doi:10.1093/bjc/azw039 BRIT. J. CRIMINOL. (2017) **57**, 1080–1100
Advance Access publication 11 April 2016

A COMPREHENSIVE EVALUATION OF THE ASSOCIATION BETWEEN PERCENT YOUNG AND CROSS-NATIONAL HOMICIDE RATES

Meghan L. Rogers* and William Alex Pridemore

Is there an association between the proportion of the population that is young and national homicide rates, and when testing other theories cross-nationally is it necessary to control for this concept? To answer these questions, we carried out an extensive review of the empirical literature and then used data for the years 1999–2004 from a sample of 55 nations to test two predominant hypotheses: Percent young is significantly associated with homicide victimization rates across nations, and percent young accounts for a significant proportion of the overall variance in homicide victimization rates across nations. The results consistently indicated no significant association between percent young and homicide victimization rates across nations. Moreover, including percent young in models of cross-national homicide rates likely has negative ramifications for model fit. We situate these findings within the larger literature and provide a discussion of the implications for future cross-national homicide research.

Key Words: cross-national homicide victimization, percent young, model fit

It has become a stylized fact that national homicide rates are largely dependent upon the relative size of their young population. About half of all studies of the structural covariates of cross-national homicide rates include some measure of the proportion of the total population that is young, suggesting the importance of percent young in cross-national homicide research. Many scholars provide little or no justification for percent young's inclusion (Pridemore and Trent 2010; Nivette 2011), however, leading readers to infer the reasons for including percent young (often defined as the percent of the entire population aged 15–24 years), which may include any one of several different individual- or population-level explanations (South and Messner 2000). Similarly, previous descriptive literature from several nations suggests the 15–24 age group does not consistently have the highest age-specific homicide victimization rates (Birkel and Dern 2012; Ganpat and Liem 2012; Lehti and Kivivuori 2012; Liem *et al.* 2012; Preti and Maccio 2012; Rogers 2014).

If percent young is not significantly associated with cross-national homicide rates, if the theoretical justification for its inclusion is uncertain and if its inclusion in models is unnecessary for other reasons, then its inclusion could be detrimental to the statistical model. Including percent young would overfit the model. Overfitting can introduce additional error in estimations and increase overall variance, reducing model fit and effect sizes of other parameters, all of which are harmful to the inferences drawn for other variables (Myers 1990; Hocking 2003; Agresti and Finlay 2008). If percent young

*Meghan L. Rogers, School of Criminal Justice, University at Albany–State University of New York, 135 Western Avenue, Albany, NY 12222, USA; mlrogers@albany.edu; William Alex Pridemore, School of Criminal Justice, University at Albany–State University of New York, 135 Western Avenue, Albany, NY 12222, USA; pridemore@albany.edu.

is a consistent structural covariate and it is excluded from a statistical model this would bias the results, which is also harmful to the inferences drawn from the model.

In this study, we systematically explore the efficacy of percent young in accounting for variation in homicide victimization across nations. We first undertake an extensive review of the empirical literature to assess what we know about the association between percent young and cross-national homicide rates. We then carry out our own tests of the theoretical arguments for the potential compositional effects of percent young on homicide rates. Next, we systematically explore whether percent young is a necessary control variable in cross-national homicide models. While it is not possible to state that percent young always or never matters, the methods we utilize can be employed by others to better construct models that are not overfitted. Therefore, the systematic study of this specific variable also allows for a more generally informed discussion of the importance of model selection in cross-national homicide research.

Literature Review

Compositional effect

A compositional effect is one key theoretical argument for why percent young might be an important predictor of national homicide rates (LaFree 1999). It proposes that having a larger proportion of the overall population that is in the young age group, which many scholars believe is most at risk for offending or victimization, will increase population-level homicide rates (Fiala and LaFree 1988; Gartner 1990; Bennett 1991; Ortega *et al.* 1992). The argument is that the population most at risk for violent offending and victimization are those aged 15–24 (Hirschi and Gottfredson, 1983; Cohen and Land 1987; Fiala and LaFree 1988), though in the cross-national homicide literature the specific age group tested sometimes varies.

Early research in the United States by the President's Commission on Law Enforcement and Administration of Justice (1967) and Wolfgang (1967) observed that younger people were more likely to be offenders and victims. One of the strongest proponents of this age effect has been Hirschi and Gottfredson (1983) who argued the association between age and crime is 'invariant across social and cultural conditions' (Hirschi and Gottfredson 1983: 554) and that no other 'variable or combination of variables currently available to criminology' could account for the association between youth population and crime (Hirschi and Gottfredson, 1983: 554). The age–crime curve literature argues that the youthful population will most often have the highest proportion of offenders (Hirschi and Gottfredson 1983; Gottfredson and Hirschi, 1990; Marvell and Moody 1991). While the reason why the youthful population is more likely to commit crime and become victims is debated (Blumstein and Nagin 1974; Greenberg 1977; Hirschi and Gottfredson 1983; Greenberg 1985; Gottfredson and Hirschi 1990), for decades most researchers agreed the proportion of the population that is young is a driving mechanism for higher crime rates within and across nations (Chilton and Spielberger 1971; Greenberg 1977; Hirschi and Gottfredson 1983; Gartner and Parker 1990; Pampel and Williamson 2001).

Within the cross-national and US homicide literature some have connected the compositional argument for the association between percent young and homicide to reasoning within the routine activities theoretical (RAT) framework (Fiala and LaFree 1988; Gartner 1990; Bennett, 1991). RAT postulates victimization will occur when a motivated offender, a suitable target and the absence of a capable guardian converge temporally and spatially (Cohen and Felson, 1979). Therefore, in nations with a higher proportion of the population at risk for offending and victimization there is greater availability of offenders and potential targets that are present relative to other nations.

Empirical support for the compositional argument is inconclusive (Nivette, 2011). Within the US literature, the invariance of the association between age and crime has been questioned and there is some support for the lack of a time-invariant association between age and crime (Shavit and Rattner 1988; Steffensmeier *et al.* 1989). Within the cross-national research, the invariance of the age–crime curve has not been directly tested and reviews show there are many more studies that find a nonsignificant association between percent young and homicide rates than there are that find the predicted positive association (Pridemore and Trent 2010; Nivette 2011; Trent and Pridemore 2011; Rogers and Pridemore 2016).

If the compositional argument is accurate we would expect the 15–24 age group to exhibit the highest age group-specific homicide rates in most nations, especially those where 15 to 24-year-olds are the largest age group within the nation's population pyramid. However, recent research from several nations provides evidence against this hypothesis. Rogers (2014) found that in most nations and years, the 15 to 24-year-old age group did not have the highest age-specific homicide victimization rate, even when they were the most populous age group. Studies of several individual nations found the 15 to 24-year-old age group did not have the highest homicide victimization rate, including Finland (Lehti and Kivivuori 2012; Liem *et al.* 2012); Sweden (Granath 2012; Liem *et al.* 2012); Lithuania (Andresen 2012); Ukraine, Belarus and Russia (Lysova *et al.* 2012; Pridemore 2003) and other East and Central European nations (Stamatel 2009).

Comprehensive review of the literature that included percent young as a variable

We searched the Internet and hard copies of relevant journals for studies that included homicide as a dependent variable, some measure of age as an independent variable and had a cross-national sample.[1] Thirty-two articles met these criteria.[2] Table 1 provides a brief summary of the associations between percent young and homicide rates across

Table 1 *Summary of results from previous research for cross-national homicide models that included a measure of age*

	N	%
Positive association	19	13
Negative association	7	5
Not significant	120	82

Total number of articles was 32, within which the total number of models that included a measure of age was 146.

