Addendum to Social Norms and Choice.

Kevin Hasker
Rice University, Houston, TX 77005

In this addendum I include three proofs that are not intended for publication. The first section proves that the simple optimal penal code is unique in the given game. The second section proves that in two player games of pure coordination a folk theorem over the payoff space can be established. The third section shows that in the given game for the strategies considered all payoffs are weakly greater than the minmax in pure strategies.

1 The Uniqueness of the Simple Optimal Penal Code.

In this section I will prove that if $\delta = \frac{1}{2}$ then for $i \in \{r, c\}$ the following path

$$\underline{w}(i) = \begin{cases} \{M_i, M_i\} \text{ if } i \text{ deviated last period} \\ \{H_i, H_i\} \quad \text{else} \end{cases} \tag{1}$$

Is the worst equilibrium path for a player of role i in the stage game:

(2)

		column			
		H_r	H_c	M_r	M_c
row	H_r	$4, 2$	$3, 3$	$0, -2$	$1, 0$
	H_c	$0, 0$	$2, 4$	$0, -1$	$0, 0$
	M_r	$0, 0$	$0, 1$	$-4, -1$	$0, 0$
	M_c	$-1, 0$	$-2, 0$	$-5, -5$	$-1, -4$

Since

$$\pi_i(M_i, M_i) + \frac{\delta}{1-\delta}\pi_i(H_i, H_i) = -4 + \frac{\frac{1}{2}}{1-\frac{1}{2}}4 = 0 \tag{3}$$

and the individually rational payoff is zero, any other worst payoff must have the same payoff. Now if $\underline{w}(r)$ is the worst equilibrium path, the only equilibrium path that can be used as a punishment if a row player deviates while playing $\underline{w}(r)$ must be $\underline{w}(r)$. Focusing on deviations in the first period this means that:

$$v_r(\underline{w}(r)) \geq \max_{A^r \in \{M_r, M_c, H_r, H_c\}} \pi_r(A^r, A^c) + \delta v_r(\underline{w}(r)) \tag{4}$$

$$0 \geq \max_{A^r \in \{M_r, M_c, H_r, H_c\}} \pi_r(A^r, A^c)$$

Thus $A^c = M_r$. Furthermore, $\{M_c, M_r\}$ is not tenable since the continuation payoff must be five if $\{M_c, M_r\}$ is played in the first period, which is impossible. As well if either $\{H_c, M_r\}$ or $\{H_r, M_r\}$ is played in the first period then the continuation payoff for the column player must be five as well, since this is the amount they can increase their payoff by deviating today, and as previously mentioned this is not possible. Thus the first period action profile must be $\{M_r, M_r\}$ and given this the continuation payoff must be $\{H_r, H_r\}$ or the total discounted payoff will be negative. Thus the worst path must be the $\underline{w}(r)$ given by 1.

2 The Folk Theorem in Two Player Coordination Games.

Here I will prove that in two player coordination games every payoff that is strictly individually rational can be supported as an equilibrium.

Lemma 1 *For all two role stage games if $\pi^* = \pi(a) >> 0$ for $a \in A$ then there exists a δ^* such that if $\delta \geq \delta^*$ then there is an equilibrium with a payoff π^* each period.*

Proof. Define a^+ as the action that gives the highest payoff, and $m = \{m_1^2, m_2^1\}$ as the mutual minmax. Let a^* be the action where players play m with probability ρ and a^+ with probability $1 - \rho$ such that $\pi^* = \pi(a^*) = \rho\pi(m) + (1-\rho)\pi(a^+)$. Let the social standard of behavior be

$$\widetilde{\sigma}^t\left(z_i^t, z_j^t\right) = \begin{cases} m \text{ if } z_i^t = t-1 \text{ or } z_j^t = t-1 \\ a^* \quad\quad \text{else} \end{cases} \tag{5}$$

then if δ^* is such that for $i \in \{1, 2\}$

$$-\pi_i(m) \leq \delta^*[\pi_i(a^*) - \pi_i(m)] \tag{6}$$

for all $\delta \geq \delta^*$ the players will cooperate. Notice that given the social standard of behavior, if the player cooperates today the only future path is a^* forever, if they deviate then they will play m next period, thus the difference on the right hand side above. For the left hand side there are actually two cases but if they are playing a^+ and the left hand side is zero. If they are playing m then the left hand side is as given, and the proof is done. ■

