

Macro-level Research on Immigration and Crime in the Contemporary USA: Problems and Simple Solutions¹

Patrick Schnapp*

Abstract

Extant research uses regression analysis with macro-level data to study the effect of immigration on crime in the contemporary USA. These studies have found mostly null or negative associations between the two variables. I point out three problems with these analyses. (i) Some studies use negative binomial regression inappropriately, in effect measuring determinants of frequencies rather than rates; (ii) all existing estimates arguably suffer from severe undercontrol, overcontrol, or both and cannot establish the direction of causality (if any) between immigration and crime; (iii) all studies present estimates that give equal weight to areas of differing population size. Taken together, these limitations render the research inconclusive. I show how to avoid these problems in a regression analysis of the effect of recent immigration on homicide rates in a sample of 91 US cities in 2000. Estimates point to a negative effect of immigration on homicide, but are not statistically significant.

Introduction

Several theories, such as social disorganization (Shaw and McKay 1931, 1974), culture conflict (Sellin 1938; Wirth 1931), and anomie (Merton 1955, 1964, 1968), suggest that more immigration may lead to more crime. The view that immigration is an important cause of crime is also prominent in US public opinion (see Simon and Sikich 2007) and discourse (see Butcher and Piehl 1998, 457-59; Martinez 2006, 2-6; Wadsworth 2010, 531-38) and has been used as an argument in favor of limiting immigration to the US (Brimelow 1995).

Recent years have seen an increase in studies that test for such an influence of immigration on crime in the contemporary USA. Most of the research presents analyses of supra-individual units, such as neighborhoods or cities. This is appropriate, as most of the purported mechanisms in extant theories refer to contextual effects. These studies assess the total effect of immigration on crime; only few of them (Martinez, Stowell, and Lee, 2010; Ousey and Kubrin 2009; Stowell 2007) try to additionally explore processes that may bring about such an influence. A consensus has emerged according to which this literature fails to support the view that immigration leads to more crime (Akins, Rumbaut, and Stansfield 2009, 307; Desmond and Kubrin 2009, 528; Martinez 2006, 10; Wadsworth 2010, 537) and perhaps points to a negative effect (Stowell et al. 2009, 608). As a consequence of such findings, Martinez, Lee and coauthors have developed the immigration revitalization perspective, which aims to explain how immigration may lower crime rates (Lee and Martinez 2002; Lee, Martinez, and Rosenfeld 2001; Martinez, Stowell, and Lee 2010), and Sampson (2006, 2008) has suggested that high immigration contributed to the crime decline of the 1990s.

The present paper points out that this consensus is based on research which suffers from a number of methodological weaknesses that are so severe that they render the research inconclusive. It also demonstrates how to overcome these weaknesses within a multivariate regression framework. The remainder of this article is structured as follows. Section 2 discusses limitations of extant research, with a particular focus on the problem of statistical overcontrol, section 3 presents an analysis of

* University of Hildesheim, p_schnapp@gmx.de

immigration and homicide rates in a sample of US cities which does not suffer from the identified weaknesses, and section 4 provides a summary and discussion.

Limitations of Extant Studies

While it is customary to present the findings of the previous literature before discussing its shortcomings, this section is structured differently. I start by pointing out the inappropriate use of negative binomial regression in some studies, as this limitation is so severe as to warrant the exclusion of these studies from the literature review. The following subsection first discusses the role of control variables in causal analysis and then presents and discusses extant findings on that basis. A third subsection points out another shortcoming of previous estimates, the failure to weight samples, and a fourth summarizes this section.

Inappropriate use of negative binomial regression

Many studies of immigration and crime use negative binomial regression, which is appropriate when rare crimes are studied using small units of analysis and hence produce highly skewed distributions that cluster at zero. While negative binomial regression is best known as a method for analyzing *counts*, an analysis of crime counts in an area can be turned into an analysis of rates by adding the natural log of the population at risk as an independent variable in a multivariate binomial regression (Osgood 2000, 27). An example of the “population at risk” variable, also known as an “offset variable”, is the number of Latinos if the dependent variable is Latino homicide victimization.

Three extant studies apply negative binomial regression in analyses of immigration and crime, but fail to include an appropriate offset variable. Shihadeh and Barranco (2010a) use black homicide offending counts as the dependent variable, controlling for the unlogged size of either the Latino or the total population. Shihadeh and Barranco (2010b) control for unlogged total population in their separate analyses of white and black non-Latino homicide offending. Shihadeh and Winters (2010) use the unlogged total population in their analysis of Latino homicide victimization counts. As a result, the estimates from these studies are not informative about the determinants of crime rates; they will hence not be discussed in the review of the literature below.

Problems of assessing effects

Background: The role of control variables in the estimation of causal influences

Most researchers working in the field agree that causality is most usefully conceptualized in terms of counterfactuals. According to the counterfactual concept, answers to questions of causality compare actual to hypothetical situations which differ in an important respect (Angrist and Pischke 2009: 52-53).

For example, we may be interested in the causal influence of a variable D (such as the rate of immigrants), called the *target variable* (Moody and Marvell 2010), on a dependent variable, Y (such as the homicide rate). In order to estimate this effect, we could run a regression of Y on D. Usually this will produce misleading estimates, because there are influences on Y other than D which are correlated with D. The vector of these variables may be referred to as X. Multivariate regression can be used in an attempt to make different units of analysis comparable with respect to X. If we have measured all relevant variables comprising X, we can hold them constant in a regression model to obtain unbiased estimates of the effect of D on Y (Heckmann 2005, 32-34; Winship and Sobel 2001, 494-495). This strategy is sometimes called the “all causes” approach (e.g., Heckmann 2005).

One may hence be tempted to control for as many variables as possible when estimating the effect of D on Y. Such a strategy is inadvisable, however. The reason for this is that some of the variables one controls for may themselves be influenced by D. If they are included in a regression, the

effect of D on Y will be estimated incorrectly. This is the problem of overcontrol. In this respect, a thorny dilemma is presented by variables which Angrist and Pischke (2009) call “proxy controls”. Proxy controls are “variables that might partially control for omitted factors but are themselves affected by the [target variable]” (Angrist and Pischke 2009, 66). The researcher then faces a tradeoff. Failure to include them may result in undercontrol, but their inclusion could cause overcontrol.

