ACI Technical Report: Initial Measures Derived from Census

Dr. Robert H. Langworthy
Justice Center
University of Alaska Anchorage

June 9, 2006

ACI Technical Report: Initial Measures Derived from Census

The decennial census provides a wealth of information about communities that has been mined by social scientist for decades. The purpose of this technical report is to describe an initial set of measures taken from or derived from the 2000 U.S. Census in an effort to develop a statistical description of Anchorage communities. The initial set of measures isolated from census are inspired by two principal bodies of work: 1) the Project on Human Development in Chicago Neighborhoods (PHDCN), an exceptionally well endowed research effort that took neighborhood measurement very seriously; and, 2) Peter Blau's (Blau) work that specifies parameter of social structure, heterogeneity and inequality.

The focus of the paper is on documenting how the measures were formed from 2000 Summary File 3 census tables. However, measures without conceptual content are of little value. Accordingly, the paper will offer a brief introduction to the derivative works (PHDCN, Blau) and then follow with a fairly detailed presentation of each measure (what concept is addressed, how it is measured, how the measure is distributed across block group and census tracts, and isolation of the census tables providing essential counts).

PROJECT ON HUMAN DEVELOPMENT IN CHICAGO NEIGHBORHOODS MEASURES FROM CENSUS

The Project on Human Development in Chicago Neighborhoods was a decade long research effort funded jointly by the National Institute of Justice and the John D. and Catherine T. MacArthur Foundation. The project was focused on explicating developmental sequences and correlates in production of criminals (see Earls and Reiss, 1994 for an early description of the scope of the project).[1] The project employed multiple research designs (an innovative staggered age cohort design, and a hierarchical analysis of community as both cause and context for social success and pathos). These multiple designs required careful specification of measures and multiple data collection methods (panel surveys of individuals to isolate developmental issues, observational surveys of communities, and reliance on numerous official records—among them census). The PHDCN thus represents a store of measures that were carefully conceptualized and assessed (reliability and validity was established). This paper isolates the PHDCN measures drawn from census and reports on their reproduction for Anchorage block groups and census tracts.[2]

[1] The thesis driving this work was built out of the social disorganization tradition. This theoretical frame asserts a relation between characteristics of neighborhood social structure and social pathos. The essence of the thesis is that structural antecedents to social disorganization lead to the inability to establish normative order and/or means of informal social control which in turn provide a fertile context of social pathos.

[2] The PHDCN measures were taken at the 'neighborhood' level. PHDCN defined neighborhoods in terms of aggregates of census tracts. The 847 Chicago census tracts were collapsed into 343 neighborhood clusters through a process jointly considering the proceeds of cluster analysis, local knowledge, and physical barriers (e.g., freeways, waterways, railroad tracks) (see Sampson, Raudenbush and Earls 1997 for a description of their operationalization of neighborhood).

PHDCN produced a number of conceptualized and assessed measures from census data tables. The table below lists the measures and the papers that introduced them.

PHDCN Census Based Measures

Measure	Study Citation
Concentrated Disadvantage	Sampson, Raudenbush and Earls (1997)
Immigrant Concentration	Sampson, Raudenbush and Earls (1997)
Residential Stability	Sampson, Raudenbush and Earls (1997)
Population Density	Sampson and Raudenbush (1999)
Concentrated Affluence	Sampson, Morenoff and Earls (1999)
Ratio of Adults to Children	Sampson, Morenoff and Earls (1999)
Index of Concentration at the Extremes	Morenoff, Sampson and Raudenbush (2001)

The most significant of these papers is the first paper which appeared in *Science* in 1997. In this paper Sampson and his colleagues introduced *concentrated disadvantage, immigrant concentration,* and *residential stability*. These three constructs were identified as parsimonious elements of the structure of Chicago neighborhoods and were developed following factor analysis of the: percent of families below poverty line, percent of families receiving public assistance, percent of families female headed, percent unemployed, percent less than 18 years of age, percent African American, percent Latino, percent foreign born, percent in same house in 1985, and percent of houses owner occupied.

The first step toward reproducing these measures for Anchorage was an attempt to reproduce the factor structure that underlay the PHDCN indices. The table below presents a comparison of the Chicago factor structure reported in the *Science* essay with that representing Anchorage.

Comparison of Chicago and Anchorage factor scores related to specification of *concentrated disadvantage, immigrant concentration*, and *residential stability*

Panel A: Factor Loadings: Census Tracts

	Chicago Neighborhoods (N=344)			Anchorage Census Tracts (N=55)		
	Concentrated disadvantage	Immigrant. concentration	Residential stability	F1	F2	F3
Below poverty level	.93	--	--	**.883**	-.110	-.633
On public assistance	.94	--	--	**.853**	-.301	-.678
Female-headed families	.93	--	--	**.938**	-.221	-.485
Unemployed	.86	--	--	.562	-.666	-.511
Less than age 18	.94	--	--	-.134	**.947**	.025
Black	.60	--	--	.466	.100	**-.884**
Latino	--	.88	--	**.820**	-.103	-.749
Foreign-born	--	.70	--	**.869**	-.276	-.287
Same house last 5 years	--		.77	-.465	.211	**.929**
Owner-occupied house	--		.86	-.585	.414	**.855**
Eigenvalues	>5			5.83	1.34	1.10

Panel B: Factor Loadings: Block Groups

	Chicago Neighborhoods (N=344)			Anchorage Block Groups (N=214)	
	Concentrated disadvantage	Immigrant concentration	Residential stability.	F1	F2
Below poverty level	.93	--	--	.794	.226
On public assistance	.94	--	--	.800	.128
Female-headed families	.93	--	--	.796	.008
Unemployed	.86	--	--	.573	-.131
Less than age 18	.94	--	--	-.016	**.940**
Black	.60	--	--	.607	310
Latino	--	.88	--	.669	.187
Foreign-born	--	.70	--	.633	-.101
Same house last 5 years	--	--	.77	-.753	.186
Owner-occupied house	--	--	.86	-.833	.318
Eigenvalues	>5			4.712	1.246

Review of the table suggests that the factor structure isolated in Chicago does not reproduce in Anchorage though it comes closer for Anchorage census tracts than block groups.[3] When the Anchorage census tract factor scores are compared to the Chicago neighborhood scores there is no evidence of an isolated *immigrant concentration* factor,[4] less than 18 years of age loads alone, and the residential stability factor is marginally isolated. When the level of aggregation shifts to block groups a single factor is isolated, again without proportion of the population less than age 18.

