Intersecting Inequalities: A Fuzzy-Set Analysis of Family Background, Test Scores, and Poverty

Inequality is a key feature of human social organization—some would say *the* key feature. In almost all known societies, inequalities coincide. Those at the top of social hierarchies do their best to fortify their advantages, while those at the bottom struggle to gain leverage. In modern societies, coinciding inequalities are reflected in the substantial correlations among individual-level aspects such as family background, education, and test scores. When studying life outcomes such as poverty, researchers typically estimate the net, independent contribution of these distinct yet correlated individual-level characteristics, treating each as an "independent" variable. In the *Bell Curve* debate, for example, scholars contest the "correct" estimate of the effect of test scores (from the Armed Forces Qualification Test) on poverty, net of the effect of family background and other correlated variables. I offer an alternative to the examination of correlations and the estimation of net effects. This alternative approach is based on the analysis of set-theoretic relations. To illustrate my approach, I present a fuzzy-set analysis of the same National Longitudinal Survey of Youth data set used by adversaries in the *Bell Curve* debate.

I. The Bell Curve Debate

The debate started in the mid 1990s following the publication of ***The Bell Curve*** by Richard Herrnstein and Charles Murray.

H&M argue that "intelligence"

- is (a) unidimensional, (b) inborn, and (c) relatively easy to measure.
- is more important than parental SES in its impact on life chances (e.g., staying out of poverty).
- has increased in importance because high cognitive ability is the key to success in an advanced, technologically sophisticated economy—a trend that is sure to continue.

The debate that H&M spawned

- is primarily about effect sizes.
- focuses mostly on the net effect of test scores (Armed Forces Qualification Test) relative to the effects of other causal conditions (e.g., parental SES).

The estimate of the net effect of test scores, like virtually all such estimates, is specification dependent. Consider, for illustration, the following logistic regression analyses of ***The Bell Curve*** data (National Longitudinal Survey of Youth), with poverty as the outcome ("in poverty" = 1).

The Bell Curve versus ***Inequality By Design*** (White sample)

	(A1)	(A2)	(A3)	(A4)	(A5)	(A6)	(A7)	(A8)	(A9)
Intercept	-2.8535**	-2.8815**	-2.7951**	-3.3355**	-1.0545	-1.5690	-.5260	.5255	2.2967
	(.0885)	(.0898)	(.2461)	(.3099)	(.9167)	(.9324)	(.9679)	(1.0577)	(.9546)
AFQT	**-.8434****	**-.7027****	**-.6839****	**-.6630****	**-.4511****	**-.4489****	**-.4494****	**-.4104****	
	(.0716)	**(.0806)**	**(.0823)**	**(.0842)**	**(.0979)**	**(.0979)**	**(.1003)**	**(.1078)**	
SES		-.2973							
		(.0802)							
Age		-.0478	-.0834	-.0210	-.0673	-.0644	-.0130	.0097	.0273
		(.0738)	(.0842)	(.0849)	(.1083)	(.1086)	(.1096)	(.1189)	(.1179)
Family Income			-.4441**	-.4416**	-.4328**	-.4199**	-.4194**	-.3781**	-.3957**
			(.1218)	(.1233)	(.1248)	(.1243)	(.1252)	(.1343)	(.1346)
Parents' SEI			-.0734	-.0736	-.0560	-.0444	-.0336	-.0475	-.0692
			(.0857)	(.0882)	(.0898)	(.0900)	(.0908)	(.1008)	(.0993)
Mother's Education			.0048	-.0345	.0048	.0058	.0297	.0651	.0342
			(.0895)	(.0908)	(.0932)	(.0935)	(.0941)	(.1019)	(.1013)
Father's Education			-.0334	-.0418	.0071	.0111	.0082	-.0134	-.0476
			(.0986)	(.1000)	(.1035)	(.1038)	(.1057)	(.1159)	(.1147)
Siblings (1979)			.1839**	.1685**	.1385*	.1429*	.1410*	.0399	.0351
			(.0650)	(.0675)	(.0687)	(.0691)	(.0699)	(.0768)	(.0765)
Farm Background			-.2593	-.2899	-.2868	-.3402	-.2153	-.2634	-.3220
			(.3120)	(.3185)	(.3184)	(.3219)	(.3229)	(.3604)	(.3563)
Two-Parent Family			-.2331	-.0919	-.1163	-.1073	-.0969	-.0544	-.0171
			(.2375)	(.2451)	(.2457)	(.2466)	(.2491)	(.2759)	(.2731)
Missing Fam. Income			.0235	.0720	.0799	.0416	-.0035	.0369	.0601
			(.2855)	(.2885)	(.2911)	(.2925)	(.2961)	(.3167)	(.3132)
Independent (Miss. Inc.)			.4364	.2048	.1107	.0952	-.0700	-.2059	-.2872
			(.3395)	(.3422)	(.3431)	(.3442)	(.3481)	(.3800)	(.3757)
Missing Parents' SEI			-.1740	-.1745	-.1546	-.1722	-.1909	-.0750	.0510
			(.3236)	(.3314)	(.3335)	(.3336)	(.3375)	(.3714)	(.3681)
Missing Mother's Ed.			-.0806	.0048	-.0741	-.0606	-.0149	.0780	.1868
			(.3556)	(.3571)	(.3570)	(.3581)	(.3622)	(.4106)	(.4027)
Missing Father's Ed.			.3615	.4141	.3616	.3906	.3576	.2448	.2415
			(.2659)	(.2675)	(.2671)	(.2692)	(.2722)	(.3079)	(.3047)
Fewer Dropout Students				-.1963**	-.1919**	-.2016**	-.2000**	-.2161**	-.2251**