[1]The criteria for selection were any published articles in any field in any year that had an outcome variable of homicide, a cross-national sample, inferential statistical analysis and a predictor for age structure. We did not include studies that used percent young within a principal component and/or as a latent variable because we were interested in the direct effect of percent young on homicide victimization.

[2]Table of articles utilized to draw this conclusion are available at https://sites.google.com/site/homicidedata/a-comprehensive-evaluation-of-the-association-between-percent-young-and-cross-national-homicide-rates

nations in the 32 articles. Few studies observed the direct positive effects proposed by the compositional theoretical arguments. Of the 146 coefficients for percent young generated within the 32 articles, it was positively and significantly associated with homicide rates only 13 per cent of the time ($n = 19$). Even in those articles that found a positive association for percent young that association was only significant under certain specifications and not across all models tested in the study (Messner 1989; Ortega *et al.* 1992; Pampel and Gartner 1995; Neumayer 2003; McCall and Nieuwbeerta 2007).

Five per cent ($n = 7$) of models showed a significant negative association between percent young and cross-national homicide rates. Eighty-two per cent ($n = 120$) of the models found a non-significant association between percent young and homicide rates, and several studies found consistent non-significant effects across all models in the study (Conklin and Simpson 1985; Avison and Loring 1986; Gartner 1990; Neapolitan 1994, 1997, 1998; Savolainen 2000; Altheimer 2008; Pridemore 2008; Cole and Gramajo 2009; Pridemore 2011; Chon 2012; Rogers and Pridemore 2013). Thus, while many scholars continue to accept as a stylized fact the impact of the relative size of the young population on national homicide rates, the evidence shows only 13 per cent of all models in the empirical literature found the expected positive association.

Hypotheses

We collected our own cross-national data to test the hypotheses that the percent of the entire population aged 15–24 will (1) be positively and significantly associated with homicide victimization rates and (2) significantly account for variation in homicide victimization rates. The subtle difference between the two hypotheses is that the first tests for an association between X and Y, while the second tests if X provides a unique contribution to the explained variance.

Methods

Sample

Table A1 in the Appendix provides a list of the 56 nations in our sample. The sample of nations is limited due to missing data on the outcome and independent variables and due to our desire to ensure the utilization of the most valid and reliable sources for measuring our theoretical concepts. Similarly small or smaller sample sizes are not uncommon in cross-national homicide research (Pridemore, 2008; Messner *et al.* 2010; Savolainen, 2000).[3]

Data

The outcome variable was the average homicide victimization rate for the 5-year period 1999–2004. We smoothed homicide rates to prevent extreme rates in a given year from influencing model estimation. When creating the average we ignored data that were missing for a particular year. The nation with the most missing years was Georgia

[3]We conducted a power analysis for all models and found that power ranged above 0.95 for all models except Models 11 and 12, for which power was 0.92.

(missing 2002–04). Seven other nations had one missing observation in the 5-year time span: Uzbekistan (2001), Hungary (2003), Armenia (2004), Australia (2004), Italy (2004), Portugal (2004) and Puerto Rico (2004).

We obtained homicide victimization and population data from the World Health Organization's (WHO) WHOSIS database (World Health Organization 2012). Homicide is defined using the International Classification of Diseases 10th revision, categories X85-Y09: 'homicides and injury purposely inflicted by another person' (World Health Organization 2012). While WHO often has homicide data available for fewer nations compared with other sources (e.g. United Nations), they are preferred because they are considered the most valid measure of homicide victimization across nations (Smit *et al.* 2004, 2012) because WHO enforces a uniform definition of homicide and includes quality checks to ensure the data obtained from nations are valid.

The key independent variable was the percent of the entire population aged 15–24, which we term 'percent young'. The 15–24 age group is the most common indicator of 'young' in the recent empirical literature. We obtained the total population within the age groups and the total population within nations from the WHOSIS database (World Health Organization 2012).

The selection of additional variables was dictated by the empirical literature. We did this to ensure we did not omit a known or speculated important structural covariate, since to claim that percent young overfits a statistical model we must guard against omitted variable bias. Table 2 provides a list of each of the variables and the data source. We used regional dummy variables in many models based on the Wald and likelihood ratio tests.

Analysis

We employed multiple statistical estimation techniques to test the hypotheses. We used ordinary least squares (OLS) regression to estimate direct effects between percent young and homicide victimization rates. We built the models such that we introduced variables based on sample attrition due to missing data. Initial models included all variables without missing data, and each variable was then included based on the number of nations lost to missing data. We did this in a stepwise process, so it is possible to

Table 2 *Variables, data sources, and means and standard deviations*

Variable	Source	Mean	SD
Homicide rate	WHO	5.96	7.28
Percent 15–24	WHO	15.43	2.66
Sex ratio	WHO	95.04	3.97
Infant mortality (poverty)	WHO	9.32	5.79
Total population	WHO	50,235,404	1,186,870
Region	Defined by WHO		
Gini coefficient	United Nations Human Development Reports	35.57	8.39
Education index	United Nations Human Development Reports	0.77	0.12
Unemployment rate	World Bank	8.63	4.25
Percent urban	World Bank	67.93	16.10
Crude divorce rate	United Nations Demographic Yearbook	1.91	1.05
GDP per capita	United Nations	11,977	11,361
Ethnic heterogeneity	Alesina *et al.* (2004)	0.31	0.20

observe whether the variable or sample loss affects the results. This aids in addressing the hypotheses because it is possible that in certain samples, with specific variables, there is a direct relationship between percent young and homicide rates.

We also employed sequential analyses of variance (ANOVAs) to test the association between percent young and homicide rates. A sequential ANOVA provides information regarding the unique contribution of a variable, given the variables included in the model before that variable. This method is also one way in which multicollinearity's effect on the model can be further explored. For instance, if the model were estimated as such:

$$y_{\text{homicide rates}} = \beta_0 + \beta X_{\text{Africa}} + \ldots + \beta X_{\ln\%15-24} + \beta X_{\ln\text{pov}} + \beta X_{\text{sex ratio}},$$

a sequential ANOVA in which percent young is the variable of interest would give information on the unique contribution of percent young, given that all variables in the equation before it are included in the model (Myers 1990). This estimation allows for a slightly different view of the relationship and helps us understand how much a variable like percent young aids the overall model in accounting for variation in cross-national homicide rates.[4] The method also provides insight into how multicollinearity affects the model and provides a method to explore whether a variable would have an association with homicide victimization despite being highly correlated with another independent variable in the model. If percent young does not overfit a model and its ability to account for variation in homicide across nations cannot be 'replaced by any other variable or combination of variables' (Hirschi and Gottfredson 1983), then it would maintain a significant association under the sequential ANOVA estimation technique because it would provide a unique contribution in accounting for variation in homicide rates.

We used the likelihood ratio test (LRT) and the Wald Test for model comparison. These techniques utilize the projection of the regression line in a model that includes all variables and subtracts out the effect of the projected regression line for a reduced model that excludes percent young (Agresti and Finlay 2008). This process provides an estimation of whether the variables removed from the full model, or left out of the reduced model, have a significant effect on the projection of the regression line and the error surrounding the projection.

Finally, we employed Mallows's C_p, which is a more precise method of model selection (Mallows 1973) that allows one to measure overfitting and underfitting of models. The value obtained from estimating Mallows's C_p is a ratio of bias-to-variance. Having too few variables in the model that contribute to explaining the variation leads to underfit biasing the model. Having too many variables in the model and including variables that do not contribute to explaining the variation results in overfit, thereby introducing more variance into the estimation of the outcome. Mallows's C_p estimates every possible combination of variables and compares the mean square error of each of the possible combinations of variables in the full model. To estimate Mallows's C_p, we used an initial model that included percent young, poverty, total population, ethnic heterogeneity, sex ratio, gini coefficient, the education index and unemployment. We removed the regional controls because Mallows's C_p cannot be estimated with multiple dummy control variables (Hocking 2003).

