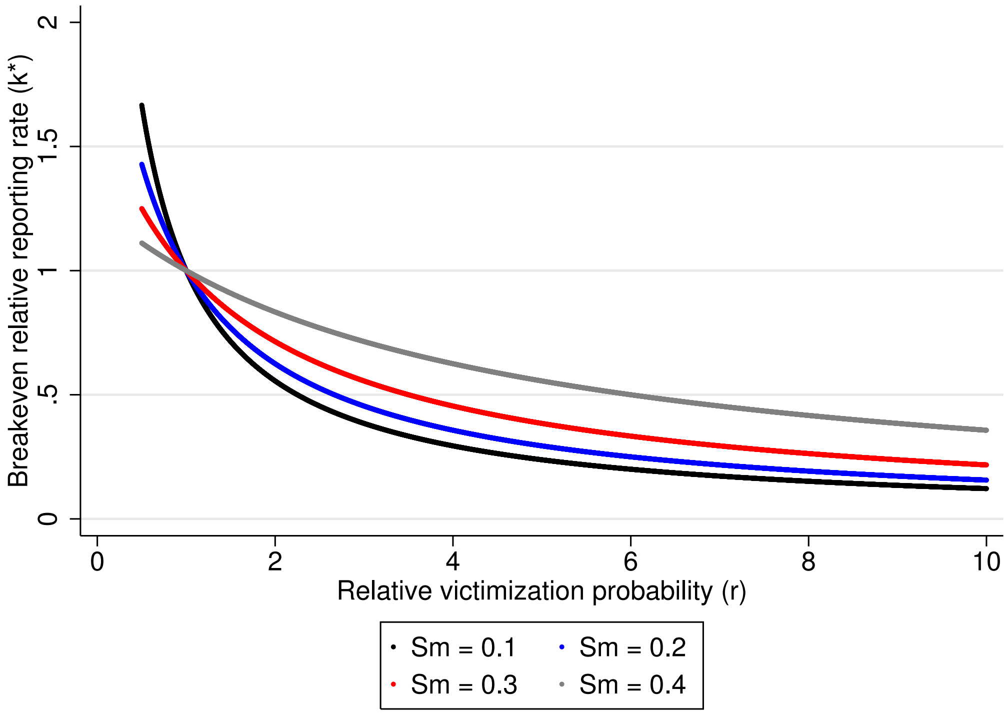


FIGURE 5. MODEL-BASED ESTIMATES OF BIAS IN AN EMPIRICAL ESTIMATE OF MEXICAN IMMIGRATION ON REPORTED CRIMES



Note: The above figure presents model-based estimates of bias in an empirical estimate of the effect of Mexican immigration on the rate of reported crimes. The figures present the "breakeven" relative reporting rate (k^*) as a function of the relative victimization probability (r) for four different values of the Mexican population share: S_m = 0.1, 0.2, 0.3 and 0.4. An empirical estimate of the effect of Mexican immigration on crime will be "conservative" if the crime reporting rate of Mexicans relative to natives is less than r^*. Given that the relative victimization ratios are equal (r=1), the breakeven relative reporting rate = 1 regardless of the Mexican population share. As r increases, the breakeven reporting rate declines. For r = 2 and a small Mexican population share (0.1), r^* = 0.56. Notably, when $r < 1$, k^* > 1 for all values of S_m indicating that when immigrants report crimes with a higher probability than natives, the breakeven relative reporting rate must be larger for immigrants.

TABLE 1. DESCRIPTIVE STATISTICS FOR ARIZONA, 2001-2010

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
DEMOGRAPHIC VARIABLES										
% white	0.696	0.687	0.676	0.673	0.654	0.648	0.625	0.621	0.626	0.612
% black	0.024	0.021	0.023	0.027	0.026	0.027	0.030	0.031	0.033	0.036
% married	0.452	0.465	0.446	0.454	0.443	0.440	0.433	0.432	0.432	0.422
% age 0-14	0.205	0.208	0.214	0.211	0.212	0.205	0.208	0.204	0.196	0.195
% age 15-24	0.123	0.122	0.126	0.115	0.122	0.126	0.127	0.124	0.126	0.125
% age 25-39	0.193	0.193	0.186	0.187	0.186	0.189	0.187	0.187	0.179	0.180
% age 40-54	0.210	0.208	0.212	0.210	0.211	0.206	0.202	0.203	0.202	0.198
% age 55+	0.269	0.270	0.262	0.277	0.269	0.274	0.276	0.282	0.296	0.302
NATIVITY VARIABLES										
% non-citizen Mexican	0.049	0.051	0.058	0.053	0.058	0.062	0.061	0.055	0.050	0.049
% non-citizen non-Mexican	0.023	0.024	0.023	0.025	0.023	0.024	0.025	0.024	0.025	0.023
% non-citizen	0.072	0.074	0.081	0.078	0.081	0.086	0.086	0.079	0.074	0.073
% immigrant	0.121	0.122	0.134	0.130	0.135	0.139	0.142	0.134	0.130	0.130
Mexican share among non-citizens	0.678	0.678	0.716	0.675	0.715	0.721	0.708	0.691	0.668	0.679
ECONOMIC VARIABLES										
% in labor force	0.603	0.600	0.612	0.601	0.610	0.605	0.599	0.616	0.597	0.588
% employed	0.566	0.556	0.569	0.563	0.574	0.576	0.565	0.580	0.538	0.519
% unemployed	0.062	0.074	0.071	0.063	0.059	0.048	0.056	0.058	0.099	0.117
INDUSTRY CONCENTRATION										
% employed in:										
Agriculture	0.013	0.013	0.007	0.006	0.006	0.006	0.006	0.005	0.006	0.006
Construction	0.042	0.041	0.041	0.044	0.049	0.054	0.052	0.050	0.043	0.041
Manufacturing	0.057	0.055	0.051	0.051	0.048	0.045	0.046	0.043	0.043	0.040
Restaurants	0.028	0.033	0.029	0.028	0.030	0.030	0.032	0.031	0.032	0.033
Retail trade	0.074	0.071	0.075	0.073	0.071	0.070	0.068	0.072	0.069	0.070

Note: Table reports means for selected characteristics of the Arizona population from 2001 to 2010.

TABLE 2. COMPOSITION OF SYNTHETIC COMPARISON GROUP BY INDEX CRIME TYPE

State	Violent crimes	Murder	Rape	Robbery	Assault	Property crimes	Burglary	Larceny	Motor vehicle theft	Foreign-born Mexican share
2 AK	-	-	-	0.014	-	-	-	-	-	-
1 AL	-	-	-	-	-	-	-	-	-	-
5 AR	-	0.172	0.182	-	-	-	-	-	-	0.314
6 CA	-	-	-	-	-	-	-	-	-	-
8 CO	-	-	-	-	-	-	-	-	-	-
9 CT	-	0.188	-	-	-	-	-	-	-	-
11 DC	0.046	0.154	-	0.077	0.127	0.693	0.077	0.190	0.535	-
10 DE	-	0.090	-	-	0.102	-	-	-	-	-
12 FL	-	-	-	-	-	-	-	-	0.175	-
13 GA	-	-	-	-	-	-	-	-	-	-
15 HI	-	-	-	-	-	0.034	0.251	0.522	-	-
19 IA	-	-	0.202	-	-	-	-	-	-	-
16	-	-	-	-	-	-	-	-	-	-
17 IL	-	-	-	-	-	-	-	-	-	-
18 IN	-	-	-	0.335	-	-	-	-	-	-
20 KS	-	-	0.255	-	-	-	-	-	-	-
21 KY	0.178	-	-	-	0.239	-	-	-	-	-
22 LA	-	0.068	-	-	-	-	0.071	-	-	-
25 MA	-	-	-	-	-	-	-	-	-	-
24 MD	0.216	-	-	-	0.294	-	-	-	-	-
23 ME	-	-	-	-	-	-	-	-	-	-
26 MI	-	-	-	-	-	-	-	-	-	-
27 MN	-	-	-	-	-	-	-	-	-	-
29 MO	0.129	-	-	-	-	-	-	-	-	-
28 MS	-	-	-	-	-	-	-	-	-	-
30 MT	-	0.035	0.094	-	-	-	-	-	-	-
37 NC	-	-	-	-	-	-	0.295	-	-	-
38 ND	-	-	0.177	-	-	-	-	-	-	-
31 NE	-	-	-	-	-	-	-	-	-	-
33 NH	-	-	-	-	-	-	-	-	-	0.141
34 NJ	-	-	-	-	-	-	-	-	-	0.168
35 NM	0.128	0.293	-	0.210	-	-	-	-	-	-
32 NV	0.216	-	-	-	0.229	-	-	-	-	-
36 NY	-	-	-	-	-	-	-	-	-	-
39 OH	-	-	-	-	-	-	-	-	-	-
40 OK	-	-	-	-	-	-	-	-	-	-
41 OR	-	-	-	-	-	-	-	0.044	-	-
42 PA	-	-	-	-	-	-	-	-	-	-
44 RI	-	-	0.091	-	-	-	-	-	-	0.377
45 SC	-	-	-	-	-	-	-	-	-	-
46 SD	-	-	-	-	0.004	-	-	-	-	-
47 TN	-	-	-	0.363	-	-	0.306	-	-	-
48 TX	-	-	-	-	-	-	-	0.219	-	-
49 UT	-	-	-	-	0.004	-	-	0.026	-	-
51 VA	-	-	-	-	-	-	-	-	-	-
50 VT	-	-	-	-	-	-	-	-	-	-
53 WA	-	-	-	-	-	0.326	-	-	0.290	-
55 WI	-	-	-	-	-	-	-	-	-	-
54 WV	-	-	-	-	-	-	-	-	-	-
56 WY	0.086	-	-	-	-	-	-	-	-	-
$\sqrt{MSE}$	0.0012	0.0044	0.0186	0.0168	0.0115	0.0307	0.0076	0.0221	0.0250	0.0010

Note: Each row reports the percentage contribution of a given state to the synthetic control region for Arizona for a given crime type. For example, for the murder specification, the synthetic control region for Arizona is comprised of: Alaska (0.3%), Colorado (7.4%), District of Columbia (11.6%), Florida (14.1%), Hawaii (24.0%), Nevada (16.0%) and Washington (26.7%). Save for rounding errors, the percentages sum to 100 along the columns.

Table 3. Synthetic "Differences-in-Differences" and IV Estimates of the Effect of LAWA on Index Crimes Reported to Police

	Pre-treatment period, 2001-2007	Post treatment period, 2008			Post-treatment period, 2009-2010		
	Mean difference relative to synthetic control group	Difference relative to synthetic control group	Difference-in-Difference estimate	Implied p-value*	Mean difference relative to synthetic control group	Difference-in-Difference estimate	Implied p-value*
A. Reduced Form Estimates							
Violent crimes	-0.002	-0.091	-0.089	0.106	-0.134	-0.132	0.128
Murder	0.018	-0.056	-0.074	0.489	-0.057	-0.075	0.319
Rape	-0.010	-0.161	-0.151*	0.064	-0.073	-0.063	0.362
Robbery	-0.000	-0.017	-0.017	0.489	-0.110	-0.110	0.340
Assault	-0.005	-0.152	-0.147*	0.085	-0.145	-0.140	0.255
Property crimes	0.023	-0.105	-0.128*	0.064	-0.189	-0.212**	0.042
Burglary	-0.001	-0.069	-0.068	0.170	-0.118	-0.117	0.170
Larceny	-0.010	-0.028	-0.018	0.447	-0.100	-0.090	0.149
Motor vehicle theft	-0.006	-0.154	-0.148*	0.085	-0.397	-0.391**	0.043
B. First Stage Estimates							
Non-citizen foreign born Mexican population share	-0.000	-0.462	-0.462	0.022	-0.961	-0.961	0.022
C. Implied IV Estimates							
Violent crimes			0.193			0.137	
Murder			0.160			0.078	
Rape			0.327			0.066	
Robbery			0.037			0.114	
Assault			0.318			0.146	
Property crimes			0.277			0.221	
Burglary			0.147			0.122	
Larceny			0.039			0.094	
Motor vehicle theft			0.320			0.409	

Note: Panel A presents synthetic"difference-in-differences" estimates of the treatment effect of the Legal Arizona Workers Act (LAWA) on seven UCR index crimes (murder, rape, robbery, aggravated assault, burglary, larceny and motor vehicle theft) and two crime aggregates (violent crimes and property crimes). These are referred to as "reduced form" estimates because they estimate the impact of the law, rather than the responsiveness of crime to Mexican immigration directy. The first column calculates the average difference between Arizona and its synthetic control region prior to the law's passage. The next set of columns present the post-treatment difference between the treatment and control group and the differences-in-differences estimate using 2008 as the post-treatment period. The implied p-value is computed by dividing Arizona's rank among the 46 states in the donor pool by 46. This is shown in Abadie, Diamond and Hainmuller (2010) to be equivalent to a p-value arising from a one-taled test of the null hypothesis. The final set of columns report identical estimates using 2009-2010 as the post-treatment period. As the dependent variable is the log of the crime rate, these coefficients cna be interpreted as semi-elasticities (e.g., LAWA is associated with a β percent change in the crime rate). Panel B presents "differences-in-differences" estimates of the effect of LAWA on the non-citizen Mexican share of the population. These "first stage" estimates estimate the effet of the treament on the endogenous regressor, Mexican population. LAWA is associated with a 0.46-0.96 percentage point decline in the Mexican population. In Panel C, the reduced form estimates are divided by the first stage estimates to yield implied instrumental variables estimates of the effect of Mexican immigration on crimes reported to the police. The coefficient estimates in Panel C provide estimates of the percent change in the crime rate arising from a one percentage point increase in the state's Mexican population share.

TABLE 4. SYNTHETIC "DIFFERENCES-IN-DIFFERENCES" AND IV ESTIMATES
OF THE EFFECT OF LAWA ON INDEX CRIMES REPORTED TO POLICE
BORDER STATES EXCLUDED FROM THE DONOR POOL

	Pre-treatment period, 2001-2007	Post treatment period, 2008			Post-treatment period, 2009-2010		
	Mean difference relative to synthetic control group	Difference relative to synthetic control group	Difference-in-Difference estimate	Implied p-value*	Mean difference relative to synthetic control group	Difference-in-Difference estimate	Implied p-value*
REDUCED FORM ESTIMATES							
Violent crimes	-0.003	-0.081	-0.078	0.140	-0.141	-0.138	0.140
Murder	0.016	-0.201	-0.217	0.209	-0.130	-0.146	0.209
Rape	-0.010	-0.161	-0.151*	0.070	-0.073	-0.063	0.349
Robbery	-0.001	-0.068	-0.067	0.326	-0.239	-0.239*	0.093
Assault	-0.009	-0.107	-0.099	0.186	-0.133	-0.125	0.209
Property crimes	0.014	-0.114	-0.128*	0.070	-0.198	-0.212**	0.023
Burglary	-0.001	-0.069	-0.068	0.186	-0.118	-0.117	0.140
Larceny	-0.002	-0.020	-0.018	0.349	-0.092	-0.090	0.140
Motor vehicle theft	-0.006	-0.154	-0.148	0.116	-0.397	-0.391**	0.047
FIRST STAGE ESTIMATES							
Non-citizen foreign born Mexican population share	-0.000	-0.462	-0.462	0.022	-0.961	-0.961	0.022
IMPLIED IV ESTIMATES							
Violent crimes			0.193			0.137	
Murder			0.160			0.078	
Rape			0.327			0.066	
Robbery			0.037			0.114	
Assault			0.318			0.146	
Property crimes			0.277			0.221	
Burglary			0.147			0.122	
Larceny			0.039			0.094	
Motor vehicle theft			0.320			0.409	

Note: Panel A presents synthetic"difference-in-differences" estimates of the treatment effect of the Legal Arizona Workers Act (LAWA) on seven UCR index crimes (murder, rape, robbery, aggravated assault, burglary, larceny and motor vehicle theft) and two crime aggregates (violent crimes and property crimes). Results in this table ae based on a donor pool that excludes California, Colorado, Nevada and Mew Mexico, states that border Arizona. These are referred to as "reduced form" estimates because they estimate the impact of the law, rather than the responsiveness of crime to Mexican immigration directy. The first column calculates the average difference between Arizona and its synthetic control region prior to the law's passage. The next set of columns present the post-treatment difference between the treatment and control group and the differences-in-differences estimate using 2008 as the post-treatment period. The implied p-value is computed by dividing Arizona's rank among the 47 states in the donor pool by 47. This is shown in Abadie, Diamond and Hainmuller (2010) to be equivalent to a p-value arising from a one-taled test of the null hypothesis. The final set of columns report identical estimates using 2009-2010 as the post-treatment period. As the dependent variable is the log of the crime rate, these coefficients cna be interpreted as semi-elasticities (e.g., LAWA is associated with a β percent change in the crime rate). Panel B presents "differences-in-differences" estimates of the effect of LAWA on the non-citizen Mexican share of the population. These "first stage" estimates estimate the effet of the treament on the endogenous regressor, Mexican population. LAWA is associated with a 0.46-0.96 percentage point decline in the Mexican population. In Panel C, the reduced form estimates are divided by the first stage estimates to yield implied instrumental variables estimates of the effect of Mexican immigration on crimes reported to the police. The coefficient estimates in Panel C provide estimates of the percent change in the crime rate arising from a one percentage point increase in the state's Mexican population share.

TABLE 5. SYNTHETIC "DIFFERENCES-IN-DIFFERENCES" AND IV ESTIMATES OF THE EFFECT OF LAWA ON INDEX CRIMES REPORTED TO POLICE USING 2007 AS A POST-TREATMENT YEAR

	Pre-treatment period, 2001-2006	Post treatment period, 2007-2008			Post-treatment period, 2007-2010		
	Mean difference relative to synthetic control group	Difference relative to synthetic control group	Difference-in-Difference estimate	Implied p-value*	Mean difference relative to synthetic control group	Difference-in-Difference estimate	Implied p-value*
REDUCED FORM ESTIMATES							
Violent crimes	0.008	-0.057	-0.065	0.395	-0.100	-0.108	0.279
Murder	0.046	0.071	0.026	0.744	-0.003	-0.049	0.465
Rape	0.032	-0.030	-0.062	0.488	0.011	-0.021	0.628
Robbery	0.002	-0.020	-0.022	0.558	-0.121	-0.123	0.349
Assault	-0.015	-0.174	-0.159	0.186	-0.171	-0.156	0.209
Property crimes	0.021	-0.053	-0.074	0.163	-0.144	-0.165	0.023
Burglary	0.001	-0.055	-0.056	0.256	-0.091	-0.092	0.209
Larceny	0.015	0.025	0.010	0.558	-0.039	-0.054	0.329
Motor vehicle theft	0.002	-0.154	-0.156	0.186	-0.389	-0.391	0.047
FIRST STAGE ESTIMATES							
Non-citizen foreign born Mexican population share	-0.009	-0.463	-0.454	0.023	-0.956	-0.947	0.023
IMPLIED IV ESTIMATES							
Violent crimes			0.143			0.114	
Murder			-0.057			0.052	
Rape			0.137			0.022	
Robbery			0.049			0.130	
Assault			0.350			0.165	
Property crimes			0.163			0.174	
Burglary			0.123			0.097	
Larceny			-0.022			0.057	
Motor vehicle theft			0.339			0.413	

Note: Panel A presents synthetic"difference-in-differences" estimates of the treatment effect of the Legal Arizona Workers Act (LAWA) on seven UCR index crimes (murder, rape, robbery, aggravated assault, burglary, larceny and motor vehicle theft) and two crime aggregates (violent crimes and property crimes). Results in this table include 2007 as a post-treatment year. These are referred to as "reduced form" estimates because they estimate the impact of the law, rather than the responsiveness of crime to Mexican immigration directy. The first column calculates the average difference between Arizona and its synthetic control region prior to the law's passage. The next set of columns present the post-treatment difference between the treatment and control group and the differences-in-differences estimate using 2008 as the post-treatment period. The implied p-value is computed by dividing Arizona's rank among the 47 states in the donor pool by 47. This is shown in Abadie, Diamond and Hainmuller (2010) to be equivalent to a p-value arising from a one-taled test of the null hypothesis. The final set of columns report identical estimates using 2009-2010 as the post-treatment period. As the dependent variable is the log of the crime rate, these coefficients cna be interpreted as semi-elasticities (e.g., LAWA is associated with a β percent change in the crime rate). Panel B presents "differences-in-differences" estimates of the effect of LAWA on the non-citizen Mexican share of the population. These "first stage" estimates estimate the effet of the treament on the endogenous regressor, Mexican population. LAWA is associated with a 0.46-0.96 percentage point decline in the Mexican population. In Panel C, the reduced form estimates are divided by the first stage estimates to yield implied instrumental variables estimates of the effect of Mexican immigration on crimes reported to the police. The coefficient estimates in Panel C provide estimates of the percent change in the crime rate arising from a one percentage point increase in the state's Mexican population share.

TABLE 6. "DIFFERENCES-IN-DIFFERENCES ESTIMATES:
AGENCY-LEVEL REGRESSION ESTIMATES USING MONTHLY DATA

	Violent crimes	Murder	Rape	Robbery	Assault	Property crimes	Burglary	Larceny	Motor vehicle theft
PANEL A. STANDARD ESTIMATES									
AZ × 2008-2010	-0.070**	-0.133	-0.057	0.004	-0.099**	-0.266**	-0.102*	-0.056	-0.432***
	(0.033)	(0.120)	(0.108)	(0.053)	(0.043)	(0.125)	(0.058)	(0.055)	(0.070)
R^2	0.880	0.910	0.902	0.868	0.807	0.873	0.848	0.873	0.863
PANEL B. PLACEBO ESTIMATES									
AZ × 2005-2006	0.041	0.086***	0.156**	0.044	0.046	-0.188	-0.106	0.001	0.008
	(0.050)	(0.025)	(0.074)	(0.037)	(0.075)	(0.167)	(0.051)	(0.025)	(0.040)
AZ × 2007	0.021	0.028	0.079	0.105	-0.009	-0.193	-0.066	0.021	-0.080
	(0.057)	(0.027)	(0.071)	(0.052)	(0.066)	(0.177)	(0.100)	(0.027)	(0.062)
AZ × 2008-2010	-0.061	-0.117	-0.024	0.025	-0.094	-0.320*	-0.126*	-0.053	-0.442***
	(0.045)	(0.120)	(0.123)	(0.062)	(0.058)	(0.173)	(0.076)	(0.058)	(0.082)
R^2	0.880	0.910	0.902	0.869	0.807	0.874	0.848	0.873	0.863
N	56,842	56,887	56,632	55,116	56,545	56,875	56,856	56,788	56,756

Note: For each crime type, Panel A reports coefficients and standard errors from a regression of the log crime rate on a state-level treatment dummy using monthly agency-level data. Panel B reports coefficients and standard errors on the estimated treatment effect (AZ × 2008-2010) as well as on two placebo dummies which measure the impact of the law on agency-months that are untreated. AZ × 2007 captures the effect of the treatment in the year of its passage but prior to its implementation. AZ × 2005-2006 captures the effect of the treatment in the two years prior to its passage. All models condition on agency and month fixed effects. Standard errors are clustered at the agency level.

TABLE 7. MODEL-BASED ESTIMATES OF THE IMPORTANCE OF LATE

	Violent crimes	Murder	Rape	Robbery	Assault	Property crimes	Burglary	Larceny	Motor vehicle theft
PANEL A. ARRESTEES									
Males, ages 15-24	195,907	5,445	8,099	68,841	113,522	525,083	139,422	350,260	35,401
All arrestees	581,765	12,418	21,407	126,725	421,215	1,716,081	299,351	1,334,933	81,797
Young adult male share	0.337	0.438	0.378	0.543	0.270	0.306	0.466	0.262	0.433
a	4.7	6.1	5.2	7.5	3.7	4.2	6.4	3.6	6.0
PANEL B. MODEL-BASED PREDICTIONS									
ΔV_p	-0.104	-0.130	-0.116	-0.152	-0.084	-0.095	-0.136	-0.082	-0.129
ΔV_a	-0.089	-0.074	-0.151	-0.017	-0.147	-0.128	-0.068	-0.018	-0.148

Note: For each Part I. crime, Panel A presents data on the total number of 2009 arrestees as well as the share of arrestees who are 15-24 year old ("young adult" males. Using an estimate of the young adult male share of the U.S. population of 7.2 percent, Panel B computes a, the degree to which young adult males are overrepresented mong arrestees relative to their share of the population. ΔV_p is the model-based predicted crime drop given M_1 = 0.11 and M_2 = 0.07, the initial and post-treatment young adult male shares among Mexican immigrants in Arizona. ΔV_a is the synthetic D-D estimate of the actual decline in crime post-LAWA using 2008 as the post-treatment period.

Chapter 3:
The Effect of Police on Crime:
New Evidence from U.S. Cities, 1960-2010

Aaron Chalfin
UC Berkeley

Justin McCrary
UC Berkeley, NBER

May 13, 2013

Abstract

Using a new panel data set on crime in medium to large U.S. cities over 1960-2010, we show that (1) year-over-year changes in police per capita are largely idiosyncratic to demographic factors, the local economy, city budgets, measures of social disorganization, and recent changes in crime rates, (2) year-over-year changes in police per capita are mismeasured, leading many estimates in the literature to be too small by a factor of 5, and (3) after correcting for measurement error bias and controlling for population growth, a regression of within-state differences in year-over-year changes in city crimes on within-state differences in year-over-year changes in police yields economically large point estimates. Our estimates are generally similar in magnitude to, but are estimated with a great deal more precision than, those from the quasi-experimental literature. Our estimates imply that each dollar spent on police is associated with approximately $1.60 in reduced victimization costs, suggesting that U.S. cities employ too few police. The estimates confirm a controversial finding from the previous literature that police reduce violent crime more so than property crime.

I. Introduction

One of the most intuitive predictions of deterrence theory is that an increase in a typical offender's chance of being caught decreases crime. This prediction is a core part of Becker's (1968) account of deterrence theory and is also present in historical articulations of deterrence theory, such as Beccaria (1764) and Bentham (1781). The prediction is no less important in more recent treatments, such as the models discussed in Lochner (2004b), Burdett, Lagos and Wright (2004) and Lee and McCrary (2009), among others.[1]

On the empirical side, one of the larger literatures in crime focuses on the effect of police on crime, where police are viewed as a primary factor influencing the chance of apprehension facing a potential offender.[2] This literature is ably summarized by Cameron (1988), Nagin (1998), Eck and Maguire (2000), Skogan and Frydll (2004) and Levitt and Miles (2006), all of whom provide extensive references.

Papers in this literature employ a wide variety of econometric approaches. Early empirical papers such as Ehrlich (1972) and Wilson and Boland (1978) focused on the cross-sectional association between police and crime. More recently, concern over the potential endogeneity of policing levels has led to a predominance of papers using panel data techniques such as first-differencing and, more recently, quasi-experimental techniques such as instrumental variables (IV) and differences-in-differences. Prominent panel data papers include Cornwell and Trumbull (1994), Marvell and Moody (1996), Witt, Clarke and Fielding (1999), Fajnzylber, Lederman and Loayza (2002), and Baltagi (2006). Some of the leading examples of quasi-experimental papers are Levitt (1997), DiTella and Schargrodsky (2004), Klick and Tabarrok (2005), Evans and Owens (2007) and Machin amd Marie (2011).

Despite their extraordinary creativity, the quasi-experimental approaches pursued in the literature are typically limited in terms of their inferences by difficulties with precision. For example, a typical finding from this literature is that the police elasticity

[1] Polinsky and Shavell (2000) provide a review of the theoretical deterrence literature that emerged since Becker (1968), with a particular focus on the normative implications of the theory for the organization of law enforcement strategies.

[2] A related literature considers the efficacy of adoption of "best practices" in policing. Declines in crime have been linked to the adoption of "hot spots" policing (Sherman and Rogan 1995, Sherman and Weisburd 1995, Braga 2001, Braga 2005, Weisburd 2005, Braga and Bond 2008, Berk and MacDonald 2010) "problem-oriented" policing (Braga, Weisburd, Waring, Mazerolle, Spelman and Gajewski 1999, Braga, Kennedy, Waring and Piehl 2001, Weisburd, Telep, Hincke and Eck 2010) and a variety of similarly proactive approaches. In this paper, we address the effect of additional manpower, under the assumption that police departments operate according to "business-as-usual" practices. As a result, the estimates we report are likely an underestimate with respect to what is possible if additional officers are hired and utilized optimally.

is larger in magnitude for violent crime than for property crime. This finding is often viewed skeptically however, as there is a common belief that violent crimes such as murder or rape are more apt to be crimes of passion than property crimes such as motor vehicle theft. However, the standard errors on the violent and property crime estimates from the previous literature have been large enough that it is unclear whether the difference in the point estimates is distinguishable from zero. Indeed, for many of the papers in the literature, estimated police elasticities for specific crimes are only statistically distinct from zero if additional pooling restrictions are imposed (e.g., equal effect sizes for all violent crime categories). Overall, the imprecision of the estimates from the quasi-experimental literature has led to substantial ambiguity regarding the substance of its findings.

Approaches based on natural variation lead to notably more precise estimates than do quasi-experimental approaches, but the former may be apt to bias because of confounding. This suggests there is merit in assessing the extent of confounding. We present evidence that confounding may be less of an issue than previously believed. In particular, using a new panel data set on crime, police, and a host of covariates for 242 large U.S. cities over the period 1960-2010, we demonstrate empirically that, conditional on standard controls, year-over-year changes in police have generally weak associations with the confounders mentioned in the literature, such as demographic factors, the local economy, city budgets, social disorganization, and recent changes in crime. This new dataset covers more cities than have been used and more years than have been used in most (but not all) of the previous literature.

The weakness of the correlations between police and confounders suggests that estimates of the effect of police on crime using natural variation in police may be only slightly biased, despite the *a priori* concerns raised in the literature. A potential problem, however, with using natural variation is that any measurement error in police could lead to bias of a different nature—measurement error bias. The "iron law of econometrics" is that, in a regression, the coefficient on a predicting variable will be too small in magnitude if it is measured with error, with the bias increasing in the amount of measurement error (Hausman 2001). Most natural experiment approaches, such as IV, do not suffer from the same bias (see, for example, Bound, Brown and Mathiowitz (2001), at least under the hypotheses of the classical measurement error model. Measurement error bias thus has the potential to explain the larger magnitude of the estimates from the quasi-experimental literature, as compared to the traditional literature using natural variation, which has not addressed the issue of measurement errors in police. We show that there is a surprisingly high degree of measurement error in the basic dataset on police used in the U.S. literature, the Uniform Crime Reports (UCR).[3] Estimates from the older panel data literature that failed to account

[3] The degree to which estimates of the total number of police nationally are compromised by

for measurement error bias were likely too small by a factor of 5.