I point out, however, that this is much weaker than the folk theorem over the action space. The usual folk theorem says that "anything is possible" here all that has been shown is that the set of equilibria is the correlated combination of the static Nash equilibria and m.

3 A Stage Game where all Constant Path Strategies have a Payoff Greater than the Minmax in Pure Strategies.

Consider the following stage game.

	L	R
U	$0,1$	$1,1$
M	$-1,4$	$-1,-4$
D	$-1,-4$	$-1,4$

(7)

A *constant path strategy* as one where given any history of play and $\left\{z_{\mu i}^{t+s}, z_{\mu i}^{t+s+1}\right\}$ that are consistent with the strategy being an equilibrium:

$$\#\left(s|s\in\{1,2,3,...\infty\},\sigma_i\left(z_i^{t+s},z_{\mu i}^{t+s}\right)\neq\sigma_i\left(z_i^{t+s+1},z_{\mu i}^{t+s+1}\right)\right)<\infty \quad (8)$$

this allows any actions in the next T periods but rules out strategies where the action profile cycles. The limiting average payoff is $\lim_{\delta\to 1}(1-\delta)v_i$. The proof can be requested from the author.

However, the reader may be interested to know that except for this qualification the author conjectures that the folk theorem does generalize for many stage games. This would require substantially more difficult strategies than used to prove the folk theorem above, as is best understood by example. Define a

This game is strategically simple, the minmax in pure strategies for the column player is one, and in mixed strategies it is zero—when the row player mixes over M and D with equal likelihood. Notice that in this game the row player has no *static* incentive to not mix over the strategies M and D, all incentive problems we will face are from the affect of choosing M and D on the matching rule. Also note that in order for the row player's limiting average payoff to be strictly less than one either $\{M,R\}$ or $\{D,L\}$ must be played with strictly positive probability.

Lemma 2 *For the above game in all the constant action equilibria the limiting average payoff for the column player is one.*

Proof. The proof is by contradiction, thus assume there is such an equilibrium. All other assumptions in this proof will be made without loss of generality. In this proof no row player will deviate, but a finite number of column players will deviate once each. This sequence of play will lead to a situation where maintaining the row player's indifference between M and D will require that her continuation value will have to be strictly greater than $\frac{1}{1-\delta}$, thus infeasible.

Assume that on the initial path the action played infinitely often is played in every period, and that $\{D, L\}$ is played with strictly positive probability. Also that the strategy following any deviation by the row player is always the same, and that this sequence of action profiles is independent of who the deviator is matched with (if there are no more deviations.)

Consider the equilibrium where the initial path gives column players' their lowest limiting average payoff of any equilibrium path, call this payoff $\underline{\pi}_c^*$. Since this is the column player's lowest initial path equilibrium payoff, on all continuation path a payoff greater than $\underline{\pi}_c^*$ must be played infinitely often, assume $\underline{\pi}_c^*$ is. Let T be the last period where the column player does not get $\underline{\pi}_c^*$ after a deviation, and let S be the number of periods in which the row player must mix over M and D. Our first step is to show that $S \geq 2$.

