Trust, Reciprocity and Favors in Cooperative Relationships

Atila Abdulkadiroğlu Kyle Bagwell*

May 9, 2005

Abstract

We study trust, reciprocity and favors in a repeated trust game with private information. In our main analysis, players are willing to exhibit trust and thereby facilitate cooperative gains only if such behavior is regarded as a favor that must be reciprocated, either immediately or in the future. Private information is a fundamental ingredient in our theory. A player with the ability to provide a favor must have the incentive to reveal this capability, and this incentive is provided by an equilibrium construction in which favors are reciprocated. We also offer the novel prediction that the size of a favor owed may decline over time, as neutral phases of the relationship are experienced. Indeed, a favor-exchange relationship with this feature offers a higher total payoff than does a simple favor-exchange relationship. We also describe specific circumstances in which a relationship founded on favor exchange may be inferior to a relationship in which an infrequent and symmetric punishment motivates cooperative behavior. Finally, we show that a hybrid relationship, in which players begin with a honeymoon period and then either proceed to a favor-exchange relationship or suffer a symmetric punishment, can also offer scope for improvement.

*Atila Abdulkadiroğlu is an Associate Professor of Economics, Department of Economics, Columbia University; Kyle Bagwell is Kelvin J. Lancaster Professor of Economic Theory (Department of Economics) and Professor of Finance and Economics (School of Business), Columbia University. We would like to thank Alberto Martin for his excellent research assistance and Rajiv Sethi, Ennio Stacchetti, Eric Verhoogen and seminar participants at Maryland and NYU for their valuable comments.

1. Introduction

A substantial experimental literature confirms that subjects exhibit trust and practice reciprocity. For example, Berg, Dickhaut and McCabe (1995) consider the *trust game*, in which one subject (the investor) has income and can invest by sending some or all of this income to another subject (the trustee), where the income sent grows en route and is received as a larger amount. The trustee may then choose to reciprocate, by returning some income to the investor. An investor that gives income to the trustee has shown trust, since the investor has incurred a cost and cannot be sure that the trustee will reciprocate. Berg, Dickhaut and McCabe find that subjects often exhibit trust and practice reciprocity. In particular, evidence of positive reciprocity is reported: many subjects reward kind behavior with a kind response. de Quervain (2004) et al. study a modified trust game, in which the investor can incur a cost and punish the trustee if the latter does not reciprocate. They observe that such punishments often occur, indicating that subjects may also practice negative reciprocity, whereby they punish unkind behavior with an unkind response.[1]

Psychological and anthropological studies also report that an important category of human social interactions emphasizes trust and reciprocity. For example, Fiske (1992) surveys ethnographic field work and experimental studies and argues that virtually all human social interactions can be described in terms of four patterns, each with a distinctive psychological basis. One pattern is called "equality matching" (EM). As Fiske (1992, p. 703) states, "The operating principle is that when people relating in an EM mode receive a favor, they feel obligated to reciprocate by returning a favor." In EM relationships, people keep track of the imbalances between them and engage in score keeping, with the restoration of balance being a primary aspiration. As Fiske (1992, p. 705) puts it, "People think about *how much* they have to give to reciprocate or compensate others or come out even with them. EM always entails some kind of additive tally of who owes what and who is entitled to what." Following Blau (1964), we may think of an initial favor as involving trust and the expectation of subsequent reciprocation.

Why do humans exhibit trust and practice reciprocity? One hypothesis is

[1] de Quervain, et al (2004) also use PET scans and investigate the neural basis of punishment, finding evidence that humans derive satisfaction from the punishment of defectors. See also King-Casas, et al (2005) for related evidence of positive and negative reciprocity in a multi-round trust game. As Fehr and Gatcher (2000a) observe, subjects in public-good games may also practice negative reciprocity. See Camerer (2003) and Fehr and Gachter (2000b) for excellent surveys of experimental work.

that trusting and reciprocal behaviors facilitate gains from cooperation. Person a may be willing to incur some cost to assist Person b when the benefit to Person b exceeds the cost to Person a, if Person a believes that this favor will one day be returned when the roles are reversed. From this perspective, these behaviors generate cooperative gains for calculating, self-interested individuals. Over time, an instinct for trust and reciprocity may thus have evolved among our ancestors, as a means to facilitate cooperation and thereby promote survival among individuals that interact frequently.[2]

In this paper, we assume self-interested players and use repeated-game theory to study trust and reciprocity. In making the assumption of self-interested players, our purpose is not to deny that individuals have social preferences that perhaps include an instinct for trust and reciprocity. Rather, our purpose is to better understand the underlying advantages that trusting and reciprocal behaviors afford when gains from cooperation are present.[3]

It is well known that repeated interaction can foster cooperation. Consider a *repeated trust game* with two players in which, in each period, one player is randomly selected as the investor. Given that the investment technology is value enhancing, the players recognize a gain from cooperation: if in each period the investor were to send all income to the trustee, then the players would both enjoy higher payoffs than they would were instead the investor always to keep all income. For sufficiently patient players, this cooperative behavior can be enforced as a subgame perfect equilibrium, if the players threaten that any deviation induces a reversion to the autarkic Nash equilibrium of the stage game. The repeated trust game thus gives a direct interpretation of positive reciprocity: in the cooperative equilibrium, when a player is selected as the investor, that player is willing to exhibit trust provided that the other player has always done so in the past when

[2] This instinct may explain why experimental subjects often exhibit trust and practice reciprocity even in one-shot settings. At the same time, self-interested calculations remain as an influence on contemporary behavior. For example, Engle-Warnick and Slonim (2003) find that experimental subjects tend to show more trust when they understand that the relationship is on-going. As Seabright (2004, Chapter 3) emphasizes, we may regard instinctive and calculating reciprocity as complementary human virtues: an ability to calculate enables an individual to adopt the behavior that is most effective for a given social environment, while an instinct for reciprocity inspires the trust of other individuals. For further discussion of the evolutionary foundations of the human instincts for cooperation and reciprocity, see Field (2002), Pinker (2002) and Ridley (1997). See also Sethi and Somanathan (2003) for a survey of evolutionary game theory models that analyze cooperation.

[3] Our approach thus shares important themes with the sociology literature on social exchange theory. See Blau (1964) and Coleman (1988).

the roles were reversed.

But the repeated trust game fails in other respects. First, while the repeated trust game admits a large set of equilibria, the cooperative equilibrium is simple, efficient and thus focal; however, this equilibrium does not offer an interpretation of EM relationships. For all histories in which no deviation occurs, the same equilibrium behavior (namely, a transfer of all income by the current investor) is specified. Behavior in the cooperative equilibrium is thus insensitive to the balance of favors owed. Second, the cooperative equilibrium also fails to offer an interpretation of negative reciprocity. On the equilibrium path, negative reciprocity is never induced, since the investor always transfers all income and thus only behaves in a kind way. Finally, at a descriptive level, the repeated trust game is limited in that it precludes the realistic and important possibility that players are asymmetrically informed with regard to their respective abilities to provide favors at a given point in time.

We argue here that the predictive power and realism of the repeated-game analysis can be greatly enhanced, if the repeated trust game is modified to include private information. We thus study the *repeated trust game with private information.* In the stage game, either player a is given income, player b is given income, or neither player is given income. Each player is privately informed as to whether or not he is the investor. Thus, if a player does not receive income, then the player does not observe whether neither player received income or the other player received income. Next, if one player receives income, then that player may choose to exhibit trust and invest by sending some or all of his income to the other player. If a transfer is made, then the level of the investment is publicly observed; however, while the investment is value enhancing on average, the outcome is random. The investment either succeeds or fails, and the investment is completely lost when it fails. The trustee privately observes the investment outcome. If the investment is successful, then the trustee can reciprocate within the period and send some (or even all) of the returns back to the investor. Thus, if the investor exhibits trust and reciprocation does not occur within the current period, then the investor does not observe whether the trustee elected not to immediately reciprocate or the investment failed.

This game is highly stylized, but it serves to introduce two key incentive problems. First, when a player is selected as the investor, the gains from cooperation can be enjoyed only if this player has incentive to exhibit trust and thereby reveal that he is the investor. This constraint suggests a role for EM relationships. If a player reveals that he is the investor and exhibits trust, then the player has given a favor to the other player. As the investor always has the option of pretending that he has not received income, some gain must be anticipated when a favor is extended in this way. This gain may take the form of a favor that the current trustee now owes the current investor. This favor may be paid in the current period if the investment is successful, or it may be paid in the future if the players then adopt a path of play for the continuation that favors the current investor. We think of the former payment as *immediate reciprocity* and the latter payment as *dynamic reciprocity*. Second, when a successful investment is made, if the cooperative equilibrium calls for immediate reciprocity, then the trustee must be given incentive to reciprocate and thereby reveal that the investment is successful.

More generally, the repeated trust game with private information serves as a simple framework within which to explore the provision of favors among individuals in on-going relationships. A self-interested individual that extends a favor naturally hopes for some gain in return. But the individual may not be able to determine when the recipient is in a position to return the favor. The recipient may not be in a position to reciprocate immediately, and we capture this possibility by assuming that the investment may be unsuccessful. As well, while at some point in the future the recipient will be in a position to pay the favor, the individual may not be able to observe the date at which this occurs. Further, the individual may find that he is in a position to extend another favor before having been paid for his last favor. We capture these possibilities with the assumption that each player is privately informed as to whether he is the investor, where there is a chance that neither player is the investor.

In our formal analysis, we follow Abreu, Pearce and Stacchetti (1986, 1990) and use the concept of self generation. We thus look for a set of payoffs that can be enforced using only continuation payoffs that are drawn from that set. Athey and Bagwell (2001) also use this concept to examine a repeated game in which colluding firms are privately informed about their respective costs.[4] For a two-type model, they implement a self-generating line of payoffs along which total payoffs sum to the first-best value. Specifically, they show that the "corner" utility pairs can be implemented and then observe that all other payoffs along the line can be achieved using a public-randomization device. The implementation is

[4]For related contributions in the collusion literature, see Aoyagi (2003), Athey and Bagwell (2004), Athey, Bagwell and Sanchirico (2004) and Skryzpacz and Hopenhayn (2004). Related themes are also explored in macroeconomics; early contributions include Green (1987) and Wang (1995). See Fudenberg, Levine and Maskin (1994) for an analysis of a general class of repeated games with private information.

symmetric, if each corner is selected with equal probability at the start of the game. Similarly, we characterize the properties of the *highest symmetric self-generating line* (HSSGL) of the repeated trust game with private information.

A first set of findings concerns properties of symmetric self-generating lines. One finding is that a HSSGL does not achieve first-best payoffs. Another finding is that dynamic reciprocity is required for the implementation of any payoff pair along a symmetric self-generating line. In particular, a player expects a better future payoff if in the current period the player is an investor that does not receive immediate reciprocity than if in the current period that player is a trustee that does not extend immediate reciprocity.

We next consider the implementation of a HSSGL. Following Athey and Bagwell (2001), we consider the implementation of the corner utility pair that represents the lowest (highest) payoff for player a (player b) along a HSSGL. We show that this implementation initially requires that player a transfers all income if he is the investor, while player b transfers less than all income if he is the investor. If player a is selected as the investor and transfers all income, we may understand that player a's favor is paid, and the game moves to the opposite corner utility pair, at which player a (player b) receives his highest (lowest) payoff along this HSSGL. The opposite corner is implemented analogously. Here, it is player b that owes the favor. In this way, when a player owes a favor, the player is induced to admit that he is the investor and pay the favor, since the player gains the future reward of becoming the favored player. The implementation reflects an intertemporal balancing of favors that is broadly consistent with EM relationships.

A novel feature of the implementation arises following a period in which neither player receives income. In the implementation of the corner utility pair that gives the lowest payoff for player a, if neither player reports income, a new utility pair is induced in the following period, where the new pair favors player b but to a smaller extent than did the initial corner utility pair. This intermediate utility pair can be realized, in expectation, by using a public-randomization device. This feature suggests a novel prediction: the "size" of favor that is owed diminishes in expectation when a "neutral" state (i.e., a state in which neither player has income) is encountered. Another interesting feature of our implementation is that it does not require the use of immediate reciprocity.

We next construct an implementation of a HSSGL that does not require a public-randomization device. The implementation works as above, except that we now directly implement the intermediate utility pair that follows a neutral state. Our implementation indicates that the size of favor that is owed diminishes with the realization of every successive neutral state. Thus, if player a owes a favor to player b, then player a transfers all income if player a is immediately selected as the investor; however, if a neutral state is experienced first and player a is selected as the investor in the next period, then player a can fulfill his favor by transferring less than all income. Similarly, if two neutral states are encountered and then player a is selected as investor, then player a can fulfill his favor with an even smaller transfer. Intuitively, this process gives player b incentive to transfer the required amount when he is the investor, since otherwise a neutral state would be observed and in the next period player b would be favored to a smaller extent.

Thus, following several neutral periods, the disfavored player continues to acknowledge that a favor is owed but holds that less is now required to fulfill the favor. One may imagine the disfavored player remarking: "Yeah, but what have you done for me lately?" The prediction that the size of the favor owed deteriorates over time when neutral states are experienced is novel to our framework and does not appear to be a feature of the EM relationships that Fiske (1992) describes. At the same time, the prediction has some intuitive appeal. We show as well that this prediction is necessary for the implementation of a HSSGL.

We next compare the total payoffs achieved in a HSSGL with those enjoyed under other benchmarks. We first introduce the benchmark of a *simple EM relationship*: in such a relationship, once a player provides a favor, the player is unwilling to provide any further favors - of any size - until the favor is paid in full (i.e., until the other player provides a favor of equal size). We show that a simple EM relationship can be described in terms of a symmetric self-generating line; however, this line is not a HSSGL. Thus, our implementation of a HSSGL offers strictly higher total payoff than does any simple EM relationship. Intuitively, our implementation captures a *sophisticated EM relationship* in that it induces a player to transfer some income even when that player provided the most recent favor. As explained above, it is precisely this feature that implies that favors must deteriorate in size following the experience of neutral states.

Second, we relax the requirement of self-generating lines and characterize the set of *strongly symmetric equilibria* (SSE). In such equilibria, asymmetric continuation values are not allowed, and so players cannot use future favors as they do in a HSSGL. But the players can provide incentives for trust, if a period without an investor triggers a symmetric punishment. Likewise, the players can provide incentives for immediate reciprocity, if a symmetric punishment may be initiated once an investment is not reciprocated. We show that players are unable to cooperate in SSE, if both informational asymmetries are significant (i.e., if a period without an investor often occurs and investments are often unsuccessful). But, when either informational asymmetry is less significant, the players can construct SSE with payoffs that exceed those under autarky. Indeed, the optimal SSE may then even offer a total payoff exceeding that attained in a HSSGL.

