# "Dysfunctional Identities" Can Be Rational

*By* Hanming Fang And Glenn C. Loury*

January 9, 2005

Understanding the nature and sources of human identity is an important objective in the study of a variety of social problems. Scholarly and popular writing on the cultural determinants of economic disadvantage underscores this point. Some analysts (e.g., Edward Banfield 1970, Thomas Sowell 1994, John McWhorter 2000, and John Ogbu 2003, to name a varied few) have hypothesized that a causal connection exists between the poor social performance of a group of people and their "culture." That disadvantaged people harbor "dysfunctional" notions about identity has been offered as an explanation of a group's welfare dependency, or its low academic proficiency. It has been said, for instance, that people fair poorly because they focus overly much on their own victimization, or because they disassociate themselves from their more successful fellows, etc.

At the root of such cultural criticism lies the presumption that the disadvantaged should "reform" themselves: If those people would only see themselves differently, the critics hint, they could be so much better off. This mode of social explanation easily accommodates racial overtones. With the present paper we intend to raise serious doubts about such normative criticisms of the poor when applied to their conceptions of identity. We show that the identities adopted by a group of people can be perfectly consistent with rational individual choices, even though feasible alternative configurations may exist under which everyone would be better-off. Indeed, we argue that identity choice by interactive agents with ongoing economic relations has a "tragedy of the commons" quality about it: the profile of dominant strategies for the agents can yield a Pareto inferior collective outcome. Preaching "identity reform" to such people is a bit like trying to counter an over-fishing problem by lecturing fishermen on the moral need for forbearance!

We wish to be explicit and clear at the outset about what we have in mind when using the term "identity." Human identity includes both a personal and a social aspect. Social identity deals with how an individual is perceived and categorized by others (e.g., Erving Goffman 1963). In contrast, personal identity – which is the subject of this paper, and which psychologists sometimes call "ego identity" – deals with a person's answer to the question: "Who am I?" Our proposed model of personal identity posits that, to answer this question, an agent must provide a "narrative" about her personal history.

That is, she has to summarize her life experiences. Because a full personal history is (necessarily) a *very* complex object, and since their *cognitive capacities* are limited, answering the "Who Am I?" question requires agents to project elaborate personal accounts onto manageable categories of self-description. We think of an agent's identity as the mechanism she uses to convert complex personal history into a more simplified account of herself. A group's "collective identity" is any self-representional mode of this sort which has been adopted in common by (most of) the agents in that group. We formalize the problem of selective self-representation, and use the resulting framework to study the efficiency implications of the "identity" choices people make. This, we believe, is one way that economic analysis can contribute to the study of identity-related issues.[1]

To the extent that self-representation affects subsequent economic transactions, a rational agent chooses her identity to maximize the payoff from such transactions. Following Fang and Loury (2004), we embed our notion of identity in a particular economic transaction – repeated income risk sharing. More specifically, we consider a two-stage game in which identity choices are made in the first stage, and agents engage – more or less remuneratively – in an infinitely repeated income-risk-sharing game in the second stage. Given this framework, we say that a collective identity has been adopted when, in subgame perfect equilibrium, individuals make the same first stage identity choices. We show under this set-up that a group of people may rationally elect to embrace a way of thinking about themselves that inhibits their economic functioning. We refer to such an inefficient collective identity choice as "dysfunctional."

The key intuition highlighted by our approach is that embracing an identity is a social event, not merely the expression of an individual's values or preferences. In particular, autonomous agents who interact frequently may end-up adopting similar categories of self-representation because they think this leaves them better placed to manage their collective action problems. When this is so, different contexts of social interaction can foster different equilibrium identity configurations, and agents interacting within relatively closed social networks may be inclined to embrace the same or similar identities. Moreover, our analysis makes clear why there is no reason to expect the common categorical maps (collective identities) settled-upon by rational agents to be socially efficient.