				(.0729)	(.0743)	(.0740)	(.0754)	(.0831)	(.0824)
Fewer Disad. Students				-.0284	-.0207	.0232	.0079	-.0431	-.0695
				(.1257)	(.1262)	(.1269)	(.1280)	(.1375)	(.1363)
Fewer Nonwhite Students				-.3119*	-.3310*	-.3490*	-.3420*	-.3221*	-.3547*
				(.1384)	(.1395)	(.1434)	(.1441)	(.1552)	(.1554)
Missing Dropout Stud.				-.0428	.0651	.0239	-.0358	-.0963	.0363
				(.3927)	(.3907)	(.3922)	(.3941)	(.4489)	(.4316)
Missing Disad. Stud.				.1588	.1758	.2093	.2033	.1576	.1133
				(.2368)	(.2376)	(.2396)	(.2421)	(.2619)	(.2592)
Missing Nonwhite Stud.				.3189	.2138	.2274	.2183	.0902	-.0196
				(.3939)	(.3941)	(.3941)	(.3940)	(.4415)	(.4286)
West Region				.8926**	.8566**	.7905**	.8274**	.5445*	.5416*
				(.2258)	(.2291)	(.2380)	(.2401)	(.2608)	(.2593)
Northeast Region				.0397	.0947	.1081	.1266	-.0345	-.0521
				(.2760)	(.2785)	(.2860)	(.2881)	(.3104)	(.3078)
Central Region				.5798**	.5782**	.4813*	.5057*	.2806	.2907
				(.2033)	(.2054)	(.2161)	(.2183)	(.2353)	(.2349)
Years of Ed. pre-AFQT					-.1575*	-.1575*	-.2182**	-.2489**	-.3817**
					(.0736)	(.0737)	(.0761)	(.0829)	(.0760)
H.S. Academic Track					-.4899	-.4795	-.4695	-.2176	-.3502
					(.2568)	(.2578)	(.2584)	(.2737)	(.2683)
Years of Ed. post-AFQT					-.2235**	-.2326**	-.2295**	-.1910**	-.2550**
					(.0676)	(.0684)	(.0684)	(.0711)	(.0689)
Unemployment Rate (1990)						.0833*	.0809	.0701	.0710
						(.0423)	(.0426)	(.0472)	(.0467)
Central City (1990)						.5186	.5749*	.4892	.4270
						(.2696)	(.2714)	(.2868)	(.2853)
Rural (1990)						.1534	.1706	.1410	.1406
						(.1970)	(.1988)	(.2183)	(.2159)
Male							-.8219**	-.9718**	-.9973**
							(.1649)	(.2254)	(.2226)
Children (1990)								.7213**	.7141**
								(.0887)	(.0875)
Married (1990)								-3.0629**	-3.1042**
								(.2569)	(.2570)
Married Man (1990)								1.0432**	1.1044**
								(.3605)	(.3576)
Pseudo R^2	0.0948	0.1037	0.1209	0.1475	0.1623	0.1682	0.1849	0.3238	0.3145

The Bell Curve versus ***Inequality By Design*** (African-American sample)

	(B1)	(B2)	(B3)	(B4)	(B5)	(B6)	(B7)	(B8)	(B9)
Intercept	-1.1171**	-1.1513**	-1.2421**	-1.8618**	2.2521**	1.4504*	2.6730**	2.0534**	3.0998**
	(.0601)	(.0617)	(.1380)	(.1788)	(.6820)	(.7194)	(.7604)	(.8249)	(.7762)
AFQT	**-.8031****	**-.6869****	**-.6858****	**-.6705****	**-.3730****	**-.3694****	**-.3904****	**-.3638****	
	(.0674)	**(.0707)**	**(.0725)**	**(.0746)**	**(.0850)**	**(.0856)**	**(.0885)**	**(.0945)**	
SES		-.2973	-.3488**						
		(.0802)	(.0630)						
Age		-.0478	-.0462	-.0364	-.0110	.0003	.0187	-.0088	-.0094
		(.0738)	(.0582)	(.0658)	(.0821)	(.0822)	(.0844)	(.0900)	(.0894)
Family Income			-.4441**	-.3435**	-.3308**	-.3151**	-.2717**	-.2499*	-.2629*
			(.1218)	(.0994)	(.1009)	(.1019)	(.1004)	(.1050)	(.1062)
Parents' SEI			-.0734	-.2221**	-.2307**	-.2329**	-.2603**	-.2007*	-.2269**
			(.0857)	(.0780)	(.0805)	(.0810)	(.0828)	(.0873)	(.0868)
Mother's Education			.0048	-.1613*	-.0851	-.0851	-.0437	-.0173	-.0368
			(.0895)	(.0744)	(.0765)	(.0770)	(.0794)	(.0857)	(.0847)
Father's Education			-.0334	.0369	.0811	.0882	.0924	.0917	.0839
			(.0986)	(.0853)	(.0879)	(.0884)	(.0901)	(.0971)	(.0963)
Siblings (1979)			.1839**	-.0354	-.0615	-.0662	-.0732	-.0464	-.0228
			(.0650)	(.0645)	(.0661)	(.0665)	(.0684)	(.0729)	(.0719)
Farm Background			-.2593	-.0053	-.1176	-.2187	-.1888	-.2558	-.2127
			(.3120)	(.3592)	(.3658)	(.3701)	(.3857)	(.4323)	(.4262)
Two-Parent Family			-.2331	-.0843	-.0558	-.0431	-.0376	-.0143	.0022
			(.2375)	(.1477)	(.1503)	(.1516)	(.1546)	(.1643)	(.1633)
Missing Fam. Income			.0235	-.2040	-.2023	-.2531	-.2189	-.2863	-.2079
			(.2855)	(.2309)	(.2351)	(.2376)	(.2419)	(.2599)	(.2600)
Independent (Miss. Inc.)			.4364	.7011*	.6119	.6490*	.4181	.3977	.2968
			(.3395)	(.3053)	(.3137)	(.3166)	(.3254)	(.3493)	(.3457)
Missing Parents' SEI			-.1740	.1473	.1471	.1138	.3254	.1115	.1477
			(.3236)	(.1710)	(.1758)	(.1770)	(.1807)	(.1917)	(.1911)
Missing Mother's Ed.			-.0806	.0921	-.0609	-.0429	-.0123	.0490	.0817
			(.3556)	(.2167)	(.2219)	(.2229)	(.2305)	(.2434)	(.2434)
Missing Father's Ed.			.3615	.1862	.1503	.1730	.1843	.1115	.1275
			(.2659)	(.1448)	(.1472)	(.1483)	(.1530)	(.1623)	(.1610)
Fewer Dropout Students				-.1257	-.0602	-.0761	-.0571	-.0575	-.0567