The core of our paper is a series of measurement error corrected estimates of the effect of police on crime using natural variation in year-over-year changes in police at the city level in the U.S. in recent decades. Our estimated police elasticities are substantively large and, taken at face value, suggest that the social value of an additional dollar spent on police in 2010 is approximately $1.60. We introduce a conceptual framework articulating precise conditions under which such a cost-benefit test justifies hiring additional police. The results we introduce along these lines parallel the "sufficient statistic" results discussed in some of the recent public finance literature (e.g., Chetty 2009).

In addition to being significant in substantive terms, our estimated police elasticities are significant in statistical terms. The precision of our estimates allows us to confirm the common and somewhat surprising finding from the previous literature, alluded to above, that police have more of an influence on violent crime than on property crime.[4] However, prior literature has not been able to reject the null hypothesis that the violent crime elasticity is equal to the property crime elasticity, due to imprecise estimates. Our analysis is the first to demonstrate that this apparent finding is unlikely to be due to chance.

Essential to our empirical approach is the existence of two independent measures of police. We combine the standard UCR data on the number of police with data on the number of police from the Annual Survey of Government (ASG). Under the assumptions of the classical measurement error model, described below, IV using one measure as an instrument for the other is a consistent estimator for the results of least squares, were there to be no measurement error. The assumptions of the classical measurement error model are strong, but partially testable. We present the results of a battery of tests of the hypotheses of the classical measurement error model, finding little evidence in our data against them. The tests we utilize would appear to be new to the literature.

Since we focus on natural variation in policing, it is, of course, possible that our estimates are subject to simultaneity bias. It is typical in this literature to difference the data, thus removing between-city variation, and to control for national crime trends using year effects. As the quasi-experimental literature has emphasized, however, this approach may be compromised by confounders associated with growth rates in police and

measurement errors in the UCR data has been noted by Eck and Maguire (2000). However, they do not discuss the potential for measurement errors at the city level to bias estimates of the police elasticity derived from panel data.

[4]The cross-crime pattern of the police elasticity estimates could reflect relative deterrence effects, relative incapacitation effects, or non-classical measurement error. The deterrence effect of police is that some crimes will not occur, because a person notes the increase in police presence and thereby is deterred from committing the offense. The incapacitation effect of police is that some crimes will not occur because additional police will result in arrests, pre-trial detention, and jail time for those who offend (McCrary 2009). The non-classical measurement error hypothesis we have in mind is that increases in police might increase reporting of crimes to police. See Levitt (1998) for discussion.

growth rates in crime. A particular concern is that changes in regional macroeconomic conditions, shocks to regional crime markets, or changes in state-level criminal justice policies may act as important confounders, thus biasing the results from standard panel data approaches. The omission of time-varying state-level policy variables is especially concerning as the adoption of a "get tough on crime" attitude among a state's lawmakers might plausibly lead to both increases in police through increased block grants and passage of more punitive state sentencing policies. Such an attitude might be associated with harsher sentencing along both the intensive and extensive margin, changes in a state's capital punishment regime, decreases in the generosity of the state's welfare system or changes in the provision of other public services to low-income individuals.

We seek to address these potential sources of bias with the inclusion of state-by-year effects, an innovation that has not, to date, been utilized in the literature. These state-by-year effects add roughly 1,500 parameters to each set of IV estimates and control for unobserved heterogeneity in city-level crime rates that is constant within the state. Inclusion of these variables increases the R^2 in crime regressions to nearly 60 percent for most crime categories. This is a remarkably high degree of explanatory power for a panel data model specified in growth rates. To the extent that omitted variables bias remains, we note that the previous literature has emphasized that simultaneity bias would lead regression estimates to be positively biased, i.e., to understate the magnitude of the police elasticity of crime (e.g., Nagin 1978, 1998. This reasoning would suggest that our estimates are conservative in magnitude.

The police elasticity of crime is obviously an important component of any public policy discussion regarding the wisdom of changes in police staffing. However, public investments in policing may crowd out private investments in precaution, making a social welfare evaluation of police more involved than it would at first appear. In Section II, we articulate precise conditions under which the police elasticity of crime can be used as a basis for social welfare analysis when private precautions are a first-order consideration. Our framework is related to recent work in public finance emphasizing the central role of policy elasticities in social welfare analysis (e.g., Chetty 2009).

After the social welfare analysis of Section II, Section III shows police hiring is only weakly related to the usual suspected confounders and discusses institutional aspects of police hiring that limit the scope for confounding. This section also provides some comments regarding interpretation. Next, in Section IV, we present direct evidence on the degree of measurement error in survey and administrative data on the number of police. We then outline our econometric methodology in Section V, discuss our primary data in Section VI, and report estimated police elasticities of crime in Section VII. In Section VIII, we compare our results to those from the previous literature. Section IX connects the social welfare analysis of Section II with the empirical findings of Sectio VII; produces a list of the 30 most overpoliced and 30 most underpoliced cities in our sample; and discusses the robustness of our policy conclusions to incapacitation effects of police. Finally, Section X concludes.

II. Conceptual Framework

Our paper provides an empirical examination of the magnitude of the police elasticity of crime. A natural question is whether the elasticity estimates we present are large or small. We now introduce a conceptual framework designed to adress this issue.[5] The framework will provide conditions under which comparing a police elasticity of crime to the ratio of taxes for supporting public policing to the expected cost of crime is a valid basis for welfare analysis (cf., Saez 2001, Chettty 2006, 2009). That is, this section answers the question: Supposing policing passes a cost-benefit test, under what types of conditions is this sufficient to justify hiring additional police officers?

Here is the basic framework we consider. Suppose society consists of n individuals with linear utility over wealth. Each individual i faces a probability of victimization that depends on own precautions, X_i, the precautions of others, and policing, S. The probability of victimization is denoted $\phi_i \equiv \phi_i(X_1, X_2, \ldots, X_n, S)$ and ϕ_i is assumed continuous in all arguments and convex in X_i and in S. To finance policing, each individual pays a lump-sum tax, τ. We assume agents are in a Nash equilibrium, so that the beliefs of any one individual regarding the precautions of others is consistent with the beliefs of the others regarding the precaution of the one. For person i, we take expected utility to be given by

$$U_i \quad = \quad (y_i - k_i)\,\phi_i + y_i(1-\phi_i) = y_i - k_i\phi_i \tag{1}$$

where k_i is the cost of crime, $y_i = A_i - \tau - p_i X_i$ is after-tax wealth net of expenditures on precautions, A_i is initial wealth, and p_i is the price of precaution. We assume any goods that must be purchased in order to obtain precaution are produced under conditions of perfect competition, implying that the only social value of precaution is in lowering crime.[6] Our definition of expected utility can either be thought of as implying that society is comprised exclusively of potential victims or as implying that the social planner refuses to dignify the perpetrator's increased utility, as in Stigler (1970).[7]

Our social planner faces two types of constraints. The *financing* constraint is that

[5] Our analysis holds fixed the punishment schedule facing offenders and asks only how to optimally set the probability of apprehension. This can be thought of as a social welfare analysis focused on the choice of policing facing a city having little influence on state sentencing policy.

[6] Precaution may or may not involve a market transaction. For example, it could entail circumnavigating a dangerous neighborhood at the expense of extra travel time, or it could also involve the purchase of a burglar alarm. In these examples, the price of precaution is either the cost of the additional travel time or the market price of the alarm.

[7] See Cameron (1989) for a valuable discussion of these conceptual issues and extensive references to the relevant literature.

total tax receipts for policing, $n\tau$, must equal total expenditures, wS, where w is the cost of hiring an additional officer. The *liberty* constraint is that the social planner is either unwilling or unable to dictate an individual's investments in precaution. To motivate the liberty constraint, note that a person installing a burglar alarm would not be held liable in tort for the burglary of her neighbor, even if it could be shown that the cause of her neighbor's burglary was the installation of the alarm. The liberty constraint is thus one that actual governments respect. To clarify that our social planner calculations are different from an unrestricted social planner's calculations where precautions could conceivably be dictated, we refer to the constrained social planner as the *state*. We define the state's problem as the maximization of average expected utility, $\frac{1}{n}\sum_{i=1}^{n} U_i$, subject to the financing and liberty constraints. This problem can be thought of as (1) delegating to each individual the choice of precaution; and (2) maximizing the average indirect utility function over policing. To solve the state's problem, then, we begin by solving the individual's problem.

Individuals adjust precautions to maximize expected utility. The first order necessary condition for this problem, which is also sufficient under our assumptions, is $p_i = -k_i\phi_{ii}$, where the second subscript indicates a partial derivative. We assume that precautions and policing are both protective against crime, or that $\phi_{ii} < 0$ and $\phi_{iS} < 0$. Solving the first order condition for X_i leads to a reaction function, $X_i(X_{-i}, S)$, specifying the privately optimal level of precaution as a function of the precaution of others and policing, where X_{-i} is the vector of precautions for all agents other than i.[8]

Under the assumptions above, each agent has a unique best strategy for any given set of beliefs regarding the actions of other agents, and we obtain a Nash equilibrium in pure strategies (Dasgupta and Maskin 1986, Theorems 1,2). Figure 1 shows individual reaction functions for the $n = 2$ case under high and low policing.[9] The equilibrium requirement that beliefs be mutually consistent implies a set of restrictions. These restrictions lead to equilibrium demand functions, or the level of precaution demanded by person i as a function of policing, prices, taxes, and assets alone (i.e., not the precautions of others). Write equilibrium demand for precaution as $X_i(S)$. Substituting the equilibrium demand functions into the individual's utility function yields equilibrium maximized expected utility for the individual, or

$$V_i(S) \quad = \quad A_i - \tau - p_i X_i(S) - k_i\phi_i\Big(X_1(S), X_2(S), \ldots, X_n(S), S\Big) \qquad (2)$$

[8]We suppress the dependence of the reaction function on prices, taxes, and initial assets to maintain a simple presentation.

[9]The example assumes $-\ln\phi_i(X_1, X_2, S) = \alpha X_i + \beta X_{-i} + \gamma S$, with $\beta < \alpha$, which leads to linear reaction functions $X_i(X_{-i}, S) = (1/\alpha)\left(\ln(\alpha k_i/p_i) - \beta X_{-i} - \gamma S\right)$. This formulation thus echoes the traditional textbook treatment of Cournot duopoly with linear demand (e.g., Tirole 1988, Chapter 5).

The state maximizes the average $V_i(S)$ subject to the financing constraint. Define $\mathcal{V}(S) \equiv \frac{1}{n}\sum_i V_i(S)$ where $\tau = wS/n$. The first order necessary condition, which is also sufficient, is $0 = \mathcal{V}'(S) = \frac{1}{n}\sum_i(-w/n + V_i'(S))$. In this framework, police affect expected utility for individuals through five distinct mechanisms:

1. additional police lower utility by increasing the tax burden ($-w/n < 0$);
2. additional police increase utility by lowering expenditures on precaution ($-p_i X_i'(S) > 0$);
3. additional police lower utility by crowding out precaution, thereby increasing the probability of crime indirectly ($-k_i\phi_{ii}X_i'(S) < 0$);
4. additional police increase utility by reducing the probability of crime directly ($-k_i\phi_{iS} > 0$); and
5. additional police either lower or increase utility by crowding out precautions by persons $\ell \neq i$, either increasing or decreasing, respectively, the probability of crime externally (the sign of $-k_i\phi_{i\ell}X_\ell'(S)$ is ambiguous because the sign of $\phi_{i\ell}$ is ambiguous)

The first order condition for the state's problem reflects these different mechanisms. Multiplying the first order condition by S/C, where $C = \frac{1}{n}\sum_{i=1}^n k_i\phi_i$ is the *crime index*, or the average expected cost of crime, does not change the sign of the derivative and yields a convenient elasticity representation. We have

$$\mathcal{V}'(S)\frac{S}{C} = -\frac{wS}{nC} - \sum_{i=1}^n \omega_i\rho_i\eta_i - \sum_{i=1}^n \omega_i\varepsilon_{ii}\eta_i - \sum_{i=1}^n \omega_i\varepsilon_{iS} - \sum_{i=1}^n \omega_i \sum_{\ell\neq i}\varepsilon_{i\ell}\eta_\ell \quad (3)$$

where $wS/(nC) = \tau/C$ is the tax burden relative to the expected cost of crime, $\omega_i = k_i\phi_i \big/ \sum_{i=1}^n k_i\phi_i$ is the fraction of the expected cost of crime borne by person i, $\rho_i = p_iX_i(S)/(k_i\phi_i) < 1$ is the ratio of precaution expenses to the expected cost of crime, $\varepsilon_{iS} = \phi_{iS}S/\phi_i < 0$ is the partial elasticity of the probability of crime for person i with respect to policing, $\varepsilon_{i\ell} = \phi_{i\ell}X_\ell(S)/\phi_i$ is the partial elasticity of the probability of crime for person i with respect to precaution for person ℓ, and $\eta_i = X_i'(S)S/X_i(S)$ is the elasticity of precaution for person i with respect to policing. The five terms in equation (3) correspond to the five different mechanisms described above. Note that if individuals are taking optimal precautions, then the second and third mechanisms exactly offset, i.e., $-\sum_i \omega_i\rho_i\eta_i - \sum_i \omega_i\varepsilon_{ii}\eta_i = 0$, or the envelope theorem.

We now turn to the task of connecting the state's optimality condition to observable quantities, in particular the police elasticity of crime. Estimates of the police elasticity of crime are of two types. The first type is a *total* police elasticity, so called because it reflects both the direct reduction in crime due to increasing police as well as the indirect increase in crime due to reductions to precautions that result from hiring police.

The second type is a *partial* police elasticity, so called because it holds precautions fixed and thus reflects only the direct reduction in crime due to increased police. Since our study focuses on changes in crime associated with year-to-year fluctuations in policing, we believe that our study most likely identifies a partial police elasticity, at least if most precautions are fixed investments, such as deadbolts and burglar alarms, or if precautions take the form of habits of potential crime victims that are slow to evolve. Because this is plausible but not demonstrable, however, we provide empirical calibrations both under the assumption that our study identifies the partial elasticity and under the assumption that it identifies the total elasticity.

To make these ideas explicit, note that the total and partial elasticities are given by

$$\widetilde{\theta} = \left(\frac{1}{n}\sum_{i=1}^{n} k_i \left\{\phi_{ii}X_i'(S) + \phi_{iS} + \sum_{\ell\neq i}\phi_{i\ell}X_\ell'(S)\right\}\right)\frac{S}{C} = \sum_{i=1}^{n}\omega_i\left(\varepsilon_{ii}\eta_i + \varepsilon_{iS} + \sum_{\ell\neq i}\varepsilon_{i\ell}\eta_\ell\right) \tag{4}$$

and

$$\theta = \left(\frac{1}{n}\sum_{i=1}^{n} k_i\phi_{iS}\right)\frac{S}{C} = \sum_{i=1}^{n}\omega_i\varepsilon_{iS}\,, \tag{5}$$

respectively. Next, combining equations (3), (4), and (5), we have

$$\mathcal{V}'(S)\frac{S}{C} = -\frac{wS}{nC} - \sum_{i=1}^{n}\omega_i\rho_i\eta_i - \widetilde{\theta} \equiv -\frac{wS}{nC} - r - \widetilde{\theta} \tag{6}$$

$$= -\frac{wS}{nC} - \sum_{i=1}^{n}\omega_i\sum_{\ell\neq i}\varepsilon_{i\ell}\eta_\ell - \theta \equiv -\frac{wS}{nC} - e - \theta \tag{7}$$

where $r = \sum_{i=1}^{n}\omega_i\rho_i\eta_i$ is the *crowdout effect*, or the weighted average product of the ratio of precaution expenses to the expected cost of crime (ρ_i) and the elasticity of precaution with respect to policing (η_i), and $e = \sum_i\omega_i\sum_{\ell\neq i}\varepsilon_{i\ell}\eta_\ell$ is the *externality effect*, or the weighted average change in the crime index that results from policing crowding out precautions and externally impacting crime (i.e., the fifth mechanism affecting expected utility described above). The weights in the weighted average (ω_i) correspond to the fraction of the total expected cost of crime borne by person i.

The signs of the crowdout and externality effects will be important for some of our reasoning. Consider first the crowdout effect. While we can imagine that a given individual might perversely increase precaution with increased policing,[10] we believe that this is rare. We assume that, at least on average in the population, policing crowds out precautions. Since ρ_i cannot be negative, this means we assume $r \leq 0$.

[10]For example, we can imagine an individual who does not think installing a camera is worth it, because she does not believe there are enough police to follow up on any leads she might give them.

The sign of the externality effect is somewhat more ambiguous. On the one hand, if forced to guess we would say that most precautions have beggar-they-neighbor effects (i.e., for most i and ℓ, $\varepsilon_{i\ell} \geq 0$), implying a negative overall externality effect, or $e \leq 0$. On the other hand, there are of course precautions that have positive externalities, such as LoJack®. Finally, many precautions have aspects of *both* positive and negative externalities.[11] Consequently, although we have a prior view, we will calibrate our empirical analysis allowing for both positive and negative externality effects.

As noted, equations (6) and (7) are both proportional to the first order condition for the state's problem of maximizing $\mathcal{V}(S)$. Consequently, the state's solution can be recast in terms of the total and partial police elasticities, taxes relative to the expected cost of crime, the externality effect, and the crowdout effect.

Consider first the possibility that our empirical analysis identifies the total elasticity, $\widetilde{\theta}$, i.e., that precautions adjust quickly. Rearranging equation (6) shows that

$$\mathcal{V}'(S) > 0 \quad \Longleftrightarrow \quad \widetilde{\theta}\left(1 + r/\widetilde{\theta}\right) < -\frac{wS}{nC} \tag{8}$$

Suppose that increasing police is worthwhile in the provisional cost-benefit sense that

$$|\widetilde{\theta}| \Big/ \frac{wS}{nC} \equiv \widetilde{\kappa} \quad > \quad 1 \tag{9}$$

Since r and $\widetilde{\theta}$ share sign, the adjustment term $1 + r/\widetilde{\theta}$ is bigger than one, and if $\widetilde{\kappa} > 1$ then it is conservative to conclude that increasing police is welfare improving. Intuitively, this follows since increasing police under this scenario has two benefits for individuals—reduced crime and reduced expenditures on precaution—and only the first benefit is measured by the police elasticity.

Consider next the possibility that our empirical analysis identifies the partial elasticity, θ, i.e., that precautions are slow to adjust. Rearranging equation (7) shows that

$$\mathcal{V}'(S) > 0 \quad \Longleftrightarrow \quad \theta\left(1 + e/\theta\right) < -\frac{wS}{nC} \tag{10}$$

Suppose now that increasing police is worthwhile in the provisional cost-benefit sense that

[11] For example, the Club® has a negative externality in that it may displace car theft to another car (Ayres and Levitt 1998). On the other hand, each additional car using the Club® raises search costs for the car thief and provides a marginal disincentive to car theft. As a second example, consider a business installing a security camera. The camera could have a negative externality in displacing a burglary to another business and a positive externality in deterring a sidewalk robbery.

$$|\theta| \Big/ \frac{wS}{nC} \equiv \kappa \quad > \quad 1 \tag{11}$$

An analysis like that above shows that if $e \leq 0$, i.e., if precautions have beggar-thy-neighbor effects on average, then it is conservative to conclude that increasing police is welfare improving. This makes sense because under this scenario a typical person's precaution imposes a negative externality on others which government can mitigate through police hiring. Suppose instead that $e > 0$, or that precautions have positive externalities on average. In this scenario, government has an incentive to restrict public policing somewhat, in order to encourage precaution. We assume that externalities play a smaller role than the direct effect of policing, or that $e < |\theta|$.[12] We then have the bounds $0 < 1 + e/\theta < 1$, and the conclusion that

$$\mathcal{V}'(S) > 0 \quad \Longleftrightarrow \quad \theta < -\frac{wS}{nC}\frac{1}{1+e/\theta} \quad \Longleftrightarrow \quad |\theta| \Big/ \frac{wS}{nC} = \kappa > \frac{1}{1-e/|\theta|} \tag{12}$$

Consequently, the provisional conclusion that increasing police is welfare improving remains correct if

$$\kappa > \frac{1}{1-e/|\theta|} \quad \Longleftrightarrow \quad \frac{e}{|\theta|} < \frac{\kappa - 1}{\kappa} \tag{13}$$

In words, if $\kappa > 1$, hiring police improves welfare as long as the externality effect is not too big relative to the partial elasticity. For example, if $\kappa = 2$, then additional police are socially valuable unless the externality effect is half as large as the partial elasticity, and if $\kappa = 1.5$, then additional police are socially valuable unless the externality effect is one-third as large as the partial elasticity.

This basic framework is readily extended in a variety of directions. One such direction pertains to multiple crime categories, which will be relevant for our empirical calibrations. For multiple crime categories, the crime index continues to be defined as the average expected cost of crime but no longer has the simple definition from above because there is more than one crime category. However, if we redefine the crime index as

$$C \quad = \quad \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{J} k_i^j \phi_i^j \tag{14}$$

[12] Since θ is negative and $wS/(nC)$ is positive, the second inequality in (10) cannot be satisfied if $e > |\theta|$. If $e > |\theta|$ regardless of the level of S, then $\mathcal{V}'(S) > 0$ is never satisfied, and the state is at a corner solution where it is optimal to have no police.

we retain the core conclusions of the above analysis with analogous redefinition of terms and greater notational complexity. In connecting our empirical results with this normative framework, we draw on the the literature seeking to estimate the cost of various crimes (e.g., Cohen 2000, Cohen and Piquero 2008). This literature can be understood as seeking to estimate k_i^j for a "typical person". With these estimates, we can take $\widehat{C} = \sum_j k^j N^j / P$ as an approximation to the true crime index, where P is a measure of population, k^j is the cost of crime j, and N^j is the number of such crimes reported to police in a given jurisdiction in a given year, or an approximation for $\sum_i \phi_i^j$. These measurement considerations suggest that in empirical analysis one could either use the cost-weighted sum of crimes per capita as a dependent variable, or use the cost-weighted sum of crimes as a dependent variable provided there were population controls included as covariates. We follow the latter approach, as we detail below. Consequently, throughout our analysis, we will consider not just the effect of police on aggregate crime, as is typical of most crime papers, but also the effect of police on the *cost-weighted crime index*, or the weighted sum of crimes, where the weights are an estimate of the cost of the crime. We provide detail on these weights in Section VI, below.

III. Institutional Background and Identification Strategy

As noted above, the primary focus of much of the recent literature on police and crime has been the potential endogeneity of changes in police force strength. These concerns are rooted in the notion that a city ideally intertemporally adjusts its policing levels to smooth the marginal disutility of crime for the median voter, just as a consumer in a lifecycle model ideally intertemporally adjusts purchases to smooth the marginal utility of consumption. Such intertemporal adjustments to police would lead changes in police levels to be endogenous, i.e., to be correlated with unobserved determinants of changes in crime.

Our reading of the economics, political science, and public administration literatures is that the realities of city constraints and politics make intertemporal smoothing difficult, dampening the scope for endogeneity of this type. Cities labor under state- and city-level statutory and constitutional requirements that they balance their budgets annually,[13] they face tax and expenditure limitations,[14] they confront risks associated with hiring police due to legal and contractual obligations which encourages hiring as a means of solving long-term rather than short-term problems,[15] they may be operating

[13]See Cope (1992), Lewis (1994), Rubin (1997) and City of Boston (2007).

[14]See Advisory Commission on Intergovernmental Relations (1977b, 1995), Joyce and Mullins (1991), Poterba and Rueben (1995), Shadbegian (1998, 1999).

[15]Regarding legal obligations, consider two examples: during 1972-1982, the federal government began pressuring departments to hire protected class group members with threat of withholding city and department revenues (Chicago Tribune 1972), and during 1972-1973, Massachusetts municipalities

under a consent decree or court order regarding racial, ethnic, or sex discrimination which may affect hiring decisions directly or indirectly and may affect retention,[16] they may suffer from inattention regarding staffing or may utilize staffing reductions as bargaining chips (e.g., bailout-seeking),[17] and cities may be hamstrung by unilateral changes to state and federal revenue sharing funds that are difficult to anticipate.[18] In addition, state and local civil service ordinances necessitate a lengthy and transparent hiring process making it difficult to adjust policing levels quickly or in great numbers.[19] Finally, cities may suffer from important principal-agent problems with elected officials having potentially quite different objectives from those of the median voter.[20] In short, if the city is analogous to a lifecycle consumer, it is most akin to one confronting liquidity constraints, limited information, inattention, and perhaps even self-commitment problems.

To amplify these points, consider the case of Chicago over the last five decades. Figure 2 presents an annotated time series of the number of sworn officers in the

---

were unsure how to proceed with hiring in light of a constitutional challenge to a state statute allowing departments to favor city residents (Larkin 1973). Regarding contractual obligations, note that union contracts and state and local civil service ordinances may make it difficult to fire a police officer, even one who is substantially underperforming.

[16] For general background, see McCrary (2007).

[17] See, for example, LA Times (1966), Ireton (1976), or Recktenwald (1986a, 1986b). A common pattern is for police departments to have hired a large cohort of officers at some point. For some cities, this was after World War II, for other cities it was the late 1950s, and for other cities it was the 1960s crime wave. Combined with typical pension plans pegged to 20 years or 25 years of service, many departments face retirement waves roughly two decades after a hiring wave, setting the stage for a 20 to 25 year cycle unless the city exercises foresight. For example, in response to the famous Boston Police Strike of 1919, in which nearly three-quarters of the police department went on strike on September 9, then-governor Calvin Coolidge, having assumed control of the department on an emergency basis, refused to allow the strikers to return to work and replaced them all with veterans from World War I (Boston Police Department 1919, Russell 1975). This hiring burst, combined with the State-Boston Retirement System which provides for a defined benefit pension after 10 years if over 55 and after 20 years if of any age, led to a highly persistent "lumpiness" in the tenure distribution of the department (Boston Police Department 1940, Table VI).

[18] Relevant federal programs over this time period include the Law Enforcement Assistance Administration (1968-1982), the Edward Byrne Memorial State and Local Law Enforcement Assistance programs (1988-2006), the Local Law Enforcement Block Grant program (1996-2006), the Justice Assistance Grant (2006-present), and the Community Oriented Policing Services (1994-present). For background on federal programs, see Varon (1975), Hevesi (2005), Richman (2006) and James (2008). At its peak in the late 1970s, LEAA funding accounted for roughly 5 percent of state and local criminal justice expenditures (Advisory Commission on Intergovernmental Relations 1977a). Background on state programs, which are ubiquitous, is much more scarce, but see Richardson (1980).

[19] See, for example, Greisinger, Slovak and Molkup (1979) and Koper, Maguire and Moore (2001).

[20] This perspective is particularly emphasized in the political science literature; see Banfield and Wilson (1963), Salanick and Pfeffer (1977), Schwochau, Feuille and Delaney (1988) and Clingermayer and Feiock (2001).

Chicago Police Department. In 1961, there were just over 10,000 sworn officers in Chicago. Crime and, in particular, the inadequacy of law enforcement was a major theme of the 1964 presidential election (Dodd 1964, Pearson (1964). As riots broke out in many U.S. cities between 1965 and 1968 (Kerner 1968), federal revenue sharing dollars made their way into Chicago budgets and the number of police increased rapidly (Varon 1975). By 1971, the number of sworn officers had risen to just over 13,000. A 1970 suit filed by the Afro-American Patrolmen's League against Chicago alleging *inter alia* discrimination in violation of 42 U.S.C. §1981, the modern legacy of §1 of the Civil Rights Act of 1866, was later joined by the Department of Justice in 1973 after the 1972 amendments to the 1964 Civil Rights Act expanded coverage of the Act to government employers.[21] Eventually, Judge Prentice H. Marshall, a self-described activist judge, reached a now-famous standoff with Mayor Richard Daley Dardick (2004). Marshall ordered the department to use a quota system for future hiring in order to remedy discrimination in past hiring practices. Daley insisted that under such conditions, he did not intend to hire many officers. After impressive brinkmanship, Daley yielded when it became clear that failing to follow the court order would mean the loss of $100 million dollars in federal funds Enstad (1976). Thereafter, the city faced a serious budget crisis (O'Shea 1981). The early 1980s saw the initiation of a long-term hiring freeze (Davis 1985), and with attrition the number of sworn officers fell from 12,916 in February 1983 to 11,945 in May 1986.

By summer 1986, the city faced a tidal wave of upcoming retirements. The department had added a large number of officers in the late 1950s, and those officers were nearing retirement. As of early 1987, fully 4,000 officers were eligible for retirement. The city tried to get ahead of the predictable decline in manpower, but it could not hire quickly enough to replace departing officers (Recktenwald 1986a, 1986b). Consequently, the department began shrinking again from 12,809 in April 1987 to 12,055 in November 1989 as the crack epidemic was roughly three years old.[22] The department managed to return to 12,919 sworn officers by January 1992, however, and policing levels were roughly stable until the beginning of the Community Oriented Policing Services (COPS) program. Between COPS funds and improving city revenues from the strong economy, the number of sworn police officers approached 14,000, reaching a peak of 13,927 in December 1996. The numbers were then stable during the crime decline of the 2000s, but in the wake of the 2008 financial crisis, the number of officers declined to 12,244, eroding nearly all the gains in police strength since 1990. Recently released data from the UCR program suggest that the number of sworn

[21] For background on this litigation, see McCrary (2007) generally and more specifically *Robinson v. Conlisk*, 385 F. Supp. 529 (N.D. Ill. 1974), *United States v. City of Chicago*, 385 F. Supp. 543 (N.D. Ill. 1974), and *United States v. City of Chicago*, 411 F. Supp. 218 (N.D. Ill. 1976).

[22] Based on our own readings, Chicago newspapers begin mentioning the crack epidemic in 1986, and this is also the date identified more quantitatively by Evans, Garthwaite and Mooew (2012).

officers fell to 12,092 in 2011, and recent city payroll data indicate that the number of officers in October 2012 stood at 11,937.[23]

Overall, the key takeaways from Figure 2 are that: (1) Chicago's police strength has fluctuated a great deal over the past five decades, with swings of 10 percent being rather common, and (2) these fluctuations seem to respond to perceptions of lawlessness, but are also the product of political haggling, budgetary mismanagement, gamesmanship, and a seeming lack of attention on the part of city planners. By our reading, these cycles are not limited to Chicago, but are a pervasive feature of police hiring in cities across the United States (cf., Wilson and Grammich 2009).