## I. A Model of Identity Choice

**<u>Preliminaries.</u>** In the model to be presented here, agents need to "talk" about their personal experiences before realizing potential gains from trade. How they elect to represent themselves to one another affects the productivity of their subsequent economic interactions. We show how a "bad" (dysfunctional, victim-based, oppositional) collective identity can be sustained in equilibrium for one group of people and not another, notwithstanding the fact that the "values" of people in the two groups are similar. And, we illustrate why it can be difficult to shift such a problematic pattern of personal identifications using only a marginal intervention: Beneficial tacit arrangements may have evolved among the agents, the viability of which turns on their embrace in common of the prevailing identity convention.

So, imagine that two agents, indexed by $i = 1, 2$, engage in a two-stage game of identity choice and repeated risk sharing.[2] In the first stage of play, acting simultaneously, each agent makes a *once-for-all* choice of "identity." In every one of the infinite sequence of periods that constitutes the second stage, the agents receive random income endowments which they might agree to share with one another. We assume that income is perishable and cannot be stored; we also assume that random endowments are independent and identically distributed both across agents and across periods. Let $y \in Y$ denote the realization of an agent's stochastic income endowment in any second stage period, where $Y$ is a subset of $R_+$. Endowment $y$ is realized with probability $p(y) > 0$, such that $\sum_{y \in Y} p(y) = 1$. An agent's endowment realization is private information in each period.

Both agents seek to maximize their expected discounted utility from consumption over the course of the second stage. They have a common per-period utility function $u : R_+ \to R$ which is continuous, three-times differentiable, and satisfies $u' > 0$, and $u'' < 0$. Agents discount the future at a common, constant rate $\delta < 1$. Notice that, in contrast with Akerlof and Kranton (2000), "identity" is not a direct argument of an agent's utility function in our formulation. Given these preferences, consumption fluctuations are undesirable. So, gains from trade are available to the agents if they can arrange to make interpersonal income transfers in an ongoing manner over the course of the second stage. This is their collective action problem. Because their second stage interactions are repeated, by making their future dealings contingent on current behavior agents can exert leverage to enforce compliance with some agreed upon risk sharing scheme. We define a *risk sharing arrangement* to be any agreement obligating the agents to make and receive interpersonal transfers to and from one another in some specified manner. We investigate how their choices about identity affect the agents' risk sharing prospects.

"Identity" is modelled as follows: Although each agent's endowment is invisible to the other, publicly observable "indicators" are available in each period through use of which an agent must signal her endowment. Let $X$ be the set of all indicators, with $x \in X$ a generic element. We assume $|X| << |Y|$ – that is, $X$ is a "much smaller" set than $Y$. This captures the idea that there are many fewer indicators than there are income states. Thus, an agent's actual endowment realization, $y$, can be interpreted as her full experience. And, the notion that it is practically infeasible for an agent to fully describe all aspects of her experience is captured by our requirement that, in every second stage period, each agent makes a public "representation" of her income $y \in Y$ by "announcing" an indicator $x \in X$. The actual making of these announcements ought not to be thought of as a strategic act. Instead, we envision a situation where, once agents have entered the second stage, the signals given-off about their endowments are issued involuntarily, according to some formula or "code" adopted in the first stage of play. In effect, the agents use these indicators to construct a "narrative" about their (income) experience.

The crucial step in our analysis is to assume that any second-stage income risk sharing undertaken by the agents must be implemented using these "income narratives." That is, consumption smoothing transfers between the agents can depend only on what is common knowledge between them – namely, their indicators, not their endowment realizations. Thus, in this two-stage game, first-stage identity choices bind the agents to noisily signal their income realizations to one another in a particular manner. And this, in turn, limits the extent of income risk sharing and associated expected utility payoffs which the agents can achieve over the course of the second stage.

A function mapping the set of incomes into the set of indicators, $C : Y \to X$ is called an *identity code*. We restrict attention to *monotonic* codes. Under a monotonic code there is a way to assign numbers to indicators such that higher numbers invariably connote higher endowments. Formally, a code $C : Y \to X$ is *monotonic* if, for every $\{y, y', y''\} \subset Y : C(y) = C(y')$ and $\min\{y, y'\} < y'' < \max\{y, y'\}$ implies $C(y'') = C(y)$. Let $C_i$ denote the first stage choice of a monotonic code by agent $i$. We refer to the pair $(C_1, C_2)$ as a *code configuration*.