				(.0766)	(.0799)	(.0814)	(.0820)	(.0863)	(.0862)
Fewer Disad. Students				-.1519	-.1543	-.1513	-.1648	-.1482	-.1620
				(.0808)	(.0826)	(.0837)	(.0858)	(.0913)	(.0913)
Fewer Nonwhite Students				-.1096	-.1405	-.1629	-.1373	-.1832	-.1999
				(.0940)	(.0956)	(.0997)	(.1012)	(.1075)	(.1070)
Missing Dropout Stud.				-.0028	-.1006	-.0724	-.1855	-.3428	-.3495
				(.3236)	(.3272)	(.3311)	(.3408)	(.3604)	(.3577)
Missing Disad. Stud.				.1306	.1658	.1914	.1221	.1929	.2221
				(.2117)	(.2155)	(.1914)	(.2252)	(.2420)	(.2415)
Missing Nonwhite Stud.				.4012	.4267	.4366	.5533	.6538	.7071
				(.3523)	(.3560)	(.3613)	(.3693)	(.3860)	(.3825)
West Region				.6579*	.7042**	.6559*	.6821*	.6189*	.6616*
				(.2626)	(.2658)	(.2737)	(.2816)	(.2935)	(.2920)
Northeast Region				.4669*	.3560	.2505	.3367	.2105	.1208
				(.1884)	(.1944)	(.2043)	(.2065)	(.2152)	(.2137)
Central Region				.8355**	.8132**	.7523**	.7980**	.5757**	.5206**
				(.1577)	(.1599)	(.1645)	(.1689)	(.1816)	(.1792)
Years of Ed. pre-AFQT					-.3260**	-.3183**	-.3861**	-.3490**	-.4351**
					(.0571)	(.0575)	(.0602)	(.0641)	(.0600)
H.S. Academic Track					-.4329**	-.4361**	-.4025*	-.3631*	-.4513**
					(.1646)	(.1655)	(.1679)	(.1776)	(.1749)
Years of Ed. post-AFQT					-.2459**	-.2399**	-.2565**	-.2092**	-.2617**
					(.0557)	(.0560)	(.0573)	(.0609)	(.0588)
Unemployment Rate (1990)						.1200**	.1177	.1282**	.1273**
						(.0406)	(.0414)	(.0444)	(.0441)
Central City (1990)						.1553	.1689	.0372	.0613
						(.1559)	(.1598)	(.1698)	(.1684)
Rural (1990)						.2053	.2118	.2774	.2974
						(.1706)	(.1745)	(.1870)	(.1861)
Male							-.9977**	-.6644**	-.6653**
							(.1338)	(.1712)	(.1701)
Children (1990)								.4252**	.4200**
								(.0609)	(.0607)
Married (1990)								-2.0396**	-2.0829**
								(.2272)	(.2264)
Married Man (1990)								.1896	.2241
								(.3629)	(.3621)
Pseudo R^2	0.0859	0.1021	0.1175	0.1469	0.1771	0.1833	0.2129	0.2917	0.2841

Notes:

- AFQT scores have been massaged (by H&M) so that they are normally distributed (the raw test scores are not) and then converted to z scores.
- Parental SES is also converted to z scores, to permit direct comparison of its effect with that of AFQT.
- H&M avoid using AFQT test percentile scores because they are very interested in foregrounding the effect of being in the “cognitive elite.” That is, they want to make sure that their analysis assesses the impact of being in the 99th percentile versus the 99.9th percentile versus the 99.99th percentile.

Observations:

- The effect of test scores declines as the number of competing variables is increased (from 2 to more than 20).
- The standard error of the test score variable increases as the number of competing variables increases.
- The impact of disaggregating SES into its components is nontrivial (this is one of Fischer et al.’s main points).
- In Fischer et al.’s (final) analysis the independent contribution of test scores is very small (compare the pseudo R^2 values in the last two columns).
- Overall, the pseudo R^2 values increase from small (around 10% in H&M’s analysis) to moderate (around 30% in Fischer et al.’s analysis).
- The results for Whites and African-Americans are very similar.

A Middle Path Between H&M and Fischer et. al.