Call the equilibrium sequence of action profiles used after a deviation $\{\sigma^i\}_{i=1}^{\infty}$ where σ^i is used i periods after the deviation, and define $\pi_c' = \frac{1}{T}\Sigma_{i=1}^{T}\pi_c(\sigma^i)$. Then one can establish that the following conditions are necessary in any equilibrium with $\underline{\pi}_c^* < 1$.

$$\max_{a\in\{R,L\}} \pi^*\left(a, \sigma_r^k\right) < k\underline{\pi}_c^* - \Sigma_{s=1}^{k-1}\pi_c(\sigma^s) \tag{9}$$

$$\pi_c(R, D) - \pi_c(L, D) < T\left(\underline{\pi}_c^* - \pi_c'\right)$$

where σ_r^k is the row player's action in the k'th period after any deviation.[1] First the action in the first period must be a mixed strategy, since $\max_{a\in\{R,L\}} \pi^*\left(a, \sigma_r^k\right) < \underline{\pi}_c^* < 1$ if and only if the row player mixes over M and D. The constrained infimum of $\pi^*\left(a, \sigma_r^k\right)$ is -1 in this case, therefore assume $\pi_c(\sigma^1) = -1$. In the second period, the constraint is $2\underline{\pi}_c^* - \pi_c(\sigma^1) < 2+1$, and solutions to this problem require either that the row player mixes over M and D or that U is played with probability one. Since $1 > \underline{\pi}_c^*$ any finite number of periods playing U with probability one will not solve the problem. If you have $T-1$ periods of $\pi_c(\sigma^i) \geq \underline{\pi}_c^*$, then $T(\underline{\pi}_c^* - \pi_c') < 2 < \pi_c(R, D) - \pi_c(L, D)$. Thus we must at least have one more period where the row player mixes over M and D, or $S \geq 2$. Assume $S = 2$ and in the first two periods after any deviation the row player mixes over M and D.

In the second step we will show that if j mixed over M and D in $t-1$ and t, and then is matched with a player in $t+1$ with whom she must mix her discounted value must be

[1] These conditions are derived by taking the limit as $\delta \to 1$ of $\max_{a\in\{R,L\}} \pi^*\left(a, \sigma_r^k\right) \leq \delta^{T+1}\frac{1-\delta^k}{1-\delta}\underline{\pi}_c - \Sigma_{s=1}^{k-1}\delta^s\pi_c\left(\sigma^{t+s}\right)$ and $\pi_c(R, D) - \pi_c(L, D) \leq \delta\frac{1-\delta^T}{1-\delta}\underline{\pi}_c - \Sigma_{s=1}^{T}\delta^s\pi_c\left(\sigma^{t+s}\right)$

increased by one. Consider the following matching regime for a given row player j.

$$\mu_j^{t+1} = \begin{cases} \text{if possible with a player who deviated in } t-1 \text{ if } j \text{ plays } a \text{ in period } t \\ \text{otherwise someone at random.} \end{cases} \tag{10}$$

where $a \in \{M, D\}$. In period t by straightforward player j can know if some player deviated in $t-1$. On the other hand, the strategy can not be a function of the matching rule. Thus the only way to make j indifferent in t is if her payoff next period is the same independent of who she is matched with. This is done by decreasing or increasing her discounted value, and must be done in every period $t+1$ if she punished in t.

Assume that j plays M then she will be matched with someone who deviated in $t-1$. Assume that t is the first period in which j is matched with a deviator, and that if she is not matched in $t+1$ with someone who must be minmaxed then her discounted value will be decreased by $\alpha \geq 1$.

While this will work if she is matched with someone who must be punished in $t+2$ then her discounted value must be increased by α. Assume not, then if she plays M today in $t+2$ her discounted value will be decreased by α, while if she plays D since she expects not to mix in $t+1$ or $t+2$ and her discounted value will not change. Thus she will play D, contradicting the claim.

Now assume that in period $t-1$ two column players deviated and that in t j is matched with one of them. Further assume that in periods t to $t+k-1$ a column player deviates in each period. Clearly there exists a k such that if j's discounted value is increased by α each period then j's discounted value will be greater than $\frac{1}{1-\delta}$. Furthermore, there exists a positive probability that j should play M each period. Thus given a finite number of deviations with positive probability there exists a sequence of play where j will not randomize and the proposed strategy is not an equilibrium. ■