An Economic Theory of Collective Identity

Hanming Fang Glenn C. Loury
Yale University Boston University

February 2, 2004

Abstract

1 Introduction

What is identity? How do identities form? How does identity affect economic performance? The answers to these questions are vitally useful for understanding a variety of important social problems. There are numerous views about what identity is, mainly because this is a multi-layered concept. Psychologists have in fact distinguished between "social identity" and "personal identity." Goffman (1963) describes a person's social identity as "category and attributes" anticipated by others during routines of social intercourse in established settings. Personal identity, or sometimes called "ego identity," is defined as a person's sense of self, or her answer to the question "Who am I?" There is also the concept of "collective identity" for For Akerlof and Kranton (AK, "Economic and Identity," QJE, 2000)), identity is a combination of "role" and "prescription." Society is partitioned into a set of "types" – occupants of the various roles. These types have preferences that naturally favor certain actions – those consistent with their role's prescription. Both roles and prescriptions are exogenous, given by history. The analysis proceeds by delineating and describing these role/prescription pairs, and then studying the implications for behavior of the resulting bias toward prescriptive actions for role occupants.

In this paper Hanming Fang and Glenn Loury, building on ideas from Loury's recent book "The Anatomy of Racial Inequality" and on the unpublished work of Roland Fryer and Matt Jackson [FJ, "Categorical Cognition: A Pyschological Model of Categories and identification in Decision Making," mimeo, Harvard University 2003], posits an alternative approach to identity. This approach is based on the ideas of "categories" and "narratives." These are quite different from "roles" and "prescriptions," and unlike AK, Loury's proposed theory will be driven by cognitive, not utilitarian considerations.

The main idea is that identity at its root has to do with **self-understanding.** It is a psychological, not a sociological category, ultimately. It is a matter of what might be called "reflexive cognition" – how a people see themselves. I want to use the categorical approach to social cognition pioneered in FJ to consider the matter of auto-cognition.

Thus, from the perspective of this paper a person's identity will be thought of as his answer to the question, "Who am I?" This answer is based, in substantial part, on a reading of the person's history – the totality of events that the person has experienced. But, any person's history is a VERY long vector of attributes. It cannot be comprehended by agents with limited cognitive resources until it has been reduced to a relatively few summary statistics. That is, settling upon an identity for oneself means that one makes selective generalizations about one's past that highlight and retain for future reference only the most salient aspects of one's experience. The "categories" reflect what a person takes to be salient, and the "narrative" is what results when a particular complex individual history gets mapped onto the categories.

Thus, the reduction of the full historical experience to a salient (relatively) few factors is akin to the mapping of a vector of attributes into the "prototypes" of categories in FJ. Likewise, one can think of a mapping that reduces any (possibly quite complex) history to one of a small number of pat stories about who a person is: "I'm an immigrant who came up the hard way;" I'm a working class white male angry at the world for not feeling my pain;" I'm a tough-minded professional woman determined not to take a back seat to any less qualified man;" "I'm a black intellectual who grew up poor unlike those silver-spoon-fed middle class blacks who love to talk about the ghetto but have never lived there;" "I'm a black man who likes to have sex with other men, but I'm not a sissy and I'm not "gay" either;" (see a recent NYTimes Magazine article on "down low" culture), etc. Each of these hypothetical people – responding to the question, "Who am I" – tell us something selective about their personal histories – they have chosen certain categories of self-understanding, and they "read" ("narrate") their particular historical attributes (and those of others too, I imagine) through their chosen categorical lenses. So, their "narrative" is their way of reading their own personal histories, their way of seeing themselves, their terms of self-understanding.

The categories, and the narrative to which they lead, are together what I want to refer to as the person's identity. I maintain that this is a more powerful approach than AK, because it is more foundational (less ad hoc, more room to deduce a rich field of implications with relatively little structure), and because it is more readily and more insightfully testable. In my theory, people with a given identity are supposed to see the world in similar terms sort the historical data into the same bins, more or less. This can be gotten at with experimental methods in a relatively direct manner, it would seem. The same can't be said for AK.

To illustrate how the analysis will go, I want to sketch here a clean, simple but powerful example to illustrate the ideas and to provide guidance for more general modeling. I want to give analytic content to the notion of "collective identity": I start with the premise that two or more people who have adopted more or less the same categories of self-understanding are going to be able to reach viable agreements (contracts, promises) about collective action that would not be possible if they used different categorical maps to interpret experience. "Collective identity facilitates in-group cooperation" would be the upshot of such a view. If this is so, then it is likely under such a set-up for incentives militating in favor of the evolution of collective identities to exist. But, not all common categorical maps (collective identities) are created equal – some are superior to others. A "bad" (dysfunctional, self-destructive, victim-based, alienated, oppositional, anti-system) collective identity might emerge in stable equilibrium and be difficult to shift with marginal incentives precisely because people have worked out a range of tacit collective action agreements that depend for their viability on maintaining in common the currently prevailing set of categories for reading one's (and another's) personal history.

In a repeated game, this might be modeled in terms of players' actions at any point in time being functions not of the entire history of play, but only of the "prototypes" for that history that are implied by the categories in use, and the "narrative assignment functions" that map the full personal histories into the categories. The intuition would be that cooperation can be achieved as an equilibrium in such a game only when (enough) agents in a group share the same ways of organizing their histories. Bad collective identities in such a world are simply shared ways of organizing history which have the unfortunate property that the best equilibria obtainable under them are dominated by the best (or maybe even the worst!) equilibria obtainable under some alternative categorical schema. The idea would be, however, that it would pay no single player, or perhaps not even any modestly sized coalition of players, to "change their identities" from "bad" to "better", given that they anticipated all other agents to continue with the prevailing historical categorization schemes.

For concreteness, consider a simple two person risk-sharing repeated game. This will be the context in which I propose to study the emergence of "dysfunctional collective identities." Every period each of the two receives a random endowment, either high or low. The endowment is not storable, so total consumption by the two must not exceed the sum of their endowments. Utility from consumption is strictly concave, so people dislike consumption fluctuations, and they can raise their common expected discounted utility levels by agreeing to transfer some consumption from the one who's random endowment is high to the one whose endowment is low, whenever there is a "high/low" realization. This sharing is enforced in classic repeated game fashion by a grim trigger strategy – any agent who refuses to transfer loses access to this sharing arrangement in the future.

Suppose further that the endowment realization for each individual is partially private information. That is, it is not publicly and immediately clear whether a person's realization has been high or low. Rather, there is a set of indicators associated with a "high" and with a "low" realization. These indicators (of an agent's history) are complex, and must be summarized by a "narrative" that reduces the full range of experience to a relatively few categories. In plain English, both people have to agree on whether or not one of them has been "victimized" by a low endowment realization before the transfer can take place. These agreements will be easier to reach when both agents share the same categorizing behavior (i.e., when they employ a common narrative of victimization!) The may both cast a rather wide definition or both a rather narrow one, but they share the genuine endowment risk between themselves more effectively in the maximally cooperative equilibrium of their repeated game interaction only when they also share a common categorizing rule – that is, when they adopt in common a "collective identity."

