LECTURE II: Social Identity and the Ethics of Punishment

Recapitulation of Lecture I

In the previous lecture I asserted that the racially disparate incidence of the massive punitive structure which has been built-up in the US over the past thirty years is, when viewed in historical context, patently unjust, and that this situation weakens the legitimacy of the American political regime – appropriately so – in the eyes of many of its citizens and in the view of a great many people throughout the world who see this social practice, in light of our racial history, as barbaric. I offered a concise, dense and brutal "history" of the rise of the race/class/punishment nexus since the 1960's – covering basic facts about incarceration rates, about how the incidence of punishment varies by social location, and about the social and epidemiological harm that punishment inflicts on the communities from which offenders come, and to which they invariably return. I was keen to emphasize the connection of this development to the rhetoric of social discipline writ large in our recent politics: rhetoric about dependency, personal responsibility, social hygiene, and punishment as the reclamation of public order.

Moreover, I was insistent that this development must be viewed within the context of America’s often brutal, ugly and violent racial history. Slavery ended a long time ago, I acknowledged; but, I observed, the institution of chattel slavery and the ideology of racial subordination that necessarily accompanied it have cast a very long shadow. Those distant events are not unrelated to our current situation – neither as a matter of objective historical causation (the structure of our cities with their massive racial ghettos being implicated in the production of deviancy amongst those living their), nor as a matter of ethical evaluation (the decency of our institutions being dependent on the extent to which they comport with a narrative of national purpose that involves acknowledging and acting to correct history’s wrongs.)

I argued that race was not peripheral, but central, to the evolution of social policy in the US in this post-civil rights era; and that the rise of the mass incarceration state should be viewed within that broader context. I suggested that there is a reason why the United States is so exceptional among the democratic industrial societies in the severity and extent of its punitive policy, and in the paucity of its social welfare institutions. There is a reason, I suggested, why such social solidarity as we can muster seems not to extend into these marginal backwaters. There is a reason why our inclination toward forgiveness and the extension of a second chance to those who have violated our behavioral strictures is so stunted; and why our mainstream political discourses are so bereft of reflexive self-examination and searching social criticism. These phenomena, I suggested, have a great deal to do with the fact that “the faces at the bottom of the well” in American society – to borrow Derrick Bell’s phrase – are vastly disproportionately black and brown.

I claimed, and tried to support the claim with extensive empirical evidence, that the rise of the mass imprisonment state opens a new front in the historic struggle for racial justice. I was unapologetic about linking the notions of race and social justice. I rejected the purely procedural conception of racial fairness, wherein the goal is set at merely insuring equal treatment before the law, thereafter letting the chips fall as they may. Instead, I advocated the rather more demanding idea of substantive racial justice, wherein the goal is to bring about through conventional social policy and far reaching institutional reforms a situation in which the fact of historic racial oppression is no longer so evident in the disparate life experiences of those who descend from slaves. I argued
that the enormous racial disparity in our punishment policy reflects both explicit and tacit racism -- that this policy has been popular sometimes because, and sometimes despite, it having a disproportionately adverse impact on blacks. And, I suggested that all of this has occurred when feasible alternatives policies existed, and were known to exist, that might have produced much less harm. Finally, I claimed that this punishment policy complex has become the principal way that racial hierarchy is now reproduced in our society, and that this matter requires and deserves the concerted attention of the nation’s policy makers. In a rush to declare ourselves “healed” of the disease of racism which had been festering for a century after emancipation we have embraced what Michael Tonry (1996) calls a policy of "malign neglect."

**Prologue for Lecture II:** Consider this imaginary dialogue between two distinguished social scientists whom I shall refer to as Dr. Rationalist (RT) and Dr. Functionalist (FT). Or, rather than two different scholars, maybe these are two warring voices within one poor scholar’s head.

RT: (*chanting, but otherwise sitting very still*) “Relations before transactions. Relations before transactions. Relations before transactions…”

FT: (*enters with a start – alarmed*) “What’s wrong, my friend? Why are you saying that? You must be the culprit who lifted my copy of Bourdieu last week.

RT: No, I’m not. And, who’s Bourdieu, anyway? One of those airy French sociologists you’re always fawning over, I’ll bet. It’s my mantra; I’m meditating. It’s very calming. You ought to try it sometime.

FT: (*Ignoring the dig.* ) I mediate all the time. Remember: I’m the one who belongs to a profession fraught with anxiety. But what’s your excuse?

RT: Well, I’ve been having a recurrent nightmare lately. I want it to stop. My shrink thinks meditation might help.

FT: Who’s your shrink?

RT: Oh, this guy who was my roommate at Swarthmore. Brilliant dude; works a lot with gunshot victims; inner-city types involved in the drug trade, and so forth; he thinks they’re making passive suicide attempts. Writes books about that stuff – full of all this talk about fear, self-loathing, hopelessness, existential abyss, Freud, Nietzsche, de Sade. Strange guy, but brilliant. He gave me the mantra and promised it would help. Said I should repeat it slowly while taking deep breaths.

FT: Perhaps. But, remember what I told you about those pizzas – not a good idea after midnight… Did you say, de Sade?? Anyway, what’s the dream?

RT: Oh, it’s awful. I’m back in grad school. I’m sitting in my usual place right at the front of the class. The professor has posed what he says is an important question. He’s invited one of us to the board to work out an answer. I get there first, and proceed to fill the board with equations. Finally, I arrive at what must be the solution. My derivation is far too elegant not to be true. I turn to explain myself to the rest of the class. Just then, I realize that I’ve forgotten the original question! I rack my (very large) brain, but for the life of me, I can’t recall it. The class begins to snicker. They’re a ruthless bunch when they smell blood. The guffaws and catcalls grow louder. It’s humiliating, just humiliating. (Dr. Rationalist begins to tremble uncontrollably.)

FT: (*Comforting his friend*) Yeah, I can see that. It’s got to be tough – being the smartest person in the room, but without a clue as to what’s the point. You ought to stick with this shrink. Dreams can be very revealing, you know. But, I’m not sure I get the mantra. And, what was the professor’s question, anyway?