	White Males		White Females		Black Males		Black Females	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	1.527	2.927***	1.953*	2.945***	3.748***	4.291***	5.533***	6.477***
	(.971)	(.864)	(.993)	(.921)	(.863)	(.820)	(.926)	(.878)
AFQT (percentile)	-.021***		-.018**		-.016*		-.021**	
	(.006)		(.006)		(.008)		(.007)	
Parental Income	-.151**	-.166***	-.051	-.063	-.121*	-.140*	-.121**	-.142**
	(.051)	(.051)	(.038)	(.038)	(.059)	(.058)	(.046)	(.046)
Parental Education	-.036	.010	-.019	-.003	-.028	-.044	-.034	-.022
	(.056)	(.055)	(.052)	(.051)	(.044)	(.043)	(.040)	(.040)
Respondent Education	-.212*	-.362***	-.256**	-.373***	-.340***	-.389***	-.480***	-.572***
	(.085)	(.073)	(.084)	(.074)	(.068)	(.064)	(.076)	(.070)
Married	-1.544***	-1.567***	-2.855***	-2.885***	-1.767***	-1.813***	-2.093***	-2.125***
	(.341)	(.332)	(.271)	(.269)	(.338)	(.337)	(.244)	(.244)
Children	.738*	.699*	1.745***	1.777***	.569*	.592*	.769***	.753***
	(.343)	(.332)	(.273)	(.271)	(.271)	(.270)	(.225)	(.223)
N	1363	1363	1315	1315	732	732	775	775
Pseudo R^2	0.182	0.164	0.329	0.317	0.173	0.167	0.286	0.277

Notes:

- The analysis is by race and gender, not just by race. This approach reveals the stronger impact of marriage (positive) and children (negative) on poverty status for females.
- This difference aside, the results are remarkably similar for the four subsamples.
- The effect of AFQT scores, when viewed from the perspective of the pseudo R^2 increment, is (again) very modest.
- Pseudo R^2 values are somewhat higher for females than for males.

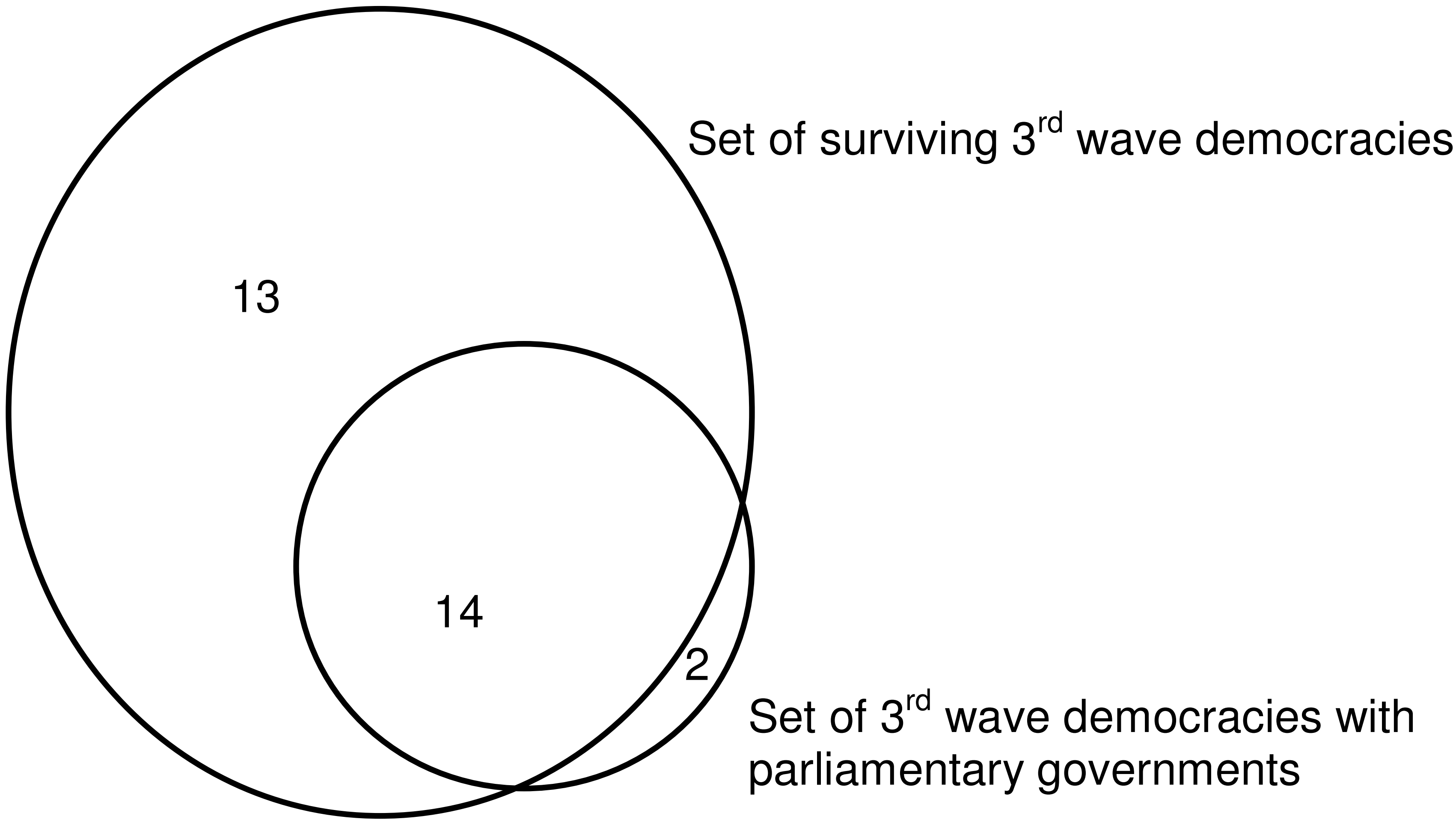
Set Coincidence

Set coincidence combines and bridges consistency and coverage. Set coincidence focuses on the degree to which two sets overlap—that is, the degree to which they are one and the same set.