Now, let there be some costly action they can both take a priori that affects the probability of their actual endowments being high in any period. Suppose further that the incentive to take this action is undermined when they both adopt the wide definition of what constitutes evidence of victimization. (My intent here is to model the repeated risk sharing interaction as the second stage of a two-stage game. In the first stage each agent selects a costly action that is not observed by the other agent, and this choice generates a probability of that agent receiving low (and high) endowment realizations in each subsequent period of the second stage. The infinitely repeated risk sharing game is played in the second stage, and the grim trigger sustains cooperation there. Moreover, signals are available in each second stage period giving partial indication as to what action an agent took in the first stage (that is, of whether or not a low income realization is due to "bad luck" or to "poor preparation.") These signals must be interpreted via some categorical assignments that reflect the identity considerations I just discussed.

In this context one might seek conditions under which the following proposition obtains: An overall equilibrium to the two stage game can support a high level of costly effort from the players in the first stage only if their choice of collective identity leads them to adopt a **narrow** interpretation of what constitutes victimization during their second stage play. Such a result might follow from the fact that, for some parameter values, to adopt a **wide** sense of what constitutes victimization would imply that the "best" second stage risk sharing equilibrium would entail a degree of insurance advanced to the "victim" such as to create a situation in which there would not be a sufficiently large difference in the post-transfer consumption levels across the two states of nature (ie., high vs. low endowment realizations) in each second stage period. But, it is this post-transfer difference in consumption that provides the incentive to put in high effort in the first stage.

The intuition here is as follows: To adopt a collective identity that places too much emphasis on victimization leads to a second stage subgame in which there is "too much" insurance – too much, that is, to sustain high effort from both players in the first stage. In other words, I would seek conditions in the model under which the moral hazard problem is ”too great” to engender high effort from agents, given that they have adopted a collective identity which implies a definition of being a victim (of bad luck) that is ”too broad.”

In this way, one could construct a theory of ”dysfunctional collective identity” – namely, a collective identity is ”dysfunctional” when it reflects a pareto dominated scheme of categorical interpretation that nevertheless is an equilibrium in the two stage game just described!!

2 The Basic Model

There are two agents, indexed by $i = 1, 2$. At each period, agent i receives an income $y_i \in \{l, m, h\}$ with $l < m < h$. The probability that y_i is l, m and h are respectively given by p_l, p_m and p_h with $\sum_{k \in \{l,m,h\}} p_k = 1$. We assume that the income realizations are i.i.d across time and agents. For the sake of simplicity, we assume that the incomes are not storable. An agent’s income realization is her private information. However, we assume that she is subject to a mental coding mechanism which she chooses. To make things interesting and to capture our idea that it is practically infeasible for us to fully describe all aspects of our experience, we assume that an agent can handle two codes $\{G, B\}$. In other words, each agent has to code her realized income into G or B. Formally, a code is simply a map $C : \{l, m, h\} \to \{G, B\}$. We assume that once the mental coding mechanism is chosen, the agent will translate her income realization to an affect via her coding mechanism. The equilibrium coding mechanism is interpreted as collective identity.

Agents are risk averse with utility function $u(\cdot) : R_+ \to R$ that satisfies $u' > 0, u'' < 0$ and it is contiuous and differentiable.

The time line of the model is as follows. Before all interactions start, both agents choose their coding mechanisms. Agents observe each other’s coding mechanism and then will be engaged in an infinitely repeated risk sharing.

We consider two coding mechanisms, indexed by C_P (for pessimistic coding) and C_O (for optimistic coding). Code C_P is given by $C_P(l) = C_P(m) = B$ and $C_P(h) = G$; while $C_O(l) = B$, $C_O(m) = C_O(h) = G$.

3 Analysis

3.1 Equilibria under $\langle C_P, C_P \rangle$

Suppose that both choose code C_P. $C_P(l) = C_P(m) = B$ and $C_P(h) = G$. We then ask the degree of risk sharing between the two agents following such code configurations that can be sustained as repeated game equilibrium. We assume that any deviation from the prescribed equilibrum strategy will be followed by reverting to autarky.

Consider a symmetric equilibrium in which transfer occurs only when one agent's state is G and the other agent's state is B; and the the agent with state is G transfers $t \geq 0$ to the agent with state B. There is no transfer when both agents have the same state. We will focus on the most sustainable efficient equilibrium. We first ask what levels of t can be sustained in equilibrium.

Consider a candidate equilibrium with transfer level t. An agent's continuation utility in equilibrium is

$$\begin{aligned} V_P(t) &= \frac{1}{1-\delta} \times [p_l(p_l + p_m) u(l) + p_l p_h u(l+t) \\ &\quad + p_m(p_l + p_m) u(m) + p_m p_h u(m+t) \\ &\quad + p_h(p_l + p_m) u(h-t) + p_h^2 u(h)]. \end{aligned} \tag{1}$$

The autarky utility is

$$V_0 = \frac{1}{1-\delta} \times [p_l u(l) + p_m u(m) + p_h u(h)]. \tag{2}$$

Note that

$$V_P(t) - V_0 = \frac{1}{1-\delta} \left\{ \begin{array}{c} p_l p_h [u(l+t) - u(l)] + p_m p_h [u(m+t) - u(m)] \\ + p_h(p_l + p_m)[u(h-t) - u(h)] \end{array} \right\}.$$

The only incentive condition we need to check is that of an agent with state G who is forced to transfer t to the other agent. The incentive condition is simply

$$u(h-t) + V_P(t) \geq u(h) + V_0 \tag{3}$$

where the left hand side is the agent's discounted utility from participating in the equilibrium, and the right hand side is her discounted utility if she deviates from it (it is easy to verify that if she were to deviate from the candidate equilibrium, she will deviate to zero transfer). Combining (1), (2) and (3), we know that a necessary and sufficient condition for sustain level of transfer under C_P must satisfy

$$\begin{aligned} & p_l p_h \left[u\left(l+t\right)-u\left(l\right)\right]+p_m p_h \left[u\left(m+t\right)-u\left(m\right)\right] \\ \geq \;& \left[p_h\left(p_l+p_m\right)+\left(1-\delta\right)\right]\left[u\left(h\right)-u\left(h-t\right)\right] \end{aligned}$$

or equivalently,

$$\begin{aligned} F_P\left(t\right) &= \left\{p_l p_h u\left(l+t\right)+p_m p_h u\left(m+t\right)+\left[p_h\left(p_l+p_m\right)+\left(1-\delta\right)\right]u\left(h-t\right)\right\} \\ &\quad -\left\{p_l p_h u\left(l\right)+p_m p_h u\left(m\right)+\left[p_h\left(p_l+p_m\right)+\left(1-\delta\right)\right]u\left(h\right)\right\} \qquad (4) \\ &\geq 0 \end{aligned}$$

Note first $F_P\left(0\right)=0$; second, $F_P\left(t\right)$ is concave in t; third, $F_P'\left(t\right)$, evaluated at $t=0$, is given by

$$\begin{aligned} & p_l p_h u'\left(l\right)+p_m p_h u'\left(m\right)-\left[p_h\left(p_l+p_m\right)+\left(1-\delta\right)\right]u'\left(h\right) \\ = \;& p_h p_l\left[u'\left(l\right)-u'\left(h\right)\right]+p_h p_m\left[u'\left(m\right)-u'\left(h\right)\right]-\left(1-\delta\right)u'\left(h\right) \end{aligned}$$

Thus, for a fix set of parameter values $\langle p_l, p_m, p_h, l, m, h\rangle$ such that $l<m<h$, there exists

$$\bar{\delta}_P = 1-\frac{p_h p_l\left[u'\left(l\right)-u'\left(h\right)\right]+p_h p_m\left[u'\left(m\right)-u'\left(h\right)\right]}{u'\left(h\right)}<1, \qquad (5)$$

such that $F_P'\left(0\right)>0$ if and only if $\delta>\bar{\delta}_P$.

