Are Smarter Groups More Cooperative? Evidence from Prisoner's Dilemma Experiments, 1959-2003

Abstract

Are more intelligent groups better at cooperating? A meta-study of repeated prisoner's dilemma experiments run at numerous universities suggests that students cooperate 5% to 8% more often for every 100 point increase in the school's average SAT score. This result survives a variety of robustness tests. Axelrod (1984) recommends that the way to create cooperation is to encourage players to be patient and perceptive; experimental evidence suggests that more intelligent groups implicitly follow this advice.

"The 'Prisoner's Dilemma'—to cooperate or to defect, to trust or to distrust, to sacrifice immediate profit for the sake of long-run gain—is a situation which occurs time and again in the marketplace, in the smoke-filled rooms of a political convention, on Wall Street, and on Capitol Hill. But 'pure' game theory assumes completely rational players, and as every reader will attest, human beings can be anything but rational…."

Lave (1965), p. 26.

"The key human capital externality is not technological but political…."

Glaeser et al. (2004), p. 14.

I. Introduction

Are more intelligent groups of people better at cooperating? Repeated prisoner's dilemma (RPD) experiments run at numerous universities since 1959 may hold the answer. Overall, the tendency is clear: Students at schools with higher average SAT and ACT scores cooperate much more often—yielding about 5% more cooperation for every 100 combined SAT points (See Figures 1-3). Students at the best schools cooperate about 15% more often than typical college students.

These results may be of special interest to growth and development economists: Recent research has demonstrated that national average IQ as measured by Lynn and Vanhanen (2002, 2006) is a strong predictor of national economic performance.[1] Lynn and Vanhanen summarize hundreds of IQ studies from across the 20th and 21st centuries and find that the IQ gap between countries in the 5th and 95th percentiles is 38 IQ points—equivalent to 2.3 standard deviations

[1] Deary (2001) is recommended as a brief scholarly primer into modern intelligence research; Jensen (1998) is highly recommended as the standard lengthy treatment.

within the U.S. population (n.b. The miracle economies of East Asia have had the highest average IQ scores for as long as such tests have been available.). Jones and Schneider (2006b) demonstrate a correlation of 0.8 between national average IQ and log GDP per capita, and in a Bayesian model averaging test, Jones and Schneider (2006a) found that in steady state 1 IQ point is associated with 6% more output per person (for similar results cf., *inter alia*, Ram (2006), Weede and Kampf (2002), Lynn and Vanhanen (2006)).

However, Jones and Schneider (2006b) report that in the labor literature, the correlation between IQ and log wages is only about 0.3, and 1 IQ point is associated with only about 1% higher wages. Thus, IQ seems to matter more at the macro level than at the micro level. The question arises: Why the difference between individual and aggregate outcomes?

One possibility is that smarter individuals are more patient and therefore have higher savings rates. In the aggregate, if the economy is closed or if there are other reasons why investment and savings rates would be tightly linked (e.g., Aghion, Comin, and Howitt (2006)), then economies with smarter workers would also have larger capital stocks. Another possibility is that, at least in democracies, more intelligent voters are more likely to choose productive economic policies; for example, Caplan (2007) shows that more educated voters are more likely to agree with professional economists on a variety of topics even after controlling for income and political affiliation, while Caplan and Miller (2007) shows that the same is true for more intelligent voters when intelligence is measured through a simple vocabulary test. Thus, both neoclassical channels such as capital accumulation and Northian channels of institutional choice could explain causation running from individual cognitive abilities to strong national economic performance.

Reverse causation is also an important possibility: As Jones and Schneider (2006) and Lynn and Vanhanen (2002, 2006) note, in the world's poorest countries, poor public health and nutrition quite likely damage the cognitive performance of the typical citizen. In the recent Copenhagen Consensus project a panel of eight prominent economists including Nobel laureates Robert Vogel, Thomas Schelling, and Vernon Smith evaluated various proposals to improve the well-being of humanity; their number-two-rated proposal was to include micronutrients in the food supply. They concluded that including simple micronutrients would, among other things, improve the cognitive functioning of the world's poorest people (Berhman et al. (2004)).

The fact that causation can run through multiple channels *and* in both directions makes it all the more important to use detailed micro-level evidence to identify which channels are quantitatively crucial. The results presented here begin to do just that. Cooperation is key to most definitions of social capital (Putnam, 2000) and trust (Fukuyama, 1995), both of which have been explored as possible drivers of income differences across countries (*inter alia*, Knack and Keefer (1997), Francois and Zabojnik (2005), Temple and Johnson (1998), and especially Landes (2000); though see Miguel, Gertler, and Levine (2005) for contrary evidence). Further, a Google search yields 22,000 hits for ["economic development" and "prisoner's dilemma"] and 26,000 hits for ["economic growth" and "prisoner's dilemma"]. Clearly, economists are used to thinking of agents within countries as playing a prisoner's dilemma, and are used to thinking that difficulty cooperating is likely to hurt national economic performance. The evidence presented here provides one explanation for why some countries are better at resolving this dilemma: Because their populations are, on average, more intelligent.

I am only aware of one paper that explicitly looks at whether smarter groups are more cooperative in a repeated prisoner's dilemma: Segal and Hershberger (1999).[2, 3] They report how pairs of identical and fraternal twins play a repeated prisoner's dilemma; and unlike any other study in the literature, they gave the players an IQ test beforehand. Their finding corroborates the results reported herein: "[P]airs scoring higher in IQ were more likely to be mutually cooperative….(p.43)." The correlation between the pair's average IQ and joint cooperation was 0.31, significant at conventional levels. However, since all of Segal and Hershberger's games involved two genetically related individuals *knowingly* playing with each other, it would be difficult to generalize their results.

The present paper instead surveys dozens of repeated prisoner's dilemma (RPD) experiments where undergraduates from the same school are randomly matched with one another. This would appear to provide a closer approximation to neighborhoods, firms, or nations, where individuals are randomly placed in situations where they must choose whether or not to cooperate. Thus, there is greater reason to believe that these results could generalize.

The paper proceeds as follows: Section II summarizes the data and Sections III and IV report results and robustness tests. Section V places these results within the broader academic literature and concludes.

---

[2] Millet and Dewitte (2006) should also be mentioned. They run a complicated public goods experiment. In a one-shot setting, they find that higher-IQ players (whether measured through an IQ test or through the speed with which students press lighted buttons) are much more likely to contribute high amounts. Since public goods experiments are simply multi-person prisoner's dilemmas, their study provides some evidence that smarter individuals cooperate more often in PD-like settings.

