Supplementary Materials for

"Collateral Consequences of Violence in Disadvantaged Neighborhoods"

Social Forces, December 2009

David J. Harding

University of Michigan

March 2009

(available at http://www-personal.umich.edu/~dharding)

CONTENTS:

1. Scales for Individual Violence and Neighborhood Violence, Intergenerational Closure, and Social Cohesion

Construction of the individual violence, neighborhood violence, neighborhood intergenerational closure, and neighborhood social cohesion scales is based on methods described in Raudenbush and Sampson (1999). In the case of the neighborhood scales, these methods provide a way to aggregate survey data collected from individual respondents to the neighborhood level. Each scale combines data from multiple indicators of the concept. There are seven binary indicators for the individual violence measure, nine binary indicators for the neighborhood violence measure, three five-category ordinal measures for the intergenerational closure scale, and three binary indicators for the social cohesion scale (these indicators are described in the main text). The scales are constructed using all wave 1 Addhealth cases, not just those respondents used in this analysis (which is limited to those followed through wave 3).

For each scale, the multiple indicators can be thought of as hierarchically nested in a three level model: items nested within individuals nested within neighborhoods. I modeled these items using three-level hierarchical models. The dependent variable in each model is the value of the particular indicator. The level 1 model includes a constant and dummy variables for each item (minus one to allow for an omitted category). The level two model includes a constant with variance component and age and gender indicators. These age and gender variables remove age and gender variation from the scale. The omitted category for the age indicators is 15, and the gender indicator is one for male and zero for female. The level three model includes only a constant with a variance component. The estimated models used to construct the scales are displayed in Tables S1-S4.

After estimating this model, the predicted value of the constant for either the individual (for the individual violence scale) or the neighborhood (for the neighborhood scales) is the measure of the scale in the logit metric (known as empirical Bayes estimates). These values are the sum of the constant and either the individual-specific or neighborhood-specific random effect. The variables are then standardized for easier interpretation. Coefficients on the item indicators can be interpreted as item "severity" relative to the omitted category. The more negative a coefficient, the rarer the indicator. The age and gender indicators capture differences by age and gender in the indicators, and they allow the resulting scales to be independent of differences across neighborhoods in the age and gender of sampled individuals. An additional advantage of this framework is that individuals with missing data on some items do not need to be excluded from the model as long as they have data on at least one item.

Reliability of the neighborhood scales based on binary indicators can be calculated for neighborhood *k* as (Raudenbush and Sampson 1999, Equation 10):

$$\frac{\omega_{pp}}{\omega_{pp} + \frac{\tau_{pp}}{J_k} + \frac{1}{n_k J_k w_k}}$$

Where ω_{pp} is the between neighborhood variance from the model, τ_{pp} is the within neighborhood variance from the model, J_k is the number of individuals in neighborhood *k*, n_k is the mean number of items per individual in neighborhood *k*, and w_k is the average (over individuals in neighborhood *k*) of the product of the proportion of positive items (those coded one) times proportion negative items (those coded zero). Thus the primary determinants of neighborhood level reliability are the proportion of variance between neighborhoods and the number of individuals per neighborhood (number of items per individual varies little across neighborhoods). Reliability will also be higher when w_k is near its maximum, which occurs when the proportion

of items that are positive is one half. When more of the variance is between neighborhoods, reliability is higher for all neighborhoods, and neighborhoods with more individuals have higher reliability. Reliability of the neighborhood violence scale has a mean of 0.48 and a standard deviation of 0.28. Reliability of the neighborhood social cohesion scale has a mean of 0.29 and a standard deviation of 0.23. These reliabilities are uncorrelated with neighborhood disadvantage and are largely a function of the small numbers of respondents in some neighborhoods. Though these reliabilities are somewhat lower than those of neighborhood scales in prior research (e.g. Sampson et al. 1999), we can assess the impact of these low reliabilities by estimating alternative models. Models that weight by the reliability of the neighborhood violence scale, essentially privileging observations for which measurement reliability is higher, produce estimates that are substantively similar to those presented in the main text.

Reliability of the individual violence scale for individual *j* is given by (Raudenbush and Sampson 1999, Equation 11):

$$\frac{\omega_{pp} + \tau_{pp}}{\omega_{pp} + \tau_{pp} + \frac{1}{n_{jk} w_{jk}}}$$

Where n_{jk} is the number of items for individual *j* in neighborhood *k* and w_{jk} is the product of the proportion of positive items times proportion negative items for individual *j* in neighborhood *k*. Thus the reliability of an individual's measurement on the violence scale is determined by the number of items for which the individual provides data and the proportion of positive items. The larger each of these is, the higher the reliability. Reliability of the individual violence scale has a mean of 0.58 and a standard deviation of 0.14.