Intuitively, if the probability that neither player is selected as the investor is small, then the players may impose a severe and symmetric punishment when neither player reports income. This punishment gives each player a great incentive to be honest when he is the investor; furthermore, the punishment is rarely experienced along the equilibrium path. It is then possible to use such a construction to generate equilibrium payoffs that lie above the HSSGL. One interesting feature of this construction is that it offers an equilibrium interpretation of negative reciprocity. If neither player is "nice" to the other, then the relationship runs the risk of deteriorating, with both players being "mean" to each other in the future.

Third, we build from the HSSGL and SSE constructions and introduce the benchmark of a *hybrid equilibrium*. In such an equilibrium, players begin with a "honeymoon" period that is characterized by a high level of trust. If in the first period some player is chosen as the investor and makes the appropriate transfer, then the players proceed in the next period and thereafter to implement a HSSGL. The player that made the first-period investment begins as the favored player. Alternatively, if no income is reported in the first period, then the players suffer a symmetric punishment ("break up"). Thus, in a hybrid equilibrium, sophisticated EM relationships and negative reciprocity are both predicted.

We first compare the optimal hybrid equilibrium with equilibria that implement a HSSGL. For a large set of parameters, we show that a honeymoon period is valuable: the optimal hybrid equilibrium offers a greater total payoff than is achieved in a HSSGL. The underlying insight here is that the first period is unique, since then players are not encumbered by obligations that are derived from past favors; hence, players may exhibit full trust in the first period. In a second comparison, we show that a large set of parameters also exists over which the optimal hybrid equilibrium offers a greater total payoff than is obtained in the optimal SSE. We show, however, that the optimal SSE can offer a greater total payoff if the probability that neither player is selected as the investor is sufficiently small.

Our paper is related to two recent papers that provide theoretical analyses of favors. First, in a collusion model, Athey and Bagwell (2001) provide a theory of "future market share favors." In the present paper, we study a different repeated game and provide equilibrium interpretations for favors that decline in size as neutral phases are experienced, negative reciprocity and honeymoon phases. Second, Mobius (2001) also studies equilibrium favor provision when the ability to provide a favor is private information. Mobius studies a continuous-time game and focuses on the existence of particular equilibria that specify intuitive rules for favor provision. Our paper is also related to Watson's (1999, 2002) recent work on long-term partnerships with persistent and two-sided incomplete information. In this setting, a role for learning is present, and players may "start small;" by contrast, in our model, a role for learning does not arise, and indeed players may "start big" with an initial honeymoon period.

The paper is organized as follows. Section 2 presents the model. Section 3 provides our findings for HSSGL. Section 4 considers simple EM relationships. Section 5 contains our analysis of SSE. Section 6 characterizes optimal hybrid equilibria. Section 7 concludes. Remaining proofs are located in the Appendix.

2. The Model

We study a stylized model with two players, a and b. In the stage game, either player a is given an income of \$1, player b is given an income of \$1, or neither player is given an income. The former two events each occur with probability $p \in (0, 1/2)$ and the latter event thus occurs with probability $1 - 2p$. In any period, a player who receives income becomes an investor. Each player is privately informed as to whether or not he is the investor. Thus, if a player does not receive income, then the player does not observe whether neither player received income or the other player received income. If a player receives income, then that player may choose to exhibit trust and invest by sending any $x \in [0, 1]$ to the other player. The transfers between players are publicly observed. The outcome of the investment is random. The investment either succeeds or fails, where success occurs with probability $q < 1$. The investment produces kx when it is successful, and the investment is completely lost otherwise. We assume $qk > 1$; that is, the investment is value enhancing on average. The trustee is the player to whom an investment is sent. The trustee privately observes the investment outcome. If the investment is successful, then the trustee can reciprocate within the period and send some (or even all) of the returns back to the investor. Thus, if the investor exhibits trust and reciprocation does not occur within the current period, then the investor does not observe whether the trustee elected not to reciprocate in the current period or the investment failed. We assume risk neutral players in order to abstract from insurance arrangements, and we let $\beta \in (0, 1)$ denote the players' common discount factor.

Let t denote the time index. For $i \in \{a, b\}$, let $w_t^i = 1$ if player i receives income and $w_t^i = 0$ otherwise. Player i privately observes $W_t^i = \{w_z^i\}_{z=1}^t$. Let $\tau_t = (j, x)$ if player j invests in the amount of $x > 0$ in period t and $\tau_t = 0$ otherwise. Both players observe $T_t = \{\tau_z\}_{z=1}^t$. Let $\kappa_t^i = 1$ if player j invests in player i and the investment succeeds, $\kappa_t^i = 0$ if player j invests in player i and the investment fails, and $\kappa_t^i = \emptyset$ if player j does not invest in player i. The trustee privately observes $K_t^i = \{\kappa_z^i : \kappa_z^i \neq \emptyset\}_{z=1}^t$. Since κ_z^i is relevant only when player j invests, we do not consider $\kappa_z^i = \emptyset$ as part of player i's private history. Let $\theta_t = (i, r)$ if player j invests and player i reciprocates in the amount of $r > 0$, and $\theta_t = 0$ otherwise. Both players observe $R_t = \{\theta_z\}_{z=1}^t$. Note that $\theta_t = 0$ when $\tau_t = 0$; that is, if there is no investment, then there is no reciprocity by the other player either.

Thus, the private history of player i at time t is denoted $h_t^i = (W_t^i, K_t^i)$, and the public history is denoted $H_t = (T_t, R_t)$. Let $\mathcal{H}_t^i$ denote the set of possible private histories, and $\mathcal{H}_t$ denote the set of public histories at t.

A strategy σ_i for player i consists of an investment decision $I_t^i : \mathcal{H}_t^i \times \mathcal{H}_{t-1} \to [0, 1]$, such that $I_t^i(h_t^i, H_{t-1}) = 0$ when $w_t^i = 0$, and $I_t^i(h_t^i, H_{t-1}) \in [0, 1]$ if $w_t^i = 1$; and a reciprocity decision $R_t^i : \mathcal{H}_t^i \times (\mathcal{H}_{t-1}, \tau_t) \times [0, 1] \to [0, k]$ such that $R_t^i(h_t^i, H_t, \tau_t) = 0$ if $\tau_t \neq (j, I_t^j)$ or $\kappa_t^i = 0$ and $R_t^i(h_t^i, H_t, \tau_t = (j, I_t^j)) \in [0, kI_t^j]$. Note that $\tau_t = (j, I_t^j)$ if and only if $I_t^j > 0$, and $\theta_t = (i, R_t^i)$ if and only if $R_t^i > 0$.

Following Fudenberg, Levine and Maskin (1994), we use the solution concept of *perfect public equilibrium* (PPE). A strategy for player i is *public* if at every period t, it depends only on player i's current-period private information, (w_t^i, κ_t^i), and the public history, H_{t-1}. A PPE is a profile of public strategies that forms a Nash equilibrium at any date, given any public history.

Following Abreu, Pearce and Stacchetti (1990), we can define an operator B which yields the set of PPE values, Ψ^*, as the largest self-generating set.[5] This

[5] To this end, let us note that players' strategy spaces are effectively finite. Using terminology provided by Athey, Bagwell and Sanchirico (2004), we say that a deviation is an off-schedule deviation (i.e., observable, as a deviation, to other players) if it contains a positive investment or positive reciprocity that differs from the equilibrium value. Such deviations can be avoided by the threat of reverting to autarky. Thus, a deviation is relevant to our analysis only if it is an on-schedule deviation (i.e., unobservable, as a deviation, to other players). In such a deviation, a player selects zero investment or zero reciprocity, even though the equilibrium strategy calls for a positive value. A player effectively chooses between the action that is suggested by his equilibrium strategy and an on-schedule deviation with zero investment or zero reciprocity. Therefore, a player reveals his income or the investment outcome truthfully in a PPE. Equivalently, if an income level or investment outcome represents the player's type, then

operator is defined as follows:

For any set $\Psi \subset \Re^2$, consider the following mapping: $B(\Psi) = \{(u,v) : \exists$ $(u_{i\theta}, v_{i\theta}) \in \Psi$, for $i \in \{a,b\}$ and $\theta \in \{0,1\}; (u_o, v_o) \in \Psi; x, y \in [0,1]$, $r \in [0, kx]$ and $s \in [0, ky]$ such that:

$$IR : u, v, u_{i\theta}, v_{i\theta}, u_o, v_o \geq \frac{p}{1-\beta} \tag{2.1}$$

$$IC_x^a : 1 - x + q(r + \beta u_{a1}) + (1-q)\beta u_{ao} \geq 1 + \beta u_o \tag{2.2}$$

$$IC_y^b : 1 - y + q(s + \beta v_{b1}) + (1-q)\beta v_{bo} \geq 1 + \beta v_o \tag{2.3}$$

$$IC_\theta^a : ky - s + \beta u_{b1} \geq ky + \beta u_{bo} \tag{2.4}$$

$$IC_\theta^b : kx - r + \beta v_{a1} \geq kx + \beta v_{ao} \tag{2.5}$$

$$\begin{aligned} PK^a \quad : \quad & u = p[1 - x + q(r + \beta u_{a1}) + (1-q)\beta u_{ao}] \\ & + p[q(ky - s + \beta u_{b1}) + (1-q)\beta u_{bo}] + (1-2p)\beta u_o \end{aligned} \tag{2.6}$$

$$\begin{aligned} PK^b \quad : \quad & v = p[1 - y + q(s + \beta v_{b1}) + (1-q)\beta v_{bo}] \\ & + p[q(kx - r + \beta v_{a1}) + (1-q)\beta v_{ao}] + (1-2p)\beta v_o\}. \end{aligned} \tag{2.7}$$

Observe that we use u to denote player a's payoff, x to denote investment level by player a, and r to denote the amount that player b reciprocates when the investment is successful. Similarly, we use v to denote player b's payoff, y to denote the investment level by player b, and s to denote the amount that player a reciprocates when the investment is successful. The utility pairs that are induced may depend on the public path of play: we use (u_o, v_o) to denote the continuation values that are induced when neither player reports income, and we use $(u_{i\theta}, v_{i\theta})$ to denote the continuation values that are induced when player $i \in \{a,b\}$ invests and the other player reciprocates ($\theta = 1$) or not ($\theta = 0$). For a given Ψ, we will say that $\{x, y, r, s, u_{i\theta}, v_{i\theta}, u_o, v_o\}$, for $i = a, b$ and $\theta = 0, 1$, *implements* a utility pair (u, v) if all of the constraints above are satisfied.

We now mention two important benchmarks. First, the Nash equilibrium of the static game is autarky: no player invests, and so each player expects a payoff of p. In the Nash benchmark, in every period, the players use the Nash equilibrium

a player's action space consists of this finite type space. We can thus directly apply the dynamic programming techniques of Abreu, Pearce and Stacchetti (1990).

of the stage game. The payoffs for the repeated game are then $u = v = \frac{p}{1-\beta}$, and so $u + v = \frac{2p}{1-\beta}$. The Nash benchmark payoff is used in the IR constraint above, since autarky is the worst punishment. Second, given our assumption that $qk > 1$, the first-best benchmark occurs when each player invests all of his income. The players' joint per-period payoff is then $2pqk$. Thus, in the first-best benchmark, $u + v = \frac{2pqk}{1-\beta}$.

We observe that the first-best benchmark could be achieved by patient players, if either informational asymmetry were absent. If some player always receives income (i.e., $p = 1/2$), then in any period it is common knowledge among the two players as to which player received income. When the players are sufficiently patient, they can then support an equilibrium with first-best payoffs, by threatening an infinite reversion to the autarky equilibrium of the static game in the event that a player with income does not invest all income. Likewise, if an investment is always successful (i.e, $q = 1$), then in any period it is common knowledge among the two players that the trustee has received $k > 1$ and is thus able to reciprocate immediately this entire quantity. If the players are sufficiently patient, they can again support an equilibrium with first-best payoffs, by threatening an infinite reversion to the autarky equilibrium of the static game in the event that the trustee does not immediately reciprocate the quantity $k > 1$.

3. Self-Generating Lines

In this section, we consider PPE that can be characterized in terms of a self-generating line. We characterize the necessary features of such equilibria, finding that they involve trust and dynamic reciprocity. We then characterize some necessary features of equilibria that are constructed along a highest self-generating line. Next, we provide an implementation of the utility pairs that rest upon a highest self-generating line. Finally, we characterize the unique features of such implementations.

3.1. Self-Generating Lines: Necessary Features

A *line* (segment) is defined by a closed and convex set of utility pairs, (u, v), that sum to the same total; thus, a line is defined by $(\underline{u}, \overline{v}) \to (\overline{u}, \underline{v})$ where $T \equiv u + v$ along the line. A *self-generating line* is a line such that, for any utility pair (u, v) on the line, the pair can be implemented using some (x, y, r, s) and continuation values, $(u_{i\theta}, v_{i\theta}, u_o, v_o)$, where the continuation values are all drawn from the given line. Thus, if a pair (u, v) is on a self-generating line, with $u + v = T$, then it is necessary that $u_o + v_o = T$ and $u_{i\theta} + v_{i\theta} = T$, for all i and θ. Our focus in this subsection is on the necessary features of self-generating lines.

We first define the notions of trust and reciprocity in our model. Fix an implementation of a utility pair, (u, v), that rests on a self-generating line. We define the *level of trust* in the implementation as $x + y$, and we say that player a (b) exhibits *more trust* if $x > y$ ($x < y$). Likewise, we will say that player a (b) exhibits *immediate reciprocity* if $s > 0$ ($r > 0$). Finally, we say that the implementation embodies *dynamic reciprocity* if $u_{ao} > u_{bo}$ and $v_{bo} > v_{ao}$.

We begin by considering the level of trust along a self-generating line. Our first finding is that the level of trust is fixed along a self-generating line.

Lemma 3.1. *Along a self-generating line, total payoff is given as*

$$T = \frac{p[2 + (x + y)(qk - 1)]}{1 - \beta}, \tag{3.1}$$

and so the same level of trust, $x + y$, is used when implementing any pair on the self-generating line.