While degree of set coincidence can be assessed using multiple sets (i.e., more than two), it is easiest to grasp the basic principles using two sets. For example, the degree to which the set of *surviving* 3rd wave democracies and the set of 3rd wave democracies *with parliamentary governments* are "one and the same" is indicated by the degree to which the cases than have *both* of these two traits embraces the set of cases that have *either* trait. In other words, set coincidence is the number of cases found in the intersection of two sets, expressed relative to the number of cases found in their union:

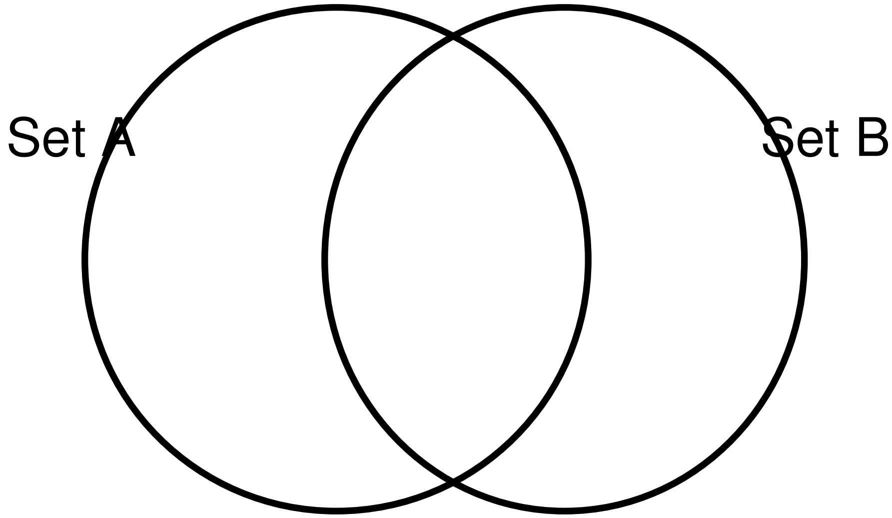
(# of cases in intersection)/(# of cases in union)

In the next figure, the coincidence of "Parliamentary" and "Democracy Survived" is 14/29 = 0.483 (i.e., relatively modest).

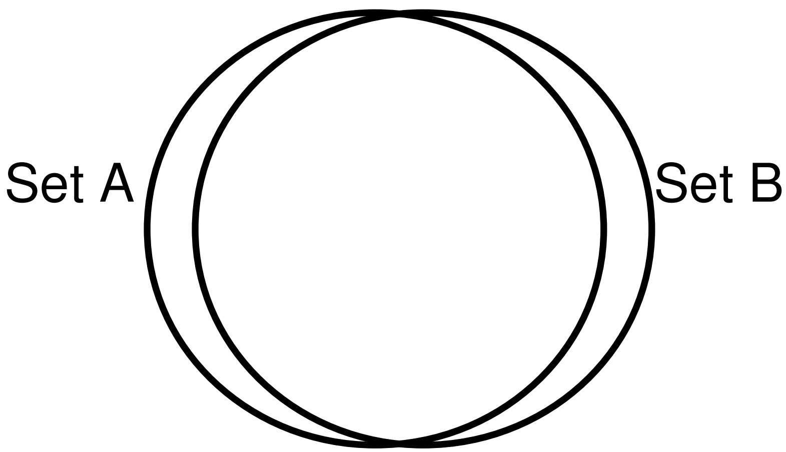


Coincidence of sets = 14/(13+ 14+ 2) = 14/29 = 0.483

Here are two graphic examples, showing the contrast between high and low coincidence:



Low set coincidence: set intersection is a small fraction of set union.

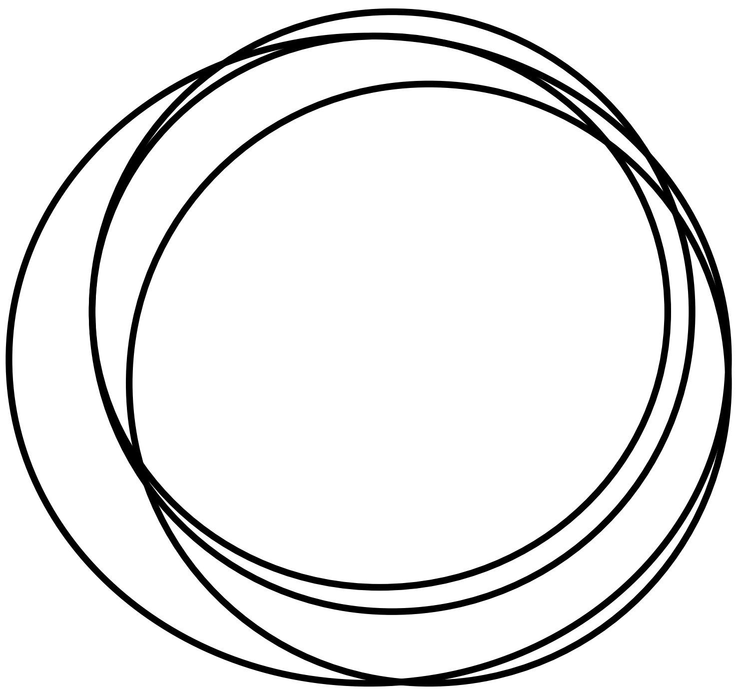


High set coincidence: set intersection almost fully “covers” set union.