A *risk sharing arrangement* is a way to move resources between agents that depends on what they have to "say" to each other about their incomes, not the incomes themselves. We say that such an arrangement is *feasible under a given code configuration* if it can be supported as a subgame perfect equilibrium continuation for the infinitely repeated interactions in the second stage. To keep things simple, we assume that second-stage agents only consider period-stationary risk sharing arrangements. That is, they restrict attention to those arrangements where transfers depend, in the same manner each period, on that period's indicators alone.[3]

Thus, we have a dynamic game in two stages, with the second stage extending over an infinite sequence of periods. We assume that the agents play non-cooperatively, and that their first stage identity choices are common knowledge when they enter the second stage. The time line of the model is as follows: Before all interactions start, both agents choose their identities. After observing each other's choices in this regard, they engage in an infinitely repeated risk sharing interaction. We adopt subgame perfection as an equilibrium concept.

Given this set-up, the analysis might proceed by working backwards in two steps: First, fixing the identity code configuration, we would derive the agents' discounted sums of expected utility associated with some feasible transfer arrangement chosen by them in the second stage continuation. Then, we would study first stage code choice as equilibrium behavior in a symmetric, two-player, normal form game, where actions for player $i$ are the alternative codes $\{C_i : Y \to X\}$, and payoffs are the agents' discounted expected utility levels in the implied continuation. Obviously, many feasible second-stage continuations are possible for each code configuration. So, to pursue this two-step program we need to associate a *unique* second-stage welfare level for the agents with each configuration, thereby specifying how the expected utility surplus (relative to autarky) generated by the prospect of risk sharing is to be divided among agents. *Accordingly, we assume that, given their first-stage choices of identity codes, second-stage agents select a feasible, stationary risk sharing arrangement so as to maximize the sum of their expected discounted utilities.*[4]

Now, it is obvious that the maximal punishment available for deviators from any proposed risk sharing arrangement is a reversion to autarky. And in light of the uniform discounting, if no one-shot deviation from a proposed arrangement is beneficial, taking the ensuing punishment into account, then neither can any finite or infinite sequence of deviations be beneficial. Let $t \in \Re$ denote a (possibly negative) transfer from agent one to agent two. The discussion to this point motivates the following formal definitions:

**Definition 1** *A* ***risk sharing arrangement*** *is a period-stationary function, $T : X^2 \to \Re$, such that whenever the agents' signals are $(x_1, x_2)$, the income transfer between agents is given by: $t = T(x_1, x_2)$.*

**Definition 2** *A risk sharing arrangement $T$ is* ***feasible under a given code configuration*** *if, for both agents $i = 1, 2$, in every second stage period and for all possible income realizations $(y_1, y_2) \in Y \times Y$, no net gain is anticipated for a one-shot deviation from the arrangement that is followed by a reversion to autarky.*

**The Special Case: $|X| = 2$.** Fang and Loury (2004) study this dynamic game in some generality. In this paper, in order to illustrate the main ideas and for the sake of concreteness, we will investigate a special case which is already sufficiently rich to capture the key trade-off at work in our model. Suppose that only two indicators are available: $X = \{B, G\}$. Then, in effect each second stage period involves the agents involuntarily signalling to one another whether that period's endowment realization has been "good" or "bad." Subsequent transfers between the agents must be based on these binary signals. With $|X| = 2$, to choose a code $C$ is necessarily to partition the endowment space into realizations with "good" and with "bad" signals: $Y = C^{-1}(B) \cup C^{-1}(G)$. Moreover, monotonic codes are always of the threshold form: {for some $y^* \in Y$, $C(y) = B$ if and only if $y \leq y^*$}. To choose a code is thus to decide both about *the frequency of* and *the disparity between* good and bad announced indicator states. As we shall see, there is reason to think that decentralized choices of this kind made by rational agents will generally not be Pareto efficient. That is, there is reason to suppose that the identity configurations emergent in decentralized equilibrium will generally be dysfunctional.

