Lecture Note: Group Differences in Economic Outcomes: Market and Non-Market Factors

David Autor
MIT 14.663 Spring 2009

May 7, 2009

1 Introduction: Group Differences in Economic Outcomes

By any measure, the disparities in economic outcomes between Black and White (non-Hispanic) Americans are enormous. Measured in terms of average education, college completion, earnings, residual earnings, labor force participation, marriage, imprisonment, life expectancy, and death by disease or violence, there is a huge gulf in the average outcomes of Blacks and Whites. The 2004 paper by Western and Petit, focusing on mass incarceration, paints a particularly stark portrait. Among Black, non-Hispanic males born between 1965 and 1969, W&P estimate that 22.4% were incarcerated in prison at some point prior to 1999, and this percentage rises to 31.9 percent among men who have not attended college. As W&P document, these disparities in incarceration greatly increased in the 1980s and 1990s as Black relative earnings stagnated. In 1979, 17.1 of black male high school dropouts who did not hold a GED ages 30 to 34 had spent some time in prison. By 1999, this share had risen to 61.8 percent.

Although the Black-White disparity in life outcomes is a uniquely American social problem. (for example, it doesn't afflict Africans in America, only African-Americans), study of this topic has generated a wealth of important economic research. A controversial 1994 book by Richard Herrnstein and Charles Murray, *The Bell Curve*, brought this topic to public prominence. The publication of *The Bell Curve* was itself a watershed event—not primarily because the book was a superb piece of scholarship (it has major flaws that have since been documented), but because it sparked a fire-storm of responses and research, which substantially advanced study of group disparities in economic outcomes. I highly recommend Jim Heckman's feature length review of *Bell Curve* in the *JPE* in 1995 (you don't have to have read *Bell Curve* to profit from reading this article).

This lecture will focus on understand the B-W gap. The classes of explanations are considered are:

1. Pre-market factors that affect human capital acquisition
2. Genetic and environmental differences (these are in some sense a subset of 2)
3. Discrimination reflecting accurate but self-fulfilling expectations

4. Cultural differences

5. Self-fulfilling prophesies ('stereotype threat')—related but not identical to item (3).

There is far more excellent material on all of these topics than I can possibly cover in the time allotted; it would be easy and worthwhile to teach a semester class on this subject. With luck, my lectures will give you some sense of the depth and breadth of the subject matter.

2 Is labor market discrimination the problem? Neal and Johnson (1996)

- This is the rare modern regression study that had a major impact. Why? Because what it give up in cleverness it more than compensates for in substance.
- This work is related to *Bell Curve* (1994), which generate a huge response literature by economists and sociologists in the late-1990s.
- See also Heckman's superb review of this book in the *JPE*.
- Basic question: How much of the B/W earnings gap is explained by differences in skills that are formed prior to market entry? This is a great question because so much of literature is focused on market discrimination. Is this focus misplaced? Neal and Johnson say yes.
- Benchmark test: Look at identically skilled Black and White teens before market entry and then again later in life: What is the initial B/W earnings gap and does it grow over time. Assuming there are no diffs in tastes or costs of skill investment, one could potentially attribute differences to discrimination.
- Since can't do that, use NLSY. Sample 15 to 23 years old, who took AFQT prior to age 18. Regress age effects out of AFQT.
- See Table 1: The basic result. 'Pre-market' skill appears to explain a large part of racial earnings gap for currently employed workers.
- This is the main message of the paper. Now, they ask a bunch of good questions.

2.1 Is the AFQT racially biased?

- Military has done validation studies using objective performance measures. Finds no evidence that tests systematically under-predict performance of blacks.

- See also the excellent chapter by Christopher (AKA, Sandy) Jencks on Testing Bias in the 1998 Jencks-Phillips volume, *The Black-White Test Score Gap.* I quote a few paragraphs to peak your interest:

"Another way to assess content bias is to abandon a priori judgments and simply ask how much the black-white gap varies from one question to the next. The Peabody Picture Vocabulary Test (PPVT), for example, tries to measure the size on an individual's English vocabulary. The tester reads a word aloud in standard English and asks the test taker which of four pictures best approximates the word's meaning. If black and white children spoke relatively distinct languages, as some experts on black English claim, one might expect black and white children with vocabularies of equal size to know somewhat different words. In that case, the PPVT should show much larger black-white gaps for some words than for others. But that is not what psychologists have found. The percentage of white children who know a PPVT word correlates .98 with the percentage of slightly older black children who know the same word. Black children, in other words, learn pretty much the same words as white children. The black children just tend to learn these words somewhat later. Black children never catch up, however, so black adults know fewer words that white adults. These finding suggest that black and white English differ more in their pronunciation and syntax than in their vocabulary. (I return to pronunciation differences below).