If $\delta<\bar{\delta}_P$, then $F_P'\left(0\right)<0$. Since $F_P''<0$, it implies that there exists no $t>0$ that satisfies incentive condition (4). If $\delta>\bar{\delta}_P$, then $F_P'\left(0\right)>0$. Since $t=0$ satisfies the incentive condition, it implies that there exists $\bar{t}_P>0$ such that IC condition is satisfied for all $t\in\left[0,\bar{t}_P\right]$, where $\bar{t}_P$ is the solution (other than 0) to equation $F_P\left(t\right)=0$.

We can also look at the conditions on $\mathbf{p}\equiv\left(p_l,p_m,p_h\right)$, for a fixed δ (sufficiently close to 1), under which $F_P'\left(0\right)>0$. Rewrite

$$\begin{aligned} & p_h p_l\left[u'\left(l\right)-u'\left(h\right)\right]+p_h p_m\left[u'\left(m\right)-u'\left(h\right)\right]-\left(1-\delta\right)u'\left(h\right) \\ = \;& p_h p_l\left[u'\left(l\right)-u'\left(h\right)\right]+p_h\left(1-p_l-p_h\right)\left[u'\left(m\right)-u'\left(h\right)\right]-\left(1-\delta\right)u'\left(h\right) \\ = \;& p_l p_h\left[u'\left(l\right)-u'\left(m\right)\right]+p_h\left(1-p_h\right)\left[u'\left(m\right)-u'\left(h\right)\right]-\left(1-\delta\right)u'\left(h\right)>0 \end{aligned}$$

Fix $p_h \in (0,1)$, it requires p_l to satisfy

$$p_l > \frac{(1-\delta)\, u'(h) - p_h (1-p_h) \left[u'(m) - u'(h)\right]}{p_h \left[u'(l) - u'(m)\right]} = \bar{p}_l^P$$

The most efficient risk sharing equilibrium under code C_P is the solution to the following problem

$$\max_{\{t\}} V_P(t) \;\; s.t. \; F_P(t) \geq 0 \tag{6}$$

Let $\hat{t}_P$ be the unconstrained maximizer of $V_P(t)$, i.e., $\hat{t}_P = \arg\max V_P(t)$. That is $\hat{t}_P$ satisfies the first order condition

$$p_l u'\left(l + \hat{t}_P\right) + p_m u'\left(m + \hat{t}_P\right) - (p_l + p_m)\, u'\left(h - \hat{t}_P\right) = 0. \tag{7}$$

Lemma 1 $l + \hat{t}_P < h - \hat{t}_P < m + \hat{t}_P.$

Proof. This follows from the fact that (7) implies that $u'\left(h - \hat{t}_P\right)$ is a convex combination of $u'\left(l + \hat{t}_P\right)$ and $u'\left(m + \hat{t}_P\right)$. ■

This is an interesting result: when agents code income according to C_P, the optimal risk sharing (without constraint) transfer will make the income-m agent consume the highest amount. This clearly will have important implications when we endogenize the realization of incomes later on.

Then, following from the concavity of $V_P(\cdot)$, we immediately have the following characterization:

Lemma 2 *The solution to problem (6), t_P^*, is given as follows:*

$$t_P^* = \begin{cases} \hat{t}_P \;\; \textit{if}\; \hat{t}_P \leq \bar{t}_P \\ \bar{t}_P \;\; \textit{if}\; \hat{t}_P > \bar{t}_P. \end{cases}$$

We write the discounted payoff from the most efficient equilibrium under code C_P as $V_P^* \equiv V_P(t_P^*)$. Alternatively, we may be interested in the whole utility possibility set under C_P:

$$\Omega_P = \{V_P(t) : F_P(t) \geq 0\}.$$

3.2 Equilibria under $\langle C_O, C_O \rangle$

Suppose that both agents choose code C_O, that is, $C_O(l) = B, C_O(m) = C_O(h) = G$. Analogous to the previous case, we consider symmetric equilibria in which an agent whose state is G will be asked to transfer $t \geq 0$ to an agent with state B; and no transfer occurs in other situations. We first ask what levels of t can be sustained as repeated game equilibrium; and then ask what will be the most efficient equilibrium under C_O.

Consider a candidate equilibrium with transfer level t. An agent's continuation utility in equilibrium is

$$\begin{aligned} V_O(t) &= \frac{1}{1-\delta}[p_l^2 u(l) + p_l(p_m+p_h)u(l+t) \\ &\quad + p_m p_l u(m-t) + p_m(p_m+p_h)u(m) \\ &\quad + p_h p_l u(h-t) + p_h(p_m+p_h)u(h)]; \end{aligned} \tag{8}$$

and the autarky utility is still given by (2). Thus

$$V_O(t) - V_0 = \frac{1}{1-\delta}\left\{ \begin{array}{c} p_l(p_m+p_h)[u(l+t)-u(l)] + p_m p_l[u(m-t)-u(m)] \\ + p_h p_l[u(h-t)-u(h)]. \end{array} \right\}$$

There are two incentive conditions for the repeated game equilibrium, one for agent with income m and one for agent with income h. Now we show that, if the incentive condition for income-m agent is satisfied, then that for income-h agent will be satisfied as well. The incentive condition for income-m agent is

$$u(m-t) + V_O(t) \geq u(m) + V_0; \tag{9}$$

while that for income-h agent is

$$u(h-t) + V_O(t) \geq u(h) + V_0; \tag{10}$$

Since $u(h-t) - u(h) > u(m-t) - u(m)$,[1] we have that (9) implies (10).

Now let us write out (9) more explicitly as follows:

$$\begin{aligned} F_O(t) &= \{p_l(p_m+p_h)u(l+t) + [p_m p_l + (1-\delta)]u(m-t) + p_h p_l u(h-t)\} \\ &\quad - \{p_l(p_m+p_h)u(l) + [p_m p_l + (1-\delta)]u(m) + p_h p_l u(h)\} \\ &\geq 0. \end{aligned} \tag{11}$$

---

[1] To see this, consider $G(x) = u(x-t) - u(x)$. $G'(x) = u'(x-t) - u'(x) > 0$. Thus $G(h) > G(m)$.

Again $F_O(0) = 0$; $F_O''(t) < 0$; and

$$\begin{aligned} F_O'(0) &= p_l(p_m + p_h) u'(l) - [p_m p_l + (1-\delta)] u'(m) - p_h p_l u'(h) \\ &= p_l p_m [u'(l) - u'(m)] + p_l p_h [u'(l) - u'(h)] - (1-\delta) u'(m). \end{aligned}$$

Thus $F_O'(0) > 0$ if and only if

$$\delta > \bar{\delta}_O = 1 - \frac{p_l p_h [u'(l) - u'(h)] + p_l p_m [u'(l) - u'(m)]}{u'(m)} < 1. \tag{12}$$

For any $\delta > \bar{\delta}_O$, there exists $t_O > 0$ such that IC condition (9) is satisfied for all $t \in [0, \bar{t}_O]$, where $\bar{t}_O$ is the solution (other than 0) to equation $F_O(t) = 0$.