[4]It is not problematic the sequential ANOVA allows results to vary depending on the location in the model in which percent young is included. The key to this analysis is that no matter the variables used or a variable's location in the model, if the variable 'matters' in accounting for variation in the outcome then the variable will maintain a significant association.

Additional statistics

We used additional statistics to aid in understanding the efficacy of percent young in accounting for variation in homicide rates across nations, including the R^2, adjusted R^2, F statistic (i.e. the ratio of explained to unexplained variance), and mean squared residuals. When the rate of increase in R^2s is small, any additional variable added to the model is not beneficial, meaning the researcher should reconsider inclusion of the variable. Including too many variables overfits the model and increases the variance, thereby decreasing the ability of all the variables to account for variation in homicide across nations. Thus, we can gauge if including percent young is beneficial to the model based on the differences in R^2s in the full and reduced model. If the adjusted R^2 increases, stays the same or decreases only marginally when a variable is removed then that variable was not helping the model account for variation.

The final statistics were the F test for model fit and the mean square residuals (MSRs). The F statistic is the ratio of the variation accounted for (means of square regression) and not accounted for (MSRs) or the left over error in the model. The MSR is of interest beyond its value to calculating F statistics because it provides information on the amount of error in the model that cannot be accounted for by variables in the model (Myers 1990). It is expected that with each additional variable MSR will decrease. When the rate of decrease becomes incremental, the model will likely be overfitted by introducing additional variables, thereby introducing more variance into the model and the estimation. Specifically, if the model were to have a small increase in the MSR or even a smaller MSR when percent young is removed compared with a full model that includes percent young, then the model would be overfitted. This means the inclusion of the variable is not only not helping but actually hurting the model in accounting for variation in homicide victimization across nations.

Results

Descriptive statistics

We transformed several variables—homicide victimization rate, percent young, infant mortality, total population, ethnic heterogeneity, Gini coefficient and divorce rate—to normalize distributions. We used multiple techniques to test for normality, including box plots, QQ plot, Shapiro–Wilk and Box-Cox power transformation. Tukey's ladder transformations are preferred because they aid in maintaining the ability to interpret results (Hoaglin *et al.* 1983). We used a natural logarithmic transformation for homicide, percent young, infant mortality, total population and the Gini coefficient. We used a square root transformation for ethnic heterogeneity and divorce rate. Table 3 provides descriptive statistics and a correlation matrix. Percent young and homicide victimization rates were positively correlated ($r = 0.71$).

Model estimation

Table 4 provides results for the first set of analyses using OLS regression. The first model shows results for logged homicide victimization rates regressed on all variables with no missing data. At each step, percent young was removed from the 'full' model to

Table 3 *Descriptive statistics and correlation matrix*

	1	2	3	4	5	6	7	8	9	10	11
1. Ln homicide	–										
2. Ln % 15–24	**0.71**	–									
3. Ln poverty	**0.78**	**0.84**	–								
4. Ln population	0.02	−0.06	0.00	–							
5. % Urban	−0.05	**−0.33**	**−0.32**	**0.31**	–						
6. Sex ratio	−0.25	0.11	−0.11	0.16	0.11	–					
7. Sq ethnic heterogeneity	**0.62**	**0.49**	**0.59**	−0.13	−0.06	−0.10	–				
8. Ln Gini	**0.46**	**0.40**	**0.35**	0.20	0.15	0.22	**0.37**	–			
9. Education	**−0.52**	**−0.61**	**−0.62**	0.03	**0.39**	−0.13	**−0.35**	**−0.60**	–		
10. Unemployment	**0.33**	0.27	0.24	−0.13	−0.06	**−0.37**	0.26	−0.01	−0.13	–	
11. Sq divorce	0.27	−0.27	−0.18	0.09	**0.43**	**−0.35**	−0.11	−0.11	**0.49**	0.00	–
Mean	1.13	2.72	2.05	16.34	67.93	95.04	0.53	3.54	0.77	8.63	1.33
SD	1.14	0.17	0.61	1.28	16.10	3.97	0.20	0.23	0.12	4.25	0.39

$p < 0.05$ for the values given in boldface.

generate a reduced model for comparison. The first model had a sample of 56 nations and included percent young, poverty, total population, percent urban, sex ratio, GDP and regional controls. Both the LRT and Wald Test indicated regional controls were necessary because they significantly increased the ability to account for variation in homicide rates ($\chi^2 = 13.47$, $p = 0.036$; $F = 3.34$, $p = 0.008$).

In Model 1, percent young and homicide victimization were not significantly associated ($b = 1.74$, $p = 0.146$). There were only small differences between the full (Model 1) and reduced model (Model 2) in the fit indicators for the first pair of models in Table 4. The first difference was that the R^2 decreased only slightly in Model 2 ($R^2 = 0.777$) compared with Model 1 ($R^2 = 0.788$), though the difference within the adjusted R^2s was negligible in Model 2 (adjusted $R^2 = 0.721$) compared with Model 1 (adjusted $R^2 = 0.728$). The MSRs did increase slightly in the reduced model. The LRT and Wald Test both failed to reject the null hypothesis that $\beta_{\ln\%15\text{–}24}$ is equal to 0; i.e. the tests indicated that in this model percent young did not significantly contribute in increasing the amount of explained variation in homicide victimization rates ($\chi^2 = 2.79$, $p = 0.094$; $F = 2.19$, $p = 0.146$). This does not mean we confirm or accept the null hypothesis but instead that we fail to reject the null hypothesis of no association. Table 5 provides results for the LRT and Wald Test for the comparison of the full and reduced models we describe in Table 4.

Since $b = 0.00$ ($p = 0.985$) for GDP, we estimated another model using the full sample. The additional model accomplished two goals. First, to see whether GDP was necessary since it increases the variance inflation factor (VIF) of both percent young and poverty.[5] Second, to see whether results of the LRT and Wald test for the full models changed when the complication introduced by GDP was removed. In the model excluding GDP (Model 3), percent young was not significantly associated with homicide ($b = 1.74$, $p = 0.127$). The comparison of Model 3 to Model 1 indicated GDP did not significantly contribute in accounting for variation in homicide in the models ($\chi^2 = 0$,

[5]All sensitivity analyses are available upon request in HTML format from *R*'s notebook function.

Table 4 *Stepwise models for percent young's effect on homicide victimization rates*