To see the key trade-off at work here, the following two observations are useful: First, notice that the more widely disparate are the agents' endowment states associated with a given indicator pair, the more profitable are their risk sharing trades conditional on those signals. Secondly, observe that the more frequent are the encounters between unequally endowed agents, the greater are their opportunities to engage in profitable risk sharing. Hence, two traits of a code configuration – which we refer to as "mismatch frequency" and "endowment disparity" – are socially desirable. When $|X| = 2$, both traits are simultaneously determined by the choice for each agent of a dividing line between "good" and "bad" endowments, $y_i^*$, $i = 1, 2$. Therefore, in the neighborhood of an optimal choice, one of these desiderata is being traded-off against the other at the margin.

## II. Three Endowment Realizations

To fix ideas, suppose further that only three endowment realizations are possible: $y \in Y = \{l, m, h\}$, $l < m < h$. Denote the endowment probabilities $p(y)$ by $p_l, p_m$ and $p_h$ respectively, where $\sum_{k \in \{l,m,h\}} p_k = 1$. A code is simply a map, $C : \{l, m, h\} \rightarrow \{G, B\}$. Under monotonicity, and without further loss of generality, we can restrict attention to the codes $C^P$ ("pessimistic") and $C^O$ ("optimistic"), where: $C^P(l) = C^P(m) = B$, $C^P(h) = G$; and $C^O(m) = C^O(h) = G$; $C^O(l) = B$. Thus, with two indicators and three endowment levels, only three code configurations are possible in this two-person society: either both are "pessimists" $\langle C^P, C^P \rangle$; or both "optimists" $\langle C^O, C^O \rangle$; or the codes are mixed $\langle C^P, C^O \rangle$ or $\langle C^O, C^P \rangle$. In each second stage period the agents' incomes $y_i \in \{l, m, h\}$ are mapped to their signals $x_i \in \{B, G\}$ via one of the two codes, so:

$$x_i = C_i(y_i), \text{ for } C_i \in \{C^P, C^O\}, \ i \in \{1, 2\}.$$

Risk sharing transfers are then carried out in each period according to that period-stationary function of the announced indicators, $T(x_1, x_2)$ which maximizes the sum of the agents' payoffs.

**Identity Choice in a $2 \times 2$ Normal Form.** With these conventions in hand, we can characterize first-stage play with a reduced $2 \times 2$ normal form game as follows:[5]

Agent 2

| Agent 1 | | $C^P$ | $C^O$ |
|---|---|---|---|
| | $C^P$ | $V_P^*, V_P^*$ | $V_M^{P*}, V_M^{O*}$ |
| | $C^O$ | $V_M^{O*}, V_M^{P*}$ | $V_O^*, V_O^*$ |

Our interpretation of this example goes like this: the indicators $x \in X$ connote that an agent has experienced either a "good" or a "bad" realization, although in actuality endowments can be either "high," "medium" or "low." So, given the requirement of monotonicity, an agent's choice of "identity" amounts to a decision about how to code an intermediate income realization (whether to react as if this were a "good" or a "bad" event.) One way to talk about this is that, in effect, the agents must choose between being "pessimists" or "optimists." Alternatively, we could envision them as deciding whether, in the event of a middling endowment realization, to view themselves as a "victim" – that is, as someone who needs a helping hand but who is not in position to lend one.[6] Whatever the interpretation, we can ask whether the "optimistic" configuration $\langle C^O, C^O \rangle$ is better than the "pessimistic" one $\langle C^P, C^P \rangle$, in terms of the potential gains from second stage risk sharing that it engenders. And, we can inquire whether a mixed configuration – $\langle C^P, C^O \rangle$, say – is inferior to either "collective identity."[7]

Thus, this example permits us to discuss our ideas about dysfunctional collective identities using the basic notions of elementary game theory. If the normal form depicted above is a coordination game (i.e., $V_P^* > V_M^{O*}$ and $V_O^* > V_M^{P*}$), then strategic forces favor the adoption of *some* collective identity and multiple, Pareto-ranked equilibria exist. Avoiding a dysfunctional identity then becomes a coordination problem for the agents. Alternatively, if this game is a Prisoners' Dilemma (i.e., $V_M^{P*} > V_O^* > V_P^* > V_M^{O*}$, for instance, so that, although a pessimistic configuration is Pareto inferior to an optimistic one, it is nevertheless a dominant strategy for the agents to be pessimistic), then the two-stage strategic interaction has a "tragedy of the commons" quality about it, and the adoption by rational agents of a dysfunctional identity is all but guaranteed!