Though the Chicago factor structure did not reproduce with the Anchorage census data each of the Chicago measures (*concentrated disadvantage, immigrant concentration,* and *residential stability*) are reproduced as described below. Because the Anchorage data suggests a single construct another measure, *multiform disadvantage*, is computed and described below. Each of these measures is described below.

Concentrated Disadvantage is computed as an indicator of relative neighborhood poverty (see Sampson, Raudenbush and Earls, 1997). The original measure developed by Sampson and his colleagues sought to tap into multiple indicators of economic disadvantage such that the resulting composite measure was an indicator of multiform disadvantage. The resulting measure included the proportion of families below the poverty line, the proportion of families receiving public assistance, the proportion of families that were female headed, the proportion of the population 16 years and older unemployed, the proportion of the population under 18 years, and the proportion of the population Black or African American. The *concentrated disadvantage* measure computed for Anchorage includes those same variables except for the proportion of the population less than 18 years which did not load with the other variables. The table below presents the principal components factor loadings for *concentrated disadvantage* including all six variables and demonstrates the failure of the proportion of the population less than 18 years to relate to the single factor.[5]

[3] This may be the result of differing levels of aggregation. The Chicago neighborhood clusters were clusters of a few census tracts producing neighborhoods about twice the size of census tracts and many times the size of block groups.

[4] It is possible that the composition of a measure of *immigrant concentration* differs from one region to another. That is, the dominant immigrant population may vary from one region to another. To test the possibility that other immigrant nationalities might co-vary with proportion foreign born to form an *immigrant concentration* proxy for Anchorage the proportion Latino was replace with proportion Asian, and proportion Pacific Islander—neither altered the resulting factor structure.

[5] The factor score measures of *concentrated disadvantage* in the data files were computed without the proportion of the population less than 18 years.

Table Concentrated Disadvantage Factor Loadings, Anchorage, Block Groups and Census Tracts

Proportion of:	Block groups	Census tracts
Families receiving public assistance	.848	.936
Families below poverty line	.854	.899
Families female headed	.820	.873
Population 16 years and older unemployed	.542	.735
Population Black/African American	.642	.679
Population under 18 years old	.122	-.222

The measures of *concentrated disadvantage* in the data base are factor scores saved from the principal components factor analysis of the proportions of: families receiving public assistance, families below poverty line, families female headed, population 16 years and older unemployed, and population Black/African American. This represents a substantial technical departure (from PHDCN) in the calculation of this measure but not a significant departure in interpretation of the conceptual content of the composite score. Indeed both the PHDCN and Anchorage measures are defined by the same empirical elements. The factor score that represent *concentrated disadvantage* is distributed as:

Table Concentrated Disadvantage (factor scores)

	Block group	Census tract
Mean	0.000	0.000
Standard error	.068	.135
Standard deviation	1.000	1.000
Minimum	-1.320	-1.314
Maximum	3.289	3.034

The data to compute variables used to construct *concentrated disadvantage* were drawn from the following 2000 census Summary File 3 tables:

- Proportion of families below poverty ((P90_2)*Family income in 1999 below poverty level*/(P90_1)*Families*)
- Proportion of families receiving public assistance ((P64_2*Households with public assistance income*/(P64_1)*Households*)
- Proportion of families female headed ((P15_15)*Female householder, no husband present*/(P15_1)*Families*
- Proportion of persons over 16 years unemployed(((P43_7)*Males over 16 unemployed* + (P43_14)*Females over 16 unemployed*)/(P43_1)*Population 16 years and older*
- Proportion of population less than 18 years((P8_3 thru P8_20 + P8_42 thru P8_59)*Males and Females >1 year thru 17 years*/(P8_1)*Total population* Factor loading—did not load on factor not included in measure.
- Proportion of population Black/African American((P6_3)*Black or African American alone*/(P6_1)*Total Population*

Immigrant Concentration is computed as an indicator of "…areas of the city undergoing immigration…" (Sampson, Raudenbush and Earls, 1997:920). It is suggested that "Because it {immigrant concentration} describes neighborhoods of ethnic and linguistic heterogeneity, there is reason to believe that immigrant concentration may impede the capacity to realize common values and to achieve informal social control…" (Sampson, Raudenbush and Earls, 1997:920).[6]

Sampson and his colleagues measured *immigrant concentration* as factor scores. As noted above, this factor did not materialize in the Anchorage data. Therefore, in Anchorage the proxy for *immigrant concentration* is computed as the simple sum of two proportions: proportion Latino and the proportion foreign born. Theoretically, this index could vary from as little as 0 if there are no Latinos or persons foreign born in an area to 2 if the entire population were Latino and foreign born. The index that measures *immigrant concentration* is distributed as follows:

Table Immigrant Concentration

	Block group	Census tract
Mean	.145	.142
Standard error	.006	.009
Standard deviation	.090	.065
Minimum	.000	.047
Maximum	.574	.291

The data to compute variables used to construct *immigration concentration* were drawn from the following 2000 census Summary File 3 tables:

- Proportion Latino((P7_10)*Hispanic or Latino*/(P7_1)*Total Population*
- Proportion foreign born((P21_13)*Foreign Born*/(P21_1)*Total Population*

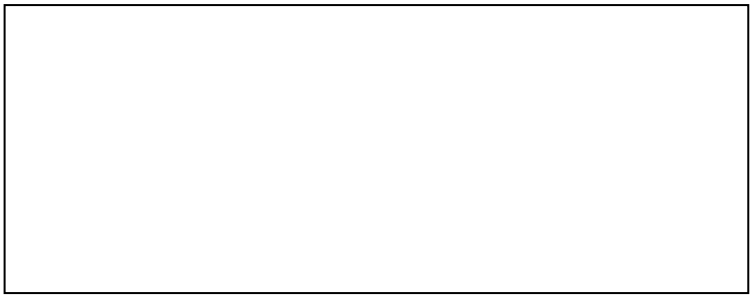
Residential Stability is computed as an indicator of the degree to which neighborhoods are stable. There is reason to believe that *residential stability* promotes evolution of common values and the capacity for informal control. The two variables that Sampson, Raudenbush and Earls isolated as constituting a measure of *residential stability* include: the proportion of the population five years and older living in the same house for five years, and the proportion of housing units that are owner occupied.