	Model 1			Model 2			Model 3			Model 4			Model 5			Model 6		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*	*b*	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Ln % 15–24	1.74	1.17	0.146	–	–	–	1.74	1.12	0.127	–	–	–	1.90	1.11	0.093	–	–	–
Ln poverty	0.75	0.39	0.062	1.00	0.35	0.007	0.75	0.33	0.029	1.11	0.25	<0.001	0.55	0.34	0.117	0.95	0.26	0.001
Ln pop	−0.04	0.08	0.639	−0.06	0.08	0.480	−0.04	0.08	0.635	−0.06	0.08	0.467	0.02	0.08	0.819	−0.01	0.08	0.927
% Urban	0.01	0.01	0.231	0.01	0.01	0.286	0.01	0.01	0.208	0.01	0.01	0.316	0.01	0.01	0.377	0.00	0.01	0.524
Sex ratio	−0.10	0.03	0.001	−0.09	0.03	0.003	−0.10	0.03	0.001	−0.08	0.03	0.003	−0.10	0.03	0.002	−0.08	0.03	0.006
Ln GDP	0.00	0.18	0.985	−0.07	0.17	0.688	–	–	–	–	–	–	–	–	–	–	–	–
Sq ethnic heterogeneity													1.07	0.59	0.076	1.00	0.60	0.104
Ln Gini																		
Education																		
Unemployment																		
Sq divorce																		
Africa	−0.23	0.71	0.749	−0.31	0.72	0.668	−0.23	0.65	0.721	−0.43	0.65	0.512	−0.24	0.64	0.710	−0.45	0.65	0.489
South America	1.02	0.40	0.016	1.16	0.40	0.006	1.01	0.34	0.005	1.08	0.34	0.003	0.94	0.35	0.011	1.02	0.36	0.006
Asia	0.18	0.47	0.712	0.33	0.47	0.486	0.17	0.44	0.695	0.27	0.44	0.544	0.24	0.43	0.588	0.34	0.44	0.444
Pacific	−0.14	0.56	0.803	−0.15	0.56	0.790	−0.14	0.52	0.785	−0.23	0.52	0.664	−0.07	0.51	0.885	−0.17	0.52	0.741
Western Europe	−0.18	0.39	0.646	−0.25	0.39	0.526	−0.18	0.33	0.587	−0.33	0.32	0.308	−0.11	0.33	0.737	−0.28	0.32	0.391
North America	0.67	0.56	0.238	0.78	0.56	0.171	0.67	0.46	0.152	0.66	0.46	0.163	0.33	0.49	0.505	0.34	0.50	0.493
Constant	4.40	4.19	0.299	7.98	3.47	0.026	4.36	3.38	0.204	7.22	2.88	0.016	2.61	3.53	0.464	5.89	3.03	0.059
R^2	0.788			0.777			0.788			0.776			0.794			0.780		
Adj R^2	0.728			0.721			0.735			0.726			0.735			0.723		
F	13.29	<0.001		13.92	<0.001		14.83	<0.001		15.58	<0.001		13.51	<0.001		13.84	<0.001	
MSR	0.351			0.361			0.343			0.354			0.333			0.348		
N	56			56			56			56			55			55		

Note: Comparison group for regional controls is Eastern Europe.

Table 4 continued

	Model 7			Model 8			Model 9			Model 10			Model 11			Model 12		
	b	SE	*p*	*b*	SE	*p*	*b*	SE	*p*	*b*	SE	*p*	*b*	SE	*p*	*b*	SE	*p*
Ln % 15–24	1.82	1.13	0.114	–	–	–	1.71	1.24	0.175	–	–	–	2.68	1.33	0.054	–	–	–
Ln poverty	0.54	0.35	0.126	0.92	0.27	0.001	0.66	0.37	0.082	0.99	0.28	0.001	0.74	0.40	0.072	1.32	0.29	<0.001
Ln pop	0.02	0.08	0.832	−0.01	0.08	0.927	0.04	0.09	0.651	0.01	0.09	0.873	−0.05	0.09	0.613	−0.10	0.09	0.287
% Urban	0.01	0.01	0.353	0.01	0.01	0.469	0.00	0.01	0.817	0.00	0.01	0.976	0.01	0.01	0.175	0.01	0.01	0.343
Sex ratio	−0.09	0.03	0.004	−0.07	0.03	0.013	−0.07	0.03	0.030	−0.06	0.03	0.072	−0.13	0.04	0.001	−0.09	0.03	0.007
Ln GDP	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Sq ethnic heterogeneity	0.99	0.62	0.117	0.89	0.63	0.164	1.14	0.63	0.079	1.06	0.63	0.104	0.55	0.65	0.403	0.38	0.67	0.577
Ln Gini	0.30	0.55	0.583	0.42	0.55	0.453	0.61	0.77	0.435	0.95	0.74	0.203	−0.25	0.72	0.736	0.28	0.70	0.693
Education index							0.88	1.37	0.525	1.15	1.37	0.406	0.07	1.25	0.955	0.45	1.29	0.731
Unemployment													−0.01	0.03	0.689	0.01	0.03	0.751
Sq divorce																		
Africa	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
South America	0.78	0.46	0.099	0.79	0.47	0.099	0.60	0.52	0.252	0.58	0.53	0.276	0.96	0.50	0.061	0.81	0.51	0.123
Asia	0.13	0.48	0.797	0.18	0.49	0.718	0.11	0.50	0.829	0.13	0.51	0.795	0.44	0.50	0.385	0.55	0.52	0.299
Pacific	−0.18	0.55	0.748	−0.31	0.55	0.578	−0.35	0.60	0.562	−0.54	0.59	0.364	0.02	0.56	0.966	−0.22	0.58	0.704
Western Europe	−0.18	0.36	0.615	−0.37	0.34	0.294	−0.22	0.37	0.555	−0.40	0.35	0.263	0.09	0.37	0.810	−0.13	0.37	0.730
North America	0.23	0.52	0.667	0.20	0.53	0.704	−0.03	0.57	0.959	−0.10	0.57	0.856	0.64	0.60	0.295	0.62	0.63	0.330
Constant	1.41	4.16	0.736	4.04	3.90	0.306	−2.05	5.09	0.689	−0.53	5.03	0.917	5.26	4.92	0.294	6.56	5.11	0.208
R^2	0.796			0.783			0.815			0.805			0.863			0.846		
Adj R^2	0.736			0.726			0.749			0.742			0.802			0.783		
F	13.3	<0.001		13.75	<0.001		12.22	<0.001		12.76	<0.001		13.99	<0.001		13.48	<0.001	
MSR	0.338			0.351			0.340			0.348			0.266			0.291		
N	54			54			50			50			46			46		

Note: Comparison group for regional controls is Eastern Europe.

Table 4 continued

	Model 13			Model 14		
	b	SE	*p*	*b*	SE	*p*
Ln % 15–24	1.32	1.42	0.364	–	–	–
Ln poverty	1.05	0.41	0.018	1.28	0.33	0.001
Ln pop	−0.14	0.10	0.156	−0.16	0.09	0.090
% Urban	0.00	0.01	0.796	0.00	0.01	0.919
Sex ratio	−0.07	0.05	0.214	−0.04	0.04	0.369
Ln GDP	–	–	–	–	–	–
Sq ethnic heterogeneity	0.05	0.68	0.944	−0.03	0.67	0.959
Ln Gini	0.99	1.16	0.402	1.60	0.96	0.108
Education index	0.29	1.75	0.868	0.87	1.63	0.600
Unemployment	0.03	0.03	0.303	0.05	0.03	0.104
Sq divorce	0.82	0.44	0.075	0.92	0.42	0.041
Africa	–	–	–	–	–	–
South America	–	–	–	–	–	–
Asia	0.62	0.51	0.236	0.66	0.50	0.204
Pacific	−0.26	0.62	0.685	−0.47	0.57	0.421
Western Europe	0.14	0.37	0.710	0.02	0.35	0.963
North America	0.67	0.58	0.257	0.66	0.57	0.260
Constant	−1.29	7.29	0.861	−2.98	7.04	0.676
R^2	0.882			0.877		
Adj R^2	0.804			0.805		
F	11.22	<0.001		12.10	<0.001	
MSR	0.219			0.217		
N	36			36		

Note: Comparison group for regional controls is Eastern Europe.

$p = 0.983$; $F = 0$, $p = 0.985$). In addition, the R^2 remained the same in Model 1 compared with Model 3, despite the exclusion of GDP ($R^2 = 0.788$), the adjusted R^2 increased, and the MSRs decreased. The decrease in the MSR and increase in the adjusted R^2 indicate the model was overfitted. Thus, the inclusion of GDP falsely increased the variance in the estimation of the regression line. Given these results, we took the more parsimonious approach of excluding GDP from all future models.