RT: He asks us to explain how durable racial inequality in the US can be squared with the premises of rational choice theory, without assuming any innate racial inferiority or postulating any unexplained preference for own-group associations.
FT: That’s a damned good question! It’s a tough one, too. You’re telling me you ran to the board to take that one on? Brave man. (Fools jump in where angels fear to tread, he thinks.)

RT: Well, to be honest, in the dream I always start to the board before he finishes posing the question. Happens the same way every time. I can’t stop myself... (The trembling returns.)

FT: (In a bright tone, hoping to shift to a happier subject.) So, what was your elegant solution?

RT: I’d tell you but, what with the math and all, you’d never understand it. (At this, Dr. Functionalist takes offense and storms off angrily. Dr. Rationalist yells after him…)

RT: Besides, I’m not sure I believe it anymore, myself. Anyway, my shrink gave me this mantra and it seems to be helping. (He returns to his chanting... “Relations before transactions. Relations before transactions. Relations before transactions...”)

Plan of Attack

This afternoon I, a humble non-philosopher, will nevertheless try to employ the tools of my trade in the service of a normative argument. I will endeavor to defend some of the value statements embodied in yesterday’s lecture (which, I confess, veered-off at times into vituperation and the rhetoric of moral outrage. For this, I offer no apology.) As an economist, I am aware that I tread on thin ice here. My professional training and practice have not prepared me for the task at hand. I must confess that, when I’m feeling like Dr. Rationalist in that opening dialogue, I am beset by doubts: what can a mere economist offer to a debate about fundamental social values – beyond getting the numbers straight, and ensuring that the logic of causal argument is devoid of internal contradiction? How can the rational choice-style of reasoning that is natural to an economist shed light on the moral issues at hand? If criminals must be held responsible for their wrongful deeds, if blacks are found more frequently to have acted wrongly, and if one wants to argue that their disproportionate punishment is somehow unjust – which is to say, that the sum of individual cases judged more or less fairly nevertheless amounts in the aggregate to morally unfounded social judgment – then what good is a mere reckoning of costs and benefits?

Frankly, I am not sure I have a good answer for that question – which is why I identify so closely with our troubled Dr. Rationalist in that opening dialogue. But then, I also have great sympathy for the position of Dr. Functionalist. In any event, I feel compelled to step out onto the proverbial limb in an effort try to bridge the gap between the utilitarian worldview of my profession, and the deontological outlook of contemporary moral theory. And so, here goes.

Summary of the Moral Argument

While much of the discussion to come is abstract and theoretical, it is not divorced from the reality that I tried vividly to render in yesterday’s more descriptive and empirical exposition. The connection between theoretical abstraction and concrete social reality lies in our need to elaborate foundational principles by means of which we can critically assess inequality in this society and, in particular, the racially unequal allocation of punishment. I will pose this task of ethical assessment as a problem in the theory of distributive justice, and approach that problem from a Rawlsian perspective, invoking his famed Difference Principle – that departures from equality are only justified to the extent that they work to improve the circumstances of the least advantaged members of society – except that, now, the object of moral inquiry is not the distribution among individuals
of wealth and income, but instead the distribution of a negative good, punishment, among individuals and, importantly, between racial groups.

So, let us put ourselves in Rawls' argument original position and imagine a situation where each of us could occupy any rank in the social hierarchy, including that of black American male outcast shuffling between prison and the labor market on his way to an early death to the chorus – courtesy of Fox News – of NIGGER or CRIMINAL or DUMMY. What set of social rules would we devise if we actually thought that THEY could be US? What set of institutions should we create so as to maximize human developmental potential? And, most pointedly, what set of punishment institutions would we create in order to contain bad behavior and protect society, to be sure, but in ways that respect the humanity of each individual and of those they are connected to via the bonds of social and psychic affiliation. If it were, indeed, the case that anyone of us had a real chance of being one of those faces looking up from the bottom of the well, then how would we choose to order our criminal justice practice? How would we talk publicly about those who break our laws? What would we do with juveniles who go awry, who roam the streets with guns and sometimes commit acts of violence? What weight would we give to various elements in the deterrence-retribution-incapacitation-rehabilitation calculus, if we thought that calculus could end-up being applied to our own children, or to us? How would we apportion blame and affix responsibility for the cultural and social pathologies evident in some quarters of our society, if we envision that we ourselves might well have been born into the social margins where such pathology flourishes?

Without making a full-fledged philosophical argument (of which, I fear, I am not capable), I nevertheless wish to gesture – in the spirit of John Rawls – toward some answers to these questions. Specifically, I will be trying to show the limits of a pure ethic of personal responsibility as the basis for distributing the negative good of punishment in America’s contemporary, racially hierarchical and class-stratified society. My aim is to shift the boundary toward a greater acknowledgement of social responsibility – even for the wrongful acts freely chosen by individual persons. In pursuing this aim, I am not so much making a “root causes” argument – he did the crime, but only because he had no choice – as I am arguing that the society at large is implicated in his choices because we have acquiesced in structural arrangements which work to our benefit and his detriment, and which shape his consciousness and sense of identity in such a way that the choices he makes, which we may condemn, are nevertheless compelling to him. I am interested in the moral implications of what Loic Wacquant has called the “double-sided production of urban marginality.” (Wacquant ??; Sidanius and Pratto ??) As a social scientist I will approach this problem of moral philosophy by emphasizing that closed and bounded social structures – like racially homogeneous urban ghettos – create contexts where “pathological” and “dysfunctional” cultural forms emerge; but, these forms are not, I will suggest, intrinsic to the people caught in these structures. Nor are they independent of the behavior of the people who stand outside of them.