Set Coincidence and the Study of Social Inequality

A. Multiple Set Coincidence

Sometimes the intersection of several sets can occupy a large proportion of the union of these same sets. This would occur, for example, in a situation where social advantages (or disadvantages) strongly overlap. The formula for the calculation of the coincidence of multiple sets is the same as it is for two sets: *(sum of membership in the intersection)/(sum of membership in the union).*



As the number of sets increases, the possibility of strong overlap decreases, unless the pattern of set coincidence is very strong.

Implications of Multiple Set Coincidence for the Study of Social Inequality

A basic sociological principle is that people try to compound their advantages and try to avoid having multiple disadvantages.

This notion of “compounding” is directly captured by the concept of multiple set coincidence. If advantages or disadvantages tend to cohere (i.e., to be multiple), compounding will be reflected in the relative number of people who combine multiple traits. In other words, if compounding is present, the *intersection* of the relevant sets will “cover” much of the *union* of these same sets.

The analysis of set coincidence, therefore, is central to the analysis of social inequality.

The Asymmetry of Set Coincidence

At first glance, it may appear that set coincidence is roughly the same as correlation. It is not. Set coincidence is asymmetric and thus sensitive to the specification of the sets in question. Consider the following table:

	Supports Reform	Opposes Reform
Republican	50	250
Democrat	100	50

Focusing on the coincidence of “Republican” with “Opposes Reform,” the calculation is intersection/union (# of cases in both sets / # of cases in either set):

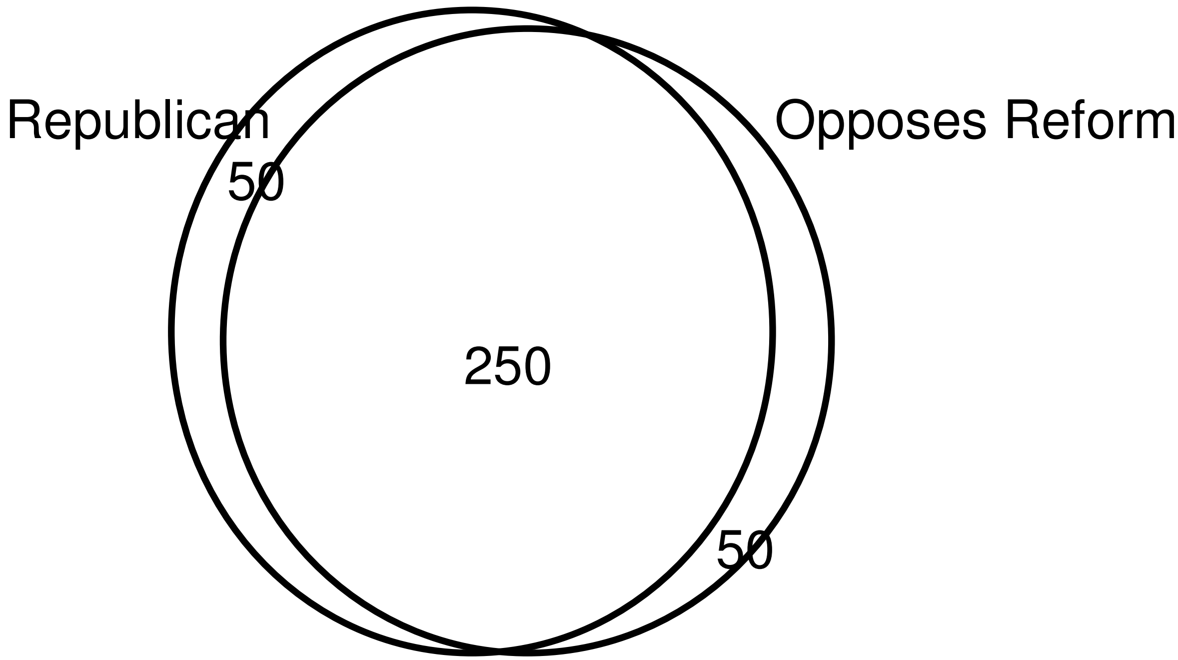
$$250 / (250 + 50 + 50) = 250 / 350 = 0.71$$

However, shifting the focus to the coincidence of “Democrat” with “Supports Reform” yields a different calculation:

$$100 / (100 + 50 + 50) = 100 / 200 = 0.50$$

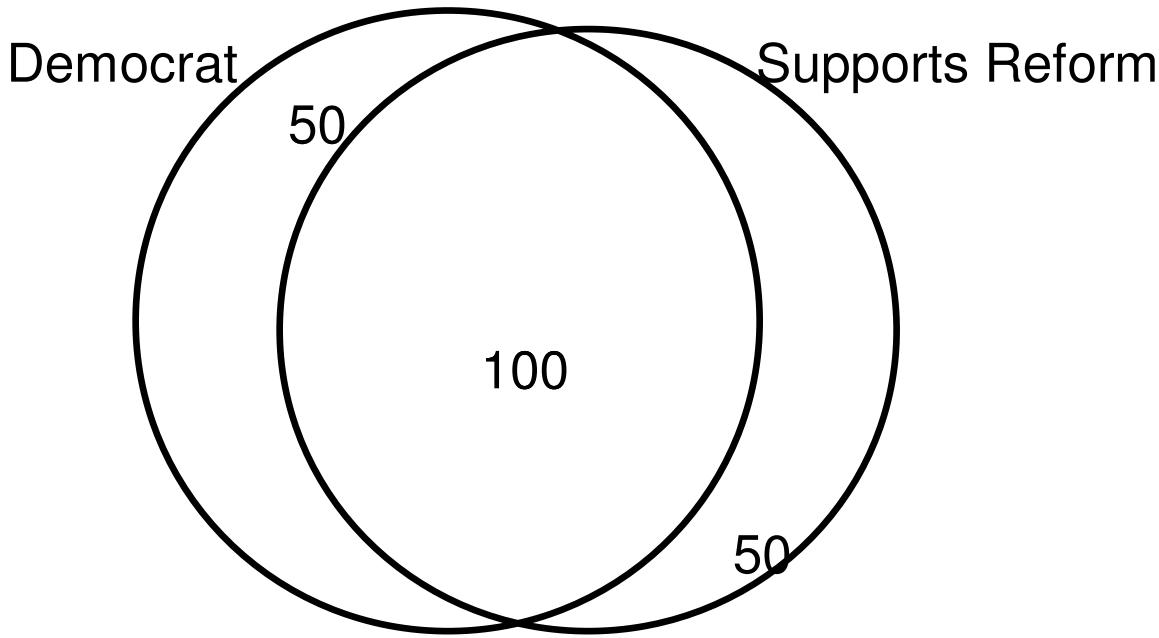
Here are the contrasting emphases as Venn diagrams:

Coincidence of “Republican” and “Opposes Reform”



Set coincidence = 250/(250 + 50 + 50) = 0.71

Coincidence of "Democrat and Supports Reform"



Set coincidence = 100/(100+ 50 + 50) = 0.50

Implications of the Asymmetry of Set Coincidence for the Study of Social Inequality

Because set coincidence is asymmetric, the assessment of the degree to which advantages are combined is distinct from the assessment of the degree to which disadvantages are combined. This feature distinguishes set coincidence from correlational measures and provides the opportunity to differentiate the compounding of advantages from the compounding of disadvantages. In the language of set theory, the issue is which is stronger, the coincidence of sets A and B or the coincidence sets of ~A and ~B?

If advantages coincide strongly and disadvantages do not (or at least not as strongly), the implication is that people strive to fortify their position by seeking to combine and reinforce their advantages.

If disadvantages coincide strongly and advantages do not (or at least not as strongly), the implication is that people may succumb to downward social forces and be subject to an accumulation of disadvantages.

The relative importance of coinciding advantages versus coinciding disadvantages may differ by race and gender.

C. The Divergence of Set Coincidence and Correlation

It is possible for two sets to display **strong coincidence**, yet as variables exhibit only a **weak correlation**.

For example, a researcher might want to assess the degree to which respondents combine moderate-to-high parental income with moderate-to-high AFQT test scores. In the following table, most respondents (500) combine the two advantages. A moderate number of respondents (200) have one advantage but not both, and a small number have neither advantage (50).

Parental Income	**low AFQT scores**	**medium-high AFQT scores**	*Total*
medium-high	100	500	600
low	50	100	150
Total	150	600	750

Set coincidence of medium-high parental income with medium-high test scores:
500/700 = **0.714** (i.e., a high level of confounding)

Correlation = **0.167** (i.e., a low level of confounding)

Implications of the Divergence of Set Coincidence from Correlation for the Study of Social Inequality

From the perspective of correlational/net effects analysis, a weak correlation provides an opportunity to estimate net effects without much concern for confounding. However, strong set coincidence may coexist with weak correlations.

From a set theoretic point of view, strong set coincidence raises questions about the utility of analyses that seek to disentangle the effects of overlapping characteristics.

V. Concept of Set Coincidence Applied to *Bell Curve* Data

The outcome is

> ***not-in poverty,*** *a fuzzy set based on the ratio of the respondent's household income to the poverty level for households of that type.*

The four main causal conditions (advantages/disadvantages) are

> *1.* ***parent educated****, a fuzzy set based on the years of education (for the parent with more years of education);*
> *2.* ***not-low income parents****, a fuzzy set based on the ratio of parental household income to the poverty level for households of that type;*
> *3.* ***not-low AFQT score****, a fuzzy set based on AFQT percentile scores;*
> *4.* ***respondent educated****, a fuzzy set based on respondent's years of education.*

Question #1: Which are stronger, coinciding advantages or coinciding disadvantages?

Sets	black females	black males	white female	white males
nlpinc nlafqt	0.491	0.482	0.861	0.844
lpinc lafqt	*0.460*	*0.405*	*0.131*	*0.150*
nlafqt edu	0.538	0.523	0.745	0.727
lafqt neduc	*0.450*	*0.460*	*0.259*	*0.286*
nlpinc peduc	0.514	0.522	0.662	0.669
lpinc npeduc	*0.479*	*0.426*	*0.108*	*0.119*

Advantages
nlpinc = not-low parental income
nlafqt = not-low test scores (AFQT)
educ = educated respondent
peduc = educated parent

Disadvantages
lpinc = low parental income
lafqt = low test scores (AFQT)
neduc = not educated respondent
npeduc = not educated parent

In general, advantages coincide more than disadvantages. Also, there is a very striking racial difference—whites enjoy much stronger coinciding advantages and have a very low level of coinciding disadvantages.

Question #2: How strongly do *multiple* advantages coincide?