Proof: Using (2.6) and (2.7), if we can implement a pair (u, v) on a self-generating line, then

$$\begin{aligned} T \equiv\ & u + v = p\{2 - x - y + q(r + s + \beta(u_{a1} + v_{b1})) + (1 - q)\beta(u_{ao} + v_{bo}) \\ & +q[k(x + y) - (r + s) + \beta(u_{b1} + v_{a1})] + (1 - q)\beta(u_{bo} + v_{ao})\} \\ & +(1 - 2p)\beta(u_o + v_o) \end{aligned}$$

Rearranging terms and using $u_o + v_o = T$ and $u_{i\theta} + v_{i\theta} = T$, we may solve for T and confirm (3.1). ■

We now consider whether a self-generating line can take the form of a self-generating point. In other words, can we implement a single utility pair, (u, v), using continuation values that satisfy $(u_{i\theta}, v_{i\theta}) = (u, v)$ and $(u_o, v_o) = (u, v)$? Our next finding confirms that the opportunities for such an outcome are quite limited.

Lemma 3.2. *A point (u, v) constitutes a self-generating line if and only if $u = v = \frac{p}{1-\beta}$.*

Proof: Suppose $u_{i\theta} = u_o = u$ and $v_{i\theta} = v_o = v$. Using (2.2), it follows that $qr \geq x$. Likewise, (2.3) implies that $qs \geq y$. Next, (2.4) and (2.5) respectively imply that $0 \geq s$ and $0 \geq r$, from which it follows (from feasibility) that $s = 0 = r$. It thus follows that $0 \geq y$ and $0 \geq x$, from which it follows (from feasibility) that $x = 0 = y$. Using $u_{i\theta} = u_o = u$ and $s = r = x = y = 0$, we may solve (2.6) for u, finding that $u = \frac{p}{1-\beta}$. ■

This finding indicates that a point is self-generating only if it entails *no trust* (i.e., $x = y = 0$) and thus results in the Nash (autarky) payoff.

We consider now the implementation of the corner of a self-generating line, $(\underline{u}, \overline{v})$. We focus here on *symmetric self-generating lines*, where $\underline{u} = \underline{v}$, and $\overline{u} = \overline{v}$. Our finding places some structure on x and y.

Lemma 3.3. *Consider any symmetric self-generating line with $T > \frac{p}{1-\beta}$. Let $(\underline{u}, \overline{u})$ denote the point on the line at which player a's utility is minimized. The implementation of $(\underline{u}, \overline{u})$ requires $x > y$, and so player a exhibits more trust.*

Proof: Given $T > \frac{p}{1-\beta}$, the line must not be a point (by Lemma 3.2). Thus, $\underline{u} < T/2$. Using Lemma 3.1, it follows that

$$\underline{u} < \frac{p[2 + (x + y)(qk - 1)]}{2(1 - \beta)}. \tag{3.2}$$

Next, using (2.2) and (2.4), we have from (2.6) that

$$\begin{aligned} \underline{u} &\geq p(1 + \beta u_o) + p\{q(ky + \beta u_{bo}) + (1 - q)\beta u_{bo}\} + (1 - 2p)\beta u_o \\ &= p + pqky + (1 - p)\beta u_o + p\beta u_{bo} \\ &\geq p + pqky + (1 - p)\beta \underline{u} + p\beta \underline{u} \\ &= p + pqky + \beta \underline{u}, \end{aligned}$$

where in the second inequality we use $u_o \geq \underline{u}$ and $u_{bo} \geq \underline{u}$. It follows that

$$\underline{u} \geq \frac{p + pqky}{1 - \beta}. \tag{3.3}$$

Using (3.2) and (3.3), it is clearly necessary that

$$f(x, y) \equiv \frac{p[2 + (x + y)(qk - 1)]}{2(1 - \beta)} - \frac{p + pqky}{1 - \beta} > 0. \tag{3.4}$$

Calculations confirm the following inequalities: $f_x > 0 > f_y$ and $f(x,x) \leq 0$. By the latter inequality and (3.4), $x = y$ is not possible. Likewise, if $x < y$, then a contradiction is reached with (3.4), since the inequalities just stated then imply that $f(x,y) < 0$. ■

Thus, a player's utility can be driven to its minimum level along a self-generating line only if that player exhibits more trust. In essence, the trust that the player shows is the means through which that player's utility is reduced.

In our model, players can achieve a first-best outcome only if they exhibit *total trust* ($x = y = 1$). Building on Lemma 3.3, we now establish that players are not able to use a symmetric self-generating line to achieve a first-best outcome.

Corollary 3.4. *There does not exist a symmetric self-generating line that yields first-best total payoffs.*

The argument is simple. By Lemma 3.1, if a self-enforcing line generates first-best total payoffs, then $x + y = 2$ is required, so that total payoff is $T = 2pqk/(1-\beta)$. Given $x \in [0,1]$ and $y \in [0,1]$, this means that each utility pair on the self-generating line is implemented using $x = y = 1$. By Lemma 3.2, this total payoff cannot be achieved with a self-generating point. Further, as shown in Lemma 3.3, when a symmetric line is used, we can implement the corner only if $x < y$.

We consider next a necessary condition that is associated with the implementation of *any* (u, v) along a symmetric self-generating line. This condition establishes a key relationship between the level of trust and dynamic reciprocity.

Proposition 3.5. *Consider any symmetric self-generating line and associated value* $x + y$*. For any* (u, v) *on this line to be implemented, it is necessary that*

$$u_{ao} - u_{bo} \geq \frac{x+y}{\beta}. \tag{3.5}$$

Proof: Consider the implementation of any utility pair (u, v) along a symmetric self-generating line. Using $u_o + v_o = T$ and $u_{i\theta} + v_{i\theta} = T$, we may rewrite (2.3) as

$$1 - y + q[s - \beta u_{b1}] - (1-q)\beta u_{bo} \geq 1 - \beta u_o. \tag{3.6}$$

We may now add (2.2) and (3.6) to obtain

$$u_{ao} - u_{bo} + q[u_{a1} - u_{b1} - u_{ao} + u_{bo}] \geq \frac{x + y - q(r+s)}{\beta}. \tag{3.7}$$

In similar fashion, using $u_{i\theta} + v_{i\theta} = T$, we may rewrite (2.5) as

$$kx - r - \beta u_{a1} \geq kx - \beta u_{ao}. \tag{3.8}$$

We may now add (2.4) and (3.8) to obtain

$$\frac{-(r+s)}{\beta} \geq u_{a1} - u_{b1} - u_{ao} + u_{bo}. \tag{3.9}$$

Combining (3.7) and (3.9), we see that implementation of (u, v) is possible only if

$$u_{ao} - u_{bo} - \frac{q(r+s)}{\beta} \geq \frac{x + y - q(r+s)}{\beta}.$$

Equivalently, (3.5) must hold. ■

This proposition reveals two important lessons. First, if players achieve a positive level of trust, then dynamic reciprocity is necessary for the implementation of *any* utility pair along a symmetric self-generating line. In words, when the two players are cooperating along a line, player a must do better tomorrow when player a made an investment today and player b did not reciprocate than when player b made an investment today and player a did not reciprocate. It is perhaps surprising that dynamic reciprocity is required. After all, players have available some instruments (r, s) with which to achieve immediate reciprocity. The important point, though, is that these instruments are useless when the investment is unsuccessful, as then there is no money to return, and the incentive problem associated with this state (IC_θ^a and IC_θ^b) thus requires reciprocity through the remaining instruments - future continuation values. Second, as the players increase the level of trust (i.e., as they implement larger values for $x + y$), incentive compatibility implies that the degree of dynamic reciprocity (i.e., $u_{ao} - u_{bo}$) must also grow. Greater trust is associated with greater dynamic reciprocity.

3.2. Highest Symmetric Self-Generating Lines: Necessary Features

We turn next to consider some necessary conditions for implementing a *highest symmetric self-generating line* (*HSSGL*). A HSSGL is a symmetric self-generating line that achieves the highest value for $T = u + v$. Following Athey and Bagwell (2001), we focus on the implementation of a corner utility pair, $(u, v) = (\underline{u}, \overline{u})$, of a HSSGL. By the symmetry of the environment, if we can implement the corner pair $(\underline{u}, \overline{u})$, then we can also implement the other corner pair, $(\overline{u}, \underline{u})$. Assuming that players have access to a public randomization device, we can then implement any utility pair along the HSSGL as a convex combination of the two corners. Further, as shown in the next subsection, once we can implement a corner utility pair of a HSSGL, we can devise an implementation of a HSSGL that does not require a pubic randomization device.

Let $\{x, y, r, s, u_{i\theta}, v_{i\theta}, u_o, v_o\}$ implement $(\underline{u}, \overline{u})$ on a HSSGL. Our game allows for a rich set of instruments, and a given utility pair on a HSSGL may have multiple implementations. In particular, $(\underline{u}, \overline{u})$ on a HSSGL may admit distinct implementations. It is also possible that multiple HSSGL's may exist. By definition, all such lines achieve the same value for $T = u + v$ and thus $x + y$; but one HSSGL may be wider than another, and so the corner utility pairs, $(\underline{u}, \overline{u})$ and $(\overline{u}, \underline{u})$, may differ across HSSGL's. Our characterizations of necessary features thus take different forms. Our strongest characterizations hold for any HSSGL and for any implementation of the associated $(\underline{u}, \overline{u})$. But it is also useful to offer characterizations of necessary features that apply only to certain HSSGL's. In particular, by characterizing the necessary features of an implementation of the *widest HSSGL* (i.e., the HSSGL for which $\overline{u} - \underline{u}$ is greatest), we acquire insights that then enable us to construct a HSSGL.[6]

Henceforth, we maintain the assumption that β is sufficiently large, so that

$$\beta \geq \beta^* \equiv \frac{1}{1 + p(qk - 1)}. \tag{3.10}$$

For any $\beta > 0$, this constraint is sure to hold for qk sufficiently large. At the other extreme, this constraint can only hold for β near unity when qk is close to unity.

We begin by confirming that a HSSGL must achieve some trust (i.e., $x+y > 0$) and thus generate a total payoff that exceeds the Nash autarky payoff (i.e., $T > 2p/(1 - \beta)$). To establish these points, we construct a symmetric self-generating line in which $x + y = 1$.[7]

Lemma 3.6. *There exists a symmetric self-generating line, in which $x + y = 1$ and thus $T = p[1 + qk]/(1 - \beta) > 2p/(1 - \beta)$.*

[6] Given a symmetric self-generating line, it is straightforward to use the techniques of Abreu, Pearce and Stacchetti (1990) and establish the existence of a widest self-generating line that contains the given line. The existence of a widest HSGGL is used below in the proof of Lemma 3.9, for example.

[7] See also Proposition 4.1 below, which provides a deterministic implementation of a symmetric self-generating line with the same total payoff.

Proof: We implement the corner utility pair $(\underline{u}, \overline{u})$ for a symmetric self-generating line with $x = 1 > y = 0$. As discussed above, the opposite corner utility pair, $(\overline{u}, \underline{u})$, then can be implemented in symmetric fashion (with $y = 1 > x = 0$), and all utility pairs on the line between the corners can be implemented using a public randomization device. Consider then the following specifications: $x = 1, y = 0, r = s = 0, u_{ao} = u_{a1} = \underline{u} + 1/\beta, u_{bo} = u_{b1} = u_o = \underline{u}, v_{ao} = v_{a1} = \overline{u} - 1/\beta$ and $v_{bo} = v_{b1} = v_o = \overline{u}$, where $\underline{u} = p/(1-\beta)$ and $\overline{u} = pqk/(1-\beta)$. Observe that $\underline{u}+\overline{u} = T = p[1+qk]/(1-\beta) = u_{a\theta}+v_{a\theta} = u_o+v_o$, for all $\theta \in \{0,1\}$. It is direct to confirm that the specifications satisfy the IR and IC constraints, (2.1)-(2.5), and also the promise keeping constraints, (2.6) and (2.7). Finally, given that $\beta \geq \beta^*$, simple calculations confirm that $\underline{u} + 1/\beta \leq \overline{u}$, and so every specified utility pair indeed falls on the line that connects $(\underline{u}, \overline{u})$ and $(\overline{u}, \underline{u})$. ■

Using Lemmas 3.2 and 3.6, we may conclude that a HSSGL cannot be a point (i.e., $\overline{u} > \underline{u}$ on a HSSGL).

We now report a finding that holds for any HSSGL and implementation of the associated $(\underline{u}, \overline{u})$.

Lemma 3.7. *Fix any HSSGL. For any implementation of the associated* $(\underline{u}, \overline{u})$, $x = 1$ *and thus* $y < 1$.

Proof: Assume to the contrary that $(\underline{u}, \overline{u})$ is implemented on a HSSGL with $x < 1$. Recall from Lemma 3.1 that $T = \frac{p[2+(x+y)(qk-1)]}{1-\beta}$. We obtain a contradiction by constructing an alternative self-generating line with $u + v = T' > T$. To construct this alternative line, it is sufficient to implement a new corner pair, $(\underline{u}', \overline{u}')$, on a line with $T' > T$. The rest of the alternative line can be implemented using convex combinations of $(\underline{u}', \overline{u}')$ and $(\overline{u}', \underline{u}')$.

Starting from the implementation of $(\underline{u}, \overline{u})$, we implement the new corner pair $(\underline{u}', \overline{u}')$ by making several changes. First, we increase x by a small amount, $\varepsilon > 0$. This change leads to a higher value for T, which increases in amount $\frac{p(qk-1)}{1-\beta}\varepsilon \equiv \gamma$. To place our new continuation pairs on this higher line, we must ensure that $\underline{u}'+\overline{u}'$ is higher than $\underline{u}+\overline{u}$ by γ; likewise, we must ensure that the values for $u_o + v_o$ and $u_{i\theta} + v_{i\theta}$ increase by γ, for all i and θ. To this end, we leave u_o, u_{bo} and u_{b1} at their original levels, increase u_{ao} and u_{a1} by ε/β, increase v_{ao} and v_{a1} by $\gamma - \varepsilon/\beta$, and increase v_{b1}, v_{bo} and v_o by γ. Note that $\gamma - \varepsilon/\beta \geq 0$ if and only if $\beta \geq \beta^*$. We leave s, r and y unaltered. Given that $(\underline{u}, \overline{u})$ was originally implemented, it is straightforward to confirm that the new specifications satisfy the IR and IC constraints, (2.1)-(2.5). Referring to (2.6), we calculate that $\underline{u}$ is unchanged (i.e.,

$\underline{u} = \underline{u}'$). We may use (2.7) to confirm that $\overline{u}$ has increased by γ (i.e., $\overline{u}' - \overline{u} = \gamma$). Thus, all new continuation values are at or above $\underline{u}'$ and at or below $\overline{u}'$, given $\beta \geq \beta^*$, and thus rest on the new - and strictly higher - self-generating line. This is a contradiction, and so $x = 1$ is necessary. Finally, given Corollary 3.4, it follows immediately that $y < 1$. ■

Thus, when implementing the worst value on any HSSGL for player a, player a must exhibit *full trust* (i.e., $x = 1$) even though player b does not (i.e., $y < 1$).