[3] The one academic literature to investigate group cognitive ability through numerous experiments is the management literature on the determinants of team productivity. It is well-summarized by Devine and Phillips (2001) who find that a team's average IQ has slightly more predictive power than either the maximum or minimum IQ of individual team members (correlation=0.29 between team IQ and team productivity). But while this literature is valuable for showing that IQ tests are useful predictors of team productivity—and thus of great practical value for business managers—it does not address the question likely of greatest interest to economists: Which specific channels actually drive the higher productivity of high-IQ groups? The results presented here point to the ability to escape prisoner's dilemmas as being one such channel.

II. Data

With the aid of a research assistant, data were collected from 36 studies that reported results from repeated prisoner's dilemma experiments. All 36 studies used large numbers of undergraduates at colleges and universities within the U.S. as test subjects. Two-thirds of the studies involved male and female students mixed together; the remainder were evenly divided between all-male and all-female studies. Studying the gender breakdown of any link between IQ and cooperation would be an important task for future work.

In a few cases, multiple "studies" were published in one paper; a typical example is Lave (1962) which reported separate results from studies conducted at Reed College and MIT. The earliest of these 36 studies was in 1959 and the last in 2003, and only data on these 36 schools are used in the statistics reported below. All but three of the studies were published between 1959 and 1977; the remaining three were performed in 1993, 1999 and 2003.[4]

Of course, each of these studies was designed to test a particular hypothesis about RPD behavior, and studies differed in their structure and their payoffs. So a natural question is whether I can legitimately combine such different studies into one database. On this question, three points should be noted: First, as long as the modifications to the RPD are uncorrelated with the average test score at a given school, then I will have only introduced noise, not bias. Second, one might expect *a priori* for faculty at the strongest schools to create tests that were more difficult, testing for more sophisticated forms of cooperation. This would tend to bias my estimates downward from the true estimate.[5] Third, I control for a number of study-specific

[4] No studies with confederates were included in the database—both players are actual undergraduate students, (presumably) attempting to play the game as well as possible. The scarcity of post-1980 prisoner's dilemma experiments is noted by Camerer and Weigelt (1988, p. 2).

[5] The data provide mixed support for this hypothesis. Schools with higher SAT and ACT 2006 scores (defined below) held experiments that were more likely to play RPD's with actual cash (r=.37 between ACT 2006 and the use of money) and were more likely to play more repeated trials (r = .19 between ACT 2006 and number of trials). But at the same time, the correlations between ACT 2006 and (respectively) whether students are separated from

factors in robustness tests below, and find that doing so doesn't impact the test score/cooperation relationship.

Data were collected on the percentage of the time that both players cooperated on average (mean 43%, std. dev. 15%). The cooperation measure is the key variable in the paper: it is the fraction of the time that *both* players chose the "cooperate" option during a given round of the RPD. In most cases, this number, the average rate of cooperation, was precisely reported in the study; in six cases, the author provided charts that reported the trial number on the X-axis and the rate of cooperation on the Y-axis. In the six cases when a precise number for the average rate of cooperation was not reported, the approximate rate of cooperation was estimated as carefully as possible from the charts in the paper (results do not substantially change when the six are omitted). Certainly, some measurement error is inevitable when inferring the average rate of cooperation from a graph, but as long as the errors are independently distributed across studies, this will have no impact on the expected value of the coefficient estimates. In other words, inferring the rate of cooperation from graphs introduces, once again, noise not bias.

This paper uses school-level average SAT and ACT scores as proxies for the intelligence of the students. But is this choice justified? Recent research suggests the answer is yes. Frey and Detterman (2003) found that the correlation between an individual student's SAT score and her level of general intelligence (extracted as the first principal component of the student's Armed Forces Qualifying Test score) to be quite large. In two different specifications, they found correlations between SAT scores and intelligence of 0.72 and 0.86 after routine corrections. They found that for the post-1996 SAT test, a person's IQ could be estimated by the following formula:

---

each other or whether students make sequential rather than simultaneous decisions are .26 and .16. The correlations have the same signs if SAT 2006 scores are used.

$$0.096*\text{SAT Math} + 0.003*\text{SAT Verbal} + 50.241 = \text{IQ}$$

If a person's math and verbal scores rise by 50 points each, this would imply that a 100 point increase in the SAT score would equal a 5 point increase in IQ (about 1/3 of a standard deviation within the U.S.). With appropriate caution, I use this equation to interpret the results below; but in a reply to a criticism of the above paper, Frey and Detterman (2005) emphasize their main point: "[W]e think that focusing on the equations might cause readers to overlook the essential finding: that the SAT is, essentially, an intelligence test."

The SAT and ACT score data used herein were collected from a variety of online sources for the 2006 scores—primarily from the Princeton Review's website—and from the American Association of College and Universities for the 1966 and 1970 entering classes (1968, 1973). All data and citations are available at my website in Excel format. I use only aggregate scores rather than math and verbal subtests, since the subtest estimates reinforced the main results and shed little additional light. The ACT data for 1966 and 1970 were so scanty (n=10 and 11, respectively) that these results are not reported.

SAT and ACT scores are more widely available for the year 2006 than for 1966 and 1970. The informational gain from having this larger 2006 sample size may well be greater than the cost of using estimates that are separated by decades from the time of the study. The key cost of using such temporally distant estimates is that a school's student quality could change over the decades, which would introduce noise into the econometric estimates.[6] However, as I note below, test scores across the decades have a correlation ranging between 0.7 and 0.8, so this is unlikely to be a major cause for concern. In any case, I report results using both the early and the late test scores. Summary statistics are reported in Table 1; summary data on a dummy variable

[6] I thank an anonymous referee for emphasizing this point.

for whether the school is private (1) or public (0) is also reported. All data, including bibliographic information for each study, are available upon request in Excel format.

III. Results

Tables 2 and 3 contain the main results. Appendix 1 reports nonparametric results—including kernel estimates, nearest neighbor estimates, and rank correlations—that tell much the same story. Table 2, the correlation matrix, demonstrates that higher rates of cooperation tend to occur at schools with more intelligent students, regardless of which test one uses to measure intelligence. The lowest correlation between rates of cooperation and a school's average test score is 0.36, while the highest is 0.67. Thus, a substantial fraction of the variance in the rates of cooperation in an RPD can be predicted just by knowing the average SAT score at a given school. As Figures 1 through 4 indicate, these results are not driven by one or two outliers.