Again we can look at the conditions on $\mathbf{p}$ such that $F_O'(0) > 0$. Rewrite

$$\begin{aligned} & p_l p_m [u'(l) - u'(m)] + p_l p_h [u'(l) - u'(h)] - (1-\delta) u'(m) \\ = \; & p_l (1 - p_l - p_h) [u'(l) - u'(m)] + p_l p_h [u'(l) - u'(h)] - (1-\delta) u'(m) \end{aligned}$$

The most efficient risk sharing equilibrium under code C_O is the solution to the following problem

$$\max_{\{t\}} V_O(t) \;\; s.t. \; F_O(t) \geq 0. \tag{13}$$

Let $\hat{t}_O$ be the unconstrained maximizer of $V_O(t)$, i.e., $\hat{t}_O = \arg\max V_O(t)$, which is the unique solution to the FOC:

$$(p_m + p_h) u'\left(l + \hat{t}_O\right) - p_m u'\left(m - \hat{t}_O\right) - p_h u'\left(h - \hat{t}_O\right) = 0. \tag{14}$$

Again noting that $u'\left(l + \hat{t}_O\right)$ is a convext combination of $u'\left(m - \hat{t}_O\right)$ and $u'\left(h - \hat{t}_O\right)$.

Lemma 3 $m - \hat{t}_O < l + \hat{t}_O < h - \hat{t}_O$.

Following the concavity of $V_O(\cdot)$, we have the following characterization to problem (13):

Lemma 4 *The solution to problem (13), t_O^*, is given as follows:*

$$t_O^* = \begin{cases} \hat{t}_O \;\; if \; \hat{t}_O \leq \bar{t}_O \\ \bar{t}_O \;\; if \; \hat{t}_O > \bar{t}_O. \end{cases}$$

We write the discounted payoff from the most efficient equilibrium under code C_O as

$$V_O^* \equiv V_O\left(t_O^*\right).$$

And write the whole utility possibility set under C_O as

$$\Omega_O = \left\{V_O\left(t\right) : F_O\left(t\right) \geq 0\right\}.$$

Let us now do some comparisons between the degree of risk sharing under C_P and C_O. First, compare $\hat{t}_O$ and $\hat{t}_P$. We have the following observations:

- if $m \approx h$, then $\hat{t}_O > \hat{t}_P$;
- if $l \approx m$, then $\hat{t}_O < \hat{t}_P$.

To see the first, recall that $\hat{t}_P$ satisfies FOC (7):

$$p_l u'\left(l + \hat{t}_P\right) + p_m u'\left(m + \hat{t}_P\right) - \left(p_l + p_m\right) u'\left(h - \hat{t}_P\right) = 0.$$

Evaluate FOC (14) at $\hat{t}_P$, we have

$$\begin{aligned} & \left(p_m + p_h\right) u'\left(l + \hat{t}_P\right) - p_m u'\left(m - \hat{t}_P\right) - p_h u'\left(h - \hat{t}_P\right) \\ > \quad & p_m u'\left(l + \hat{t}_P\right) - p_m u'\left(m - \hat{t}_P\right) \end{aligned}$$

since $u'\left(l + \hat{t}_P\right) > u'\left(h - \hat{t}_P\right)$. If $m \approx h$, then $u'\left(l + \hat{t}_P\right) > u'\left(m - \hat{t}_P\right)$ as well. The second case is shown analogously.

Now we provide interpretable conditions under which $\bar{\delta}_O < \bar{\delta}_P$. From (5) and (12), we have

$$\begin{aligned} \bar{\delta}_P &= 1 - \frac{p_h p_l \left[u'\left(l\right) - u'\left(h\right)\right] + p_h p_m \left[u'\left(m\right) - u'\left(h\right)\right]}{u'\left(h\right)} \\ \bar{\delta}_O &= 1 - \frac{p_l p_m \left[u'\left(l\right) - u'\left(m\right)\right] + p_l p_h \left[u'\left(l\right) - u'\left(h\right)\right]}{u'\left(m\right)} \end{aligned}$$

Since $u'\left(m\right) > u'\left(h\right)$, we immediately have the following three observations:

- If $p_l \left[u'\left(l\right) - u'\left(m\right)\right] \leq p_h \left[u'\left(m\right) - u'\left(h\right)\right]$, then $\bar{\delta}_P < \bar{\delta}_O$.
- if $m \approx h$, then $\bar{\delta}_P > \bar{\delta}_O$;

- if $l \approx m$, then $\bar{\delta}_P < \bar{\delta}_O$;

The second and third observations above follows from continuity. Note that $\bar{\delta}_P < \bar{\delta}_O$ means that the *possibility* of risk sharing, but not necessarily the level of risk sharing, is higher under the coding C_P.

3.3 Equilibria under $\langle C_P, C_O \rangle$

Suppose that agent 1 chooses code C_P and 2 code C_O. Now we investigate the risk sharing under the repeated game equilibria in this situation. We will consider the easiest possibility first and then discuss the implications of the other two possibilities later.

Possibility One: Equal transfers from G to B We first suppose that an agent will be asked to transfer $t \geq 0$ when her state is G and the other agent's state is B. (note that in this scenario, we are restricting the transfers to be the same for the two agents despite their different choice of coding mechanisms. Agent 1's discounted expected utility in the candidate equilibrium, $V_M^P(t)$ where M stands for "mixed codes", 1 stands for possibility 1; and P stands for pessimistic coding, is

$$V_M^P(t) = \frac{1}{1-\delta} \left\{ \begin{array}{c} p_l^2 u(l) + p_l(p_m + p_h) u(l+t) \\ +p_m p_l u(m) + p_m(p_m + p_h) u(m+t) \\ +p_h p_l u(h-t) + p_h(p_m + p_h) u(h) \end{array} \right\}.$$

Agent 2's discounted expected utility in the candidate equilibrium is

$$V_M^O(t) = \frac{1}{1-\delta} \left\{ \begin{array}{c} p_l(p_l + p_m) u(l) + p_l p_h u(l+t) \\ +p_m(p_l + p_m) u(m-t) + p_m p_h u(m) \\ +p_h(p_l + p_m) u(h-t) + p_h^2 u(h) \end{array} \right\}$$

Clearly $V_M^P(t) > V_M^O(t)$ for all $t > 0$. The autarky utility is as before.