We next report two simple conditions that characterize any implementation of $(\underline{u}, \overline{u})$ along the widest HSSGL.

Lemma 3.8. *Consider the widest HSSGL. For any implementation of the associated* $(\underline{u}, \overline{u})$, $u_{bo} = \underline{u}$ *and (2.4) binds.*

Proof: Consider any implementation of $(\underline{u}, \overline{u})$ along the widest HSSGL and suppose to the contrary that $u_{bo} > \underline{u}$. Then $v_{bo} < \overline{u} = T - \underline{u}$. Starting with this implementation, let us now decrease u_{bo} by $\varepsilon > 0$ and increase v_{bo} by ε. Making no other changes, we observe that the new specifications satisfy the IR and IC constraints (2.1)-(2.5). Referring to (2.6) and (2.7), we see that the new corner utility pair, $(\underline{u}', \overline{u}')$, satisfies $\underline{u}' < \underline{u}$ and $\overline{u}' > \overline{u}$, contradicting the assumption that the original implementation corresponded to the widest HSSGL.

Next, consider any implementation of the widest HSSGL and suppose to the contrary that (2.4) is slack. Then $\beta(u_{b1} - u_{bo}) > s \geq 0$, and it follows that $u_{b1} > \underline{u}$ and $u_{bo} < \overline{u}$. Starting with this implementation, let us now decrease u_{b1} by $\varepsilon > 0$ and increase v_{b1} by ε. We note that (2.2) and (2.5) are unaffected by this change and thus continue to hold. Further, (2.3) is now sure to hold with slack, and (2.4) holds provided that ε is sufficiently small. Once again, we refer to (2.6) and (2.7) and observe that the new corner utility pair, $(\underline{u}', \overline{u}')$, satisfies $\underline{u}' < \underline{u}$ and $\overline{u}' > \overline{u}$, contradicting the assumption that the original implementation corresponded to the widest HSSGL. ■

As this result confirms, when implementing the worst value for player a along the widest HSSGL, player a's continuation value remains at this worst value in the event that player a fails to reciprocate in the current period.

We now consider specific implementations of the corner utility pair for the widest HSSGL. In particular, we posit an implementation of the widest HSSGL and then show that an implementation must exist that satisfies useful properties.

Lemma 3.9. *Consider the widest HSSGL. There exists an implementation of the associated* $(\underline{u}, \overline{u})$ *in which (i). (2.5) binds, (ii).* $r = s = 0, u_{a1} = u_{ao}$ *and* $u_{b1} = u_{bo}$*, (iii). (2.3) and (2.2) bind, (iv).* $u_{ao} = u_{bo} + (x+y)/\beta$*, and (v).* $x = 1$*,* $u_{bo} = \underline{u}$ *and (2.4) binds.*

Proof: To prove part (i), we fix any symmetric self-generating line and implementation of the associated $(\underline{u}, \overline{u})$. Suppose that (2.5) is slack. Then $\beta(v_{a1} - v_{ao}) = \beta(u_{ao} - u_{a1}) > r \geq 0$, and it follows that $u_{a1} < \overline{u}$ and $u_{ao} > \underline{u}$. Starting with this implementation, let us now decrease u_{ao} by $\varepsilon > 0$ and increase u_{a1} by $(1-q)\varepsilon/q$. Correspondingly, we increase v_{ao} by $\varepsilon > 0$ and decrease v_{a1} by $(1-q)\varepsilon/q$. For ε sufficiently small, (2.5) continues to hold; furthermore, all other constraints are unaffected by this change. Thus, the new specification also implements $(\underline{u}, \overline{u})$ along the same self-generating line. We can proceed in this way until (2.5) binds.

For part (ii), we consider the widest HSSGL. By Lemma 3.8, we know that (2.4) binds in the implementation of $(\underline{u}, \overline{u})$. Further, as just established, there exists an implementation of $(\underline{u}, \overline{u})$ under which (2.5) binds. Thus, $(\underline{u}, \overline{u})$ can be implemented with a specification under which (2.4) and (2.5) bind. For this implementation, we thus have that $r + \beta u_{a1} = \beta u_{ao}$ and $s + \beta v_{b1} = \beta v_{bo}$. Given u_{ao} and v_{bo}, any values for r, s, u_{a1} and v_{b1} that satisfy these latter two equations and feasibility constraints can also be used to implement $(\underline{u}, \overline{u})$. Thus, there exists an implementation in which $u_{a1} = u_{ao}$, $v_{b1} = v_{bo}$, $r = 0$ and $s = 0$.

For part (iii), we consider the widest HSSGL. We know that there exists an implementation of $(\underline{u}, \overline{u})$ in which (2.4) and (2.5) bind, and $u_{a1} = u_{ao}$, $v_{b1} = v_{bo}$, $r = 0$ and $s = 0$. Further, by Lemma 3.8, we also know that $u_{bo} = \underline{u}$. Since $x = 1$ by Lemma 3.7, we may use Proposition 3.5 and further conclude that $u_{a1} = u_{ao} \geq u_{bo} + (x+y)/\beta > \underline{u}$. Finally, we know from Corollary 3.4 that $y < 1$.

Let us now suppose that (2.3) is slack in this implementation. Using the properties just reported, we then find that $\beta[u_o - u_{bo}] > y \geq 0$, and so it follows that $u_o > \underline{u}$. We now derive a contradiction, by implementing an alternative utility pair, $(\underline{u}', \overline{u}')$, such that $T' = \underline{u}' + \overline{u}' > \underline{u}' + \overline{u} = T$. Starting with the original implementation, we first increase y by $\varepsilon > 0$, where ε is small. This change generates an increase in T in amount $\gamma = \frac{p(qk-1)}{1-\beta}\varepsilon$. It also increases the right-hand side of (2.6) by $pqk\varepsilon$. Second, we decrease u_{ao}, u_{a1} and u_o in amount δ, where δ satisfies $p\beta\delta + (1-2p)\beta\delta = pqk\varepsilon$ and is thus given by $\delta = \frac{pqk\varepsilon}{\beta(1-p)}$. Third, we increase v_{ao}, v_{a1} and v_o in amount $\delta + \gamma$. Finally, we increase v_{b1} and v_{bo} in amount δ, while leaving u_{b1} and u_{bo} unaltered. It is straightforward to confirm that our new specifications satisfy the IR and IC constraints (2.1)-(2.5), where (2.3) continues to hold if ε is sufficiently small. Referring to (2.6), we see that $\underline{u}' = \underline{u}$. Since u_{a1}, u_{ao} and u_o all exceed $\underline{u}$, all continuation values under our new specification continue to exceed $\underline{u}' = \underline{u}$, provided that ε is small. Referring to (2.7), we see that $\overline{u}' = \overline{u} + \gamma$. Since v_{ao}, v_{a1} and v_o are all less than $\overline{u}$, all continuation values under our new specifications rest below $\overline{u}'$, if ε is sufficiently small. The contradiction is now established.

Last, we suppose that (2.2) is slack in this implementation. Recalling the properties reported above, we know that $u_{a1} = u_{ao} > \underline{u}$ and hence $v_{a1} = v_{ao} < \overline{u}$. We now derive a contradiction by constructing a wider HSSGL. To this end, we start with the original implementation, and then decrease u_{a1} by ε and increase v_{a1} by ε. Making no other changes, we observe that the new specifications satisfy the IR and IC constraints (2.1)-(2.5). Referring to (2.6) and (2.7), we see that the new specification implements $(\underline{u}', \overline{u}')$, with $\underline{u}' < \underline{u}$ and $\overline{u}' > \overline{u}$. Thus, we can implement a wider line without changing T, which is a contradiction.

For part (iv), we observe from above that there exists an implementation of $(\underline{u}, \overline{u})$ on the widest HSSGL, in which all four incentive constraints (i.e., (2.2)-(2.5)) bind, and $u_{a1} = u_{ao}$, $v_{b1} = v_{bo}$, $r = 0 = s$, $u_{bo} = \underline{u}$ and $x = 1 > y$. Given that all four incentive constraints bind, we may follow the steps in the proof of Proposition 3.5 and confirm that the necessary condition (3.5) then must hold with equality: $u_{ao} - u_{bo} = (x + y)/\beta$. Thus, $u_{ao} = \underline{u} + (x + y)/\beta$.

Finally, part (v) simply lists properties (identified and used above) which we establish in Lemmas 3.7 and 3.8 as being true in any implementation of the widest HSSGL. ■

According to part (ii) of this result, if we can implement a corner utility pair and thereby construct the widest HSSGL, then we can do so with an implementation for which (2.5) binds and in which neither player exhibits immediate reciprocity. Referring to Proposition 3.5 and Lemma 3.9, we are now able to summarize some key findings on dynamic and immediate reciprocity.

Corollary 3.10. *For any symmetric self-generating line, any utility pair on the line can be implemented only if the implementation embodies dynamic reciprocity. In particular, for any HSSGL, the associated* $(\underline{u}, \overline{u})$ *can be implemented only if the implementation embodies dynamic reciprocity. In the widest HSSGL, there exists an implementation of the associated* $(\underline{u}, \overline{u})$ *such that neither player exhibits immediate reciprocity.*

In short, dynamic reciprocity is necessary for constructing a HSSGL, but immediate reciprocity is not.

We are now in position to derive an upper bound for y.

Proposition 3.11. *Fix any HSSGL. For any implementation of the associated* $(\underline{u}, \overline{u})$, $x = 1$ *and* $y \leq \frac{\beta - \beta^*}{\beta + \beta^*}$.

Proof: Consider any HSSGL and the implementation of the associated $(\underline{u}, \overline{u})$. By Lemma 3.7, $x = 1$. Suppose to the contrary that $y > \frac{\beta - \beta^*}{\beta + \beta^*}$. Let us now consider the widest HSSGL. (Recall that x and y are invariant across all HSSGL's.) By Lemma 3.9, we can implement the associated $(\underline{u}, \overline{u})$ with all four incentive constraints (i.e., (2.2)-(2.5)) binding, $u_{a1} = u_{ao} = \underline{u} + (1+y)/\beta$, $u_{b1} = u_{bo} = \underline{u}$, and $r = 0 = s$. Referring to the binding (2.3), we find that u_o may be expressed as $u_o = \underline{u} + y/\beta$. Using this expression, we may derive from (2.6) that

$$\underline{u} = \frac{p + y[p(qk-1)+1]}{1-\beta}. \tag{3.11}$$

Using as well that $u_o + v_o = u_{i\theta} + v_{i\theta} = \underline{u} + \overline{u}$, we may derive from (2.7) that

$$\overline{u} = \frac{pqk - y}{1-\beta}. \tag{3.12}$$

Recalling that $u_{a1} = u_{ao} = \underline{u} + (1+y)/\beta$ and using (3.11), we may derive that

$$u_{ao} = \frac{1 + y - \beta[1 - p - yp(qk-1)]}{\beta(1-\beta)}. \tag{3.13}$$

Finally, we may use (3.12) and (3.13) to find that $\overline{u} \geq u_{ao}$ if and only if $y \leq \frac{\beta - \beta^*}{\beta + \beta^*}$. Thus, under our assumption that $y > \frac{\beta - \beta^*}{\beta + \beta^*}$, it follows that $\overline{u} < u_{ao}$, and so a contradiction is obtained. ■

We note that Proposition 3.11 implies an upper bound for the total level of trust; in particular, this proposition establishes that, in the HSSGL,

$$x + y \leq \frac{2\beta}{\beta + \beta^*}. \tag{3.14}$$

Thus, Proposition 3.11 provides important guidance as we go forward and attempt to construct a HSSGL: if we can implement a symmetric self-generating line with $x + y = 2\beta/(\beta + \beta^*)$, then we can be assured that we have constructed a HSSGL.

3.3. Highest Self-Generating Line: Implementation

We now construct a HSSGL. We do this in two ways. First, we assume the existence of a public-randomization device and achieve the construction by implementing the corner utility pair $(\underline{u}, \overline{u})$ along a HSSGL. Using Proposition 3.11, we can be assured that we have a HSSGL if $x = 1$ and $y = \frac{\beta-\beta^*}{\beta+\beta^*}$. Under this approach, when the implementation calls for an intermediate utility pair following an event in which neither player reports income (i.e., when $\underline{u} < u_o, v_o < \overline{u}$), we may require that the players use the device to achieve a distribution over the corner utility pairs, $(\underline{u}, \overline{u})$ and $(\overline{u}, \underline{u})$, that generates the desired intermediate pair in expectation. Second, we construct a HSSGL when players do not have a public-randomization device. In this case, when the implementation of $(\underline{u}, \overline{u})$ requires that an intermediate utility pair is used following an event in which neither player reports income, then that pair itself must be implemented. This approach has the benefit of offering predictions about the evolution of trust between players as a succession of events are experienced in which neither player reports income.

We begin with the situation in which players have access to a public randomization device.