**Segregation, Spillovers and Persistent Inequality: Theory**

Disparity in the incidence of punishment in America is just one facet of a persisting racial hierarchy. Persistent group inequality is nothing new. Many societies have sustained long periods of hierarchical organization characterized by distinctly unequal opportunity for members of different social groups (Tilly 1998; Sidanius and
Pratto 2001.) Examples include the United States during slavery and the Jim Crow period, where hierarchy and status were based in an informal system of racial classification; South Africa during Apartheid, where a person’s social identity was based on a formal system of racial classification; and, the Indian subcontinent, where one finds caste-based hierarchies that have been in place for centuries. I will argue that there is a common structural elements the spans these otherwise very different cases – the deleterious effects on human development that derive from social segregation. In each of these societies a transition has taken place from an explicitly hierarchical order to what is at least a notionally egalitarian one. Yet, when one looks closely, the consequences of these transitions have been less than fully satisfying. Evidently, historical discrimination against a marginal identity group implies that formal equality of opportunity need not result in the convergence of group outcomes.

One of the principal reasons for this, I argue, is that informal social networks remain segregated under the newly reformed order. Enforcement of anti-discrimination laws may eradicate discrimination, but because many important non-market interactions lie outside the scope of such laws the reform process cannot undo the harmful affects of stigma. The law can have only the most modest impact on individual choices of a date, a spouse, an adopted child, a role model, residence in a neighborhood of membership in a voluntary association. Early childhood development takes place within the spheres of family, neighborhood and peer-group. Racial segregation in the formation of these social networks can have important implications for the perpetuation of group inequality across generations. Discrimination in contact can give rise to persistent group inequality even in the absence of discrimination in contract.

At the most basic conceptual level, one expects to see a link between social segregation and the dynamics of inequality, because of what an economist would call interpersonal spillovers in human capital accumulation. Human development always and everywhere takes place within a social context, and can be greatly facilitated by access to a social network that is rich in human capital. In his classic article on the mechanics of economic development (1988), Robert Lucas observes that, “human capital accumulation is a social activity, involving groups of people in a way that has no counterpart in the accumulation of physical capital.” The economists William Brock and Steven Durlauf (2001) have hypothesized that a given individual's cost of investing effort in education falls as the level of investment by his social affiliates rises. When this is so, two persons with identical ability who belong to different social groups will generally make different investment decisions, and group bias in social ties can cause historical group disparities to become locked-in. This can happen even though human capital investments are not impeded by financial constraints.
I have explored this issue, at the level of theory, in a paper with Sam Bowles and Rajiv Sethi (2007). The paper develops a model where people belong to one of two groups and parents invest in the human capital of their children. We study the case where there are two occupational categories, one of which requires a costly human capital investment and the other not. The cost of human capital depends on ability and on the level of human capital in a person’s social network. Under market competition wages are determined by the overall distribution of human capital in the economy, and investment decisions are assumed to depend on anticipated wages. Discrimination in the labor market is assumed not to exist, due to the perfect enforcement of laws guaranteeing formal equality of opportunity. Nevertheless, we show that even if ability is identically distributed within the two groups, when the initial state is one of inequality, when human capital spillovers are important, and when the degree of informal segregation of social networks is sufficiently great, then the members of different groups will invest in human capital at different rates. Our findings are illustrated in Figure 2 above, where a threshold level of network segregation above which persistent inequality is assured is shown to vary with the population share of the disadvantaged group. Furthermore, we show that when segregation is too high, groups are not likely to achieve similar socioeconomic outcomes even if they do not differ in underlying potential.

This finding is relevant to the debate over the appropriate policy response to a history of overt discrimination. And, in a general way, it is relevant to a discussion of racial inequality in the burdens imposed by law enforcement. Rule-oriented approaches emphasize the equality of individual persons before the law. This procedural approach to defining equal citizenship is characterized by a focus on individuals and their rights – but necessarily gives short shrift to group interactions and large-scale social processes. And, it leaves little room for the advocacy of group-redistributive remedies. That inequality can persist indefinitely between racially self-aware population aggregates, even in the

![Figure 2. Effects of segregation and population shares on persistent inequality](image-url)
face of formal equality of opportunity need not imply that group-redistributive policies ought to be adopted in every case. But it does mean that a failure to adopt them can result in persistent divergence across groups in economic and social outcomes. Put differently, if group equality is a policy goal, equal opportunity may not be enough to secure it.

Conditional on the income of their parents, African-Americans earn substantially (about a third) lower incomes than do whites, and this intergenerational race gap has not diminished appreciably over the past two decades (Hertz, 2005). Similarly, Derek Neal of the University of Chicago reports that convergence in years of schooling attained and cognitive scores at given levels of schooling, which was being celebrated prior to 1980, appears not to have continued thereafter. Significant racial differences in mortality, wealth, subjective well being, and other indicia also persist (Deaton and Lubotsky, 2003, Wolff, 1998, Blanchflower and Oswald, 2004).

Enduring discriminatory practices are no doubt part of the explanation (Bobo et al., 1997, Greenwald et al., 1998, Bertrand and Mullainathan, 2004, Quillian, 2006). These may be motivated by racial prejudice or hostility (Becker, 1957), or by more subtle psychic mechanisms. Although there is considerable survey evidence to suggest that the principle of equal opportunity is now widely accepted (Bobo, 2001, Schuman et al., 1997), it is also clear from the evidence that racial stereotypes persist to the disadvantage of blacks and, as I suggested yesterday, that the policy preferences of whites on facially non-racial matters (like welfare and crime) are nevertheless influenced, again to the disadvantage of blacks, by racial factors. (Bobo and Johnson 2004.)

The Importance of Social Affiliation: Evidence

Racial sorting within social networks represents a mechanism other than explicit formal discrimination through which group inequality may be sustained: Segregation of friendship networks, mentoring relationships, neighborhoods, workplaces and schools places the less affluent group at a disadvantage in acquiring the things -- contacts, information, cognitive skills, behavioral attributes -- that contribute to economic success. Preferentially associating with members of one's own kind (known as homophily) is a common human trait that has been well-documented for race and ethnic identification, as well as religion, and other characteristics (Tajfel, Billig, Bundy, and Flament, 1971). A recent survey of empirical work in this area (McPherson, Smith-Lovin, and Cook, 2001) reported that:

“We find strong homophily on race and ethnicity in a wide range of relationships, ranging from the most intimate bonds of marriage and confiding, to the more limited ties of schoolmate friendship and work relations, to the limited networks of discussion about a particular topic, to the mere fact of appearing in public or 'knowing about' someone else... Homophily limits peoples' social worlds in a way that has powerful implications for the information they receive, the attitudes they form, and the interactions they experience.” (pp. 415, 420).