Table S1: Multi-level Logit Model Used in Construction of Individual Violence Scale

Term	*Coefficient (Standard Error)*
Constant	-2.986 (0.047)
Item 1 (in physical fight)	omitted
Item 2 (pulled knife/gun)	-3.001 (0.041)
Item 3 (shot/stabbed someone)	-4.079 (0.055)
Item 4 (serious fight)	-0.032 (0.022)
Item 5 (caused injury requiring treatment)	-1.058 (0.026)
Item 6 (use/threaten w/weapon)	-3.174 (0.043)
Item 7 (in group fight)	-0.957 (0.028)
Age 11	-0.085 (0.397)
Age 12	-0.140 (0.107)
Age 13	-0.022 (0.056)
Age 14	0.063 (0.052)
Age 15	omitted
Age 16	-0.125 (0.047)
Age 17	-0.297 (0.052)
Age 18	-0.483 (0.054)
Age 19	-0.362 (0.108)
Age 20	-0.366 (0.280)
Age 21	-0.680 (1.028)
Male	1.051 (0.033)
Variance Components:	
Neighborhood	0.199
Individual	2.580
N items	142,555
N individuals	20,399
N neighborhoods	2,431

Table S2: Multi-level Logit Model Used in Construction of Neighborhood Violence Scale

Term	*Coefficient (Standard Error)*
Constant	-2.645 (0.034)
Item-Level Variables:	
Item 1 (saw shooting/stabbing)	omitted
Item 2 (had weapon pulled)	0.042 (0.031)
Item 3 (shot)	-2.501 (0.065)
Item 4 (stabbed)	-1.098 (0.039)
Item 5 (was jumped)	-0.102 (0.031)
Item 6 (injured in fight)	-0.429 (0.033)
Item 7 (n'hood not safe)	-0.116 (0.031)
Item 8 (>= 50-50 chance getting killed)	0.262 (0.030)
Item 9 (drug problem in n'hood)	-0.347 (0.035)
Individual-Level Variables:	
Age 11	-0.415 (0.532)
Age 12	-0.430 (0.081)
Age 13	-0.356 (0.045)
Age 14	-0.129 (0.040)
Age 15	omitted
Age 16	0.009 (0.035)
Age 17	-0.011 (0.036)
Age 18	-0.028 (0.038)
Age 19	0.070 (0.076)
Age 20	0.241 (0.176)
Age 21	0.100 (0.390)
Male	0.502 (0.022)
Variance Components:	
Neighborhood	0.302
Individual	0.775
N items	180,158
N individuals	20,531
N neighborhoods	2,449

Table S3: Multi-level Ordinal Logit Model Used in Construction of Neighborhood Intergenerational Closure Scale

Term	*Coefficient (Standard Error)*
Threshold 1	-2.773 (0.038)
Threshold 2	1.064 (0.015)
Threshold 3	2.243 (0.020)
Threshold 4	2.667 (0.025)
Item 1 (tell neighbor)	omitted
Item 2 (neighbor tell you)	0.731 (0.019)
Item 3 (friends parents spoken to)	2.410 (0.048)
Age 11	0.335 (0.484)
Age 12	-0.054 (0.067)
Age 13	-0.172 (0.040)
Age 14	-0.071 (0.037)
Age 15	omitted
Age 16	0.121 (0.037)
Age 17	0.202 (0.034)
Age 18	0.218 (0.037)
Age 19	0.433 (0.090)
Age 20	0.465 (0.200)
Age 21	0.277 (0.297)
Male	-0.056 (0.021)
Variance Components:	
Neighborhood	0.122
Individual	0.533
N items	52,108
N individuals	17,752
N neighborhoods	2,261

Table S4: Multi-level Logit Model Used in Construction of Neighborhood Social Cohesion Scale