Sometimes, these cycles are driven by fiscal crisis and bad luck. For example, in 1981, Boston confronted a sluggish to recessionary economy, Proposition $2\frac{1}{2}$, and a major Massachusetts Supreme Court decision that led to large reductions in Boston's property tax revenue.[24] Massachusetts, like other states, requires municipalities to balance their budgets annually.[25] Forced to balance its budget, the city reduced the police department budget by over 27 percent, eliminated all police capital expenditures, closed many police stations, and reduced the number of sworn officers by 24 percent (Boston Police Department (1982).

Other times, these cycles are driven by mayoral objectives that are unrelated to crime. For example, in the mid 1970s, Mayor Coleman Young sought to aggressively hire officers under an affirmative action plan (Deslippe 2004). The department hired 1,245 officers under the plan in 1977, increasing the size of the police force by some 20 percent, and the next year, a further 227 officers were hired under the plan. After Detroit hired those officers, the city confronted a serious budget crisis, forcing the city to lay off 400 and 690 officers in 1979 and 1980, respectively. In 1981 and 1982, the city was able to recall 100 and 171 of the laid off officers, respectively, but a new round of cuts in 1983 undid this effort, and 224 officers were again laid off. In 1984, 135 of those officers were recalled.[26] These sharp changes indicate liquidity problems or perhaps bargaining.

These anecdotal considerations suggest that short-run changes in police are, to a great extent, idiosyncratic. That case is strengthened by establishing that changes in police are only weakly related to changes in observable variables.[27] We now

[23]See `http://www.fbi.gov/about-us/cjis/ucr/` and `http://data.cityofchicago.org`, both accessed on October 24, 2012. The financial crisis led to force reductions in many cities, most famously Camden, which laid off 45 percent of its sworn officers in early 2011 (Katz and Simon 2011).

[24]*Tregor v. Assessors of Boston*, 377 Mass. 602, *cert. denied* 44 U.S. 841 (1979). For background on Proposition $2\frac{1}{2}$, see MDRC (2007).

[25]*General Laws of Massachusetts*, Chapter 59, Section 23. Note that these cuts were partially offset by intervention from state government. See in this regard footnote 27 and Figure 3D, below.

[26]*NAACP v. Detroit Police Officers Association*, 591 F. Supp. 1194 (1984).

[27]Note that we are *not* arguing that police levels fail to respond to crime in the medium- to long-run. Over a longer time horizon, cities may be able to overcome transaction costs and reoptimize, particularly when confronting severe crises. For example, cities facing a difficult crime problem may be able

present statistical evidence on the exogeneity of changes in police to several key social, economic and demographic factors, conditional on some basic controls. Each column of Table 1 presents coefficients from 13 separate regressions of the growth rate in the UCR or ASG measure of police on the growth rate in a potential confounder, conditional on the growth rate in city population and either year effects or state-by-year effects, and weighted by 2010 city population.[28] We motivate and describe in greater detail these controls below. For now, it is sufficient to understand that these are the key covariates we will condition on later in the paper, where we model crime growth rates as a function of police growth rates and other covariates. Standard errors, in parentheses, are robust to heteroskedasticity.[29]

The table is divided into three panels, each of which addresses a different class of potential confounders. Panel A explores the relationship between police and the local economy, as measured by personal income, adjusted gross income, wage and salary income, county-level total employment, and the city's municipal expenditures exclusive of police. There are four sets of models, corresponding to the UCR or ASG measure of police and to year effects or state-by-year effects. The estimates in Panel A give little indication that police hiring is strongly related to local economic conditions. While the estimates based on year effects are all positive, they are generally small in magnitude. For example, the largest estimated elasticity is that of police with respect to total county employment for the UCR measure given city population growth and

to obtain "emergency" funding from the state or federal government. Describing the situation in Washington, D.C., around 1994, (Harriston and Flaherty 1994) note that the "hiring spree [in police] was a result of congressional alarm over the rising crime rate and the fact that 2,300 officers—about 60 percent of the department—were about to become eligible to retire. Congress voted to withhold the $430 million federal payment to the District for 1989 and again for 1990 until about 1,800 more officers were hired." As another example, in response to the 1980-1981 Boston police staffing crisis, "the Massachusetts Legislature enacted the Tregor Act [in 1982]... [providing] the city of Boston with new revenues... This legislative action terminated all layoffs and greatly diminished the risk that future layoffs might take place." *Boston Firefighters Union Local 718 v. Boston Chapter NAACP, Inc.*, 468 U.S. 1206, 1207 (1984). To the extent that even short-run fluctuations in police are partly responding to crime, it is likely that our estimates understate the effect of police on crime.

[28]As discussed in greater detail below, we include two separate measures of city population growth in these regressions to mitigate measurement error bias associated with errors in measuring city population. The UCR and ASG measures of police are described in Section VI, below. For details on the other variables used in this table, see the Data Appendix. We prefer not to control for these variables directly in our main analyses because they are missing for many years. However, after presenting our main results, we conduct a robustness analysis for the 1970-2002 subsample. During that time period, we can control for most of the potential confounders. These results, given in Table 7, show that our main effects are essentially unaffected by the inclusion of further covariates.

[29]For this and all other tables in the paper, we have additionally computed standard errors that are clustered at the level of the city. These are scarcely different from, and often smaller than, those based on Huber-Eicker-White techniques. The similarity in the standard errors suggests small intra-city residual correlations.

year effects, where the elasticity estimate is 0.10. This would imply that a large 10 percent increase in total employment would result in a 1 percent increase in police. The estimates based on state-by-year effects are slightly smaller in magnitude and of varying signs. In addition to being economically small, these estimates are generally statistically indistinguishable from zero at conventional significance levels, or just barely on the cusp of statistical significance.

In Panel B of Table 1 we relate police hiring to several measures of social disorganization and a measure of the population at risk of arrest. The measures of social disorganization are the fraction of births where the mother is a teenager, for all babies as well as for African American babies, the fraction of births where the baby is low birthweight (less than 2500 grams, or about 5.5 pounds), and a proxy for the 12th grade dropout rate.[30] The first three measures are only available at the county level, while the fourth is measured at the city level. While not directly linked to crime, all four of these variables capture changes in local conditions which may correlate with a need for increased police. Notably, a casual examination of the city-specific time series for each of these variables shows that they are often strongly related to the onset of the "crack epidemic" that swept through U.S. cities during the late 1980s and early 1990s.[31] The results for these four variables are similar to those in Panel A, with little indication of a strong relationship between police and social disorganization, at least conditional on controls. The majority of the elasticities are negative, indicating that more social disorganization is associated with less policing rather than more, and are extremely small in magnitude. The elasticity of largest magnitude, that pertaining to the fraction of births where the mother is a teenager, among black births, for the ASG measure with year effects, is just -0.04.

The last variable in Panel B shows the relationship between police and a proxy for the number of persons at risk for arrest. This variable is constructed using Census data on county population for 16 detailed age-race-gender groups, weighted by the 2009 share of each of these groups among arrestees nationally.[32] For the UCR and ASG series, we estimate elasticities pertaining to the at-risk population of 0.11-0.12 and 0.05-0.10, respectively. As with the other estimates we have seen so far, controlling for year effects or even state-by-year effects matters little. These estimates suggest that a 10 percent increase in the at-risk population leads to a 1 to 1.3 percent increase in police, controlling for the covariates described. We note that a 10 percent increase in the population at risk of arrest is an extreme hypothetical representing 2.6 standard deviations for this variable. Overall, while there is some evidence that police hiring is responsive to demographic changes, we would characterize the relationship as fairly weak.

[30] The dropout rate proxy in year t is one minus the number of 12th graders in the city in year t relative to the number of 11th graders in the city in year $t-1$. See Data Appendix for discussion.

[31] The degree to which the "crack epidemic" represents an exogenous shock to local crime markets has been subject to considerable debate in the literature on the effect of state-level abortion policy on crime.

[32] Arrest data from the FBI were available through 2009 at the time of this writing.

Finally, Panel C presents elasticities of police with respect to three lagged crime aggregates: violent crimes, property crimes and a cost-weighted crime index which weights the prevalence of each crime by an estimate of the social damages associated with that crime (we describe these weights in detail in Section VI, below). These elasticities range from 0.004-0.012 for violent crimes to 0.008-0.019 for property crimes, suggesting that a 10 percent increase in crime would lead to no more than a 0.2 percent increase in police.[33] While the weak association reported here is somewhat at odds with the sense one gets from the existing academic literature, it is consistent with the limited available reports from interviews with police chiefs.[34]

Taken as a whole, the estimates in Table 1 suggest a more limited link between police hiring and potential confounders than the literature has presupposed, at least conditional on our preferred controls. While the results do not indicate a large discrepancy between models controlling for year effects and those controlling for state-by-year effects, in the remainder of the paper, we focus on models for crime that include state-by-year effects. These models are robust to any possible confounder that varies over time at the state level. This includes, for example, state-level policies that may affect crime, such as poor support, education policy, or penal policy. Most papers in the literature focus on models for year-over-year growth rates in crimes at the city level include year effects. Evans and Owens (2007) are unusual in focusing on a more flexible model involving group-specific year effects, where the groups are defined according to population and pre-COPS program crime trends. However, as far as we know, no paper in the literature has used state-by-year effects and thereby completely isolated the effect of police from the effect of state-level policies.

Before closing this section, we would like to address one final issue pertaining to interpretation. After reviewing the literature on police staffing fluctuations, including a non-random sampling of newspaper coverage for specific cities in specific years, we have the impression that policing increases are sometimes associated with the city council or mayor being pleased with the direction of the department. This may mean that the number of police partially proxies for changes in what police are doing, as well. For example, it is possible that increases in police are associated with the hiring of a popular new police chief or with the department being willing to transition

[33]We note that a limitation of our analysis of possible confounders is that there are few variables which are collected systematically for a large number of cities for a long period of time. One suspected confounder in particular–calls for service–is sometimes reported in police department annual reports, but is not collected on a systematic basis by any organization. Consequently, we are unable to completely address the issue of possible confounders. On the other hand, calls for service likely does correlate strongly with other measures we do observe, such as crime.

[34]For example, Police Executive Research Forum (2005b) discusses the results of a focus group with four police chiefs (Largo, Scottsdale, Omaha, and Baltimore County) and a deputy chief (Charlotte-Mecklenburg). "Participants pointed out that the crime rate is usually not a major factor in budget success." (p. 18).

to a community policing model. This might mean that our estimates, and those in the previous literature utilizing natural variation, capture an effect of police that is somewhat broader than just the effect of police manpower, *per se*.

IV. Evidence on the Extent of Measurement Error in the Number of Police

A. Direct Evidence

We begin our discussion of the nature and extent of measurement errors in police personnel data using the case of New York City in 2003. The UCR data reports that the New York Police Department employed 28,614 sworn police officers in 2003.[35] Relative to the 37,240 sworn officers employed in 2002 and the 35,513 officers employed in 2004, this is a remarkably low number. If these numbers are to be believed, then the ranks of sworn officers in New York City fell by one-quarter in 2003, only to return to near full strength in 2004.

An alternative interpretion is that the 2003 number is a mistake. Panel A of Figure 3 compares the time series of sworn officers of the New York Police Department based on the UCR reports with that based on administrative data from 1990-2009.[36] These data confirm that the 2003 measure is in error and additionally suggest that the 1999 measure may be in error.[37]

Administrative data on the number of officers are difficult to obtain. More readily available are numbers from departmental annual reports. However, even these are not easy to obtain; annual reports are largely internal municipal documents and historically did not circulate widely. In recent years, many departments have begun a practice of posting annual reports online, but only a few cities post historical annual reports. Moreover, the annual report may or may not report the number of officers employed by the police department.[38]

Nonetheless, we have been able to obtain scattered observations on the number of sworn officers from annual reports for selected years for selected other cities: Los Angeles, Chicago, Boston, and Lincoln, Nebraska. The numbers for Chicago have been

[35]As discussed below, the UCR measurement protocol is a snapshot of the stock of officers as of October 31 of the survey year.

[36]See the Data Appendix for details on these data. Special thanks to Frank Zimring for pointing us towards public domain information on New York police staffing based on his work on the New York City crime drop (Zimring 2011).

[37]We have discussed the 2003 measure of police with other scholars of crime and police, both in economics and in criminology. To date, we have not heard a plausible account for this number, other than that it is a data entry error.

[38]For example, the annual reports for the Boston Police Department are available online beginning in 1885, but the reports stop detailing the number of officers between 1972 and 1981, when the number of officers fell by 40 percent. See `http://www.bpl.org/online/govdocs/bpd_reports.htm`.

further augmented by the strength report data reported in Siskin and Griffin 1997).[39] The time series of sworn officers for these cities is given in Figure 3 in panels B through E. Treating the administrative and annual report data as the true measure, it seems that there is a broad range of fidelity in reporting to the UCR program, with Los Angeles being the most faithful, New York the least, and the others somewhere between those two bookends. While the series are highly correlated in levels or logs, the correlation is notably lower after taking first differences (results unreported). This is important because most of the recent literature analyzes the data in first-differences or with city effects.

Many people are suprised that there are errors in measuring the number of police officers. Errors could arise due to (1) transitory movements within the year in the number of police, (2) conceptual confusion, or (3) data entry errors. Regarding the first source of error, Figure 4 gives information on transitory movements in police for Chicago for the period 1979-1997. The figure displays the monthly count of the number of sworn officers, with the count for October superimposed as horizontal lines.[40,41] October is chosen because this is the reference month for the UCR data on police used in the literature. The figure demonstrates that there is a great deal of within-year volatility in the number of sworn officers. Overall, the series is characterized by hiring bursts followed by the gradual decline associated with losses due to retention or retirement. Transitory movements in police officers are relevant because surveys typically ask for a point-in-time measure, and the snapshot date differs across surveys. Among those we have been able to examine, internal police department documents use different reporting conventions, typically corresponding to the end of the municipal fiscal year, which varies across municipalities and over time. Perhaps responding to the ambiguities of point-in-time measures, the New York City Police Department uses average daily strength in internal documents.

In addition to transitory movements, there may also be conceptual ambiguity over who counts as a sworn police officer. First, there may be confusion between the number of total employees, which includes civilians, and the number of sworn officers. Second, newly hired sworn officers typically attend Police Academy at reduced pay for roughly 6 months prior to swearing in, and there may be ambiguity regarding whether those students count as sworn officers prior to graduation. Third, due to frictions

[39] See the Data Appendix for details on the annual report and strength report data.

[40] We are not aware of any public-use data sets containing information on within-year fluctuations in police. During the period 1979-1997, a unique non-public dataset on sworn officers in Chicago is available to the authors, however, that allows the construction of monthly counts. These data are discussed in Siskin and Griffin (1997) and were previously used in McCrary (2007). See the Data Appendix for details.

[41] A natural question is whether there is seasonality to police hiring, particularly since summer months are typically high crime months. A regression of log police on an exhaustive set of year and month dummies over the period 1979 to 1997, where monthly data are available, yield an R-square of 0.95. This regression gives little indication of seasonality. While the set of 18 year dummies have an F-statistic of over 193, the set of 11 month dummies have an F-statistic of 1.08, with a p-value of 0.38.

associated with the hiring process, there is often a discrepancy between the number of officers the department has authority from city government to employ ("authorized strength") and the number of officers currently employed ("deployed strength").[42] For our main sample of cities, we have measures of the number of authorized and deployed sworn officers for selected recent years from the Law Enforcement Management and Administrative Statistics (LEMAS). These data show that the number of deployed sworn officers ranges from 62 to 128 percent of authorized strength.[43]

Finally, the UCR measure of sworn police has errors (e.g., New York in 2003) that are inconsistent with transitory movements within the year in the number of sworn police officers and inconsistent with conceptual confusion. For such errors, we have no other explanation than typographical or data entry error.[44]

B. Comparison of Two Noisy Measures

Police department internal documents are presumably more accurate than the information police departments report to the UCR program. However, as discussed, these are only available in selected cities and selected years. Trading off accuracy for coverage, we now present a comparison of the UCR series on the number of sworn officers with a series based on the ASG. We use the ASG data to construct an annual series on full-time sworn officers for all the cities in our main analysis sample. We define this sample and give background on the ASG data in Section VI, below.

Figure 5 provides visual evidence of the statistical association between the UCR and ASG series for sworn officers, measured in logs (panel A) and first differences of logs ("growth rates", panel B). In panel A, we observe a nearly perfect linear relationship between the two measures, with the majority of the data points massed around the 45° line. The regression line relating the log UCR measure to the log ASG measure is nearly on top of the 45° line, with a slope of 0.99. Panel B makes it clear that

[42] Typical steps include a written examination, a drug test, a background check, an interview, and a series of physical and psychological tests, among others Police Executive Research Forum (2005b), Wilson and Grammich (2009).

[43] Numbers refer to a pooled analysis of data from 1987, 1990, 1993, 1997, 1999, 2000, and 2003. Population weighted mean and standard deviation are 97 percent and 5 percent, respectively. The LEMAS data also allow us to discount the possibility that there is error due to ambiguities among sworn officers, full-time sworn officers, or full-time-equivalent sworn officers, as only 1 to 2 percent of officers appear to work part-time.

[44] It is worth noting that the crime data are the focus of the UCR system, with notably less attention paid to the police numbers. It is common to see a discussion of UCR crime figures in the local news and for local politicians to be under fire for any spikes in those numbers. However, neither of us have ever seen a local news discussion of the UCR measure of the number of officers. Perhaps because the release cycle used by the FBI for the UCR system involves releasing the numbers for police well after they release the numbers for crimes, reporters seem to ask cities directly for figures on police. This suggests that any lack of care in preparing the UCR police numbers would usually go unnoticed.

differencing the data substantially reduces the association between the two series; the slope coefficient for the data in growth rates is just 0.22.

To appreciate the implications of these findings for quantification of the police elasticity of crime, we turn to a simple statistical model. Suppose the two observed series on police are related to true police as

$$\begin{aligned} S_i &= S_i^* + u_i && (15) \\ Z_i &= S_i^* + v_i && (16) \end{aligned}$$

and suppose the outcome of interest, Y_i, is related to the true number of police and covariates X_i as

$$Y_i = \theta S_i^* + \gamma' X_i + \varepsilon_i \qquad (17)$$

Here, S_i is the log UCR measure in a given city and year, Z_i is the log ASG measure, S_i^* is the "true" log police or *signal*, X_i are other covariates measured without error, u_i and v_i are mean zero measurement errors that are mutually uncorrelated and uncorrelated with ε_i, S_i^*, and X_i, and ε_i is mean zero and uncorrelated with S_i^*, X_i, u_i, and v_i. Equations (15) through (17) and the stochastic restrictions just named constitute what is known as the classical measurement error model (Fuller 1987).

A famous result from the econometric literature on measurement errors relates the probability limit of the least squares regression estimate of θ, based on using covariates S_i and X_i, to the scope of measurement errors and the relationship between the signal and the included covariates, under the assumptions of the classical measurement error model:

$$\begin{aligned} \operatorname{plim}_{n\to\infty} \widehat{\theta}_{\text{OLS}} &= \frac{\mathbb{C}[\mathcal{M}S_i, \mathcal{M}Y_i]}{\mathbb{V}[\mathcal{M}S_i]} = \frac{\mathbb{C}[\mathcal{M}S_i^* + \mathcal{M}u_i, \mathcal{M}Y_i]}{\mathbb{V}[\mathcal{M}S_i]} = \frac{\mathbb{C}[\mathcal{M}S_i^*, \mathcal{M}Y_i]}{\mathbb{V}[\mathcal{M}S_i]} && (18) \\ &= \theta\, \frac{\mathbb{V}[\mathcal{M}S_i^*]}{\mathbb{V}[\mathcal{M}S_i]} = \theta\, \frac{\mathbb{V}[\eta_i]}{\mathbb{V}[\eta_i + \mathcal{M}u_i]} = \theta\, \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_u^2} && (19) \\ &= \theta\, \frac{\sigma_*^2(1-R^2)}{\sigma_*^2(1-R^2) + \sigma_u^2} && (20) \end{aligned}$$

where for a random variable A_i, we define $\mathcal{M}A_i = A_i - \mathbb{V}[X_i]^{-1}\mathbb{C}[X_i, A_i]$, i.e., applying $\mathcal{M}$ purges a random variable of its linear association with X_i, η_i is the associated linear projection residual, σ_η^2 is the variance of η_i, σ_u^2 is the variance of u_i from equation (15), and R^2 is the population R-squared from a regression of the signal S_i^* on the covariates X_i.[45]

This formula stands for three ideas. First, since $\sigma_u^2 > 0$, OLS will be too small in magnitude, or attenuated. Second, while it is a staple of empirical work to see whether a regression estimate is robust to the inclusion of various control variables, equation (20) indicates that the cure of additional covariates may be worse than the disease of omitted variables bias. Adding more controls increases the R^2, exacerbating any attenuation bias. Third, since the estimates of θ and γ will generally covary, the bias in the estimate of θ will spill over to result in bias in the estimate of γ.

Now return to equation (17) and suppose that X_i is measured without any errors. Under the models in equations (15) and (16) and the associated assumptions on u_i and v_i, it is straightforward to estimate the reliability ratio. The probability limit of the coefficient on Z_i in a regression of S_i on Z_i and X_i is

$$\frac{\mathbb{C}[\mathcal{M}S_i, \mathcal{M}Z_i]}{\mathbb{V}[\mathcal{M}S_i]} = \frac{\mathbb{C}[\mathcal{M}S_i^* + \mathcal{M}u_i, \mathcal{M}S_i^* + \mathcal{M}v_i]}{\mathbb{V}[\mathcal{M}S_i]} = \frac{\mathbb{V}[\mathcal{M}S_i^*]}{\mathbb{V}[\mathcal{M}S_i]} \equiv \pi \qquad (21)$$

This implies that the ratio of the least squares estimate of the police elasticity of crime, relative to the estimate of π, is consistent for θ, suggesting a role for IV. This also implies that, in the context of the discussion of Figure 5, a regression of log crime on log police will not be importantly compromised by measurement errors in police, because in logs the reliability ratio is 0.99. However, a regression of growth rates in crime on growth rates in police and other covariates will be compromised, because in growth rates the reliability ratio is 0.22. Indeed, as we show below, once population growth rates and state-year effects are included in the model, the reliability ratio falls to 0.16. Consequently, even setting aside problems with simultaneity bias of the type discussed in the literature, measurement errors in police suggest that least squares estimates of the police elasticity in the literature are too small by a factor of 5 or more.

V. Econometric Approach

The three equation model introduced in Section IV.B leads naturally to a simultaneous equations model. Substituting equation (15) into equation (17) and linearly projecting S_i onto Z_i and X_i yields

$$Y_i = \theta S_i + \gamma' X_i + \varepsilon_i \qquad (22)$$

$$S_i = \pi Z_i + \phi' X_i + \nu_i \qquad (23)$$

[45]Recall that $R^2 = 1 - \sigma_\eta^2/\sigma_*^2$, so that $\sigma_\eta^2 = \sigma_*^2(1-R^2)$. The formula for $\mathcal{M}$ assumes X_i is a vector with no linear dependencies. More generally, $\mathcal{M}A_i$ is A_i less the linear projection onto the column space of X_i.

where we now interpret Y_i as the year-over-year change in log crime in a given city and year, S_i as the year-over-year difference in observed log police, and X_i as a vector of control variables such as log revenues per capita, log population, the demographic structure of the population, all measured in first differences, as well as year effects or state-by-year effects. In this model, $\varepsilon_i = \varepsilon_i - \theta u_i$, and ν_i is a linear projection error. This is then a standard simultaneous equations model where Z_i is potentially an instrument for S_i.

Estimation of the parameters in equations (22) and (23) proceeds straightforwardly by IV since the model is just-identified, and 2010 city population is used as a weight to obtain a police elasticity estimate representative of the typical person. Sufficient conditions for excluding Z_i from equation (22) are

$$
\begin{array}{lr}
(A1) & \mathbb{C}[u_i, \varepsilon_i] = \mathbb{C}[v_i, \varepsilon_i] = 0 \\
(A2) & \mathbb{C}[u_i, (S_i^*, X_i')'] = \mathbb{C}[v_i, (S_i^*, X_i')'] = 0 \\
(A3) & \mathbb{C}[u_i, v_i] = 0 \\
(A4) & \mathbb{C}[\varepsilon_i, (S_i^*, X_i')'] = 0
\end{array}
$$

where u_i and v_i are the measurement errors from equations (15) and (16) and ε_i is the structural error term from equation (17).[46]

Assumptions $(A1)$ through $(A3)$ assert that the measurement error in the UCR and ASG measures of police are not associated with the structural error term in equation (17), and are not associated with the true growth rate in police and the covariate vector X_i, and that the UCR and ASG measurement errors are mutually uncorrelated, respectively. We discuss empirical implications of assumptions $(A1)$ through $(A3)$ below. Assumption $(A4)$ is innocent if we maintain that we would be interested in running a regression of crime growth rates on police growth rates and controls X_i, were police growth rates observed without error. On the other hand, $(A4)$ may reasonably be called into question. In particular, city population growth rates are measured with error. City population growth is a sufficiently important confounder that we feel the (infeasible) regression model implied by equation (17) and assumption $(A4)$ would not be of interest unless X_i included it.[47] We discuss the challenges of mismeasurement

[46] Assumptions $(A1)$ through $(A4)$ together imply that $\mathbb{E}[Z_i\varepsilon_i] = \mathbb{E}[Z_i u_i] = 0$, which implies that $\mathbb{E}[Z_i\varepsilon_i] = 0$. Assumptions $(A2)$ and $(A4)$ imply that $\mathbb{E}[X_i\varepsilon_i] = 0$. Of course, $\mathbb{E}[(Z_i, X_i')'\varepsilon_i] = 0$ is one of the two familiar conditions for consistency of IV using Z_i as an excluded instrument and X_i as an included instrument. The other familiar condition, that the excluded instrument predict the endogenous regressor, i.e., that $\pi \neq 0$, is unremarkable in this context.

[47] In times of population growth, police force size and crime both grow mechanically. For our sample, a population-weighted regression of the growth rate in a typical crime category on the growth rates of population as measured in the UCR and ASG yields a sum of population elasticities of roughly one or even higher. Replacing the dependent variable by the growth rate in police yields a sum of the population elasticities of roughly four-fifths. The resulting positive bias in the estimated police

of city population growth in greater detail below.

Under the classical measurement error, the exact same steps we used to motivate the simultaneous equations model in equations (22) and (23) can be used to motivate a second simultaneous model with the roles of S_i and Z_i reversed and identical parameters in equation (22).[48] We refer to IV models that use the ASG measure of police as an instrument for the UCR measure as *forward* IV estimates and to models that use the UCR measure of police as an instrument for the ASG measure as *reflected*. As noted, both IV estimates are consistent for the police elasticity of crime. This raises the possibility of pooling the estimates to increase efficiency. To do so, we stack the orthogonality conditions for the forward and reflected IV programs into the broader set of moments

$$g_i(\beta) = W_i \begin{pmatrix} Z_i(Y_i - \theta_1 S_i - \gamma_1' X_i) \\ X_i(Y_i - \theta_1 S_i - \gamma_1' X_i) \\ S_i(Y_i - \theta_2 Z_i - \gamma_2' X_i) \\ X_i(Y_i - \theta_2 Z_i - \gamma_2' X_i) \end{pmatrix} \tag{24}$$

where W_i is 2010 city population in levels and all other variables are as defined before, and we estimate the parameters using generalized method of moments (GMM). When the parameters θ_1 and θ_2 and γ_1 and γ_2 are allowed to differ, estimating those same parameters by GMM is equivalent to estimating them separately by IV and correcting the standard errors for the common dependent variable. We can also estimate the system imposing the restriction $\theta_1 = \theta_2 = \theta$.[49] This leads to an implicit averaging of

elasticity for specifications that omit population growth is quite large economically.

[48]Some well-known papers utilizing IV strategies to address measurement error have focused on the estimated return to education among samples of twins (see Card (1999) for a review of this literature). The set of econometric issues raised in those papers is slightly different than in our context, simply because twin number is randomly assigned in those studies. In our context, the labels "UCR" and "ASG" carry substantive meaning in a way that the twin labels do not.

[49]A somewhat technical issue arises if we additionally seek to impose the restriction that $\gamma_1 = \gamma_2 = \gamma$: redundancy of moments. When we do not impose any restrictions, we have a just-identified system with $2K$ parameters and $2K$ moments, all of which are linearly independent, where X_i has $K-1$ elements. However, once the restrictions $\theta_1 = \theta_2 = \theta$ and $\gamma_1 = \gamma_2 = \gamma$ are imposed, we have K parameters and only $K+1 < 2K$ linearly independent moments. This suggests two obvious approaches to estimation: (1) impose only the restriction $\theta_1 = \theta_2 = \theta$, in which case there is no moment redundancy; or (2) impose both sets of restrictions and drop $K-1$ moments, in which case GMM will embarassingly differ depending on which set of $K-1$ moments are dropped. An involved solution to the difficulty posed by the second approach is to estimate the models by empirical likelihood (EL; see Imbens (1993), Qin and Lawless (1994), and Imbens (2002) for an introduction and references to the literature), in which case estimates are invariant to the set of moments used to identify the model. EL may also be of interest for the first approach, as the model is (slightly) overidentified. We have used both approaches, using both GMM and EL for the sake of completeness, and there is hardly any difference across the four total possibilities. In our discussion, we focus on the first approach using GMM to maintain a simple presentation and

the unrestricted IV estimates and potentially to efficiency gains.[50] An omnibus test of the classical measurement error model is also then available as the standard GMM test of overidentifying restrictions. Since these models are overidentified, there is *a priori* merit in considering empirical likelihood (EL) estimation as well. For overidentified models, EL has been shown to have smaller higher order bias than GMM (Newey and Smith 2004) and to enjoy other advantages as well (see, for example, Imbens, Spady and Johnson 1998). However, in our data, EL estimates and standard errors are nearly identical to two-step GMM estimates, as discussed below, and we focus on GMM.

A challenge we face in implementing the above ideas is that population growth is an important confounder, yet is also likely measured with error. As discussed above, measurement error bias may not have the attenuation bias form if more than one covariate is measured with error. Measurement errors in the population variable in the UCR data are, to the best of our knowledge, not discussed in the literature, but they are likely at least as bad as the measurement errors in police. As with police, any such problems will be particularly serious when the data are measured in growth rates. A potential solution to the measurement problems with city population growth is to again use the UCR measure as an instrument for the ASG measure since both surveys report city population. However, because the two measures are measured very similarly—both are essentially forecasts based on counts from the Census—there are good reasons to believe that the errors in the two measures are not independent of one another. Accordingly, we follow an approach suggested by Lubotsky and Wittenberg (2006) and include both the UCR and the ASG population measures in our main equation of interest. We argue below, based on an empirical comparison to models including data on alternative population controls, that this procedure is sufficient to control for the confounding influence of city population growth.