**An Illustrative Numerical Analysis.** We can illustrate these ideas by calculating equilibrium identity choices for the agents in this $3 \times 2$ example under the assumption that the utility function, $u(\cdot)$ belongs to the constant relative risk aversion (CRRA) family:

$$u(y) = \frac{y^{1-\rho}}{1-\rho}, \text{ with } \rho \in (0, 1]$$

(when $\rho = 1, u(y) \equiv \ln y$. Note that CRRA utilities satisfy $u''' > 0$.) Given this utility function, the outcome in our model is determined by the distribution of random incomes, the discount factor, $\delta$, and the risk aversion parameter, $\rho$. We will examine how equilibrium identities chosen in the first stage depend on the degree of risk aversion. We do this by calculating numerically the second stage continuation values under autarky, under the "optimistic" and "pessimistic" collective codings, and under the mixed coding. We then examine how these continuation values vary with the changes in the parameter, $\rho$. Our results are summarized in Figure 1.

[Insert Figure 1 About Here]

As $\rho$ gets larger, the agents become more risk averse, which means (of course) that risk sharing becomes more valuable to them, other things equal. Thus, one way to interpret this exercise is to identify an increase in the risk aversion parameter with a raising of the stakes for the agents in their second stage interactions. The equilibrium coding depends on the relative value of $V_M^{O*}$ as compared with $V_O^*$ and $V_P^*$. Figure 1 shows the differences between $\left(V_P^*, V_O^*, V_M^{P*}, V_M^{O*}\right)$ and the autarky value $V_A$ as $\rho$ varies. Note the figure depicts a threshold $\rho^*$ such that when for any $\rho \in (0, \rho^*)$, we have the following inequalities:

$$V_O^* - V_A > V_M^{P*} - V_A; \quad V_P^* - V_A > V_M^{O*} - V_A.$$

The first inequality implies that $\left\langle C^O, C^O \right\rangle$ is an equilibrium. Likewise, the second inequality implies that if $\left\langle C^P, C^P \right\rangle$ is an equilibrium. Therefore when $\rho \in (0, \rho^*)$, we have *multiple* collective identities which are equilibria . Moreover, these identities are Pareto-ranked: the "optimistic" equilibrium $\left\langle C^O, C^O \right\rangle$ dominates the "pessimistic" one $\left\langle C^P, C^P \right\rangle$. By contrast, when $\rho > \rho^*$, Figure 1 shows that

$$V_O^* - V_A < V_M^{P*} - V_A, \text{ but } V_P^* - V_A > V_M^{O*} - V_A.$$

Therefore, the unique (dominant strategy) equilibrium is the "pessimistic" collective identity $\left\langle C^P, C^P \right\rangle$. It is worth noting that if the agents could commit to choose the "optimistic" coding, they would both be better-off than at the equilibrium, since $V_O^* > V_P^*$. The "optimistic" coding does not constitute an equilibrium due to forces familiar from the "Prisoner's Dilemma."

### III. A Continuum of Endowment Realizations

We can readily extend this example, since the assumption of three discrete income realizations, though allowing a colorful interpretation, is incidental to the analysis. When $Y$ is an interval of real numbers and $|X| = 2$, the reduced-form game involves the agents simultaneously choosing thresholds $(y_1^*, y_2^*)$ in the first stage, and reporting a "bad" outcome whenever their endowments are at or below the chosen thresholds.[8] This continuum specification is useful because, since the set of alternative thresholds is a bounded interval, and the agents' payoffs are differentiable functions of the threshold pair (assuming a well-behaved endowment distribution), we can use calculus to study the agents' strategic interaction in the first stage.