Term	*Coefficient (Standard Error)*
Constant	0.964 (0.038)
Item 1 (know most neighbors)	omitted
Item 2 (stopped to talk to someone on street)	0.484 (0.035)
Item 3 (neighbors look out for each other)	0.049 (0.038)
Age 11	-0.085 (0.504)
Age 12	0.097 (0.091)
Age 13	0.127 (0.052)
Age 14	0.105 (0.042)
Age 15	omitted
Age 16	-0.111 (0.042)
Age 17	-0.274 (0.041)
Age 18	-0.308 (0.044)
Age 19	-0.368 (0.097)
Age 20	-0.613 (0.226)
Age 21	-0.109 (0.433)
Male	0.236 (0.026)
Variance Components:	
Neighborhood	0.313
Individual	1.085
N items	60,947
N individuals	20,432
N neighborhoods	2,432

2. Neighborhood Structural Disadvantage Scale

Table S5 shows the inter-item correlations of the variables that make up the scale. The item-rest correlation is the correlation between each item and the scale constructed without that item. Note that correlations are high for all three of the measures of presence of middle class families (college graduates, managerial/professional occupations, and family income above $75K), and reasonably high for percent black and percent youth, suggesting that these variables are strongly correlated with the scale even when they are not included in it. The average inter-item correlation is the average correlation of the items in a scale constructed without that one item. Note that there is not much variation in this column across items, suggesting that no one item is pulling down the inter-item correlation in the overall scale. Finally, Cronbach's alpha is a measure of the reliability of the scale. The alpha column shows this value for the scale constructed without each individual item (and at the bottom for the scale produced with all the items). Note that the reliability of the scale is little affected by the removal of any one item and that it is lowered slightly if any one of the items measuring the presence of middle class families is removed. This suggests that these two sets of measures capture the same underlying neighborhood SES concept but simply focus on the presence of families at opposite ends of the SES distribution as indicators of a neighborhood's position in that distribution. Removing either percent black or percent youth does not significantly change the reliability of the scale. In addition, an alternative scale constructed without these variables correlates 0.98 with the original scale, suggesting that inclusion of these variables does not impact the analysis.

Table S5: Inter-item Correlations for Variables Included in the Neighborhood Disadvantage Scale (for census tracts that contain Addhealth in-home respondents at wave 1)

	Obs	*Sign*	*Item-rest correlation*	*Average inter-item correlation*	*Alpha*
Family Poverty Rate	2438	+	0.7896	0.4915	0.8712
Female Headed Household Rate	2438	+	0.7278	0.5051	0.8772
Percent College Grads	2447	-	0.7406	0.5016	0.8757
Male Unemployment Rate	2441	+	0.7059	0.5099	0.8793
Percent Managerial/Professional Occs	2442	-	0.7349	0.5033	0.8765
Percent Familes with income > $75K	2438	-	0.6436	0.5241	0.8852
Percent Black	2447	+	0.5581	0.5434	0.8928
Percent Youth	2449	+	0.4969	0.5567	0.8979
Scale				0.5169	0.8954

3. Sample Descriptive Statistics

Table S6: Descriptive Statistics for High School Graduation Models

	Mean	*SD*	*Min*	*Max*	*Percent Imputed*
Individual Variables (*n* = 14,668)					
Graduated High School	.83	--	0	1	--
Female	.53	--	0	1	--
Individual Violence Scale	0	1	-1.89	4.79	.55%
Age	15.64	1.73	11	21	--
Hispanic	.16	--	0	1	--
Black	.22	--	0	1	--
Native American	.04	--	0	1	--
Asian	.08	--	0	1	--
Other Race	.09	--	0	1	--
Multi Racial	.05	--	0	1	--
Home Language not English	.10	--	0	1	.03%
Immigrant	.08	--	0	1	.03%
Household Size	4.61	1.65	1	21	--
Single Parent Household	.23	--	0	1	--
Other Household Type	.21	--	0	1	--
Parent Immigrant	.19	--	0	1	.72%
Parent Education: High School Grad	.30	--	0	1	1.17%
Parent Education: Some College	.28	--	0	1	1.17%
Parent Education: College Grad	.25	--	0	1	1.17%
Parent Professional Occupation	.34	--	0	1	2.02%
Parent Disabled	.05	--	0	1	1.90%
Parent Welfare Receipt	.09	--	0	1	2.61%
Log Family Income	3.56	.84	0	6.91	24.26%
Low Birth Weight	.10	--	0	1	17.78%
Mother's Age at Birth	25.91	5.35	12	53	26.50%
Neighborhood Variables (*n* = 2,013)					
Neighborhood Violence Scale	0	1	-4.16	4.76	--
Neighborhood Disadvantage Scale	0	1	-5.19	5.10	--
Neighborhood Intergenerational Closure Scale	0	1	-4.25	4.41	6.01%
Neighborhood Social Cohesion Scale	0	1	-3.60	5.00	.30%
Community/School Variables (*n* = 89)					
Urban	.31	--	0	1	--
Rural	.16	--	0	1	--
Small School Size (< 400)	.16	--	0	1	--
Large School Size (> 1000)	.47	--	0	1	--
Cumulative Dropout Rate	11.29	12.13	0	68.52	--
% College Prep Program	57.63	27.35	0	100	5.62%
Catholic School	.04	--	0	1	5.62%
Private School	.04	--	0	1	5.62%