Agent 1's incentive condition requires that

$$u(h-t) + V_M^P(t) \geq u(h) + V_0 \tag{15}$$

Agent 2's incentive condition (only income-m type is relevant) is

$$u(m-t) + V_M^O(t) \geq u(m) + V_0. \tag{16}$$

Since

$$V_M^O(t) - V_0 \geq V_M^P(t) - V_0$$

and

$$u(m) - u(m-t) \geq u(h) - u(h-t),$$

we know that constraint (16) implies (15). Now we write constraint (16) in more detail:

$$V_M^O(t) - V_0 = \frac{1}{1-\delta}\left\{\begin{array}{c} p_l p_h [u(l+t) - u(l)] \\ +p_m (p_l + p_m) [u(m-t) - u(m)] \\ +p_h (p_l + p_m) [u(h-t) - u(h)] \end{array}\right\},$$

thus constraint (16) is equivalent to

$$\begin{aligned} F_M(t) &= \{p_l p_h u(l+t) + [p_m(p_l+p_m) + (1-\delta)] u(m-t) + p_h(p_l+p_m) u(h-t)\} \quad (17) \\ &\quad - \{p_l p_h u(l) + [p_m(p_l+p_m) + (1-\delta)] u(m) + p_h(p_l+p_m) u(h)\} \\ &\geq 0. \end{aligned}$$

As before, $F_M(0) = 0$; $F_M'' < 0$; and

$$F_M'(0) = p_l p_h u'(l) - [p_m(p_l+p_m) + (1-\delta)] u'(m) - p_h(p_l+p_m) u'(h).$$

If $F_M'(0) \leq 0$ when $\delta = 1$, i.e., if

$$p_l p_h u'(l) - p_m(p_l+p_m) u'(m) - p_h(p_l+p_m) u'(h) < 0$$

then there will be no scope of risk sharing under this scenario. Thus, in this scenario, agent 2 whose code is C_O are asked to give up resources when her state is G and the opponent's state is B, an event with probability $(p_l+p_m)(p_m+p_h)$; while only receiving resources from agent 1 when 2's state is B and 1's state is G, an event with probability $p_l p_h$. Such asymmetry makes it difficult for agent 2 to benefit from risk sharing, unless her income is very uneven, i.e. l is very small relative to m and h.

If $p_l p_h u'(l) - p_m(p_l+p_m) u'(m) - p_h(p_l+p_m) u'(h) > 0$, however, then there will again be scope of risk sharing when $\delta > \bar{\delta}_M$ where

$$\bar{\delta}_M = 1 - \frac{p_l p_h u'(l) - p_m(p_l+p_m) u'(m) - p_h(p_l+p_m) u'(h)}{u'(m)}. \quad (18)$$

Now for any δ, the sustainability set of transfers in the repeated game is $\mathcal{T} = [0, \bar{t}_M]$ where $\bar{t}_M = 0$ if $F'_M(0) \leq 0$ and otherwise, it is the unique positive solution to $F_M(t) = 0$.

Now in order for $F'_M(0) > 0$, we need to have

$$p_l p_h u'(l) - [p_m(p_l + p_m) + (1-\delta)]\, u'(m) - p_h(p_l + p_m)\, u'(h) > 0$$

Lemma 5 $\bar{\delta}_M > \max\left\{\bar{\delta}_P, \bar{\delta}_O\right\}$.

Proof. Recall

$$\begin{aligned}
\bar{\delta}_P &= 1 - \frac{p_h p_l [u'(l) - u'(h)] + p_h p_m [u'(m) - u'(h)]}{u'(h)} \\
&= 1 - \frac{p_h p_l u'(l) + p_h p_m u'(m) - p_h (p_l + p_m) u'(h)}{u'(h)} \\
\bar{\delta}_O &= 1 - \frac{p_l p_m [u'(l) - u'(m)] + p_l p_h [u'(l) - u'(h)]}{u'(m)} \\
&= 1 - \frac{p_l (p_m + p_h) u'(l) - p_l p_m u'(m) - p_l p_h u'(h)}{u'(m)} \\
\bar{\delta}_M &= 1 - \frac{p_l p_h u'(l) - p_m (p_l + p_m) u'(m) - p_h (p_l + p_m) u'(h)}{u'(m)}.
\end{aligned}$$

Comparing $\bar{\delta}_M$ with $\bar{\delta}_P$, we notice that

$$\frac{p_h p_l u'(l) + p_h p_m u'(m) - p_h (p_l + p_m) u'(h)}{u'(h)} > \frac{p_l p_h u'(l) - p_m (p_l + p_m) u'(m) - p_h (p_l + p_m) u'(h)}{u'(m)}$$

because $u'(h) < u'(m)$ and the numerator is larger for the term in the left. Thus $\bar{\delta}_P < \bar{\delta}_M$.

Comparing $\bar{\delta}_M$ with $\bar{\delta}_O$, we notice that

$$\frac{p_l (p_m + p_h) u'(l) - p_l p_m u'(m) - p_l p_h u'(h)}{u'(m)} > \frac{p_l p_h u'(l) - p_m (p_l + p_m) u'(m) - p_h (p_l + p_m) u'(h)}{u'(m)}$$

because the numerator is smaller for the term in the left. Thus $\bar{\delta}_M > \bar{\delta}_O$. ■

Generalization of the Above Lemma: $F'_M(0) > 0 \Rightarrow$ both $F'_P(0) > 0$ and $F'_O(0) > 0$; but not vice versa.

The next question is, if there is scope for risk sharing, how should the total surplus be divided between agent 1 and 2? One position is as follows: Since 2's equilibrium payoff $V_M^{1O}(t)$ is always smaller than agent 1's, $V_M^{1P}(t)$, under any positive transfer. Thus it is not possible to obtain equal division of surplus. We can adopt a maxmin criterion and maximize agent 2's payoff $V_M^O(t)$. Thus we can solve

$$\max V_M^O(t) \;\; s.t. \; F_M^O(t) \geq 0. \tag{19}$$

Now the unconstrained maximizer of $V_M^O(t)$, denoted by $\hat{t}_M$, satisfies the FOC

$$p_l p_h u'\left(l+\hat{t}_M\right) - p_m \left(p_l + p_m\right) u'\left(m-\hat{t}_M\right) - p_h \left(p_l + p_m\right) u'\left(h-\hat{t}_M\right) = 0.$$

Different from the definitions of $\hat{t}_P$ and $\hat{t}_O$ where the solutions to their respective defining FOCs are necessarily strictly positive, $\hat{t}_M$ as defined above may be negative. We will take the max of zero and the solution to the above equation as the value of $\hat{t}_M$ we use below.

Lemma 6 *The solution to Problem (19), t_M^*, is given as follows:*

$$t_M^* = \begin{cases} \hat{t}_M \text{ if } \hat{t}_M \leq \bar{t}_M \\ \bar{t}_M \text{ if } \hat{t}_M > \bar{t}_M. \end{cases}$$

As before, we write thd discounted payoff from the most efficient equilibrium under codes $\langle C_P, C_O\rangle$ as

$$V_M^{O*} = V_M^O\left(t_M^*\right) \text{ and } V_M^{P*} = V_M^P\left(t_M^*\right).$$

We may also be interested in the whole utility possibility set under $\langle C_P, C_O\rangle$ given by

$$\Omega_M = \left\{\left(V_M^P(t), V_M^O(t)\right) : F_M(t) \geq 0\right\}.$$

Another possibility in terms of how t should be determined is to assume that agents engage in a Nash bargaining with equal bargaining power. Thus the objective function will be

$$\max \left[V_M^P(t) - V_0\right]^{\frac{1}{2}} \left[V_M^O(t) - V_0\right]^{\frac{1}{2}} \text{ s.t. } F_M(t) \geq 0.$$

[I have not analyzed this situatin so far].