VI. Data

In this section, we introduce our sample of cities and describe the main sources of information for our data. Our sample of 242 cities is drawn from all cities with more than 50,000 population each year from 1960-2010.[51] In Figure 6 we present a map of the United States highlighting the location of our sampled cities. The shading

additionally report EL estimates for the sake of completeness. We note that EL computation—a thorny issue—in our application was facilitated greatly by suggestions in Guggenberger and Hahn (2005).

[50] Indeed, a very good approximation to the GMM estimate is the weighted average of the forward and reflected IV estimates, with weights of the inverse squared standard errors. In most software packages, this average will be far easier to compute than GMM. However, the standard errors for GMM are notably larger than the square root of the sum of the weights, so for inference purposes the GMM computation may be necessary.

[51] We exclude approximately 30 cities due to extensive missing data and various data quality issues. See Data Appendix for details.

of states provides information on the number of sampled cities in each state. Our sample contains at least one city in 45 of 51 U.S. states, inclusive of the District of Columbia.[52] In addition, there are 10 states for which our sample contains only a single city. This feature of our data will become relevant in understanding parameter estimates that condition on state-by-year effects.[53]

For each city in our sample, we collect information from public data sources on a variety of different measures. We obtain data on crimes and sworn police officers from the UCR. We collect information on sworn police officers from the ASG and from another survey, described below, that is available for selected years since 1987. These three types of data are the core of our analysis, but we also collect auxiliary data on city revenues, police payroll, and police operating budget from the finance files of the ASG; city demographic structure from the Census Bureau; county-level economic data from the Bureau of Economic Analysis; and proxies for social disorganization from the Centers for Disease Control and the National Center for Educational Statistics. Finally, we obtain data on city population from the UCR and ASG which we supplement with data from the National Cancer Institute's Surveillance Epidemiology and End Results (SEER) dataset and some limited information on city births from the National Center for Health Statistics. We now provide more detail regarding each of these data sources. We focus our discussion on our measures of crimes, police, and population, and provide more information regarding our auxiliary data in the Data Appendix.

The UCR crime data we collect are the standard measure used in the empirical literature. These data are collected annually by the FBI. Crime measures represent the total number of offenses known to police to have occurred during the calendar year and are part of the "Return A" collection. The offenses recorded in this system are limited to the so-called index offenses—murder, forcible rape ("rape"), robbery, aggravated assault ("assault"), burglary, larceny exclusive of motor vehicle theft ("larceny"), and motor vehicle theft. Time series for each of the crime rates utilized for each of our cities are shown in Web Appendix Figure 1.

Sworn police are included in both the Law Enforcement Officers Killed or Assaulted (LEOKA) collection and the Police Employees (PE) collection and represent a snapshot as of October 31st of the given year. Because of the late date of the measurement of the number of police, it is typical to measure police in year t using the measure from year $t-1$ (cf.,Levitt 1997), and we follow that convention here. Consequently, although we have data on levels from 1960-2010, our regression analyses of growth rates pertain to 1962-2010.

As noted above, we augment data from the UCR with data from the employment

[52] Alaska, Idaho, North Dakota, Vermont and Wyoming are unrepresented in our sample.

[53] States with only a single sampled city are dropped from the analysis when unrestricted state-by-year effects are included.

files of the ASG. The ASG is an annual survey of municipal payrolls that has been administered by the Bureau of Labor Statistics and reported to the U.S. Census annually since 1952. The ASG data provide payroll data for a large number of municipal functions including elementary and secondary education, judicial functions, public health and hospitals, streets and highways, sewerage and police and fire protection, among others. The survey generally provides information on the number of full-time, part-time and full-time equivalent sworn and civilian employees for each function and for each municipal government.[54] As with the UCR system, the ASG reports a point-in-time measure of police. For 1960-1995 the reference date is November 1 and for 1997-2010 the reference date is June 30.[55]

The UCR data provide the number of full-time sworn police officers and the total number of police officers in each year. The ASG data provide the same information beginning in 1977. Prior to 1977, the ASG series reports only the number of full-time equivalent (FTE) police personnel, without differentiating between sworn officers and civilian employees. In order to extend the series, we use the UCR data to generate a city- and year-specific estimate of the proportion of police personnel who are sworn officers. This was accomplished by regressing the proportion of police personnel who are sworn on city and year indicators using the 1960-1977 sample and generating a predicted value for the sworn percentage in each city-year.[56] The ASG FTE numbers before 1977 were then multiplied by the estimated proportion.

For selected analyses we also draw upon a third measure of police. This measure is drawn from two additional sources: the Law Enforcement Management and Administrative Statistics (LEMAS) series and the Census of State and Local Law Enforcement Agencies. These data, which we refer to as the LEMAS series, have been collected at regular intervals from 1987-2008. For additional details, see the Data Appendix.

The measure of city population used in the majority of crime research is from the FBI's Return A file. While this series contains observations for nearly all city-years, it is potentially contaminated by measurement error, particularly in the years immediately prior to each decennial Census. The population entries are contemporaneous; while the FBI could retroactively correct any of the population figures used in the files, it does not. We augment the city population measure from the UCR with the city population

[54]Full-time equivalent employees represent the number of full-time employees who could have been employed if the hours worked by part-time employees had instead been dedicated exclusively to full-time employees. The statistic is calculated by dividing the number of part-time hours by the standard number of full-time hours and then adding this number to the number of full-time employees.

[55]No annual ASG survey was conducted in 1996. We impute data for 1996 using the average of the 1995 and 1997 levels. Other than this one missing year and occasional missing data, information on police is available in both the UCR data and ASG data for each of these cities for the entire study period.

[56]Time series plots of the number of full-time sworn officers according to the UCR and ASG measures for each city are provided in Web Appendix Figure 2.

measure from the ASG, as noted. As with the UCR, the ASG population measure is noisy and often not smooth across Census year thresholds. Because of the clear errors around Census years, we smooth both series using local linear regression with a bandwidth of 5 years and the triangle kernel (Fan and Gijbels 1996).[57,58,59] Intuitively, this is akin to taking a moving average of the underlying series.

In Section IX, we use data on the cost of police and the cost of crime to derive approximate benefit-cost ratios, both nationally and for specific cities. We pause here to describe these data briefly, with further detail provided in the Data Appendix. Data on the cost of police are taken from the ASG Finance and ASG Employment files from 2003-2010 and are used in conjunction with other public data to estimate a fully-loaded cost of hiring an additional police officer. As noted we also make use of the crime index, or the cost-weighted sum of crimes. The correct weight to use for connecting our empirical estimates to the framework of Section II is a measure of the *ex ante* cost of crime—i.e., the dollar amount a potential crime victim would pay to reduce their probability of victimization, relative to the change in the probability. While this is in principle a person-specific concept, we follow the literature in using an estimate for a representative person. Unfortunately, estimates of the *ex ante* cost of crime are not available except for the crime of murder, where we can take advantage of the rich literature on the value of a statistical life (VSL). For other crimes, we use estimates of the *ex post* costs of crime, which are typically derived from civil jury awards. The value of these civil jury awards captures both direct costs to crime victims arising from

[57]Appendix Figures 1A and 1B provide evidence of the importance of smoothing the raw population measures. These figures present scatterplots of the growth rate in violent and property crimes against the growth rate in the the raw and smoothed population measures from both the UCR and the ASG file. In panel A of Appendix Figure 1A, we see that a 10 percent increase in the population growth rate is associated with a 2.5 percent increase in the number of violent and property crimes. While the crime-population elasticity need not equal 1, this population elasticity is surprisingly small. Panel B plots the crime growth rate against the smoothed UCR population measure. Here, the regression slopes for violent and property crime are 0.94 and 0.84, respectively, neither of which is statistically significantly different from 1. Appendix Figure 1B reports similar results for the ASG population measure. We interpret these findings as evidence that the smoothed population measures more accurately reflect changes in city population.

[58]Below, we test empirically our notion that using both the UCR and the ASG population series adequately controls for population growth using the number of births in a city-year. This can be thought of as a proxy for the size of the population. These data are available from the National Center for Health Statistics (NCHS) for the years 1960-1993 for all 242 cities in our sample and for the years 1960-2003 for 147 of the larger cities. These data are not available electronically, but are available as a series of scanned PDF files at an NCHS website. See `http://www.cdc.gov/nchs/products/vsus.htm`. We had the data on the number of births entered by workers from Amazon's Mechanical Turk service and reviewed them for accuracy. For an introduction to this service, see `http://www.mturk.com/mturk/welcome`. We note that the manner in which we had the data entered and the screening process we undertook together persuade us that there are no data entry errors on the part of the Mechanical Turk workers.

[59]Our population imputations, as well as the raw data underneath them, are shown for each city in the sample in Web Appendix Figures 4A and 4B.

injuries sustained during the commission of the crime, as well as losses arising from reductions in a victim's quality of life.

We turn now to Table 2, which provides summary statistics for each of our two primary police measures as well as each of the seven index offenses. We additionally report summary statistics for the aggregated crime categories of violent and property crime, which simply add together the relevant corresponding individual crime categories, respectively, and for the cost-weighted crime index.

Descriptive statistics are reported for a sample of 10,589 observations, the universe of data for which measures of crime, police and population are nonmissing. The left-hand panel of Table 2 gives statistics for the levels of crime and police in per capita terms, specifically as a measure of the value per 100,000 population. The right-hand panel gives statistics for log differences of crime and police.

Several features of the data are worth noting. First, a typical city employs approximately 250 police officers per 100,000 population, one officer for every 4 violent crimes, and one officer for every 24 property crimes. There is considerable heterogeneity in this measure over time, with the vast majority of cities hiring additional police personnel over the study period. However, there is even greater heterogeneity across cities, with between city variation accounting for nearly 90 percent of the overall variation in the measure. The pattern is somewhat different for the crime data, with a roughly equal proportion of the variation arising between and within cities.

Second, the vast majority (91 percent) of crimes are property crimes with the most serious crimes (murder and rape) comprising less than 1 percent of all crimes reported to police. It is likewise important to note that each of the crime aggregates is dominated by a particular crime type, with assault comprising nearly half of all violent crimes and larceny comprising nearly sixty percent of all property crimes. This is particularly problematic since these are the two crime categories that are generally believed to be the least comparable across jurisdictions and time periods. Third, and turning to the growth rates, perhaps the most relevant feature of the data is that taking first differences of the series comes close to eliminating time invariant cross-sectional heterogeneity in log crime and log police. For each measure of crime and police, the within standard deviation in growth rates is essentially equal to the overall standard deviation. Moreover, in results not shown, the first difference of a log per capita measure exhibits essentially no cross-sectional heterogeneity.

Because of the prominence of the growth rate in police for our analysis, it is of interest to examine the marginal distribution of the growth rate in police for the UCR data and the ASG data separately. Both series exhibit a mass point at zero. In the UCR data, roughly 3.9 percent of the population-weighted observations have a growth rate of zero. The corresponding figure in the ASG data is 6.1 percent. Figure 7 presents estimates of the conditional density function for the growth rate in police, conditional on not being zero.[60] The figure indicates that the growth rate in police is roughly symmetric with a range of approximately -8 to 12 percent for both series. Compared to the UCR series, the ASG data has a greater prevalence of zero growth rates and a greater prevalence of extreme growth rates. For reference, the figure also shows normal density curves. These are generally close to the local linear density estimates.

Figure 8 highlights long-run trends in crime and police for our sample of 242 cities as well as for all cities in the United States, 1960-2010. The dotted lines in Panels A present the time series for total violent crimes per 100 thousand persons while the solid lines present the time series for cost of violent crimes per person.[61] Panel B presents the same time series evidence for property crimes while Panel C presents the time series for total sworn officers. Focusing on the trends among our sample of cities, we see that regardless of whether crimes are cost-weighted, the series show a remarkable 30 year rise in criminality from 1960 to 1990, followed by an equally remarkable 20 year decline in criminality from 1990 to 2010. These swings are spectacular in magnitude. For violent crime, costs in 2010 dollars per person rose from $500 in 1960 to $2,000 in 1990 before falling to less than $1,000 in 2010. For property crime, costs in 2010 dollars per person rose from less than $50 in 1960 to nearly $150 in 1990 before falling to just above $50 in 2010. Notably, our sample of cities, which covers approximately one third of the U.S. population over the 1960-2010 time period, closely parallels national trends.

Trends in policing in our sample of cities also closely track trends in policing nationally. The 1960s is a decade of strong gains in police strength, from roughly 160 officers per capita to just over 250 officers per capita, with some acceleration evident after the wave of riots in the period 1965-1968, followed by a slower rate of increase during the first half of the 1970s. During the second half of the 1970s, we see an era of retrenchment, perhaps related to urban fiscal problems. From 1980 to 2000, sworn police generally increase, with particularly strong increases in the 1990s. Since 2000 the numbers are roughly flat, with the exception of 2003, which is driven entirely by the erroneous estimate provided by the New York City Police Department to the UCR program (cf., Figure 5).

VII. Results

A. *Main Results*

To estimate the police elasticity of crime correcting for measurement error, we utilize IV estimates where one noisy measure of police is an instrument for another

[60]The conditional density function estimates are based on local linear density estimation (Fan and Gijbels 1996) and use a binsize of $b = 0.005$, a bandwidth of $h = 0.025$, and the Epanechnikov kernel. See McCrary (2008) for discussion of this density estimation technique and an application.

[61]This is simply the cost-weighted sum of crimes, computed for the subset of violent crimes, relative to the number of persons and is presented in units of dollars per person.

noisy measure. The logical starting point for this analysis is then an examination of the extent to which the UCR and ASG measures of the growth rate in police are related. The first two columns of Table 3 present coefficients and standard errors from models in which the growth rate in the UCR measure is regressed on the growth rate in the ASG measure. These models correspond to what we term our forward IV regressions, in which the UCR measure is the endogenous regressor and the ASG measure is the instrument. The final two columns correspond to what we term our reflected regressions, in which the roles are reversed, with the UCR measure as the endogenous regressor and the ASG measure as the instrument.

Column (1) presents a regression of the growth rate in the UCR measure on the growth rate in the ASG measure, conditional on two measures of the growth rate in the city's population (one from the UCR file and one from the ASG file) as well as a vector of year effects. In the interest of simplicity, we refer to including both population measures as "controlling for population" throughout the paper. In column (2), we condition on state-by-year effects. These capture the effect of any potential covariate that varies over time at the state-level, such as state welfare policy, penal policy, or education policy.[62]

Consistent with the scatterplots presented in Figure 5, the coefficients reported in Table 3 are relatively small in magnitude, indicating that both the UCR measure and the ASG measure contain a great deal of noise once measured in growth rates. Referring for example, to column (1) of Table 3, we observe that, conditional on the growth rate in population, a 10 percent increase in the ASG measure is associated with only a 1.8 percent increase in the UCR measure. Column (2) shows that this result is robust to the inclusion of the full set of state-by-year effects with the coefficient value falling by roughly 10 percent from 0.18 to 0.16.

Turning to columns (3) and (4), which present the results from the reflected first stage regressions, we see that these coefficients are substantially larger in magnitude than the coefficients in columns (1) and (2). These differing magnitudes are expected since the UCR measure of police growth rates exhibits less variance than the ASG measure, and since the first stage coefficient is the covariance between the two measures, relative to the variance of the predicting variable. As with the forward first stage regressions, results differ only slightly when the state-by-year effects are added.[63]

The F-statistic on the excluded police measure is reported below the coefficient estimates. Since the sample size only affects the scaled distribution of the IV estimator

[62] In Table 3, and in subsequent tables, we report Huber-Eicker-White standard errors that are robust to heteroskedasticity. We note that the heteroskedasticity robust standard errors are extremely similar in magnitude to robust standard errors, clustered at the city level. We favor the robust standard errors as they are generally slightly larger in magnitude.

[63] First stage results are extremely similar when we condition additionally on a large number of local-level control variables.

through its impact on the F-statistic, it is often said that the F-statistic is the "effective sample size" of the IV estimator (Rothenberger 1984, Section 6). Since the F-statistics we report are all above 140, standard asymptotic approximations will be highly accurate in the context of our application (cf., Bound, Jaeger and Baker 1995). That is, weak instruments are not a concern in this context.

In Table 4, we present estimates of the police elasticity. The first four columns correspond to least squares models in which we regress the growth rate in crime on the growth rate in police, conditioning on population growth and either year or state-by-year effects. The final four columns of Table 4 correspond to IV regressions that are robust to measurement errors in either of the two police series. Elasticities are estimated for each of the seven index crimes as well as three crime aggregates—violent crimes, property crimes and the cost-weighted crime index.[64]

Turning to column (1) of Table 4, we see that using the UCR measure of police officers, the police elasticity of crime is largest for murder (-0.27), motor vehicle theft (-0.19) and robbery (-0.18). All three elasticities are statistically significant at conventional significance levels. Overall, the elasticity is greater for violent crime (-0.12) than for property crime (-0.07).[65] Reflecting the large weight on murder, the cost-weighted

[64]An alternative to using one measure as an instrument for the other is to try to restrict attention to observations that do not contain obvious errors. For example, out of our primary sample of 10,589 observations, roughly 1,000 are either zero (potentially consistent with simply filling out the survey with a copy of the numbers for last year) or are consistent with a growth rate in excess of 20 percent in absolute value (potentially consistent with a gross error such as New York in 2003). This approach is only somewhat successful in our application. For example, the OLS regression of the growth rate of murder on the UCR measure of the growth rate in police is -0.204 in the primary sample and -0.359 in the restricted sample of 9,616 observations where the UCR measure is neither zero nor larger than 0.2 in magnitude. The IV estimate using the ASG as an instrument is -0.889, or more than twice as large as the estimate from the restricted sample (all three estimates control for two measures of population growth rates and state-by-year effects). If we perform the same analysis with the ASG measure as the endogenous regressor, the analogous three estimates are -0.143, -0.171, and -0.572.

[65]In a recent working paper, Solon, Haider and Wooldridge (2012) note that using weighted least squares will not necessarily estimate the average partial effect in the presence of unmodeled heterogeneous effects. They suggest an alternate procedure whereby population is interacted with the main effect of interest. As a robustness check, we re-estimate the population-weighted estimates in Table 4 using this formulation, centering population around the population of the city in which a typical individual lives in our sample, which we write as $\overline{w}$, and including the population weight as an additional regressor. Under a linear approximation to the heterogeneity, i.e., $\theta(W_i) = \theta(\overline{w}) + (W_i - \overline{w})\theta'(\overline{w})$, where the prime indicates differentiation, the coefficient on the growth rate in police represents the average partial effect. The estimates we obtain are similar to those reported in Table 4, but slightly less negative for the forward estimates and somewhat more negative for the reflected estimates. For example, for violent crime, we obtain forward and reflected estimates (standard errors) of -0.123 (0.042) and -0.092 (0.037) for violent crime and -0.049 (0.030) and -0.030 (0.026) for property crime, respectively. The degree of similarity between these results and those in columns (2) and (4) of Table 4 provide little evidence in favor of important unmodeled heterogeneity in our primary models, and since the effects

crime elasticity is -0.21 indicating that a ten percent increase in police is associated with a two percent decline in the cost of crime to victims. Referring to column (2), the estimated elasticities are largely similar when the full set of state-by-year effects are included in the model. Here, the elasticities are generally smaller though of the same order of magnitude. Conditioning on the state-by-year effects, the largest elasticities are for murder (-0.20), robbery (-0.20), and motor vehicle theft (-0.13). Elasticities for the aggregates are -0.12 for violent crimes, -0.06 for property crimes, and -0.14 for the cost-weighted crime index.

Columns (3) and (4) report results for models in which the growth rate in crimes is regressed on the growth rate in the ASG measure of police. To our knowledge, this is the first time a city-level panel data regression of crime on the ASG measure of police has been run.Marvell and Moody (1996) use the ASG police measure in regressions of the growth rate in crime on the growth rate in police at the state level. While the coefficients in columns (3) and (4) are smaller in magnitude, they are also more precisely estimated with significant coefficients for murder (-0.15), motor vehicle theft (-0.11), and robbery (-0.09). While the violent crime elasticity (-0.05) remains significant, the property crime elasticity (-0.03) is no longer significant. Note that the smaller magnitude of the reduced form coefficient in columns (6)-(10) is expected; returning to equation (20), we recall that the degree of attenuation is greater when the reliability ratio is smaller, and the reliability ratio of the ASG measure is worse than that of the UCR measure. These elasticities are largely similar when the full set of state-by-year effects are included in column (4) with the exception of motor vehicle theft which falls by roughly half.

Taken as a whole, least squares estimates of the elasticity of crime with respect to police point to a persistent but modest relationship between changes in police and criminal activity. Regardless of whether we rely on the UCR or ASG measure, a 10 percent increase in the size of a city's police force (which would correpond to a large and costly change in the policy regime) is predicted to lead to only a 1 percent reduction in the rate of violent and property crimes.

In the final four columns of Table 4 we report IV estimates of each crime elasticity that correct for measurement error. These estimates are typically five times larger in magnitude than those estimated via least squares.[66] Referring to column (5), the largest elasticites are those for murder (-0.80), motor vehicle theft (-0.59), robbery (-0.46) and burglary (-0.22). In addition, we report elasticities for each of the two crime aggregates of -0.29 for violent crimes and -0.15 for property crimes, though the latter is not precisely estimated. The elasticity with respect to cost-weighted crimes is

---

are opposite for forward and reflected models, this does not change our pooled estimates importantly.

[66]A familiar result is that the IV estimate can be recovered by dividing the "reduced form" estimate of the police elasticites in Table 4 by the first stage estimate presented in Table 3. In this context, to recover the forward IV coefficients presented in columns (5) and (6) of Table 4, we would divide the reflected least squares coefficients in columns (3) and (4) by the relevant first stage coefficient.

-0.61. The elasticities arising from the reflected IV regressions reported in column (7) exhibit a similar pattern with elasticities for murder, motor vehicle theft and robbery of -0.74, -0.51 and -0.49, respectively. Elasticities for the crime aggregates are -0.32 for violent crimes and -0.20 for property crimes.

Finally, in columns (6) and (8), we present IV results that condition on state-by-year effects. Here we report a violent crime elasticity that is approximately -0.35 and a property crime elasticity that is approximately -0.17. Depending on whether the forward or reflected estimates are used, the cost-weighted crime elasticity is between -0.40 and -0.61. With regard to the individual crimes, elasticities are largest for murder (between -0.57 and -0.89), robbery (between -0.52 and -0.57), motor vehicle theft (between -0.30 and -0.37) and burglary (between -0.17 and -0.34). While the coefficient on robbery does not change appreciably when conditioning on state-by-year effects, coefficients on motor vehicle theft are approximately 30 to 50 percent smaller with the inclusion of the unrestricted state-by-year effects as compared to the standard first differencing specification. We interpret this as evidence in favor of the presence of substantial time-varying unobserved heterogeneity at the state-level.

In Table 5, we present GMM and EL estimates of the elasticity of crime with respect to police. These estimates combine the information from the forward and reflected IV estimates presented in Table 4. For each crime type, the table reports an elasticity conditional on population growth and state-by-year effects. As before, robust standard errors are presented in parentheses.

The table shows that two-step GMM is more precise than, but hardly differs from, one-step GMM, and that EL and GMM are nearly indistinguishable. The two-step GMM estimates are -0.67 for murder, -0.56 for robbery, -0.34 for motor vehicle theft and -0.23 for burglary. With regard to the crime aggregates, we report an elasticity of -0.34 for violent crimes, -0.17 for property crimes and -0.47 for the cost-weighted crime index. These estimates represent our best guess regarding the police elasticity and are our preferred estimates.[67]

[67] An alternative approach is to specify a distributional assumption for S_i^* and the errors in the model. Under normality and mutual independence of S_i^*, u_i, v_i, and ε_i, imposing zero means for u_i, v_i, and ε_i, but allowing a non-zero mean of μ_* for S_i^*, we obtain a log likelihood function of the form

$$\begin{aligned} L_i(\beta) &= \frac{1}{2}\ln(\omega_Y) + \frac{1}{2}\ln(\omega_S) + \frac{1}{2}\ln(\omega_Z) + \frac{1}{2}\ln(\omega_*) - 2\ln(2\pi) - \frac{1}{2}\ln\left(\theta^2\omega_Y + \omega_S + \omega_Z + \omega_*\right) \\ &- \frac{1}{2}S_i^2\omega_S - \frac{1}{2}Z_i^2\omega_Z - \frac{1}{2}\mu_*^2\omega_* - \frac{1}{2}(Y_i - \gamma X_i)^2\omega_Y + \frac{1}{2}\frac{\{\theta(Y_i - \gamma X_i)\omega_Y + S_i\omega_S + Z_i\omega_Z + \mu_*\omega_*\}^2}{\theta^2\omega_Y + \omega_S + \omega_Z + \omega_*} \end{aligned}$$

where $\omega_j = 1/\sigma_j^2$ for $j \in \{Y, S, Z, *\}$, corresponding to ε_i, u_i, v_i, and S_i^*, respectively, and where $\beta = (\theta, \gamma, \omega_Y, \omega_S, \omega_Z, \omega_*, \mu_*)$. The normal likelihood approach implicitly forms an estimate of S_i^*, given by a linear combination (call it μ_i) of S_i, Z_i, and $Y_i - \gamma X_i$, and imposes orthogonality conditions akin to those for a regression of Y_i on μ_i and X_i, but adjusted for the fact that μ_i is a generated regressor. To economize

In the bottom panel of Table 5, we report Hansen's J-test of overidentifying restrictions, which provides a measure of the discrepancy between the two parameter estimates.[68] Under the null hypothesis of classical measurement error, the test statistic has a χ^2 distribution with one degree of freedom, which has a 95 percent critical value of 3.84. Table 5 reveals that we fail to reject the null hypothesis of classical measurement errors in each of ten tests. In fact, the largest of these test statistics is just 1.86. We thus interpret the differences in the IV coefficients reported in columns (6) and (8) of Table 4 as providing little evidence against the classical measurement error hypothesis.

Generally speaking, Hansen's J is an omnibus test. Some insight into what aspects of the classical measurement error model are being tested by Hansen's J can be obtained by examining the null hypothesis in more detail. Abstracting from covariates and using the indirect least squares interpretation of IV, the null hypothesis for Hansen's J in this context is that the two IV estimators share a probability limit, or

$$\frac{\mathbb{C}[Y_i, Z_i]}{\mathbb{C}[S_i, Z_i]} = \frac{\mathbb{C}[Y_i, S_i]}{\mathbb{C}[Z_i, S_i]} \tag{25}$$

Since the denominators for the two ratios in equation (25) are the same, the ratios can only be equal if the numerators are. Hansen's J-test is thus a convenient way to test equality of covariances, which is implied by the classical measurement error model since it implies that both covariances simplify to $\mathbb{C}[Y_i, S_i^*]$.

A second characterization of Hansen's J suggests other testing possibilities as well. Write the null hypothesis for Hansen's J as

$$0 = \frac{\mathbb{C}[Y_i, Z_i]}{\mathbb{C}[S_i, Z_i]} - \frac{\mathbb{C}[Y_i, S_i]}{\mathbb{C}[Z_i, S_i]} = \frac{\mathbb{C}[Y_i, Z_i - S_i]}{\mathbb{C}[S_i, Z_i]} \iff 0 = \frac{\mathbb{C}[Z_i - S_i, Y_i]}{\mathbb{V}[Y_i]} \tag{26}$$

where the logical equivalence follows since the ratio can only be zero if the numerator is zero. This latter characterization emphasizes that the null hypothesis for Hansen's J-test can also be understood as the requirement that the outcome not predict the difference in measures. This is implied by the classical measurement error model because the difference in measures is supposed to reflect only the difference in measurement

on computing time, we apply the MLE to data de-meaned by state-year, just as with EL. This approach yields point estimates (standard errors) for the 10 crime categories in Table 5 of -0.614 (0.225), -0.233 (0.212), -0.530 (0.111), -0.101 (0.122), -0.207 (0.085), -0.079 (0.064), -0.331 (0.097), -0.327 (0.085), -0.166 (0.059), and -0.433 (0.166). Between the MLE estimates and the EL estimates, we favor the EL estimates because they are consistent under a weaker set of assumptions. Between the EL and GMM estimates, we observe small enough differences that in this application the distinction seems academic.

[68] Here, the test statistic is computed via two-step GMM. The results are nearly identical when the test is computed using an EL approach. Because we are unwilling to assert that the variance matrix of the errors is spherical, the two-step GMM estimator is no longer the efficient estimator in its class, which implies that the test of over-identifying restrictions is not equal to the minimized value of the objective function. However, the proper test statistic can nonetheless be constructed; see Newey (1985) for a discussion and the proper formula for this case.

errors, and each measurement error is supposed to be uncorrelated with the signal and the controls, which is $(A2)$, and with the structural error term, which is $(A1)$.

This is a helpful characterization because it clarifies what aspects of the classical measurement error model can and cannot be tested using Hansen's *J*-test. Hansen's *J* evidently does not test the validity of $(A3)$. This makes sense, because if $(A1)$ and $(A2)$ hold, but $(A3)$ does not, both IV estimators measure the incorrect quantity of $\mathbb{C}[Y_i, S_i^*]/(\mathbb{V}[S_i^*] + \mathbb{C}[u_i, v_i])$.

However, the analysis above suggests a natural method for testing $(A3)$ that takes advantage of the fact that for some years we possess a third measure of police from the LEMAS survey. Specifically, with 3 measures of police, we can see whether the difference between any two measures is predictable using a third measure. Since a third measure, say $\widetilde{Z}_i$, can be written as $\widetilde{Z}_i = S_i^* + \widetilde{v}_i$, where the same properties are asserted to hold for $\widetilde{v}_i$ as for u_i and v_i, $S_i - Z_i$ should be unrelated to $\widetilde{Z}_i$, $S_i - \widetilde{Z}_i$ should be unrelated to Z_i, and $Z_i - \widetilde{Z}_i$ should be unrelated to S_i. This method of testing $(A3)$ is really a joint test of $(A2)$ and $(A3)$, since each measure reflects both the signal and the measurement error. Access to a fourth measure would of course make such an approach even more powerful, but that is infeasible in our application.