A different tradition in counterfactual causal analysis starts with the observation that randomized experiments are an excellent way of holding other influential variables constant and hence tries to mimic the logic of such experiments using nonexperimental data. Terms used in this context are *Rubin causal model* (Holland 1986), *potential outcomes model* (Rubin 2005), and *selection on observables* (Angrist and Pischke 2009). I will use the term *selection approach*, because the challenge for the researcher is to model the process by which units of analysis are selected into various treatments, or different intensities of a treatment.

The central insight of this tradition is that when an experiment is not randomized, it may still be possible to interpret the estimates as causal with as much confidence as those from a randomized experiment. The prerequisite for this is that the estimates are adjusted for all variables which influence D, unless they have no influence on Y (Rubin 1974). This condition is known as conditional independence or ignorability (Legewie 2012, 130). If it is fulfilled, differences in Y as a function of D may be interpreted as causal. This logic generalizes to observational cases in which a “treatment” (such as the rate of immigrants) can take on many values (Angrist and Pischke 2009, 77-80). Researchers in this tradition choose control variables that “we can think of as having been fixed at the time the [target variable] was determined” (Angrist and Pischke 2009: 64), thus avoiding overcontrol.

Although both traditions ultimately converge on the view that only variables connected to both D and Y should be held constant, the focus is different. While the all causes approach focuses the researcher’s attention on variables that may influence the values of Y (and might be correlated with D), the selection approach prescribes the control of variables that may influence D (and might be correlated with Y) (Morgan and Winship 2007, 81-83).

A second problem regarding causality is often left undiscussed in the counterfactual literature (for an exception, Winship and Sobel 2001). This is the problem of causal direction. Suppose we had established to our satisfaction that a correlation between D and Y may be interpreted as causal. We still would not know whether we are observing an influence of D on Y, of Y on D, or a mixture of both. Sometimes reasonable assumptions can be made about the direction of causality on the basis of our knowledge of the world. In other cases, this will be problematic.

To summarize, when the aim is to isolate the causal effect of D on Y, it is generally a viable strategy to choose an all causes approach and control for influences on Y other than D. However, when a variable may cause overcontrol, the cost of including it may be greater than the cost of omitting it. The alternative selection approach focuses the researcher’s attention on variables that influence D, and which cannot cause overcontrol because they were fixed before D was determined. If we are confident that we are measuring a causal connection between D and Y, we are still faced with the problem of determining the direction of the influence. In the next two subsections, I review studies that estimate the influence of immigration on crime, with a specific focus on overcontrol problems.

Cross-sectional studies

In this subsection I review macro-level studies that aim to estimate the effect of some measure of the presence of immigrants on crime rates using cross-sectional data. Excluded are studies that use immigration measures as a mere control variable, include it as one predictor of crime among many (Morenoff, Sampson, and Raudenbush 2001; Nielsen, Lee, and Martinez 2005; Phillips 2002), or present cross-sectional estimates only in order to compare them to preferred longitudinal results (Stowell et al. 2009; Wadsworth 2010).

Table 1 presents results from studies using census tracts (a proxy for neighborhoods) as their units of analysis, table 2 does the same for analyses of cities and standard metropolitan statistical areas

(SMSAs). All included studies use as their target variable a measure of immigrants as a percentage of the population; however, they differ in whether they include all immigrants or only those who arrived recently (typically, within the last 10 years). Furthermore, some studies use ethnic-specific measures. If both more general and more specific measures are used, I report only estimates for the more general measure. Likewise, when there are different measures of crime, I report results for the most general one. If researchers present various regression models using the same dependent and target variables, I report estimates from the one using *the least control variables*, because these are the ones that will be *the least prone to overcontrol*, the problem I highlight in this section.

Table 1: Overview of Cross-Sectional Studies of Immigration and Crime at the Neighborhood level

Study	Subsample	Year(s)	N	Dependent Variable	Measure Immigration	Im	RS	EH	Other demographic	Economic/educational	Other	
Lee, Martinez, and Rosenfeld 2001	El Paso	1985-1995	86	Black homicide victimization	% immigrants <10 yrs.	-	_*		3 ^{a, b, c}	2 ^{d, e}	1 ^f	
				Latino homicide		_*	_*		3 ^{a, b, c}	2 ^{d, e}	1 ^f	
	Miami		70	Black homicide victimization		_*	-		3 ^{a, b, c}	2 ^{d, e}	1 ^f	
Martinez, Lee, and Nielsen 2004	San Diego	1985-1995	196	Latino homicide		+	-		3 ^{a, b, c}	2 ^{d, e}	1 ^f	
				Black homicide victimization		+*	_*		3 ^{a, b, c}	2 ^{d, e}	1 ^f	
	Miami		70	Drug-related homicides (binary)	% 1960-69 immigrants	-	+		3 ^{a, b, c}	2 ^{d, e}	1 ^f	
Stowell 2007	Houston	1999-2001	32	Violent crime	% 1970-79 immigrants	-	+		1 ^a	2 ^{g, h}	0	
					% 1980-89 immigrants	-	+		1 ^a	2 ^{g, h}	0	
					% 1960-69 immigrants	_*	+		1 ^a	2 ^{g, h}	0	
					% 1970-79 immigrants	-	+		1 ^a	2 ^{g, h}	0	
					% 1980-89 immigrants	+*	-		1 ^a	2 ^{g, h}	0	
Stowell and Martinez 2007	Miami	1999-2001	76	Cuban violent crime	% immigrants <10 yrs.	+			1 ^a	1 ^d	0	
					Nicaraguan violent crime		+*			1 ^a	1 ^d	0
					Honduran violent crime		-			1 ^a	1 ^d	0
	Houston	516	76	Haitian violent crime		_*	-	-	1 ^a	2 ^{d, e}	1 ^f	
					Mexican violent crime		+	-	-	1 ^a	2 ^{d, e}	1 ^f
					Salvadoran violent crime		+	-	+*	1 ^a	2 ^{d, e}	1 ^f
					Vietnamese violent crime		+	_*	+*	1 ^a	2 ^{d, e}	1 ^f
Martinez, Stowell, and Cancino 2008	San Antonio	1995-2004	259	Homicide	% immigrants <10 yrs.	-	-		1 ^a	2 ^{d, e}	1 ^f	
						+*	+		1 ^a	2 ^{d, e}	1 ^f	
Akins, Rumbaut, and Stansfield 2009	San Diego	2004-2006	182	Homicide	% immigrants <10 yrs.	_*	-		1 ⁱ	2 ^{g, j}	0	
						-		+	1 ^a	1 ^g	1 ^f	
Velez 2009		1993-1995	186	Homicide	% immigrants <5 yrs.	+	? ^p	+*	3 ^{a, k, l}	1 ^g	2 ^{f, m}	
Graif and Sampson 2009		2002-2006	836	Homicide	% foreign born	-	+*		1 ⁿ	1 ^g	0	
Olson et al. 2009		1996-1997	191	Homicide (native offenders)	% foreign born	-	+	+	1 ^c	3 ^{d, e, o}	0	
				Robbery (native offenders)		+	+	+	1 ^c	3 ^{d, e, o}	0	
				Agg. assault (native offenders)		-	-	+*	1 ^c	3 ^{d, e, o}	0	
				Sexual assault (native offenders)		-	+	-	1 ^c	3 ^{d, e, o}	0	
				Homicide		_*			1 ^a	1 ^j	0	
Stowell and Martinez 2009		1997-2003	76	Homicide	% Cuban immigrants	_*			1 ^a	1 ^j	0	
					% Honduran immigrants	_*			1 ^a	1 ^j	0	
					% Nicaraguan immigrants	_*			1 ^a	1 ^j	0	
					% Haitian immigrants	+			1 ^a	1 ^j	0	