Sampson and his colleagues measured *residential stability* as factor scores. As noted above, this factor did not materialize in the Anchorage data as in Chicago.[7] Therefore, in Anchorage the proxy for *residential stability* is computed as the simple sum of two

[6] Another measure of ethnic heterogeneity in the database is *racial heterogeneity* (see description in discussion of measures of heterogeneity and inequality below). There is ample reason to believe that *racial heterogeneity* may impede the capacity to realize common values and achieve informal social control.

[7] A weak factor similar to the PHDCN measure was isolated in Anchorage at the census tract level of aggregation. It departed from the Chicago measure by the strong negative loading of proportion Black/African American. However, since the factor did not emerge in block group data the decision was made to use a summated scale instead of factor scores.

proportions: proportion owner-occupied houses and proportion in the same house last 5 years. Theoretically, this index could vary from as little as 0 if there are no owner-occupied houses and no one lived in the same house for the last 5 years to 2 if all houses are owner-occupied and all resided in the same house for the past 5 years. The index that measures *residential stability* is distributed as follows:



The data to compute variables use to compute *residential stability* were drawn from the following 2000 census Summary File 3 tables:

- Proportion same house last 5 years((P24_2)*Same house in 1995*/(P24_1)*Population 5 years and older*
- Proportion owner occupied house((H32_2)*Owner occupied housing units*/H30_1*Housing units*

Multiform Disadvantage. As noted above the factor structure reported in Sampson, Raudenbush and Earls (1997:920) did not reproduce in Anchorage. Indeed the variables that distinguished economic disadvantage (concentrated disadvantage), immigrant concentration, and residential stability from one another in Chicago form a single factor in Anchorage. We call this composite measure *multiform disadvantage* because it contains elements of economic disadvantage, and indicators associated with limited capacity to establish and enforce common values.

The table below presents the principal components factor loadings of the variables that together constitute an indicator of *multiform disadvantage*. Two factors were isolated at the census tract level but the significant loadings of the second factor were stronger on the first and the second factor eigenvalue is very low. These considerations taken together suggest that it is reasonable to abandon the second factor at the census tract level and retain factor one as an indicator of *multiform disadvantage*.

Table Multiform Disadvantage Factor Loadings, Anchorage, Block Groups and Census Tracts

Proportion of:	Block Groups	Census Tracts F1	F2
Families below poverty	.792	.872	.193
Families receiving public assistance	.796	.893	.123
Families female headed	.791	.842	.404
Over 16 years unemployed	.535	.665	.035
Population Black/African American	.608	.720	.526
Population Latino	.667	.891	.009
Population foreign born	.633	.704	.544
Same residence past 5 years	-.755	-.772	.526
Housing owner occupied	-.831	-.825	.334
Eigenvalues	4.712	5.794	1.137

The factor scores that represent *multiform disadvantage* are distributed as:

Table Multiform Disadvantage (factor scores)

	Block group	Census tract
Mean	0.000	0.000
Standard error	.068	.135
Standard deviation	1.000	1.000
Minimum	-1.567	-1.541
Maximum	2.729	2.529

Proportion of Families below Poverty is a variable in the construction of the composite measure *concentrated disadvantage*. As the table below notes, the *proportion of families below the poverty level* varies across Anchorage. On average, 5-6 percent of families in block groups or census tracts are below the poverty level, with at least one census tract and block group having no families below poverty, and at least one block group having nearly a third of families below poverty.

Table Proportion of Families below the Poverty Level

	Block group	Census tract
Mean	.06	.05
Standard error	.005	.005
Standard deviation	.07	.04
Minimum	0.00	0.00
Maximum	.31	.19

The data to compute *proportion of families below poverty level* were drawn from the following 2000 census Summary File 3 tables:

- ((P90_2)*Family income in 1999 below poverty level*/(P90_1)*Families*)

Proportion of Households Receiving Public Assistance is a variable in the construction of the composite measure *concentrated disadvantage*. On average, 7-8 percent of households in block groups or census tracts received public assistance, with at least one census tract and block group having no households (or nearly none) receiving public assistance, and at least one block group having slightly more than a third of households receiving assistance.

Table Proportion of Households Receiving Public Assistance

	Block group	Census tract
Mean	.08	.07
Standard error	.005	.007
Standard deviation	.07	.05
Minimum	0.00	.004
Maximum	.37	.23

The data to compute *proportion households receiving public assistance* were drawn from the following 2000 census Summary File 3 tables:

- ((P64_2*Households with public assistance income*/(P64_1)*Households*)

Proportion of Families Female Headed is a variable in the construction of the composite measure *concentrated disadvantage*. On average, nearly 20 percent of families in block groups or census tracts are female headed, with at least one census tract and block group having no families (or nearly none) female headed, and at least one block group having half of families female headed.

