

APPENDICES TO ACCOMPANY: GROUP REPUTATIONS, STEREOTYPES, AND COOPERATION IN A REPEATED LABOR MARKET

Paul J. Healy

The correlation structure in the distribution of types assumed in the text is formally derived in Appendix A. A literature review of experimental results on stereotyping is provided in Appendix B. The three main propositions of the text are proved in Appendix C. The function $L(\bar{w}, \bar{e})$ is derived in Appendix D, and the experimental instructions are provided in Appendix E.

A The Correlation Structure

Here we formally demonstrate how correlated beliefs affect firms' beliefs about workers' types after observing a single defection. For each worker i , let θ_i be a binary random variable representing i 's type, where $\theta_i = 1$ indicates a reciprocal type and $\theta_i = 0$ indicates a selfish type. We assume that $\Pr[\theta_i = 1] = p_1$ for each i and that $\text{Corr}[\theta_i, \theta_k] = \gamma \in [0, 1]$ for all $i \neq k$. Let $\sigma = \sqrt{p_1(1-p_1)}$. From this we wish to derive $\Pr[\theta_i = 1 | \theta_k = 0]$.

Recall that for any random variables X and Y , $\text{Corr}[X, Y] = \text{Cov}[X, Y] / (\sigma_X \sigma_Y)$ and $\text{Cov}[X, Y] = E[XY] - E[X]E[Y]$. Since θ_i and θ_k are binary random variables,

$$\text{Corr}[\theta_i, \theta_k] = \frac{\Pr[\theta_i = 1, \theta_k = 1] - p_1^2}{p_1(1-p_1)}.$$

Since $\text{Corr} [\theta_i, \theta_k] = \gamma$ we have

$$\Pr [\theta_i = 1, \theta_k = 1] = p_1^2 + \gamma p_1 (1 - p_1).$$

It must be that

$$\Pr [\theta_i = 1] = \Pr [\theta_i = 1, \theta_k = 0] + \Pr [\theta_i = 1, \theta_k = 1],$$

or

$$p_1 = \Pr [\theta_i = 1, \theta_k = 0] + p_1^2 + \gamma p_1 (1 - p_1).$$

Therefore, $\Pr [\theta_i = 1, \theta_k = 0] = (1 - \gamma) p_1 (1 - p_1)$. From this, the probability that i is reciprocal given that some $k \neq i$ is selfish is

$$\begin{aligned} \Pr [\theta_i = 1 | \theta_k = 0] &= \frac{\Pr [\theta_i = 1, \theta_k = 0]}{\Pr [\theta_k = 0]} \\ &= \frac{(1 - \gamma) p_1 (1 - p_1)}{(1 - p_1)} \\ &= (1 - \gamma) p_1. \end{aligned}$$

B Stereotyping in the Lab: A Literature Review

Previous experimental studies find support for reputation-building in repeated games (for example, see Colin F. Camerer & Keith Weigelt (1988), John Neral & John Ochs (1992), and James Andreoni & John H. Miller (1993).) Type heterogeneity is also clearly present in subject populations. Ernst Fehr, Georg Kirchsteiger & Arno Riedl (1998), Fehr & Armin Falk (1999), and others observe heterogeneous behavior in laboratory labor markets, and Nahoko Hayashi *et al.* (1999) find reciprocal behavior in 68% of second-movers in a one-

shot sequential prisoners' dilemma.¹ Actual heterogeneity is not inconsistent with the FRE model, however, since the added correlation in the model may stem from uncertainty about the base rate of selfish types, as mentioned above, or from non-Bayesian belief updating. Although either is plausible, research in economics and social psychology verify that such stereotypical beliefs, though inconsistent, can exist in a laboratory setting.

As an example, William McEvily *et al.* (2007) show that subjects make inferences about the trustworthiness of future opponents based on whether or not *past* opponents were trustworthy. This effect becomes more pronounced when the opponents are grouped together according to some unrelated criterion. This 'minimal group paradigm' effect is well documented in the social psychology literature (see Henri Tajfel (1970) or Tajfel *et al.* (1971).) Decision makers apparently use past behavior to make inferences about the future behavior of others, especially when there is any reason to think those individuals share a group identity.

The observation that people infer more correlation than is warranted is known in the social psychology literature as 'illusory correlation'. By design, labor market experiments separate firms and workers into groups before the experiment begins, creating an initial identification of group membership among the subjects. A subject acting as a firm may see the group of firms as his 'ingroup' and the group of workers as the 'outgroup'. This partitioning leads naturally to categorical thinking (i.e., stereotyping) on the part of subjects, even if it is common knowledge that the outgroup is heterogeneous. As Louise F. Pendry & C. Neil Macrae (1999, p. 926) note, "while true that outgroups are commonly perceived to be less heterogeneous in composition than ingroups, outgroup members nonetheless still display appreciable degrees of variability. Acknowledging the variability of social groups, however, is no antidote to stereotypical thinking." Thus, subjects who are aware of their

¹In Hayashi *et al.* (1999), 11 of 18 American subjects and 15 of 20 Japanese subjects play the reciprocal strategy. Zero of 13 American and 3 of 25 Japanese second-movers cooperate after observing defection, providing evidence that cooperation in this game is almost always due to reciprocity and not pure altruism.

opponents' heterogeneity may still act in a stereotype-consistent way.

Experimental psychology has established that perceivers are less likely to apply existing stereotypes when the actions of the perceived affect the outcomes of the perceiver (see Ralph Erber & Fiske (1984) or Steven L. Neuberg & Susan T. Fiske (1987).) This would suggest that competitive environments reduce stereotypical beliefs. However, the stereotyping bias returns when cognitive resources are depleted by multiple task requirements. For example, Pendry & Macrae (1994) find that subjects who are asked to memorize an 8-digit number are more likely to recall stereotype-consistent information about others. As summarized by Macrae & Galen V. Bodenhausen (2000, p. 105), "judgement becomes more stereotypic under cognitive load."