Abbreviations:

Im: Immigration
 RS: Residential Stability
 EH: Ethnic/racial heterogeneity

Variables footnotes:

a: % young men	e: % poor	i: Adult-to-child ratio	m: Lagged dependent variable
b: Population size	f: Spatial lag	j: % professional	n: Population density
c: % female-headed households	g: Economic deprivation	k: Majority Black (dummy)	o: % high-school graduates
d: % unemployed	h: % low-skill workers	l: Majority Latino (dummy)	p: coefficient not reported

Table 2: Overview of Cross-Sectional Studies of Immigration and Crime at the City/SMSA level

Study	Year(s)	Unit	N	Dep. Var.	Measure Immigration	Im	RS	EH	% Lat.	Other Demographic	Economic/Educational	Other
Hagan and Palloni 1998	1980	SMSAs	43	Arrests (any crime)	% illegal immigrants	-				1 ^a	1 ^b	
					% noncitizens	+				1 ^a	1 ^b	
Butcher and Piehl 1998	1981-1984; 1986-1990	SMSAs	339	Overall Crime	% recent (<1 yr. immigrants)	+*						
Martinez 2000	1980	Cities	111	Homicide	% foreign-born/% foreign-born < 5 yrs.	+			-*	3 ^{c, d, e}	3 ^{b, f, g}	1 ^h
Feldmeyer and Steffensmeier 2009	2000	Cities	328	Homicide	% immigrants <10 yrs.	+	+	+*		4 ^{d, e, i, j}	1 ^k	1 ^l
Reid et al. 2005	2000	SMSAs	137	Homicide	% immigrants <10 yrs.	-*				3 ^{c, m, n}	4 ^{k, o, p, q}	1 ^r
					% Asian foreign-born/% Hispanic foreign-born	-/+				3 ^{c, m, n}	4 ^{k, o, p, q}	1 ^r
					% foreign-born w/ limited English	+				3 ^{c, m, n}	4 ^{k, o, p, q}	1 ^r
				Robbery	% immigrants <10 yrs.	-				3 ^{c, m, n}	4 ^{k, o, p, q}	1 ^r
					% Asian foreign-born/% Hispanic foreign-born	-/+				3 ^{c, m, n}	4 ^{k, o, p, q}	1 ^r
					% foreign-born w/ limited English	+				3 ^{c, m, n}	4 ^{k, o, p, q}	1 ^r
				Burglary	% immigrants <10 yrs.)	-				3 ^{c, m, n}	4 ^{k, o, p, q}	1 ^r
					% Asian foreign-born/% Hispanic foreign-born	-/+				3 ^{c, m, n}	4 ^{k, o, p, q}	1 ^r
					% foreign-born w/ limited English	+				3 ^{c, m, n}	4 ^{k, o, p, q}	1 ^r
				Larceny	% immigrants <10 yrs.	+				3 ^{c, m, n}	4 ^{k, o, p, q}	1 ^r
					% Asian foreign-born/% Hispanic foreign-born	-*/+				3 ^{c, m, n}	4 ^{k, o, p, q}	1 ^r
					% foreign-born w/ limited English	+				3 ^{c, m, n}	4 ^{k, o, p, q}	1 ^r

Abbreviations:

Im: Immigration
 RS: Residential Stability
 EH: Ethnic/racial heterogeneity
 % Lat.: % Latino

Variables footnotes:

a % African American
 b % poor
 c Divorce rate
 d Population size
 e % young males
 f Income inequality
 g % high-school graduates

h Southwest (dummy)
 i Population density
 j % female-headed households
 k: Economic deprivation
 l: Police per capita
 m: Summary measure: population size/density
 n: % young (both sexes)

o: % unemployed
 p: % in low-skilled service employment
 q: % in manufacturing employment
 r South (dummy)

Which dependent, target and control variables were used can be seen from the table. In each of the tables, I display the sign of the coefficients for the immigration measure, ethnic heterogeneity, residential stability², and percent Latino (if applicable). Significant results are asterisked. Other covariates are summarized in three broad groups: other demographic, economic/educational, and other. The number of such control variables is displayed in the table, with footnotes specifying which variables were used.

The tables show that the cross-sectional literature does not show a clear pattern of positive associations between immigration and crime. However, with the exception of the bivariate estimate from Butcher and Piehl, *all* of the estimates potentially suffer from overcontrol problems. Why the variables in question may cause overcontrol problems is discussed in the following paragraphs.