Tests along these lines are presented in Table 6. These tests partially take advantage of the fact that, for selected years, we have three measures of police taken from the UCR, ASG, and LEMAS measurement systems, as discussed in Section VI, above.[69] Each column of Table 6 pertains to regressions of the the difference between the growth rates for two police measures. Column (1) pertains to the difference between the growth rate in the UCR and ASG measures for the full 1960-2010 sample. Columns (2)-(4) use only the subsample of years for which the LEMAS measure is available, with column (2) pertaining to the UCR and ASG series, column (3) pertaining to the UCR and LEMAS series, and column (4) pertaining to the LEMAS and ASG series. Each column of Table 6 presents coefficients from a regression of the growth rate in the measurement error gap on three categories of covariates: the growth rate in each of the seven index crimes (Panel A), the growth rate in the remaining police measure (Panel B) and the growth rate in each of our two population measures (Panel C).

Referring to Panel A, using the full sample in column (1), we find little evidence of a relationship between the growth rate in the measurement errors and the growth rate in crime for any of the seven index crimes. Of the seven t-ratios, only one is above 1 in magnitude. Columns (2)-(4) provide twenty-one tests of this hypothesis using only the subsample for which the LEMAS measure was collected. Each of these three columns uses a particular difference in measures as the dependent variable: $S_i - Z_i$, $S_i - \widetilde{Z}_i$, and $Z_i - \widetilde{Z}_i$. None of the 21 t-ratios in these columns in Panel A give evidence against the restrictions of the classical measurement error model. As noted, these t-ratio tests

[69]For a more detailed discussion of the LEMAS data, see the Data Appendix.

amount to joint tests of Assumptions ($A1$) and ($A2$), because crime growth rates reflect both the structural error ε_i and the signal S_i^*.

Panel B of Table 6 presents coefficients and standard errors from a regression of a difference in police measures on the police measure not involved in the difference (e.g., $S_i - Z_i$ being regressed on $\widetilde{Z}_i$). These are tests of Assumptions $(A2)$ and $(A3)$, because under the classical measurement error model, $S_i - Z_i$ is simply a difference in measurement errors, and the third measure reflects both the signal and a third measurement error. The results in this panel may contain some slight evidence against the classical measurement error model. Specifically, one of the three tests (UCR-LEMAS) rejects at the 1 percent level and this may be consistent with mean-reverting measurement error. On the other hand, the other two tests in Panel B provide little evidence against the classical measurement error model at the 5 percent level. More broadly, the magnitude of the covariance seems to be quite small—a 10 percent increase in the growth rate of a given police measure is associated with only a 0.8 percent change in the measurement error.

In Panel C, we present results from a series of regressions of the growth rate in the measurement errors on the growth rate of each of our two population measures. In all cases, we find little evidence of a systematic relationship between measurement errors and population growth rates.

Finally, in the bottom panel of Table 6, we present p-values from a series of F-tests on the joint significance of all of the variables in predicting the growth rate in the measurement errors. For the full sample, we fail to reject (p-value = 0.83) that the measurement errors are unrelated to crime, police, and population. For the LEMAS subsample, we fail to reject the null hypothesis in all three cases (p-values = 0.25, 0.17, and 0.07).

Overall, we interpret the evidence in Table 6 as furnishing little evidence against the assumptions of the classical measurement error model. There are 39 total tests presented in Table 6; only one of these tests rejects at the 5 percent level, and no joint test is significant at the 5 percent level.[70]

However, since these tests are not commonly used in the literature, there is a question regarding how powerful these tests are at detecting violations of the classical measurement error model. To address this point, we conducted a small simulation study pegged to our sample. We generate simulated data $(Y_i, S_i, Z_i, \widetilde{Z}_i)$ as

$$Y_i = \theta S_i^* + \varepsilon_i \tag{27}$$

$$S_i = \lambda_1 S_i^* + u_i \tag{28}$$

$$Z_i = \lambda_2 S_i^* + v_i \tag{29}$$

$$\widetilde{Z}_i = \lambda_3 S_i^* + \widetilde{v}_i \tag{30}$$

[70]While these tests are not independent, we note that a plot of the quantiles of the 39 t-ratios in Table 6 against the standard normal quantiles indicates similar distributions. Relatedly, the one-sample Kolmogorov-Smirnov test statistic (versus the standard normal distribution) has a p-value of 0.18.

where the vector $(S_i^*, \varepsilon_i, u_i, v_i, \widetilde{v}_i)$ is distributed jointly normal with zero mean and standard deviations calibrated to match key features of our data.[71] In the simulations, we allow five parameters of the data generating process (DGP) to vary: ρ_1, λ_1, λ_2, λ_3, and ρ_3, where ρ_1 is the (constant) correlation between u_i and ε_i, between v_i and ε_i, and between $\widetilde{v}_i$ and ε_i, and where ρ_3 is the (constant) correlation between u_i and v_i, between u_i and $\widetilde{v}_i$, and between v_i and $\widetilde{v}_i$. These parameters control the covariances among the elements of the vector $(S_i^*, \varepsilon_i, u_i, v_i, \widetilde{v}_i)$.

Note that when $\rho_1 = \rho_3 = 0$ and $\lambda_1 = \lambda_2 = \lambda_3 = 1$, the DGP is consistent with the classical measurement error hypothesis. The parameter ρ_1 indexes the extent to which Assumption ($A1$) is violated; λ_1, λ_2, and λ_3 index the extent to which Assumption ($A2$) is violated; and ρ_3 indexes the extent to which Assumption ($A3$) is violated.[72] We maintain Assumption ($A4$) throughout. For each of 10,000 simulated data sets, we construct the tests performed in Table 6 and record whether the null hypothesis was rejected.[73] This allows us to examine the power of these tests against specific alternatives.

The results of this analysis are shown in Figure 9, which contains four panels. Each panel shows the impact of a departure from the classical measurement error model on the rejection rate for two tests ("Test A" and "Test B"). Test A is a t-ratio test in a bivariate regression of either $S_i - Z_i$, $S_i - \widetilde{Z}_i$, or $Z_i - \widetilde{Z}_i$ on an outcome Y_i (i.e., a test of the type discussed in Table 6, Panel A), and Test B is a t-ratio test where the covariate is not Y_i but a third measure of police (i.e., a test of the type discussed in Table 6, Panel B). The four panels in Figure 9 vary ρ_1, λ_1 and λ_2, and ρ_3, relative to the baseline of the classical measurement error model. The curves displayed are power curves corresponding to the tests which have power against the alternative being displayed. For reference, each panel also shows the average of the simulated GMM estimates. The true parameter in all scenarios is -0.5.

The figure shows that these tests have generally good power. For example, turning to Panel A, if the correlation between a measurement error and the structural error is 0.05, the rejection probability for Test A is roughly 30 percent. This is important, because even a small degree of correlation between a measurement error and the structural error leads to bias. The power of Tests A is very good for column 1, where we have our full sample size, but it is notably lower for columns 2 through 4. Our sense is that

[71] We set $\sigma_* = 0.044$, $\sigma_\varepsilon = 0.260$, $\sigma_u = 0.047$, and $\sigma_v = 0.070$, and $\sigma_{\widetilde{v}} = 0.055$. This roughly matches the root mean squared error from IV models for the cost-weighted sum of crimes corresponding to Table 4, the first stage coefficients in Table 3 that condition on state-by-year effects, and the standard deviations of the various police measures after demeaning by state-year.

[72] Throughout, we maintain zero correlation between $(u_i, v_i, \widetilde{v}_i)$ and S_i^*. The parameters λ_j control the extent to which a composite error such as $u_i + (\lambda_1 - 1)S_i^*$ is correlated with S_i^*, where $S_i \equiv S_i^* + u_i + (\lambda_1 - 1)S_i^*$ and analogously for Z_i and $\widetilde{Z}_i$.

[73] To match our tests from Table 6, tests corresponding to column 1 are based on simulated data sets of size $n = 10,589$ and tests corresponding to columns 2 through 4 are based on $n = 1,752$.

the measurement errors are unlikely to be correlated with the structural error, because we did not observe any rejections in any of the 28 tests in Panel A of Table 6, even those in column 1 where this test has quite good power.

Turning to the results in Panels B and C, we see that mean-reverting measurement error is quite likely to be detected as λ_1 or λ_2 depart from 1. Importantly, both Test A and Test B may detect mean-reverting measurement error. The curve labeled "A-any" is the power of a test which rejects at the 5 percent level if and only if one or more of the four Tests A reject at the $0.05/4 = 0.0125$ level. For a single crime outcome, this test has power approaching 20 percent for λ_1 or λ_2 equal to 0.7. We suspect that mean-reverting measurement errors in our data would thus be detected more decisively, either by rejections in columns 2, 3, and 4 of Table 6, or by at least threshold rejections for one or more crime categories.

Finally, in Panel D, we examine the power of Test B against alternatives rooted in correlated measurement errors. It is conceivable that the same core (mismeasured) information informs both the ASG and UCR measures of police. We suspect this happens rarely, as the UCR forms are filled out by employees of the police department and signed by the police chief, whereas the ASG forms are filled out by the mayor's office or city manager's office. However, it is of course true that the mayor could contact the police department for the information, in which case any measurement errors would be positively correlated. Nonetheless, Test B has power to detect correlated measurement errors. We note that to the extent the measurement errors in police are positively associated, we would understate the true effect of police on crime (cf., the expectation of the GMM estimates presented in Panel D).[74]

B. Robustness

Before turning to a discussion of the results presented above, we consider several robustness checks. The estimates in Tables 4 and 5 assume the exogeneity of police conditional on population growth and state-by-year effects. While state-by-year effects soak

[74] Of course, as with any specification test, there will be a lack of power in specific directions. We have examined the power to detect local departures from the classical measurement error model, but one could instead imagine joint departures, and our tests will have little power against some of these joint alternatives. For example, if $(A2)$ is violated, but the λ_j parameters differ from 1 by the exact same amount, then the rejection rate for both Test A and Test B will be 5 percent. Similarly, if $(A1)$ is violated, but the measurement errors have the exact same covariance with the structural error, then the rejection rate for Test A will be 5 percent regardless of how large is the covariance with the structural error. This underscores, in our minds, the importance of validation studies based on administrative, rather than survey, data. We note that progress in labor supply and in the return to education, for example, occurred after several decades of hard work spent documenting fundamental properties of the measurement errors in the standard data sources on hours, earnings, and education. For a review of some of this literature, see Bound, Brown and Mathiowetz 2001. Much more research along these lines is needed to obtain a clear picture of the proper inferences to be drawn from the crime literature.

up important time-varying state-level variation, results will nevertheless be inconsistent if there are time-varying covariates measured at the city-level which are correlated with both growth in police and crime. In Table 1, above, we presented evidence that the growth rate in police is correlated with the growth rate in a number of city- and county-level covariates to only a limited degree. In Table 7, we explore the extent to which elasticities reported in Table 4 are robust to the inclusion of city-level covariates directly. The cost of this more direct analysis is that we are required for data availability reasons to restrict attention to the 1970-2002 subsample. The first six columns refer to estimates using the forward models while the last six refer to estimates from the reflected models.

We begin in column (1) by replicating the coefficients presented in column (6) of Table 4 for the 1970-2002 subsample of our data. These estimates condition on population growth and state-by-year effects. For the 1970-2002 subsample, the violent crime elasticity is -0.29 and the property crime elasticity is -0.26. The largest elasticites are for murder, robbery, and burglary (-1.1, -0.55, and -0.41, respectively). The elasticity for the cost-weighted crime index is -0.79. In column (2) we add a series of economic covariates that capture the growth rate in personal income and total employment as well as revenue and employment in four leading industrial sector (construction, manufacturing, wholesale trade and retail trade). We also include a variable that captures each city's public expenditures exclusive of police to capture the impact of all other municipal spending. In column (3), we include the lags of each of these variables to capture a potentially lagged response of crime to local macroeconomic conditions. In column (4), we capture changes in a city's demographic composition by adding control variables for the population share of sixteen age-gender-race groups within each city. In order to control flexibly for the effect of changes in a city's composition, in column (5) we add polynomials (to the second degree) and interactions for each of the demographic subgroups. Finally, in column (6) we add city-specific linear time trends that would capture long-standing crime trends that are independent of growth in police.

Referring to the forward models, it is apparent that the estimated elasticites change very little with the inclusion of the controls. Referring, for example, to the cost-weighted crime index, the estimated elasticity moves from -0.79 when conditioning only on population and state-by-year effects to 0.76 when economic covariates are included. Conditioning also on the lags of the economic covariates brings the estimated elasticity up to -0.82 while controlling extensively for demographics brings the elasticity back to -0.79. When time trends are included, the elasticity increase to -0.82, just 2.5 percent higher than the original elasticity. A similar pattern holds for each of the other crime types with the largest change from column (1) to column (6) occuring for assault and larceny, both of which are imprecisely estimated. Referring to columns (7)-(12), the reflected estimates follow a similar pattern with the exception of murder which appears to be somewhat sensitive to the inclusion of lagegd economic covariates, a result which drives the difference between columns (7) and (12) for violent crimes and for the cost-weighted crime index. While the estimated murder coefficient changes with the inclusion of controls, it is nevertheless similar in magnitude to the estimate that conditions only on the state-by-year effects.[75]

As discussed, the elasticities reported in the paper condition on two measures of city population growth, taken from the UCR and ASG data systems, respectively. The motivation for including both measures is that we are persuaded there is measurement error in each series individually. As discussed in Section VI, above, it is necessary to smooth both series to circumvent clear measurement problems around Census years. However, even after smoothing, it may not be the case that true population growth rates can be represented as a linear combination of the growth rates of the UCR and ASG series. To assess the extent to which measurement error in population represents a source of bias for the estimated police elasticities, we take advantage of two additional proxies for a city's population: (1) population data from the Surveillance Epidemiology and End Results (SEER) dataset which has been compiled by the National Cancer Institute to track disease incidence and (2) the number of births in a city, drawn from the National Center for Health Statistics at the Centers for Disease Control. Births are correlated with population because, other things equal, the more individuals who are living in a city, the more children will be born. Moreover, unlike the UCR and ASG series on population, there is close to no measurement error in the number of births in a city, since births are estimated from no worse than 50 percent samples of birth certificates over the sample period and since birth certificates cover an estimated 99 percent of births in the U.S. over this time period.[76] Consequently, including the growth rate in births as a covariate should pick up on any remaining association between true population growth rates and crime growth rates.

Appendix Table 1 shows the sensitivity of our resulting estimates to the inclusion of SEER population data and the births data. For the years and cities for which data on births are available, the estimates change very little when the growth rate in births is added to the model. For example, referring to Panel A for which data on all cities in the sample are available for the 1960-1993 subsample, we see that pooled estimates of each of the crime elasticites are extremely similar with and without the inclusion of the births measure. For example, the violent crime elasticity moves from -0.16 to -0.18 while the property crime elasticity changes only in the third decimal place.

Panel B which provides estimates for a sample of 147 of our cities using the 1960-

[75] We also consider whether the estimates are robust to the exclusion of the two largest cities in the sample–New York and Los Angeles–as well as whether the results are robust to the exclusion of cities with various data problems, namely those cities which have merged with their respective counties (e.g., Jacksonville, Nashville, Charlotte and Louisville) and cities which have been recently found to have misreported data to the FBI's Uniform Crime Reporting System (e.g., Milwaukee). When these cities are excluded from the sample, the estimates are nearly identical to those reported in Table 5.

[76] In the early 1970s, the NCHS transitioned to 100 percent samples of birth certificates.

2003 window. For that sample, the largest impacts are on murder and motor vehicle theft where the elasticity changes from -0.548 to -0.565 and from -0.346 to -0.369, respectively, when births are included. Finally, in Panel C, we test the sensitivity of the pooled elasticities to inclusion of the SEER population data over the 1970-2008 time period. Again, the estimates are extremely similar when the SEER population measure is added, with motor vehicle theft showing the largest change (-0.320 versus -0.340). We interpret these findings as indicating that our estimates are not importantly compromised by measurement error in population. Indeed, to the extent that our estimates do change when additional population controls are added, they tend to get larger in magnitude with additional population controls, suggesting that our full sample estimates may be conservative.[77]

VIII. Discussion

The estimates reported in the previous section of this paper can be thought of as police elasticities that are robust to errors in the measurement of police. Pooled estimates in Table 5 represent our best guess regarding crime-specific police elasticities. Pooling via GMM or EL, we obtain precisely estimated elasticities of -0.34 for violent crimes and -0.17 for property crimes, with especially large elasticities for murder (-0.67), robbery (-0.56), motor vehicle theft (-0.34) and burglary (-0.23).

In this section, we contextualize these findings by comparing our reported elasticities to those in the prior literature. Table 8 presents selected police elasticities from eight recent papers that use U.S. data. Each of the papers explicitly seeks to correct for simultaneity bias, for which our estimates do not adjust. While these papers do not discuss the possibility of measurement error in police or in population, an IV estimator using exogenous instruments will correct for both simultaneity bias and measurement error bias under the classical measurement error hypothesis.

Looking across the estimates from these papers in Table 8, four tendencies are evident. First, the estimates are generally negative. Some of the estimates are zero (e.g., Levitt 1997) for property crime), but virtually none are positive.[78] Second, the

[77] A final issue that is worth mentioning is the possibility of displacement—an increase in policing in one jurisdiction might displace crime to a nearby jurisdiction. If this is the case, then our approach will tend to overestimate the social value of policing, since part of the apparent crime reduction associated with increasing policing would stem from a simple reshuffling of criminal activity. Appendix Table 2 addresses this concern. In this table, we contrast the GMM estimates presented in Table 5 with estimates based on aggregating up to the MSA level. The estimates in Table 2 indicate that there is not enough statistical power in these data to distinguish the estimates at the city-level from those at the MSA-level. If anything, the estimates at the MSA are larger, rather than smaller, than the estimates at the city level, which is the opposite of the expected pattern if displacement were a first-order phenomenon in these data.

[78] The pooled estimates in Levitt (1997) are in error due to a mistake in the use of weights (McCrary 2002). The numbers in Table 8 listed as Levitt (1997) are actually the corrected numbers reported in McCrary (2002) that use Levitt's mayoral election year series. The numbers in Table 8 listed as McCrary

general tendency of these estimates is similar to, or perhaps slightly larger than, that of our own estimates. For example, the average of the murder elasticities is 1.18 in magnitude.[79] This is similar to the magnitude of our own estimated murder elasticity (roughly 1) when we replace our preferred state-by-year effects specification with a simpler year effects specification, which is more similar to most of the research designs employed in the previous literature. Similarly, the average of the elasticities for robbery, burglary, and auto theft is approximately 0.79, 0.35, and 0.77 in magnitude, respectively. Controlling for year effects, our estimates of the same quantities are roughly 0.50, 0.17, and 0.50, respectively. The differences between these estimates would likely not rise to the level of statistical significance, but the general tendency is for our estimates to be slightly smaller in magnitude. Some of this discrepancy stems from utilization of different time periods. For example, when we restrict our analysis to the years analyzed by Evans and Owens (2007), namely 1990 to 2001, our estimated elasticities are -0.83 and -0.31, for violent and property crimes, respectively. These are extremely close in magnitude to those in Evans and Owens (2007) (-0.99 for violent crimes and -0.26 for property crimes). Given the magnitude of the standard errors, the differences in estimates are likely consistent with the hypothesis of sampling volatility.

Third, there is a general tendency to find that police have a larger protective effect on violent crimes than on property crimes. This is a surprising finding if we conceive of the estimated effect of police on crime as being about deterrence. However, as noted in the introduction, the effect of police on crime operates through both a deterrence and an incapacitation channel. Moreover, police departments actively focus their resources on the incapacitation of individuals posing the greatest risk to society, which may make the incapacitation channel particularly important.

Fourth, the estimated elasticities tend to be quite imprecise, with estimated standard errors ranging from 0.2 to 0.7 for violent crimes and 0.2 to 0.9 for property crimes. As a result, it is often the case that even large elasticites (on the order of 1) cannot be rejected as being different from zero. Similarly, the cross-crime pattern of the elasticities is difficult to discern. For example, one of the more precise studies is that of Evans and Owens (2007). In that study, the magnitude of the estimated elasticities and standard errors suggest that it would be difficult to reject tests of the equality of various crime-specific elasticities. As a result, though the general pattern of the elasticities is suggestive, it is difficult to draw inferences about even the most basic policy questions such as the relative effectiveness of police in reducing violent versus property crimes.

The elasticities we report in this research are estimated with considerably greater precision, with standard errors that are between one-quarter and one-half the size of those

(2002) are the numbers reported in McCrary (2002) that use McCrary's mayoral election year series.

[79]To avoid double-counting research designs, we count the average of the estimates from Levitt (1997) and McCrary (2002) as a single entry.

reported by Evans and Owens (2007) and up to an order of magnitude smaller than those reported in other papers. The result is that we are able to generate considerably stronger inferences regarding the cross-crime pattern of the elasticities.

In Table 9, we formalize this idea and test the equality of all pairs of individual crime elasticities. The table reports p-values from each of these tests, operationalized by stacking up crime categories into a broader GMM system. For a given row, a given column reports the p-value associated with a test of the equality of the coefficient for the crime category on the row and the coefficient for the crime category on the column. The pattern of the resulting p-values suggests that we can be confident that police reduce murder to a greater extent than assault and larceny and perhaps burglary. Likewise, the effect of police on robbery is greater than it is for assault, burglary and larceny and the effect of police on motor vehicle theft is greater than the effect of police on larceny. Referring to the aggregates, the elasticities for murder and robbery are greater than the property crime elasticity. We can also reject, at the 10 percent level, the equality of the violent and property crime elasticities. Despite a dominant pattern in the literature that suggests that the effect of police on crime is most concentrated among violent crimes, to our knowledge, this is the first paper that offers more than suggestive evidence on this point.

Whether our estimates are similar to or different from those in the preceding literature is important for getting the magnitude of police elasticities right, but is also interesting because it speaks to the broader issue of whether simple regression techniques are compromised by simultaneity bias. If our estimates are deemed to be similar to those reported in prior research, then our research implies a smaller role for simultaneity than has been suggested by prior studies.

Overall, our suspicion is that the estimates we have presented here are compromised somewhat by simultaneity bias, despite our best efforts to control for unobserved heterogeneity. The sign of the bias, as criminologists and economists have argued for several decades now, is likely positive, leading our approach to underestimate of the magnitude of the policing elasticity. Thus, the correct magnitude is likely at least as large as what our results indicate. As we turn in the next section to connecting our estimates to the state's optimal level of policing, these considerations should be kept in mind, as they suggest that our policy conclusions may well be conservative.

IX. Cost-Benefit Analysis

A. National Estimates

The results presented in Table 5 represent our best estimate of the elasticity of each type of crime with respect to police. These elasticities allow us to predict the change in reported crimes expected to arise from a given percent increase in the size of a city's police force. However, in allocating scarce resources among a large number of critical public services a potentially more relevant parameter is the ratio of the benefits to the costs of hiring additional police personnel. In Section II, we established that even in the presence of investment in precautions with externalities, the state's optimal choice of policing can be characterized by the the parameter θ, which represents the elasticity of the cost of crime with respect to police, holding precautions fixed. In particular, the rule-of-thumb outlined in Section II is that hiring police improves welfare when

$$|\theta| / \frac{wS}{nC} \equiv \kappa > 1 \tag{31}$$

In this section, we use the GMM approach described above to estimate the ratio of the benefits (as proxied by averted costs to potential victims) to the costs of police.

For a VSL of \$7 million, we estimate a police elasticity of the cost of crime of -0.47 (standard error = 0.17). This elasticity estimate is based on a model including state-by-year effects and two controls for population, analogous to our preferred specification in Table 5. Scaling this elasticity estimate by the ratio of mean victimization costs to mean police expenditures produces an estimate of the 2010 social dollars saved from increasing spending on police by one dollar, or the benefit-cost ratio (BCR).[80] Varying the VSL from \$1 to \$28 million, our GMM estimate of the police elasticity of the cost of crime ranges from -0.32 (standard error = 0.09) to -0.55 (standard error = 0.26).

An unfortunate feature of these types of estimates is that benefit-cost calculations are often extremely sensitive to the monetized value of an averted murder.[81] Figure 10 provides a visual presentation of the findings from this analysis. The figure plots the BCR that follows from this GMM procedure on the vertical axis against possible VSL estimates on the horizontal axis. The change in the BCR is linear with respect to the VSL employed since the VSL is simply the factor by which murders are scaled in the analysis. The BCR ranges from approximately 0.4 at a VSL of \$1 million to approximately 6.0 at a VSL of \$28 million. To further narrow down these estimates, we superimpose a kernel density estimate of the density of the 64 VSL estimates for the U.S. While the estimates vary considerably, approximately 80 percent of the data lies below

[80]To obtain the cost of increasing policing, we take the average of the UCR and ASG counts and scale it by \$130,000, an estimate of the fully-loaded cost of a police officer in 2010. As discussed in detail in Section VI, above, this estimate is based on data on the operating budget per officer, i.e., the ratio of the operating budget for the police department, in 2010 dollars, to the number of sworn officers. These figures are taken from the ASG Finance and ASG Employment files for 2003-2010. We use multiple years to get a clear picture of the finances for a department. These figures fluctuate a good deal from year to year. We use the city-specific median over time, and then compute a 2010 population weighted average of the city-specific medians; this weighted average is about \$130,000. Our estimate of the cost of hiring additional officers is notably higher than those used in some of the literature (e.g., Evans and Owens 2007) use \$55,000).

[81]In the literature, it is not uncommon for the results of a benefit-cost analysis of a given policy to depend on the researcher's choice between two reasonable alternative values of the cost of a murder.

$10 million which is associated with an approximate BCR of 2. At $7 million, the mean value of the VSL, the resulting BCR is 1.63, indicating that, in a typical U.S. city, an additional dollar allocated towards policing is predicted to save $1.63 in costs to crime victims. This would be consistent with classical notions of the underprovision of public goods (Samuelson 1954). On the other hand, as noted there is substantial ambiguity regarding VSL estimates. The estimated VSL from Ashenfelter and Greenstone 2004b) implies a BCR of roughly $0.80, indicating substantial overpolicing.[82] If we revert to estimates from Cohen and Piquero (2008), the BCR is just below 1, suggesting that the political process may have arrived at the social optimum.

B. City-specific Estimates

The national benefit-cost ratios reported above answer the question: For a typical U.S. city in a typical year in our sample, what is the dollar value of crime reduction obtained by increasing spending on police by one dollar? A somewhat different question pertains to specific U.S. cities in 2010. For example, for many years Oakland, California, has had fewer police per capita than other cities, despite a relatively high crime rate. Journalists often note this fact and query whether Oakland should hire additional police (e.g., McKinley 2009). We now seek to answer the question: For specific U.S. cities in 2010, given that the value of a statistical life is $7 million, what is the dollar value of crime reduction obtained by increasing spending on police by one dollar? Despite great interest in such issues, we emphasize that investigations along these lines are necessarily somewhat speculative, as data for individual cities are less reliable than data for a few hundred cities, taken as a whole. Indeed, such an analysis may be heroic, as it involves assuming that the police elasticity of crime is constant across cities, across time, and across possible adjustments to the size of the police force.

These limitations aside, we believe it is nonetheless of interest to characterize the heterogeneity across cities in a benefit-cost ratio, as a function of the prevalence of crime, the number of officers, and the cost to the city of hiring officers. There is extraordinary heterogeneity across cities in the prevalence of crime. For example, the cost-weighted sum of crimes per capita in the most dangerous city in our sample (Gary, Indiana), is nearly 40 times that of the safest city in our sample (Waltham, Massachusetts). Similarly, cities vary quite a lot in terms of the number of officers and the expense per officer. We suspect that our approach, while flawed, captures much of the variation from city to city in the true benefit-cost ratio.

Table 10 presents the bottom and top 30 cities in our sample, based on our estimated benefit-cost ratio. The benefit-cost ratio, reported in the final column (column 9) of the table, is rooted in our overall estimated police elasticity of the crime index for all

[82] In fact, while estimates of the VSL arising from the study of individuals' labor market behavior tend to yield large values (on average, $9.5 million), estimates of the VSL arising from the study of non-labor market behaviors tend to yield much smaller values (on average, $4 million).

cities and all years, but is scaled by the mean cost of crime in the city from 2003-2010 relative to the product of the average number of officers from 2003-2010 and the estimated cost of hiring an additional officer (see Section VI and the Data Appendix for discussion of the construction of this variable). The mean cost of crime is reported in per capita terms in the table (column 6), as is the estimated cost per officer (column 8). Column 5 of the table reports per capita income in the city, and column 7 reports the cost of crime relative to income per capita ("fraction income at risk"). This last column encourages thinking of crime as a tax on the populace. For reference, we additionally report city population as of 2010 and the city's poverty rate.

The 30 cities listed in the top half of the table have the lowest benefit-cost ratios among our 242 cities, while the 30 cities listed in the bottom half have the highest benefit-cost ratios. For example, for Sunnyvale, California, we estimate that every dollar spent on policing yields only 20 cents in benefits in terms of crime reduction. In contrast, we estimate that every dollar spent on policing in Gary, Indiana, yields $14 in benefits in terms of crime reduction. The population weighted average of the city-specific benefit-cost ratio is about $1.78, or slightly higher than our estimate for the overall sample reported above.

Scanning down the table, we see several interesting patterns. Cities with low benefit-cost ratios are small, low-poverty cities, with low levels of crime and low to moderate levels of policing. Police officers in these cities often enjoy high salaries and benefits, leading to high employer costs per officer. Sunnyvale and Berkeley, for example, both have costs per officer of roughly $280,000.

Cities with high benefit-cost ratios are suprisingly representative of our broader sample in some regards. For example, cities with high benefit-cost ratios include both low and high population cities. Also, these cities sometimes have low policing levels (e.g., Oakland and Richmond, California, have 180 and 160 sworn officers per 100,000 population, respectively) and sometimes have high policing levels (e.g., Baltimore and Camden have 480 and 510 sworn officers per 100,000 population). On the other hand, cities with high benefit-cost ratios have high poverty rates and extraordinarily high crime rates—generally an order of magnitude higher than cities with low benefit-cost ratios. Crime costs residents in these cities anywhere from 5 percent of their annual income (Mobile, Alabama) to 34 percent of their annual income (Camden). In contrast, for cities with low benefit-cost ratios, crime costs resident at most 1 percent of their annual income.

Another interesting pattern is that California cities are prevalent among the lowest benefit-cost ratio cities, with 13 out of 30 spots, but also are represented among the highest benefit-cost ratio cities (Oakland, Richmond, and San Bernadino). The estimated cost per officer is very high among California cities generally. High costs per officer keeps several high-crime California cities from being among the highest benefit-cost ratio cities (Sacramento, Vallejo). Richmond's high estimated costs

($240,000) are particularly remarkable given its high poverty rate and low per capita income; wealthy Palo Alto's estimated cost per officer is slightly lower than that for Richmond. Factoring in base salary, overtime, and lump-sum payments, a police officer in Richmond makes an average of $148,000, or six times what a city resident makes.[83] It is worth noting that the estimated cost per officer does not include unfunded pension liabilities for the city, which is an ongoing issue for many cities that may lead the figures for officer expense to be understated (Gralla 2012).