Since firms in the experimental labor markets are likely using cognitive resources to watch the market, devise strategies, and compute payoffs, they may be more likely to think categorically about the group of workers even though their payoffs depend on the behavior of workers. Furthermore, Vincent Y. Yzerbyt, A. Coull & S. J. Rocher (1999) find that when subjects are exposed to information about a group member inconsistent with a formed stereotype, the stereotype shifts more dramatically when the subject is under a high cognitive load. This evidence supports a significant change in beliefs when confronted with a sudden change in behavior by a single worker.

Finally, a study by Janet B. Ruscher *et al.* (1991) shows that when groups are perceived to be in competition rather than individuals (so, if firms see themselves as *collectively* in competition with workers,) then subjects tend to pay more attention to stereotype-consistent information regarding individuals in the outgroup and stereotype-inconsistent information for members of their ingroup. Additionally, Marilyn B. Brewer, Joseph G. Weber & Barbara Carini (1995) and Hank Rothgerber (1997) show that "competition has the potential to create stereotypes where none or very few exists before," as summarized by Oliver Corneille & Yzerbyt (2002, p. 118). This emphasizes that there need not be existing

stereotypes of the group of workers for the firms to develop stereotypes when placed in a competitive market situation.

C Proofs of the Propositions

C.1 Proposition 1

Proposition 1 *Assume there are n workers and m firms with $n > m$. In the T -period repeated labor market with publicly observed random matching and public wage and effort choices, there is a full reputation equilibrium (\bar{w} in every period and \bar{e} in every period but the last) if and only if (1) firms' common prior belief about each worker's type is at least p^* , and (2) $m/n \geq \delta^*$.*

Proof. First note that we have restricted the strategy set of reciprocal workers so that they must choose \bar{e} in response to \bar{w} and \underline{e} in response to \underline{w} , regardless of the history of the game or strategies of others. We can therefore restrict attention to only the selfish worker types.

Let n and m represent the number of workers and firms, respectively, and recall that we assume $n \geq m$. Workers do not explicitly discount future payoffs, but since unmatched workers earn zero payoffs, future payoffs are implicitly discounted by m/n . Index workers by i and firms by j and denote their wage and effort choices in each period t by $w_{j,t}$ and $e_{i,t}$, respectively. For each period t , if i' is matched with j' , let $i(j', t) = i'$ and $j(i', t) = j'$. Let the firms' common probability that worker i is reciprocal at the beginning of period t be $p_{i,t}$.

Formal Description of the Full Reputation Equilibrium (FRE)

Beliefs

Workers are assumed to have complete information. In each period t , if worker i has chosen $e_{i,s} = \bar{e}$ when $w_{j(i,s),s} = \bar{w}$ and $e_{i,s} = \underline{e}$ when $w_{j(i,s),s} = \underline{w}$ for all $s < t$, then

$p_{i,t} = p_{i,1}$. Otherwise, $p_{i,t} = 0$. Thus, $p_{i,t} \in \{0, p_{i,1}\}$.

Strategies

In each period t , each reciprocal worker i chooses $e_{i,t} = \bar{e}$ if $w_{j(i,t),t} = \bar{w}$ and $e_{i,t} = \underline{e}$ if $w_{j(i,t),t} = \underline{w}$. For each selfish worker i , if $t < T$ and $p_{i,t} \geq p^*$, then i plays $e_{i,t} = \bar{e}$ if $w_{j(i,t),t} = \bar{w}$ and $e_{i,t} = \underline{e}$ if $w_{j(i,t),t} = \underline{w}$. Otherwise, $e_{i,t} = \underline{e}$ regardless of $w_{j(i,t),t}$.

In every period t , each firm j chooses $w_{j,t} = \bar{w}$ if and only if $p_{i(j,t),t} \geq p^*$.

Verifying Sequential Rationality

Note that this equilibrium is in pure strategies. Also, by construction, reciprocal workers cannot deviate from the FRE.

We now show that no selfish worker or firm will prefer to deviate in any period. If $p_{i,t} = 0$ for some selfish worker i in some period t , then i cannot affect firms' future beliefs; it is common knowledge that i is selfish. By a standard backwards induction argument for finitely repeated games with complete information, i will play her stage game dominant strategy in every period $s \geq t$. Knowing this, every firm matched with i in periods $s \geq t$ will offer \underline{w} .

We now consider only the cases where $p_{i,t} \neq 0$.

Period T

Pick an arbitrary selfish worker i . In *any* sequential equilibrium, $e_{i,T} = \underline{e}$ since it is a dominant strategy of the stage game.

By the definition of p^* , firm j will prefer to offer $w_{j,T} = \bar{w}$ if and only if $p_{i(j,T),T} \geq p^*$. Since $p_{i,T} \leq p_{i,1}$ for all i , this demonstrates that the FRE cannot exist if $p_{i,1} < p^*$ for some i . Therefore we restrict further analysis to the case of $p_{i,1} \geq p^*$ for all i .

Period $T - 1$

Workers: Case 1

Assume $w_{j(i,T-1)} = \underline{w}$ and $p_{i,T-1} = p_{i,1}$. If i conforms to the FRE by playing \underline{e} , then $p_{i,T} = p_{i,T-1}$ and her discounted payoff over the last two periods is $u(\underline{w}, \underline{e}) + \frac{m}{n}u(\bar{w}, \underline{e})$.