3.4 Equilibrium Coding

Now we consider equilibrium of the coding choice stage. We essentially have a reduced normal form game given with the payoffs given by the following matrix (where we assume that the most efficient risk sharing equilibrium in the repeated game will be the relevant payoff in the subgame defined by the choices of coding mechanisms(:

		Agent 2	
		C_P	C_O
Agent 1	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 first result, relying on Lemma 5, is the following:

Proposition 1 *For all* $\delta \in \left(\max\left\{\bar{\delta}_P, \bar{\delta}_O\right\}, \bar{\delta}_M\right)$, *there are two equilibria in the coding choices:* $\langle C_P, C_P \rangle$ *and* $\langle C_O, C_O \rangle$.

The intuition for this result is very simple. Since $\bar{\delta}_M > \max\left\{\bar{\delta}_P, \bar{\delta}_O\right\}$, the scope of risk sharing is more limited under the mixed codes than under identical codes. That is, if $\delta < \bar{\delta}_M$, both V_M^{O*} and V_M^{P*} are zero. On the other hand, if $\delta > \max\left\{\bar{\delta}_P, \bar{\delta}_O\right\}$, then $V_P^* > 0$ and $V_O^* > 0$. This creates a coordination-like situation.

We also point out that when $\delta \in \left(\max\left\{\bar{\delta}_P, \bar{\delta}_O\right\}, \bar{\delta}_M\right)$, while the two equilibria $\langle C_P, C_P \rangle$ and $\langle C_O, C_O \rangle$ could be Pareto ranked generically, we (so far) do not have general result regarding their rankings. We hope to shed light on factors that will lead to one equilibrium to be superior to the other in both the numerical examples and more general environments.

It will be desirable to study, for a given δ, what will be the equilibrium codings from the payoff matrix above. It is possible, of course, no risk sharing is possible under any codings (when δ is sufficiently small). More interesting, however, is to consider the case that $\delta > \bar{\delta}_M$, i.e., the case that some degree of risk sharing will be subsequently achieved under all coding combinations. [I am not sure if any general analytical results can be obtained yet...]

3.5 Discussions of Other Possibilities under mixed codes

So far, we have restricted our attention, when the agents choose different coding mechanisms, to a particular possible of transfers between the two agents. Such restrictions definitely faciliate delivering our message (see Proposition 1). If we allow the agents, when they have different coding mechanisms, to engage in more general transfer possibilities, we think the same message will survive. We now deal with these cases.

3.5.1 Possibility Two: Unequal transfers from G to B

First, suppose that an agent will be asked to transfer income when her state is G and the other agent's state is B, and no transfer occurs otherwise. Let t_1 denote the transfer from agent 1 to 2 when 1's state is G and 2's state is B; and t_2 denote the transfer from agent 2 to 1 when $2'$s state is G and 1's state is B. We need to check the incentive conditions for both agent 1 and 2.

Consider a candidate equilibrium with transfers (t_1, t_2). Agent 1's expected discounted utility from the candidate equilibrium is

$$V_M^P(t_1, t_2) = \frac{1}{1-\delta} \left\{ \begin{array}{c} p_l^2 u(l) + p_l(p_m + p_h) u(l + t_2) \\ +p_m p_l u(m) + p_m(p_m + p_h) u(m + t_2) \\ +p_h p_l u(h - t_1) + p_h(p_m + p_h) u(h) \end{array} \right\};$$

and agent 2's discounted expected utility from the candidate equilibrium is

$$V_M^O(t_1, t_2) = \frac{1}{1-\delta} \left\{ \begin{array}{c} p_l(p_l + p_m) u(l) + p_l p_h u(l + t_1) \\ +p_m(p_l + p_m) u(m - t_2) + p_m p_h u(m) \\ +p_h(p_l + p_m) u(h - t_2) + p_h^2 u(h) \end{array} \right\}.$$

The incentive condition for agent 1 is

$$u(h - t_1) + V_M^P(t_1, t_2) \geq u(h) + V_0 \tag{20}$$

and that for agent 2 is

$$u(m - t_2) + V_M^O(t_1, t_2) \geq u(m) + V_0. \tag{21}$$

Now we further examine these constraints. Write constraint (20) explicitly as

$$\begin{aligned} G_M^P(t_1, t_2) &= p_l(p_m + p_h)[u(l + t_2) - u(l)] + p_m(p_m + p_h)[u(m + t_2) - u(m)] \\ &\quad + [p_h p_l + (1-\delta)][u(h - t_1) - u(h)] \\ &\geq 0. \end{aligned}$$

Note that, for any $t_2 \geq 0$, and for a fixed δ, there exists $\tilde{t}_1$ such that

$$\begin{aligned}
u(h) - u\left(h - \tilde{t}_1\right) &= \frac{p_l(p_m + p_h)\left[u(l + t_2) - u(l)\right] + p_m(p_m + p_h)\left[u(m + t_2) - u(m)\right]}{p_h p_l + (1 - \delta)} \\
u\left(h - \tilde{t}_1\right) &= u(h) - \frac{p_l(p_m + p_h)\left[u(l + t_2) - u(l)\right] + p_m(p_m + p_h)\left[u(m + t_2) - u(m)\right]}{p_h p_l + (1 - \delta)} \\
&\equiv \Delta_1(t_2) \\
h - \tilde{t}_1 &= u^{-1}(\Delta_1(t_2)) \equiv \phi(\Delta_1(t_2)) \\
\tilde{t}_1(t_2) &= h - \phi(\Delta(t_2)).
\end{aligned}$$

That is, agent 1's IC condition is satisfied for all $t_1 \leq \tilde{t}_1(t_2)$. It can be shown that $\tilde{t}_1(\cdot)$ is a concave function of t_2; and $\tilde{t}_1(0) = 0$. Whether $\tilde{t}_1(t_2)$ is larger or smaller than t_2 depends on the discount factor δ.

[Glenn, I will have to continue on this later ...]

3.5.2 Possibility Three: General Transfers

The most general possibility is of course, to allow agents to transfer incomes even when they have the same state - because the state G for agent 1 means differently from the state G for agent 2; so there is scope of risk sharing even in the event that both agents have states G, etc.

4 Moral Hazard, Risk Sharing and Identity

As we have learned from previous lemmas, different coding mechanisms imply different scope for risk sharing. So far, we have assumed that agents' income generating process is exogenous, that is, p_l, p_m and p_h are exogenous. It is interesting to examine two issues. First, how different codings provide different incentives in exerting effort? Second, if agents anticipate the moral hazard implications of coding mechanisms, what will be the equilibrium codings?

Point to make: the production function that represents how efforts affect incomes will play a crucial factor in equilibrium coding mechanisms. In essence, there will be a tradeoff between overcoming moral hazard and encouraging risk sharing. In the one extreme, if effort is completely useless in changing the income generating process, then agents will simply focus on the best codings for risk sharing; on the other hand, if moral hazard problem is severe, then agents will be better off forgo some subsequent risk sharing to encourage ex ante effort.