Table Proportion of Families Female Headed

	Block group	Census tract
Mean	.19	.17
Standard error	.008	.012
Standard deviation	.12	.09
Minimum	0.00	.03
Maximum	.50	.39

The data to compute *proportion of families female headed* were drawn from the following 2000 census Summary File 3 tables:

- ((P15_15)*Female householder, no husband present*/(P15_1)*Families*

Proportion of Population Over 16 Years Unemployed is a variable in the construction of the composite measure *concentrated disadvantage*. On average, 5 percent of the adult labor force population in block groups or census tracts were unemployed, with at least

one census tract and block group where there was virtually no unemployment, and at least one block group where nearly 20 percent of adults were unemployed.[8]

Table Proportion of Population Over 16 Years Unemployed

	Block group	Census tract
Mean	.05	.05
Standard error	.002	.004
Standard deviation	.03	.03
Minimum	0.00	.01
Maximum	.17	.16

The data to compute *proportion of population 16 and older unemployed* were drawn from the following 2000 census Summary File 3 tables:

- (((P43_7)*Males over 16 unemploy* + (P43_14)*Females over 16 unemployed*)/(P43_1)*Population 16 years and older*

Proportion of Population less than 18 Years is a variable in the construction of the composite measure *concentrated disadvantage*.

Table Proportion of Population less than 18 Years

	Block group	Census tract
Mean	.28	.28
Standard error	.005	.007
Standard deviation	.07	.05
Minimum	.02	.06
Maximum	.44	.38

The data to compute *proportion of population less than 18 years* were drawn from the following 2000 census Summary File 3 tables:

- ((P8_3 thru P8_20 + P8_42 thru P8_59)*Males and Females >1 year thru 17 years*/(P8_1)*Total population*

Proportion of Population Black/African American is a variable in the construction of the composite measure *concentrated disadvantage*. On average, 6 percent of the population in block groups or census tracts were African American, with at least one census tract and block group having no African American residents, and at least one block group or census tract composed of more than 20 percent African Americans.

[8] To be considered unemployed, individuals had to be available for work, but not engaged in it. Adults who were primary caregivers for young children or elderly relatives were not considered to be in the labor force and their lack of engagement in paid labor was not recorded as unemployment.

Table Proportion of Population Black/African American

	Block group	Census tract
Mean	.06	.06
Standard error	.004	.006
Standard deviation	.05	.04
Minimum	0.00	0.00
Maximum	.24	.20

The data to compute *proportion of population black/African American* were drawn from the following 2000 census Summary File 3 tables:

- ((P6_3)*Black or African American alone*/(P6_1)*Total Population*

Proportion of Population Latino is a variable in the construction of the composite measure *immigrant concentration*. On average, 6 percent of the population in block groups or census tracts were Latino, with at least one census tract and block group having virtually no Latino residents, and at least one block group composed of more than 20 percent Latinos.

Table Proportion of Population Latino/Hispanic

	Block group	Census tract
Mean	.06	.06
Standard error	.003	.004
Standard deviation	.04	.03
Minimum	0.00	.006
Maximum	.24	.15

The data to compute *proportion latino* were drawn from the following 2000 census Summary File 3 tables:

- ((P7_10)*Hispanic or Latino*/(P7_1)*Total Population*

Proportion of Population Foreign Born is a variable in the construction of the composite measure *immigrant concentration*. On average, nearly 10 percent of the population in block groups or census tracts were foreign born, with at least one census tract and block group having virtually no foreign born residents, and at least one block group a third foreign born.

Table Proportion of Population Foreign Born

	Block group	Census tract
Mean	.09	.08
Standard error	.004	.006
Standard deviation	.06	.04
Minimum	0.00	.02
Maximum	.33	.20

The data to compute *proportion foreign born* were drawn from the following 2000 census Summary File 3 tables:

- ((P21_13)*Foreign Born*/(P21_1)*Total Population*

Proportion Same House Last 5 Years is a variable in the construction of the composite measure *residential stability*. On average, slightly more than 40 percent of the population in block groups or census tracts were in the same home for the past five years, with at least one census tract and block group having less than 5 percent in the same home, and at least one block group or census tract composed of more than two-third of residents reside in the same home.

Table Proportion of Population 5 Years and Older in Same Residence 1995

	Block group	Census tract
Mean	.42	.41
Standard error	.01	.02
Standard deviation	.15	.13
Minimum	.04	.04
Maximum	.83	.67

The data to compute *proportion same house last 5 years* were drawn from the following 2000 census Summary File 3 tables:

- ((P24_2)*Same house in 1995*/(P24_1)*Population 5 years and older*

Proportion Owner Occupied House is a variable in the construction of the composite measure *residential stability*. On average, 56 percent of the population in block groups or census tracts live in owner occupied residences, with at least one census tract and block group having virtually no one in owner occupied residences, and at least one block group or census tract composed of nearly all residents owner occupied homes.

Table Proportion of Housing Units Owner Occupied

	Block group	Census tract
Mean	.56	.56
Standard error	.018	.033
Standard deviation	.27	.25
Minimum	0.00	.01
Maximum	1.00	.92

The data to compute *proportion housing owner occupied* were drawn from the following 2000 census Summary File 3 tables:

- ((H32_2)*Owner occupied housing units*/H30_1*Housing units*

Index of concentration at the extremes (ICE) is computed as a measure of inequality. It is computed as the difference between the number of affluent and poor households

divided by the number of households. Affluent households are those with incomes greater than $100,000. Poor households are those with incomes less than $20,000. These thresholds were established following the logic in the Morenoff, Sampson, and Raudenbush (2001) paper but modified the location of the thresholds used. In the original the thresholds were:

- poor families include those with incomes below the poverty level;[9] and,
- affluent families include those with incomes over $50,000.

These thresholds are not appropriate for Alaska. First, federal poverty levels obscure geographical variation in cost of living and minimally required income that are especially poignant outside the contiguous 48 states. Second, the median family income was about $50,000 in Alaska in 2005—and accordingly does not suggest that this grouping specifies an affluent group.

The $20,000 poor household threshold was set after review of a U.S. Department of Health and Human Services report titled "The 2005 HHS Poverty Guidelines." This research note suggests poverty guidelines for different sized families and in different geographies (separate estimates for 48 contiguous states and DC, Alaska, and Hawaii). This report establishes an Alaska poverty threshold for a three person family at $20,110.[10] This level (three person household) was chosen because the average Anchorage family size in is 3.19 and the average household size is 2.67. The poor household threshold, set as those earning less than $20,000, defines 12.4 percent of households as poor. The threshold for affluence was taken as the complement of poor—includes just the top earning households. The affluence threshold was set at households earning more than $100,000 and defines 18.8 percent of households as affluent.