If i deviates by playing $e_{i,T-1} = \bar{e}$, then $p_{i,T} = 0$, leading to $w_{j(i,T),T} = \underline{w}$. Worker i 's discounted expected payoff over the last two periods becomes $u(\underline{w}, \bar{e}) + \frac{m}{n}u(\underline{w}, \underline{e})$, which is strictly lower than conforming to the FRE.

Workers: Case 2

Assume $w_{j(i,T-1)} = \bar{w}$ and $p_{i,T-1} = p_{i,1}$. If i conforms to the FRE by playing \bar{e} , then $p_{i,T} = p_{i,T-1}$ and she receives $u(\bar{w}, \bar{e}) + \frac{m}{n}u(\bar{w}, \underline{e})$. If she deviates, then $p_{i,T} = 0$ and she receives $u(\bar{w}, \underline{e}) + \frac{m}{n}u(\underline{w}, \underline{e})$. By the definition of δ^* , deviating is strictly worse if and only if $m/n \geq \delta^*$. This demonstrates that the FRE does not exist if $m/n < \delta^*$. Therefore we restrict further analysis to the case of $p_{i,1} \geq p^*$ for all i and $m/n \geq \delta^*$. It only remains to be verified that a FRE will always exist under these conditions.

Firms

The only remaining case is for $p_{i(j,T-1),T-1} = p_{i(j,T-1),1} \geq p^*$. If the firm conforms by offering \bar{w} , his expected payoff over the last two periods is

$$\pi(\bar{w}, \bar{e}) + P_{T-1}\pi(\bar{w}, \bar{e}) + (1 - P_{T-1})\pi(\bar{w}, \underline{e}),$$

where $P_t = |\{i : p_{i,t} = p_{i,1}\}|/n$. If the firm defects by offering \underline{w} , his expected payoff is

$$\pi(\underline{w}, \underline{e}) + P_{T-1}\pi(\bar{w}, \bar{e}) + (1 - P_{T-1})\pi(\bar{w}, \underline{e}).$$

The firm's period $T - 1$ decision does not affect his period T beliefs or expected outcomes, so he prefers to act myopically and choose \bar{w} .

Period $T - k$

Workers: Case 1

Assume $w_{j(i,T-k),T-k} = \underline{w}$ and $p_{i,T-k} = p_{i,1}$. If i conforms to the FRE by playing \underline{e}

then $p_{i,T-k+1} = p_{i,1}$ and her discounted payoff over the last $k + 1$ periods is

$$(1) \quad u(\underline{w}, \underline{e}) + \sum_{s=1}^{k-1} \left(\frac{m}{n}\right)^s u(\bar{w}, \bar{e}) + \left(\frac{m}{n}\right)^k u(\bar{w}, \underline{e}).$$

If she defects, her payoff is

$$(2) \quad u(\underline{w}, \bar{e}) + \sum_{s=1}^k \left(\frac{m}{n}\right)^s u(\underline{w}, \underline{e}),$$

which is lower in every period.

Workers: Case 2

Assume $w_{j(i,T-k),T-k} = \bar{w}$ and $p_{i,T-k} = p_{i,1}$. If i conforms by playing \bar{e} , her payoff is

$$(3) \quad u(\bar{w}, \bar{e}) + \sum_{s=1}^{k-1} \left(\frac{m}{n}\right)^s u(\bar{w}, \bar{e}) + \left(\frac{m}{n}\right)^k u(\bar{w}, \underline{e}).$$

If she defects, her payoff is

$$(4) \quad u(\bar{w}, \underline{e}) + \sum_{s=1}^k \left(\frac{m}{n}\right)^s u(\underline{w}, \underline{e}).$$

Since $m/n \geq \delta^*$, she prefers to conform for all $k \geq 1$.

Firms

Assume $p_{i(j,T-k),T-k} = p_{i(j,T-k),1} \geq p^*$. As in period $T - 1$, the firm's actions in period $T - k$ cannot affect his beliefs or others' actions in the next period. He therefore acts myopically by choosing $w_{j,T-k} = \bar{w}$.

Verifying Belief Consistency

It only remains to show that this system of beliefs is the limit of beliefs that are Bayes-consistent with totally mixed strategies converging to the pure strategy equilibrium.

Since reciprocal workers are *unable* to play \underline{e} in response to \bar{w} or \bar{e} in response to \underline{w} ,

$p_{i,t}$ must drop to zero after such actions are observed. (Formally, these actions lead to singleton information sets which are roots of subgames with complete information in which reciprocal workers do not participate.) Therefore we need only to consider information sets following sequences of actions of the form (\bar{w}, \bar{e}) or $(\underline{w}, \underline{e})$.

For some period $t < T$, Assume $p_{i,t} \geq p^*$. Equilibrium strategies require $w_{j(i,t),t} = \bar{w}$ and $e_{i,t} = \bar{e}$ with probability one for both worker types. If these actions are observed, no information is gained about the worker and $p_{i,t+1} = p_{i,t}$ by Bayes's Law. If instead \underline{w} and \underline{e} are observed, a zero-probability information set is reached. If we perturb strategies to be totally mixed, so that $\Pr [w_{j(i,t),t} = \underline{w}] = \varepsilon$ and $\Pr [e_{i,t} = \underline{e} | w_{j(i,t),t} = \underline{w}] = 1 - \delta$, then Bayes's Law requires

$$\begin{aligned} p_{i,t+1} &= \frac{p_{i,t} \varepsilon}{p_{i,t} \varepsilon + (1 - p_{i,t}) \varepsilon (1 - \delta)} \\ &= \frac{p_{i,t}}{p_{i,t} + (1 - p_{i,t}) (1 - \delta)}. \end{aligned}$$

As δ (and ε) shrink to zero, $p_{i,t+1}$ converges to $p_{i,t}$. Thus, $p_{i,t+1} = p_{i,t}$ is consistent.