Black females

	Coincidence
nlpinc*nlafqt	0.491
educ*nlpinc*nlafqt	0.371
educ*peduc*nlpinc*nlafqt	0.283

Black males

	Coincidence
nlpinc*nlafqt	0.482
educ*nlpinc*nlafqt	0.358
educ*peduc*nlpinc*nlafqt	0.284

White females

	Coincidence
nlpinc*nlafqt	0.861
nlpinc*nlafqt*educ	0.676
educ*peduc*nlpinc*nlafqt	0.573

White males

	Coincidence
nlpinc*nlafqt	0.844
educ*nlpinc*nlafqt	0.654
educ*peduc*nlpinc*nlafqt	0.559

Question #3: Do correlations and coincidence scores agree?

White males: coincidence scores

	peduc	nlpinc	nlafqt
nlpinc	0.669	--	--
nlafqt	0.672	0.844	--
educ	0.767	0.710	0.727

Average coincidence score = 0.731

White males: correlations

	peduc	nlpinc	nlafqt
nlpinc	0.231	--	--
nlafqt	0.373	0.256	--
educ	0.477	0.227	0.518

Average correlation = 0.347

nlpinc = not-low parental income
educ = educated respondent

nlafqt = not-low test scores (AFQT)
peduc = educated parent

White females: coincidence scores

	peduc	nlpinc	nlafqt
nlpinc	0.662	--	--
nlafqt	0.669	0.861	--
educ	0.775	0.716	0.745

Average coincidence score = 0.738

White females: correlations

	peduc	nlpinc	nlafqt
nlpinc	0.219	--	--
nlafqt	0.339	0.229	--
educ	0.501	0.181	0.477

Average correlation = 0.324

nlpinc = not-low parental income
educ = educated respondent

nlafqt = not-low test scores (AFQT)
peduc = educated parent

Black males: coincidence scores

	peduc	nlpinc	nlafqt
nlpinc	0.522	--	--
nlafqt	0.457	0.482	--
educ	0.623	0.559	0.523

Average coincidence score = 0.528

Black males: correlations

	peduc	nlpinc	nlafqt
nlpinc	0.351	--	--
nlafqt	0.330	0.279	--
educ	0.283	0.192	0.489

Average correlation = 0.321

nlpinc = not-low parental income
educ = educated respondent

nlafqt = not-low test scores (AFQT)
peduc = educated parent

Black females: coincidence scores

	peduc	nlpinc	nlafqt
nlpinc	0.514	--	--
nlafqt	0.451	0.491	--
educ	0.620	0.573	0.538

Average coincidence scores = 0.531

Black females: correlations

	peduc	nlpinc	nlafqt
nlpinc	0.388	--	--
nlafqt	0.321	0.335	--
educ	0.387	0.336	0.533

Average correlation: 0.383

nlpinc = not-low parental income
educ = educated respondent

nlafqt = not-low test scores (AFQT)
peduc = educated parent

VI. Fuzzy Set Analysis of Poverty Status

How consistently do respondents with coinciding advantages avoid poverty?

	Consistency	Coverage
Black Females	0.793	0.401
Black Males	0.848	0.339
White Females	0.878	0.653
White Males	0.898	0.619

The consistency scores show the degree to which those who combine all four advantages are able to avoid poverty (i.e., the degree to which they constitute a subset of those avoiding poverty).

The coverage scores show how common the combination of advantages is among those who successfully avoid poverty. Because the combination is not as common among Blacks, it is also not as common among Blacks who successful avoid poverty.

How consistently do respondents with coinciding disadvantages experience poverty/low income

	Consistency	Coverage
Black Females	0.861	0.346
Black Males	0.646	0.329
White Females	0.681	0.098
White Males	0.679	0.148

The consistency scores show the degree to which those who combine all four disadvantages experience poverty (i.e., the degree to which they constitute a subset of those in poverty).

The coverage scores show how common the combination of disadvantages is among those who are in poverty.

The only high consistency scores for the link between “coinciding disadvantages” and in-poverty is for black females. This finding indicates that the problem of reinforcing disadvantages applies only to this group.

VII. Conclusions

Herrnstein and Murray argue, based on their research, that if a person could chose between being born into a high SES family or being born with a high level of “intelligence,” it would be better to choose “intelligence.” They base this statement on the stronger net effect of AFQT scores, compared to parental SES, on life outcomes such as poverty.

The set coincidence analysis I have presented shows clearly, for whites especially, that choosing either high SES or “intelligence” usually involves choosing the other. The set coincidence scores are very high, so much so that the whole idea of calculating the “net effect” of either seems hazardous, from a set theoretic perspective.

More generally, the striking racial differences in coinciding advantages is invisible to correlational /net effects analysis. Both the logistic regression results and the correlational analysis show striking similarities across racial groups. This homogeneity contradicts both everyday experience and set theoretic analysis.

For whites, advantages cohere and appear to reinforce; disadvantages do not. For blacks, there is evidence of both reinforcing advantages and reinforcing disadvantages. However, the prevalence of reinforcing advantages is much lower for blacks than for whites.

Calculating Set Coincidence Using fsQCA

Unfortunately, fsQCA does not produce set coincidence calculations automatically. However, they can be calculated using the program. Here are the steps:

1. Select the two causal conditions of interest.

2. Using the compute variables function in the data window, calculate a new variable, which is the maximum of the two selected causal conditions. Use the "fuzzyor" function (fuzzyor is set union) to get the max.

3. Use the subset/superset procedure to calculate degree of coincidence:
    (a) specify the max variable you created in step 2 as the outcome
    (b) specify the two conditions joined via the "fuzzyor" function as the causal conditions
    (c) the "coverage" of the first recipe in the output window is the degree of set coincidence