We estimated an additional reduced model, Model 4, removing both GDP and percent young to allow for comparison of the effect of percent young in a model without GDP. When comparing Model 3 with Model 4, the LRT and Wald test both indicated a failure to reject the null hypothesis that the contribution of percent young in accounting for variation in homicide victimization rates is different from 0 ($\chi^2 = 3$, $p = 0.083$; $F = 2.42$, $p = 0.127$).

Models 5 and 6 introduced ethnic heterogeneity, reducing the sample size to 55 nations due to missing data from Puerto Rico. Percent young was not significantly associated with homicide ($b = 1.90$, $p = 0.093$) in the full model (Model 5). When Model 5 and reduced Model 6 were compared using the LRT and Wald Test they indicated no significant difference between the models ($\chi^2 = 3.73$, $p = 0.053$; $F = 2.95$, $p = 0.093$). Models 7 and 8 introduced the Gini coefficient, reducing the sample size to 54 nations (excluding Puerto Rico and Mauritius). In the full Model 7, percent young was not significantly associated with homicide ($b = 1.82$, $p = 0.114$). When comparing Model 7 to its reduced Model 8, the LRT and Wald Test both indicated the exclusion of percent young did not

Table 5 *Likelihood ratio and Wald test results for full versus reduced model comparisons for homicide victimization rates*

	Likelihood ratio test		Wald test	
	χ^2	p	F	p
Model 1 vs. Model 2	2.79	0.094	2.19	0.146
Model 1 vs. Model 3	0.00	0.983	0.00	0.985
Model 3 vs. Model 4	3.00	0.083	2.42	0.127
Model 5 vs. Model 6	3.73	0.053	2.95	0.093
Model 7 vs. Model 8	3.33	0.068	2.61	0.114
Model 9 vs. Model 10	2.59	0.108	1.91	0.175
Model 11 vs. Model 12	5.62	0.019	4.03	0.054
Model 13 vs. Model 14	1.45	0.229	0.86	0.364

significantly change the model ($\chi^2 = 3.33$, $p = 0.068$; $F = 2.61$, $p = 0.114$). Models 9 and 10 introduced the education index and reduced the sample size to 50 (excluding Puerto Rico, Azerbaijan, Belarus, Georgia, Mauritius and Uzbekistan). Percent young was not significantly associated with homicide ($b = 1.71$, $p = 0.175$) in the full Model 9. The LRT and Wald Test comparing Models 9 and 10 indicated the exclusion of percent young did not significantly change the model ($\chi^2 = 2.59$, $p = 0.108$; $F = 1.91$, $p = 0.175$).

Models 11 and 12 introduced the unemployment rate, reducing the sample size to 46 nations. Missing nations in this model were Albania, Armenia, Azerbaijan, Brazil, Belarus, Bulgaria, Georgia, Mauritius, Puerto Rico and Uzbekistan. In the full Model 11, percent young was significantly and positively associated with homicide ($b = 2.68$, $p = 0.054$). The LRT and Wald test showed percent young significantly contributed to accounting for variation in homicide ($\chi^2 = 5.62$, $p = 0.019$; $F = 4.03$, $p = 0.054$) when comparing the reduced Model 12 to the full Model 11. However, given multiple comparisons and the borderline p value, the results of this single model showing an association between percent young and homicide rates should be interpreted with caution.

Models 14 and 15 introduced the crude divorce rate, resulting in a sample size of 36 (excluding Albania, Argentina, Armenia, Azerbaijan, Brazil, Belarus, Bulgaria, Chile, Dominican Republic, Ecuador, Guatemala, Georgia, Hong Kong SAR, Mauritius, Puerto Rico, Venezuela, United States, Ukraine and Uzbekistan). In Model 14, percent young was not significantly associated with homicide ($b = 1.32$, $p = 0.364$). The LRT and Wald test both indicated percent young did not significantly contribute to accounting for variation in homicide ($\chi^2 = 1.45$, $p = 0.229$; $F = 0.86$, $p = 0.364$) when comparing the reduced Model 15 to the full Model 14.

Results of OLS models suggest inclusion of percent young may overfit models of cross-national homicide rates. In all but one model we failed to reject the null hypothesis of no association between percent young and homicide victimization rates. In addition, percent young is not accounting for a significant proportion of the variance in homicide rates. The results suggest that by including percent young in models of cross-national homicide rates researchers are overfitting models and thereby increasing the variance and removing degrees of freedom, which are often precious in such models, given small sample sizes and the sometimes large number of independent variables.

Sequential ANOVA estimation

While VIFs are low, there are indications multicollinearity may be present in the models above. In many models, poverty is not significant when percent young is included, and percent young is significant only when sex ratio is included. There are a few ways to address this problem. The first is to fit the model sequentially that fits the variables into the model in a stepwise process and provides an *F* statistic that is a ratio of the variation accounted for to the residuals (i.e. the error). The *F* statistics in the sequential ANOVA provide the specific contributions of each variable, given all variables included before it in the model. The difference between the sequential *F* statistics and the *F* statistics provided by a standard statistical package is that the standard *F* statistic is an overall fit of the OLS model when all variables are estimated simultaneously, whereas the sequential *F* statistic is a sequential fitting of the values and is specific to each variable's ability to account for variation in the outcome variable. The alternative is putting all variables into a principal component analysis and losing the ability to differentiate effects of each variable. The first method is the best option here because we are specifically addressing whether or not the individual variable percent young is necessary given other variables.

Overall, the results of the sequential ANOVAs were consistent across model specification. Table 6 summarizes results, providing sequential ANOVA *F* statistics when percent young was included in the model before (in table: 'Percent young + poverty') and after poverty ('Poverty + percent young'). Percent young did not account for a significant proportion of the variation in homicide victimization when poverty was included before percent young, indicating poverty accounts for all of the variation in homicide victimization rates percent young accounts for without poverty included. Poverty is preferred for inclusion over percent young because no matter its location in the model it accounts for a significant proportion of variation in homicide rates.

Mallows's C_p

We used Mallows's C_p as a more precise method of model selection (Mallows 1973). We used an initial model that included percent young, poverty, total population, ethnic heterogeneity, sex ratio, Gini coefficient, education index and unemployment. Estimating Mallows's C_p for all models described above is redundant. Of the 255 models Mallows's C_p estimated, the best was a five-variable model including poverty, total population, ethnic heterogeneity, education and unemployment. Mallows's C_p for this model

Table 6 *Sequential ANOVA F statistics for percent young*

	Percent young + poverty		Poverty + percent young	
	F	*p*	*F*	*p*
Model 1	2.04	0.161	1.63	0.209
Model 3	10.03	0.003	0.06	0.805
Model 5	7.77	0.008	0.27	0.608
Model 7	6.38	0.016	0.23	0.634
Model 9	6.87	0.013	0.32	0.578
Model 11	9.54	0.004	0.00	0.976
Model 13	6.08	0.022	0.05	0.819

is 5.12, meaning it is fitted nearly perfectly. A perfect fit five-variable model would have a Mallows's C_p of 5.0 (Mallows 1973). Thus, the ratio of bias-to-variance is small in the model with a Mallows's C_p of 5.12, with only slightly more variance (Hotelling 1940; Mallows 1973). The R^2 and adjusted R^2 are high ($R^2 = 0.82$, adjusted $R^2 = 0.79$) for this model. Thus, according to Mallows's C_p, the best fitting model—i.e. the one that has almost an equal ratio of bias-to-variance (underfitting compared with overfitting)—is a five-variable model that does not include percent young.

Sensitivity analysis

We explored influential and outlier observations using dfbetas, dffits, the covariance ratio, Cook's distance and the diagonal of the hat matrix. We reestimated all models using all techniques (OLS, LRT, Wald and sequential ANOVA) excluding any influential or outlying observations. We then estimated an additional set of models excluding sex ratio due to a pattern emerging in which percent young was significant only when sex ratio was included in the model after outliers and influential observations were excluded. Without sex ratio included in the model, percent young was no longer significantly associated with homicide victimization across nations.