Assume now that $p_{i,t} < p^*$. Equilibrium strategies require $w_{j(i,t),t} = \underline{w}$ and $e_{i,t} = \underline{e}$ for both worker types. If these actions are observed, no information is revealed and $p_{i,t+1} = p_{i,t}$. If instead \bar{w} and \bar{e} are observed, a zero-probability information set is reached. The calculation of $p_{i,t+1}$ under 'perturbed' strategies is identical to the above case (observing \underline{w} and \underline{e} when $p_{i,t} \geq p^*$), so, by an identical argument, $p_{i,t+1} = p_{i,t}$ is consistent.

C.2 Proposition 2

Proposition 2 *In the T -period repeated labor market with completely anonymous random matching and public wage and effort choices, there is a full reputation equilibrium (\bar{w} in*

every period and \bar{e} in every period but the last) if and only if (1) the firms' common prior belief (p_1) satisfies

$$p_1 \in \left[p^*, \frac{n}{n-1} p^* \right),$$

and (2) $m/n \geq \delta^*$.

Proof. Along the equilibrium path, selfish workers' strategies are identical to those of reciprocal workers, so $p_t = p_1$. Assume play has proceeded along the equilibrium path to period T , so that $p_T = p_1$. If $p_1 < p^*$, then each firm j will prefer to deviate by offering $w_{j,T} = \underline{w}$. Thus a FRE does not exist when $p_1 < p^*$.

Assume play has proceeded along the equilibrium path to period $T - 1$, so that $p_{T-1} = p_1$. Assume $p_1 \geq \frac{n}{n-1} p^*$ and $w_{j,T-1} = \bar{w}$ for each firm j . If worker i conforms to the FRE by choosing \bar{e} (and all other workers and firms conform as well,) then $p_T = p_1$ and worker i expects to earn $u(\bar{w}, \bar{e}) + \frac{m}{n} u(\bar{w}, \underline{e})$ in the last two periods. If instead the worker deviates to \underline{e} (but all others conform to the FRE,) then the firms' period- T belief becomes $p_T = \frac{n-1}{n} p_1 \geq p^*$. All firms will offer $w_{j,T} = \bar{w}$, so worker i expects to earn $u(\bar{w}, \underline{e}) + \frac{m}{n} u(\bar{w}, \underline{e})$ in the last two periods. This is greater than the payoff of conforming the FRE, so at least one worker will defect. A FRE does not exist when $p_1 \geq p^*$.

Finally, assume play reaches period $T - 1$ and $p_1 \in [p^*, \frac{n}{n-1} p^*)$. By an argument identical to Case 2 of Period $T - 1$ in the proof of Proposition 1, selfish workers will prefer to defect by choosing \underline{e} in response to \bar{w} .

It remains to verify that an FRE exists when conditions (1) and (2) are both satisfied.

Beliefs are specified as follows. If all workers' actions in all periods $s < t$ are consistent with cooperation (\bar{e} given \bar{w} and \underline{e} given \underline{w} .) then $p_t = p_1$. If only one deviation from cooperation has previously occurred, then $p_t = \frac{n-1}{n} p_1$. We do not need to explicitly specify beliefs after observing multiple deviations, other than to require that $p_t \leq \frac{n-1}{n} p_1$. Note that $p_{t+1} \leq p_t$ for all $t < T$ regardless of players' actions.

Selfish workers choose reciprocal actions in all periods t such that $t < T$ and $p_t \in [p^*, \frac{n}{n-1}p^*)$, and choose \underline{e} regardless of $w_{j(i,t),t}$ otherwise. Firms offer \bar{w} in all periods t such that $p_t \in [p^*, \frac{n}{n-1}p^*)$, and \underline{w} otherwise.

Period T

As the last movers, selfish workers will conform to the equilibrium by playing \underline{e} regardless of $w_{j,T}$.

Since $p_T \geq p^*$, firms will prefer to offer \bar{w} in the last period.

Period $T - 1$

Case 1: $p_{T-1} \in [p^*, \frac{n}{n-1}p^*)$

Pick a worker i and assume $w_{j(i,T-1),T-1} = \bar{w}$. If i conforms, she expects to earn $u(\bar{w}, \bar{e}) + \frac{m}{n}u(\bar{w}, \underline{e})$ over the last two periods. If she defects, then $p_T = \frac{n-1}{n}p_{T-1} < p^*$, so $w_{j(i,T),T} = \underline{w}$ and she expects to earn $u(\bar{w}, \underline{e}) + \frac{m}{n}u(\underline{w}, \underline{e})$. The assumption that $m/n \geq \delta^*$ guarantees that i will conform.

Assume $w_{j(i,T-1),T-1} = \underline{w}$. If i conforms, she expects to earn $u(\underline{w}, \underline{e}) + \frac{m}{n}u(\bar{w}, \underline{e})$. If she defects, $p_T = \frac{n-1}{n}p_{T-1} < p^*$ and she earns $u(\underline{w}, \bar{e}) + \frac{m}{n}u(\underline{w}, \underline{e})$. Defecting makes i strictly worse off in both periods, so she conforms.

Since $p_{T-1} \geq p^*$, an argument identical to that from used in the case of public matching shows that firms prefer to offer \bar{w} .

Case 2: $p_{T-1} < p^*$

Since p_t is non-increasing in t regardless of workers' actions, $w_{j,T} = \underline{w}$ for all j regardless of workers' period $T - 1$ effort choices. Therefore workers act myopically by choosing \underline{e} regardless of the wage offers. Knowing this (and since $p_{T-1} < p^*$), firms offer \underline{w} .