Residential stability, ethnic heterogeneity and percent Latino. Given the overrepresentation of Latinos among the foreign-born (Schmidley and Gibson 1999, 25), immigration will tend to increase the proportion of these as a percentage of the population. Immigration also seems likely to increase ethnic/racial heterogeneity and local residential stability.

Other demographic variables. Immigration will increase population size and hence density, unless the influx of immigrants is compensated by an outflux of at least equal size. Given the underrepresentation of black people among the foreign-born (Schmidley and Gibson 1999, 25), immigration will typically decrease the proportion of these as a percentage of the population. It may be expected to have a statistically-negative impact on measures of family disruption (% female-headed households, % divorced), as divorce is less common among immigrants (USCB 2000b); Ousey and Kubrin (2009) discuss revitalization of traditional family structures as a mechanism by which immigration may decrease crime. They also point out that recent immigrants tend to be younger than the population average (Ousey and Kubrin 2009, 449), which is why it seems problematic to control for age structure when assessing the impact of *recent* immigration on crime. In the short run, immigration may be expected to increase the adult-to-child ratio, as the foreign-born are underrepresented among the underage (USCB 2000c).

Economic and educational variables. Immigrants are economically disadvantaged compared to US natives (USCB 2000a, b). Immigration may also have contextual effects on economic outcomes. Studies tend to show that immigration has a beneficial effect on economic outcomes in the current US, although there may be a negative impact on the incomes of low-skilled workers (for reviews, Borjas 2008; Peri 2010). Controlling for economic measures hence seems likely to cause overcontrol (Sampson 2008, 32). Some studies include black people as a percentage of the population and/or female-headed households as a percentage of all households in their summary measure of economic disadvantage. Both are problematic for the reasons outlined above. Immigrants are underrepresented among those holding a high school or equivalent diploma (Schmidley and Gibson 1999, 5), hence we should expect a negative correlation between the two measures. They are overrepresented among those working in low-skilled jobs and will hence be drawn to places where such jobs are available (Shihadeh and Barranco 2010a).

Other variables. Most of the other control variables used, such as spatial lags and region dummies, cannot cause overcontrol. However, Ousey and Kubrin (2009) speculate that immigration may foster fear among residents and hence lead to increases in the size of the police force. If so, the use of this control variable (Feldmeyer and Steffensmeier 2009) is problematic.

With the exception of Butcher and Piehl's (1998), all estimates of the effect of immigrant presence on crime come from regressions that control for some of the variables discussed above. It is hence unclear what to make of these results. If we nonetheless assume that neither under- nor overcontrol is a problem, we are still faced with the problem of ascertaining the direction of causality. The cross-sectional literature always assumes that an association between immigration and crime represents an influence of the former on the latter. But it is reasonable to hypothesize that immigrants prefer to live in cities low in crime, other things equal. This view is supported by a study of Latin Americans living in their countries of birth by Wood et al. (2010), who estimate that criminal victimization increases intentions to migrate to the US by about 30%, suggesting a strong preference

for low-crime environments among potential migrants. There may hence be an influence of crime on immigration rates.

Longitudinal studies

A smaller number of studies use longitudinal designs. Most of these control for unit fixed effects. Fixed effects designs are attractive in terms of assessing causality because they allow for the control of *unmeasured* influences on the outcome variable insofar as these influences are stable within units over time (Brüderl 2010; Halaby 2004), but cannot cause overcontrol. Other variables can be added to control for time-varying influences on Y, but these may cause overcontrol. Conversely, when important time-varying control variables are omitted, this can cause bias due to undercontrol (Bjerk 2009). Using a fixed effects design does not solve the problem of causal direction (Brüderl 2010, 992). Table 3 displays results from longitudinal studies estimating the effect of immigration on crime, again omitting analyses that use immigration measures as mere control variables (Jensen 2001; Moody and Marvell 2010).³

Table 3: Overview of Longitudinal Studies of Immigration and Crime

Study	Years	Units	Dependent	Measure Immigrants	Bivariate			Full Model			
					Im	Im	RS	% Lat.	Other demographic	Economic/Educational	Other
<i>Random effects estimates</i> Stowell et al. 2009	1995-2004	SMSAs	Violent Crime	% born outside US/% Latino	-*	-*	?af		5 ^{a, b, c, d, e}	2 ^{f, g}	5 ^{h, i, j, k, l}
<i>Fixed effects estimates</i> Butcher and Piehl 1998	1981-1984; 1986-1990	SMSAs	Overall crime	% foreign born < 1yr./%foreign born	-	-/+		+*	4 ^{b, c, m, n}	2 ^{f, o}	0
Ousey and Kubrin 2009	1981; 1991; 2001	Cities	Violent crime Property crime	“immigration index” ^{aae}	-*	-	-		4 ^{b, c, e, p}	4 ^{f, q, r, s}	2 ^{i, t}
Martinez, Stowell, and Lee 2010	1980; 1990; 2000	Neighborhoods	Homicide	% foreign born	-*	-	-		1 ^d	2 ^{g, r}	0
Wadsworth 2010	1990; 2000	Cities	Homicide	% foreign born/% foreign born < 5 yrs.		+/-	-		4 ^{a, b, c, e}	4 ^{u, v, w, x}	0
			Robbery	% foreign born/% foreign born < 5 yrs.		-	+	+*	4 ^{a, b, c, e}	4 ^{u, v, w, x}	0
Kreager, Lyons, and Hays 2011	1990-2000	Neighborhoods	Overall crime	% foreign born		*/+			2 ^{b, e}	2 ^{y, z}	6 ^{aa, ab, ac, ad}

Abbreviations:

Im: Immigration
 RS: Residential Stability
 EH: Ethnic/racial heterogeneity
 % Lat.: % Latino

Variables footnotes:

a Divorce rate
 b % African American
 c % young males
 d Adult-to-child-ratio
 e Population size
 f Unemployment rate
 g Summary measure: economic deprivation
 h Firearms availability
 i Police per capita
 j Dummy census region South

k: West
 l: Northeast (dummy)
 m: % female
 n: % of population in central city
 o: Mean wage
 p: Family instability
 q: % poor
 r: % professional
 s: % in manufacturing
 t: Drug markets
 u: "Job access"

v: "Young adult education"
 w: White/black economic dissimilarity
 x: White/Latino economic dissimilarity
 y: % with college degree
 z: Mean household income
 aa: % homes built recently
 ab: Mean mortgage investment
 ac: Dummy "gentrifying tracts"
 ad: Dummy "appreciating tracts"
 ae: % immigrants < 10 yrs./% poor English/% Latino
 af: Direction of effect is unclear, as variable comprises % who moved and % living in owner-occupied buildings

One of the studies (Stowell et al. 2009) uses a random, rather than a fixed, effects model, thus failing to exploit the central advantage of longitudinal data (see, generally, Halaby 2004). However, the fixed effects estimates are not without problems either. Again, the control variables used are displayed in the table. It is apparent that the full models from the papers represented in the table are subject to potential overcontrol problems. The arguments are analogous to those made with respect to the cross-sectional research and need not be repeated here.