A few words are unique to black English. In the early 1970s Robert Williams developed the Black Intelligence Test of Racial Homogeneity (BITCH), which measured knowledge of black street language. As one would expect, black high school students scored higher than white ones on this test. But blacks' scores on the BITCH were negatively correlated with their scores on other tests. Since the number of words unique to black street language is very small relative to the total number of English words, this negative correlation implies that blacks who did well on the BITCH had smaller total vocabularies than blacks who did poorly on the BITCH. Thus when the goal is to measure the total size of an individual's English vocabulary, tests designed by and for whites are probably not very biased against blacks. The main reason why blacks score lower than whites on these tests is probably that black children have fewer opportunities and incentives to expand their vocabulary, not that conventional tests significantly underestimate the number of words that black children know.

A third approach to assessing content bias in intelligence and aptitude tests is to ask whether the black-white gap is larger on tests that measure familiarity with the content of American culture on "nonverbal" tests that do not measure familiarity with a particular culture. Larry Hedges and Amy Nowell have reviewed all the national surveys that have tested high school seniors (see chapter 5). These surveys no longer include nonverbal tests, but in 1965 and 1972 the black-white gap was marginally larger on tests of "nonverbal reasoning" than on tests dealing with vocabulary, reading and mathematics.

The large difference between black and whites on many different tests and items suggests that race must play a much broader and deeper role in depressing black scores than many critics assume. There is nothing distinctively white or middle class about the content of the Block Design test, in which a tester asks a child to reproduce a two-dimensional red and white design using a set of colored blocks. Race must affect performance on this test by affecting black and white children's approach to learning, solving problems, and taking tests, not by affecting their familiarity with the content of the test.

These findings underscore a fact that is easily forgotten. Culture is not merely a body of knowledge and skills. It is also a set of strategies for dealing with the unknown and with tasks that seem difficult. Culture can affect people's willingness to think about unfamiliar questions, their strategies for seeking answers that are not obvious, their motivation to persist in the face of frustration, their confidence that such persistence will be rewarded, and their interest in figuring out what the tester thinks is the right answer. No one knows how large a role cultural differences of this kind play in the black-white test score gap. But we do know that the size of the gap does not depend in any simple, predictable way on the nominal content of the test."

2.2 Do blacks under-invest in skills b/c return is lower?

- This is exceedingly difficult to test.
- In Table 2 and 3, no evidence of differences in return to AFQT. But this is an endogenous comparison – conditional on having attained a given level of skill.
- Moreover, return may have been lower in prior eras, and this could affect beliefs of black parents, who choose how much to invest in kids' skills.
- How would you convincingly test this? What does this imply that we should see for IV estimates of returns to school for blacks versus whites?

2.3 Selection into labor force participation

- See Figure 1. Low scoring blacks are noticeably less likely to participate in labor market.
- How do we deal with this? Two approaches:

1. Median regressions. If nonparticipants have wage offers less than median for group with similar observables (race, test score) and at least half participate, then the median is identified. Bingo. See Table 4. This table suggests that less is explained of the gap once we condition on participation.

   - Smith and Welch method (this is slightly cool):

$$E(w) = LFPR \cdot E\left(w|\text{participate}\right) + (1 - LFPR) \cdot E\left(w|\text{don't participate}\right).$$

   Therefore

$$\frac{E\left(w_i\right)}{E\left(w_j\right)} = \frac{LFPR_i \cdot E\left(w_i|\text{participate}\right) + (1 - LFPR_i) \cdot E\left(w_i|\text{don't participate}\right)}{LFPR_j \cdot E\left(w_j|\text{participate}\right) + (1 - LFPR_j) \cdot E\left(w_j|\text{don't participate}\right)},$$

$$\frac{E\left(w_i\right)}{E\left(w_j\right)} = \left[\frac{E\left(w_i|\text{participate}\right)}{E\left(w_j|\text{participate}\right)}\right] \cdot b,$$

   where b is a bias term equal to

$$b = \left[\frac{\left(1 - k_i\right) LFPR_i + k_i}{\left(1 - k_j\right) LFPR_j + k_j}\right],$$

with

$$k_i = \frac{E\left(w_i | \text{don't participate}\right)}{E\left(w_i | \text{participate}\right)}.$$