As noted above, there is substantial ambiguity regarding some of the inputs to the city-specific benefit-cost ratios. However, we note that in many cases the benefit-cost ratios are sufficiently extreme that only gross errors in the inputs would alter the conclusion that the benefit-cost ratio was on the wrong side of 1. Our sense is that cities with benefit-cost ratios between 0.5 and 1.5 may well be near the optimal level of policing, but that the many cities outside this band are unlikely to be.

C. Police Incapacitation Effects and the Benefit-Cost Ratio

The estimates in the preceeding sub-sections are valid under the assumption that either (i) the decline in crime resulting from increased police is entirely due to deterrence or that (ii) the cost of incarcerating offenders is fixed in the short run so that the downstream cost of incapacitating offenders need not be counted as a cost of increased police personnel. Here, we re-frame the national benefit-cost analysis, treating the expected increase in incarceration resulting from more police as an additional cost of hiring a new officer. Because we are interested in the short-run costs and benefits of new police hiring, we count only the costs of incarceration that are borne in the first year. We begin with an estimate of the number of arrests per officer. Using our sample of 242 cities, an average officer made between 18.7 and 20.2 arrests in 2010, depending on whether the UCR or ASG officer count is employed.[84] Next, we employ an estimate of the conditional probability of a conviction given an arrest. In 2010, there were 13,120,947 arrests made by police officers in the United States while there were 1,132,290 convictions in state courts and another 81,934 convictions in federal courts. Dividing convictions by arrests yields an estimated conditional probability of a conviction of 9.3 percent. Of defendants sentenced in state courts, 40 percent were sentenced to state prison (with a mean sentence length of 4 years and 11 months), 28 percent were sentenced to a term in local jail (with a mean sentence of 6 months) and the remaining 32 percent were sentenced to a term of probation or an alternate penalty that did not involve incarceration. On average, offenders serve approximately 55 percent of their sentence. Thus, in steady state, a typical officer is associated with 20

[83] Average take-home pay for officers based off of data pulled from the *San Jose Mercury News*. See footnote **??**.

[84] The working assumption here is that a new officer's productivity, and the lost productivity associated with laying off an officer, can be approximated using the productivity of an average officer.

new arrests, 1.85 new convictions and 0.87 incarceration-years.[85] At an incarceration cost of $25,000 per year, each new officer is thus associated with $21,738 in additional costs. Augmenting the salary figure with this estimate yields a benefit-cost estimate of $1.40 using the $7 million estimate of the value of a statistical life.

X. Conclusion

In this paper, we have presented estimates of the elasticity of crime with respect to police for index offenses: murder, rape, robbery, aggravated assault, burglary, larceny excluding motor vehicle theft, and motor vehicle theft. These estimates are based on annual data on year-over-year growth rates in crime and police in a panel data set of 242 cities observed from 1960-2010. Our primary specifications model growth rates in crime as a function of the growth rate in the number of sworn officers, population growth and a full set of state-by-year effects which render our estimates robust to arbitrary changes in state policy, such as penal policy, and other factors affecting cities in the same state similarly. In auxiliary regressions we show that our results are also robust to a wide array of local-level confounders.

We argue that a central problem in estimating the police elasticity of crime is measurement error in the number of police. These measurement errors are unimportant for specifications involving the level or log of police, but are first-order for specifications involving growth rates or city fixed effects. Problems with measurement errors in police have gone unaddressed in the crime literature, but add to a long list of literatures where measurement errors have been shown to be important (Ashenfelter and Krueger 1994, Bound, Brown, Duncan and Rogers 1994, Kim and Solon 2005, Bollinger 2003, Black and SMith 2006, Edlin and Karaca 2006, Nunn 2008).

A recent literature has focused on quasi-experimental estimates of the police elasticity of crime, and many of these estimates solve both for problems with measurement errors in police and for simultaneity bias (for a review, see Levitt and Miles 2006). We add to this literature by addressing the measurement error bias directly, utilizing independent measurements of the number of police departments collected annually by the Census Bureau. Correcting for measurement error increases least squares estimates of the police elasticity of crime by roughly a factor of 5 and results in estimates just slightly smaller than those estimated in the previous quasi-experimental literature. This may suggest a smaller role for simultaneity bias than has previously been emphasized. An advantage of bracketing the issue of simultaneity bias, and focusing instead on correcting for measurement errors, is that we obtain estimates that are demonstrably more precise than those from the previous literature and arguably conservative in

[85] Regarding pre-trial detention, since the average length of time between arrest and sentence is approximately 5 months, we operate under the assumption that if an arrestee does not receive bail, their expected sentence amounts to time served.

magnitude. The received wisdom in this literature, going back to Nagin (1978) and before, is that police departments hire officers during and perhaps even in anticipation of crime waves, leading the police elasticity of crime to be too small in magnitude.

Our best guess regarding the police elasticity of crime is -0.34 for violent crime and -0.17 for property crime. Crime categories where police seem to be most effective are murder, robbery, burglary, and motor vehicle theft, with estimated elasticities of -0.67, -0.56, -0.23, and -0.34, with standard errors of roughly 0.2 for murder and 0.1 for other crimes. The elasticity of the cost-weighted crime index is -0.47 with a standard error of 0.17.

To assess whether these magnitudes are small or large, we introduced a framework for assessing whether policing levels are socially desirable. This framework delivers a rule-of-thumb for optimal policing that pertains to a social planner unwilling to monitor precautions undertaken by individuals, where the social planner pays for policing using lump-sum taxes. This analysis, together with our empirical estimates of the police elasticity of the cost of crime, is suggestive of substantial underpolicing. However, as with many analyses pertaining to public investments and safety, our normative conclusions turn to a great extent on the price society is willing to pay for reductions in the probability of fatalities, or the value of a statistical life (VSL).

Despite the ambiguity regarding the appropriate quantity for the VSL, we note that federal and state regulatory authorities frequently undertake investments where the same tradeoff is confronted. Pegging policing investments to the typical federal standard suggests that society would receive approximately \$1.60 in benefits from an additional 2010 dollar spent on policing. This estimate is likely conservative if simultaneity bias in the police elasticity of crime is important.

This policy conclusion is most strongly justified if externalities in precautions are unimportant or if there is little scope for policing crowding out precautions. Precautions may however be important. Nonetheless, the framework we have introduced allows us to be clear about the assumptions supporting our policy conclusions in such a scenario. If precautions have negative externalities on average, with one individual's precaution displacing criminal activity to her neighbors, then the policy conclusion is conservative. If precautions have positive externalities on average, with one individual's precaution protecting her neighbors against criminal activity, then this channel has to be fully one-third as large as the direct effect of police on crime in order to overturn the policy conclusion.

Figure 1. Private Precaution Reaction Functions: Two Person Case, Low and High Public Policing

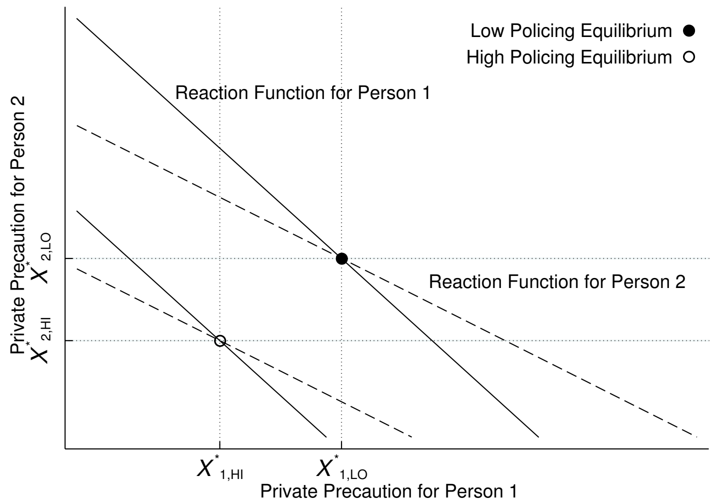
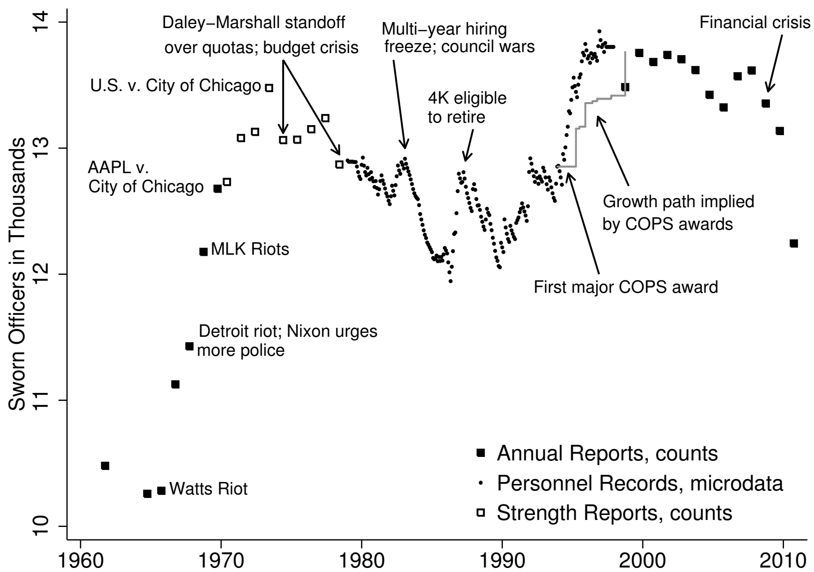


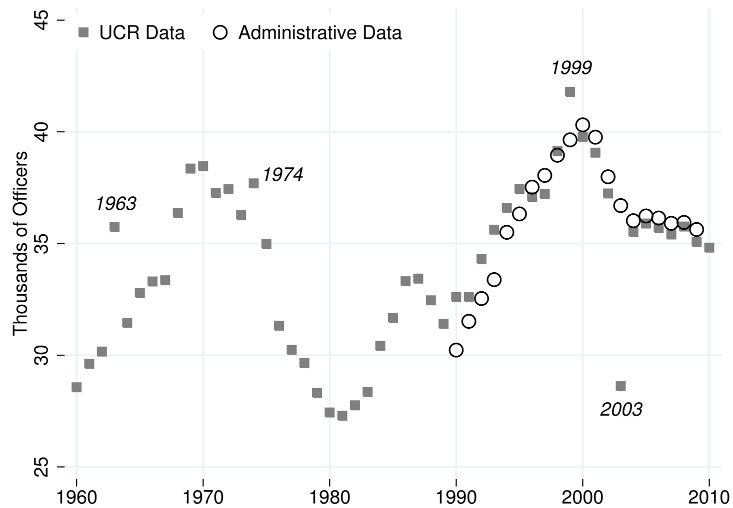
Figure 2. Police Hiring in Chicago, 1960-2010



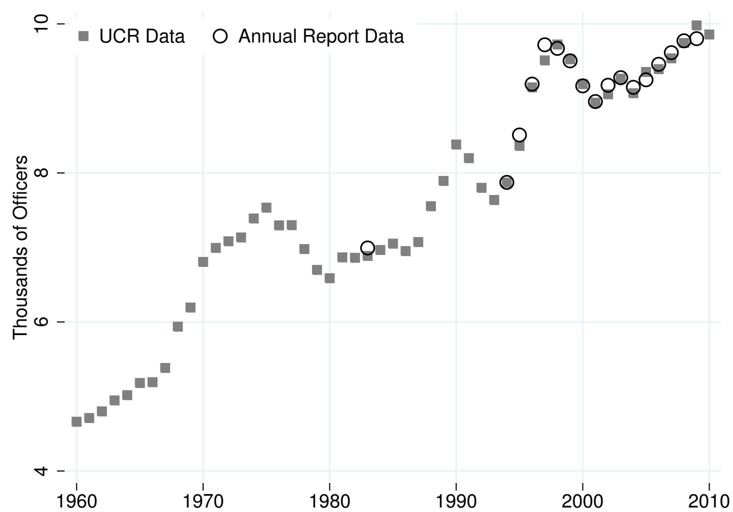
Note: See text for details.

Figure 3. Sworn Officers in Five Cities: the Uniform Crime Reports and Direct Measures from Departments

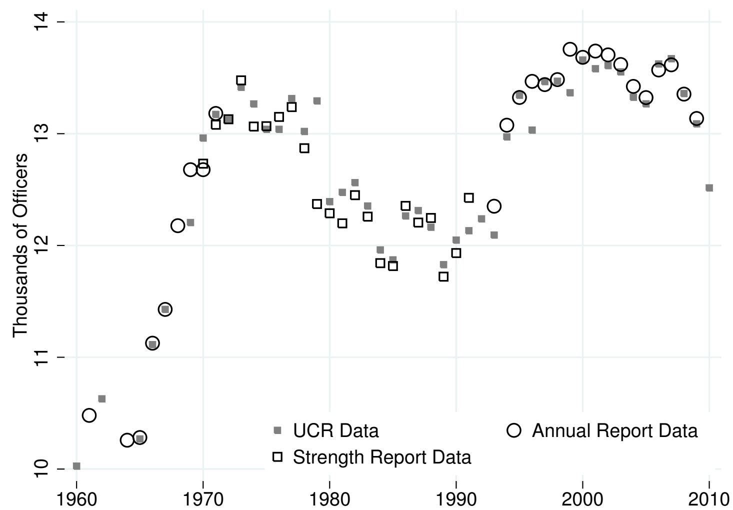
A. New York



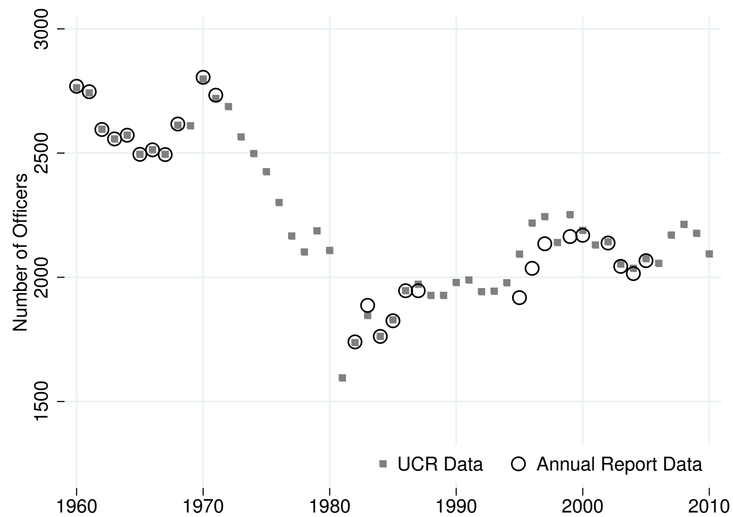
B. Los Angeles



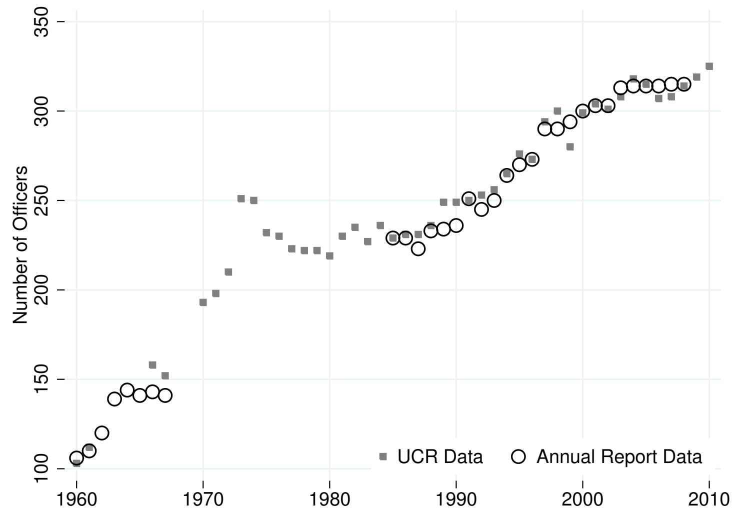
C. Chicago



D. Boston

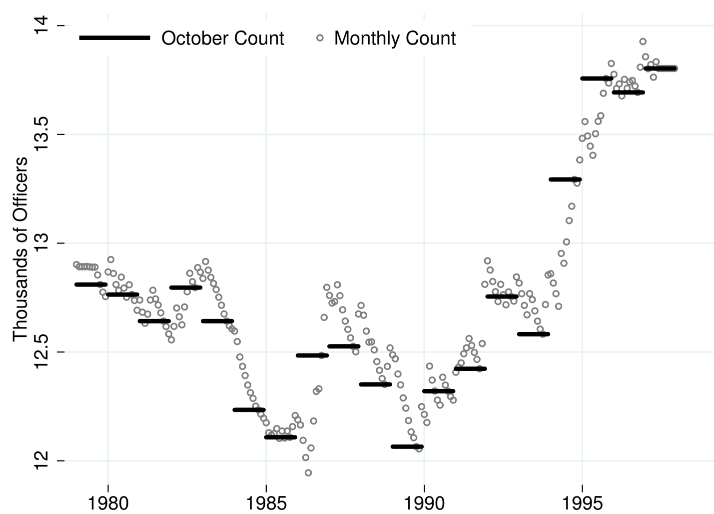


E. Lincoln, Nebraska



Note: In panel A, numbers for 1960-1994 are adjusted to account for the 1995 merger of NYPD with housing and transit police. See Data Appendix for details.

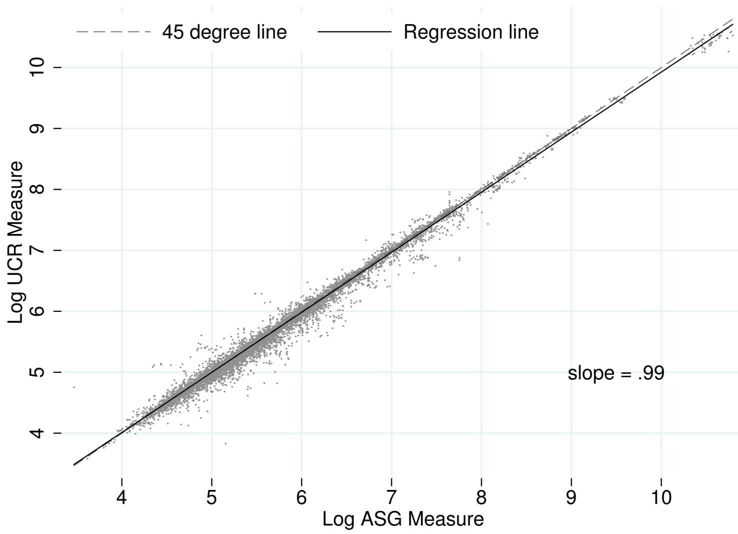
Figure 4. Sworn Officers in Chicago 1979-1997, by Month



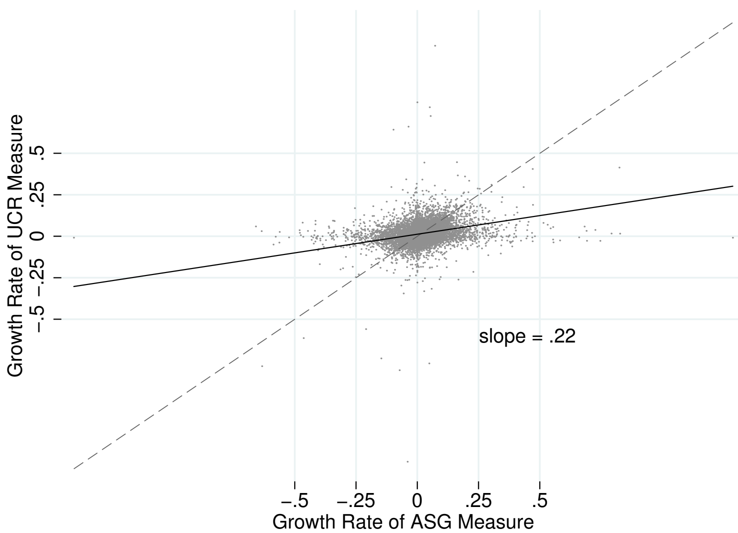
Note: See text for details.

Figure 5. Two Leading Measures of Sworn Officers: the Uniform Crime Reports and the Annual Survey of Government

A. Logs



B. Log Differences



Note: See text and Data Appendix for details.

FIGURE 6. LOCATION OF CITIES IN SAMPLE

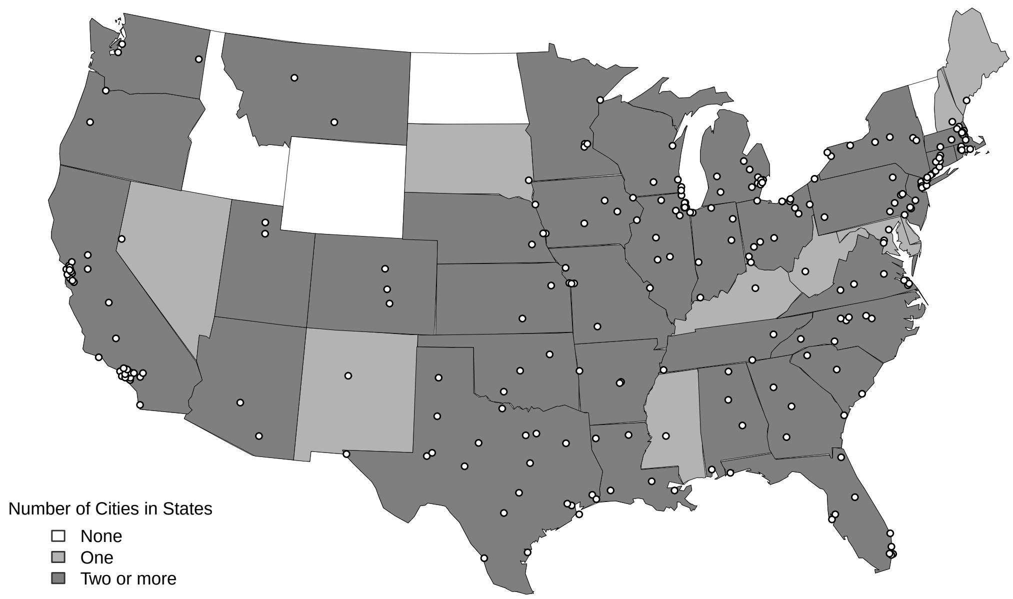
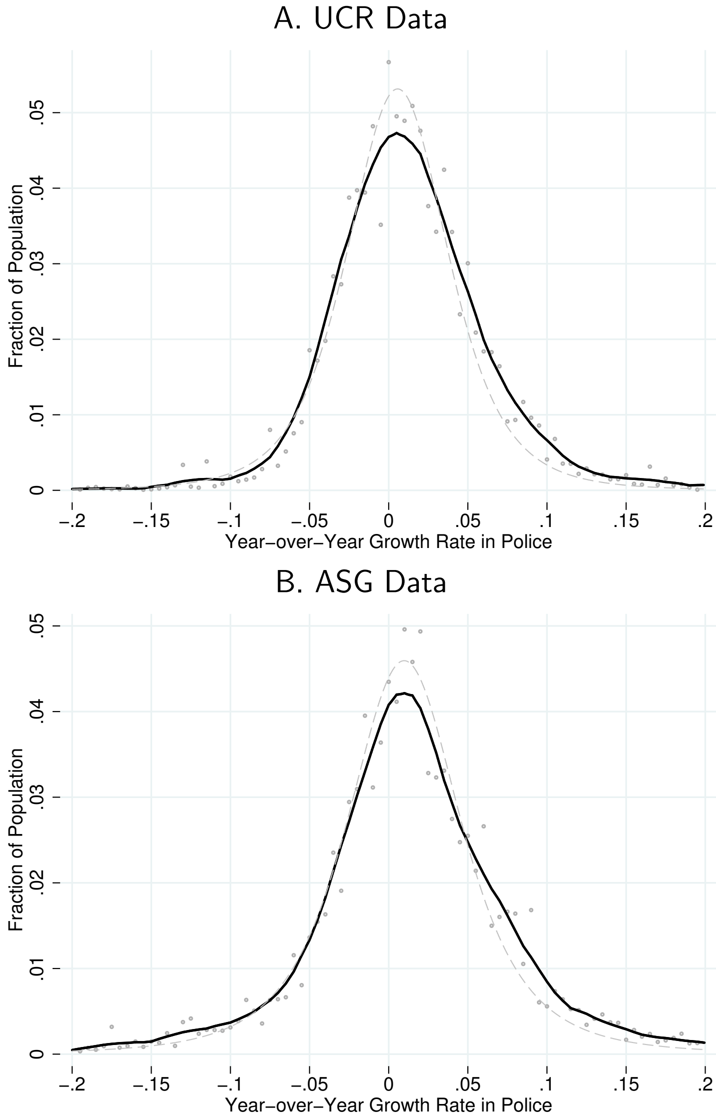


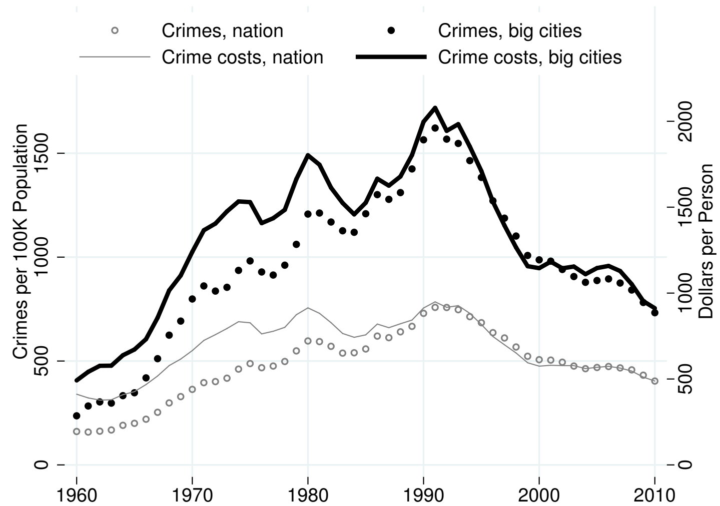
FIGURE 7. DISTRIBUTION OF GROWTH RATES IN POLICE



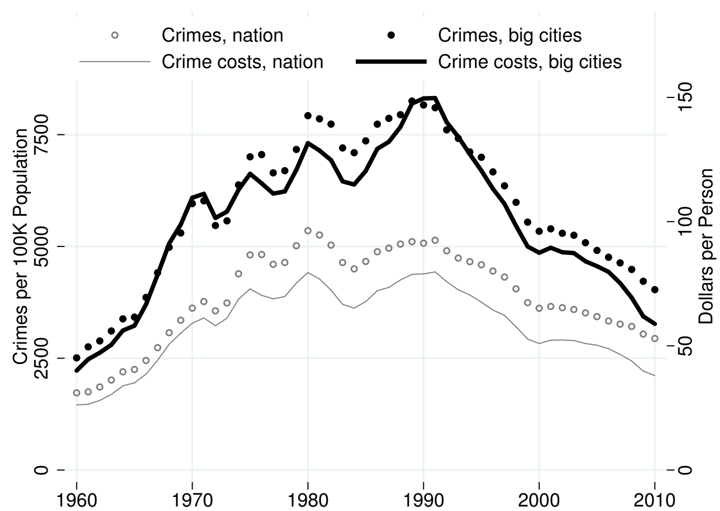
Note: Curves are proportional to the population weighted conditional density function of growth rates in police, conditional on not having been equal to zero. In the UCR (ASG) data, 3.8 (6.1) percent of person-weighted city-years have exactly zero growth rate. Gray circles are undersmoothed histogram heights. The gray dashed line is represents the normal density plot. The scale for the y-axis is percent of person-weighted city-years. See text for details of density estimation.

FIGURE 8. AGGREGATE TRENDS IN VIOLENT AND PROPERTY CRIME AND POLICE:
EVIDENCE FROM THE UNIFORM CRIME REPORTS

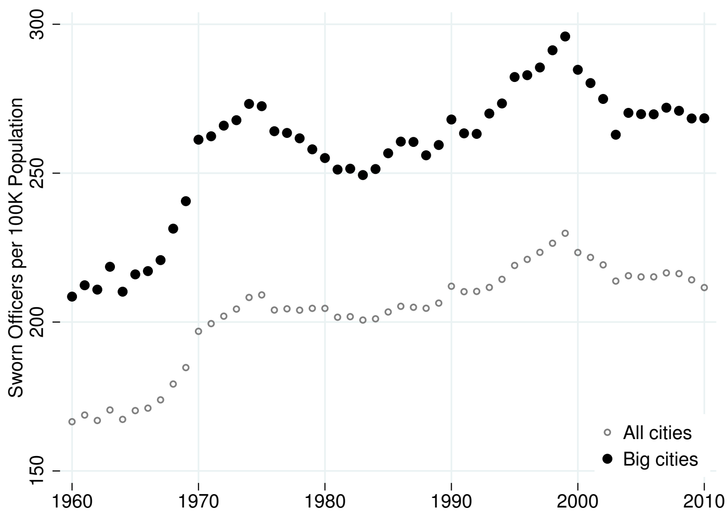
A. Violent Crime: Murder, Forcible Rape, Robbery, Aggravated Assault



B. Property Crime: Burglary, Larceny, Motor Vehicle Theft



C. Sworn Police

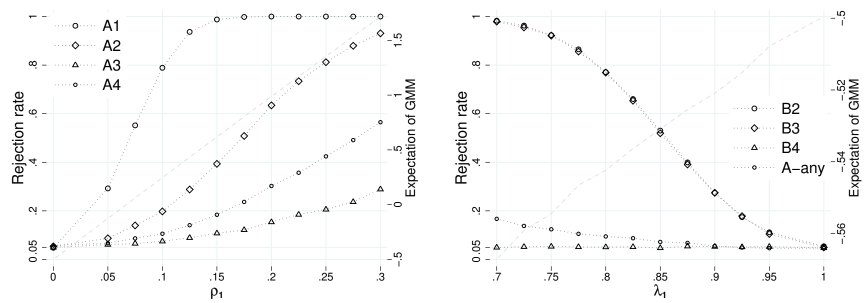


Note: In panels A and B, data on crimes nationally are taken from `http://www.ucrdatatool.gov`. In panel C, no such data are available, and we construct an index using all municipalities ever reporting to the UCR system 1960-2010 and imputation. See text and Data Appendix for details.