Proposition 3.12. *There exists a HSSGL, in which* $x + y = 2\beta/(\beta + \beta^*)$ *and thus* $T = p[2 + \frac{2\beta}{\beta+\beta^*}(qk - 1)]/(1 - \beta)$. *In particular, the corner utility pair* $(\underline{u}, \overline{u})$ *can be implemented using the following specifications:* $x = 1, y = (\beta - \beta^*)/(\beta + \beta^*), r = s = 0, u_{ao} = u_{a1} = \underline{u} + (1 + y)/\beta = \overline{u}, u_{bo} = u_{b1} = \underline{u}, u_o = \underline{u} + y/\beta$, $v_{ao} = v_{a1} = \underline{u}, v_{bo} = v_{b1} = \overline{u}$, *and* $v_o = \overline{u} - y/\beta$, *where*

$$\underline{u} = \frac{p + \frac{\beta-\beta^*}{\beta+\beta^*}\frac{1}{\beta^*}}{1 - \beta}, \text{ and} \tag{3.15}$$

$$\overline{u} = \underline{u} + \frac{2}{\beta + \beta^*}. \tag{3.16}$$

The corner utility pair $(\overline{u}, \underline{u})$ *can be implemented symmetrically, by interchanging* x *with* y *and* u *with* v *in the above specification. Finally, any utility pair on the line between the corners - and specifically the utility pair* (u_o, v_o) *- can be implemented using a public randomization device so that each corner utility pair is selected for implementation with appropriate probability.*

Proof: By Proposition 3.11, if a symmetric self-generating line exists for which $x + y = 2\beta/(\beta + \beta^*)$, then this line is a HSSGL. Thus, the proof is complete if we show that the specifications above implement the corner utility pair $(u,v) = (\underline{u}, \overline{u})$ for a symmetric self-generating line. First, we observe that $\underline{u} + \overline{u} = T = p[2 + \frac{2\beta}{\beta+\beta^*}(qk-1)]/(1-\beta) = u_{i\theta} + v_{i\theta} = u_o + v_o$, for all $i \in \{a, b\}$ and $\theta \in \{0, 1\}$. Second, we observe that $\overline{u} = u_{ao} = u_{a1} > u_o \geq \underline{u}$, where the final inequality is strict when $\beta > \beta^*$. Third, it is direct to confirm that the specifications satisfy the IR and IC constraints, (2.1)-(2.5), and also the promise keeping constraints, (2.6) and (2.7). In particular, the IC constraints all bind. Finally, as explained in the statement of the proposition, it is now direct to implement the opposite corner utility pair, $(\overline{u}, \underline{u})$, and we may then implement (u_o, v_o) by using a public randomization device. ■

To interpret this implementation, let us suppose that the players begin the game with a coin toss, where with probability 1/2 they implement $(\underline{u}, \overline{u})$ and with probability 1/2 they implement $(\overline{u}, \underline{u})$. Each player then expects an ex ante payoff of $(1/2)(\underline{u}+\overline{u})$. For simplicity, suppose that player b wins the toss, and the players thus start by implementing $(\underline{u}, \overline{u})$. Player b is now the *favored player*, since $v = \overline{u} > \underline{u} = u$. In this case, player a is expected to exhibit more trust: if player a receives the income, then player a exhibits full trust ($x = 1$); however, if player b receives the income, then player b does not exhibit full trust ($y < 1$). Three events may happen. First, if player a receives the income and exhibits full trust, then player a's "favor" is paid, and in the next period the players implement $(\overline{u}, \underline{u})$. Player a then becomes the favored player, since $u = \overline{u} > \underline{u} = v$, and so it then becomes player b's turn to pay a favor. Second, if player b receives the income and exhibits (partial) trust by sending y to player a, then player a's favor has not yet been paid, and so in the next period the players again implement $(\underline{u}, \overline{u})$, with player b thus again the favored player. Third, if neither player reports income, then in the next period the players implement the utility pair (u_o, v_o), by using a public randomization device and randomizing over $(\underline{u}, \overline{u})$ and $(\overline{u}, \underline{u})$. Notice that $v = v_o > u_o = u$, and thus player b remains the favored player. However, if $\beta > \beta^*$ so that $y > 0$, then player a's expected utility following the event in which no income is reported is strictly greater than player a's expected utility at the beginning of the period (or following an event in which player b receives the income). In expectation, the "size" of the favor that player a owes is thus reduced when the neutral event (neither player reports income) is experienced. The key intuition derives from the IC_y^b constraint. As (2.3) reveals, when the players are attempting to implement player b's preferred utility pair $(\underline{u}, \overline{u})$, they must be sure to give player b the incentive to report income (and thus send y to player a) when player b receives income. This is accomplished by penalizing player b somewhat when no income is reported.

The behavior described in Proposition 3.12 is relatively simple and yet it generates the highest possible expected utility for players when self-generating lines are used. We emphasize that players do not need to use immediate reciprocity: the implementation of a HSSGL is achieved with $r = s = 0$. We note, though, that alternative implementations of a HSSGL exist in which immediate reciprocity is used. Consider the following specifications: $x = 1, y = (\beta - \beta^*)/(\beta + \beta^*), u_{b1} - s/\beta = u_{bo} = \underline{u}, u_{ao} = u_{a1} + r/\beta = \underline{u} + (1+y)/\beta = \overline{u}, u_o = \underline{u} + y/\beta$, $v_{ao} = v_{a1} - r/\beta = \underline{u}, v_{bo} = v_{b1} + s/\beta = \overline{u}$, and $v_o = \overline{u} - y/\beta$. These specifications satisfy the IR and IC constraints, (2.1)-(2.5), and also the promise keeping constraints, (2.6) and (2.7). Further, it is direct to confirm that $\overline{u} \geq u_{b1} = \underline{u} + s/\beta$ if $s \leq \beta[\overline{u} - \underline{u}] = 1 + y$; likewise, we see that $\underline{u} \leq u_{a1}$ if $r \leq 1 + y$. Recalling that s and r are feasible if and only if $s \in [0, ky]$ and $r \in [0, kx]$, we may conclude that these specifications also implement a HSSGL provided that $s \in [0, \min(ky, 1+y)]$ and $r \in [0, \min(k, 1+y)]$, where $y = (\beta - \beta^*)/(\beta + \beta^*)$. We note that this family of implementations includes the implementation featured in Proposition 3.12 as a special case. Based on this discussion, we see that the practice of immediate reciprocity implies that a player that extends trust enjoys a less valuable future when some of that trust is reciprocated in the immediate period; for example, if the players seek to implement $(\underline{u}, \overline{u})$ and player a receives income, we see that $u_{a1} < u_{ao}$ when $r > 0$. By contrast, as Proposition 3.5 suggests, our analysis indicates that the extent of dynamic reciprocity, which we define as $u_{ao} - u_{bo}$, remains at the value $\overline{u} - \underline{u}$ whether or not players exhibit immediate reciprocity.

We now suppose that players do not use a public-randomization device.

Proposition 3.13. *There exists a HSSGL that can be implemented without a public-randomization device and in which $x + y = 2\beta/(\beta + \beta^*)$ and thus $T = p[2 + \frac{2\beta}{\beta+\beta^*}(qk-1)]/(1-\beta)$. In particular, let $(\underline{u}, \overline{u})$ be defined by (3.15) and (3.16) and consider any utility pair (u, v) along the line connecting $(\underline{u}, \overline{u})$ and $(\overline{u}, \underline{u})$. This pair can be implemented using the following specifications: $r = s = 0, u_{ao} = u_{a1} = \underline{u} + (x+y)/\beta = \overline{u}, u_{bo} = u_{b1} = \underline{u}, v_{ao} = v_{a1} = \underline{u}, v_{bo} = v_{b1} = \overline{u}$, and*

$$x = \beta\beta^*[\frac{v-p}{\beta} - \underline{u}], \tag{3.17}$$

$$y = \beta\beta^*[\frac{u-p}{\beta} - \underline{u}], \tag{3.18}$$

$$u_o = \beta^*[\frac{u-p}{\beta} + \frac{\underline{u}(1-\beta^*)}{\beta^*}], \tag{3.19}$$

$$v_o = \beta^*[\frac{v-p}{\beta} + \frac{\underline{u}(1-\beta^*)}{\beta^*}]. \tag{3.20}$$

Proof: Pick any utility pair (u, v) such that $u \in [\underline{u}, \overline{u}], v \in [\underline{u}, \overline{u}]$ and $u+v = \underline{u}+\overline{u}$. From Proposition 3.12, we know that $\underline{u} + \overline{u} = T = p[2 + \frac{2\beta}{\beta+\beta^*}(qk-1)]/(1-\beta)$. Simplifying, we have that $\underline{u}+\overline{u} = \frac{2p}{1-\beta} + \frac{2\beta(1-\beta^*)}{(\beta+\beta^*)(1-\beta)\beta^*}$. We also know that $\overline{u} - \underline{u} = 2/(\beta+\beta^*)$, where $\underline{u}$ is given by (3.15). Using these facts, we may use (3.17) and (3.18) to confirm that $x + y = 2\beta/(\beta+\beta^*)$. Thus, by setting $u_{ao} = u_{a1} = \overline{u}$, we also set $u_{ao} = u_{a1} = \underline{u} + (x+y)/\beta$. We now proceed as follows. First, using (3.19) and (3.20), we may confirm that $\underline{u}+\overline{u} = u_{i\theta} + v_{i\theta} = u_o + v_o$, for all $i \in \{a, b\}$ and $\theta \in \{0, 1\}$. Second, we may use (3.17)-(3.20) to confirm that the values for x, y, u_o and v_o are feasible. In particular, using (3.17), we find that $x \geq 0$ since $v \geq \underline{u}$ and $\beta \geq \beta^*$, where $x > 0$ if $v > \underline{u}$ or $\beta > \beta^*$; and we find that $x \leq 1$ since $v \leq \overline{u}$, where $x < 1$ if $v < \overline{u}$. Similarly, using (3.18), we find that $y \geq 0$ since $u \geq \underline{u}$ and $\beta \geq \beta^*$, where $y > 0$ if $u > \underline{u}$ or $\beta > \beta^*$; and we find that $y \leq 1$ since $u \leq \overline{u}$, where $y < 1$ if $u < \overline{u}$. Next, we may use (3.19) to confirm that $u_o \leq \overline{u}$ since $u \leq \overline{u}$ and $\beta \geq \beta^*$, where $u_o < \overline{u}$ if $u < \overline{u}$ or $\beta > \beta^*$; and we find that $u_o \geq \underline{u}$ since $u \geq \underline{u}$ and $\beta \geq \beta^*$, where $u_o > \underline{u}$ if $u > \underline{u}$ or $\beta > \beta^*$. Finally, given that $\underline{u} + \overline{u} = u_o + v_o = u + v$, it now follows that $v_o \geq \underline{u}$ since $v \geq \underline{u}$ and $\beta \geq \beta^*$, where $v_o > \underline{u}$ if $v > \underline{u}$ or $\beta > \beta^*$; and it follows as well that $v_o \leq \overline{u}$ since $v \leq \overline{u}$ and $\beta \geq \beta^*$, where $v_o < \overline{u}$ if $v < \overline{u}$ or $\beta > \beta^*$. Third, it is direct to confirm that the specifications satisfy the IR and IC constraints, (2.1)-(2.5), and also the promise keeping constraints, (2.6) and (2.7). In particular, the IC constraints all bind. Thus, any (u, v) along the line connecting $(\underline{u}, \overline{u})$ and $(\overline{u}, \underline{u})$ can be implemented using only continuation values drawn from that line. ■

In the implementation featured in Proposition 3.13, any utility pair (u, v) on the line that connects $(\underline{u}, \overline{u})$ and $(\overline{u}, \underline{u})$ can be implemented using only continuation values drawn from that line. For example, at the start of the game, the players might seek to implement a symmetric utility pair corresponding to the midpoint of this line. Let $(\widetilde{u}, \widetilde{u})$ denote the midpoint:

$$\widetilde{u} \equiv \frac{\underline{u}+\overline{u}}{2} = \frac{p + \frac{\beta(1-\beta^*)}{\beta+\beta^*}\frac{1}{\beta^*}}{1-\beta}. \tag{3.21}$$

Notice from (3.17) and (3.18) that $x = y$ when $u = v$; thus, since $x + y = 2\beta/(\beta+\beta^*)$, we have that $x = y = \beta/(\beta+\beta^*)$ in the first period. Suppose, for example, that player b receives income in the first period. The implementation then calls for player b to exhibit trust and send $y = \beta/(\beta + \beta^*)$ to player a. Play then moves to the second period, at which point the players seek to implement the (asymmetric) utility pair $(\underline{u}, \overline{u})$. This asymmetric pair rewards player b for reporting income and showing trust toward player a in the first period. To implement the corner pair $(\underline{u}, \overline{u})$, the players use the corresponding values for x and y that are given by (3.17) and (3.18). When $(u, v) = (\underline{u}, \overline{u})$, it is direct to confirm that these values are given by $x = 1$ and $y = (\beta - \beta^*)/(\beta + \beta^*)$, indicating that player a now exhibits more trust than player b. Notice that this is the same implementation for the corner utility pair that is used above in Proposition 3.12.

Continuing with this example, suppose next that no income is reported in the second period. Going into the third period, the players now seek to implement the utility pair (u_o, v_o), as given by (3.19) and (3.20) when $(u, v) = (\underline{u}, \overline{u})$. As Proposition 3.13 shows, this pair may be implemented deterministically (i.e., without a public randomization device). To determine the implementation for the pair $(u, v) = (u_o, v_o)$, we again refer to (3.17)-(3.20). At this point, it is important to use the notation with care. Given $(u, v) = (u_o, v_o)$, we may think of the two left-hand-side variables determined by (3.19) and (3.20) as a pair $(\widetilde{u}_o, \widetilde{v}_o)$ that represents the utilities that the players seek to implement at the start of the fourth period, in the event that no income is reported in the third period. In this general manner, for any given path of income realizations for the infinite game, we may refer to (3.17)-(3.20) and determine the path of trust (i.e., the amounts of income that are given from one player to another) for the infinite game.

As the discussion above suggests, one interesting possibility is that the players report no income over successive periods. Following the example above, suppose that player b sends income to player a in the first period, so that player b is the favored player in the second period, and suppose that neither player reports income in the second, third, etc., periods. Does player b remain the favored player, until a period finally arrives in which player a has income? Is the size of the favor that player a owes reduced in each successive period that no income is reported?

These questions are readily answered using (3.17)-(3.20). To this end, we may use (3.17) and (3.18) to find that

$$x - y = \beta^*[v - u]. \tag{3.22}$$

Equation (3.22) captures a basic relationship between the utility pair that the players seek to implement and the extent to which each player exhibits trust. In particular, if the players seek to implement a utility pair in which player b is favored (i.e., in which $v > u$), then player a must exhibit more trust (i.e., $x > y$).

Next, given the expressions for $\underline{u}$ and $\widetilde{u}$ presented in (3.15) and (3.21), respectively, we may use (3.19) to derive that

$$u_o - \widetilde{u} = \frac{\beta^*}{\beta}[u - \widetilde{u}]. \tag{3.23}$$

Of course, given that $2\widetilde{u} = u_o + v_o = u + v = \underline{u} + \overline{u}$, we may equivalently restate (3.23) as

$$\widetilde{u} - v_o = \frac{\beta^*}{\beta}[\widetilde{u} - v]. \tag{3.24}$$

Equations (3.23) and (3.24) indicate key relationships between the utility pair (u, v) that the players seek to implement in a given period and the utility pair (u_o, v_o) that they seek to implement in the next period in the event that no income is reported in the given period.