In a nationally representative sample of 130 schools and 90,118 students, controlling for school racial composition, same-race friendships were found to be almost twice as likely as cross-race friendships (Moody, 2001). Compared to the friends of white students
in this sample, the friends of African American students had significantly lower grades, attachment to school, and parental socioeconomic status. There is also evidence that peer effects such as penalties for ‘acting white’ among African American students can provide disincentives for academic achievement (Fryer and Torelli, 2005).

While there are many channels through which the racial bias in social networks can disadvantage members of the less well-off group, statistical identification of these effects often is an insurmountable challenge because networks are selected by individuals and, as a result, exogenous variation in the composition of an individual's networks is seldom observed. Some scholars [Hoxby (2000); Hanushek, Kain, and Rivkin (2002)] have used the year-to-year cohort variation in racial composition within grade and school to attempt to identify racial network effects. They find large negative effects of racial sorting on the academic achievement of black students. Moreover, studies that use the randomized assignment of college roommates have also found important behavioral and academic peer effects (Kremer and Levy, 2003, Sacerdote, 2001, Zimmerman and Williams, 2003). Another study (of annual work hours), that used longitudinal data with individual fixed effects, found strong neighborhood effects especially for the least educated and those in the poorest neighborhoods (Weinberg et al., 2004). An experimental study documents strong peer effects in a production task, particularly for those with low productivity in the absence of peers (Falk and Ichino, 2004). Our model shows that socioeconomic outcomes for blacks and whites can continue to diverge across the generations because of racial differences in the social networks to which individuals have access.

Stigma versus Discrimination: More Theory

Social networks are also social products – the result of the choices to affiliate that people make. In turn, choices about informal association and imagined connectivity are at least partially guided by how people perceive and value those with whom they might or might not connect. So, we need to consider why, if the era of racism and discrimination has passed, are the social networks important for human development so consistently segregate by race in American society.

In this regard, I believe that it is crucially important to distinguish between racial discrimination and racial stigma. Discrimination has to do with how blacks are treated, while stigma is concerned with how black people are perceived. I do not claim these two kinds of bias to be mutually exclusive. Still, I think this distinction is useful for, whereas the bias emerging from formal discrimination presents a straightforward moral problem, and calls forth the nearly universally embraced remedy of anti-discrimination law, the barriers to human development deriving from informal racial bias is a subtler, more insidious moral problem – one that may be impossible to remedy in any manner likely to garner a majority’s support. This difficulty has both a cognitive and an ethical dimension. From a cognitive point of view, many observers, when seeking an explanation for a group’s poor social performance, may not distinguish between limited opportunities for development and limited capacities to develop. From an ethical point of view, citizens who find the overt discrimination by race associated with formal bias to be noxious may nonetheless be unmoved by the tacit discrimination that blocks access for blacks to the informal social networks that facilitate human development.

In my book, The Anatomy of Racial Inequality, I suggested that we think about “race” as a social phenomenon which results from the combination of two processes –
categorization and signification. Categorization entails the sorting of persons into a cognitively manageable number of subgroups, based on their physical appearance, in order to differentiate one’s dealings with such persons. Signification involves the mental activity of associating certain connotations or “meanings” with these categories. My argument was that, at bottom, “race” is all about “embodied social signification.” With my core concept – biased social cognition – I attempted to move from the fact that people make use of racial classifications, to some understanding of how this alters the causal accounts settled-upon for what they observe in the social world. I asked: how does the “race” of those experiencing some difficult circumstance affect whether powerful observers come to understand the plight of the disadvantaged as constituting a societal, as distinct from a communal, problem?

Causal Misattribution and Its Consequences

My argument may be illustrated as follows:

This figure models an abstract, dynamic process where the prevalence of some activity within a social group evolves over time. For concreteness, let us suppose that \( X(t) \) represents the prevalence of gun-carrying among young males in some population, at stage \( t \) of an adjustment process. Each young man decides whether or not to carry at each stage of the process, based on how many others he expects to be carrying. That is, a given young man carries only if he thinks the overall carry rate is sufficiently great. But this threshold (above which a given young man carries) varies from one person to the next. The lazy-S curve in the figure, labeled \( F(x) \), depicts the proportion of young men who feel obliged to carry a gun when they expect that the overall carry rate to be \( x \).

The figure illustrates what is sometimes called “the tipping property”: For \( x < x^* \), few want to carry; but for \( x > x^* \) most want to carry. The dynamic behavior of this little system depends on initial conditions. If the initial expectation, \( X(0) \) lies above the threshold \( x^* \) then the system eventually settles-in at a relatively low rate of gun carrying in the population. While, if the initial expectation exceeds the threshold, the system “tips” – that is, it gravitates toward a relatively high carry rate. If there were two racially distinct and socially isolated population sub-groups, then even though these racial groups may be characterized by an identical distribution of gun-carrying proclivities (represented by the function \( F(x) \)), different initial conditions would lead to radically different long-run behavioral patterns within the groups. (Even though I illustrate this argument by
reference to gun ownership, the same logic could apply to many other behaviors where
the value of an action to each individual depends on how many others in their social
network are acting in a similar way.)

Consider now the inferences about group inequality that someone observing this
process might be inclined to make. The structure here is a bit complex – multiple self-
sustaining behavior patterns are consistent with a common population distribution of
some trait (here, the proclivity to carry a gun). If whites are observed mainly near the low
equilibrium in the figure, and if blacks are observed mainly near the high one, an
observer, particularly one already inclined in this direction, will be tempted to conclude
that the racial populations are essentially different, characterized by distinct distribution
functions [denoted FW(x) and FB(x) in the figure]. That is, they might settle upon what I
call a racially essentialist account of the inequality. The choice between the complex
structural account and the simplistic essentialist one is a choice of specification (about the
causal processes underlying one’s observations) which can easily be impacted by the
social meanings associated with the racial markers characteristic of a particular society.