Four papers also present bivariate estimates. These are unproblematic in terms of overcontrol, but may be subject to undercontrol. One might nonetheless trust these estimates more than the multivariate ones. Indeed, this is the logic implicit in Ousey and Kubrin (2009). Although they never say so explicitly, these authors appear to interpret the bivariate association as the best estimate of the effect; other independent variables subsequently entered into the regression are interpreted as mediators. If we accept this reasoning, the direction of the effect remains unclear: do rising rates of immigrants drive down crime rates, do drops in crime rates attract immigrants, or does the estimate represent a mixture of both effects? This question cannot be answered by the studies reviewed here.

Use of unweighted estimates

The studies on immigration and crime reviewed in the previous sections all use data from units that differ in population size, but none presents results weighted by this variable.⁴ This means, for example, that the crime rate of Erie, Pennsylvania (population ~ 100,000) contributes as much to the estimate as New York City's (population ~ 8,000,000). This seems inappropriate. The extrapolation of such findings to the national level (Sampson 2006, 2008) is particularly problematic.

Summary of problems with extant studies

Studies have failed to find consistent positive effects of the presence of immigrants on offending. However, the shortcomings of these studies are so severe that they are hardly informative about the effect of immigration on crime, for three reasons. *First*, some studies apply negative binomial regression inappropriately, leaving the results uninterpretable. *Second*, all studies fall well short of fulfilling the requirements for the causal interpretation of a statistical association. They almost certainly suffer from a mix of undercontrol and overcontrol problems. In particular, many variables that are, in part, outcomes of immigration are controlled for in the vast majority of models. Even if we are willing to overlook this problem, it must be noted that the direction of causality cannot be established by these studies. *Third*, no study uses a weighting procedure that takes into account the differences in population size between the units of analysis. Given the quality of the available evidence, it seems prudent to declare ignorance regarding the effect of immigration on crime. This state of affairs is particularly regrettable because simple solutions are available to increase the quality of the statistical analyses. These are implemented in the following section.

Estimating the concurrent effect of recent immigration on crime in a weighted sample using the selection approach

In this section I present an estimate of the concurrent effect of immigration on crime using a statistical model which solves the aforementioned problems. The first problem discussed above is immaterial given that linear least squares rather than negative binomial regression is employed. The second problem – overcontrol – is solved by using a selection approach. In line with previous practice, I use *recent* rather than overall immigration as the target variable and estimate its effect on homicide. Departing from previous practice, I control for variables which predict recent immigration, were fixed before that immigration happened, and which may also covary with crime. The third problem – not weighting the sample – is solved by weighting all regressions with a variable that is proportional to the population size. The resulting regression is longitudinal in the sense that data from two years is used, but cross-sectional in the sense that the dependent and the target variable are measured concurrently. Estimation of a “true” longitudinal model is beyond the scope of this paper.

Data and methods

I estimate the effect of recent immigration, measured in the year 2000, on 2000 homicide rates in a sample of US cities. All other independent variables refer to the year 1990 or are time-invariant. Cities are included in the sample if (i) their population in 2000 is at least 100,000, (ii) the population estimate given in census data is in reasonable agreement with that given in the Uniform Crime Reports⁵, and (iii) they exhibit complete data on all variables of interest. All regressions are run using weights which are proportional to the population size of the city in question. These are calculated by dividing a city's population size by the mean of the population size. Unless noted otherwise, variables are taken from the US census (see USCB 2007, n.d.).

The dependent variable, homicides per 100,000 population, is calculated on the basis of homicide counts and population data from the FBI's Uniform Crime Reports. The target variable represents foreign-born residents who migrated to the USA during the ten years prior to the census date as a percentage of the total population. The following 1990 or time-invariant variables serve as controls.

The presence of immigrants is higher in areas located close to typical points of entry to the US, namely the Mexican border, New York, and Los Angeles (Ottaviano and Peri 2006). I hence include the logged distance to these points, as well as a dummy variable indicating whether a city is located in a state that shares a border with Mexico. Additionally, I use dummy variables for a city's location in the Southern, Northeastern, or Midwestern census region. Population size is also included, as immigrants to the USA settle predominantly in more urban areas (Gibson and Jung 2006).

As discussed above, immigrants may prefer low-crime environments. The 1990 homicide rate is included as a measure of the most severe and most widely reported crime. A measure of street markets for crack cocaine is also used; the public sale of crack is widely seen as having led to public violence and the deterioration of neighborhoods (Bowling 1999; Levitt and Venkatesh 2000) and may thus have discouraged migrants from settling in cities which exhibit such markets. On the other hand, Ousey and Kubrin (2009, 451) speculate that immigrants may be drawn to selling drugs, which suggests settlement in areas where such markets exist. To approximately measure the presence of markets for crack, I use the data developed by Fryer et al. (2005).

It seems reasonable to assume that immigrants are pulled towards localities rich in economic opportunities (Borjas 2008). I hence control for the percentage of the population that is poor, the unemployment rate, the Gini index of income inequality, and the median household income. I also include the percentage of the population that is black, as immigrants typically compete with African Americans for jobs in the low-skilled sector (Shihadeh and Barranco 2010a).

The presence of immigrants already in a city fosters further immigration to it if migrants settle close to spouses, friends, or relatives and in proximity to ethnic networks. The rate of the population that is foreign-born is hence controlled for, as well as the rates of the populations that share the same racial/ethnic backgrounds with most recent immigrants, Asians and Hispanics.