Discussion

Previous research on the structural covariates of cross-national homicide rates consistently included a measure of the percent of the total population that is young (Pampel and Gartner 1995; Pampel and Williamson 2001; Messner *et al.* 2010). The crux of the argument is that there is a positive association between percent young and homicide rates because young people are most at risk of becoming offenders and victims of violence (Hirschi and Gottfredson 1983; Gottfredson and Hirschi 1990).

Our review of the empirical literature did not support the contention that percent young has an effect on cross-national homicide rates. First, our review showed that of the 146 models in the 32 published studies that utilized a cross-national sample, homicide as an outcome variable, and included a measure of percent young, only 19 models (or 13 per cent) found the expected significant positive association between percent young and homicide rates. In total, 87 per cent of the models containing percent young revealed either no statistically significant association with (82 per cent) or a negative effect on (5 per cent) national homicide rates.

The results of this review, together with the strong belief about this association among scholars, led us to employ several techniques to systematically examine the impact of percent young on models of cross-national homicide rates. We collected our own data to test the dominant hypotheses that percent young (1) is significantly associated with cross-national homicide rates net of the other structural covariates and (2) significantly accounts for variation in models examining cross-national homicide rates. Neither hypothesis was supported. A series of OLS models revealed a consistent lack of support for the first hypothesis. Across all model specifications, we failed to reject the null hypothesis of no association between percent young and homicide victimization rates across nations. These findings spanned differences in sample composition, sample size and variables included in the statistical model.

Despite the failure to reject the null hypothesis of no association, it may be that percent young is a necessary control when testing other theories, which is why we tested the second hypothesis that percent young meaningfully contributes to the explained variance in models of cross-national homicide rates. Results from F ratios, Wald tests, likelihood ratio tests and other techniques all showed a lack of support for this hypothesis. All tests indicated including percent young may be overfitting the model and thus increasing the error in the estimation.

The globalization and demography literature provide reasons why percent young may be a proxy for poverty (Caldwell 1976; Nielsen 1994; Moller *et al.* 2003). Therefore, it may not be that percent young never matters, but it may not matter in the way criminologists proposed. Specifically, by including a measure of poverty and percent young, researchers may be measuring a similar underlying trait with two different variables. This would overfit the model as we have observed throughout all of our analyses. Additional evidence is in the overlap in the variation in homicide victimization across nations accounted for by percent young and the proxy for poverty, infant mortality. Future research should explore whether percent young operates as a proxy for poverty within cross-national criminological research.

The final step in our examination of the association between percent young and cross-national homicide rates was to use model selection techniques to assess the impact of the inclusion of percent young on overall models of cross-national homicide rates. The results suggest the efficacy of the models lessened when percent young was included. Another way to determine whether a variable is important is to employ Mallows's C_p, which is a more explicit model selection test, and we did so for 255 models that estimated every possible combination of the variables in our initial model. Mallows's C_p estimates the bias-to-variance ratio, or underfitting-to-overfitting of models. The best model according to these criteria was one that did not include percent young. The models that did include percent young often performed poorly on Mallows's C_p, and if they did not perform poorly they had a smaller R^2 or adjusted R^2.

In short, our systematic examination of the association between percent young and cross-national homicide rates had four main findings. First, many scholars continue to accept as a stylized fact the proportion of the entire population that is young has an effect on cross-national homicide rates even though 87 per cent of models in prior studies showed either a null or negative association. Second, our models consistently failed to reject the null hypothesis of no association between percent young and homicide rates. Third, sequential ANOVAs showed percent young did not account for appreciable explanation of variance when included in the models. Finally, model selection techniques revealed percent young's inclusion is actually counterproductive as it harms the overall performance of models of cross-national homicide rates.

Limitations

The age–crime curve speaks specifically to criminal offending. It is possible that those aged 15–24 are the most active offenders across nations but their victims are on average outside their age range. Our data measured homicide victimization rates. Prior research in multiple nations showed homicide victims tend to be slightly older than offenders (e.g. Wolfgang 1967). Unfortunately, when data are available for homicide offenders across nations their validity is questionable because the definition of homicide may vary by nation and data are derived only from offenders who have been arrested (Marshall and Block 2004; Smit *et al.* 2004, 2012). Thus, our findings may only apply to homicide victimization (and not offending) rates, and we are unable to draw strong conclusions about the efficacy of the age–crime curve because we are not examining criminal offending. Nevertheless, our findings are germane, given the choice of data source (homicide victimization rates from the WHO) of most recent studies of cross-national homicide rates.

Second, we do not claim that percent young never matters or that it never accounts for a significant amount of variation in homicide victimization across nations. We investigated whether percent young has a direct effect on homicide victimization as proposed by the compositional argument. As part of a factor or a principal component representing an underlying construct, percent young may affect homicide victimization. For example, scholars testing modernization theory often include percent young as one variable in a multivariable principal component or factor (DiCristina 2004; Pridemore and Trent 2010; Nivette 2011). Since we tested the direct effects of the individual variable it could be that while percent young does not matter individually it may play a role in the ability of a 'modernization' principal component or factor to explain the variation in cross-national homicide rates.

Third, failing to reject the null hypothesis of no association is not the same as confirming the null hypothesis. Thus, we cannot say with certainty there is no association between percent young and cross-national homicide victimization rates. Within the criminological literature more broadly, it is necessary to discuss at what point it is safe to say there is no association or it is no longer necessary to include a variable in a statistical model as a control after decades of studies and dozens of tests. Weisburd *et al.* (2003) offer a nice example of this in their discussion of when we can conclude that interventions do not work. However, other criminologists (including those examining cross-national homicide rates) must consider the conditions under which they are willing to conclude a variable is no longer necessary in a model or that a theory is not a viable explanation of the phenomenon under study. We do not claim to have met these criteria. Instead, it is the case that under the conditions of our sample the variables we utilized and the assumptions of the statistical techniques we employed (1) we consistently failed to reject the null hypothesis of no association between percent young and cross-national homicide victimization rates and (2) inclusion of percent young consistently failed to account for variation in cross-national homicide victimization and appeared to overfit the models. Thus, while there is reason to be sceptical of including percent young in a model meant to account for variation in cross-national homicide victimization rates, given our systematic approach we must remain cautious and not assume percent young does not matter.

Fourth, it is possible there is a difference in the effect of percent young on homicide victimization rates when within and across nation differences are estimated (Gartner 1990; Gartner and Parker 1990; Pampel and Gartner 1995; Pampel and Williamson 2001). While the OLS models provide some insight into the across-nation effect of percent young on homicide rates, this is only an average homicide rate and does not take into account any potential covariation over time. There may be differences over time as suggested by within-nation time series plots (not shown here). Mixed effects models provide a time-series component that allow for the estimation of this relationship over time, across nations, and within nations (Agresti and Finlay 2008). Future research should address this limitation.

Within the OLS models, sample size ranged from 56 to 36 nations. This sample size is small, especially given the number of variables in the model. However, the size is dictated by the number of nations for which relevant data are available (specifically homicide victimization rates from WHO) and thus such small sample sizes are not uncommon in the cross-national literature (Avison and Loring 1986; Krahn and Hartnagel 1986; Messner 1989; Gartner 1990; Bennett 1991; Savolainen 2000; Chamlin and Cochran 2006; Pridemore 2008; Messner *et al.* 2010).[6] A similar limitation is the sample of nations included in the study. While we include a broad range of regions in the sample, there are only a few nations from some of the regions. The vast majority of the sample consists of East and West European nations. As with sample size, this limitation is a result of data availability on both homicide victimization and key structural covariates.