I am suggesting here that tacit association in the American public's imagination of
“blackness” with "unworthiness" or with “dangerousness” affects cognitive processes and
promotes essentialist causal misattributions. When confronted by the facts of racially
disparate achievement, racially disproportionate transgression of legal strictures, and
racially unequal development of productive potential, observers will have difficulty
identifying with the plight of a group of people whom they (mistakenly) think are simply
“reaping what they have sown.”

**Boys and Girls, Blacks and Whites**

I can make this same point with a less abstract example, one that does not involve
race. Consider gender inequality. We know that there is disparity in the social outcomes
for boys and girls in two different venues – the schools and the jails. Thus suppose that,
when compared to the girls, boys are over-represented among those doing well in math
and science in the schools, and also among those doing poorly in society at large by
ending-up in jail. There is some evidence to support both suppositions, but only the first
is widely perceived to be a problem for public policy. Why? My answer is that it
offends our basic intuition about the propriety of underlying social processes that boys
and girls have different levels of achievement in the technical curriculum of our schools.
Although we may not be able to put our fingers on exactly why this outcome occurs, we
instinctively know that it is not right. In the face of this disparity we are inclined to
interrogate our institutions – to search the record of our social practice and examine
myriad possibilities in order to see where things might have gone wrong. Our base-line
expectation is that equality should prevail here. Our moral sensibility is offended when it
does not. And so, an impetus to reform is spurred thereby. We cannot easily envision a
wholly legitimate sequence of events that would produce the disparity, so we set
ourselves the task of solving a problem.

On the other hand, gender disparity in rates of imprisonment occasions no such
disquiet. This is because, tacitly if not explicitly, we are “gender essentialists.” That is,
we think boys and girls are different in some ways relevant to explaining the observed
disparity – different either in their biological natures, or in their deeply ingrained
socializations. (Note well, the essentialism with which I am concerned need not be based
solely or even mainly in biology. It can be grounded in (possibly false) beliefs about profound cultural difference as well.) As “gender essentialists,” our intuitions are not offended by the fact of vastly higher rates of imprisonment among males than females. We seldom ask any deeper questions about why this disparity has come about. In this sense, we do not perceive there to be a problem, and so no solution is sought.

Now, we may be right or wrong to act as we do in these gender disparity matters, but my point with the example is to show that the bare facts of gender disparity do not, in themselves, suggest any course of action. To act, we must marry the facts we observe to some model of social causation. This model need not be explicit in our minds. It can and usually will lurk beneath the surface of our conscious reflections. Still, it is the facts plus the model that lead us to perceive a given circumstance as indicative of some as yet undiagnosed failing in our social interactions, or not. This kind of reflection on the deeper structure of our social-cognitive processes, as they bear on the issues of racial disparity, is what I have in mind when I speak of “biased social cognition.” And, the role of “race” in such processes is what I am alluding to when I talk about “racial stigma.” Such stigma could be of great political moment, because if one attributes an endogenous difference (a difference produced within a system of interactions) to an exogenous cause (a cause located outside that system), then one is unlikely to see any need for systemic reform.

This distinction between endogenous and exogenous sources of social causation, I am arguing, is the key to understanding the difference in our reformist intuitions about gender inequalities in the schools and in the jails: Because we think the disparity of school outcomes stems from endogenous sources, while the disparity of jail outcomes is tacitly attributed in most of our “causal models” to exogenous sources, we are differentially moved to do something about the observed disparities. So, the effect I am after when I talk about “racial stigma” and the reason I employ an apparently loaded phrase like “biased social cognition” is this: It is a politically consequential cognitive distortion to understand the observably disadvantageous position of a racially defined population subgroup as having emerged from qualities taken to be intrinsic to the group when, as a matter of actual social causation, that disadvantage is the product of a system of social interactions. I argue that a given instance of social disparity is less likely to be thought to constitute a social problem when people see the disparity as having been caused by what they take to be the deficiencies of those (e.g., the boys in the jails, but not the girls in the schools) who lag behind. I reiterate that it hardly matters whether the internal qualities mistakenly seen as source of some group’s observed laggardly status are biological or deeply cultural. What matters, I argue, is that something has gone wrong if observers fail to see systemic, endogenous interactions that lead to bad social outcomes for blacks, and instead attribute those results to exogenous factors taken as internal to the group in question.

Campaigning for Respect (and Being Judged for It)

I turn now to another analytic example. The Reputation Game models an ethically significant interaction between Bob and John. (See diagram below.) In this dynamic game of incomplete information, John can be either of two types – soft or tough [one thinks of Elijah Anderson’s “decent” vs. “street” distinction (Anderson 2001).] Bob chooses whether to Attack or Not; John responds, choosing whether to Fight or Not. The
game is ‘ethically significant’ because it shows that the link between character and behavior depends on the social context. The payoffs, specified in parentheses in the diagram, are such that both the soft and the tough versions of John would much prefer not to be attacked, though if attacked a soft John would want not to fight and a tough John would want to fight. (This is reflected in the fact that both soft and tough versions of John receive payoff = +1 from not fighting, while fighting gives soft John a payoff = -1 but tough John’s payoff from fighting =+2.) Also, the payoffs indicate that Bob wants not to attack if he will be fought (payoff = -1 versus 0), and wants to attack if he will not be fought (payoff = +1 versus 0).

THE REPUTATION GAME

Payoffs = (Bob, John)

If this interaction between Bob and John were to occur only once, then rational agents would play as follows: if attacked, a tough John fights and a soft John does not. So, Bob computes the average of his payoffs over the two outcomes that could occur and attacks only if he thinks the probability that John is tough is less than ½. The soft and tough types of John react naturally. The game is trivial.

But, should these players interact twice in succession, the outcome of the game is more interesting. Now, John’s action at the first stage can serve as a signal to Bob about his type, thereby affecting how the second stage is played. John knows this. Bob knows that John knows. John knows that Bob knows that he knows…, etc.

Given this structure, I make the following two claims:

Claim 1: In the twice-played game it is inconsistent with rationality for the soft John never to fight when first attacked.