Additionally, if one glances at the correlations between the various test scores themselves, one sees that the values are always greater than 0.7, implying that average test scores at the same school have, for the most part, remained stable over the decades. This means that getting the precise test score for the precise year in which each study was conducted is unlikely to change the major results.

Indeed, the fact that these test scores are somewhat noisy measures of the true average intelligence of the students in a given study introduces a classic errors-in-variables problem, implying that the coefficients are biased downward. Therefore, one can reasonably expect the true relationship between intelligence and cooperation to be even stronger than reported here.

The regression results in Table 3 show quantitatively how important higher intelligence is likely to be. Regardless of the SAT measure used, a 100 point increase in the average SAT score at a school is associated with between 4.6% and 8% more cooperation. All results are significant at the 5% level. Even if the true coefficient value is the lowest of these estimates, then cognitive ability appears to be an important predictor of cooperation.

So how large are these effects? Using our weakest results—those from the 2006 SAT regression—one sees that moving from "typical" American universities in the database such as Kent State and San Diego State (with SAT scores around 1000) to elite schools like Pomona College and MIT (with scores around 1450) implies a rise in cooperation from around 30% to around 51%, a 21% increase. Thus, substantially more cooperation is likely to occur in RPD games played at the best schools. It indeed appears that smarter groups are better at cooperating in the RPD environment.

With appropriate caution, one can use the Frey and Detterman equation to convert SAT scores into IQ points to give a back-of-the-envelope estimate of the relationship between IQ and cooperation. If a school's math and verbal scores each increased by 50 points, then a 5 point increase in average IQ would predict a 5% increase in cooperation—a simple one-for-one relationship between IQ points and the rate of cooperation. Since Lynn and Vanhanen document a 38 IQ point gap between countries in the 5^{th} and 95^{th} IQ percentiles, one can reasonably expect sizable increases in cooperative behavior as one moves from low-average-IQ to high-average-IQ countries.

IV. Robustness tests

Changing the specifications so that rates of cooperation were regressed on SAT scores that were closest in time to the relevant study had no substantial impact on the results. This result is unsurprising given the high correlation of test scores at the same school across the decades. In addition, the timing of the experiments themselves do not seem to matter: Omitting the earliest or latest 10% of the experiments (e.g., pre-1962 or post-1977) had no substantial impact on the results, and inclusion of a linear or quadratic time trend likewise had no substantial impact on the results.

Further, for each study information was collected on the number of rounds in each study's RPD, whether actual money was involved, and whether the interaction was face-to-face versus across some kind of screen. When all were included simultaneously in broader regression specifications simultaneously, none of these four features had any impact on the results; SAT 2006 was significant at the 2% level, the rest at 1%. Results are reported in Table 4. Only the rounds variable was statistically significant in any specification[7], but the inclusion of these three variables impacted the statistical or quantitative significance of the test score variables.

An additional robustness test checked to see whether the effect is simply due to students at private schools being more cooperative, perhaps due to their smaller class sizes and smaller campus populations, or perhaps due to cultural differences associated with the higher average socioeconomic status of private school students.

In these regressions, in addition to the money, interaction, and trial number variables a dummy variable was included for private schools: a 0.5 was coded for two studies that included a mix of public and private school students (the main regressions likewise used a simple average of the test scores for these two mixed-school studies; omission of these studies had no substantial impact on the results). Results are reported for Table 5. Controlling for test score, the private school dummy was never significant at conventional levels—its p-values ranged from 22% to

[7] As in Sally (1995), longer games had *less* cooperation.

95% depending on the test score measure included, and the dummy's value implied that private school students cooperated from -0.5 to +8% more often.[8]

The cognitive ability coefficient estimates were little changed. The SAT 1966, SAT 1970, and ACT 2006 results, the results were statistically significant at the 5% level; and for SAT 2006 results, at the 10% level. Out of all the specifications reported here, this SAT is the only one that falls below the 5% level. Since SAT 2006 performs well across many specifications while the private school dummy always performs poorly, one may plausibly take this single estimate as only weak evidence against the robustness of cognitive ability.[9]

Finally, log and semilog specifications had no substantial impact on the results, though the linear specification used in the previous section was slightly more statistically significant in general. The semi-elasticity of cooperation with respect to SAT score (omitting other controls, which, as noted, never quantitatively impact the coefficient estimate) was estimated to equal .094 per 100 year 2006 SAT points, and equal to 0.17 and 0.16 for the 1966 and 1970 SAT scores, respectively. The ACT 2006 semi-elasticity of cooperation was 4.6 percent more cooperation per ACT point. The 2006 SAT results were significant at the 10% level, while the 1966 and 1970 SAT estimates were well under the 5% value; the ACT estimate was significant at the 1% level. One might summarize these semi-elasticity results by saying that a 400-point increase in a school's SAT score (or a 20-point increase in a group's average IQ) is associated with roughly 50% more cooperation. Overall, these robustness tests appear to confirm the result from the

[8] Perhaps paying with cash matters more at public schools, where poorer students might value cash more; accordingly, I also ran specifications that included a private school/money interaction dummy, which was never statistically significant.

[9] One may be concerned that with such small sample sizes, adding so many controls is statistically unwarranted; but even when only one, two or three controls are added at a time, the cognitive ability measures still perform well. SAT 2006 always is always the weakest, but that coefficient's magnitude changes little across specifications.

preceding section: Groups with high cognitive abilities appear to cooperate more often when playing a repeated prisoner's dilemma.

By way of comparison, consider the widely-cited meta-study of prisoner's dilemma experiments, Sally (1995). His survey of 89 repeated game experiments used data on 22 explanatory variables in addition to a cooperation measure. Sally included many multi-player games and other features that disqualified them from inclusion in this paper. In his regressions involving only repeated prisoner's dilemmas, six variables were statistically significant at the 5% level or better across most specifications. Sally found that if students played for money or played in smaller groups, then they cooperated more often (while money is always positive in the robustness tests reported above, it is never significant in this sample). The other four factors mentioned by Sally relate to the ability to see and communicate with other players, and how the experimenter described the game to the players.

It is difficult to make a direct comparison between Sally's results and those presented here—his sample size is more than twice as large (possibly making it easer to get statistical significance), but he also includes many explanatory variables in his regressions (possibly making it harder to get statistical significance). With these caveats in mind, it should nevertheless be noted that no single variable from Sally's (1995) repeated prisoner's dilemma regressions is significant at the 1% level in the majority of specifications. The only variables *ever* significant at the 1% level are whether the game is played for actual money and how often players are allowed to communicate; but such results are sensitive to the specification.[10]

The results presented here indicate that cognitive ability is often significant at the 1% level across specifications and almost always significant at the 5% level, with slightly weaker

[10] Technical exceptions: Some communications-related variables are included in only one or two of Sally's four specifications, but are always robust at the 1% level.

results occurring only when 2006 SAT scores are used as the ability measure. One might conservatively conclude that cognitive ability appears as robust as the communication, money, and group-size variables reported in Sally (1995). It appears that the average intelligence of game participants should be considered among the most robust factors driving cooperation in a repeated prisoner's dilemma.