Figure 9. Power of Tests of Classical Measurement Error Model

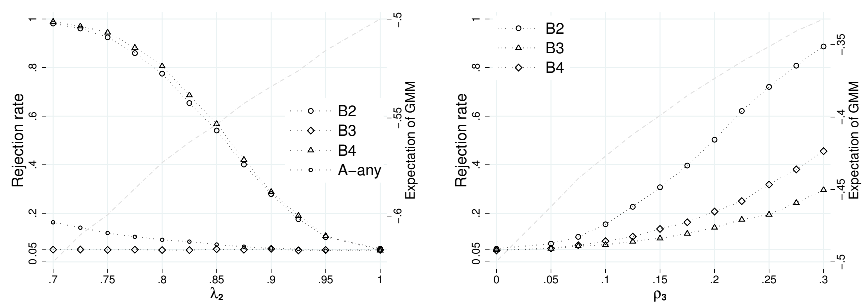
A. Between m.e. and structural error in UCR

B. Mean-reverting m.e.



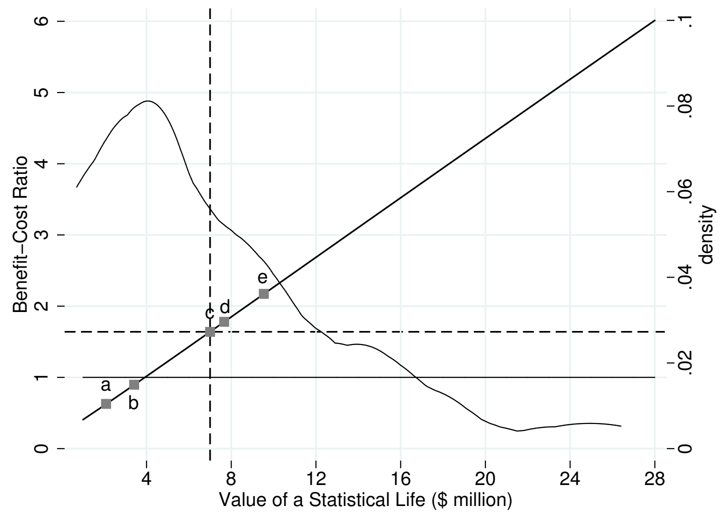
C. Mean-reverting m.e. in ASG

D. Between UCR m.e. and ASG m.e.



Note: m.e. = measurement error. Each panel presents the fraction of 10,000 simulated data sets for which a t-ratio test rejects, as a function of a particular parameter. The parameter ρ_1 indexes the correlation between the measurement error and the structural error (panel A). The parameters λ_1 and λ_2 index the degree of mean reversion in the mesaurement errors (panels B, C). The parameter ρ_3 indexes the degree of correlation between the measurement errors themselves (panel D). Two types of t-ratio tests are presented, corresponding to the tests presented in the top two panels of Table 6. Test A is the t-ratio on the outcome in a regression of the difference in two measures of the variable of interest on the outcome and corresponds to Panel A of Table 6. Test B is the t-ratio on a third measure in a regression of the difference in measures on the third measure and corresponds to Panel B of Table 6. There are four such tests examined corresponding to the columns of Table 6. For example, "A1" corresponds to the t-ratio tests of Panel A of Table 6 for the first column, i.e., the full sample, whereas "B2" corresponds to the t-ratio tests of Panel B of Table 6 for the second column, i.e., the LEMAS subsample. The curves with open circles, diamonds, and triangles correspond to rejection rates for the given scenario. The curve labeled "A-any" is a rejection rate assuming "reject" occurs if any of the four Tests A reject at the 0.05/4 level. The dashed line with no symbols overlaid is the simulation estimate of the expectation of two-step GMM (right axis). The true parameter in all scenarios is -0.5, so departures from -0.5 capture estimator bias. See text for details.

Figure 10. Cost-Benefit Analysis
Benefit-Cost Ratio as a Function of the Value of a Statistical Life



Note: The table plots the value of the benefit-cost ratio calculated using the two-step GMM procedure that pools the "forward" and "reflected" IV regressions of the growth rate in each of nine crime rates on the first lag of the growth rate in the number of sworn police officers, conditional on both the UCR and the ASG measure of the growth rate in the population size and a vector of unrestricted state-by-year dummies. For each measure of police, expenditures on personnel are estimated by multiplying the number of personnel by $130,000, an estimate of the "fully-loaded" annual salary of a police officer. Victimization costs for rape, robbery, assault, burglary, larceny and motor vehicle theft are drawn from Cohen and Piquero (2008). As there is a great deal of variation in extant estimates of the value of a statistical life, the cost of murder is allowed to vary. Using the solid black line, we plot the benefit-cost ratio on the vertical axis as a function of the value of a statistical life, plotted on the horizontal axis in millions of dollars. The horizontal line correpsonds to a benefit-cost ratio of 1. In addition, we superimpose a kernel density function that plots the distribution of the extant estimates of the value of a statistical life. Key estimates include the $2.1 million VSL estimated by Ashenfelter and Greenstone (2004) ("a"), $3.4, the mean VSL among studies of non-labor market behavior ("b"), $7 million, the mean VSL in our sample ("c"), $7.7 million, the mean VSL used by various federal agencies for the 2004-2010 period ("d") and $9.5 million, the mean VSL among studies of U.S. labor market behavior ("e"). The dotted lines show the BCR ($1.63) at the mean value of a statistical life ($7 million). The majority of these estimates are drawn from Viscusi and Aldy (2003). We supplement these estimates with several that are drawn from the more recent literaure.

Table 1. Elasticity of Police with Respect to Selected Covariates

Variable	Source	Years	UCR measure	ASG measure	UCR measure	ASG measure
A. Economic Characteristics						
Personal income	Bureau of Economic Analysis (county-level)	1969-2010	0.053	0.055	-0.002	-0.052
			(0.040)	(0.046)	(0.034)	(0.050)
Total employment			0.097	0.079	0.030	-0.042
			(0.047)	(0.065)	(0.043)	(0.064)
Adjusted gross income	Internal Revenue Service (county-level)	1990-2009	0.012	-0.012	-0.018	-0.039
			(0.035)	(0.034)	(0.036)	(0.050)
Wage and salary income			0.044	0.037	-0.009	-0.047
			(0.042)	(0.042)	(0.043)	(0.058)
Municipal expenditures exclusive of police	Annual Survey of Gov't Finances (city-level)	1960-2010	0.019	0.016	0.011	0.018
			(0.007)	(0.008)	(0.006)	(0.008)
B. Social Disorganization and Demographics						
Share of births to teenage mothers, overall	Centers for Disease Control (county-level)	1968-2002	-0.019	-0.033	-0.022	-0.026
			(0.011)	(0.015)	(0.012)	(0.015)
Share of births to teenage mothers, black births only			-0.012	-0.038	-0.018	-0.029
			(0.011)	(0.015)	(0.012)	(0.015)
Share of births that are low birthweight			0.012	0.004	0.011	-0.007
			(0.012)	(0.016)	(0.011)	(0.017)
12th grade dropout rate	Nat'l Center for Educational Statistics (city-level)	1986-2008	0.006	0.005	-0.012	0.006
			(0.015)	(0.012)	(0.009)	(0.015)
Estimated population at risk of arrest	U.S. Census (county-level)	1970-2010	0.113	0.050	0.124	0.096
			(0.026)	(0.042)	(0.038)	(0.051)
C. Lagged Crimes						
Violent crimes	Federal Bureau of Investigation, Uniform Crime Reports (city-level)	1960-2010	0.012	0.006	0.009	0.004
			(0.005)	(0.006)	(0.005)	(0.006)
Property crimes			0.019	0.012	0.013	0.008
			(0.011)	(0.011)	(0.009)	(0.012)
Cost-weighted crimes			-0.001	0.002	0.001	-0.000
			(0.003)	(0.003)	(0.002)	(0.003)
year effects			yes	yes	—	—
state-by-year effects			no	no	yes	yes

Note: Each column reports coefficients from 13 separate regressions of the growth rate in a measure of the number of police officers on the the growth rate in a covariate, controlling for two measures of the growth rate in city population and either year effects or state-by-year effects. The coefficient reported corresponds to the variable given in the first column, and Huber-Eicker-White standard errors are given in parentheses. The first police measure is from the Uniform Crime Reports and the second is from the Annual Survey of Government Employment. All regressions are weighted by 2010 city population. The municipal budget cycle, the 12th grade dropout rate, and the lagged crime rates are all measured at the city level, but city level data was unavailable for other variables, which were measured instead at the county level. See Data Appendix for details. Population weighted means (standard deviations) for variables in Panel A are 0.068 (0.042), 0.015 (0.029), 0.040 (0.050), 0.041 (0.038), and 0.030 (0.119); for those in Panel B are -0.014 (0.083), -0.016 (0.086), -0.002 (0.065), 0.007 (0.111), and 0.011 (0.038); and for those in Panel C are 0.037 (0.160), 0.017 (0.110), and 0.023 (0.270).

Table 2. Summary Statistics on Police and Crime

		Levels				Log Differences			
Variable		**Mean**	**S.D.**	**Min.**	**Max.**	**Mean**	**S.D.**	**Min.**	**Max.**
Sworn police, UCR	O	245.5	111.6	54.4	786.6	0.016	0.058	-1.359	1.148
(per 100K pop)	B		105.7				0.012		
	W		36.0				0.056		
Sworn police, ASG	O	257.7	128.0	50.1	779.8	0.016	0.078	-1.401	1.288
(per 100K pop)	B		120.4				0.012		
	W		42.4				0.078		
Violent crimes	O	972.7	630.5	8.2	4189.0	0.035	0.162	-1.804	1.467
(per 100K pop)	B		440.3				0.019		
	W		451.4				0.161		
Murder	O	14.6	10.6	0.0	110.9	0.014	0.382	-2.792	2.446
(per 100K pop)	B		8.4				0.021		
	W		6.5				0.382		
Rape	O	49.0	29.6	0.0	310.5	0.035	0.291	-4.384	4.199
(per 100K pop)	B		17.4				0.028		
	W		23.9				0.289		
Robbery	O	438.0	344.5	1.1	2,358.0	0.035	0.202	-1.792	1.946
(per 100K pop)	B		257.5				0.019		
	W		228.9				0.201		
Assault	O	471.1	329.5	1.2	2,761.4	0.037	0.213	-2.833	3.129
(per 100K pop)	B		209.5				0.024		
	W		254.4				0.212		
Property crimes	O	6,223.4	2,355.0	667.3	18,345.2	0.015	0.113	-1.304	1.248
(per 100K pop)	B		1,366.2				0.014		
	W		1,918.2				0.112		
Burglary	O	1,671.9	810.9	143.0	6,713.5	0.010	0.149	-1.549	1.411
(per 100K pop)	B		433.8				0.018		
	W		685.1				0.148		
Larceny	O	3,655.4	1,500.2	84.2	11,590.7	0.017	0.122	-1.435	2.146
(per 100K pop)	B		982.6				0.015		
	W		1,133.7				0.121		
Motor vehicle	O	896.0	574.5	42.5	5,294.7	0.014	0.169	-1.516	1.447
theft	B		428.6				0.016		
(per 100K pop)	W		435.3				0.169		
Cost-Weighted	O	1,433.9	904.9	15.36	8,909.2	0.019	0.271	-2.363	3.033
Crimes	B		699.6				0.018		
($ per capita)	W		573.9				0.270		

Note: This table reports descriptive statistics for the two measures of sworn police officers used throughout the article as well as for each of the seven crime categories and three crime aggregates. For each variable, we report the overall mean, the standard deviation decomposed into overall ("O"), between ("B"), and within ("W") variation, as well as the minimum and maximum values. Summary statistics are reported both in levels per 100,000 population and in growth rates. All statistics are weighted by 2010 city population. The sample size for all variables is N=10,589.

Table 3. First Stage Estimates

	Forward Models		Reflected Models	
Endogenous Regressor	(1)	(2)	(3)	(4)
UCR measure	0.184 (0.014)	0.161 (0.013)		
ASG measure			0.364 (0.029)	0.356 (0.029)
F-statistic	169.1	144.7	154.2	146.4
Instrument:	ASG		UCR	
year effects	yes	—	yes	—
state-year effects	no	yes	no	yes

Each column reports results of a least squares regression of the growth rate in a given measurement of the number of police officers on the the growth rate in the other measurement of police. Columns (1) and (2) report results for the forward regressions in which the UCR measure is employed as the endogenous covariate and the ASG measure is employed as the instrumental variable while columns (3) and (4) report results for the reflected regressions in which the ASG measure is employed as the endogenous covariate and the UCR measure is employed as the instrumental variable. For each set of models, the first column reports regression results, conditional on both the UCR and the ASG measures of the growth rate in the city's population and a vector of year dummies. The second column adds a vector of state-by-year effects. All models are estimated using 2010 city population weights. Huber-Eicker-White standard errors are reported in parentheses below the coefficient estimates.

Table 4. Estimates of the Effect of Police on Crime

	Least Squares Estimates				2SLS Estimates			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	UCR Measure		ASG Measure		*Forward Models* UCR Measure		*Reflected Models* ASG Measure	
Violent crimes	-0.117 (0.037)	-0.120 (0.040)	-0.053 (0.024)	-0.058 (0.023)	-0.289 (0.128)	-0.361 (0.143)	-0.321 (0.100)	-0.336 (0.106)
Murder	-0.270 (0.071)	-0.204 (0.097)	-0.148 (0.047)	-0.143 (0.059)	-0.804 (0.260)	-0.889 (0.364)	-0.742 (0.198)	-0.572 (0.262)
Rape	-0.066 (0.069)	-0.074 (0.092)	-0.038 (0.043)	-0.054 (0.050)	-0.208 (0.234)	-0.339 (0.301)	-0.181 (0.188)	-0.208 (0.248)
Robbery	-0.180 (0.048)	-0.204 (0.047)	-0.085 (0.032)	-0.084 (0.029)	-0.459 (0.176)	-0.521 (0.177)	-0.493 (0.128)	-0.572 (0.125)
Assault	-0.052 (0.044)	-0.037 (0.050)	-0.010 (0.030)	-0.013 (0.035)	-0.052 (0.164)	-0.079 (0.209)	-0.143 (0.120)	-0.104 (0.136)
Property crimes	-0.071 (0.028)	-0.059 (0.026)	-0.028 (0.020)	-0.030 (0.015)	-0.152 (0.109)	-0.189 (0.090)	-0.195 (0.077)	-0.167 (0.068)
Burglary	-0.061 (0.043)	-0.062 (0.037)	-0.041 (0.027)	-0.054 (0.021)	-0.222 (0.144)	-0.339 (0.128)	-0.166 (0.118)	-0.174 (0.098)
Larceny	-0.038 (0.031)	-0.025 (0.027)	-0.002 (0.021)	-0.018 (0.017)	-0.012 (0.115)	-0.113 (0.103)	-0.103 (0.085)	-0.070 (0.074)
Motor vehicle theft	-0.187 (0.049)	-0.131 (0.043)	-0.109 (0.031)	-0.047 (0.025)	-0.592 (0.169)	-0.292 (0.151)	-0.514 (0.130)	-0.367 (0.115)
Cost-Weighted crime	-0.213 (0.054)	-0.144 (0.071)	-0.112 (0.034)	-0.099 (0.041)	-0.605 (0.184)	-0.614 (0.250)	-0.583 (0.147)	-0.403 (0.192)
year effects	yes	—	yes	—	yes	—	yes	—
state-year effects	no	yes	no	yes	no	yes	no	yes
Instrument:	—	—	—	—	ASG		UCR	

Note: Columns (1)-(4) reports results of a least squares regression of the growth rate in each of ten crime rates on the first lag of the growth rate in the number of sworn police officers. For each set of models, the first column reports regression results, conditional on both the UCR and the ASG measures of the growth rate in the city's population and year effects. The second column adds state-by-year effects. Columns (5)-(8) reports results from a series of 2SLS regressions of the growth rate in each of nine crime rates on the first lag of the growth rate in the number of per capita sworn police officers. All models are estimated using 2010 city population weights. Huber-Eicker-White standard errors that are robust to heteroskedasticity are reported in the second row below the coefficient estimates.

Table 5. Pooled Estimates of the Effect of Police on Crime: Within-State Differences

	Violent Crimes				Property Crimes			Aggregates		
Estimator	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft	Violent Crime	Property Crime	Cost-Weighted Crimes
One-step GMM	-0.727 (0.242)	-0.272 (0.231)	-0.547 (0.123)	-0.092 (0.142)	-0.255 (0.093)	-0.091 (0.072)	-0.331 (0.105)	-0.348 (0.103)	-0.178 (0.066)	-0.506 (0.173)
Two-step GMM	-0.666 (0.238)	-0.255 (0.219)	-0.559 (0.117)	-0.099 (0.127)	-0.225 (0.089)	-0.083 (0.067)	-0.343 (0.101)	-0.344 (0.096)	-0.174 (0.062)	-0.473 (0.171)
Empirical Likelihood	-0.667 (0.236)	-0.256 (0.221)	-0.559 (0.117)	-0.099 (0.127)	-0.221 (0.087)	-0.082 (0.067)	-0.341 (0.103)	-0.344 (0.096)	-0.173 (0.061)	-0.473 (0.170)
Test statistic:	0.78	0.19	0.09	0.02	1.86	0.18	0.24	0.03	0.06	0.71

Note: Each column reports generalized method of moments (GMM) or empirical likelihood (EL) estimates of the growth rate in each of ten crime rates on the first lag of the growth rate in the number of sworn police officers, conditional on both the UCR and ASG measures of the growth rate in population and a vector of unrestricted state-by-year effects. Computationally, however, we de-mean the data by state-year rather than computing the fixed effects parameters, even when this is an approximation (e.g., EL). The one-step GMM estimator uses the identity weighting matrix while the two-step GMM estimator uses a weighting matrix based on the updated variance estimator from the one-step procedure. Following Guggenberger and Hahn (2005), we interpret EL as a particular just-identified GMM estimator and compute it using 5 Newton iterations using the two-step GMM estimates as starting values. All models use 2010 city population weights, and Huber-Eicker-White standard errors are reported in parentheses. Below the parameter estimates and the standard errors, we report the value of of the overidentification statistic from the two-step GMM procedure. The test statistic corresponds to the pooling restriction that we estimate a common parameter on the growth rate in police and refers to a test of the equality of the "forward" and "reflected" IV coefficients. The test statistic is distributed χ_1 under the null hypothesis of classical measurement error. The 95 percent critical value of the test is 3.84.

Table 6. Further Tests of Classical Measurement Errors

	Full Sample	LEMAS Subsample		
	UCR-ASG (1)	UCR-ASG (2)	UCR-LEMAS (3)	ASG-LEMAS (4)
A. Growth Rate in Crimes				
Murder	-0.002	-0.009	0.003	0.010
	(0.002)	(0.007)	(0.005)	(0.007)
Rape	0.001	-0.009	-0.002	0.008
	(0.004)	(0.016)	(0.010)	(0.014)
Robbery	-0.004	-0.021	0.009	0.026
	(0.006)	(0.020)	(0.016)	(0.022)
Assault	-0.002	0.004	-0.010	-0.014
	(0.005)	(0.018)	(0.013)	(0.018)
Burglary	0.014	-0.001	-0.013	-0.011
	(0.010)	(0.027)	(0.022)	(0.028)
Larceny	0.002	-0.012	0.006	0.020
	(0.012)	(0.035)	(0.023)	(0.033)
Motor vehicle theft	-0.007	-0.021	0.025	0.040
	(0.008)	(0.023)	(0.017)	(0.023)
B. Growth Rate in Police				
LEMAS police measure		-0.083		
		(0.049)		
ASG police measure			-0.077	
			(0.026)	
UCR police measure				-0.082
				(0.049)
C. Growth Rate in Population				
UCR population measure	-0.055	0.571	0.364	-0.124
	(0.124)	(0.308)	(0.213)	(0.201)
ASG population measure	-0.035	-0.406	-0.200	0.172
	(0.108)	(0.348)	(0.236)	(0.236)
D. Joint Tests of Significance				
F-test: all variables	0.83	0.25	0.17	0.07
F-test: crime variables	0.82	0.39	0.67	0.08
F-test: police variable		0.09	0.00	0.09
F-test: population variables	0.11	0.39	0.11	0.76

Note: Each column corresponds to a particular difference between measures of the growth rate in police. The heading for each column gives the sources of the two measures being differenced. The LEMAS data are only available for years 1987, 1990, 1992, 1993, 1996, 1997, 1999, 2000, 2003, 2004, 2007 and 2008, and so columns involving the LEMAS data correspond to a limited subsample. Each column reports coefficients and heteroskedasticity-robust standard errors (parentheses) from a single regression of the difference in measures reported in the column heading on the variables listed in the first column. The coefficients are grouped substantively into panels A, B, and C. Panel D gives p-values from a series of heteroskedasticity-robust F-tests on the joint significance of each set of variables. Each of the models controls for state-by-year effects and is weighted by 2010 city population.

Table 7. Robustness of Results to the Inclusion of Covariates
1970-2002 Sample

	"Forward" Models						"Reflected" Models					
	Endogenous Covariate: UCR Instrument: ASG						Endogenous Covariate: ASG Instrument: UCR					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Violent crimes	-0.286	-0.291	-0.319	-0.293	-0.316	-0.332	-0.238	-0.233	-0.174	-0.159	-0.159	-0.173
	(0.152)	(0.154)	(0.160)	(0.161)	(0.166)	(0.168)	(0.114)	(0.115)	(0.114)	(0.115)	(0.117)	(0.117)
Murder	-1.068	-1.029	-1.086	-1.058	-1.044	-1.079	-0.523	-0.467	-0.361	-0.336	-0.342	-0.366
	(0.392)	(0.394)	(0.411)	(0.411)	(0.419)	(0.424)	(0.298)	(0.301)	(0.296)	(0.297)	(0.304)	(0.305)
Rape	-0.165	-0.144	-0.264	-0.235	-0.180	-0.193	0.121	0.127	0.156	0.176	0.187	0.168
	(0.258)	(0.259)	(0.266)	(0.269)	(0.274)	(0.277)	(0.210)	(0.212)	(0.211)	(0.214)	(0.219)	(0.221)
Robbery	-0.550	-0.556	-0.547	-0.554	-0.567	-0.578	-0.686	-0.686	-0.602	-0.601	-0.594	-0.603
	(0.185)	(0.186)	(0.190)	(0.193)	(0.197)	(0.200)	(0.135)	(0.136)	(0.134)	(0.135)	(0.138)	(0.139)
Assault	0.095	0.080	0.032	0.082	0.040	0.022	0.138	0.139	0.189	0.215	0.205	0.197
	(0.205)	(0.207)	(0.217)	(0.219)	(0.224)	(0.228)	(0.144)	(0.146)	(0.148)	(0.150)	(0.150)	(0.151)
Property crimes	-0.262	-0.281	-0.295	-0.285	-0.298	-0.314	-0.152	-0.148	-0.122	-0.117	-0.119	-0.124
	(0.102)	(0.102)	(0.107)	(0.107)	(0.108)	(0.108)	(0.086)	(0.087)	(0.088)	(0.089)	(0.090)	(0.090)
Burglary	-0.408	-0.414	-0.450	-0.437	-0.440	-0.461	-0.118	-0.114	-0.099	-0.095	-0.078	-0.081
	(0.146)	(0.147)	(0.156)	(0.157)	(0.161)	(0.160)	(0.117)	(0.119)	(0.123)	(0.125)	(0.127)	(0.126)
Larceny	-0.174	-0.204	-0.224	-0.222	-0.249	-0.261	-0.074	-0.071	-0.040	-0.035	-0.047	-0.049
	(0.121)	(0.121)	(0.126)	(0.126)	(0.127)	(0.128)	(0.093)	(0.093)	(0.095)	(0.095)	(0.097)	(0.095)
Motor vehicle	-0.357	-0.337	-0.328	-0.303	-0.301	-0.320	-0.385	-0.369	-0.348	-0.341	-0.351	-0.356
	(0.160)	(0.160)	(0.164)	(0.163)	(0.166)	(0.168)	(0.147)	(0.147)	(0.147)	(0.147)	(0.151)	(0.155)
Cost-Weighted Crimes	-0.786	-0.763	-0.816	-0.783	-0.790	-0.815	-0.345	-0.312	-0.208	-0.183	-0.187	-0.206
	(0.257)	(0.259)	(0.267)	(0.267)	(0.271)	(0.274)	(0.207)	(0.209)	(0.200)	(0.201)	(0.206)	(0.208)
state-by-year effects	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
economic covariates	no	yes	yes	yes	yes	yes	no	yes	yes	yes	yes	yes
lagged economic covariates	no	no	yes	yes	yes	yes	no	no	yes	yes	yes	yes
demographic variables	no	no	no	yes	yes	yes	no	no	no	yes	yes	yes
polynomials and interactions	no	no	no	no	yes	yes	no	no	no	no	yes	yes
linear time trends	no	no	no	no	no	yes	no	no	no	no	no	yes

Note: Each column reports results of a 2SLS regression of the growth rate in each of ten crime rates on the first lag of the growth rate in the number of sworn police officers. Columns (1)-(6) report results for the "forward" regressions in which the UCR measure is employed as the endogenous covariate and the ASG measure is employed as the instrumental variable while columns (7)-(12) report results for the "reflected" regressions in which the ASG measure is employed as the endogenous covariate and the UCR measure is employed as the instrumental variable. For each set of models, the first column reports regression results, conditional on both the UCR and the ASG measures of the growth rate in the city's population and a vector of unrestricted state-by-year dummies. The second column adds a vector of economic covariates while the third column adds the first lag of each of these covariates. In the fourth column, we add demographic controls which capture the proportion of a city's population that is comprised of each of sixteen age-gender-race groups. In the fifth column, we add polynomial terms and interactions of the demographic variables. Finally, in column (6), we add city-specific linear time trends. All models are estimated using 2010 city population weights. Huber-Eicker-White standard errors that are robust to heteroskedasticity are reported below the coefficient estimates.

Table 8. Extant Estimates of the Effect of Police on Crime
Implied Elasticities

Article	**Country**	**Years**	**Cross-Sectional Units**	**Research Design**	**Violent Crime**	**Property Crime**
Marvell and Moody (1996)	USA	1973-1992	56 cities	lags as control variables	-0.13* (murder) -0.22* (robbery)	-0.15* (burglary) -0.30* (auto theft)
Levitt (1997)	USA	1970-1992	59 cities	mayoral elections	-0.79 -3.03 (murder) -1.29 (robbery)	0.00 -0.55 (burglary) -0.44 (auto theft)
McCrary (2002)	USA	1970-1992	59 cities	mayoral elections	-0.66 -2.69 (murder) -0.98 (robbery)	0.11 -0.47 (burglary) -0.77 (auto theft)
Levitt (2002)	USA	1975-1995	122 cities	number of firefighters	-0.44* -0.91* (murder) -0.45 (robbery)	-0.50* -0.20 (burglary) -1.70* (auto theft)
DiTella and Schargrodsky (2004)	Argentina	4/1994 -12/1994	876 city blocks	redeployment of police following a terrorist attack	n/a	-0.33* (auto theft)
Klick and Tabarrok (2005)	USA	3/12/2002 - 7/30/2003	7 districts	high terrorism alert days	0.0	-0.30* (burglary) -0.84* (auto theft)
Evans and Owens (2007)	USA	1990-2001	2,074 cities	COPS grants	-0.99* -0.84* (murder) -1.34* (robbery)	-0.26 -0.59* (burglary) -0.85* (auto theft)
Our preferred estimates	USA	1960-2010	242 cities	measurement error correction	-0.34* -0.58* (murder) -0.56* (robbery)	-0.14* -0.20* (burglary) -0.33* (auto theft)

Note: This table reports implied elasticities that arise from six recent articles each of which employs a novel identification strategy to estimate a *causal* effect of police on crime. Elasticities are reported for the violent and property crime aggregates as well as for murder, robbery, burglary and auto theft. In place of the original elasticities reported in Levitt (1997), we have included elasticity estimates from McCrary (2002) which correct for a coding error in the original paper. Our preferred estimates which account for the presence of measurement error in the Uniform Crime Reports police series are shown below. Asterisks denote results that are significant, at a minimum, at the 10% level.

Table 9. Tests of the Equality of Cross-Crime Elasticities

Type	Murder	Rape	Robbery	Assault	Burglary	Larceny	Motor Vehicle Theft	Violent Crimes	Property Crimes
Murder	-	.213	.649	.036	.058	.015	.181	-	.035
Rape	-	-	.181	.485	.917	.452	.689	-	.731
Robbery	-	-	-	.002	.008	.001	.120	-	.001
Assault	-	-	-	-	.382	.922	.114	-	.554
Burglary	-	-	-	-	-	.109	.287	.295	-
Larceny	-	-	-	-	-	-	.010	.010	-
Motor vehicle theft	-	-	-	-	-	-	-	.997	-
Violent crimes	-	-	-	-	-	-	-	-	.075

Note: Each element of the table reports a p-value for a test of the equality between the two-step GMM parameters reported in Table 5 for an exhaustive combination of any two crime categories. For example, the p-value arising from a test of the equality of the pooled murder and burglary elasticities is 0.058. The p-values are generated using a GMM procedure in which we stack data pertaining to each of the two crime categories. All models are estimated using 2010 city population weights and condition on two measures of population as well as an unrestricted vector of state-by-year effects.

This dissertation has considered the role of several potential determinants of the aggregate level of crime in the United States. With regard to police, I find strong evidence that increases in police manpower reduce crime with a 10 percent increase in the size of a city's police force predicting a 4 percent in the reduction of the cost of crimes to victims. In addition, this chapter is the first paper in the literature to generate estimates that are sufficiently precise to lend credence to the claim that police are more effective in reducing violent crimes than property crimes. The implication for public policy is the hiring additional manpower is cost-beneficial. Accordingly, the United States can be characterized as being "underpoliced."

With regard to immigration, the evidence is mixed. Chapter 1 isolates plausibly exogenous variation in the assignment of Mexican immigrants to U.S. cities using rainfall in Mexico to construct an instrumental variable. Using panel data on forty-six U.S. metropolitan areas, I find little consistent evidence of an effect of Mexican immigration on crime. Chapter 2 considers a natural experiment in Arizona in which a large share of the state's foreign-born Mexican population exited the state in the aftermath of a broad-based "E-Verify" law. Here, there is evidence of sizeable reductions in property crime when this population leaves though the effects are largely compositional (i.e., explained by the share of the exiting population that is young and male). While the majority of the literature finds little effect of immigration on crime, Chapter 2 underscores the need to consider that the effects of immigration and context-dependent and likely to be heterogenous.

References

Abadie, Alberto, Alexis Diamond and Jens Hainmuller (2010). "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association* 105(490): 493-505.