- We can use this formula to adjust the observed earnings ratio by a selection factor to get the ratio of the means of the unconditional (offer) wage distributions.
- Notice that OLS wage gap is -0.072 and median gap is -0.134, implying:

$$\left[\frac{E\left(w_i | \text{participate}\right)}{E\left(w_j | \text{participate}\right)}\right] = 0.93$$

and

$$\frac{E\left(w_i\right)}{E\left(w_j\right)} = \left[\frac{E\left(w_i | \text{participate}\right)}{E\left(w_j | \text{participate}\right)}\right] \cdot \left[\frac{\left(1 - k_i\right) LFPR_i + k_i}{\left(1 - k_j\right) LFPR_j + k_j}\right] = 0.875.$$

- Using $k_b = k_w = 0.1$ and noting that $LFPR_b = 0.91$ and $LFPR_w = 0.975$, we have

$$\left[\frac{(1 - 0.1) \cdot 0.91 + 0.1}{(1 - 0.1) \cdot 0.975 + 0.1}\right] = 0.94,$$

and

$$0.93 \cdot 0.94 = 0.874.$$

- So, by making a *quite* conservative assumption on potential earnings of non-employed, we get back the observed ratio. Of course, this calculation does not account for AFQT scores (it's treated as unconditional). If score gap is much larger among non-employed, this suddenly does not seem so conservative.
- Historical perspective: If this paper had been written recently, it would make much greater use of quantile regression.

2.4 Determinants of AFQT scores

- Figures 2 and 3.
- Tables 5 and 6.
- There is a 1.0 SD test score gap. This is very large. (But both groups have gained more than this amount over century due to 'Flynn effect.')

- Adding a bunch of family background, home environment, and school quality covariates reduces this considerably. Suggests that pre-market factors have a lot of potential explanatory power for this gap.

- *Bell Curve* argued these score differences are genetic – inherent ability.

- But score gap is larger in older cohorts, suggesting skills investment (Table A3)

- When quarter-of-birth used as instrument for schooling, significantly increases AFQT (not shown). Again suggests some part of cognitive skill acquired.

3 Understanding differences in ability: Heritability and environment

The weight of the Neal and Johnson paper and the evidence we have already discussed on marketplace discrimination suggests (to your instructor) that labor market discrimination cannot be the entire explanation—perhaps not even the predominant explanation—for the observed disparities in Black-White outcomes. Neal's 2006 *Handbook of Education* chapter makes the argument that the bulk of the differences are due to skill deficits, and these can be traced in large part to environmental differences. We will now evaluate some evidence on the role of environment in skill deficits.

The starting point of this discussion is the 1996 book the *Bell Curve*, which argues strenuously that race differences in IQ are largely the result of genetic differences among races. *Bell Curve* authors Herrnstein and Murray draw heavily on the work of Arthur Jensen, an influential psychologist at UC Berkeley who has been a strong exponent of the view that race differences in IQ are largely hereditary. A chief piece of evidence in support of this view is the fact that intergenerational correlations in IQ among parents and children are very high (on the order of 0.8 or above). Many psychologists take this strong correlation as evidence that IQ is highly heritable (read: genetic) though it should be obvious why this conclusion is suspect.

3.1 Behavioral-genetics framework

The high intergenerational correlation of IQ also poses a puzzle, explored by the 2001 paper by Dickens and Flynn on your syllabus. Let's start with the standard Behavioral Genetics framework (reviewed by Sacerdote 2007, also on your syllabus).

- Let Y be a child outcome like IQ or education.

$$Y = G + F + S,$$

  where G is nature (genes), F is shared family environment, and S is non-shared or separate environment factors (which we can think of as the error term).