Consider first the possibility that $\beta = \beta^*$. Using (3.23) and (3.24), we see then that $(u, v) = (u_o, v_o)$. In this case, when the players seek to implement (u, v) and neither player reports income, then the players again seek to implement $(u, v) = (u_o, v_o)$ at the beginning of the next period. As (3.22) confirms, the trust levels that players are expected to exhibit are then unchanged. Put differently, the favor that is owed does not diminish as successive no-income states are encountered. Consider next the case in which $\beta > \beta^*$. If $u = v = \widetilde{u}$, then once again the favor owed does not diminish as successive no-income states are experienced. In this case, if no income is reported in the given period, then the players again seek to implement the same utility pair, $(u, v) = (u_o, v_o) = (\widetilde{u}, \widetilde{u})$, in the next period. In particular, x and y both remain at the symmetric level, $\beta/(\beta + \beta^*)$.

The final possibility is that $\beta > \beta^*$ and $(u, v) \neq (\widetilde{u}, \widetilde{u})$. In this case, patient players seek to implement an asymmetric utility pair. For simplicity, let us focus on the situation in which player b is favored: $v > \widetilde{u} > u$. We thus have from (3.22) that $x > y$. Now suppose that neither player reports income in the current period. Referring to (3.23) and (3.24), we see then that the players proceed to the next period and seek to implement (u_o, v_o), where $u_o < \widetilde{u} < v_o$. Given $\beta^*/\beta < 1$, we may further observe that $u < u_o$ and $v_o < v$. Thus, when $\beta > \beta^*$ and the players seek to implement (u, v) such that $v > \widetilde{u} > u$, if no income is reported, then in the next period the players seek to implement (u_o, v_o) such that $u < u_o < \widetilde{u} < v_o < v$. Applying (3.22), we see that in the next period player a continues to exhibit more trust than does player b; however, the extent of the trust differential is reduced (i.e., x remains larger than y, but $x - y$ is lower). Recalling the two questions posed above, we thus conclude that player b remains the favored player until a period occurs in which player a has income. But the size of the favor that player a owes is reduced in each successive period that no income is reported.

Thus, when player b is the favored player and a period is experienced in which neither player reports income, player a acknowledges that a favor is still owed but insists that the favor is now smaller in size. We may imagine player a exclaiming, "Yeah, but what have you done for me lately?" As suggested previously, the key intuition is associated with the IC_y^b constraint. As (2.3) indicates, when the players are attempting to implement a utility pair that favors player b, they must be sure to give player b the incentive to report income. To accomplish this, they use a utility pair (u_o, v_o) that penalizes player b somewhat when no income is reported. Recall that in the implementation with public randomization, the players always end up at a corner utility pair, with one or the other player showing full trust, and the intermediate utility pair (u_o, v_o) is achieved in expectation through an appropriate selection of probabilities. By contrast, in the implementation without public randomization, (u_o, v_o) is directly implemented; as a consequence, both players may exhibit partial trust, and we can analyze how these trust levels evolve as successive no-income periods are experienced.

We may summarize the discussion above as follows:

Corollary 3.14. *Consider the implementation of a HSSGL that is specified in Proposition 3.13. If $\beta = \beta^*$ or $(u, v) = (\widetilde{u}, \widetilde{u})$, then $(u_o, v_o) = (\widetilde{u}, \widetilde{u})$ and so the values for x and y are not altered following a period in which no income is reported. If $\beta > \beta^*$ and $v > \widetilde{u} > u$, then $u < u_o < \widetilde{u} < v_o < v$ and so $x - y$ remains positive but is reduced following a period in which no income is reported. Likewise, if $\beta > \beta^*$ and $u > \widetilde{u} > v$, then $v < v_o < \widetilde{u} < u_o < u$ and so $y - x$ remains positive but is reduced following a period in which no income is reported.*

3.4. Highest Symmetric Self-Generating Line: Unique Implementation

In Propositions 3.12 and 3.13, we present implementations of a HSSGL. The implementations are different, in that the former uses a public-randomization device. Further, as explained in the discussion following Proposition 3.12, alternative implementations that feature immediate reciprocity also can be provided for a HSSGL. Despite these findings, we show in this subsection that, for *any* utility pair on the widest HSSGL, *every* implementation is characterized by the same values for x, y, u_o and v_o.

We begin by defining our notion of uniqueness. Fix any (u, v) on the widest HSSGL. Let $\{x, y, r, s, u_{i\theta}, v_{i\theta}, u_o, v_o\}$ and $\{x', y', r', s', u'_{i\theta}, v'_{i\theta}, u'_o, v'_o\}$ be two implementations of (u, v), where each implementation uses only continuation values that are drawn from a HSSGL. We then say that (u, v) is *implemented uniquely* (up to $\{r, s, u_{i\theta}, v_{i\theta}\}$) if, for any such two implementations, we have $x = x', y = y', u_o = u'_o$ and $v_o = v'_o$. Otherwise, we say that there exists *multiple implementations* for (u, v). Thus, we define uniqueness in terms of the trust relationship (i.e., the values of x and y) and the manner in which utility pairs evolve following neutral states (i.e., the values of u_o and v_o).

Our result is as follows:

Proposition 3.15. *Every (u, v) on the widest HSSGL is implemented uniquely.*

We prove this proposition in the appendix. With this proposition, we have a uniqueness result for our prediction that the size of favor that is owed diminishes in expectation when a neutral state is encountered.

This concludes our characterization of HSSGL's. In the next three sections, we compare the total payoff achieved along the HSSGL with alternative benchmarks.

4. EM Relationships

In this section, we introduce the benchmark of a simple EM relationship. After defining and characterizing a simple EM relationship, we interpret the implementation featured in Proposition 3.13 in terms of a sophisticated EM relationship. Assuming $\beta > \beta^*$, we then show that the total payoff induced along a HSSGL by a sophisticated EM relationship is strictly greater than the total payoff achieved in a simple EM relationship.

In a *simple EM relationship*, once a player provides a favor, the player is unwilling to provide any further favors - of any size - until the favor is paid in full (i.e., until the other player provides a favor of equal size). We now construct a symmetric self-generating line that specifies a simple EM relationship. The construction yields the highest total payoff possible in such a relationship, since players exchange favors that are as large as possible. In the specification, the players move deterministically between two corner utility pairs, $(\underline{u}, \overline{u})$ and $(\overline{u}, \underline{u})$. The utility pair $(\underline{u}, \overline{u})$ is implemented when player a owes the favor. The players implement this utility pair as follows: (i) if player a receives income, then player a transfers all income $(x = 1)$; (ii) if player b receives income, then no transfer $(y = 0)$ is required; and (iii) if neither player receives income, then no transfer is feasible. In case (i), player a's favor is paid, and it is then player b's turn to provide a favor. The players thus implement the other corner utility pair, $(\overline{u}, \underline{u})$, in the next period. In case (ii), player a's favor is not yet paid, and the players thus implement $(\underline{u}, \overline{u})$ again in the next period. Finally, in case (iii), player a 's favor is also regarded as unpaid, and so the players implement $(\underline{u}, \overline{u})$ again in the next period. The utility pair $(\overline{u}, \underline{u})$ is implemented in similar fashion, except here it is player b that owes the favor.

We now provide a formal characterization of a simple EM relationship.

Proposition 4.1. *There exists a symmetric self-generating line that specifies a simple EM relationship, in which $x + y = 1$ and thus $T = p[1 + qk]/(1 - \beta)$. The line can be implemented without a public-randomization device. In particular, let*

$$\underline{u} = \frac{p^2\beta(1+qk)}{(1-\beta)(1-\beta+2\beta p)} \tag{4.1}$$

$$\overline{u} = \frac{p(1+qk)(1-\beta+\beta p)}{(1-\beta)(1-\beta+2\beta p)}. \tag{4.2}$$

The pair $(\underline{u}, \overline{u})$ can be implemented using the following specification: $r = s = 0$, $u_{ao} = u_{a1} = \overline{u}, u_{bo} = u_{b1} = u_{_o} = \underline{u}, v_{ao} = v_{a1} = \underline{u}, v_{bo} = v_{b1} = v_o = \overline{u}$, and $x = 1 > 0 = y$. The corner utility pair $(\overline{u}, \underline{u})$ can be implemented symmetrically, by interchanging x with y and u with v in the above specification.

Proof: We show that the specifications above implement the corner utility pair $(u, v) = (\underline{u}, \overline{u})$ for a symmetric self-generating line. First, we observe that $\underline{u} + \overline{u} = T = p[1 + qk]/(1 - \beta) = u_{i\theta} + v_{i\theta} = u_{_o} + v_{_o}$, for all $i \in \{a, b\}$ and $\theta \in \{0, 1\}$. Second, we observe that $\overline{u} - \underline{u} = p(1 + qk)/(1 - \beta + 2\beta p) \geq 1/\beta$, where the inequality is strict if $\beta > \beta^*$. Third, it is now direct to confirm that the specifications satisfy the IR and IC constraints, (2.1)-(2.5), and also the promise keeping constraints, (2.6) and (2.7). Finally, as explained in the statement of the proposition, we may now implement the opposite corner utility pair, $(\overline{u}, \underline{u})$. ■

We note that the symmetric self-generating line implemented in Proposition 4.1 achieves the same total payoff, T, as does the symmetric self-generating line implemented in Lemma 3.6. The implementation in Proposition 4.1, however, specifies a simple EM relationship, and thus does not use a public-randomization device.[8]

[8] We also note that the implementation given in Proposition 4.1 corresponds to a narrower symmetric self-generating line than does the implementation given in Lemma 3.6. In particular, $\underline{u}$ is greater and $\overline{u}$ is lower in the implementation given in Proposition 4.1, with the difference being strict when $\beta > \beta^*$.

The simple EM relationship generates payoffs that exceed the autarkic payoffs that arise under repeated play of the Nash equilibrium of the stage game. Thus, this intuitive relationship can be interpreted as efficiency enhancing. An important limitation of the simple EM relationship, however, is that the benefit of investment if foregone when the same player receives income in successive periods. By contrast, in the implementation that we feature in Proposition 3.13, players adopt a *sophisticated EM relationship*, in which a player transfers some income even when that player provided the most recent favor. As explained above, an incentive for such behavior is provided, if favors deteriorate in size following the experience of neutral states. When $\beta > \beta^*$, a sophisticated EM relationship offers a strictly higher total payoff than does any simple EM relationship, since the self-generating line that describes the former is strictly above that which describes the latter. In other words, a sophisticated EM relationship is then described by a strictly higher level of trust than is any simple EM relationship.

We may summarize the comparison as follows:

Corollary 4.2. *If $\beta > \beta^*$, the sophisticated EM relationship characterized in Proposition 3.13 generates a strictly higher level of trust and thus offers a strictly higher total payoff than does any simple EM relationship.*

5. Strongly Symmetric Equilibria

In the analysis above, we allow that players can promise future favors through asymmetric continuation values, but we do not allow that players may threaten a symmetric punishment whereby $u = v$ is lowered following certain public outcomes. We now consider strongly symmetric equilibria (SSE) and thus adopt the opposite emphasis: players' utilities are no longer allowed to move asymmetrically along a negatively sloped line, but players' utilities are now allowed to move symmetrically along the 45-degree line. We characterize the set of SSE and, in particular, identify specific circumstances under which SSE generate a symmetric payoff for the game that exceeds that obtained in a HSSGL.

5.1. A Simple Example

To illustrate the main ideas, we begin with a simple example. Consider the following simple strategy profile for each player: if there has been a period with an investment level of less than one unit, then each player consumes his income and makes no transfer; otherwise, a player that receives a unit of income invests it all in the other player. Further, a player that receives a transfer does not offer any reciprocation in the current period. We demonstrate that this strategy profile may constitute an SSE in which players' total payoff exceeds that in a HSSGL.

Let v be the utility level that each player derives if both players use this strategy. We may calculate v as follows:

$$v = p(0 + \beta v) + p(qk + \beta v) + (1 - 2p)(0 + \beta u_{aut}),$$

where $u_{aut} = \frac{p}{1-\beta}$. The first (second) term is a player's utility if he (his opponent) receives the income, while the third term is the player's utility if neither player receives income. Solving for v, we obtain

$$v = \frac{1}{1 - 2p\beta}(pqk + (1 - 2p)\beta u_{aut}).$$

This simple strategy profile is an equilibrium strategy if a player with income would prefer not to deviate. The optimal deviation would be not to invest at all. Therefore, v would be an equilibrium payoff if

$$1 + \beta u_{aut} \leq \beta v$$

or equivalently

$$\beta \geq \frac{1}{p(qk + 1)}. \tag{5.1}$$

Thus, if $\frac{1}{p(qk+1)} < 1$, or equivalently $p > \frac{1}{qk+1}$, then there exists $\beta \geq \frac{1}{p(qk+1)}$ for which the simple strategy profile is an equilibrium. At the start of the game, the players' total payoff is then $2v$.

We now compare this total payoff with that received in a HSSGL. By (3.1) and Proposition 3.12, the players' total payoff on a HSSGL is given by

$$T^* = \frac{p[2 + (x + y)(qk - 1)]}{1 - \beta},$$

where $x + y = \frac{2\beta}{\beta+\beta^*}$. Comparing $2v$ with T^*, we obtain $2v > T^*$ if and only if

$$\frac{1 - \beta}{\beta^2} > (1 - 2p)(1 + p(qk - 1)). \tag{5.2}$$

Therefore, we may refer to (5.1) and (5.2) and draw the following conclusion: if qk is sufficiently large and p is close to $\frac{1}{2}$, then an SSE exists that yields a total payoff that is higher than that received in a HSSGL (i.e., $2v > T^*$).

This example suggests that SSE may be attractive when it is rare that an investor fails to exist (i.e., when p is close to $\frac{1}{2}$) and investments are very productive (i.e., qk is large). Intuitively, the players may then threaten a severe symmetric punishment in the event that neither player reports that he is the investor. This severe punishment ensures that a player has incentive to be honest when he is the investor; furthermore, the punishment is "almost off the equilibrium path" when it is rare that an investor fails to exist, and thus incentives are provided at little cost to expected equilibrium payoffs. Such a severe punishment is not possible when payoffs must sum to a fixed total, as along a HSSGL, and it is thus intuitive that SSE can offer a higher total payoff than does a HSSGL in such circumstances.