To study how moral hazard and risk sharing interact in determining equilibrium identity choices, we first specify how effort affects income. Now we simply assume that the probabilities at which the various income realizations are drawn can be affected by an individual's effort, that is, there are functions $p_l(e), p_m(e), p_h(e)$ such that $\sum_{k\in\{l,m,h\}} p_k(e) = 1$ for all e; and also we need to specify an effort cost function $c(e)$. [Before imposing any shape restrictions on $p_k(\cdot)$, it is w.l.o.g to assume that $c(e) = e$.] [note we assume that labor and consumption, e and y, enter the agent's utility function separately]

Instead of considering all possible links between effort and income, we consider three special cases:

1. $p_l'(e) < 0, p_m'(e) = 0$ and $p_h'(e) = -p_l'(e) > 0$. $p_h''(e) < 0$.
2. $p_l'(e) < 0, p_m'(e) = -p_l'(e) > 0$ and $p_h'(e) = 0$; $p_m''(e) < 0$
3. $p_l'(e) = 0, p_m'(e) < 0$ and $p_h'(e) = -p_m'(e) > 0$. $p_h''(e) < 0$.

The timing of the model is now as follows. First, agents choose a mental code, anticipating its effects on the subsequent effort choice and risk sharing; then in every stage afterwards, they would choose an effort and income will be realized, and then they would engage in risk sharing.

4.1 Equilibria under $\langle C_P, C_P\rangle$

The autarky utility, which is the same regardless of the code choices, is given by

$$V_0(e) = \frac{1}{1-\delta}\left[p_l(e)\,u(l) + p_m(e)\,u(m) + p_h(e)\,u(h) - c(e)\right].$$

The optimal effort choice e_0^* is characterized by the first order condition

$$p_l'(e)\,u(l) + p_m'(e)\,u(m) + p_h'(e)\,u(h) = c'(e)$$

The optimal autarky utility is simply $V_0\left(e_0^*\right)$.

Allowing for risk sharing, suppose that in a symmetric equilibrium, each agent is suppose to exert effort e and agent with state G is supposed to transfer t to the other agent with state B. The discounted expected utility in equilibrium is then

$$\begin{aligned} V_P\left(\tilde{e},e,t\right) &= \frac{1}{1-\delta}\times\{p_l\left(\tilde{e}\right)\left[p_l\left(e\right)+p_m\left(e\right)\right]u\left(l\right)+p_l\left(\tilde{e}\right)p_h\left(e\right)u\left(l+t\right) \\ &\quad +p_m\left(\tilde{e}\right)\left[p_l\left(e\right)+p_m\left(e\right)\right]u\left(m\right)+p_m\left(\tilde{e}\right)p_h\left(e\right)u\left(m+t\right) \\ &\quad +p_h\left(\tilde{e}\right)\left[p_l\left(e\right)+p_m\left(e\right)\right]u\left(h-t\right)+p_h\left(\tilde{e}\right)p_h\left(e\right)u\left(h\right)-c\left(\tilde{e}\right)\}. \end{aligned}$$

However, we now need to check two possible deviations, first, given that the other agent is putting in effort e, and given that the transfer will be t, she herself would have incentive to put in effort e as well. That is, if we maximize

$$V_P\left(\tilde{e},e,t\right)$$

with respect to $\tilde{e}$, taking e and t as given, we need to have e back as the maximizer. That is, (e,t) has to satisfy

$$\begin{aligned} \left(1-\delta\right)\left.\frac{\partial V_P\left(\tilde{e},e,t\right)}{\partial\tilde{e}}\right|_{\tilde{e}=e} &= p_l'\left(e\right)\left[p_l\left(e\right)+p_m\left(e\right)\right]u\left(l\right)+p_l'\left(e\right)p_h\left(e\right)u\left(l+t\right) \\ &\quad +p_m'\left(e\right)\left[p_l\left(e\right)+p_m\left(e\right)\right]u\left(m\right)+p_m'\left(e\right)p_h\left(e\right)u\left(m+t\right) \\ &\quad +p_h'\left(e\right)\left[p_l\left(e\right)+p_m\left(e\right)\right]u\left(h-t\right)+p_h'\left(e\right)p_h\left(e\right)u\left(h\right)-c'\left(e\right) \\ &= 0 \end{aligned}$$

Second condition is the usual condition that an agent with state G will have incentive to transfer t to the other agent with state B :

$$u\left(h-t\right)+V_P\left(e,e,t\right)\geq u\left(h\right)+V_0\left(e_0^*\right).$$

Now we consider the three cases above:

- Case 1: $p_l'\left(e\right)<0, p_m'\left(e\right)=0$ and $p_h'\left(e\right)=-p_l'\left(e\right)>0$. $p_h''\left(e\right)<0$. Substituting these into the FOC with respect to e, we have

$$\begin{aligned} &-p_h'\left(e\right)\left[p_l\left(e\right)+p_m\left(e\right)\right]u\left(l\right)-p_h'\left(e\right)p_h\left(e\right)u\left(l+t\right) \\ &+p_h'\left(e\right)\left[p_l\left(e\right)+p_m\left(e\right)\right]u\left(h-t\right)+p_h'\left(e\right)p_h\left(e\right)u\left(h\right)=c'\left(e\right) \end{aligned}$$

$$[p_l(e) + p_m(e)][u(h-t) - u(l)] + p_h(e)[u(h) - u(l+t)] = c'(e)/p'_h(e)$$

Since the left hand side is increasing in e [note

$$\frac{\partial RHS}{\partial e} = p'_h(e)\{[u(l+t) - u(l)] - [u(h) - u(h-t)]\} > 0$$

and the right hand side is decreasing in e,

$$\frac{\partial LFS}{\partial e} = \frac{c''(e)\,p'_l(e) - c'(e)\,p''_l(e)}{[p'_l(e)]^2} < 0.$$

There is at most one $e^*(t)$. Moreover, $e^*(t)$ is decreasing in t.

- Consider the case $p'_m(e) < 0$ and $p'_h(e) = -p'_m(e) > 0$ but $p'_l(e) = 0$.

$$[1 - p_h(e)][u(h-t) - u(m)] + p_h(e)[u(h) - u(m+t)] = c'(e)/p'_h(e)$$

The point we can make is that under $\langle C_P, C_P\rangle$, effort is decreasing in transfers. To see this, note that

$$\begin{aligned} & -\{[1 - p_h(e)]\,u'(h-t) + p_h(e)\,u'(m+t)\}\,dt + [u(h) - u(m+t) - u(h-t) + u(m)]\,p'_h(e)\,de \\ = & \ \frac{c''(e)\,p'_h(e) - c'(e)\,p''_h(e)}{[p'_h(e)]^2}de \end{aligned}$$

Thus

$$\frac{\partial e}{\partial t} < 0.$$

- If the effort-income link is given by $p'_l(e) < 0, p'_m(e) = -p'_l(e) > 0$ and $p'_h(e) = 0$, then the FOC for effort is

$$\begin{aligned} p'_l(e)[p_l(e) + p_m(e)]\,u(l) + p'_l(e)\,p_h(e)\,u(l+t) & \\ +p'_m(e)[p_l(e) + p_m(e)]\,u(m) + p'_m(e)\,p_h(e)\,u(m+t) &= c'(e) \end{aligned}$$

$$[p_l(e) + p_m(e)][u(m) - u(l)] + p_h[u(m+t) - u(l+t)] = c'(e)/p'_m(e)$$

How does e change with t?

$$p_h[u'(m+t) - u'(l+t)]\,dt = \frac{c''(e)\,p'_m(e) - c'(e)\,p''_m(e)}{[p'_m(e)]^2}de$$