Conclusion

We addressed one overarching research question in this study: Does the relative size of the young population influence cross-national homicide victimization rates? Based on the results of several techniques, we conclude the likely answer to this question is no. While one should appreciate the possibilities of an association between percent young and cross-national homicide rates provided by the theoretical literature, initial hypotheses are often incorrect. Our results (1) consistently failed to reject the null hypothesis of no association between percent young and cross-national homicide rates and (2) suggested the inclusion of percent young in models of cross-national homicide rates likely overfits the models and thus hinders the estimation process.

Our findings present three broad challenges. First, it has become a stylized fact to many that national homicide rates are closely related to the relative size of the young population. This perception is confusing, given the overwhelming evidence against this belief in the existing literature. Our study helps clarify the scientific record thus far. Second, there are multiple theoretical explanations for a possible association between percent young and national homicide rates, and we discussed the main one in this article. Scholars are unable to test these substantive explanations in a meaningful way with unstable models. Third, translating empirical findings into policy and practice is dependent upon the validity of the results. Unstable models mean little trust in the results, which means we should not depend on the results to create policy.

Acknowledgement

We thank Rick Rosenfeld for comments on and critiques of an earlier draft of this manuscript.

References

Agresti, A. and Finlay, B. (eds) (2008), *Statistical Methods for Social Sciences*. Pearson Education Limited.

[6]It is possible our results are biased by the East European nations in our sample and a lack of power due to a small sample size. Previous research shows that in Eastern Europe homicide offenders and victims are much older than in other countries (Pridemore 2003, 2005; Stamatel 2008). While we attempt to address this limitation in the OLS model estimation by including regional controls, these controls may not be sufficient.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S. and Wacziarg, R. (2004), 'Fractionalization', *Journal of Economic Growth*, 8: 155–94.

Altheimer, I. (2008), 'Social Support, Ethnic Heterogeneity, and Homicide: A Cross-National Approach', *Journal of Criminal Justice*, 36: 103–14.

Andresen, M. A. (2012), 'Homicide in Lithuania', in M. C. A. Liem and W. A. Pridemore, eds, *Handbook of European Homicide Research*, 437–49. Springer.

Avison, W. R. and Loring, P. L. (1986), 'Population Diversity and Cross-National Homicide: The Effects of Inequality and Heterogeneity', *Criminology*, 24: 733–49.

Bennett, R. (1991), 'Development and Crime: A Cross-National Time-Series Analysis of Competing Models', *Sociology Quarterly*, 32: 343–3.

Birkel, C. and Dern, H. (2012), 'Homicide in Germany', in M. C. A. Liem and W. A. Pridemore, eds, *Handbook of European Homicide Research*, 313–28. Springer.

Blumstein, A. and Nagin, D. S. (1974), 'Analysis of Arrest Rates for Trends in Criminality', *Socio-Economic Planning Sciences*, 9: 221–27.

Caldwell, J. C . (1976), 'Towards a Restatement of Demographic Transition Theory', *Population and Development Review*, 2: 321–66.

Chamlin, M. B. and Cochran, J. K. (2006), 'Economic Inequality, Legitimacy, and Cross-National Homicide Rates', *Homicide Studies*, 10: 231–52.

Chilton, R. and Spielberger, A. (1971), 'Is Delinquency Increasing? Age Structure and the Crime Rate', *Social Forces*, 49: 487–93.

Chon, D. S. (2012), 'The Impact of Population Heterogeneity and Income Inequality on Homicide Rates: A Cross-National Assessment', *International Journal of Offender Therapy and Comparative Criminology*, 56: 730–48.

Cohen, L. E. and Felson, M. (1979), 'Social Change and Crime Rate Trends: A Routine Activity Approach', *American Sociological Review*, 44: 588–609.

Cohen, L. E. and Land, K. C. (1987), 'Age Structure and Crime: Symmetry Versus Asymmetry and the Project of Crime Rates Through the 1990s', *American Sociological Review*, 52: 170–83.

Cole, J. H. and Gramajo, A. M. (2009). 'Homicide Rates in a Cross-Section of Countries : Evidence and Interpretations', *Population and Development Review*, 35: 749–76.

Conklin, G. and Simpson, M. (1985), 'A Demographic Approach to Cross-National Study of Homicide', *Comparative Social Research*, 8: 171–85.

DiCristina, B. (2004). 'Durkheim's Theory of Homicide and the Confusion of the Empirical Literature', *Theoretical Criminology*, 8: 57–91.

Fiala, R. and LaFree, G. (1988), 'Cross-National Determinants of Child Homicides', *American Sociological Review*, 53: 432–45.

Ganpat, S. M. and Liem, M. C. A. (2012), 'Homicide in the Netherlands', in M. C. A. Liem and W. A. Pridemore, eds, *Handbook of European Homicide Research*, 329–41. Springer.

Gartner, R. (1990), 'The Victims of Homicide: A Temporal and Cross-National Comparison', *American Sociological Review*, 55: 92–106.

Gartner, R. and Parker, R. N. (1990), 'Cross-National Evidence on Homicide and the Age Structure of the Population', *Social Forces*, 69: 351–71.

Gottfredson, M. R. and Hirschi, T. (1990), *A General Theory of Crime*. Stanford University Press.

Granath, S. (2012), 'Homicide in Sweden', in M. C. A. Liem and W. A. Pridemore, eds, *Handbook of European Homicide Research: Patterns, Explanations, and Country Studies*, 405–14. Springer.

Greenberg, D. F. (1977), 'Delinquency and the Age Structure of Society', *Contemporary Crisis*, 1: 189–223.

—— (1985), 'Age, Crime, and Social Explanation', *American Journal of Sociology*, 91: 1–21.

Hirschi, T. and Gottfredson, M. R. (1983), 'Age and the Explanation of Crime', *American Journal of Sociology*, 89: 552–84.

Hoaglin, D. C., Mosteller, F. and Tukey, J. W. (1983), *Understanding Robust and Exploratory Data Analysis*. John Wiley and Sons, Inc.

Hocking, R. R. (eds) (2003), *Methods and Applications of Linear Models. Regression and the Analysis of Variance*. Wiley-Interscience.

Hotelling, H. (1940), 'The Selection of Variants of Use in Prediction With Some Comments on the General Problems of Nuisance Parameters', *Annals of Mathematical Statistics*, 11: 271–83.

Krahn, H. and Hartnagel, T. (1986), 'Income Inequality and Homicide Rates: Cross-National Data and Criminological Theories', *Criminology*, 24: 269–95.

LaFree, G. (1999), 'A Summary and Review of Cross-National Comparative Studies of Homicide', in M. D. Smith and M. A. Zachn, eds, *Homicide: A Sourcebook of Social Research*, 125–45. Sage.

Lehti, M. and Kivivuori, J. (2012), 'Homicide in Finland', in M. C. A. Liem and W. A.. Pridemore, eds, *Handbook of European Homicide Research*, 391–404. Springer.

Liem, M., Ganpat, S., Granath, S., Hagstedt, J., Kivivuori, J., Lehti, M. and Nieuwbeerta, P. (2012), 'Homicide in Finland, the Netherlands, and Sweden: First Findings From the European Homicide Monitor', *Homicide Studies*, 17: 75–95.

Lysova, A. V., Shchitov, N. G. and Pridemore, W. A. (2012), 'Homicide in Russia, Ukraine, and Belarus', in M. C. A.. Liem and W. A.. Pridemore, eds, *Handbook of European Homicide Research*, 451–69. Springer.

Mallows, C. L. (1973). 'Some comments on C_p', *Technometrics*, 15: 661–75.