5.2. Characterization of Optimal Strongly Symmetric Equilibria

We proceed now to characterize optimal SSE. To begin, we follow Abreu, Pearce and Stacchetti (1990) and define an operator B^{ss} which yields the set of strongly symmetric PPE values, ψ_s^*, as the largest self-generating set. This operator is defined as follows:

For any $\psi_s = [u_{aut}, u]$ consider the following mapping: $B^{ss}(\psi_s) = \{v : \exists\, x \in [0,1], r \in [0, kx], v_o, v_{10}, v_{11} \in \psi_s$ such that:

$$IC_x : 1 - x + q(r + \beta v_{11}) + (1-q)\beta v_{10} \geq 1 + \beta v_o \tag{5.3}$$

$$IC_\theta : kx - r + \beta v_{11} \geq kx + \beta v_{10} \tag{5.4}$$

$$\begin{aligned} PK \quad : \quad & v = p[1 - x + q(r + \beta v_{11}) + (1-q)\beta v_{10}] \\ & +p[q(kx - r + \beta v_{11}) + (1-q)\beta v_{10}] + (1-2p)\beta v_o\}. \end{aligned} \tag{5.5}$$

Let $\psi_s^* = [u_{aut}, u_{\max}]$ be the maximal fixed point of B^{ss}. That is, if $[u_l, u_h]$ is a fixed point of B^{ss}, then $[u_l, u_h] \subset [u_{aut}, u_{\max}]$.

Observe that this operator requires symmetry across players, with u denoting the payoff enjoyed by each player, x denoting the investment that a player makes in the current period if that player receives income, r denoting the reciprocity that the trustee then offers in the current period if the investment is successful, and v_o,v_{11} and v_{10} denoting the continuation values that each player receives in the future if the current period has no investor, a successful investment and an unsuccessful investment, respectively. For a given $\psi_s = [u_{aut}, u]$, we thus say that $\{x, r, v_{10}, v_{11}, v_o\}$ *implements* v if all of the constraints above are satisfied.

We refer to a pair (q, p) as an *information structure*. Consider the set $I = \{(q,p) : q \in [\frac{1}{k}, 1], p \in [0, \frac{1}{2}]\}$, which is the set of all feasible information structures. The characterization of optimal SSE reveals that behavior differs depending upon which of three different information-structure regions is in place. The respective regions are illustrated in Figure 1. We now describe the behavior that emerges in each region. The proofs are contained in the Appendix.

5.2.1. Region I_1: Low q and not so high p

Let $q^* = \frac{k+\sqrt{k^2+8k}}{4k} \in (\frac{1}{2}, 1)$. Consider $I_1 = \{(q,p) \in I : q \leq q^* \text{ and } p \leq \frac{1}{qk+1}\}$. In this region, we find that $u_{\max} = u_{aut}$. Thus, under this information structure, the players are unable to cooperate using SSE. Intuitively, given that p is small, no-investor states are common. Hence, if players attempt to provide incentives for trust by using the threat of a symmetric punishment, then this punishment often would be experienced on the equilibrium path. Further, with q being small as well, the value of future cooperation is not huge. The players are thus unable to enforce a strongly symmetric equilibrium in which trust is exhibited. Clearly, if β is sufficiently high that a HSSGL exists, then the players earn a higher total payoff in a HSSGL than in the optimal (autarkic) SSE.

5.2.2. Region I_2: Not so high q but high p

Consider now $I_2 = \{(q,p) \in I : q \leq \frac{1}{2(1-p)} \text{ and } p > \frac{1}{qk+1}\}$. For $\beta > \frac{1}{p(qk+1)}$, we find that $u_{\max} = u_{aut} + \frac{p(qk+1)-1}{1-\beta}$. The following implements the optimal SSE: $x = 1, v_{10} = v_{11} = u_{\max}, r = 0$ and $v_o = u_{\max} - \frac{1}{\beta} > u_{aut}$.

We observe that implementation of $u_{\max}$ is achieved without use of immediate reciprocity (i.e., $r = 0$), and that players incur a moderate punishment when the neutral (no-investor) state is experienced (i.e., $u_{\max} > v_o > u_{aut}$). We also find that $\lim_{p \to 1/2} u_{\max} = \lim_{p \to 1/2} u_{eff}$, where $u_{eff} = \frac{pqk}{1-\beta}$ is the payoff that a player enjoys in the first-best benchmark. This implies that, when p is sufficiently close to $\frac{1}{2}$, patient players achieve a higher total payoff in the optimal SSE than they do on a HSSGL. Intuitively, when p is close to $\frac{1}{2}$, the neutral state is rare; thus, the players can use the threat of a symmetric punishment in this state to provide incentives for trust while only rarely experiencing the punishment on the equilibrium path.

5.2.3. Region I_3: High q but not so high p

Finally, consider $I_3 = \{(q,p) \in I : q > q^* \text{ and } p \leq 1 - \frac{1}{2q}\}$. Define $\widehat{\beta} = \frac{1}{1+p(2kq^2-qk-1)}$. Then $\widehat{\beta} < 1$ if and only if $q > q^*$. For $\beta \geq \widehat{\beta}$, we find that $u_{\max} = u_{aut} + \frac{\lambda}{1-\beta}$, where $\lambda = \frac{p(2kq^2-qk-1)}{2q-1} \geq 0$ since $q \geq q^*$. The following implements the optimal SSE: $x = 1, v_o = v_{11} = u_{\max}, v_{10} = u_{\max} - \frac{1}{\beta(2q-1)} \geq u_{aut}$ and $r = \frac{1}{2q-1} > 0$.

We observe that implementation of $u_{\max}$ is achieved without punishment in the neutral (no-investor) states (i.e., $v_o = u_{\max}$). Instead, players punish one another when there is no immediate reciprocity (i.e., $v_{10} < u_{\max}$). Thus, in this implementation, immediate reciprocity plays an important role (i.e., $r > 0$). We also find that $\lim_{q \longrightarrow 1} u_{\max} = \lim_{q \longrightarrow 1} u_{eff}$. This implies that, when q is close to 1, patient players achieve a higher total payoff in the optimal SSE than they do on a HSSGL. Intuitively, when q is close to 1, investment is almost always successful; thus, the players can use the threat of a symmetric punishment when immediate reciprocity is not offered to provide incentives for trust while only rarely experiencing the punishment on the equilibrium path.

5.2.4. Comparisons

It is interesting to compare regions I_2 and I_3. Start with $(q,p) \in I_2$, where immediate reciprocity plays no role. As we increase q, we reach I_3, where immediate reciprocity begins playing a role. Also, as we move from I_2 to I_3, the punishment phase shifts from following a neutral (no-investor) state to following the state in which trust is exhibited but immediate reciprocity is not offered. As suggested above, the intuition is that players provide incentives most efficiently by emphasizing the information asymmetry for which the "bad" outcome (no investor, unsuccessful investment) is unlikely. Further, it is precisely in those circumstances where a bad outcome is very unlikely that the optimal SSE offers a total payoff that exceeds that in a HSSGL.

Finally, we have not specified whether a punishment-phase utility is itself implemented or if instead it is achieved in expectation via a public-randomization device that induces a lottery over $u_{\max}$ and u_{aut}. The latter interpretation is immediate and requires no further analysis. Under this interpretation, any punishment phase entails the risk of permanent autarky. Similarly, it is possible to implement a punishment-phase utility with a lottery in which the players risk temporary autarky, whereby in each period the players leave autarky (return to $u_{\max}$) with a constant hazard rate. To implement in expectation a given punishment-phase utility, the lottery must place a higher probability on going to autarky when the autarky relationship is temporary.

For future reference, we now collect our findings for payoffs:

Proposition 5.1. *Let $u_{\max}$ represent the utility achieved in the optimal SSE. (i). For $(q,p) \in I_1$, $u_{\max} = u_{aut}$. (ii). For $(q,p) \in I_2$, if $\beta \geq \frac{1}{p(qk+1)}$, then $u_{\max} = u_{aut} + \frac{p(qk+1)-1}{1-\beta}$. (iii). For $(q,p) \in I_3$, if $\beta \geq \widehat{\beta}$, then $u_{\max} = u_{aut} + \frac{\lambda}{1-\beta}$, where $\lambda = \frac{p(2kq^2-qk-1)}{2q-1} \geq 0$.*

Thus, throughout region I_1, the optimal SSE offers a strictly lower payoff than does a HSSGL. For β sufficiently high, however, the optimal SSE offers a strictly higher payoff than does a HSSGL in subsets of region I_2 and I_3 within which p is sufficiently close to $\frac{1}{2}$ and q is sufficiently close to 1, respectively.

As Proposition 5.1 confirms, our analysis of the optimal SSE in regions I_2 and I_3 imposes additional restrictions on β beyond our maintained assumption that $\beta \geq \beta^*$. The restrictions are important. For example, consider the subset of region I_2 in which $q \leq \frac{1}{2}$ and $\beta < \frac{1}{p(1+qk)}$. Letting $p_s \equiv \frac{1}{\beta(1+qk)}$, we may state the latter inequality as $p_s > p$. We observe that $p_s < \frac{1}{2}$ when $q = 1$ if and only if $\beta > \frac{2}{1+k}$. But simple calculations confirm that $\beta^* > \frac{2}{1+k}$. Given $\beta \geq \beta^*$, we thus conclude that $p_s < \frac{1}{2}$ when $q = 1$. As Figure 2 illustrates, Proposition 5.1 part (ii) refers to that portion of region I_2 that lies above the $p_s = p$ curve. In contrast, our present interest is in the subset of region I_2 that rests below the $p_s = p$ curve and in which $q \leq \frac{1}{2}$.[9]

We now provide our main finding for this subset.

Proposition 5.2. *Let $u_{\max}$ represent the utility achieved in the optimal SSE. For $(q,p) \in I_2$, if $q \leq \frac{1}{2}$ and $\beta < \frac{1}{p(1+qk)}$, then $u_{\max} = u_{aut}$.*

Thus, in this subset of region I_2, the optimal SSE corresponds to autarky and therefore offers a strictly lower payoff than does a HSSGL.

[9] We observe that $p_s \geq \frac{1}{2}$ when p_s is evaluated at $q = \frac{1}{2}$, if $k \in (1,2]$. Thus, in the special case where $k \in (1,2]$, the subset of interest is defined simply as that portion of region I_2 in which $q \leq \frac{1}{2}$.

6. Hybrid Equilibria

Our discussion above characterizes HSSGL's and optimal SSE. With these constructions established, we are now able to consider the possibility of hybrid equilibria. In such equilibria, players begin the game by exhibiting a high level of trust in period one. If some player receives and transfers income in the first period, then the players thereafter exchange favors by implementing a HSSGL, with that player being the favored player in the second period. Alternatively, if no player receives income in the first period, then the players may revert to a symmetric punishment in the second period. In broad terms, such equilibria are thus characterized by an initial "honeymoon" period, after which the players either continue with a sophisticated favor-exchange relationship or experience a breakdown. In this section, we characterize the optimal hybrid equilibria and compare the associated payoffs with those achieved in HSSGL's and optimal SSE.

6.1. Characterization of Optimal Hybrid Equilibria

Recall the definition of implementation in Section 2. For a given $\psi = [u_{aut}, u]$, we now say that a pair $\{x, u_o\}$ *implements* u *in a hybrid equilibrium* if $\{x, y, r, s, u_{i\theta}, v_{i\theta}, u_o, v_o\}$, for $i = a, b$ and $\theta = 0, 1$, implements the utility pair $\{u, u\}$ when $x = y, r = s = 0, u_{a1} = u_{ao} = v_{b1} = v_{bo} = \overline{u}, u_{b1} = u_{bo} = v_{a1} = v_{ao} = \underline{u}$ and $u_o = v_o \in [u_{aut}, u]$, where $\underline{u}$ and $\overline{u}$ are defined by (3.15) and (3.16). In an *optimal hybrid equilibrium*, x and u_o are chosen to deliver the maximal value for u. Thus, in a hybrid equilibrium, the players exhibit equal trust in the first period (*i.e.*, $x = y$). If some player receives income and transfers the amount x, then in period two the players implement a HSSGL. At this point, the player that made the period-one transfer is favored and thus enjoys a continuation value of $\overline{u}$ while the other player's continuation value is $\underline{u}$. If instead neither player received income in period one, then in period two the players implement a symmetric utility pair, (u_o, u_o).

We begin with the following lemma:

Lemma 6.1. *In any implementation of an optimal hybrid equilibrium, (2.2) and (2.3) bind and*

$$u = p + \beta u_o[1 - p(1 + qk)] + p\beta[qk\overline{u} + \underline{u}]. \tag{6.1}$$

Proof: Suppose (2.2) is slack. If $x = y < 1$, then we can raise x and y by a small amount while keeping $u_o = v_o$ fixed. This new implementation satisfies all constraints and generates a higher utility, contradicting the hypothesis that the original specification implemented an optimal hybrid equilibrium. Likewise, if $x = y = 1$ and $u_o = v_o < u$, then we can obtain a contradiction by increasing $u_o = v_o$ a small amount while keeping $x = y = 1$. Finally, if $u_o = v_o = u$ and $x = y = 1$, then $u < \overline{u} - 1/\beta < \overline{u} - 1/(\beta + \beta^*) = (\overline{u} + \underline{u})/2 \equiv \widetilde{u}$, where the first inequality follows from the supposition that (2.2) is slack. A contradiction is now obtained, since players may implement a hybrid equilibrium that generates the higher utility $\widetilde{u}$, by using the implementation of a HSSGL that is specified in Proposition 3.13 when $(u, v) = (\widetilde{u}, \widetilde{u}) = (u_o, v_o)$. (See also Corollary 3.14.) Thus, (2.2) is binding, and by symmetry so is (2.3). Next, given that (2.2) and (2.3) bind, we may substitute for x and y in (2.6) and thereby derive (6.1). ■

Our next finding indicates that the characterization of optimal hybrid equilibria is sensitive to the sign of $1 - p(1 + qk)$.

Lemma 6.2. *Suppose $\{x, u_o\}$ implements u in an optimal hybrid equilibrium. If $1 < p(1 + qk)$, then $x = 1$, $u_o = \overline{u} - 1/\beta$ and*

$$u = [p(qk + 1) - 1] + p + \beta(1 - p)\overline{u} + \beta p\underline{u}. \tag{6.2}$$

If $1 > p(1 + qk)$, then $x = \frac{\beta}{\beta+\beta^*}$ and $u_o = u = \widetilde{u}$. If $1 = p(1 + qk)$, then $u = \widetilde{u}$.

Proof: First, suppose $\{x, u_o\}$ implements u in an optimal hybrid equilibrium and that $1 < p(1 + qk)$. Using (6.1), we see that u is greater when u_o is lower. Lemma 6.1 indicates that (2.2) must bind; thus, it is necessary that $u_o = \overline{u} - x/\beta$. Suppose $x < 1$. We then have that $u_o = \overline{u} - x/\beta > \overline{u} - 1/\beta \geq \underline{u} \geq u_{aut}$, where the weak inequalities are strict if $\beta > \beta^*$. With $x < 1$ and $u_o > u_{aut}$, we may thus increase $x = y$ by ε and decrease $u_o = v_o$ by ε/β. All constraints remain satisfied. Using (2.6), we see that utility is increased by $-p\varepsilon + pqk\varepsilon + (1 - 2p)\beta(-\varepsilon/\beta) = \varepsilon[p(qk + 1) - 1] > 0$, a contradiction. Thus, if $\{x, u_o\}$ implements u in an optimal hybrid equilibrium, then $x = 1$ and $u_o = \overline{u} - 1/\beta$. Using (6.1), we may then confirm that u is given as in (6.2).