V. Discussion and Conclusion

The game theory literature—both behavioral and theoretical—has devoted enormous attention to the puzzle of cooperation (*inter alia*, empirical results reviewed in Camerer (2003) and Dawes and Thaler (1998) and the theoretical literature growing out of Kreps, et al., (1982)), but has spent little time on the question of whether differences in average intelligence are important drivers of cooperative behavior. The meta-experimental results summarized here indicate that this has been a glaring omission.

Now that there is prima facie evidence that smarter groups are indeed more cooperative, we can turn our attention to the question of why this may be so. Axelrod's (1984) classic work may hold the answer. Of the five pieces of advice he offers (c. 7) on how to promote cooperation in an RPD, the first is "Enlarge the shadow of the future," and the last is "Improve recognition abilities." I provide evidence—from other studies, not my own—that intelligent individuals are more likely to possess these traits of patience and perceptiveness that Axelrod argues are useful in creating cooperation. Thus, there are strong reasons based on external evidence for believing that smarter groups will be more cooperative.

Patience is surely one way of enlarging the shadow of the future—and recent research (Warner and Pleeter, 2001; Fredrick, 2005; Benjamin, Brown, and Shapiro, 2006) has shown that persons with higher cognitive ability tend to be more patient and less impulsive. Warner and Pleeter show this in the context of military personnel choosing a lump-sum severance payment versus an annuity; personnel with higher test scores had a lower implied discount rate. Fredrick and Benjamin et al. both showed that smarter individuals (as measured by cognitive tests) were more patient in experimental settings, and Fredrick in particular refers to a number of other studies from the psychology and economics literatures illustrating the link between cognitive ability and patience.

In the psychology literature, the raw correlation between cognitive ability and a preference for immediate rewards is widely recognized (inter alia, Jensen (1998) *passim*); this IQ-impulsivity link is analyzed in particularly thorough manner in De Wit et al. (2006). Using a large sample of middle-aged Americans, they found that IQ was a robust predictor of impulsive experimental behavior even after controlling for socioeconomic status, race, gender, and survey measures of impulsiveness. All told, research by economists and psychologists alike indicate that smarter groups are likely to be patient groups.

Additionally, the ability to recognize patterns in an ambiguous situation is a strong positive correlate of intelligence (Deary, 2001; Jensen, 1998, pps. 34-38; Gottfredson, 1997, 2004). As Gottfredson (1997, p. 94) notes,

> "For practical purposes, g [the general intelligence factor] is the ability to deal with cognitive complexity--in particular, with complex information processing."

In an RPD, players are sending difficult-to-interpret signals about one's type. Indeed, there is a small game theoretic literature that "hinge[s]..on the ability of 'cooperative types' to use some type of 'secret hand-shake' to recognize players of the same type" (Anderlini and

Sabourian, 1995, p. 1341). A high level of intelligence is likely to prove useful in accurately recognizing these ambiguous signals.

Axelrod emphasizes the need to perceive the identity of one's opponents as well as to perceive what truly counts as a move of "cooperate" or "defect." But in both experimental settings and real-world prisoner's dilemmas, the most important kind of "perceptiveness" offered by intelligent players may simply be that they have a better chance of accurately and quickly perceiving the rules of the game, as well as remembering the payoffs and the game's recent history. One well-established fact about higher-IQ individuals is that they tend to have larger working memories (*inter alia,* Conway et al., 2002), thus giving them a better chance of remembering precisely what kind of game they're playing and what has happened in the course of the game. All told, there is abundant evidence that smarter groups are generally more patient and more perceptive---traits that Axelrod recommends as keys to cooperative behavior.

If intelligence differences are indeed key drivers of differences in cooperation, this makes it all the more important for economists to explore how to raise the intelligence of the world's low-IQ populations. Jones and Schneider (2006a) survey the literature on how intelligence can be improved through better nutrition, healthier environments, and better education in the world's poorest countries. One can only hope that future research is able to uncover promising, practical methods for raising the average intelligence of the world's poorest countries. By so doing, researchers will, it appears, help to create a more cooperative world.

Bibliography

Aghion, Phillipe, Diego Comin, and Peter Howitt (2006). "When does Domestic Saving Matter for Economic Growth?" working paper, available at http://www.nyu.edu/econ/user/comin/saving.pdf

Anderlini, Luca and Hamid Sabourian (1995). "Cooperation and Effective Computability," *Econometrica*, 63: 6 (Nov), 1337-1369.

Axelrod, Robert (1984). *The Evolution of Cooperation*. New York: Basic Books.

American Council on Education (1968). *American Colleges and Universities.* Washington, DC: American Council on Education.

________________, (1973). *American Colleges and Universities.* Washington, DC.: American Council on Education.

Behrman, J., Alderman, H. & Hoddinott, J. (2004). Copenhagen Consensus – Challenges and Opportunities: Hunger and Malnutrition, available at http://www.copenhagenconsensus.com/Default.asp?ID=223

Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro (2006). "Who is Behavioral? Cognitive Ability and Anomalous Preferences." Working Paper, University of Chicago.

Camerer, Colin (2003). *Behavioral Game Theory*. Princeton, NJ: Princeton University Press.

Camerer, Colin and Keith Weigelt (1988). "Experimental Tests of a Sequential Equilibrium Reputation Model," *Econometrica*, 56:1 (January), 1-36.

Caplan, Bryan (2007). *The Myth of the Rational Voter,* Princeton, NJ: Princeton University Press.

Caplan, Bryan and Stephen C. Miller (2007). "Economic Beliefs, Intelligence, and Ability Bias: Evidence from the General Social Survey." working paper, available at http://www.gmu.edu/departments/economics/bcaplan/iqbeliefej.doc

Conway, Andrew R.A., Nelson Cowan, Michael F. Bunting, David J. Therriault, Scott R.B. Minkoff, (2002), "A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence," *Intelligence,* 30, 163-183.

Dawes, Robyn and Richard Thaler (1988). "Cooperation." *Journal of Economic Perspectives* 2:3, 187-197.