Marshall, I. H. and Block, C. R. (2004). 'Maximizing the Availability of Cross-National Data on Homicide', *Homicide Studies*, 8: 267–310.

Marvell, T. B., and Moody, C. E. (1991), 'Age Structure and Crime Rates: The Conflicting Evidence', *Journal of Quantitative Criminology*, 7: 237–73.

McCall, P. L. and Nieuwbeerta, P. (2007), 'Structural Covariates of Homicide Rates: A European City Cross-National Comparative Analysis', *Homicide Studies*, 11: 167–88.

Messner, S. F. (1989), 'Economic Discrimination and Societal Homicide Rates: Further Evidence on the Cost of Inequality', *American Sociological Review*, 54: 597–611.

Messner, S. F., Raffalovich, L. E. and Sutton, G. M. (2010), 'Poverty, Infant Mortality, and Homicide Rates in Cross-National Perspective: Assessments of Criterion and Construction Validity', *Criminology*, 48: 509–37.

Moller, S., Huber, E., Stephens, J. D., Bradley, D. and Nielsen, F. (2003), 'Determinants of Relative Poverty in Advanced Capitalist Democracies', *American Sociological Review*, 68: 22–51.

Myers, R. H. (eds) (1990), *Classical and Modern Regression With Applications*. Duxbury Press.

Neapolitan, J. L. (1994), 'Cross-National Variation in Homicides: The Case of Latin America', *International Criminal Justice Review*, 4: 4–22.

—— (1997), 'Homicides in Developing Nations: Results of Research Using a Large and Representative Sample', *International Journal of Offender Therapy and Comparative Criminology*, 41: 358–74.

—— (1998), 'Cross-National Variation in Homicides: Is Race a Factor?', *Criminology*, 36: 139–56.

Neumayer, E. (2003), 'Good Policy Can Lower Violent Crime: Evidence From a Cross-National Panel of Homicide', *Journal of Peace Research*, 40: 619–40.

Nielsen, F. (1994), 'Income Inequality and Industrial Development: Dualism Revisited', *American Sociological Review*, 59: 654–77.

Nivette, A. E. (2011), Cross-National Predictors of Crime: A Meta-Analysis', *Homicide Studies*, 15: 103–31.

Ortega, S. T., Corzine, J., Burnett, C. and Poyer, T. (1992), 'Modernization, Age Structure, and Regional Context: A Cross-National Study of Crime', *Sociological Spectrum*, 12: 257–77.

Pampel, F. C. and Gartner, R. (1995), 'Age Structure, Socio-Political Institutions, and National Homicide Rates', *European Sociological Review*, 11: 243–60.

Pampel, F. C. and Williamson, J. B. (2001), 'Age Patterns of Suicide and Homicide Mortality Rates in High-Income Nations', *Social Forces*, 80: 251–82.

Preti, A. and Maccio, A. (2012), 'Homicide in Italy', In M. C. A. Liem and W. A. Pridemore, eds, *Handbook of European Homicide Research*, 373–89. Springer.

President Commission on Law Enforcement and Administration of Justice. (1967), *The Challenge of Crime in Free Society*. United States Government Printing Office.

Pridemore, W. A. (2003). 'Demographic, Temporal, and Spatial Patterns of Homicide Rates in Russia', *European Sociological Review*, 19: 41–59.

—— (2005). 'Social Structure and Homicide in Post-Soviet Russia', *Social Science Research*, 34: 732–56.

—— (2008), 'A Methodological Addition to the Cross-National Empirical Literature on Social Structure and Homicide: A First Test of the Poverty-Homicide Thesis', *Criminology*, 46: 133–54.

—— (2011). 'Poverty Matters: A Reassessment of the Inequality-Homicide Relationship in Cross-National Studies', *British Journal of Criminology*, 51: 739–72.

Pridemore, W. A. and Trent, C. L. S. (2010), 'Do the Invariant Findings of Land, McCall, and Cohen Generalize to Cross-National Studies of Social Structure and Homicide?', *Homicide Studies*, 14: 296–335.

Rogers, M. L. (2014), 'A Descriptive and Graphical Analysis of the (Lack of) Association Between Age and Homicide Cross-Nationally', *International Criminal Justice Review*, 24: 235–53.

Rogers, M. L. and Pridemore, W. A. (2013), 'The Effect of Poverty and Social Protection on National Homicide Rates: Direct and Moderating Effects', *Social Science Research*, 42: 584–95.

—— (2016). 'The (Null) Effects of Percent Young on 15 to 24 Age-Specific and Male- and Female-Specific Cross-National Homicide Rates', *Homicide Studies,* in press.

Savolainen, J. (2000), 'Inequality, Welfare State, and Homicide: Further Support for the Institutional Anomie Theory', *Criminology*, 38: 1021–42.

Shavit, Y. and Rattner, A. (1988), Age, Crime, and the Early Life Course', *American Journal of Sociology*, 93: 1457–70.

Smit, P. R., De Jong, R. and Bijleveld, C. (2012), 'Homicide Data in Europe: Definitions, Sources and Statistics', in M. Liem and W. A. Pridemore, eds, *Handbook of European Homicide Research: Patterns, Explanations, and Country Studies*, 5–24. Springer.

Smit, P. R., Meijer, R. F. and Goroen, P. P. J. (2004), Detection Rates, an International Comparison', *European Journal on Criminal Police and Research*, 10: 225–53.

South, S. J. and Messner, S. F. (2000), Crime and Demography: Multiple Linkages, Reciprocal Relations', *Annual Review of Sociology*, 26: 83–106.

Stamatel, J. P. (2008), 'Using Mortality Data to Refine our Understanding of Homicide Patterns in Select Postcommunist Countries', *Homicide Studies*, 12: 117–35.

—— (2009), Correlates of National-Level Homicide Variation in Post-Communist East-Central Europe', *Social Forces*, 87: 1423–48.

Steffensmeier, D. J., Allan, E. A., Harer, M. D. and Streifel, C. (1989), Age and the Distribution of Crime', *American Journal of Sociology*, 94: 803–31.

Trent, C. L. S. and Pridemore, W. A. (2011), 'A Review of the Cross-National Empirical Literature on Social Structure and Homicide', in M. C. A. Liem and W. A. Pridemore, eds, *Handbook of European Homicide Research: Patterns, Explanations, and Country Studies*, 111–136. Springer.

Weisburd, D., Lum, C. M. and Yang, S.-M. (2003), 'When can we conclude that treatments or programs "Don't Work"?' *The American Academy of Political and Social Science*, 587: 31–48.

Wolfgang, M. E. (1967), *Studies in Homicide*. Harper & Row.

World Health Organization. (2012), World Health Organization Information System, available online at http://www.who.int/whosis/en/

Appendix

Table A *Nations in sample by region*

Africa	Asia	Eastern Europe	North America	South Pacific	South America	Western Europe
Mauritius	Hong Kong SAR	Albania	Canada	Australia	Argentina	Austria
	Israel	Armenia	Mexico	New Zealand	Brazil	Belgium
	Japan	Azerbaijan	United States		Chile	Finland
	Republic of Korea	Belarus			Costa Rica	France
	Singapore	Bulgaria			Dominican Republic	Germany
		Croatia			Ecuador	Greece
		Czech Republic			Guatemala	Ireland
		Estonia			Puerto Rico	Italy
		Georgia			Venezuela	Netherlands
		Hungary				Norway
		Kyrgyz Republic				Portugal
		Latvia				Spain
		Lithuania				Sweden
		Moldova				Switzerland
		Poland				United Kingdom
		Romania				
		Russia				
		Slovakia				
		Slovenia				
		Ukraine				
		Uzbekistan				