The measure was computed by summing across tallies of households in income groups to develop counts of households that are poor and those that are affluent. The count of poor households sums across three groups (household incomes in 1999 under $10,000 (p52_2); $10,000 to 15,000 (p52_3); and $15,000 to $20,000 (p52_4)). The count of affluent households sums across four groups ($100,000 to $125,000 (p52_14); $125,000 to $150,000 (p52_15); $150,000 to $200,000 (p52_16); and more than $200,000 (p52_17)). The total number of household, the denominator, is in table p52_1.

[9] Though Morenoff, Sampson, and Raudenbush do not define what they mean by "poverty level," one may presume they mean the federal poverty level, which is calculated as a function of income and the number of individuals in the household. The U.S. Department of Health and Human Services publishes yearly poverty levels "for the contiguous 48 states and the District of Columbia" (http://aspe.hhs.gov/poverty/figures-fed-reg.shtml accessed September 19, 2006).

[10] In 1990, the federal poverty level for a family of three in the contiguous 48 states was $10,560.

The *index of concentration at the extremes* (ICE), as measured, is distributed as:

Table Index of Concentration at the Extremes

	Block group	Census tract
Mean	.043	.061
Standard error	.016	.028
Standard deviation	.227	.210
Minimum	-.502	-.259
Maximum	.632	.534

Proportion of Households Earnings less than $20,000 is an element in the construction of the composite measure *concentrated disadvantage*.

Table Proportion of Households with Incomes less than $20,000

	Block group	Census tract
Mean	.13	.13
Standard error	.007	.012
Standard deviation	.11	.09
Minimum	0.00	.01
Maximum	.56	.36

The data to compute *proportion of households earning less than $20,000* were drawn from the following 2000 census Summary File 3 tables:

- ((P52_2+P52_3+P52_4)/P52_1) P52_2 thru P52_4 are counts of households with incomes of less than $10,000, $10,000-$15,000, and $15,000-$20,000. P52_1 is the total number of households.

Proportion of Households Earning more than $100,000 is a variable in the construction of the composite measure *concentrated affluence*. On average, nearly 20 percent of families in block groups or census tracts earn more than $100,000, with at least one census tract and block group having no families (or nearly none) earning that much, and at least one block group or census tract having well over half of the families earning more than $100,000.

Table Proportion of Households with Incomes more than $100,000

	Block group	Census tract
Mean	.18	.19
Standard error	.01	.02
Standard deviation	.15	.14
Minimum	0.00	.03
Maximum	.64	.57

The data to compute *proportion of households earning more than $100,000* were drawn from the following 2000 census Summary File 3 tables:

- ((P52_14+P52_15+P52_16+P52_17)/P52_1) P52_14 thru P52_17 are counts of households with incomes between $100,000-$125,000, $125,000-$150,000, $150,000-$200,000, and more than $200,000. P52_1 is the total number of households.

Concentrated Affluence (also Socio-Economic-Status) is computed as an indicator that "…taps the upper end of the SES distribution…" (see Sampson, Morenoff and Earls, 1999:640). Sampson, Morenoff and Earls define concentrated affluence as "…as the percentage of families with incomes higher than $75,000, the percentage of adults with a college education, and the percentage of the civilian labor force employed in professional and managerial occupations" (1999:640). Our measure departs from theirs because we use a higher income threshold $100,000 household rather than $75,000 family, and we specified the proportion of adults with college education to mean those with baccalaureate degrees.[11] The table below presents the principal components factor loadings for *concentrated affluence* using the three variables suggested by Sampson and his colleagues.

Table Concentrated Affluence Factor Loadings, Anchorage, Block Groups and Census Tracts

Proportion of:	Block groups	Census tracts
Population 25 years plus Baccalaureate plus	.948	.978
Employed civilian in prof. or mgmt	.944	.969
Households >$100,000 income	.897	.948

The data to compute variables used to construct *concentrated affluence* were drawn from the following 2000 census Summary File 3 tables:

- Proportion of household with income >$100K((P52_14 thru P52_17)*Household incomes in 1999*/(P52_1)*Households*
- Proportion in prof. and mgmt. occupations ((P50 tables)*No. males and females in professional and management occupations*/(P50_1)*Employed civilian population 16 years and over*
- Proportion 25 years and older with baccalaureate degree or higher ((P37_15 thru P37_18 + P37_32 thru P37_35)*No. males and females with baccalaureate degree or higher*/(P37_1)*Population 25 years and over* (Note: proportion with some college was compared with proportion with baccalaureate or higher—the latter loaded stronger on the factor than the former.)

Proportion Employed in Professional and Management Occupations is a variable in the construction of the composite measure *concentrated affluence*. On average, about 40 percent of the adult population in block groups or census tracts were employed in

[11] We tested 'some college' or higher against baccalaureate degree or higher and the latter loaded more strongly on the factor than the former.

professional or management occupations, with at least one census tract and block group having less than 20 percent in those occupations, and at least one block group or census tract where over 60 percent were employed in professional or management occupations.

Table Proportion of Civilian Population 16 Years and Older Employed in Profession and Management Occupations

	Block group	Census tract
Mean	.35	.36
Standard error	.009	.015
Standard deviation	.13	.11
Minimum	.10	.14
Maximum	.73	.63

- The measure was computed from 2000 census tables that provide counts of males and females employed in the following occupations (census tables in parentheses):
    1. Management, business, and financial operations (p50_4, p50_51)
    2. Professional and related (p50_10, p50_57)

  ((P50 tables)*No. males and females in professional and management occupations/*(P50_1)*Employed civilian population 16 years and over)*100*

Proportion 25 years and older with baccalaureate degree or higher is a variable in the construction of the composite measure *concentrated affluence.* On average, about a quarter of the adult population in block groups or census tracts had 4-year college degrees, with at least one census tract and block group very few had degrees, and at least one block group or census tract where over 60 percent had degrees.