Proof: If soft John were never to fight when first attacked, then because tough John always fights, John’s first stage action would be a perfect signal of his type: “fight at the first stage” would mean he’s tough, “not fight” would means he’s soft. But then, Bob’s rational response to this situation is to attack at the second stage only if he is fought at the first stage. Yet, this response on Bob’s part means that soft John could avoid being attacked at the second stage by fighting at the first stage which, if he is rational, he would want to do since (-1+4)--his payoff from fighting and thereby avoiding subsequent attack, exceeds (1+1)--his payoff from not fighting at either stage. This contradicts the supposition that he never fights when first attacked.
Claim 2: In the twice-played game it is inconsistent with rationality for Bob to always attack at the second stage after being fought at the first stage.

Proof: If Bob always attacks after being fought, soft John can gain nothing by fighting at the first stage and so, being rational, soft John would never fight when first attacked. But, Claim 1 asserts that this can’t occur when the players are rational.

Taken together, Claims 1 and 2 imply that the only outcome of the twice-played game consistent with player rationality has soft John mixing his behavior at the first stage between fighting and not fighting, while Bob mixes at the second stage between attacking and not, given that he has been initially fought. (If Bob is not fought at the first stage then he knows John is soft and so definitely attacks at the second stage.) Indeed, if $P < 1/2$ is Bob’s assessment of likelihood that John is tough, then the unique equilibrium of the twice-played game entails soft John fighting when first attacked with probability $P/(1-P)$, and Bob attacking at the second stage if fought at the first with probability $1/3$. John’s personal “values” may reflect a disdain for fighting (soft John), and yet his rational adaptation to his circumstances leads him to behave in a way that is inconsistent with those “values.”

‘Acting White’ or Blacks and Whites Acting?

The position I am trying to illustrate here is that we are not looking down into the souls or at the predetermined qualities of groups when we observe these disparate patterns of behavior. We’re looking at the end product, the final outcome, of a process of social interaction that is embedded in a larger structure of social relations, where ideas about race and racial identity play a prominent role. To see what I have in mind, consider the so-called “acting white” problem (Ogbu 2003; Ferguson 2005; Fryer and Torelli 2005): black students are said to ostracize those of their peers who value academic success, labeling such behavior as “acting white.” The graphs below, showing the relationship between popularity and grades among racially defined adolescent peer groups in a sample of American high schools (taken from the research of Harvard economist, Roland Fryer), tells the tale.

In response to data such as these, I would observe that “acting white” probably has as much to do with the way white people act, as it has to do with the way blacks act. The evidence presented in Fryer and Torelli (2005) supports this claim. In the schools where the population and therefore, one would have to say, the culture is predominantly black, Fryer observed no “acting white” behavior (student popularity varies positively with grades, no less so for blacks than whites). While, in the schools where blacks are fewer than 20% of the student body, the correlation between grades and popularity can be negative, at least when grades are above the level of B+ (see Figure 1B. below, where the middle line depicts the relationship between popularity and grades for black students, controlling for other observed student characteristics.). What this says to me is that the way black kids are interacting with each other has something to do with the way that black kids and white kids interact with each other. So, to talk about a pathological black youth culture, without reference to the larger structure of social interaction within which black students are situated, would be to make a mistake.
Popularity and Grades (1994-95)

Here is another example to illustrate my point -- marriage. Given that there is a high racial endogamy rate (i.e., most married blacks are married to other blacks; most married whites are married to other whites), one might imagine that the huge disparity in marital and child-bearing behavior between blacks and whites provides a compelling example of racially distinct cultural practices. Is one not forced to conclude that a higher out-of-wedlock birthrate among blacks than whites, or a lower marriage rate, indicates a racially disparate cultural dynamic?

Yet, consider that, since there are six times as many whites as blacks in America, were the percentage of white men marrying black women to increase slightly, the nature of the “marriage market” among blacks could be dramatically changed. As an economist, my model of marriage involves bargaining between negotiating parties over the division of the surplus from cooperation within the household. The threat point in that negotiation is “how well I could do if I didn't have you.” Were that threat point to move a little bit -- because, say, 5% of white men were inter-marrying with black women, then the leverage of black women in the intra-racial marriage market could be considerably enhanced, and as a result a very different outcome might arise in the equilibrium of marital interactions between black men and black women: Perhaps black men would be less “trifling” if they had more competition.

Of course, this speculation is all a bit “tongue in cheek.” My point, however, is not a factual one; nor am I making a policy recommendation. My point is a conceptual one: It is a mistake to say of black men and women that their marital behavior reveals inadequacies of black culture, when what one observes is the equilibrium outcome from a system of interactions within which a change of behavior across the group boundary (in the case at hand, behavior between white men and black women), partially determines the observed outcome.

Social Causation and Social Responsibility

Given the social meaning of “blackness” in contemporary American public culture (Bobo 2004; Fluery-Steiner 2004), I conclude that a racial disparity in the incidence of punishment is less likely to be seen as a societal problem, and more likely to be rationalize as a communal and individual problem (as in the conservative’s argument [e.g., Bennett, Dilulio and Walters 1996] that crime is a failing of personal character that has been caused by “moral poverty.”) And so it is that, in our time, an enormous racial disparity in the imposition of social exclusion, civic ex-communication and life-long
disgrace – an inequality engendered by mechanisms of imprisonment so extensive in their reach and so punitive in their practice as to dwarf any the modern world has heretofore known – can come to seem legitimate, even necessary: this disparity cannot be seen as having been caused by OUR failures; the inadequacy here must be seen as THEIRS.

Because this entire dynamic has its roots in past unjust acts that were perpetrated on the basis of race, present-day racial inequality constitutes a gross historical injustice. And, because the instrumentalities of exclusion, degradation, humiliation, shaming, and stigmatization associated the politics and the practice of mass incarceration are so central to the perpetuation of this unjust hierarchical order, their legitimacy and that of the state which promulgates them is called into question – at least, in the eyes of those bearing the brunt of this world-historic internal mobilization of state power.