Note: See variable descriptions in Appendix A

Table S7: Descriptive Statistics for Teenage Pregnancy Models

	Mean	*SD*	*Min*	*Max*	*Percent Imputed*
Individual Variables (*n* = 13,975)					
Teenage Pregnancy	.13	--	0	1	--
Female	.52	--	0	1	--
Individual Violence Scale	0	1	-1.9	4.79	.57%
Age	15.57	--	11	21	--
Hispanic	.16	--	0	1	--
Black	.22	--	0	1	--
Native American	.04	--	0	1	--
Asian	.08	--	0	1	--
Other Race	.09	--	0	1	--
Multi Racial	.05	--	0	1	--
Home Language not English	.11	--	0	1	.03%
Immigrant	.08	--	0	1	.03%
Household Size	4.61	1.63	1	21	--
Single Parent Household	.23	--	0	1	--
Other Household Type	.20	--	0	1	--
Parent Immigrant	.19	--	0	1	.52%
Parent Education: High School Grad	.30	--	0	1	.91%
Parent Education: Some College	.28	--	0	1	.91%
Parent Education: College Grad	.25	--	0	1	.91%
Parent Professional Occupation	.35	--	0	1	1.52%
Parent Disabled	.05	--	0	1	1.40%
Parent Welfare Receipt	.09	--	0	1	2.13%
Log Family Income	3.56	.84	0	6.91	23.89%
Low Birth Weight	.10	--	0	1	17.29%
Mother's Age at Birth	25.99	5.33	12	53	26.07%
Neighborhood Variables (*n* = 1,955)					
Neighborhood Violence Scale	0	1	-4.16	4.76	--
Neighborhood Disadvantage Scale	0	1	-5.19	5.10	--
Neighborhood Intergenerational Closure Scale	0	1	-4.45	4.26	5.36%
Neighborhood Social Cohesion Scale	0	1	-3.59	4.97	.31%
Community/School Variables (*n* = 89)					
Urban	.31	--	0	1	--
Rural	.16	--	0	1	--
Small School Size (< 400)	.16	--	0	1	--
Large School Size (> 1000)	.47	--	0	1	--
Cumulative Dropout Rate	11.29	12.13	0	68.52	--
% College Prep Program	57.63	27.35	0	100	5.62%
Catholic School	.04	--	0	1	5.62%
Private School	.04	--	0	1	5.62%

Note: See variable descriptions in Appendix A

4. Additional Sensitivity Analysis Tables

Table S8: Decomposition of Effects of Neighborhood Disadvantage on High School Graduation and Teenage Pregnancy (based on linear probability models)

	Males		***Females***	
High School Graduation	Effect	%	Effect	%
N'hood Dis.→ HS Grad	-.0049	39%	-.0007	9%
N'hood Dis.→ N'Hood Viol. → HS Grad	-.0078	61%	-.0073	91%
Sum	-.0127	100%	-.0080	100%
	Males		***Females***	
Teenage Pregnancy	Effect	%	Effect	%
N'hood Dis.→ Pregnancy	.0109	79%	.0174	84%
N'hood Dis.→ N'Hood Viol. → Pregnancy	.0029	21%	.0033	16%
Sum	.0138	100%	.0207	100%

Table S9: Sensitivity Analysis for Role of Neighborhood Violence in Explaining the Effect of Neighborhood Disadvantage on High School Graduation (females)

	β_{YU}					
r_{VU}	.00	.01	.02	.03	.04	.05
.00	-.0199* *91%*	-.0199* *91%*	-.0199* *91%*	-.0199* *91%*	-.0199* *91%*	-.0199* *91%*
-.05	-.0199* *91%*	-.0194* *89%*	-.0189* *87%*	-.0184* *84%*	-.0179* *82%*	-.0174* *80%*
-.10	-.0199* *91%*	-.0189* *87%*	-.0179* *82%*	-.0169* *78%*	-.0159* *73%*	-.0149* *68%*
-.15	-.0199* *91%*	-.0184* *84%*	-.0169* *78%*	-.0154* *71%*	-.0139* *64%*	-.0124* *57%*
-.20	-.0199* *91%*	-.0179* *82%*	-.0159* *73%*	-.0139* *64%*	-.0119* *55%*	-.0099* *45%*
-.25	-.0199* *91%*	-.0174* *80%*	-.0149* *68%*	-.0124* *57%*	-.0099* *45%*	-.0074 *34%*