- If these factors enter linearly and there is no correlation among them, then we can write:

$$\sigma_Y^2 = \sigma_G^2 + \sigma_F^2 + \sigma_S^2.$$

  Dividing both sides by the variance of the outcome measure, and defining $h^2 = \sigma_G^2/\sigma_Y^2$, $c^2 = \sigma_F^2/\sigma_Y^2$, and $e^2 = \sigma_S^2/\sigma_Y^2$ yields the standard BG equation:

$$1 = h^2 + c^2 + e^2,$$

  where h stands for heritability and c stands for family environment (not sure why) and e is the error term.

- Note that the correlation between two adoptive siblings is (assuming all variables are standardized): $\text{Corr}(Y_1, Y_2) = \text{Cov}(Y_1, Y_2) = \text{Cov}(F_1, F_2) = \text{Var}(F_1) = c^2$.

- Thus, if the same family environment were randomly assigned to individuals with different genetic endowments (due to adoption, for example), we could use their covariance in outcomes to estimate c^2 (and by comparison with biological children in the same families, we could back out h^2).

It is commonly estimated that h^2 is between 0.6 and 0.8 for IQ. As Sacerdote points out, this consensus estimate is non-experimental and so probably includes elements of the covariance between c and h. But take this as a starting point.

A very high level of heritability makes it difficult for environmental factors to substantially affect IQ across generations. If $h^2 = 0.7$ and, let's say $c^2 = 0.2$ and $e^2 = 0.1$, then it would take a full standard deviation improvement in family environment to raise average IQ by 0.2 standard deviations across a single generation (approx 25 years).

3.2 The Flynn effect (the 'Flynn paradox?')

The paradox of this observation is, as James Flynn famously documented in 1987, IQs have been rising rapidly across essentially all advanced countries at the rate of 0.5 to 1.0 points per year. Given a standard deviation in IQ of 15 (these tests are standardized with mean 100 and SD 15), this phenomenon implies into a rate of improvement of more than 1 SD per generation. This would imply huge gains in family environment (or the error term or both). This does not appear credible.

The Dickens and Flynn paper offers a model to explain this paradox. Their model has three elements:

1. First, G and F are assumed to be positively correlated. People born into high IQ households also have better environments. This is surely correct, and suggests that to credibly estimate c^2, we'd need to randomly assign environment. (See Sacerdote).

2. Second, they posit that F responds endogenously to G. People of high ability seek out better environments, and society intrinsically rewards talent with further investment. So, an individual who is born with a talent for basketball (their example) will receive a 'basketball environment' that is potentially many standard deviations above the average prevailing in that period.

3. They also allow for a social multiplier by which high IQ individuals come in contact with other high IQ individuals and, more generally, that a secular rise in societal IQ augments the effect of individual gains (this presumably works through S).

These arguments are intriguing and seemingly plausible. It is perhaps premature to conclude that the paradox is 'resolved'—as Dickens and Flynn claim—without some evidence on these channels. The papers by Sacerdote and by Bjorklund, Lindahl and Plug provide intriguing evidence.

4 Nature vs. Nurture: Lessons from Swedish adoptions

There are a large number of adoption studies that attempt to distinguish the effects of environment from heredity. What makes the Bjorklund, Lindahl and Plug study ('BLP'—not to be confused with the Barry, Levinsohn, Pakes) unique is that it exploits detailed data on both adoptive *and* biological parents (hey, it's Sweden!). Under some strong assumptions, one can use these data to simultaneously estimate the effect of genes and environment on adult outcomes of adopted children.

More precisely, the question of the paper is how much of the association between parental characteristics and their children's outcomes is due to genes and prenatal environment vs. postbirth factors (such as childhood environment). As per Dickens and Flynn, parental genes very likely affect both the child's birth characteristics and her home environment. Thus, to separate the contribution of genes from post-birth factors, an experiment is needed.

The prototypical intergenerational association model estimated in the literature is:

$$Y_j^{bc} = \beta_0 + \beta_1 Y_j^{bp} + \nu_j^{bc},$$

where *bc* and *bp* refer to birth-child and birth-parent. The coefficient β_1 includes the effect of birth parent characteristics operating through heredity and environment (whatever loads onto the specific Y^{bp}, which might be, for example, income or education).