Table Proportion Population 25 Years and Older with Baccalaureate Degree or Higher

	Block group	Census tract
Mean	.27	.28
Standard error	.010	.017
Standard deviation	.14	.12
Minimum	.01	.07
Maximum	.66	.57

The data to compute *proportion baccalaureate* were drawn from the following 2000 census Summary File 3 tables:

- ((P37_15 thru P37_18 + P37_32 thru P37_35)*No. males and females with degree higher than baccalaureate*/(P37_1)*Population 25 years and over*

Population Density is treated as a structural antecedent (along with concentrated disadvantage, residential stability, concentrated immigration, and adults per child) in the production of collective efficacy for child monitoring and support (Sampson, Morenoff

and Earls, 1999:640. In Sampson and Raudenbush (1999) it is suggested that *population density* "Neighborhoods with more people per unit of space may generate greater anonymity and persons in public, making harder for residents to maintain informal social control over public space" (622). The measure is simply the population divided by the number of square miles in either the block group or census tract. The table below provides descriptive statistics for *population density* in Anchorage block groups and census tracts.

Table Population density

	Block group	Census tract
Mean	5,142.59	3,530.86
Standard error	255.61	317.97
Standard deviation	3,730.57	2,358.11
Minimum	.883	3.970
Maximum	18,753.73	8,772.51

The data to compute *population density* were drawn from the following 2000 census Summary File 3 tables:

- Population per square mile. Computed from ((P1_1)*Total population/Land Area(in square miles)*. Land area in data base expressed in square meters-conversion to square miles requires dividing #square meters by 2,589,988 (see p. 407 of SF3 Technical Document)—The land area in the data file is divided by 2,589.988 to produce number of people per square mile.

Ratio of Adults to Children is computed as a measure of "…structural imbalance across neighborhoods in the relative number of adults" to children (Sampson, Morenoff and Earls, 1999:640). The measure is computed as the number of adults divided by the number of children (those less than 18 years). The ratio of adults to children is distributed as follows:

Table Ratio of Adults to Children

	Block group	Census tract
Mean	3.21	2.88
Standard error	.28	.27
Standard deviation	4.08	2.01
Minimum	1.27	1.61
Maximum	56.50	16.51

- Ratio of those over 17 to those under 18. Computed as (1-Proportion of population less than 18 years)/Proportion of population less than 18 years. (See Concentrated Disadvantage above for measurement of proportion of the population less than 18 years)

Housing Density is not used in PHDCN studies. However, it is included here as a neighborhood feature implicated in crime, delinquency, and other forms of social disorder.

Table Housing Units per Square Mile

	Block group	Census tract
Mean	2,157.73	1,452.16
Standard error	116.60	141.80
Standard deviation	1,701.79	1,051.63
Minimum	0.45	2.75
Maximum	10,105.31	4,065.70

The data to compute *housing density* were drawn from the following 2000 census Summary File 3 tables:

- (H3_1) *No. of housing units /square miles*.

Proportion of Housing Vacant is not used in PHDCN studies but is included in the data base as an important marker of social disorganization especially relevant to studies of crime and delinquency.

Table Proportion of Housing Units Vacant

	Block group	Census tract
Mean	.06	.06
Standard error	.004	.007
Standard deviation	.05	.05
Minimum	0.00	.02
Maximum	.53	.36

The data to compute *proportion housing vacant* were drawn from the following 2000 census Summary File 3 tables:

- ((H31_1)*Vacant housing units*/(H30_1)*Housing units*

Proportion 18-64 Military is not used in PHDCN studies. It is included in the present study, however, because military presence is a visible and salient aspect of community life in Anchorage with Elmendorf Air Force Base and Fort Richardson's Army and National Guard Base in the city.

Table Proportion Population 18-64 Years Military

	Block group	Census tract
Mean	.03	.04
Standard error	.004	.013
Standard deviation	.06	.11
Minimum	0.00	0.00
Maximum	.61	.61

The data to compute *proportion military* were drawn from the following 2000 census Summary File 3 tables:

- (P39_4+P39_15)/(P39_3+P39_14) The number of men 18-64 who are military (P39_4) plus the number of women 18-64 who are military (P39_15) divided by the number of people 18-64.

PETER BLAU'S MEASURES OF SOCIAL STRUCTURE: HETEROGENEITY AND INEQALITY MEASURES

Structural analysis of communities is predicated on the availability of measures that describe social structures. Blau (1974) in his Presidential address to the American Sociological Association asserts that:

> Social structures are defined by their parameters—the criteria underlying the differentiation among people and governing social interaction… Two generic types of differentiation are heterogeneity and status inequality. Nominal parameters divide people into subgroups and engender heterogeneity. Graduated parameters differentiate people in terms of status rankings and engender inequality (615).

In his seminal presentation of his primitive theory of social structure Blau describes several measures of both inequality and heterogeneity. He suggests the following as examples of parameters of social structure (1977:8):

Nominal Parameters (heterogeneity)	Graduated Parameters (inequality)
Sex	Education
Race	Income
Religion	Wealth
Ethnic affiliation	Prestige
Clan	Power
Occupation	Socioeconomic origin
Place of work	Age
Place of residence	Administrative authority
Industry	Intelligence
Marital status	
Political affiliation	
National origin	
Language	

The following is a description of several measures of social structure computed for Anchorage census block groups and tracts.

HETEROGENEITY MEASURES

Measures of heterogeneity capture structural diversity by taking into consideration both the number of different classes of a "socially salient" characteristic and the distribution of individuals across those classes. The argument behind these measures asserts that diversity is greater when more classes are present and individuals are evenly distributed across those classes (see Blau, 1977, 1994; Langworthy, 1986).