This specification can also be used to show why inefficient collective identity choices are to be expected: *The private evaluation of benefits and costs associated with alternative code configurations is likely to differ from a social assessment.* Two countervailing factors can cause private and social valuations to differ in our model:

(i) One factor involves *endowment disparity.* When contemplating the choice of a higher (more pessimistic) threshold in the first stage of play, an individual (agent one, say), takes into account that the second stage transfer policy will become marginally less attractive for her because raising her threshold makes her endowment distribution more favorable at every indicator pair,thereby lowering the transfer she receives, or raising the transfer she gives. *But this private cost to agent one is not* *a social cost.* Because the chosen risk sharing arrangement maximizes the sum of agents' welfare the Envelope Theorem implies that, in the neighborhood of an optimal configuration, the net social impact of an induced shift in the transfer arrangement is zero. So, due to this pecuniary externality, agent one may tend to set $y_1^*$ *below* its socially optimal level.

(ii) The other factor involves *mismatch frequency.* Since agent one's likelihood of giving a transfer declines as $y_1^*$ rises, raising her threshold has a negative effect on her trading partner. *But this social cost is not a private cost to agent one.* When choosing their thresholds, each agent ignores this impact on the other agent. So, due to this external diseconomy, agent one may tend to set $y_1^*$ *above* its socially optimal level.

In general, how the equilibrium and the socially optimal configurations compare depends on the relative magnitude of these two wedges between private and social valuation. In particular, the symmetric equilibrium threshold will exceed the socially optimal level if, when considering a marginal increase in $y_1^*$, the external diseconomy on agent two due to agent one's lowered frequency of giving a transfer [specified in (ii) above] exceeds the pecuniary externality on agent one due to the induced decline in her net transfer receipts [specified in (i).] But, using the Envelope Theorem again, any induced negative impact on agent one is just offset by an induced positive impact on her trading partner. We conclude that the equilibrium threshold exceeds the socially optimal one (too much pessimism!) if the direct plus the induced impact on agent two of a marginal increase in agent one's threshold is negative.

We can make this point somewhat more formally. With $X = \{B, G\}$ and $Y$ an interval on the non-negative real line, denote by $U(y_1, y_2)$ player one's payoff at the threshold pair, $(y_1, y_2)$, and let $W(y) = U(y, y)$. Moreover, let $U_i$ be the partial derivatives of $U$ with respect to $y_i$, $i = 1, 2$; let $y^e = y_1^* = y_2^*$ be the agents' common threshold in a symmetric equilibrium; and, let $y^o$ be the socially optimal (sum-of-discounted–utility-maximizing) common threshold. Then, we have the first-order conditions: $U_1(y^e, y^e) = 0$, and $W'(y^o) = U_1(y^o, y^o) + U_2(y^o, y^o) = 0$. It follows that $U_2(y^e, y^e) \lesseqqgtr 0$ implies $W'(y^e) \lesseqqgtr 0$ which, in turn, implies $y^e \gtreqqless y^o$ (assuming the relevant second-order condition.)

We conclude (in the context of this example) that the following Proposition holds:

**Proposition 1** *The symmetric equilibrium identity configuration is a spoiled collective identity involving too much pessimism (too much optimism) whenever the net effect of raising one agent's threshold marginally from its equilibrium level is to reduce (increase) the payoff of the other agent!*

Thus, the kinship of the "identity coordination problem" being posed here with the classical "tragedy of the commons" is easy to see in the case $|X| = 2$.

## IV. Conclusion

This paper has presented a novel choice-theoretic model of "identity" based on the notions of *categories* and *narratives*. Identity is conceived as a matter of "reflexive perception" – how people understand themselves. Choosing an identity is equivalent to making a generalization about one's past that highlights the most salient aspects of experience. When many individuals make a common choice in this regard, they embrace a *collective identity*. We embed this conceptualization of identity into a repeated risk sharing game. We show, in the leading example of this model, that different collective identities permits different degree of risk sharing among agents, but that rational agents acting independently of one another will generally not settle upon the most effective identity configuration.