Abadie, Alberto and Javier Gardeazabal (2003). "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review* 93(1): 113-132.

Adcock, R.J. (1878). "A problem in least squares," *Analyst* 53-54.

Advisory Commission on Intergovernmental Relations (1977). *Block Grants: A Comparative Analysis*, Washington DC: U.S. Advisory Committee on Intergovernmental Relations.

Agnew, Robert and Helene Raskin White (1992). "An Empirical Test of General Strain Theory," *Criminology* 30(4): 475-500.

Altonji, Joseph and David Card (1991). "The Effects of Immigration on the Labor Market Outcomes of Less Skilled Natives" in John Abowd and Richard Freeman (eds.) *Immigration, Trade and the Labor Market*, University of Chicago Press.

Angrist, Joshua D. and Guido W. Imbens (1994)."Identification and Estimation of Local Average Treatment Effects," *Econometrica* 62(2): 467-475.

Ashenfelter, Orley and Alan Krueger (1994). "Estimates of the Economic Returns to Schooling from a New Sample of Twins," *American Economic Review* 84(5): 1157-1173.

Ashenfelter, Orley and Michael Greenstone (2004). "Using Mandated Speed Limits to Measure the Value of a Statistical Life," *Journal of Political Economy* 112(S1): S226-S267.

Aydemir, Abdurrahaman and George J. Borjas (2011). "Attenuation Bias in Measuring the Wage Impact of Immigration," *Journal of Labor Economics* 29(1): 69-112.

Ayres, Ian and Steven D. Levitt (1998). "Measuring Positive Externalities from Unobservable Victim Precaution: An Empirical Analysis of Lojack," *The Quarterly Journal of Economics* 113(1): 43-77.

Bailey, W.C. (1994). "Poverty, Inequality and City Homicide Rates," *Criminology* 22: 531-550.

Baltagi, Badi H. (2006). "Estimating an Economic Model of Crime Using Panel Data from North Carolina," *Journal of Applied Econometrics* 21(4): 543-547.

Banfield, Edward and James Q. Wilson (1963). *City Politics*, Cambridge: Harvard University Press and MIT Press.

Beccaria, Cesare (1764). *On Crimes and Punishments*, Oxford: Clarendon Press.

Becker, Gary S. (1968). "Crime and Punishment: An Economic Approach," *Journal of Political Economy* 76(2): 169-217.

Bentham, Jeremy (1789). *An Introduction to the Principles of Morals and Legislation*, Oxford: Clarendon Press.

Berk, Richard and John MacDonald (2010). "Policing the Homeless: An Evaluation of Efforts to Reduce Homeless-Related Crime," 'emphCriminology & Public Policy 9(4): 813-840.

Black, Dan A. and Jeffrey Smith (2006). "Returns to College Quality with Multiple Proxies for Quality," *Journal of Labor Economics* 24(3): 701-728.

Blau, Judith R. and Peter M. Blau (1982). "The Cost of Inequality: Metropolitan Structure and Violent Crime," *American Sociological Review* 47(1): 114-129.

Bohn, Sarah, Magnus Lofstrom and Steven Raphael (2013). "Did the 2007 Arizona Legal Workers Act Reduce the State's Unauthorized Immigrant Population?," forthcoming in *The Review of Economics and Statistics*

Bollinger, Christopher R. (2003). "Measurement Error in Human Capital and the Black-White Wage Gap," *The Review of Economics and Statistics* 85(3): 578-585.

Borjas, George J. (1999). "The Economics Analysis of Immigration," in *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card. Elsevier B.V.

Borjas, George J. (2003). "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *The Quarterly Journal of Economics* 118(4): 1335-1374.

Borjas, George J. (2006). "Native Internal Migration and the Labor Market Impact of Immigration," *Journal of Human Resources* 41(2): 221-258.

Boston Police Department (1919). *Annual Report*.

Boston Police Department (1982). *Annual Report*.

Bound, John and Alan B. Krueger (1991). "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?," *Journal of Labor Economics* 9(1): 1-24.

Bound, John, Charles Brown and Nancy Mathiowitz (2001). "Measurement Error in Survey Data," in James J. Heckman and Edward Leamer, eds., *Handbook of Econometrics*, Vol. 5, New York: Elsevier: 3705-3843.

Bound, John, Greg J. Duncan and Willard L. Rogers (1994). "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data," *Journal of Labor*

Economics 12(3): 345-368.

Bound, John, David Jaeger and Regina M. Baker (1995). "Problems with Instrumental Variables Estimation when the Correlation between the Instrument and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association* 90(430): 443-450.

Buonanno, Paolo, Bianchi, Milo and Paolo Pinotti (2011)."Do Immigrants Cause Crime," *Journal of the European Economic Association*.

Braga, Anthony A. (2005). "Hot Spots Policing and Crime Prevention: A Systematic Review of Randomized Controlled Trials," *Journal of Experimental Criminology* 1(3): 317-342.

Braga, Anthony A. (2001). "The Effects of Hot Spots Policing on Crime," *The Annals of the American Academy of Political and Social Science* 578: 104-125.

Braga, Anthony A. and Brenda J. Bond (2008). "Policing Crime and Disorder Hot Spots: A Randomized Controlled Trial," *Criminology* 46(3): 577-607.

Braga, Anthony A., David L. Weisburd, Elin J. Waring, Lorraine Green Mazerolle, William Spelman and Francis Gajewski (1999). "Problem-Oriented Policing in Violent Crime Places: A Randomized Controlled Experiment," *Criminology* 37(3): 541-580.

Braga, Anthony A., David M. Kennedy, Elin J. Waring and Anne Morrison Piehl (2001). "Problem-Oriented Policing, Deterrence, and Youth Violence: An Evaluation of Boston's Operation Ceasefire," *Journal of Research in Crime and Delinquency* 38(3): 195-225.

Burdett, Kenneth, Ricardo Lagos and Randall Wright (2004). "An On-th-Job Search Model of Crime, Inequality, and Unemployment," *International Economic Review* 45(3): 681-706.

Butcher, Kristin F. and Anne Morrison Piehl (1998). "Cross-City Evidence on the Relationship Between Immigration and Crime," *Journal of Policy Analysis and Management* 17(3): 457-493.

Butcher, Kristin F. and Anne Morrison Piehl (1998). "Recent Immigrants: Unexpected Implications for Crime and Incarceration," *Industrial and Labor Relations Review* 51(4): 654-679.

Butcher, Kristin F. and Anne Morrison Piehl (2006). "Why are Immigrants Incarceration Rates so Low? Evidence on Selective Immigration, Deterrence and Deportation," NBER Working Paper #13229.

Cameron, Samuel (1988). "The Economics of Crime Deterrence: A Survey of Theory and Evidence," *Kyklos* 41(2): 301-323.

Card, David (1990). "The Impact of the Mariel Boatlift on the Miami Labor

Market," *Industrial and Labor Relations Review* 43(2): 245-257.

Card, David (1999). "The Causal Effect of Education on Earnings," in Orley Ashenfelter and David Card, eds., *The Handbook of Labor Economics*, Vol. 3A, Amsterdam: Elsevier.

Card, David (2001). "Immigrant Inflows, Native Outflows and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics* 19: 22-64.

Card, David and Ethan G. Lewis "The Diffusion of Mexican Immigrants During the 1990s: Explanations and Impacts" in *Mexican Immigration to the United States*, edited byGeorge J. Borjas. Chicago: University of Chicago Press, 2007.

Cardoso, Lawrence A. Mexican Emigration to the United States, 1897-1931: SocioEconomic Patterns. Tucson, Arizona: University of Arizona Press, 1980.

Chalfin, Aaron (2013). "What is the Contribution of Mexican Immigration to U.S. Crime Rates? Evidence from Rainfall Shocks in Mexico," Working Paper.

Chalfin, Aaron and Morris Levy (2013). "The Effect of Mexican Immigration on the Wages and Employment of U.S. Natives: Evidence from the Timing of Mexican Fertility Shocks," Working Paper.

Chalfin, Aaron, Charles Loeffler and Elina Treyger (2013). "Estimating the Effects of Immigration Enforcement on Local Policing and Crime: Evidence from the Secure Communities Program," Working Paper.

Chalfin, Aaron and Steven Raphael (2011). "Work and Crime" in Michael Tonry (ed.), *Oxford Handbook on Crime and Criminal Justice*, Oxford University Press.

Chetty, Raj (2006). "A General Formula for the Optimal Level of Social Insurance," *Journal of Public Economics* 90, 1879-1901.

Chetty, Raj (2009). "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-From Moethods," *Annual Review of Economics, Annual Reviews* 1:(1): 451-488.

Chicago Tribune (1972). "Freeze U.S. Funds to Police: Robinson," *Chicago Tribune*, September 8, 1972.

City of Boston (2007). "Financial management of the City," *FY07 Adopted Budget by Cabinet (Volume I)*.

Clingermeyer, James C. and Richard C. Feiock (2001). *Institutional Constraints and Policy Choice: An Exploration of Local Governance,* Albany: State University of New York Press.

Cochran, W.G. (1968). "Errors of Measurement in Statistics," *Technometrics* 10(4): 637-666.

Cohen, Mark A. (2000). "Measuring the Costs and Benefits of Crime and Justice," *Criminal Justice* 4: 263-315.

Cohen, Mark A. and Alex R. Piquero (2008). "New Evidence on the Monetary Value of Saving a High Risk Youth," *Journal of Quantitative Criminology* 25(1): 25-49.

Cope, Glenn H. (1992). "Walking the Fiscal Tightrope: Local Budgeting and Fiscal Stress," *International Journal of Public Administration* 15(5): 1097-1120.

Cornwell, Christopher and William H. Trumbull (1994). "Estimating the Economic Model of Crime with Panel Data," *The Review of Economics and Statistics* 76(2): 360-366.

Cragg, J.G. and S.G. Donald. "Testing Identifiability and Specification in Instrumental Variable Models," *Econometric Theory* 9 (1993): 222-240.

Dardick, Hal (2004). "Prentice H. Marshall Sr., 77: Activist Judge and Proud of It," *Chicago Tribune*, May 26, 2004.

Dasgupta, Partha and Eric Maskin (1986). "The Existence of Equilibrium in Discontinuous Economic Games, I: Theory," *Review of Economic Studies* 53(1): 1-26.

Davis, Robert (1985). "Police Boot Camp Recopens: 1st Crop of City Recruits n 2 Years at Academy," *Chicago Tribune*, March 3, 1985.

DeJong, Gordon F. and Anna B. Madamba (2001). "A Double Disadvantage? Minority Group, Immigrant Status, and Underemployment in the United States," *Social Science Quarterly* 82(1): 117-130.

Deslippe, Dennis A. (2004). "Do Whites Have Rights?: White Detroit Policeman and 'Reverse Discrimination' Protests in the 1970s," *The Journal of American History* 91(3): 932-960.

Desmond, S. and C.E. Kubrin (2009). "The Power of Place: Immigrant Communities and Adolescent Violence," *Sociology Quarterly* 50: 581-607.

DiTella, Rafael and Ernesto Schargrodsky (2004). "Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack," *The American Economic Review* 94(1): 115-133.

Dodd, Philip (1964). "Goldwater Calls for a Nation-wide War on Crime," *Chicago Tribune*, October 8, 1964.

Duncan, Brian and Stephen Trejo (2013). "Selectivity and Immigrant Employment," Working Paper.

Eck, John E. and Edward R. Maguire (2000). "Have Changes in Policing Reduced Violent Crime? An Assessment of the Evidence," in Alfred Blumstein and Joel Wallman, eds., *The Crime Drop in America*, New York: Cambridge University Press, p. 207-265.

Edlin, Aaron S. and Pinar Karaca-Mandic (2006). "The Accident Externality from Driving," *Journal of Political Economy* 114(5): 931-955.

Ehrlich, Issac (1972). "The Deterrent Effect of Criminal Law Enforcement," *Journal of Legal Studies* 1(2): 259-276.

Ehrlich, Issac (1996). "Crime, Punishment, and the Market for Offenses," *Journal of Economic Perspectives* 10(1): 43-67.

Enstad, Robert (1976) "Daley Bows to Expediency in Agreeing to Police Quotas," *Chicago Tribune*, June 27, 1976.

Espenshade, Thomas J. and Charles A. Calhoun. "An Analysis of Public Opinion Toward Undocumented Immigration," *Population Research and Policy Review* 12 (1993): 189-224.

Evans, William N. and Emily G. Owens (2007) "COPS and Crime," *Journal of Public Economics* 91(1): 181-201.

Evans, William N., Craig Gaithwaite and Timothy J. Moore (2012). "Stalled Progress: Black-White Education and Labor Market Differences and the Long-term Effects of the Crack Epidemic," Working Paper.

Fajnzylber, Pablo, Daniel Lederman and Norman Loayza (2002). "What Causes Violent Crime?," *European Economic Review* 46(7): 1323-1357.

Fan, Jianqing and Irene Gijbels (1996). *Local Polynomial Modelling and Its Applications*, New York: Chapman and Hall.

Forrester, Robert F. "The Racial Problems Involved in Immigration from Latin America and the West Indies to the United States," United States Department of Labor, 1925.

Freeman, Richard (1996). "Why Do So Many Young Americans Commit Crimes and What Might We Do About It?," *Journal of Economic Perspectives* 10(1): 25-42.

Fuller, Wayne A. (1987). *Measurement Error Models*, New York: Wiley.

Gralla, Joan (2012). "U.S. Munis Face $2 Trillion in Unfunded Pension Costs," *Reuters*, July 2, 2012.

Greisinger, George, Jeffrey Slovak and Joseph J. Molkup (1979). *Civil Service Systems: Their Impact on Police Administration*, Washington DC: U.S. GPO.

Grogger, Jeffrey (1991). "Certainty Versus Severity of Punishment," *Economic Inquiry* 29(2): 297-309.

Grogger, Jeffrey (1998). "Market Wages and Youth Crime," *Journal of Labor Economics* 16(4): 756-791.

Guggenberger, Patrik and Jinyong Hahn (2005). "Finite Sample Properties of the Two-Step Empirical Likelihood Estimator," *Econometric Reviews* 24(3): 247-263.

Hagan, John and Alberto Palloni (1999). "Sociological Criminology and the Mythology of Hispanic Immigration and Crime," *Social Problems* 46(4): 617-632.

Hanson, Gordon (2001). "U.S.-Mexico Integration and Regional Economies: Evidence from Border-City Pairs," *Journal of Urban Economics* 50: 259-287.

Hanson, Gordon and Craig McIntosh (2010). "The Great Mexican Emigration," *The Review of Economics and Statistics* 92(4): 798-810.

Harriston, Keith A. and Mary Pat Flaherty (1994). "D.C. Police Paying for Hiring Binge," *Washington Post*, August 28, 1994.

Hausman, Jerry A. (2001). "Mismeasured Variables in Econometric Analysis: problems from the Right and Problems from the Left," *Journal of Economic Perspectives* 15(4): 57-67.

Hevesi, Alan G. (2005). *Revenue Sharing in New York State*, Albany: Office of the New York State Comptroller.

Imbens, Guido W. (1993). "A New Approach to Generalized Method of Moments Estimation," *Harvard Institute for Economic Research Working Paper* No. 1633.

Imbens, Guido W. (2002). "Generalized Method of Moments and Empirical Likelihood," *Journal of Business and Economic Statistics* 20(4): 493-506.

Imbens, Guido W., Richard H. Spady and Phillip Johnson (1998). "Information Theoretic Approaches to Inference in Moment Condition Models," *Econometrica* 66(2): 333-357.

Ireton, Gabriel (1976). "Retirement Surge Grips City Police, Concerns Council," *Pittsburgh Post-Gazette*, May 10, 1976.

James, Nathan (2008). *Edward Byrne Memorial Justice Assistance Grant Poro-gram: Legislative and Funding History*, Washington DC: Congressional Research Service.

Joyce, Philip G. and Daniel R. Mullins (1991). "The Changing Fiscal Structure of the State and Local Public Sector: The Impact of Tax and Expenditure Limitations," *Public Administration Review* 51(3): 240-253.

Katz, Matt and Darran Simon (2011). "Daytime Camden Police Patrols Hit Hard by Layoffs," *Philadelphia Inquirer*, January 20, 2011.

Kim, Bonggeun and Gary Solon (2005). "Implications of Mean-Reverting Measurement Error for Longitudinal Studies of Wages and Employment," *The Review of Economics and Statistics* 87(1): 193-196.

Kleemans, Marieke and Jeremy Magruder. "Labor Market Changes in Response to Immigration: Evidence from Internal Migration Driven by Weather Shocks in Indonesia," Working Paper (2011).

Klick, Jonathan and Alex Tabarrok (2005). "Using Terror Alert Levels to Estimate the Effect of Police on Crime," *Journal of Law and Economics* 48(1): 267-280.

Koper, Christopher S., Edward R. Maguire and Gretchen E. Moore (2001).

Hiring and Retention Issues in Police Agencies: Readings on the Determinants of Police Strength, Hiring and Retention of Officers, and the Federal COPS Program, Washington DC: Urban Institute.

Larkin, Al (1973). "Challenge to Law on Police Hiring Nearing Finish," *Boston Globe*, November 25, 1973.

Lauritsen, J.L. "The Social Ecology of Violent Victimization: Individual and Contextual Effects in the NCVS," *Journal of Quantitative Criminology* 17 (2001): 3-32.

Lee, David S. and Justin McCrary (2009). "The Deterrence Effect of Prison: Dynamic Theory and Evidence," Working Paper.

Lee, M.T., J.R. Martinez and Richard Rosenfeld (2001). "Does Immigration Increase Homicide? Negative Evidence from Three Border Cities," *The Sociological Quarterly* 42: 559-580.

Levitt, Steven D. (1996). "Understanding Why Crime Fell in the 1990s: Four Factors that Explain the Decline and Six that Do Not," *Journal of Economic Perspectives* 18(1), 163-190.

Levitt, Steven D. (1997). "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," *The American Economic Review* 87(3): 270-290.

Levitt, Steven D. (1998). "The Relationship Between Crime Reporting and Police: Implications for the Use of Uniform Crime Reports," *Journal of Quantitative Criminology* 14: 61-81.

Levitt, Steven D. and Thomas J. Miles (2006). "Economic Contributions to the Understanding of Crime," *Annual Review of Law and Social Science* 2: 147-164.

Lewis, Carol W. (1994). "Budgetary Balance: The Norm, Concept, and Practice in Large U.S. Cities," *Public Administration Review* 54(6): 515-524.

Lochner, Lance (2004). "Education, Work and Crime: A Human Capital Approach," *International Economic Review* 45(3): 811-843.

Logan, John R., Brian J. Stults and Reynolds Farley. "Segregation on Minorities in the Metropolis: Two Decades of Change," *Demography* 41(1) (2004): 1-22.

Los Angeles Times (1966). "Early Retirements Thin Police Force," *Los Angeles Times*, February 20, 1966.

Lubotsky, Darren H. and Martin Wittenberg (2006). "Interpretation of Regressions with Multiple Proxies," *The Review of Economics and Statistics* 88(3).

MacDonald, John M., John Hipp and Charlotte Gill (2012). "The Effects of Immigrant Concentration on Changes in Neighborhood Crime Rates," *Journal of Quantitative Criminology*

MacDonald, John M. and Jessica Saunders (2012). "Are Immigrants Less Violent? Specifying the Reasons and Mechanisms," *The Annals of the American*

Academy of Political and Social Science 641: 125-147.

Machin, Stephen and Olivier Marie (2011). "Crime and Police Resources: The Street Crime Initiative," *Journal of the European Economic Association* 9(4): 678-701.

Maltz, Michael D. and Joseph Targonski. "A Note on the Use and Quality of County-Level UCR Data," *Journal of Quantitative Criminology* 18(3) (2002): 297-318.

Martinez, Ramiro Jr., Jacob I. Stowell and Matthew T. Lee (2010). "Immigration and Crime in an Era of Transformation: A Longitudinal Analysis of Homicides in San Diego Neighborhoods," *Criminology* 48: 797-830.

Marvell, Thomas B. and Carlisle E. Moody (1996). "Specification Problems, Police Levels, and Crime Rates," *Criminology* 34(4): 609-646.

Massachusetts Department of Revenue (1997). *Levy Limits: A Primer on Proposition 2.5*.

Massey, Douglas S., Jorge Durand, and Nolan J. Malone. Beyond Smoke and Mirrors: Mexican Immigration in an Era of Economic Integration. New York: Russell Sage Foundation, 2002.

McCrary, Justin (2002). "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment," *The American Economic Review* 92(4): 1236-1243.

McCrary, Justin (2007). "The Effect of Court-Ordered Hiring Quotas on the Composition and Quality of Police," *The American Economic Review* 97(1): 318-353.

McCrary, Justin (2008). "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics* 142(2).

McCrary, Justin (2009). "Dynamic Perspectives on Crime," in Bruce Benson, ed., *Handbook on the Economics of Crime*, Northhampton, MA: Edward Elgar.

McKinley, Jesse (2009). "New Oakland Police Chief Inherits a Force, and a City, in Turmoil," *New York Times*, October 15, 2009.

Mears, Daniel P. "Immigration and Crime: What's the Connection?," *Federal Sentencing Reporter* 14(5): 284-288.

Moehling, Carolyn and Anne Morrison Piehl (2009). "Immigration, Crime, and Incarceration in Early 20th Century America," *Demography* 46(4): 739-763.

Munshi, Kaivan. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," *Quarterly Journal of Economics* 118(2) (2003): 549-597.

Mustard, David B. (2010). "How Do Labor Markets Affect Crime? New Evidence on an Old Puzzle," Working Paper.

Nagin, Daniel (1978). "General Deterrence: A Review of the Empirical Evidence," in Alfred Blumstein, Jacqueline Cohen, and Daniel Nagin, eds., *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates*, Washington DC: National Academy of Sciences, p. 95-139.

Nagin, Daniel (1998). "Criminal Deterrence Research at the Outset of the Twenty-First Century," in Michael Tonry, ed., *Crime and Justice: A Review of the Research*, Vol. 23, Chicago: University of Chicago Press, p. 1-42.

National Advisory Commission on Civil Disorders (1968). *Report of the National Advisory Commission on Civil Disorders*, Washington DC: GPO.

Nielsen, Amie L., Matthew T. Lee and Ramiro Martinez Jr. (2005). "Integrating Race, Place and Motive in Social Disorganization Theory: Lessons From a Comparison of Black and Latino Homicide Types in Two Immigrant Destination Cities," *Criminology* 43(3): 837-872.

Newey, Whitney (1985). "Generalized Method of Moments Specification Testing," *Journal of Econometrics* 29(3): 229-256.

Newey, Whitney and Richard J. Smith (2004). "Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators," *Econometrica* 72(1): 219-255.

Nunn, Nathan (2008). "The Long-term Effects of Africa's Slave Trades," *The Quarterly Journal of Economics* 123(1): 139-176.

O'Shea, Robert (1981). "State, Federal Budget Cuts Mean Trouble for Chicago," *Chicago Tribune*, March 8, 1981.

Ousey, Graham C. and Charles E. Kubrin (2009). "Exploring the Connection Between Immigration and Violent Crime Rates in U.S. Cities, 1980-2000," *Social Problems* 56(3): 447-473

Passel, Jeffrey and D'Vera Cohn (2009). "A Portrait of Unauthorized Immigrants in the United States," Washington DC: Pew Hispanic Center.

Police Executive Research Forum (2005). "The Cop Crunch: Identifying Strategies for Dealing with the Recruiting and Hiring Crisis in Law Enforcement," Grant Final Report.

Police Executive Research Forum (2005). "A Guide for Law Enforcement Chief Executives," Final Draft.

Polinsky, A. Mitchell and Steven Shavell (2000). "The Economic Theory of Public Enforcement of Law," *Journal of Economic Literature* 38(1): 45-76.

Poterba, James M. and Kim S. Reuben (1995). "The Effect if Property-Tax Limits on Wages and Employment in the Local Public Sector," *The American Economic Review* 85(2): 384-389.

Pugatch, Todd and Dean Yang (2011). "The Impact of Mexican Immigration on U.S. Natives: Evidence from Migrant Flows Driven by Rainfall Shocks," Working Paper.

Qin, Jing and Jerry Lawless (1994). "Empirical Likelihood and General Estimating Equations," *The Annals of Statistics* 22(1): 300-325.

Recktenwald, William (1986). "1 in 5 Hopefuls Gets Into Police Training," *Chicago Tribune*, November, 12, 1986.

Reid, Lesley W., Harold E. Weiss, Robert M. Adelman and Charles Jaret (2005). "The Immigration-Crime Relationship: Evidence Across U.S. Metropolitan Areas," *Social Science Research* 34(4): 757-780.

Richardson, Charles (1980). *The State of State-Local Revenue Sharing*, Washington DC: U.S. Advisory Commission on Intergovernmental Relations.

Richman, Daniel (2006). "The Past, Present, and Future of Violent Crime Federalism," in Michael Tonry, ed., *Crime and Justice: A Review of Research*, Vol. 34, Chicago: University of Chicago Press, p. 377-439.

Rosenblum, Marc R. and Lang Hoyt (2011). "The Basics of E-Verify, the U.S. Employer Verification System," *Migration Information Source*.

Rothenberg, Thomas J. (1984). "Approximating the Distributions of Econometric Estimators and Test Statistics," in Zvi Grilliches and Michael D. Intrilligator, eds., *The Handbook of Econometrics*, Vol. 2, Amsterdam: North-Holland, p. 882-935.

Rubin, Irene S. (1997). *The Politics of Public Budgeting: Getting and Spending, Borrowing and Balancing*, Chatham, NJ: Chatham House Publishers.

Russell, Rosalind (1975). *A City in Terror: Calvin Coolidge and the 1919 Boston Police Strike*, Boston: Beacon Press.

Saez, Emmanuel (2001). "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies* 68(1): 205.

Saiz, Albert. "Room in the Kitchen for the Melting Pot: Immigration and Rental Prices," *Review of Economics and Statistics* 85(3) (2003), 502-521.

Salanick, Gerlad R. and Jeffrey Pfeffer (1977). "Constraints on Administrator Discretion: The Limited Influence of Mayors on City Budgets," *Urban Affairs Review* 12(4): 475-498.

Samuelson, Paul (1954). "The Pure Theory of Public Expenditure," *The Review of Economics and Statistics* 36(4): 387-389.

Schwochau, Susan, Peter Feuille and John Thomas Delaney (1988). "The Resource Allocation Effects of Mandated Relationships," *Administrative Science Quarterly* 33(3): 418-437.

Sellin, Thorsten (1938). "Culture Conflict and Crime," *American Journal of*

Sociology 44(1): 97-103.

Shadbegian, Ronald J. (1998). "Do Tax and Expenditure Limitations Affect Local Government Budgets? Evidence from Panel Data," *Public Finance Review* 26(2): 1-18.

Shadbegian, Ronald J. (1999). "The Effect of Tax and Expenditure Limitations on the Revenue Structure of Local Government, 1962-1987," *National Tax Journal* 52(2): 221-228.

Shaw, Clifford R. and Henry D. McKay (1969). *Juvenile Delinquency and Urban Areas*, University of Chicago Press.

Sherman, Lawrence W. and David Weisburd (1995). "General Deterrent Effects of Police Patrol in Crime 'Hot Spots': A Randomized Controlled Trial," *Justice Quarterly* 12(4): 625-647.

Sherman, Lawrence W. and Dennis P. Rogan (1995). "Effects of Gun Seizures on Gun Violence: 'Hot Spots' Patrol in Kansas City," *Justice Quarterly* 12(4): 673-693.

Siskin, Bernard R. and David W. Griffin (1997). *Analysis of Distributions by Rank, Race, and Gender: City of Chicago Police Department, 1987-1991*, Philadelphia: Center for Forensic Economic Studies.

Skogan, Wesley and Kathleen Frydll (2004). *Fairness and Effectiveness in Policing: The Evidence*, Washington DC: National Academies Press.

Solon, Gary, Steven J, Haider and Jeffrey M. Wooldridge (2012). "What Are We Weighting For?," Working Paper.

Spenkuch, Jorg (2011). "Understanding the Impact of Immigration on Crime," Working Paper.

Stigler, George J. (1970). "The Optimum Enforcement of Laws," *Journal of Political Economy* 78(3): 526-536.

Stock, James and Motohiro Yogo. "Testing for Weak Instruments in Linear IV Regression," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, edited by James Stock and Donald Andrews. Cambridge: Cambridge University Press.

Stowell, Jacob I., Steven F. Messner, Kelly F. McGeever and Lawrence E. Raffalovich (2009). "Immigration and the Recent Violent Crime Drop in the United States: A Pooled, Cross-Sectional Time-Series Analysis of Metropolitan Areas," *Criminology* 47(3): 889-928.

Tirole, Jean (1988). *The Theory of Industrial Organization*, Cambridge: MIT Press.

Tonry, Michael (1997). "Ethnicity, Crime and Immigration," *Crime and Justice* 21: 1-29.

Varon, Jay N. (1975). "A Reexamination of the Law Enforcement Assistance Administration," *Stanford Law Review* 27(5): 1303-1324.

Wadsworth, Tim (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.

Weisburd, David (2005). "Hot Spots Policing Experiments and Criminal Justice Research: Lessons from the Field," *The Annals of the American Academy of Political and Social Science* 599: 220-245.

Weisburd, David, Cody W. Telep, Joshua C. Hinckle and John E. Eck (2010). "Is Problem-Oriented Policing Effective in Reducing Crime and Disorder?," *Criminology & Public Policy* 9(1): 139-172.

Westat (2007). "Findings of the Web Basic Pilot Evaluation."

Westat (2009). "Findings of the E-Verify Program Evaluation."

Wilson, James Q. and Barbara Boland (1978). "The Effect of Police on Crime," *Law and Society Review* 12(3): 367-390.

Wilson, Jeremy M. and Clifford A. Grammich (2009). *Police Recruitment and Retention in the Contemporary Urban Environment: A National Discussion of Personnel Experiences and Promising Practices from the Front Lines*, Santa Monica, CA: RAND.

Witt, Robert, Alan Clarke and Nigel Fielding (1999). "Crime and Economic Activity: A Panel Data Approach," *British Journal of Criminology* 39(3): 391-400.

Woodruff, Christopher and Rene Zenteno. "Migration networks and microenterprises in Mexico." *Journal of Development Economics* 82 (2007): 509528.

Zimring, Franklin E. (2006). *The Great American Crime Decline,* Oxford University Press.

Zimring, Franklin E. (2011). *The City That Became Safe: New York's Lessons for Urban Crime and Its Control*, New York: Oxford University Press.