Results

A preliminary WLS regression of the recent immigrants rate, measured in the year 2000, on 1990 and invariant predictor variables shows that these explain 92% of the variance in the immigrants rate (detailed results available upon request). The central results from the regressions of interest are shown in table 4. To preserve space, and so as not to mislead the reader into interpreting coefficients for control variables as causal, only the results for the recent immigrants measure are shown (detailed results available upon request). Model 1 includes the full set of controls discussed above. The coefficient is substantial and negative, implying that an increase in the rate of recent immigrants by about 4 percentage points leads to a reduction in the homicide rate of almost 1 per 100,000 population. However, the coefficient is not statistically significant.

An additional analysis is performed to accommodate a possible concern with the inclusion of the 1990 homicide rate as a control variable. Using a lagged outcome variable as a predictor is

equivalent to an analysis of *changes* in the dependent variable (Finkel 2008, 486-487). This means that model 1 shows the prediction of 1990-2000 *changes* in the homicide rate by 1990-2000 immigrant rates, conditional on control variables other than the 1990 homicide rate. This posits a problem for the interpretation of the results. While one may be satisfied that the coefficient for the recent immigrants rate represents its causal connection to the outcome, it is reasonable to interpret it as an influence of 1990-2000 changes in homicide rates in specific cities on decisions to migrate to these cities. That is, this analysis leaves us uncertain about the causal direction.

Table 4: Results of WLS Regressions of 2000 Homicide Rates on the 2000 Recent Immigrants Rate and 1990/Time Invariant Control Variables

Model	Excluded variables	B	Beta	t
1	none	-0.26	-0.15	-0.66
2	1990 homicide rate	-0.16	-0.09	-0.39

Note: Weighted N=91. All models include a constant. Both models control for the following variables, which are time-invariant of refer to the year 1990: population size, logged distance to Los Angeles, New York, and the Mexican border, location in a state that borders on Mexico, in a Southern, Northeastern or Midwestern state (dummies), Fryer et al.'s crack index, % poor, % unemployed, Gini, median household income in the previous year, % black, % foreign born, % Asian, % Hispanic. Model 1 additionally controls for the homicide rate.

This problem can be rectified by excluding the 1990 homicide rate as a predictor. Results for the preliminary regression explaining the recent immigrants rate show that the explained variance is 92% (detailed results available upon request). The regression of interest is model 2. The result is reasonably similar to the one obtained using model 1.

General summary and discussion

Available studies of immigration and crime in the contemporary USA suffer from severe methodological shortcomings. Some use negative binomial regression incorrectly, and none present the results of regressions which take into account the fact that the units of analysis differ in population size. Extant studies also suffer from a mixture of under- and overcontrol. The latter problem is particularly prominent. Researchers routinely control for variables which theory, empirical research, or both suggest are themselves influenced by immigration. This makes this body of research an inappropriate basis for drawing conclusions about the effect of immigration on crime with a desirable degree of confidence. I have pointed out that there is an approach to statistical control in the regression model that circumvents the overcontrol problem. This is the selection approach, in which variables that may influence the target variable (and that might also be linked to the outcome) are controlled for.⁶ It is hoped that students of crime will consider this technique more often when designing their analyses.

The present study finds substantial but insignificant negative associations between rates of recent immigrants and homicide. While these estimates replicate the typical finding from previous research, it may be interpreted as representing a causal connection between the two variables with considerably more confidence than should be attached to previous findings. Because the processes represented by the recent immigrants variable took place before those represented by the homicide rate, we can also be quite certain that the causal arrow runs from recent immigrants to homicide, not *vice versa*. The findings imply that an increase in the rate of recent immigrants by about 3 percentage points leads to a reduction of the homicide rate of about 1 per 100,000. The findings thus lend weak

support to perspectives that aim to explain why immigration might decrease crime. This article may hence be seen as an encouragement to further pursue such perspectives.

References

- Akins, Scott, Rúben G. Rumbaut, and Richard Stansfield. 2009. "Immigration, Economic Disadvantage, and Homicide: A Community-level Analysis of Austin, Texas." *Homicide Studies* 13:307-14.
- Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton and Oxford: Princeton University Press.
- Bjerk, David. 2009. "How Much Can We Trust Causal Interpretations of Fixed-Effects Estimators in the Context of Criminality?" *Journal of Quantitative Criminology* 25:391-417.
- Borjas, George J. 2008. "Immigration." in: *Concise Encyclopedia of Economics*, edited by David R. Henderson. Accessed August 7, 2012. <http://www.econlib.org/library/Enc/Immigration.html>.
- Bowling, Benjamin. 1999. "The Rise and Fall of New York Murder: Zero Tolerance or Crack's Decline?" *British Journal of Criminology* 39:531-54.
- Brimelow, Peter. 1995. *Alien Nation: Common Sense about America's Immigration Disaster*. New York: Random House.
- Brüderl, Josef, (2010). "Kausalanalyse mit Paneldaten." In *Handbuch der Sozialwissenschaftlichen Datenanalyse*, edited by Christof Wolf and Henning Best, 963-94. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Butcher, Kristin F., and Anne Morrison Piehl. 1998. "Cross-city Evidence on the Relationship between Immigration and Crime." *Journal of Policy Analysis and Management* 17:457-93.
- Desmond, Scott, and Charis E. Kubrin. 2009. "The Power of Place: Immigrant Communities and Adolescent Violence." *The Sociological Quarterly* 50:581-607.
- Elwert, Felix, and Christopher Winship. 2011. *Endogenous Selection Bias*. Accessed November 7, 2012. <http://www.google.de/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&ved=0CCMQFjAA&url=http%3A%2F%2Fipl.econ.duke.edu%2Fseminars%2Fsystem%2Ffiles%2Fseminars%2F138.pdf&ei=fiaUNvEiIbP0QWYpIHwAw&usg=AFQjCNHP004FlmSkvBiehn9Q4WQolo2qA>
- FBI [Federal Bureau of Investigation]. 2002. *Uniform Crime Reporting Program Data [United States]: Offenses Known and Clearances by Arrest, 2000*. Ann Arbor: Inter-university Consortium for Political and Social Research.
- Finkel, Steven E. 2008. "Linear Panel Analysis." In *Handbook of Longitudinal Research*, edited by Scott Menard, 474-504. Boston: Elsevier Press.
- Gibson, Campbell, and Kay Jung. 2006. *Historical Census Statistics of the Foreign-born Population of the United States, 1850 to 2000*. Washington, D.C.: US Census Bureau. Accessed August 8, 2012. <http://www.census.gov/population/www/documentation/twps0081/twps0081.pdf>.
- Hagan, John, and Alberto Palloni. 1998. "Immigration and Crime in the United States." In *The Immigration Debate: Studies on the Economic, Demographic, and Fiscal Effects of Immigration*, edited by James P. Smith and Barry Edmoston, 367-87. Washington, D.C.: National Academy Press.
- Halaby, Charles N. 2004. "Panel Models in Sociological Research." *Annual Review of Sociology* 30:507-44.
- Heckman, James J. 2005. "The Scientific Model of Causality." *Sociological Methodology* 35:1-97.
- Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81:945-60.
- Jensen, Gary F. 2001. "The Invention of Television as a Cause of Homicide: The Reification of a Spurious Relationship." *Homicide Studies* 5:114-30.
- Kreager, Derek A., Christopher Lyons, and Zachary R. Hays. 2011. "Urban Revitalization and Seattle Crime, 1982-2000." *Social Problems* 58:615-39.
- Lee, Matthew T., and Ramiro Martinez. 2002. "Social Disorganization Revisited: Mapping the Recent Immigration and Black Homicide Relationship in Northern Miami." *Sociological Focus* 35:363-80.
- Lee, Matthew T., Ramiro Martinez, jr., and Richard Rosenfeld. 2001. "Does Immigration Increase Homicide? Negative Evidence from Three Border Cities." *Sociological Quarterly* 42:559-80.