Using data on adoptees and their birth and adoptive parents, we can estimate the following model:

$$Y_j^{ac} = \alpha_0 + \alpha_1 Y_j^{bp} + \alpha_2 Y_j^{ap} + \nu_j^{ac},$$

where *ac* stands for adoptive child. Note again that Y^{bp} and Y^{ap} capture all things correlated with the 'environments' supplied by the birth and adoptive parents, respectively. So, if Y is income, this might be correlated with the in-utero environment of the birth parent (nutrition, smoking, drug use) and the home environment of the adoptive parent (books, school quality, etc). It is perhaps better to think of Y^{bp} and Y^{ap} as pre- and post-birth factors rather than nature and nurture per se.

In theory, one can further decompose Y_j^{bp} into genetic and environmental factors by separately controlling for birth-mother and birth-father characteristics, under the assumption that mother influences both genes and environment and that father influences only genes. This assumption may be problematic, however, since the father's income might affect the fetus' nutritional environment.

To consistently estimate the second equation, one needs two key assumptions:

1. Birth parents are randomly assigned

2. Adoptees are removed at birth so the effect of *bp* is limited to prebirth factors.

If we further want to compare β_1 with α_1 and α_2, we also need to assume that adoptive and own-birth children are not treated differently. Otherwise, we cannot make externally valid inferences to outcomes for own-birth children.

To make useful comparisons between estimates of β_1 for own-birth children and estimates of α_1 and α_2 for adoptive children, we will want to estimate β_1 with α_1 and α_2 on comparable populations. Swedish adoptees are quite disadvantaged (by Swedish standards). BLP draw two samples of comparison kids: (1) kids who are born in comparable circumstances to the adoptive kids; (2) kids who are raised in comparable circumstances to the adoptive kids. Note that there is a very limited sample of non-adoptive kids who are both born into and raised in similar circumstances to adoptive children since adoptive children, on average, experience a large jump in home environment relative to birth environment at the time of adoption.

The models above assume that there is no interaction between genetics and environment. This is quite restrictive. BLP will further estimate interactive models to allow for complementarity or substitutability between pre and post-birth environments:

$$Y_j^{ac} = \alpha_0 + \alpha_1 Y_j^{bp} + \alpha_2 Y_j^{ap} + \alpha_3 Y_j^{bp} \times Y_j^{ap} + \nu_j^{ac}.$$

Notice that this implies the following model for own-birth children:

$$Y_j^{bc} = \beta_0 + \beta_1 Y_j^{bp} + \beta_2 \left(Y_j^{bp}\right)^2 + \nu_j^{bc},$$

which suggests a non-linear relationship between *bp* characteristics and child outcomes. However, this non-linearity is an artifact of the functional form chosen. We clearly have little information on the true (structural) functional form linking the $Y's$ of parents and of children. It will therefore make sense to flexibly control for Y_j^{bp} and Y_j^{ap} while testing for interactions between these variables.

See the paper for institutional details on Swedish adoptions.

Some important notes on the data:

- The income measure for fathers is averaged over a twenty year earnings period. This probably provides a relatively good measure of 'permanent' earnings.

- Children's outcomes are measured at ages 33-37. At this age, children will have completed school. And their present earnings should be reasonably highly correlated with their lifetime earnings.

Table II contains the first main results:

1. Biological parents do matter for adopted children. Moreover, the fact that impacts of bio mothers are not much larger than impacts of bio fathers on education of adopted children suggests that prenatal environment not that important on average in these data. (Recall assumption that mother influences prenatal environment + genes whereas father only affects genes)

2. Adoptive parents also matter for education.

3. Biological mother has larger influence on education of child than does adoptive mother. But for fathers, biological and adoptive parent are about equally important. Not clear why this should be true.

4. The sum of estimates for bio + adoptive parent impacts on education of adopted children is in all cases about equal to that for bio parent for own-birth children. This is consistent with the idea that there is no adoption effect per se.

5. Post-birth factors appear more important for earnings and income than do pre-birth factors.

Table III contains many sensitivity tests that we won't discuss in class. One interesting thing is that when the authors estimate own-birth transmission coefficients on samples that are matched to have the same characteristics of adoptees' biological backgrounds (that is, low income families), the transmission coefficients are smaller. This may suggest that transmission is stronger in more affluent families.

Table IV suggests that intergenerational transmission is higher for more affluent and educated families. This is potentially consistent with the Dickens and Flynn argument.