Case 3: $p_{T-1} \geq \frac{n}{n-1}p^*$

Since we have assumed $p_1 < \frac{n}{n-1}p^*$ and $p_{t+1} \leq p_t$ for all $t < T$, this case cannot occur.

Period $T - k$

Case 1: $p_{T-k} \in [p^*, \frac{n}{n-1}p^*)$

Assume $w_{j(i,T-k),T-k} = \underline{w}$ for some i . Equations (1) and (2) represent i 's payoff for conforming ($e_{i,T-k} = \underline{e}$, causing $p_{T-k+1} = p_{T-k}$) and defecting ($e_{i,T-k} = \bar{e}$, causing $p_{T-k+1} < p^*$), respectively. Conforming generates strictly higher payoffs.

Assume $w_{j(i,T-k),T-k} = \bar{w}$ for some i . Equations (3) and (4) represent i 's payoff for conforming ($e_{i,T-k} = \bar{e}$, causing $p_{T-k+1} = p_{T-k}$) and defecting ($e_{i,T-k} = \underline{e}$, causing $p_{T-k+1} < p^*$), respectively. Conforming generates strictly higher payoffs.

Regardless of the firms' wage offers, $p_{T-k+1} = p_{T-k}$, so continuation payoffs do not depend on wage offers in period $T - k$. Since $\pi(\bar{w}, \bar{e}) > \pi(\underline{w}, \underline{e})$, all firms offer \bar{w} .

Case 2: $p_{T-k} < p^$*

Since p_t is non-increasing in t regardless of workers' actions, $w_{j,T-k+1} = \underline{w}$ for all j regardless of workers' period $T - k$ effort choices. Therefore workers act myopically by choosing \underline{e} regardless of the wage offers. Knowing this (and since $p_{T-k} < p^*$), firms offer \underline{w} .

Case 3: $p_{T-k} \geq \frac{n}{n-1}p^$*

Again, this case is ruled out by assumption.

Verifying that beliefs are consistent along the equilibrium path is identical to the argument in the proof of Proposition 1. After one defection in period t , firms know that exactly one worker is the selfish type with certainty but have learned nothing about the other workers. Thus, $p_{t+1} = \frac{1}{n}(0 + (n-1)p_t) = \frac{n-1}{n}p_t$. Clearly additional defections will not increase p_{t+1} , so $p_{t+1} \leq \frac{n-1}{n}p_1$ as was assumed.

C.3 Proposition 3

Proposition 3 *In the T -period repeated labor market with completely anonymous random matching, public wage and effort choices, and a common knowledge stereotyping param-*

ter γ , there is a full reputation equilibrium (\bar{w} in every period and \bar{e} in every period but the last) if and only if (1) the firms' common prior belief (p_1) satisfies

$$p_1 \in \left[p^*, \frac{1}{1-\gamma} \frac{n}{n-1} p^* \right),$$

and (2) $m/n \geq \delta^*$.

Proof. The proof is similar to that of Proposition 2, except the equilibrium is specified as follows. If all workers' actions in all periods $s < t$ are consistent with cooperation (\bar{e} given \bar{w} and \underline{e} given \underline{w}), then $p_t = p_1$. If only one deviation from cooperation has previously occurred, then $p_t = \frac{1}{1-\gamma} \frac{n-1}{n} p_1$. We do not need to explicitly specify beliefs after observing multiple deviations, other than to require that $p_t \leq \frac{1}{1-\gamma} \frac{n-1}{n} p_1$. Note that $p_{t+1} \leq p_t$ for all $t < T$ regardless of players' actions.

Selfish workers choose reciprocal actions in all periods t such that $t < T$ and $p_t \in \left[p^*, \frac{1}{1-\gamma} \frac{n}{n-1} p^* \right)$, and choose \underline{e} regardless of $w_{j(i,t),t}$ otherwise. Firms offer \bar{w} in all periods t such that $p_t \in \left[p^*, \frac{1}{1-\gamma} \frac{n}{n-1} p^* \right)$, and \underline{w} otherwise.

The proof of Proposition 2 can now be replicated by setting $p_t = (1 - \gamma) p_1 (n - 1) / n$ (instead of $p_t = p_1 (n - 1) / n$) if a single defection has been observed in some period $s < t$ and $p_t \leq (1 - \gamma) p_1 (n - 1) / n$ after multiple defections in previous periods.

D Measure of Equilibrium Existence

Here we derive the Lebesgue measure of the set of parameters on which a full reputation equilibrium exists. Since the two parameters (p_1 and γ) are drawn from $[0, 1]^2$, the Lebesgue measure of a set is equivalently its probability assuming p_1 and γ are independent random variables with uniform distributions over $[0, 1]$. In general, assume p_1 and γ

have densities f and g (with cumulative distribution functions F and G) on $[0, 1]$, respectively. Given m and n (the number of firms and workers, respectively,) the probability that both conditions of Proposition 3 are satisfied at some (\bar{w}, \bar{e}) is given by

$$L(\bar{w}, \bar{e}) = \begin{cases} \int_0^{1-p^* \frac{n}{n-1}} \left[F\left(p^* \frac{n}{n-1} \frac{1}{1-\gamma}\right) - F(p^*) \right] g(\gamma) d\gamma \\ + \int_{1-p^* \frac{n}{n-1}}^1 [1 - F(p^*)] g(\gamma) d\gamma & \text{if } \delta^* \leq m/n \quad , \\ 0 & \text{if } \delta^* > m/n \end{cases}$$

where p^* and δ^* are functions of \bar{w} and \bar{e} and depend on the choices of π and u .² The integral expression can be simplified to

$$1 - F(p^*) - G\left(1 - p^* \frac{n}{n-1}\right) + \int_0^{1-p^* \frac{n}{n-1}} F\left(p^* \frac{n}{n-1} \frac{1}{1-\gamma}\right) g(\gamma) d\gamma.$$

Note that if $g(\gamma) = 0$ for all $\gamma < 1 - p^*n/(n-1)$, then this expression reduces to $1 - F(p^*)$.