Three measures of heterogeneity (industrial, occupational and racial) are constructed from 2000 census data in the SF-3 report. The measures are computed using an adaptation of Gibbs-Martin (Gibbs and Martin, 1962) measure of heterogeneity, $1 - \Sigma P_i^2$ where P_i^2 is the squared proportion of persons in group i (in this case industrial groups, occupational groups and racial groups). The census tables break categories out by gender so computation of the measures require several steps: 1) the tallies in male and female tables must be summed; 2) the sums in each category are divided by the number of employed civilian individuals 16 years and older and the proportion squared; and, 3) the squared proportions are summed to measures of heterogeneity.

Industrial heterogeneity is a measure of industrial employment diversity in an area (block groups, census tracts). The measure captures the degree to which individuals in the community are employed in varied industries and the degree to which those individuals are evenly spread across those industries. The value of the measure indicates more or less diversity and theoretically ranges from zero, when all employed person are employed in a single industry, to .923, when employed persons are evenly distributed across 13 industrial categories.[12] Industrial heterogeneity, as measured, is distributed as:

Table Industrial heterogeneity

	Block group	Census tract
Mean	.872	.885
Standard error	.002	.002
Standard deviation	.025	.016
Minimum	.734	.833
Maximum	.908	.902

This industrial diversity mean score is the equivalent to equal distribution across 7.8 industries in block groups and 8.7 industries in census tracts.

The measure was computed from 2000 census tables that provide counts of males and females employed in each of the following industries (census tables in parentheses):

1. Agriculture, forestry, fishing and hunting, mining (p49_3, p49_30)
2. Construction (p49_6, p49_33)
3. Manufacturing (p49_7, p49_34)
4. Wholesale trade (p49_8, p49_35)
5. Retail trade (p49_9, p49_36)

[12] See Langworthy (Appendix A, 1986) for discussion of the distribution of heterogeneity scores by number of categories and for interpretation of computed scores.

6. Transportation and warehousing, and utilities (p49_10, p49_37)
7. Information (p49_13, p49_40)
8. Finance, insurance, real estate and rental and leasing (p49_14, p49_41)
9. Professional/scientific/management/administrative/waste management services (p49_17, p49_44)
10. Educational, health and social services (p49_21, p49_48)
11. Arts/entertainment/recreation/accommodation and food service (p49_24, p49_51)
12. Other services (p49_27, p49_54)
13. Public administration (p49_28, p49_55)

Occupational heterogeneity is a measure of occupational employment diversity in an area (block groups, census tracts). The measure captures the degree to which individuals in the community are employed in a variety of occupations and the degree to which those individuals are evenly spread across those occupational categories. The value of the measure indicates more or less diversity and theoretically ranges from zero, when all employed person are employed in the same occupation, to .929 when employed persons are evenly distributed across 14 industrial categories. Occupational heterogeneity, as measured, is distributed as:

Table Occupational heterogeneity

	Block group	Census tract
Mean	.850	.860
Standard error	.002	.004
Standard deviation	.025	.033
Minimum	.734	.761
Maximum	.908	.906

These occupational diversity mean scores are equivalent to equal distribution across 6.7 occupational categories in block groups and 7.1 occupations in census tracts.

The measure was computed from 2000 census tables that provide counts of males and females employed in the following occupations (census tables in parentheses):

1. Management, business, and financial operations (p50_4, p50_51)
2. Professional and related (p50_10, p50_57)
3. Health care support (Service occupations p50_24, p50_71)
4. Protective services (p50_25, p50_72)
5. Food preparation and serving (p50_28, p50_75)
6. Building and grounds cleaning and maintenance (p50_29, p50_76)
7. Personal care and service (p50_30, p50_77)
8. Sales (p50_32, p50_79)
9. Office and administrative support (p50_33, p50_80)
10. Farming, fishing, and forestry (p50_34, p50_81)
11. Construction and extraction (p50_36, p50_83)
12. Installation, maintenance, and repair (p50_40, p50_87)
13. Production (p50_42, p50_89)

14. Transportation and material moving (p50_43, p50_90)

Racial heterogeneity is a measure of racial diversity in an area (block groups, census tracts). The measure captures the degree to which individuals in the community self associate with varied racial groups and the degree to which those individuals are evenly spread across those racial categories. The value of the measure indicates more or less diversity and theoretically ranges from zero, when all employed person are of the same race, to .857, when the population is evenly distributed across 6 racial categories. Racial heterogeneity, as measured, is distributed as:

Table Racial heterogeneity

	Block group	Census tract
Mean	.445	.442
Standard error	.013	.024
Standard deviation	.188	.175
Minimum	.000	.082
Maximum	.815	.795

These racial diversity mean scores are equivalent to equal distribution across 1.8 racial categories in both block groups and census tracts.

The measure was computed from 2000 census tables that provide counts of individuals in the following racial groups (census tables in parentheses):

1. White alone (P6_2)
2. Black or African American alone (P6_3)
3. American Indian and Alaska Native alone (P6_4)
4. Asian alone (P6_5)
5. Native Hawaiian and Other Pacific Islander alone (P6_6)
6. Some other race alone (P6_7)
7. Two or more races (P6_8)

INEQUALITY MEASURES

Income Inequality The measure is computed on the distribution of household incomes against the equal distribution of household incomes. The ***gini coefficient*** is a measure of the departure of the cumulative distributions of household incomes against theoretical equality. The gini coefficient from grouped data computation strategy is outlined in Rodrigue, J.P., et al. (2005). The measure is computed on group data as follows:

1. Compute estimated total household income for income groups. Census provides the number of households across 16 categories (<10, 10-15, 15-20, 20-25, 25-30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-75, 75-100, 100-125, 125-150, 150-200, >200). Estimates of total income by income group were computed as (#households in group (P52_*n*))*(income mid-point of the group *n*) for groups *1 thru 15.* Group 16 income is taken directly from census as the "aggregate household income for households earning over $200,000 (P54_3). The total of the estimated group

household incomes (for groups 1-15) plus the aggregate household incomes for household earning more than $200,000 is taken as the total household income for the geographic area.