Deary, I. J. (2001). *Intelligence: A very short introduction*. NY: Oxford University Press.

Devine, Dennis J. and Jennifer L. Phillips (2001). "Do Smarter Teams Do Better: A Meta-Analysis of Cognitive Ability and Team Performance," *Small Group Research*, 32(5): 507-532.

Harriet de Wit, Janine D. Flory, Ashley Acheson, Michael McCloskey, and Stephen B. Manuck. (2006). "IQ and nonplanning impulsivity are independently associated with delay discounting in middle-aged adults," *Personality and Individual Differences*, 42:1, 111-121.

Francois, Patrick and Jan Zabojnik. (2005). "Trust, Social Capital, and Economic Development." *Journal of the European Economic Association,* 3:1, 51-94.

Fredrick, Shane (2005). "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives*. 19:4, 24-42.

Frey, Meredith C., & Detterman, Douglas K. (2004). "Scholastic assessment or g? The relationship between the Scholastic Assessment Test and general cognitive ability." *Psychological Science,* 15, 373–378.

Frey, Meredith C. & Detterman, Douglas K. (2005) “Regression Basics.” *Psychological Science* 16:9, 747.

Fukuyama, Francis (1995). *Trust*. New York: Free Press.

Glaeser, Edward L., Rafael La Porta, Florencio Lopez-de-Silane, Andrei Shleifer (2004). “Do Institutions Cause Growth?” NBER Working Paper #10568, June.

Gottfredson, Linda S. (1997). “Why g matters: The complexity of everyday life,” *Intelligence,* 24, 79-132.

Gottfredson, Linda S. (2006). “Social consequences of group differences in cognitive ability (Consequencias sociais das diferencas de grupo em habilidade cognitiva)”, in C. E. Flores-Mendoza & R. Colom (Eds.), *Introducau a psicologia das diferencas individuais* (pp. 433-456). Porto Allegre, Brazil: ArtMed Publishers.

Jensen, A. R. (1998). *The g-factor: The science of mental ability*. Westport, CT: Praeger.

Jones, Garett and W. Joel Schneider (2006a). “Intelligence, Human Capital, and Economic Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach,” *Journal of Economic Growth*, 11, 71-93.

Jones, Garett and W. Joel Schneider (2006b). “IQ and Productivity: Evidence from Immigrant Earnings.” Working paper, Southern Illinois University Edwardsville.

Knack, Stephen and Philip Keefer (1997), “Does Social Capital Have An Economic Payoff? A Cross-Country Investigation,” *Quarterly Journal of Economics,* 112:4, 1251-1288.

Kreps, David M., Paul R. Milgrom, John Roberts, and Robert J. Wilson. (1982). “Rational cooperation in the finitely repeated prisoners' dilemma,” *Journal of Economic Theory*, 27, 245-252.

Landes, David (2000). “Culture makes almost all the difference,” in L. Harrison and S. Huntington, eds., *Culture Matters*, New York: Basic Books.

Lave, Lester B. (1962) “An Empirical Approach to the Prisoner's Dilemma,” *Quarterly Journal of Economics*, 76, 424-436

Lave, Lester B. (1965) “Factors Affecting Cooperation in the Prisoner's Dilemma,” *Behavioral Science*, 10, 26-38.

Lynn, Richard and Tatu Vanhanen. (2002). *IQ and the Wealth of Nations.* Westport, CT: Praeger Publishers.

Lynn, Richard and Tatu Vanhanen. (2006). *IQ and Global Inequality.* Augusta, GA: Washington Summit Publishers.

Miguel, Edward, Paul Gertler, David I. Levine. (2005) “Does Social Capital Promote Industrialization? Evidence from a Rapid Industrializer.” *Review of Economics and Statistics* 87:4, 754-762.

Miller, Gary (1993). *Managerial Dilemmas.* Cambridge: Cambridge UP.

Millet, Kobe, and Siegfried Dewitte (2006). “Altruistic behavior as a costly signal of general intelligence,” *Journal of Research in Personality.* 41:2, 316-326.

Putnam, Robert D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster.

Ram, Rati (2006). “IQ and Economic Growth: Further Augmentation of the Mankiw-Romer-Weil Model,” *Economics Letters,* 94:1, 7-11.

Sally, David (1995). “Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992,” *Rationality and Society*, 7:1(Jan): 58-92.

Segal, Nancy L. and Scott M. Hershberger (1999). "Cooperation and Competition Between Twins: Findings from a Prisoner's Dilemma Game," *Evolution and Human Behavior*, 20:29–51.

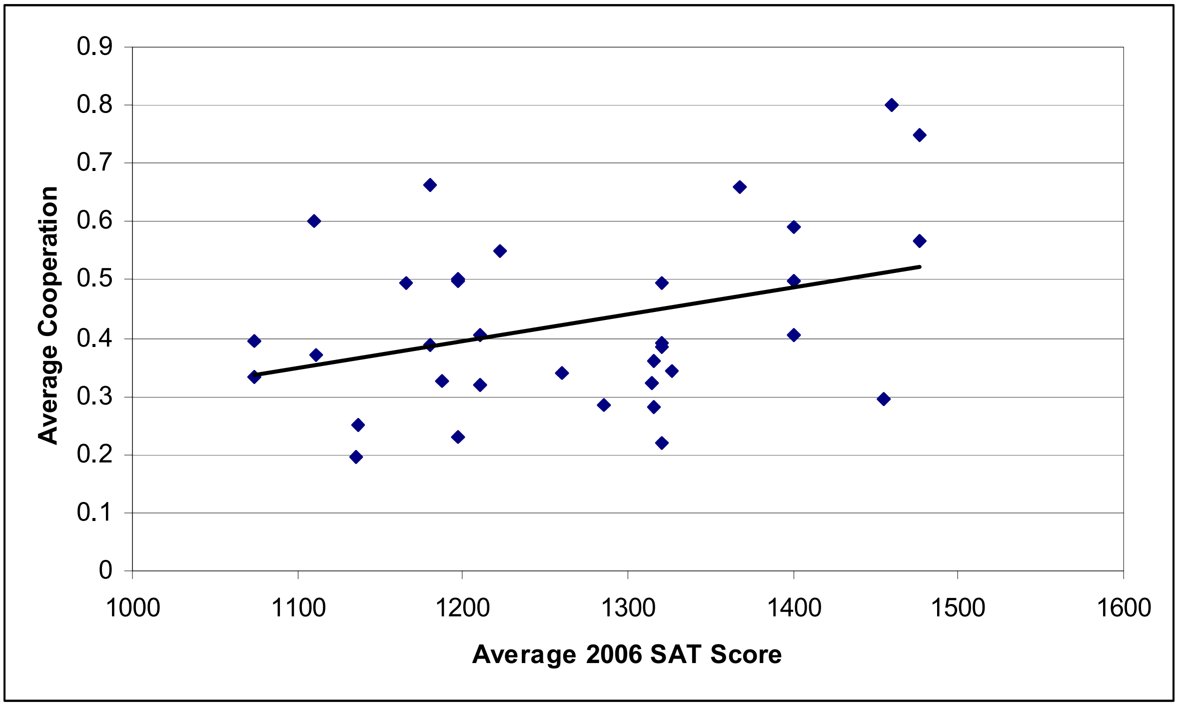
Temple, Jonathan and Paul A. Johnson (1998). "Social Capability and Economic Growth," *Quarterly Journal of Economics* 113:3, 965-990