Our theory (but also, common sense) emphasizes that "identity" is endogenous, and is shaped by social contacts. So, the question arises: What kind of social networks in which people might be embedded lead to what kinds of choices about identity? This is a particularly interesting question for someone studying race, culture and social inequality in the U.S. One implication of our theory, in a slightly expanded model allowing for the assortative matching of agents from distinct groups before playing the second stage repeated game, is that distinctive patterns of identity choices by individuals in distinct groups is more likely if patterns of social interaction are more group-segregated. This leads us to speculate that anyone who believes "culture" is important in sustaining racial inequality in a society like the US should look seriously at the linkages between identity and social integration. Casual empiricists make much of the observable differences in "values" between distinct groups. But, our analysis points toward a recognition of the fact that such cultural difference can be parasitic upon pre-existing disparities in the structures of social interaction. If group inequality is partly due to cultural differences, if cultural variation is partly a matter of distinct identity choices, and if identity choices diverge in part because of segregated social networks, then social *integration* of some sort might be an antidote for inequality.

## REFERENCES


**Akerlof, George and Kranton, Rachel.** "Economics and Identity." Quarterly Journal of Economics, 2000, 115(3), pp. 715-753.

——. "Identity and Schooling: Some Lessons for the Economics of Education." Journal of Economic Literature, 2002, 40(4), pp. 1167-1201.

**Banfield, Edward.** The Unheavenly City: The Nature and Future of Our Urban Crisis. Boston: Little Brown, 1970.

**Fang, Hanming and Loury, Glenn C.** "Toward an Economic Theory of Dysfunctional Identity." Cowles Foundation Discussion Paper No. 1483, 2004.

**Fryer, Roland and Jackson, Matthew.** "A Categorical Model of Cognition and Biased Decision-Making" unpublished manuscript, Harvard University, 2004.

**Goffman, Erving.** Stigma: Notes on the Management of Spoiled Identities. Englewood Cliffs, NJ : Prentice-Hall, 1963.

**Thomas, Jonathan and Worral, Tim.** "Self-Enforcing Wage Contracts." Review of Economic Studies, 1988, 55(4), pp. 541-553.

**Loury, Glenn C.** The Anatomy of Racial Inequality. Cambridge, MA: Harvard University Press, 2002.

**McWhorter, John.** Losing the Race: Self-Sabatoge in Black America. New York: The Free Press, 2000.

**Ogbu, John.** Black Students in Affluent Suburbs: A Study in Academic Disengagement. Mahwah, NJ: Lawrence Erlbaum, 2003.

**Sowell, Thomas.** Race and Culture: A World View. New York: Basic Books, 1994.

## FOOTNOTES

* Fang: Department of Economics, Yale University, 37 Hillhouse Avenue, New Haven, CT 06520-8264; Loury: Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215.

[1] This conceptualization expands upon ideas about racial classification, social cognition and identity introduced in Loury (2002, chapter 2). Our model has also been inspired by the related work of Roland Fryer and Matthew Jackson (2004). The present approach may be contrasted with that of George Akerlof and Rachel Kranton (2000, 2002), who offer what might be called a utilitarian theory of identity: that is, they posit a non-standard utility function meant to incorporate the value of conforming to the norms and expectations associated with a decision-maker's exogenously assigned social position.

[2] We focus on this two-agent setup for simplicity. It is more intuitive, however, to envision agents as living in a community, and as being randomly paired with one another in each second stage period when endowments might be shared. Assuming that the outcome from such pairings is public information, it would be a straightforward exercise to extend our model in this way.

[3] This restriction to stationary arrangements, while keeping things tractable, definitely entails some loss of generality. Jonathan Thomas and Tim Worral (1988) have shown that an optimal self-enforcing risk sharing contract in this setting generally entails non-stationarity. Since our focus here is on identity choice, and not risk sharing *per se*, we think this restrictive assumption is acceptable under the circumstances.