Returning to the case of independent uniform distributions, the expression reduces to

$$1 - p^* - \left(1 - p^* \frac{n}{n-1}\right) + \int_0^{1-p^* \frac{n}{n-1}} p^* \frac{n}{n-1} \frac{1}{1-\gamma} d\gamma.$$

Integrating and simplifying gives the formula

$$p^* \left[\frac{n}{n-1} \left(1 - \ln \left[p^* \frac{n}{n-1} \right] \right) - 1 \right].$$

²If the upper limit of integration is smaller than the lower limit, we assume the integral takes a value of zero.

Including the case of $\delta^* > m/n$, we have

$$(5) \quad L(\bar{w}, \bar{e}) = \begin{cases} p^* \left[(1 - \ln [p^* \frac{n}{n-1}]) \frac{n}{n-1} - 1 \right] & \text{if } \delta^* \leq m/n \\ 0 & \text{if } \delta^* > m/n \end{cases} .$$

For a given $n > 1$, the p^* that maximizes this strictly concave function is $p^* = \frac{n-1}{n} e^{-\frac{n-1}{n}}$. The value of $L(\bar{w}, \bar{e})$ at this point is simply $e^{-\frac{n-1}{n}}$, which is strictly less than one. Since $e^{-\frac{n-1}{n}}$ is the least upper bound on the measure, we only graph $L(\bar{w}, \bar{e})$ between zero and $e^{-\frac{n-1}{n}}$.

E Experiment Instructions

General Instructions [All Treatments]

The experiment you will participate in is part of a research project used to analyze the decision behavior in markets. The instructions are simple, and if you read them carefully and make appropriate decisions, you can earn a considerable amount of money. At the end of the whole experiment, all the profits you have made by your decisions will be added up and paid to you in cash. The experiment you will participate in consists of two stages. In the first stage six of you act as buyers, and nine of you as sellers. In the second stage, the sellers will determine the value of the goods for the buyers (for details of the second stage see below). We have distributed two kinds of instructions – information for the buyers, and information for the sellers, respectively. This information is for private use only – you are not allowed to reveal this information to anyone. Furthermore, you will find at the end of these instructions a second sheet (sheet 2) that is used to document your decisions. Insert your buyer or seller number there. [In later sessions,

Specific Instructions for Buyers [LA & LP Treatments]

At the market, a good is traded, and each seller sells the same good. A seller can sell this good to any buyer, and a buyer can buy it from any seller. The market is organized in the following way: we open the market for a trading period (a “trading day”), and each trading day lasts three minutes. As a buyer you can offer a price that must be divisible by 5, for example, prices like 15, 60, 80, 275 are allowed, but prices like 48, 67, 124, 83 are not. These offers will be announced to the sellers by us over the telephone. The sellers will not know your identity, that is, your buyer number; they will only know the price offered. If a seller accepts your offer, all buyers are informed about this acceptance. In this case, an agreement is concluded, and the good is bought by you at the offered price. During each trading day you can buy one unit of the good. Therefore, a trading day ends for you when your offer is accepted. Note also that each seller can sell one unit of the good per day at most. If your offer is not accepted, you are free to change your offer, that is, to make a new offer. But the new price you offer must be higher than all the prices that have not been accepted. Each seller may accept an offer or not, but he cannot make a counteroffer.

After three minutes the day ends, and you cannot buy any more of the good. Then the second stage of the experiment will be conducted. After this, a new trading day is opened. On the whole, there will be twelve trading days. In the second stage of the experiment, the seller who has sold the good to you on this day can fix the value that the good will have for you. You as a buyer get a certain amount of experimental money (reselling price) from us for each unit you have bought. This reselling price is noted in the upper part of sheet 2. Your profit (measured in experimental money) is the difference between the reselling price and the price at which you have bought the good. If you bought the good for 20 and the reselling price is 30, you make a profit of $30 - 20 = 10$ (measured in experimental money). How much one unit of experimental money is worth to you depends on “your” seller. By the choice of a conversion rate, he decides how much real money you receive from us for one unit of experimental money. Which conversion rates he is allowed to choose are noted

on the lower part of sheet 2. If he chooses, for example, the rate 0.5, you will get \$5 for 10 units of experimental money.

Sellers have two kinds of costs: production costs and decision costs. The latter are associated with the decision about the conversion rate. Production costs are noted in the middle of sheet 2, and decision costs on the lower part of sheet 2. As you can see from sheet 2, the higher the conversion rate “your” seller chooses, the greater are his decision costs. The profit of the sellers paid in dollars is given by the formula: profit = (price - production costs - decision costs). Suppose, for example, that you have bought the good for 75. The production costs of the seller are 60, and he chooses a conversion rate of 0.6 (which is associated with decision costs of 5), the profits of “your” seller are given by $75 - 60 - 5 = \$10$. Do you have any questions?

Specific Instructions for Sellers [LA & LP Treatments]

At the market, a good is traded, and each seller sells the same good. A seller can sell this good to any buyer, and a buyer can buy it from any seller. The market is organized in the following way: we open the market for a trading period (a “trading day”), and each trading day lasts three minutes. Every buyer can offer a price that will be relayed to us by telephone. We list these offers on the blackboard, and you can accept one of these offers. If, e.g., a price of 50 is offered and you as seller number 5 want to accept this offer, you just say: “Number 5 sells for 50.” In this case, the transaction is concluded. The good is sold to the buyer who made the offer of 50. The buyer will not know your identity. He will just know that his offer is accepted. You have to note your accepted price on sheet 2.