Warner, John T. and Saul Pleeter (2001). "The Personal Discount Rate: Evidence from Military Downsizing Programs." *American Economic Review*, 91:1, 33-53.

Weede, Erich (2003)

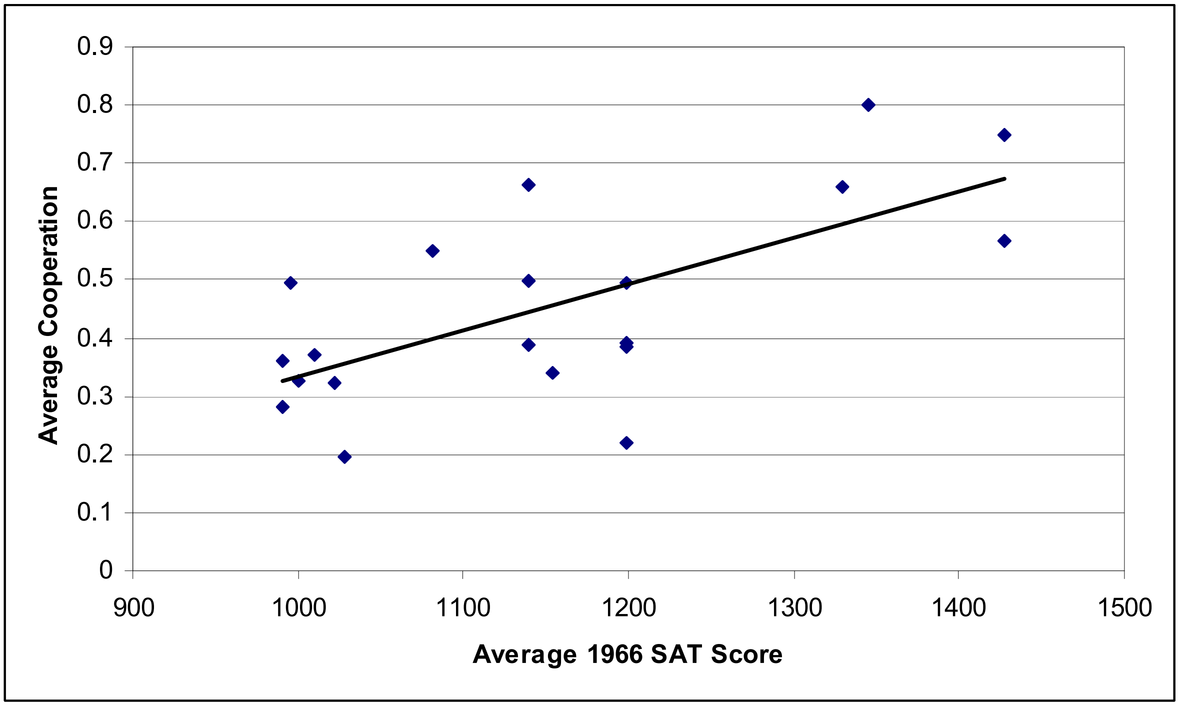
Weede, Erich and Sebastian Kämpf (2002). "The Impact of Intelligence and Institutional Improvements on Economic Growth," *Kyklos* 55:3, 361–380.

Figure 1
2006 SAT Scores and RPD Cooperation



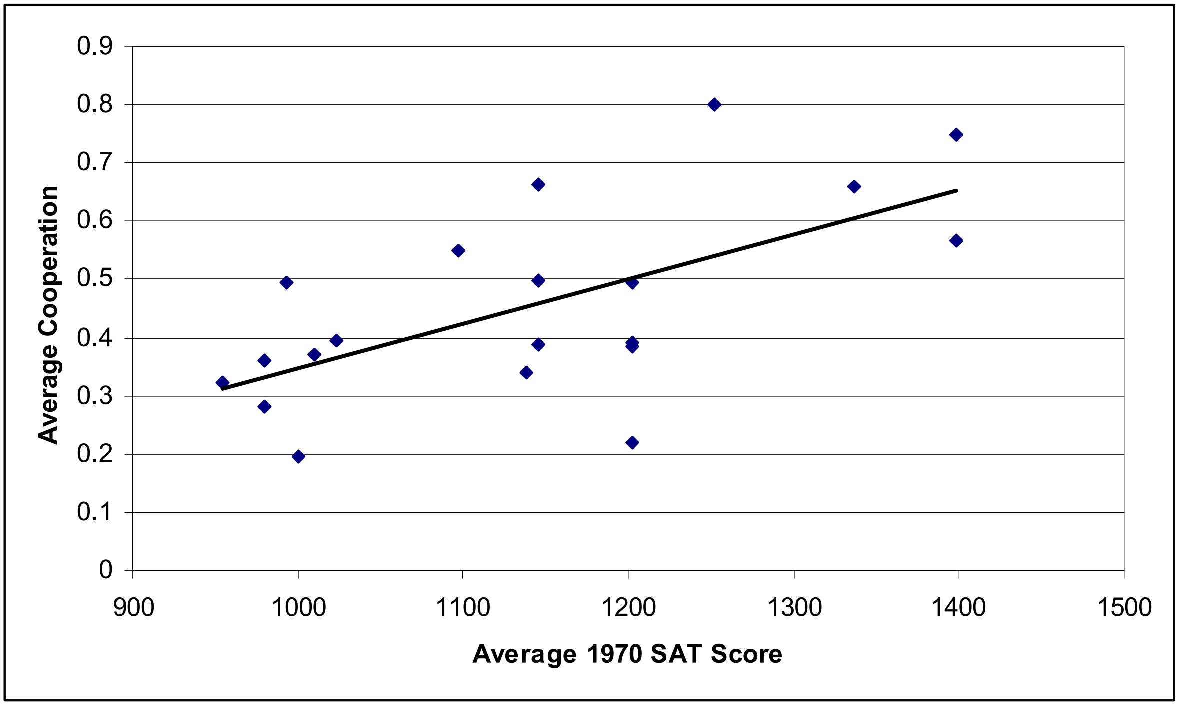
Notes: Each point represents one repeated prisoner's dilemma (RPD) study conducted between 1959 and 2003. The X-axis reports the average SAT score at the university where the study was conducted, and the Y-axis reports the percentage of the time that both players chose to cooperate rather than defect. The slope of the OLS trendline reflects 4.6% more cooperation for every 100 SAT points. $R^2 = 13\%$.