Even conditional on this non-linear relationship, there appears to be a positive interaction between birth mother and adoptive mother characteristics in education outcomes, and a positive interaction between birth and adoptive father characteristics in earnings and income. Thus, these estimates imply a positive nature-nurture interaction.

Thus, poor children born into poor backgrounds who are *not* adopted suffer from poor heredity, poor environment *and* the interaction of the two. Wealthy children benefit from good heredity, good environment and the interaction of the two. This creates a clear dilemma for social policy because it implies that the marginal return to post-birth investments in children born in poor circumstances may be *lower* than the returns to investments in children born into better circumstances (for a given level of post-birth environment).

Might we reach different conclusions about the importance of nature versus nurture if we did this study in the U.S. instead of Sweden?

5 Additional evidence on nature vs. nurture: Sacerdote 2007

The paper by Sacerdote on your syllabus offers a related analysis of the effect of adoptive family environments on adult outcomes of Korean American adoptees. Unlike Bjorklund et al., this study does not have information on birth parents. It does, however, study a richer set of outcomes (including overweight, smoking, number of children) and also makes use of a richer set of variables on adoptive families (including the number of siblings present, which is negatively correlated with adoptees' outcomes).

It also makes good use of the standard Behavioral Genetics framework. Returning to the exposition above, Sacerdote observes that one can calculate the full BG breakdown from just observing the correlations between biological and adoptive siblings. In particular, under the assumption that biological siblings hold half of their genes in common and adoptive siblings hold no genes in common, we obtain that:

$$Corr\left(Y_1, Y_1\right)_{adopt} = c^2,$$

$$
\begin{aligned}
Corr\left(Y_1, Y_1\right)_{bio} &= Cov\left(G_1 + F_1 + S_1, G_2 + F_2 + S_2\right) \\
&= Cov\left(G_1 + F_1, \frac{1}{2}G_1 + F_1\right) \\
&= \frac{1}{2}h^2 + c^2
\end{aligned}
$$

Thus

$$
\hat{h}^2 = 2 \times \left[Corr\left(Y_1, Y_1\right)_{bio} - Corr\left(Y_1, Y_1\right)_{adopt}\right]
$$

See Table V for main results of this exercise. For outcomes similar to those studied by Bjorklund et al., Sacerdote's estimates suggest that 'nature' is about twice as important as 'nurture.' Is this consistent with Bjorklund? Roughly, it is. Observe that the correlation in outcomes between a *birth parent* and a child who is given up for adoption is:

$$
\begin{aligned}
Corr\left(P_b, Y_a\right) &= Cov\left(G_1 + F_1 + S_1, G_2 + F_2 + S_2\right) \\
&= Cov\left(G_1, \frac{1}{2}G_2\right) \\
&= \frac{1}{2}h^2.
\end{aligned}
$$

In the Bjorklund estimates, the size of the transmission coefficients for birth and adoptive parents is roughly equal. This is consistent with Sacerdote's finding that $\frac{1}{2}h^2 \approx c^2$, $h^2 \approx 2c^2$.

The one flaw in this argument (my argument, not Sacerdote's) is that Bjorklund et al. don't really estimate c^2, but rather the transmission from adoptive parents' outcomes to adopted children's outcomes. This link can only work through environment. But the Bjorklund transmission coefficient is only equal to c^2 if the parents' outcome is a perfect measure of the child's home environment. Certainly, the two will be positively related. But they are certainly not identical. Thus, it's unclear how these results may be exactly compared.

So, it's useful to examine Sacerdote Table VIII which estimates transmission coefficients directly. These look reasonably similar to Bjorklund et al.

Sacerdote's conclusion makes a very nice observation. In US data, the black-white gap in mean years of schooling is 0.78 and the gap in the college completion rate is 15.4 percentage points. Table VII estimates that these gaps could be produced by a 1 standard deviation change in family environment, and are equal to effect size of random assignment of an adoptee to a small, high-education family versus a large, less-educated family. Thus, if the mean black-white gap in family environments is one standard-deviation, then this would be sufficient to explain black-white gaps in educational attainment with *no differences* in genetic endowment. This is an important, albeit intuitive, observation. Although educational attainment or IQ are highly heritable, the presence of large group differences in these outcomes does not imply the existence of group differences in underlying heredity.