You can sell one unit of the good on each trading day. Therefore, the trading day ends for you after the acceptance of an offer. Note also that each buyer can buy, at most, one unit of the good per trading day. Each seller may accept an offer or not, but the sellers cannot make counteroffers. After three minutes the trading day ends, and the second stage

of the experiment is conducted. After this, a new trading day is opened. In total there will be twelve trading days. At the second stage of the experiment, you can fix the value the good will have for the buyers. Buyers receive a certain amount of experimental money (reselling price) from us for each unit that they have bought. This reselling price is noted in the middle of sheet 2.

The profit of a buyer (measured in experimental money) is the difference between the reselling price and the price at which he has bought the good from you. If “your” buyer has bought the good for 20 and the reselling price is 30, he makes a profit of $30 - 20 = 10$ (measured in experimental money.) How much one unit of experimental money is worth for “your” buyer depends on you. By the choice of a conversion rate, you decide how much real money “your” buyer gets from us for one unit of experimental money. If you choose, e.g., the rate 0.5, your buyer gets \$5 for 10 units of experimental money. Which conversion rates you are allowed to choose, is noted on the lower part of sheet 2. You have to write down your decision on the upper part of sheet 2. Do not announce your decision publicly.

You, as a seller, have two kinds of costs: production costs and “decision costs.” The latter are associated with your decision about the conversion rate. Of course, you incur costs only in the case of a deal. If you do not trade on a certain day, your costs are zero for this day. Production costs are noted on the upper part of sheet 2. Decision costs depend on your choice of the conversion rate. The higher the conversion rate you decide to give “your” buyer, the greater are your decision costs. The costs, which are associated with the conversion rate, are noted in the lower part of sheet 2.

Your profit paid in dollars is given by the formula $\text{profit} = \text{price} - \text{production costs} - \text{decision costs}$. If, for example, you sell your good for 75, while your production costs are 60, and you choose a conversion rate of 0.6 which leads to a decision cost of 5, your profit is given by $75 - 60 - 5 = \$10$. Do you have any questions?

Specific Instructions for Buyers [HP Treatment]

At the market, a good is traded, and each seller sells the same good. A seller can sell this good to any buyer, and a buyer can buy it from any seller. The market is organized in the following way: we open the market for a trading period (a “trading day”), and each trading day lasts three minutes. As a buyer you can offer a price that must be divisible by 5, for example, prices like 15, 60, 80, 275 are allowed, but prices like 48, 67, 124, 83 are not. These offers will be announced to the sellers by us over the computer and projected in their room along with the buyers’ ID numbers. If a seller accepts your offer, all buyers are informed about this acceptance and the ID number of the seller who accepted it. At that point, the seller then chooses a number ‘ x ’ that affects how valuable the good is to you. Higher values of ‘ x ’ make the good more valuable, but cost the seller more money. This choice is then transmitted, along with the seller’s ID number, back to this room and the transaction is concluded. The good is bought by you at the offered price and your value is affected by ‘ x ’. You have to note the accepted price and the seller’s choice of ‘ x ’ on your record sheet.

During each trading day you can buy one unit of the good. Therefore, a trading day ends for you when your offer is accepted. Note also that each seller can sell one unit of the good per day at most. If your offer is not accepted, you are free to change your offer, that is, to make a new offer. But the new price you offer must be higher than all the prices that have not been accepted. Each seller may accept an offer or not, but he cannot make a counteroffer. After three minutes the day ends, and you cannot buy any more of the good. After this, a new trading day is opened. On the whole, there will be twelve trading days.

Your profit is the fixed value of the good (which is shown on your record sheet,) multiplied by the number ‘ x ’ that your seller will determine, minus the price you pay to buy the

good. Mathematically, your profit is given by the formula

$$\text{buyer profit} = \text{value} * x - \text{price}.$$

The seller's profit is the price they get for the good, minus a fixed production cost, minus an 'additional cost' based on their choice of ' x '. The formula for their profit is

$$\text{seller profit} = \text{price} - \text{production cost} - \text{additional cost}.$$

Your record sheet lists the value of the good to the buyers, the production cost to the sellers, and what the 'additional cost' for the seller is for each choice of ' x '. The higher the choice of ' x ', the greater are the 'additional costs.'

If, for example, your value for the good is 400, the seller chooses ' x ' to be 0.49, then your value times ' x ' equals 196. If the price you paid was 175, then your profit is $196 - 175 = 21$. If the seller's production cost is 100 and his additional cost from choosing $x = 0.49$ is 6, then the seller's profit is $175 - 100 - 6 = 69$. This example appears on your record sheet.

At the end of the experiment, your earnings will be converted to dollars at a rate of 12 to 1. Do you have any questions?

Specific Instructions for Sellers [HP Treatment]

At the market, a good is traded, and each seller sells the same good. A seller can sell this good to any buyer, and a buyer can buy it from any seller. The market is organized in the following way: we open the market for a trading period (a "trading day"), and each trading day lasts three minutes. Every buyer can offer a price (in multiples of 5) that will be relayed to us by computer and projected at the front of the room, along with the ID number of the buyer. We list these offers on the screen, and you can accept one of these offers. If,

e.g., a price of 50 is offered and you as seller number 5 want to accept this offer, you just say: “Seller 5 sells for 50.” At that point, you then choose a number ‘ x ’ that affects how valuable the good is to the buyer. Higher values of ‘ x ’ make the good more valuable to the buyer, but cost you more money. This choice is then transmitted, along with the your ID number, back to the buyers and the transaction is concluded. The good is sold by you at the offered price and you pay an additional cost for your choice of ‘ x ’. You have to note your accepted price and your choice of ‘ x ’ on your record sheet.