Figure 2
1966 SAT Scores and RPD Cooperation



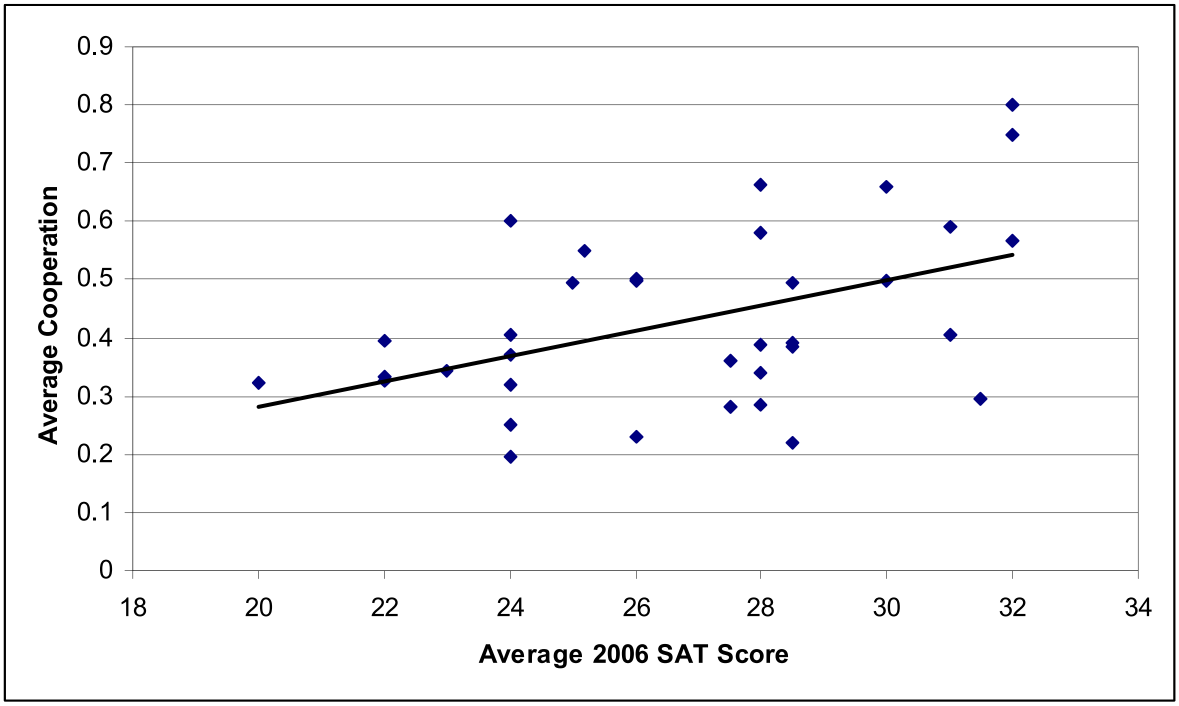
Notes: Each point represents one repeated prisoner's dilemma (RPD) study conducted between 1959 and 2003. The X-axis reports the average SAT score at the university where the study was conducted, and the Y-axis reports the percentage of the time that both players chose to cooperate rather than defect. The slope of the OLS trendline reflects 8% more cooperation for every 100 SAT points. $R^2 = 45\%$.

Figure 3
1970 SAT Scores and RPD Cooperation



Notes: Each point represents one repeated prisoner's dilemma (RPD) study conducted between 1959 and 2003. The X-axis reports the average SAT score at the university where the study was conducted, and the Y-axis reports the percentage of the time that both players chose to cooperate rather than defect. The slope of the OLS trendline reflects 7.7% more cooperation for every 100 SAT points. R^2 = 40%. (n.b.: If all scores above 1200 are omitted, R^2 = 31%)

Figure 4
2006 ACT Scores and RPD Cooperation



Notes: Each point represents one repeated prisoner's dilemma (RPD) study conducted between 1959 and 2003. The X-axis reports the average ACT score at the university where the study was conducted, and the Y-axis reports the percentage of the time that both players chose to cooperate rather than defect. The slope of the OLS trendline reflects 2.2% more cooperation for every additional ACT point. R^2 = 22%.

Table 1
Summary Statistics

	Cooperation	ACT 2006	SAT 1966	SAT 1970	SAT 2006	Private School
Mean	43%	27	1151	1140	1268	29%
Median	39%	28	1140	1145	1272	0%
Maximum	80%	32	1428	1398	1477	1
Minimum	19%	20	991	955	1074	0
Std. Dev.	15%	3.3	143	138	118	44%
n	35	35	20	20	34	35

Table 2
Correlation Matrix

	COOP	SAT 1966	SAT 1970	SAT 2006	ACT 2006	PRIVATE
COOP	1.00	0.67	0.63	0.36	0.47	0.36
SAT 1966		1.00	0.98	0.76	0.82	0.48
SAT 1970			1.00	0.70	0.83	0.41
SAT 2006				1.00	0.79	0.68
ACT 2006					1.00	0.65
PRIVATE						1.00

Table 3
Regression Results
Dependent variable: Rate of Cooperation

	Units: 100 SAT points			
	β	Std. Err.	p-val.	N
SAT 1966	8.0%	2.1%	0.1%	20
SAT 1970	7.7%	2.2%	0.3%	20
SAT 2006	4.6%	2.1%	3.9%	34
	Units: 1 ACT point			
ACT2006	2.2%	0.7%	0.5%	35

Notes: In the top section, β represents the expected impact of a 100-point increase in a school's average SAT score on the rate of cooperation in a repeated prisoner's dilemma experiment. The bottom section provides similar results for a one-point rise in the school's average ACT score.