The minimalist ethical principle that I wish to advance here is that historical racial injustice establishes a general presumption against indifference to contemporary racial inequality. This implies that the assignment of personal responsibility for wrongful acts is, in and of itself, insufficient to the task of doing justice. The sum of a million cases, each one rightly judged on its merits to be individually fair, may nevertheless constitute a great historic wrong. Thus, a central reality of our time is the fact that there has opened a wide racial gap in the acquisition of cognitive skills, the extent of law-abidingness, the stability of family relations, attachment to the workforce and the like. This disparity in human development is, as an historical matter, rooted in political, economic, social and cultural factors peculiar to this society, and reflective of its unlovely racial history. That is, the inequality of human development that is reflected in widely disparate rates of criminal offending by race in this country is a societal, not a communal or personal, achievement. At the level of the individual case we must, of course, act as if this were not so. There could be no law, and thus no civilization, without the imputation to particular persons of responsibility for their wrongful acts. This is the age-old problem of autonomy versus heteronomy that bedeviled Kant. But, the state does not only deal with individual cases. It also makes policies in the aggregate, and the consequences of these policies are more or less knowable. The morality of public action can therefore be assessed in these terms, if not by the lower courts when trying individual cases, then by the appellate courts when interpreting the Constitution, and by legislatures when promulgating public law.

Even if the current racial disparity in the incidence of punishment in our country gave evidenced of no overt discrimination against blacks – and, perhaps needless to say, I view that as a wildly optimistic supposition – it would still be true that powerful forces are at work to perpetuate the consequences of a universally acknowledged wrongful past. So, combating such racism as continues to exist must be seen as insufficient to achieve racial justice. I propose to see the problem in interpretative terms, seeking the assumption of collective responsibility for what has ensued. So doing should serve to inhibit the adoption of punishment policies which engender large racial disparities. This argument does not rest merely on identifying processes of historical causation (by arguing, for instance, that enforced racial segregation under Jim Crow impeded the accumulation of financial and social assets), and by linking said processes to contemporary disparities. Rather, this is in the first instance a matter of how we choose to interpret the facts – of the narrative overlay that we impose upon the facts. I am arguing on the basis of historical memory for a common national narrative through which the past racial injury and its continuing significance can enter into our discourses and, in doing so, complicate the
assignment of responsibility. To repeat: given our history, producing a racially defined nether caste through the ostensibly neutral application of law should be profoundly offensive to our ethical sensibilities.

**Ultimately, We’re All in the Same Fix**

Several summers ago, I took some time to read the non-fiction writings of the great 19th century Russian novelist, Leo Tolstoy. As you may know, he became an eccentric pacifist and radical Christian critic at the end of his life. I was stunned at the force of some of his arguments (though, I confess, I was not entirely persuaded on one of his key points -- that a true Christian must be absolutely celibate!)

What struck me most was Tolstoy’s provocative claim that the core of Christianity lies in Jesus’ Sermon on the Mount: You see that other fellow committing some terrible sin? Well, Jesus preaches, if you have ever lusted, or allowed jealousy, or envy or hatred to enter your own heart, then you are to be equally condemned! This, Tolstoy claims, is the central teaching of the Christian faith: namely, that we’re all in the same fix!

Now, you mustn’t worry that I am about to launch into my own sermon on this august occasion. Still, it seems to me that this religious sentiment is quite relevant to the problem at hand: while the behavioral pathologies and cultural threats that we see in society -- the moral erosions “out there” -- the crime, drug addiction, sexually transmitted disease, idleness, violence and all manner of deviance -- while these are worrisome, nevertheless, our moral crusade against these evils can take on a pathological dimension of its own. We can become self-righteous, legalistic, ungenerous, stiff-necked, and hypocritical. We can fail to see the mote in our own eye. We can neglect to raise questions of social justice. We can blind ourselves to the close relationship that actually exists between, on the one hand, behavioral pathology in the so-called urban underclass of our country and, on the other hand, society-wide factors — like our greed-driven economy, our media-encouraged worship of the self, our endemic culture of materialism, our vacuous political discourses, our declining civic engagement, and our aversion to sacrificing private gain on behalf of much needed social investments. We can fail to see, in other words, that the problems of the so-called underclass are but an expression, at the bottom of the social hierarchy, of a more profound and widespread moral deviance – one that involves all of us.

Taking this position does not make me a moral relativist. I merely hold that, when thinking about the lives of the disadvantaged in our society, the fundamental premise that should guide us is that we are all in this together. *Those* people languishing in the corners of our society are *our* people – they are *us* – whatever may be their race, creed or country of origin, whether they be the crack-addicted, the AIDS-infected, the mentally ill homeless, the juvenile drug sellers, or worse. Whatever the malady, and whatever the offense, we’re all in the same fix. We’re all in this thing together. *This* is the point that Tolstoy, and Jesus before him, were making. And, this is the point is wish to urge upon your consideration at this moment.

Just look at what we have wrought. We Americans have established what, to many an outside observer, looks like a system of racial caste in the center of our great cities. I refer here to millions of stigmatized, feared, and invisible people. The extent of disparity in the opportunity to achieve their full human potential, as between the children
of the middle class and the children of the disadvantaged – a disparity that one takes for granted in America – is virtually unrivaled elsewhere in the industrial, advanced, civilized, free world.

Yet, too many Americans have concluded, in effect, that those languishing at the margins of our society are simply reaping what they have sown. Their suffering is seen as having nothing to do with us – as not being evidence of systemic failures which can be corrected through collective action. Thus, as I noted yesterday afternoon, we have given up on the ideal of rehabilitating criminals, and have settled for simply warehousing them. Thus, we accept – despite much rhetoric to the contrary – that it is virtually impossible effectively to educate the children of the poor. Despite the best efforts of good people and progressive institutions – despite the encouraging signs of moral engagement with these issues that I have seen in my students over the years, and that give me hope – despite these things, it remains the case that, speaking of the country as a whole, there is no broadly-based demand for reform – no sense of moral outrage, no anguished self-criticism and public reflection – in the face of this massive, collective failure.