During each trading day you can sell one unit of the good. Therefore, a trading day ends for you when you accept an offer. Note also that each buyer can buy one unit of the good per day at most. If a buyer’s offer is not accepted, the buyer is free to change his offer, but the new price must be higher than all the prices that have not been accepted. Each seller may accept an offer or not, but you cannot make a counteroffer. After three minutes the day ends, and you cannot accept any offers. After this, a new trading day is opened. On the whole, there will be twelve trading days.

The profit of a buyer is the buyer’s fixed value of the good (which is shown on your record sheet,) multiplied by the number ‘ x ’ that you will determine, minus the price they pay to buy the good. Mathematically, the buyer’s profit is given by the formula

$$\text{buyer profit} = \text{value} * x - \text{price}.$$

Your profit is the price you get for the good, minus a fixed production cost, minus an ‘additional cost’ based on your choice of ‘ x ’. The formula for your profit is

$$\text{seller profit} = \text{price} - \text{production cost} - \text{additional cost}.$$

In the second stage of the experiment, your job is to choose the value ‘ x ’, and your ‘addi-

tional cost' depends on this decision. Your record sheet lists the value of the good to the buyers, your production cost, and what the 'additional cost' is for each choice of ' x '. The higher the choice of ' x ', the greater are your 'additional costs.

If, for example, you sell your good for 175, while your production costs are 100, and you choose the value of ' x ' as 0.49 which leads to a decision cost of 6, your profit is given by $175 - 100 - 6 = 69$. This example appears on your record sheet.

At the end of the experiment, your earnings will be converted to dollars at a rate of 12 to 1. Do you have any questions?

References

- Andreoni, James and John H. Miller, "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence," The Economic Journal, 1993, 103 (418), 570–585.
- Brewer, Marilyn B., Joseph G. Weber, and Barbara Carini, "Person memory in intergroup contexts: Categorization versus individuation," Journal of Personality and Social Psychology, 1995, 69 (1), 29–40.
- Camerer, Colin F. and Keith Weigelt, "Experimental Tests of a Sequential Equilibrium Reputation Model," Econometrica, 1988, 56 (1), 1–36.
- Corneille, Olivier and Vincent Y. Yzerbyt, "Dependence and the formation of stereotyped beliefs about groups: from interpersonal to intergroup perception," in Craig McGarty, Vincent Y. Yzerbyt, and Russel Spears, eds., Stereotypes as Explanations: The formation of meaningful beliefs about social groups, Cambridge: Cambridge University Press, 2002.
- Erber, Ralph and Susan T. Fiske, "Outcome dependency and attention to inconsistent information," Journal of Personality and Social Psychology, 1984, 47 (4), 709–726.

- Fehr, Ernst and Armin Falk, "Wage Rigidity in a Competitive Incomplete Contract Market," Journal of Political Economy, 1999, 107 (1), 106–134.
- ___, Georg Kirchsteiger, and Arno Riedl, "Gift Exchange and Reciprocity in Competitive Experimental Markets," European Economic Review, 1998, 42 (1), 1–34.
- Hayashi, Nahoko, Elinor Ostrom, James Walker, and Toshio Yamagishi, "Reciprocity, Trust, and the Sense of Control A Cross-Societal Study," Rationality and Society, 1999, 11 (1), 27–46.
- Macrae, C. Neil and Galen V. Bodenhausen, "Social Cognition: Thinking Categorically about Others," Annual Review of Psychology, 2000, 51 (1), 93–120.
- McEvily, William, Roberto A. Weber, Cristina Bicchieri, and Violet Ho, "Can groups be trusted? An experimental study of collective trust," 2007. In *The Handbook of Trust*, forthcoming.
- Neral, John and Jack Ochs, "The Sequential Equilibrium Theory of Reputation Building: A Further Test," Econometrica, 1992, 60 (5), 1151–1169.
- Neuberg, Steven L. and Susan T. Fiske, "Motivational Influences on Impression Formation: Outcome dependency, accuracy-driven attention, and individuating processes," Journal of Personality and Social Psychology, 1987, 53 (3), 431–444.
- Pendry, Louise F. and C. Neil Macrae, "Stereotypes and Mental Life: The Case of the Motivated but Thwarted Tactician," Journal of Experimental Psychology, 1994, 30 (4), 303–325.
- ___ and ___, "Cognitive Load and Person Memory: The Role of Perceived Group Variability," European Journal of Social Psychology, 1999, 29 (7), 925–942.
- Rothgerber, Hank, "External intergroup threat as an antecedent to perceptions of in-group and out-group homogeneity," Journal of Personality and Social Psychology, 1997, 73 (6), 1206–1211.
- Ruscher, Janet B., Susan T. Fiske, Hiromi Miki, and Scott F. van Manen, "Individuating

processes in competition: interpersonal versus intergroup,” Personality and Social Psychology Bulletin, 1991, 17 (6), 595–605.

Tajfel, Henri, “Experiments in Intergroup Discrimination,” Scientific American, 1970, 223 (5), 96–102.

___, Michael Billig, R.P. Bundy, and Claude Flament, “Social Categorization and Intergroup Behavior,” European Journal of Social Psychology, 1971, 1 (2), 149–178.

Yzerbyt, Vincent Y., A. Coull, and S. J. Rocher, “Fencing off the deviant: the role of cognitive resources in the maintenance of stereotypes,” Journal of Personality and Social Psychology, 1999, 77 (3), 449–462.

(All references in this appendix appear in the bibliography of the published manuscript.)