Notes:

β_{YU} is hypothetical coefficient on standardized unobserved variable in linear probability model predicting graduation from high school and controlling for all other control variables.

r_{VU} is hypothetical partial correlation between standardized unobserved variable and violence scale.

In each cell, the top number is the expected coefficient on the violence scale variable if the unobserved variable where included in the linear probability model predicting graduation from high school, and the bottom number is the resulting proportion of the neighborhood disadvantage effect accounted for by neighborhood violence.

Note: 2 X SE of β on violence coefficient in linear probability model predicting high school graduation = 0.0096

* coefficient on violence scale would remain statistically significant

Table S10: Sensitivity Analysis for Role of Neighborhood Violence in Explaining the Effect of Neighborhood Disadvantage on Teenage Pregnancy (males)

	β_{YU}					
r_{VU}	.00	.01	.02	.03	.04	.05
.00	.0079* *21%*	.0079* *21%*	.0079* *21%*	.0079* *21%*	.0079* *21%*	.0079* *21%*
.05	.0079* *21%*	.0074* *20%*	.0069* *18%*	.0064 *17%*	.0059 *16%*	.0054 *14%*
.10	.0079* *21%*	.0069* *18%*	.0059 *16%*	.0049 *13%*	.0039 *10%*	.0029 *8%*
.15	.0079* *21%*	.0064 *17%*	.0049 *13%*	.0034 *9%*	.0019 *5%*	.0004 *1%*
.20	.0079* *21%*	.0059 *16%*	.0039 *10%*	.0019 *5%*	--	--
.25	.0079* *21%*	.0054 *14%*	.0029 *8%*	.0004 *1%*	--	--

Notes:

β_{YU} is hypothetical coefficient on standardized unobserved variable in linear probability model predicting teenage pregnancy and controlling for all other control variables.

r_{VU} is hypothetical partial correlation between standardized unobserved variable and violence scale.

In each cell, the top number is the expected coefficient on the violence scale variable if the unobserved variable where included in the linear probability model predicting teenage pregnancy, and the bottom number is the resulting proportion of the neighborhood disadvantage effect accounted for by neighborhood violence.

Note: 2 X SE of β on violence coefficient in linear probability model predicting teenage pregnancy = 0.0068

* coefficient on violence scale would remain statistically significant

Table S11: Sensitivity Analysis for Role of Neighborhood Violence in Explaining the Effect of Neighborhood Disadvantage on Teenage Pregnancy (females)

	β_{YU}					
r_{VU}	.00	.01	.02	.03	.04	.05
.00	.0091 *16%*	.0091 *16%*	.0091 *16%*	.0091 *16%*	.0091 *16%*	.0091 *16%*
.05	.0091 *16%*	.0086 *15%*	.0081 *14%*	.0076 *13%*	.0071 *13%*	.0066 *12%*
.10	.0091 *16%*	.0081 *14%*	.0071 *13%*	.0061 *11%*	.0051 *9%*	.0041 *7%*
.15	.0091 *16%*	.0076 *13%*	.0061 *11%*	.0046 *8%*	.0031 *5%*	.0016 *3%*
.20	.0091 *16%*	.0071 *13%*	.0051 *9%*	.0031 *5%*	.0011 *2%*	--
.25	.0091 *16%*	.0066 *12%*	.0041 *7%*	.0016 *3%*	--	--

Notes:

β_{YU} is hypothetical coefficient on standardized unobserved variable in linear probability model predicting teenage pregnancy and controlling for all other control variables.

r_{VU} is hypothetical partial correlation between standardized unobserved variable and violence scale.

In each cell, the top number is the expected coefficient on the violence scale variable if the unobserved variable where included in the linear probability model predicting teenage pregnancy, and the bottom number is the resulting proportion of the neighborhood disadvantage effect accounted for by neighborhood violence.

Note: 2 X SE of β on violence coefficient in linear probability model predicting teenage pregnancy = 0.0093

* coefficient on violence scale would remain statistically significant