To assess the adequacy of the estimated total it was compared to total aggregate household income. The comparisons were as follows ((estimate total)-(aggregate total))/(aggregate total)—an estimate of proportionate departure of the estimate from the aggregate. The tests for block group and census tract level estimates departures from aggregate totals are presented below:

Table Difference between estimated block group and census tract household income and aggregated household incomes

	Block group	Census tract
Mean	.0124	.0124
Standard error	.0008	.0007
Standard deviation	.0117	.0051
Minimum	-.0279	.0003
Maximum	.0496	.0210

These data indicate that estimated total income is about 1 percent departure on average for both block group and census tract estimates.

2. Compute the proportion of households in each income group as (# households in income group *n*)/(# households)---tables P52_2 thru P52_17 divided by P52_1.
3. Compute the proportion of estimated total household income that is associated with each household income group (estimated household income in each group)/(estimated total household income for the geographic unit)----both numerator and denominator were estimated in step 1 as outlined above.
4. Compute cumulative proportion of *estimated income* as income group increases. For group 1 (less than $10,000) the cumulative proportion is the proportion in group 1; for group 2 ($10-$15,000) the cumulative proportion is the proportion for group 2 plus the proportion for group 1; for group 3 ($15-$20,000) the cumulative proportion is sum of proportions for groups 1, 2, and 3; and so on.
5. Compute cumulative proportion of *households* as income group increases. Computed as in step 4 except proportion of households is aggregated rather than proportions of household income.
6. Cumulative proportion of estimated income for group *n* plus cumulative proportion of estimated income for group *n-1*.
7. Multiply step 6 and step 2.
8. Total group scores from step 7.
9. The gini is 1 minus the proceed of step 8

The attached table (Gini ct1.01.xls) provides intermediate statistics used to calculate the Gini coefficient for the inequality of household income for census tract 1.01 in Anchorage. Numbers in the header of the table correspond to the steps (1-9) outlined above

Inequality of Household Income (gini), as measured, is distributed as:

Table Income inequality

	Block group	Census tract
Mean	.353	.363
Standard error	.004	.007
Standard deviation	.065	.051
Minimum	.199	.265
Maximum	.531	.502
Range	.332	.237

References:

Blau, Peter (1974) "Presidential Address: Parameters of Social Structure," *American Sociological Review*, 39(5):615-35.

Blau, Peter (1977) *Inequality and Heterogeneity: A Primative Theory of Social Structure.* New York: Free Press

Blau, Peter (1994) *Structural Contexts of Opportunity.* Chicago: University of Chicago Press.

Earls, Felton and Albert Reiss, Jr. (1994) *Breaking the Cycle: Predicting and Preventing Crime*. Washington, DC: USGPO.

Gibbs, Jack and Walter Martin (1962) "Urbanization, Technology, and the Division of Labor: International Patterns." *American Sociological Review*, 27:667-77.

Langworthy, Robert (1986) *The Structure of Police Organizations.* New York: Praeger.

Morenoff, Jeffery, Robert Sampson and Stephen Raudenbush (2001) "Neighborhood Inequality, Collective Efficacy, and the Spatial Dynamics of Urban Violence." *Criminology*, 39(3):517-59.

Rodrigue, Jean-Paul, et al., (2005) *Transport Geography on the Web.* Hofstra University, Department of Economics and Geography, http://people. hofstra.edu/geotrans.

Sampson, R., S. Raudenbush and F. Earls (1997) "Neighborhood and Violent Crime: A Multilevel Study of Collective Efficacy," *Science*, 277:918-924.

Sampson, R., J. Morenoff and F. Earls (1999) "Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children," *American Sociological Review*, 64:633-660.

Sampson, R. and S. Raudenbush (1999) "Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods," *American Journal of Sociology*, 105(3):603-51.

U.S. Department of Health and Human Services (2005) "The 2005 HHS Poverty Guidelines: One Version of the [U.S] Federal Poverty Measures." Washington, DC: Author.

Attachment: Gini ct1.01.xls

Computation of Gini for Census Tract 1.01 (header numbers are step outined in documentation.									
Income group	# hhlds	1, hhld inc.(thou)	2, Phhld	3, Pinc	4, Cpinc	5, Cphhld	6	7, 2*6	8, Gini
<10K	13	65	0.0081	0.0005	0.0005	0.0081	0.0005	0	
10 to 15	36	450	0.0224	0.0032	0.0036	0.0305	0.0041	0.0001	
15 to 20	38	665	0.0237	0.0047	0.0083	0.0542	0.012	0.0003	
20 to 25	28	472.5	0.0131	0.0033	0.0117	0.0672	0.02	0.0003	
25 to 30	57	1567.5	0.0355	0.0111	0.0228	0.1027	0.0344	0.0012	
30 to 35	93	3022.5	0.0579	0.0214	0.0441	0.1606	0.0669	0.0039	
35 to 40	33	1237.5	0.0205	0.0087	0.0529	0.1812	0.097	0.002	
40 to 45	84	3570	0.0523	0.0252	0.0781	0.2335	0.131	0.0069	
45 to 50	58	2755	0.0361	0.0195	0.0976	0.2696	0.1757	0.0063	1-.677=.323
50 to 60	92	5060	0.0573	0.0358	0.1334	0.3269	0.2309	0.0132	
60 to 75	191	12892.5	0.1189	0.0911	0.2245	0.4458	0.3578	0.0426	
75 to 100	350	30625	0.2179	0.2165	0.441	0.6638	0.6655	0.145	
100 to 125	243	27337.5	0.1513	0.1932	0.6342	0.8151	1.0752	0.1627	
125 to 150	160	22000	0.0996	0.1555	0.7897	0.9147	1.424	0.1419	
150 to 200	81	14175	0.0504	0.1002	0.8899	0.9651	1.6797	0.0847	
>200	56	15570.1	0.0349	0.1101	1	1	1.8899	0.0659	
Total, Census tract 1.01	1606	141465.1						0.677	
Note: Figures taken from original calculation at seven decimal points. Totals may differ from sums due to rounding.									