- Legewie, Joscha. 2012. "Die Schätzung von kausalen Effekten: Überlegungen zu Methoden der Kausalanalyse anhand von Kontexteffekten in der Schule." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 64:123-53.
- Levitt, Steven D., and Sudhir Alladi Venkatesh. 2000. "An Economic Analysis Of A Drug-Selling Gang's Finances." *Quarterly Journal of Economics* 115:755-89.
- Martinez, Ramiro jr. 2000. "Immigration and Urban Violence: The Link between Immigrant Latinos and Types of Homicide." *Social Science Quarterly* 81:363-74.
- 2006. "Coming to America: The Impact of the New Immigration on Crime." In *Immigration and Crime: Race, Ethnicity, and Violence*, edited by Ramiro Martinez jr. and Abel Valenzuela jr., 1-19. New York and London: New York University Press.
- Martinez, Ramiro jr., and Matthew T. Lee. 1998. "Immigration and the Ethnic Distribution of Homicide in Miami, 1985-1995." *Homicide Studies* 2:291-304.
- 2000a. "On Immigration and Crime." In *Criminal Justice 2000*. Vol. 1: *The Nature of Crime: Continuity and Change*, edited by Gary LaFree, 485-524. Washington, D.C.: US Department of Justice, Office of Justice Programs.
- 2000b. "Comparing the Context of Immigrant Homicides in Miami: Haitians, Jamaicans and Mariels." *International Migration Review* 34:794-812.
- Martinez, Ramiro jr., Matthew T. Lee, and Amie L. Nielsen. 2004. "Segmented Assimilation, Local Context and Determinants of Drug Violence in Miami and San Diego: Does Ethnicity and Immigration Matter?" *International Migration Review* 38:131-57.
- Martinez, Ramiro jr., Jacob I. Stowell, and Jeffrey M. Cancino. 2008. "A Tale of Two Border Cities: Community Context, Ethnicity, and Homicide." *Social Science Quarterly* 89:1-16.
- 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, 1980-2000." *Criminology* 48:797-829.
- Merton, Robert K. 1955: "The Socio-cultural Environment and Anomie [with discussion]." In *New Perspectives for Research on Juvenile Delinquency*, edited by Helen L. Witmer and Ruth Kotinsky, 24-50. Washington, D.C.: US Department of Health, Education, and Welfare.
- 1964. "Anomie, Anomia, and Social Interaction: Contexts of Deviant Behavior." In *Anomie and Deviant Behavior: A Discussion and Critique*, edited by Marshall B. Clinard, 213-42. New York: The Free Press/London: Collier Macmillan.
- 1968. *Social Theory and Social Structure*. 1968 enlarged edition. New York: The Free Press.
- Moody, Carlisle E., and Thomas B. Marvell. 2010. "On the Choice of Control Variables in the Crime Equation." *Oxford Bulletin of Economics and Statistics* 72:696-715.
- Morenoff, Jeffrey D., Robert J. Sampson, and Stephen W. Raudenbush. 2001. "Neighbourhood Inequality, Collective Efficacy, and the Spatial Dynamics of Urban Violence." *Criminology* 39:517-59.
- Morgan, Stephen Lawrence, and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge: Cambridge University Press.
- Nielsen, Amie, and Ramiro Martinez, jr. 2009. "The Role of Immigration for Violent Deaths." *Homicide Studies* 13:274-87.
- Nielsen, Amie, 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 Communities." *Criminology* 43:837-72 .
- Olson, Christina Polczynski, Minna K. Laurikkala, Lin Huff-Corzine, and Jay Corzine. 2009. "Immigration and Violent Crime: Citizenship Status and Social Disorganization." *Homicide Studies* 13:227-41.
- Osgood, D. Wayne. 2000. "Poisson-Based Regression Analysis of Aggregate Crime Rates." *Journal of Quantitative Criminology* 16:21-43.
- Ottaviano, Gianmarco I.P., and Giovanni Peri. "The Economic Value of Cultural Diversity: Evidence from US Cities." *Journal of Economic Geography* 6:9-44.
- Ousey, Graham C., and Charis E. Kubrin. 2009. "Exploring the Connection between Immigration and Violent Crime Rates in U.S. Cities, 1980-2000." *Social Problems* 56:447-73.