I am reminded in this context of Bill Cosby's widely reported criticisms of the black poor (“the lower economic types are not living-up to their end of the deal," Cosby has famously said.) Cosby has been praised for his willingness to speak frankly about such matters. But, courage, forthrightness and the willingness to "call a spade a spade" are no substitutes for a cogent social analysis. Bill Cosby’s well-publicized lament is long on outrage, but it falls woefully short of being a cogent moral argument or a serious guide to action. Black leaders, he insists, should stop making excuses for these “knuckleheads,” and should speak to today’s youth more forthrightly, insisting on the continued relevance of tried and true virtues like self-restraint, self-discipline, hard work and moral decency.

Saying all of this may be fair enough, but it is quite far from being enough said. A serious treatment of this issue would have to look beyond the "culture" of black ghetto dwellers and the public performances of so-called “black leaders,” so as to reckon with larger social, political and economic developments that have taken place in American society. Any serious prescription for how to cure what ails ghetto America must look beyond the rhetoric of an aging civil rights leadership class to consider why the social policy making process at all levels of government has failed to respond more helpfully to the increasingly dire condition of our nation's most disadvantaged persons. That is to say, morally and intellectually serious engagement with these problems must deal not just with THEM, but also with US.

So much would seem, upon reflection, to be obvious. I am aware of no instance in human history where behavioral change on the scale that would be needed to reverse the dire straights of America's urban poor has come about through the exhortation of communal leaders. Social pathologies were not unknown in the European immigrant communities of early twentieth century America. Scholars (like the Columbia University historian, Ira Katznelson [2005]) seem to agree that they were overcome not by hortatory campaigns of behavioral reform, but rather by massive social policy innovations like the New Deal and the G.I. Bill. If it is obvious that some blacks need to clean up their acts, it is equally obvious that curing what ails America's racial ghettos will require the active involvement, through our public institutions, of all Americans. Where Cosby castigates black leadership for failing to instill proper values in our youth, I am left wondering, as I
look back on the debacle that was Hurricane Katrina, where is American leadership on this problem?

The core of the problem, I have suggested, is that the socially marginal are not seen as belonging to the same general public body as the rest of us. At least implicitly, our political community acts as though some are different from the rest and, because of their culture -- because of their bad values, their self-destructive behavior, their malfeasance, their criminality, their lack of responsibility, their unwillingness to engage in hard work -- they deserve their fate.

Lest you think I exaggerate, here is what the current domestic policy advisor to President George W. Bush – a man named Karl Zinsmeister – wrote in the pages of the American Enterprise Magazine some years ago (Zinsmeister 1996):

The troubling reality in our ghettos today is that the hellish torments are being inflicted by their own residents. If only some identifiable outside force were creating the siege conditions, nearly any American would gladly swing a battle-axe against such an enemy. But the harder, more tragic reality is that inner-city Americans are being brutalized by their own neighbors, their own reproductive partners, their own teenagers, their own mothers even. And ultimately, by themselves. Who is forcing the crack pipe between those many lips?

What we in the United States have failed to recognize – not merely as individuals, I stress, but as a political community – is that these ghetto enclaves and marginal spaces of our cities are products of our own making: Precisely because we do not want those people near us, we have structured the space in our urban environment (as was revealed by the aftermath of Hurricane Katrina in New Orleans in 2005) so as to keep THEM away from US. Then, when they fester in their isolation and their marginality, we tut-tut-tut, hypocritically point a finger, saying in effect: “Look at those people. They are a threat to the civilized body They must be expelled, imprisoned, controlled.” It is not WE who must institute responsible social reforms, but THEY who must embrace personal responsibility. It is not we who must set our social affairs aright, but they who must get their individual acts together. This posture, I wish to suggest, is inconsistent with the attainment of a just distribution of benefits and burdens in society if we look at the problem of social justice in terms of a rational deliberation conducted from behind Rawls’s veil of ignorance.

So, I have endeavored to develop an argument in this spirit, applying the rational choice-style of reasoning that is natural to an economist. If you like, I am “splitting the difference” between Drs. Rationalist and Functionalist in that opening dialogue. Using a series of “models” – which, after all, is the hammer that comes in an economic theorist’s tool kit so, for my sake, this darned problem had better be a nail (!) -- I tried to illustrate the limits of an ethic of personal responsibility by graphically highlighting the extent to which a person’s actions, including their law-breaking actions, are shaped by social forces beyond that person’s control.

Civic inclusion has been the historical imperative in Western political life for 150 years. And yet – despite our self-declared status as a light unto the nations, as a beacon of hope to freedom-loving peoples everywhere – despite these lofty proclamations which were belied by images from the roof tops in flooded New Orleans in September 2005,
and are contradicted by our overcrowded prisons – the fact is that this historical project of civic inclusion is woefully incomplete in these United States.

The futility of pursuing civic inclusion has been declared by reactionary political forces, like those represented by Karl Zinsmeister, at every step along the way. Yet, in every instance, these forces have been proven wrong. At one time or another, the goal has been derided of including women, landless peasants, former serfs and slaves, or immigrants more fully in the civic body. Extending to them the franchise, educating their children, providing health and social welfare to them has always been controversial. But, this has been the direction in which the self-declared “civilized” and wealthy nations have been steadily moving since Bismarck, since the revolutions of 1848 and 1870, since the American Civil War with its Reconstruction Amendments, since the Progressive Era and through the New Deal on to the Great Society. This is why we have a progressive federal income tax and an estate tax in this country, why we feed, clothe and house the needy, why we (used to) worry about investing in our cities’ infrastructure, and in the human capital of our people. What those images out of New Orleans in 2005 showed is that this American project of civic inclusion remains incomplete. Nowhere is that incompleteness more evident than in the prisons and jails of America. And, this as yet unfulfilled promise of American democracy reveals a yawning chasm between an ugly American reality, and our nation’s exalted image of herself.
References:

Becker and Tomes 1979
*Journal of Economic Perspectives* 16:3, pp. 3-30.


Brock and Durlauf 2001


Galor and Zeira 1993


Mookherjee and Ray 2003


Tilly, Charles. 1998. *Durable Inequality*. ??


Weaver, Vesla M. “Dark Prison: Race, Rights, and the Politics of Punishment.” (draft Ph.D. dissertation, Department of Political Science, Harvard University)


Zinsmeister, Karl. 1996. [his American Enterprise piece]