6 Stereotype Threat—Another form of self-fulfilling prophesy

The 'stereotype threat' hypothesis originates with psychologist Claude Steele and coauthors. This hypothesis says that members of groups that are 'stereotypically' believed to have negative attributes may behave in ways that confirm these attributes when the 'stereotype threat' is made salient. For example, female mathematicians may perform badly on math tests when they are reminded that many people believe that women are not as capable as males at mathematics. Or, blacks may perform poorly on IQ tests when the tester subtly suggests that blacks are not as intellectually capable as whites.

This hypothesis may sound far-fetched, but in fact it is easy to think of cases where it *could* be relevant. Steele gives the example of a black male sociologist who feels anxious when he is waiting in line at an ATM if the customer ahead of him happens to be a white woman. Although this sociologist has no criminal intent, he is aware that white female ATM customers may be made anxious by his presence, believing that black males pose a criminal threat. Presumably, the sociologist reacts to this 'stereotype threat' by trying to appear especially non-threatening, perhaps by keeping exaggerated physical distance from other customers. (It's conceivable that, opposite of the intention, this makes other customers more nervous.)

In this example, the stereotype threat does not make the sociologist more likely to engage in a criminal act (which the Steele hypothesis might suggest it should). But it does support the idea that members of discriminated groups may be acutely aware of stereotypes—at least in situations that make these stereotypes salient—and that this awareness could potentially affect behavior and outcomes. (For a fictional example in which stereotype threat experienced by black males *does* directly lead to a criminal act, see the car-jacking scene in the 2005 movie *Crash*.)

Anecdotes are not evidence, but, as we will see, the experimental evidence favoring the stereotype threat hypothesis is somewhat remarkable.

Note that the stereotype threat hypothesis does not fall under the other categories we've studied. It is neither animus-based nor statistical discrimination; it is self-fulfilling prophesy. One could write an economic model about this, but it would be difficult in such a model to motivate the idea that an individual would choose to behave in a way that is individually self-destructive in response to a stereotype. (Note that in the Coate-Loury paper, everyone is playing his or her optimal strategy.)

According to the Steele 1997 article, Stereotype Threat has two testable implications:

1. For domain-identified students, stereotype threat may interfere with their domain-related intellectual performance. Translation: For groups reputed to be worse at a given task, they may perform worse at that task if stereotype threat is activated.

2. "Reducing this threat in the performance setting, by reducing its interfering pressure, should improve the performance of otherwise stereotype-threatened students" [It's not clear to me that this implication is distinct from the prior implication.]

This idea may strike you as far-fetched (it does me). But it is readily confirmed by experimentation. Its economic importance is, however, unknown.

7 'Acting White'

There is a long-standing hypothesis, attributed to Fordham and Ogbu, that part of the deficit in Black academic achievement is due to the cultural stigma that Blacks face for 'acting White,' that is, conforming with the mainstream culture. Under the AW hypothesis, Blacks economic benefits but social costs of achieving, thus reducing skill investment. By contrast, Whites and other groups that do not have an oppositional culture, receive both economic and social benefits from achieving.

The relevance of this hypothesis—which has probably attained a bit of the character of urban myth—has come in for criticism at various points. A widely discussed article by Cook and Ludwig in 1998 (originally published in the *APPAM*) documented using the NLSY that

Blacks who are academically successful are as popular as academically successful whites. C&L view this as evidence that Acting White is not in reality stigmatized. However, the 2005 article on your syllabus by Austen-Smith and Fryer counters that C&L confuse the level with the derivative. That is, the AW hypothesis says that the derivative of social status WRT academic success if negative (or less positive) for Blacks than Whites, and this is not at odds with academically successful members of both groups being popular.

Because the AW hypothesis has received so much play, it is potentially valuable to see how it can be framed in economic terms. The paper by Austen-Smith and Fryer offers an example. The set the problem up as one of 'dual audience signaling.' Specifically, agents are attempting to signal their desirability to two groups simultaneously: employers and peers. Employers do not care about peer group membership and peers do not care about skills. But nevertheless, the model identifies a potential tension between the objectives of pleasing both audiences.

A nice insight of the model: In environments in which "acting White" is salient, improved external labor markets have the effect of encouraging more individuals to leave the group, while causing those in the group to invest less in education."