This implies that

$$\frac{\partial e}{\partial t} < 0.$$

4.2 Equilibria under $\langle C_O, C_O \rangle$

The discounted expected utility in equilibrium is then

$$\begin{aligned} V_O(\tilde{e}, e, t) &= \frac{1}{1-\delta}[p_l(\tilde{e})\, p_l(e)\, u(l) + p_l(\tilde{e})\,[p_m(e) + p_h(e)]\, u(l+t) \\ &\quad + p_m(\tilde{e})\, p_l(e)\, u(m-t) + p_m(\tilde{e})\,[p_m(e) + p_h(e)]\, u(m) \\ &\quad + p_h(\tilde{e})\, p_l(e)\, u(h-t) + p_h(\tilde{e})\,[p_m(e) + p_h(e)]\, u(h)] - c(\tilde{e})\,; \qquad (22) \end{aligned}$$

We now need to check two possible deviations, first, given that the other agent is putting in effort e, and given that the transfer will be t, she herself would have incentive to put in effort e as well. That is, if we maximize

$$V_O(\tilde{e}, e, t)$$

with respect to $\tilde{e}$, taking e and t as given, we need to have e back as the maximizer. That is, (e, t) has to satisfy

$$\begin{aligned} (1-\delta)\left.\frac{\partial V_O(\tilde{e}, e, t)}{\partial \tilde{e}}\right|_{\tilde{e}=e} &= p_l'(e)\, p_l(e)\, u(l) + p_l'(e)\,[p_m(e) + p_h(e)]\, u(l+t) \\ &\quad + p_m'(e)\, p_l(e)\, u(m-t) + p_m'(e)\,[p_m(e) + p_h(e)]\, u(m) \\ &\quad + p_h'(e)\, p_l(e)\, u(h-t) + p_h'(e)\,[p_m(e) + p_h(e)]\, u(h) - c'(e) \\ &= 0 \end{aligned}$$

Second condition is the usual condition that an agent with state G will have incentive to transfer t to the other agent with state B :

$$u(m-t) + V_O(e, e, t) \geq u(m) + V_0(e_0^*)\,.$$

- Case 1: $p_l'(e) < 0, p_m'(e) = 0$ and $p_h'(e) = -p_l'(e) > 0$. $p_h''(e) < 0$. In this case, the FOC is

$$\begin{aligned} & p_l'(e)\, p_l(e)\, u(l) + p_l'(e)\,[p_m(e) + p_h(e)]\, u(l+t) \\ & + p_h'(e)\, p_l(e)\, u(h-t) + p_h'(e)\,[p_m(e) + p_h(e)]\, u(h) - c'(e) \\ = {} & 0 \end{aligned}$$

$$p_l(e)\,[u(h-t) - u(l)] + [p_m(e) + p_h(e)]\,[u(h) - u(l+t)] = \frac{c'(e)}{p_h'(e)}.$$

- Consider the case $p_m'(e) < 0$ and $p_h'(e) = -p_m'(e) > 0$ but $p_l'(e) = 0$. The FOC condition becomes

$$\begin{aligned}&-p_h'(e)\,p_l(e)\,u(m-t) - p_h'(e)\,[p_m(e)+p_h(e)]\,u(m)\\&+p_h'(e)\,p_l(e)\,u(h-t) + p_h'(e)\,[p_m(e)+p_h(e)]\,u(h) = c'(e)\end{aligned}$$

$$p_l\,[u(h-t)-u(m-t)] + (1-p_l)\,[u(h)-u(m)] = c'(e)/p_h'(e)$$

In contrast to our earlier result, under $\langle C_O, C_O\rangle$, effort is increasing in t. To see this, note the following:

$$p_l\,[u'(m-t)-u'(h-t)]\,dt = \frac{c''(e)\,p_h'(e) - c'(e)\,p_h''(e)}{[p_h'(e)]^2}de$$

thus

$$\frac{\partial e}{\partial t} > 0.$$

The FOC under $\langle C_P, C_P\rangle$ is given by

$$[1-p_h(e)]\,[u(h-t)-u(m)] + p_h(e)\,[u(h)-u(m+t)] = c'(e)/p_h'(e)$$

The first best effort under autarky is given by FOC

$$u(h)-u(m) = c'(e)/p_h'(e)$$

- If the effort-income link is given by $p_l'(e) < 0, p_m'(e) = -p_l'(e) > 0$ and $p_h'(e) = 0$, then the FOC for effort is

$$\begin{aligned}&p_l'(e)\,p_l(e)\,u(l) + p_l'(e)\,[p_m(e)+p_h(e)]\,u(l+t)\\&+p_m'(e)\,p_l(e)\,u(m-t) + p_m'(e)\,[p_m(e)+p_h(e)]\,u(m)\\&-c'(e)\\=\;&0\end{aligned}$$

$$p_l(e)\,[u(m-t)-u(l)] + [p_m(e)+p_h]\,[u(m)-u(l+t)] = \frac{c'(e)}{p_m'(e)}$$

How does e depend on t?

$$\begin{aligned} & -p_l(e)\, u'(m-t)\, dt - [p_m(e) + p_h]\, u'(l+t)\, dt \\ & +p_l'(e)\,[u(m-t) - u(l)]\, de + p_m'(e)\,[u(m) - u(l+t)]\, de \\ = & \ \frac{c''(e)\, p_m'(e) - c'(e)\, p_m''(e)}{[p_m'(e)]^2} de \end{aligned}$$

$$\begin{aligned} & -p_l(e)\, u'(m-t)\, dt - [p_m(e) + p_h]\, u'(l+t)\, dt \\ & +p_m'(e)\,[u(m) - u(l+t) - u(m-t) + u(l)]\, de \\ = & \ \frac{c''(e)\, p_m'(e) - c'(e)\, p_m''(e)}{[p_m'(e)]^2} de \end{aligned}$$

Thus

$$\frac{\partial e}{\partial t} < 0.$$

5 Numerical Examples

We provide a parametric class of utility functions $u(\cdot)$ to illustrate our main results; and to generate more comparative statics under this class of utility functions.

Consider the constant relative risk aversion family of utility functions

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

when $\rho = 1, u(y) = \ln y$. Given this parameterization on the utility function, our model will be simply defined the following list of parameters:

$$\left\langle \left(\{k, p_k\}_{k\in\{l,m,h\}}, \delta, \rho \right) : l < m < h, p_k \in (0,1), \sum_{k\in\{l,m,h\}} p_k = 1, \delta \in (0,1), \rho \in (0,1] \right\rangle.$$

We analyze

5.1 Equilibrium Coding with Exogenous Incomes

We first consider how the equilibrium coding choices are affected by the relative risk aversion parameter ρ. As ρ gets larger, the agent becomes more risk averse. Figure 1 shows

Figure 1: The Values as Functions of $\rho : p_l = 0.5, p_m = 0.2, l = 1, m = 6, h = 10.$

the differences of the four value functions and the autarky value V_0 as ρ varies. It is interesting to note that there exists ρ^* such that when $\rho \in (0, \rho^*)$, there are multiple equilibria (moreover, the equilibria are Pareto-ranked: the optimistic equilibrium Pareto dominates the pessimistic equilibrium) and when $\rho > \rho^*$, there is a unique equilibrium in which agents choose the pessimistic identity.

6 Conclusion

Figure 2: Values as Functions of δ :