Table 4
Robustness Tests: Type of Experiment
Dependent variable: Rate of Cooperation

	β	Std. Err.	p-val.	N
Units: 100 SAT points				
SAT 1966	7.2%	2.1%	0.5%	18
SAT 1970	7.3%	2.2%	0.5%	18
SAT 2006	5.6%	2.2%	1.6%	30
Units: 1 ACT point				
ACT2006	2.5%	0.8%	0.4%	31

Notes: In the top section, β represents the expected impact of a 100-point increase in a school's average SAT score on the rate of cooperation in a repeated prisoner's dilemma experiment. The bottom section provides similar results for a one-point rise in the school's average ACT score. Controls include dummies for whether the participants personally interacted, whether money was involved, as well as the number of trials.

Table 5
Robustness Tests: Adding Private School Dummy
Dependent variable: Rate of Cooperation

	B	Std. Err.	p-val.	N
	Units: 100 SAT points			
SAT 1966	6.2%	2.3%	2.0%	18
Private	6.4%	6.3%	33.1%	
SAT 1970	6.2%	2.3%	1.9%	18
Private	8.0%	6.2%	21.8%	
SAT 2006	5.7%	3.2%	8.5%	30
Private	5.6%	8.4%	51.1%	
	Units: 1 ACT point			
ACT2006	2.6%	1.1%	2.2%	31
Private	-.5%	0.8%	95.0%	

Notes: In the top section, β represents *either* the expected impact of a 100-point increase in a school's average SAT score on the rate of cooperation *or* the impact of holding the experiment at a private school, *ceteris paribus*. The bottom section provides similar results for a one-point rise in the school's average ACT score. Controls include all controls from Table 4 and a dummy for whether the study was conducted at a private school.

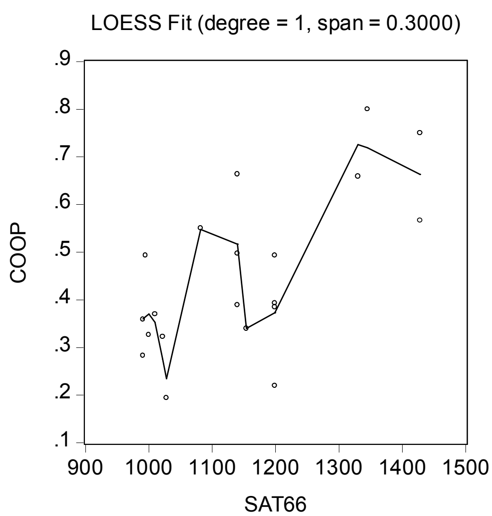
Appendix 1
Nonparametric Estimations

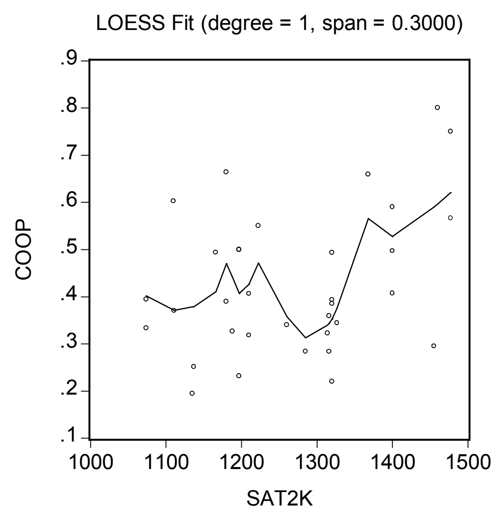
To address the possible question of nonlinearities in the data, I provide some basic results from nonparametric estimations. I provide estimates below for the test score measures with the strongest (SAT 1966) and the weakest (SAT 2006) correlation with average cooperation. I use a nearest neighbor method and kernel method; in order to avoid the perception of data-mining, I only report results from the default settings used in EViews 5.1.

The SAT 2006 results are broadly consistent with the hypothesis that only the most intelligent students (SAT > 1350) are highly cooperative—in the vast middle, all students are about alike. That would imply that whether or not a society has cooperative leaders could turn largely on the question of whether that society is able to select *highly* intelligent leaders.

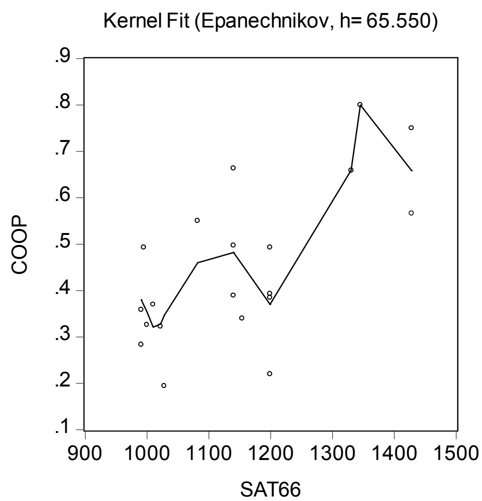
However, the SAT 1966 measures—which are typically closer in time to the actual experiments—provide less support for this hypothesis; this relationship is broadly linear. In either case, however, one would predict highly intelligent groups to be more cooperative than the less intelligent groups.

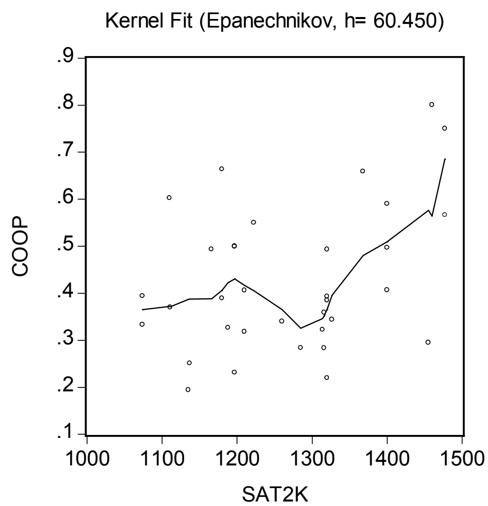
<u>Nearest Neighbor:</u>





<u>Kernel:</u>





As a final non-parametric measure, I report the Spearman rank correlation. All but the 2006 SAT score are statistically significant at conventional levels.

N.B.: If we omit the single observation of the University of Miami—a fast-rising school whose 2006 SAT score was over 300 points higher than when its cooperation study was published in the late 1960's—the 2006 SAT rank correlation rises to 0.30 with a t-statistic of 1.9.

Spearman Rank Correlation with Average Cooperation

	ACT 2006	SAT 1966	SAT 1970	SAT 2006
Rank Corr.	0.41	0.53	0.56	0.26
t-stat	2.6	2.6	2.9	1.6