- Peri, Giovanni. 2010. *The Effect of Immigrants on U.S. Employment and Productivity*. San Francisco: Federal Reserve Bank of San Francisco. Accessed August 8, 2012. <http://www.frbsf.org/publications/economics/letter/2010/el2010-26.pdf>.
- Phillips, Julie A. 2002. "White, Black, and Latino Homicide Rates: Why the Difference?" *Social Problems* 49:349-74.
- Reid, Lesley Williams, Harald E. Weiss, Robert M. Adelman, and Charles Jaret. 2005. "The Immigration-Crime Relationship: Evidence across US Metropolitan Areas." *Social Science Research* 34:757-80.
- Rubin, Donald B. 1974 "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66:688-701
- 2005. "Causal Inference Using Potential Outcomes: Design, Modeling, Decisions." *Journal of the American Statistical Association* 100:322-31.
- Sampson, Robert J. 2006. "Open Doors Don't Invite Criminals." *New York Times*, March 11.
- 2008. "Rethinking Crime and Immigration." *Contexts* 7:28-33.
- Schmidley, A. Dianne and Campbell Gibson. 1999. *Profile of the Foreign-Born Population of the United States: 1997*. Washington, D.C.: U.S. Government Printing Office.
- Sellin, Thorsten, 1938. *Culture Conflict and Crime*. New York: Social Science Research Council.
- Shaw, Clifford R., and Henry D. McKay. 1931. *Social Factors in Juvenile Delinquency: A Study of the Community, the Family, and the Gang in Relation to Delinquent Behavior*. Washington, D. C.: United States Government Printing Office.
- 1974. *Juvenile Delinquency and Urban Areas: A Study of Rates of Delinquency in Relation to Differential Characteristics of Local Communities in American Cities*. Revised edition, 2nd impression. Chicago: University of Chicago Press.
- Shihadeh, Edward S., and Raymond E. Barranco. 2010a. "Latino Employment and Black Violence: The Unintended Consequence of U.S. Immigration Policy." *Social Forces* 88:1393-1420.
- 2010b. "Latino Employment and Non-Latino Homicide in Rural Areas: The Implications of U.S. Immigration Policy." *Deviant Behavior* 31:411-39.
- Shihadeh, Edward S., and Lisa Winters. 2010. "Church, Place, and Crime: Latinos and Homicide in New Destinations." *Sociological Inquiry* 80:628-49.
- Simon, Rita J., and Keri W. Sikich. 2007. "Public Attitudes toward Immigrants and Immigration Policies across Seven Nations." *International Migration Review* 41:956-62.
- Stowell, Jacob I. 2007. *Immigration and Crime: The Effects of Immigration on Criminal Behavior*. New York: LFB.
- Stowell, Jacob I., and Ramiro Martinez, jr. 2007. "Displaced, Dispossessed, or Lawless? Examining the Link between Ethnicity, Immigration, and Violence." *Aggression and Violent Behavior* 12:564-81.
- 2009: "Incorporating Ethnic-Specific Measures of Immigration in the Study of Lethal Violence." *Homicide Studies* 13:315-24.
- 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* 47:889-928.
- USCB [United States Census Bureau]. 2000a. "Medians and Standard Errors for Various Characteristics by Citizenship Status: March 1995." Accessed August 8, 2012. <http://www.census.gov/population/foreign/files/cps1995/tab0116.pdf>.
- 2000b. "Marital Status of the Population by Citizenship Status, Age and Sex: March 1995." Accessed August 8, 2012. <http://www.census.gov/population/foreign/files/cps1995/tab0102.pdf>.
- 2000c. "Population by Citizenship Status, Age and Sex: March 1995". Accessed August 8, 2012. <http://www.census.gov/population/foreign/files/cps1995/tab0111.pdf>.
- 2007. *Census 2000 Summary File 3 Technical Documentation*. Washington, D.C.: US Census Bureau.
- n.d. *Summary Tape File 3 Technical Documentation on CD-ROM*. Washington, D.C.: US Census Bureau.

- 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* 91:531-53.
- Winship, Christopher, and Stephen L. Morgan. 1999. "The Estimation of Causal Effects from Observational Data." *Annual Review of Sociology* 25:659-707.
- Winship, Christopher, and Michael Sobel. 2001. "Causal Inference in Sociological Studies." In *Handbook of Data Analysis*, edited by Melissa Hardy and Alan Bryman, 481-504. London: Sage.
- Wirth, Louis. 1931. "Culture Conflict and Delinquency I: Culture Conflict and Misconduct." *Social Forces* 9:484-92.
- Wood, Charles H., Chris L. Gibson, Ludmila Ribeiro and Paula Hamsho-Diaz. 2010. "Crime Victimization in Latin America and Intentions to Migrate to the United States." *International Migration Review* 44:3-24.
- Wright, Emily M. and Michael L. Benson. 2010. "Immigration and Intimate Partner Violence: Exploring the Immigrant Paradox." *Social Problems* 57:480-503.

Endnotes

¹ Parts of this article are loosely based on my dissertation, supervised by Thomas Ohlemacher at the University of Hildesheim. I am also grateful to Robert Putnam for helpful comments on this paper's topic. All remaining errors are mine.

² If studies used measures of *instability*, the sign of the coefficient was reversed, so that the table reports results for stability in all cases.

³ Kreager, Lyons and Hays (2011) is a borderline case. The authors' focus is the influence of gentrification on crime and they introduce their immigration measure in the section titled "control variables." However, they write: "Sampson [...] recently made the provocative claim that the 1990s crime drop resulted from substantial increases in first-generation immigrants residing in inner-city ethnic enclaves. We test this assertion." This study is hence included.

⁴ Wadsworth (2010, 547) reports rerunning his regressions weighting by *logged* population size, with results "virtually identical" to his preferred specifications.

⁵ In most cases, population sizes given by the Uniform Crime Reports and the census match exactly. Deviations up to 3% were tolerated; in these cases, population estimates from the Uniform Crime Reports were used.

⁶ Elwert and Winship (2011) discuss cases in which controlling for pre-treatment variables may increase rather than decrease the bias of the estimate. They recommend that researchers who aim to minimize this problem choose pre-treatment controls on the basis of a theoretical understanding of the processes involved. I have hence chosen from the almost infinite array of possible pre-treatment variables controls which may reasonably be assumed to influence both immigration and homicide. Nonetheless, the reader should be aware that controlling for pre-treatment variables does not come without risk.