[4] While other methods of surplus-splitting can be imagined (e.g., Nash Bargaining), our assumption here seems quite plausible. Given the *ex ante* symmetry of this setting, rational agents viewing the surplus division problem from behind a 'veil of ignorance' well might agree to adopt the equilibrium selection method we have proposed.

[5] Here we are using the obvious notation: $V_M^{O*}$ is the payoff to the optimistic agent under a mixed configuration, while $V_P^*$ is either agent's payoff under a pessimistic configuration, etc.

[6] On this interpretation the example permits us to ask, in the habit if not in the spirit of McWhorter (2000), whether an expansive sense of one's victimization constitutes a "dysfunctional collective identity!"

[7] Stating this more provocatively, the example permits us to investigate whether the agents spread their joint income risks more effectively when they embrace a common "narrative of victimization!"

[8] It is natural here, in keeping with the intuition from the $3 \times 2$ case, to associate a higher threshold $y_i$ with a "more pessimistic" identity choice by agent $i$ (or, with the agent adopting a "more expansive sense of her victimization"), since a higher threshold makes it less likely that a "good" signal is announced.

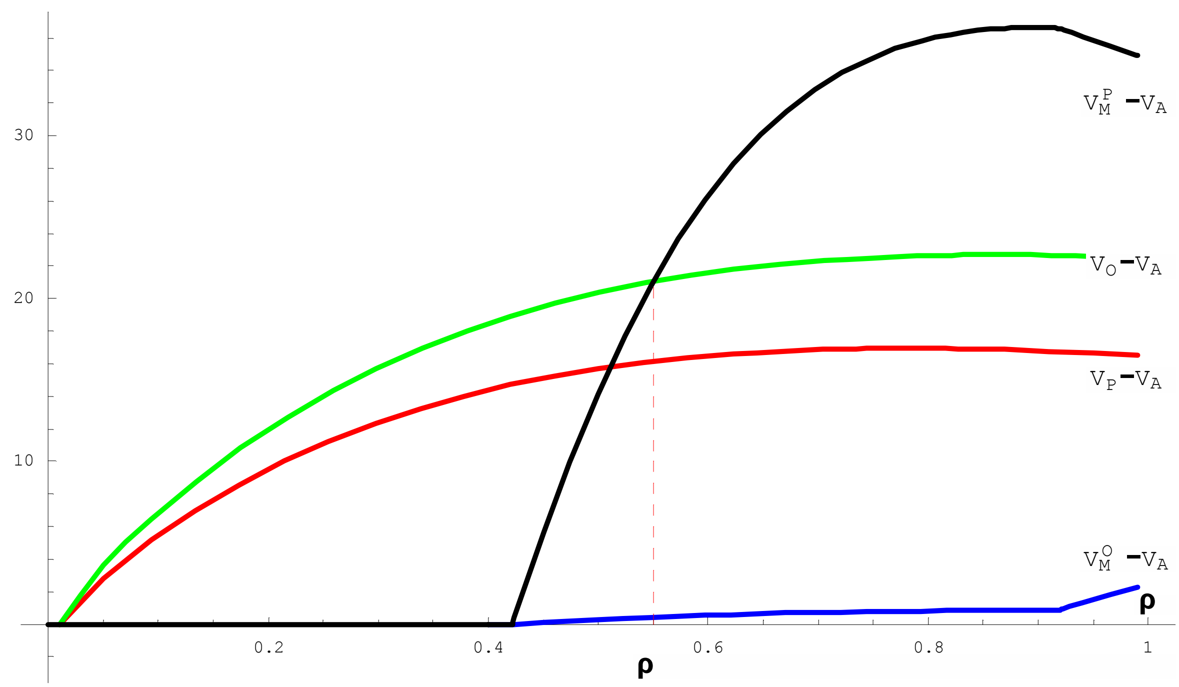


FIGURE 1: THE VALUE DIFFERENCES FROM THE AUTARKY VALUE $V_A$ AS FUNCTIONS OF $\rho : p_l = 0.5, p_m = 0.2, l = 1, m = 6, h = 10, \delta = 0.99$.