Second, suppose $\{x, u_o\}$ implements u in an optimal hybrid equilibrium and that $1 > p(1 + qk)$. As noted in the proof of Lemma 6.1, we may implement a hybrid equilibrium that generates the payoff $\widetilde{u}$. Thus, it is necessary that $u \geq \widetilde{u}$. Suppose $u > \widetilde{u}$. Since (2.2) and (2.3) bind in the implementation of $\widetilde{u}$, we may reason as in the proof of Lemma 6.1 and conclude that $\widetilde{u}$ satisfies

$$\widetilde{u} = p + \beta\widetilde{u}[1 - p(1 + qk)] + p\beta[qk\overline{u} + \underline{u}]. \tag{6.3}$$

Likewise, u and u_o must satisfy (6.1). Subtracting (6.3) from (6.1) and using $\beta[1-p(1+qk)] \in (0,1)$, we obtain $u - \widetilde{u} = \beta[1-p(1+qk)](u_o - \widetilde{u}) < u_o - \widetilde{u}$, and so it follows that $u_o > u$. This contradicts the requirement that $u_o = v_o \in [u_{aut}, u]$. It follows that the optimal hybrid equilibrium utility is $\widetilde{u}$, when $1 > p(1+qk)$. Correspondingly, we then have that $x = \beta(\overline{u} - \widetilde{u}) = \frac{\beta}{\beta+\beta^*}$.

Finally, suppose $\{x, u_o\}$ implements u in an optimal hybrid equilibrium and that $1 = p(1+qk)$. Using (6.1), we see that u is then independent of u_o, when (2.2) binds. Using (6.1), the corresponding payoff is $u = p + p\beta[qk\overline{u} + \underline{u}]$. By (6.3), when $1 = p(1+qk)$, $u = \widetilde{u}$. ■

To see the intuition, suppose that $1 < p(1+qk)$. If we increase the punishment that follows an event in which no income is reported (i.e., if $u_o = v_o$ is lowered), then the players can be motivated to transfer a greater income (i.e., $x = y$ can be raised). The benefit of an increase in the size of the transfer is measured by $qk - 1$ and happens with probability p. On the other hand, the players then suffer a greater punishment when, in fact, neither player has income. This cost is experienced with probability $1 - 2p$. Thus, the net gain is positive if $1 - 2p < p(qk - 1)$, or equivalently, if $1 < p(qk + 1)$.

We now confirm that an optimal hybrid equilibrium exists.

Proposition 6.3. *There exists an optimal hybrid equilibrium. If $1 < p(1+qk)$, then $x = 1$ and $u_o = \overline{u} - 1/\beta$ implement the optimal hybrid equilibrium, and the corresponding equilibrium utility is given in (6.2). If $1 > p(1+qk)$, then $x = \frac{\beta}{\beta+\beta^*}$ and $u_o = \widetilde{u}$ implement the optimal hybrid equilibrium, and the corresponding equilibrium utility is given by $\widetilde{u}$. If $1 = p(1+qk)$, in all implementations of optimal hybrid equilibria, the corresponding equilibrium utility is given by $\widetilde{u}$.*

Proof: Suppose $1 < p(1+qk)$. The proposed implementation satisfies all constraints, provided that $u_o = \overline{u} - 1/\beta \in [u_{aut}, u]$, where u is given in (6.2). To this end, we observe that $u_o = \overline{u} - 1/\beta \geq \underline{u} \geq u_{aut}$, where the weak inequalities are strict if $\beta > \beta^*$. Next, $u_o = \overline{u} - 1/\beta < \overline{u} - 1/(\beta+\beta^*) = \widetilde{u} < u$. When $1 \geq p(1+qk)$, we may implement $\widetilde{u}$ by using $x = \frac{\beta}{\beta+\beta^*}$ and $u_o = \widetilde{u}$, as explained in the proof of Lemma 6.1. ■

6.2. Comparisons

We next compare the payoffs in optimal hybrid equilibria with those in HSSGL's and optimal SSE. As above, we use $u_{\max}$ to represent the payoff that a player expects at the beginning of the game, when players use an optimal SSE. Similarly, if players begin the game by implementing the symmetric utility pair on a HSSGL, then $\widetilde{u} \equiv (\overline{u} + \underline{u})/2$ represents a player's payoff. Finally, if players implement an optimal hybrid equilibrium, we let u_H represent the corresponding payoff that a player expects at the beginning of the game.

We first compare optimal hybrid equilibria and HSSGL's. Using Proposition 6.3, we have the following corollary:

Corollary 6.4. *If $1 < p(1 + qk)$, then the optimal hybrid equilibrium offers a strictly higher total payoff than does any HSSGL, and thus $u_H > \widetilde{u}$. If $1 \geq p(1 + qk)$, then all optimal hybrid equilibria offer the same total payoff as does any HSSGL, and thus $u_H = \widetilde{u}$.*

This finding follows directly from Proposition 6.3. When $1 < p(1 + qk)$, we may use (6.2), (3.21), (3.16) and (3.15) to compute the explicit expression for the payoff difference:

$$\begin{aligned} u_H - \widetilde{u} &= \{[p(qk+1) - 1] + p + \beta(1-p)\overline{u} + \beta p \underline{u}\} - \{\underline{u} + \frac{1}{\beta + \beta^*}\} \\ &= \frac{[p(1+qk) - 1]\beta^*}{\beta + \beta^*} > 0. \end{aligned}$$

As discussed above, the key point is that, when $1 < p(1+qk)$, players can benefit by using the threat of a symmetric punishment to enforce an initial "honeymoon" period in which the level of trust is very high. Provided that some player receives and transfers income in the first period, the players then use a sophisticated EM relationship (i.e., move along a HSSGL) in all future periods.

If $1 < p(qk + 1)$, we may easily verify that $\overline{u} > u_H > \widetilde{u} > \underline{u}$. Thus, as Corollary 6.4 indicates, when a honeymoon period is included, the players earn a higher symmetric payoff at the start of the game ($u_H > \widetilde{u}$). One perspective on this result is that the first period is a special period, since players are not encumbered by obligations that are derived from past favors; hence, they may set $x = y = 1$ and exhibit full trust in the first period. We observe as well that the player that made a period-one transfer emerges as the favored player in period two and in fact then enjoys a higher continuation value than at the start of the game ($\overline{u} > u_H$). Correspondingly, the player that enters period two as the disfavored player experiences a reduced continuation value ($\underline{u} < u_H$).

We next compare optimal hybrid equilibria and optimal SSE. We focus on region I_2, where $1 < p(1+qk)$. We provide two results. First, recall from Proposition

5.2 that the optimal SSE generates the autarky payoff, u_{aut}, in the subset of region I_2 in which $q \leq \frac{1}{2}$ and $\beta < \frac{1}{p(1+qk)}$. In Figure 2, members of this subset satisfy $q \leq \frac{1}{2}$ and rest below the $p_s = p$ curve, where $p_s \equiv \frac{1}{\beta(1+qk)}$. Using Proposition 5.2 and Corollary 6.4, we may thus conclude that:

Corollary 6.5. *If $q \leq \frac{1}{2}$ and $1 < p(1+qk) < 1/\beta$, then the optimal hybrid equilibrium offers a strictly higher total payoff than does the optimal SSE and any HSSGL. In fact, under these conditions, $u_H > \widetilde{u} > u_{\max} = u_{aut}$.*

We have thus identified a subset of region I_2 in which the optimal hybrid equilibrium offers a strict improvement over HSSGL's and optimal SSE.

To develop our second result, we recall Proposition 5.1. As indicated there, when p is sufficiently close to $\frac{1}{2}$, players can achieve a higher total payoff in the optimal SSE than in any HSSGL: $u_{\max} > \widetilde{u}$. We now confirm that, under similar circumstances, the optimal SSE also improves upon the optimal hybrid equilibrium: $u_{\max} > u_H$. Interestingly, this ranking obtains even though the optimal hybrid equilibrium also employs symmetric punishments after neutral states.

Following Proposition 5.1, we focus on the subset of region I_2 for which $\frac{1}{\beta} < p(1+qk)$, or equivalently $p_s < p$. As established previously and depicted in Figure 2, $p_s < \frac{1}{2}$ when $q = 1$. The subset thus exists. Over this subset, we have from Proposition 5.1 that $u_{\max} = u_{aut} + \frac{p(qk+1)-1}{1-\beta}$. Next, since $1 < \frac{1}{\beta} < p(1+qk)$, we may use (6.2) and write

$$u_H - u_{\max} = [p(qk+1) - 1] + p + \beta(1-p)\overline{u} + \beta p\underline{u} - u_{aut} - \frac{p(qk+1)-1}{1-\beta}.$$

After further manipulations, we find that $sign\{u_{\max} - u_H\} = sign\{p - p^*\}$, where

$$p^* \equiv \frac{\sqrt{qk}-1}{qk-1}.$$

Simple calculations reveal that $1 < p^*(qk+1)$, $\frac{\partial p^*}{\partial q} < 0$ and $\lim_{q \to 1/k} p^* = 1/2$. Using these facts, we may draw the following conclusion: For all $q \in (\frac{2-\beta}{k\beta}, 1)$, there exists $p_L(q)$ satisfying $\max\{p_s, p^*\} \leq p_L(q) < \frac{1}{2}$ and such that, for all $p \in (p_L(q), \frac{1}{2})$, $u_{\max} > u_H$.[10]

We may now summarize as follows:

[10] For higher values of q, it is possible that the relevant constraint is that $p > 1 - \frac{1}{2q}$. In Figure 1, this inequality corresponds to the positively sloped line that separates regions I_2 and I_3. It is thus possible that $\max\{p_s, p^*\} < p_L(q)$.

Corollary 6.6. *There exists a subset of region I_2 for which the optimal SSE offers a strictly higher total payoff than does the optimal hybrid equilibrium, and thus* $u_H < u_{\max}$.

Finally, we note that the payoffs may also be easily compared in region I_1. In this region, the optimal SSE yields autarkic payoffs: $u_{\max} = u_{aut}$. Throughout this region, the optimal hybrid equilibrium corresponds to a HSSGL and thus yields the higher payoff $u_H = \widetilde{u} > u_{\max} = u_{aut}$.

7. Conclusion

Using a repeated game model with self-interested and privately informed players, we develop an equilibrium theory of trust and reciprocity. In our main analysis, players are willing to exhibit trust and thereby facilitate cooperative gains only if such behavior is regarded as a favor that must be reciprocated, either immediately or in the future. Private information is a fundamental ingredient in our theory. A player with the ability to provide a favor must have the incentive to reveal this capability, and this incentive is provided by an equilibrium construction in which favors are reciprocated.

Our study offers new predictions with respect to the social interactions of calculating, self-interested individuals. It may thus provide partial insight into the underlying sources of any instinctive preference for reciprocity and trust. We also offer an interpretation of an important category of human relationships, which Fiske (1992) refers to as equality-matching relationships. In such relationships, individuals exchange favors and keep track of the balance of favors so that equality may be achieved. In addition, we offer the novel prediction that the size of a favor owed may decline over time, as neutral phases of the relationship are experienced.

We compare our featured favor-exchange relationship with other benchmarks. We show it offers a higher total payoff than does a simple favor-exchange relationship. We also describe specific circumstances in which a relationship founded on favor exchange may be inferior to a relationship in which an infrequent and symmetric punishment (e.g., a risk of temporary or permanent autarky) keeps players honest. Finally, we show that a hybrid relationship, in which players begin with a honeymoon period and then either proceed to a favor-exchange relationship or suffer a symmetric punishment, can also offer scope for improvement.

Much work remains. First, we hope that some of our predictions can be tested in the lab. In part for this reason, we use the popular trust model. Second, future work might consider whether other behavioral regularities might be interpreted using the theory of repeated games with private information.

8. References

Abreu, D., D. Pearce and E. Stacchetti (1986), "Optimal Cartel Equilibria with Imperfect Monitoring," *Journal of Economic Theory*, 39.1, June, 251-69.

Abreu, D., D. Pearce and E. Stacchetti (1990), "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58.5, September, 1041-63.

Aoyagi, M. (2003), "Bid Rotation and Collusion in Repeated Auctions," *Journal of Economic Theory*, 112.1, September, 79-105.

Athey, S. and K. Bagwell (2001), "Optimal Collusion with Private Information," *Rand Journal of Economics*, 32.3, Autumn, 428-65.

Athey, S. and K. Bagwell (2004), "Collusion with Persistent Cost Shocks," mimeo.

Athey, S., K. Bagwell and C. Sanchirico (2004), "Collusion and Price Rigidity," *The Review of Economic Studies*, 71.2, April, 317-49.

Berg, J., J. Dickhaut, and K. McCabe (1995), "Trust, Reciprocity and Social History," *Games and Economic Behavior*, 10, 122-42.

Blau, P. M. (1964), **Exchange and Power in Social Life**, New York: Wiley.

Camerer, C. (2003), **Behavioral Game Theory: Experiments in Strategic Interaction**, Princeton: Princeton University Press.

Coleman, J. S. (1988), "Social Capital in the Creation of Human Capital," *The American Journal of Sociology*, 94, S95-S120.

de Quervain, D., U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck and E. Fehr (2004), "The Neural Basis of Altruistic Punishment," *Science*, 305, 1254-8.

Engle-Warnick, J. and R Slonim (2003), "The Evolution of Strategies in a Repeated Trust Game," mimeo.

Fehr, E. and S. Gachter (2000a), "Cooperation and Punishment in Public Good Experiments," *American Economic Review*, 90, 980-94.

Fehr, E. and S. Gachter (2000b), "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives*, 14, 159-81.

Field, A. (2002), **Altruistically Inclined?: The Behavioral Sciences, Evolutionary Theory, and the Origins of Reciprocity**, Ann Arbor: University of Michigan Press.

Fiske, A. P. (1992), "The Four Elementary Forms of Sociality: Framework for a Unified Theory of Social Relations," *Psychological Review*, 99.4, 689-723.

Fudenberg, D., D. Levine and E. Maskin (1994), "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62.5, September, 997-1039.

Green, E. J. (1987), "Lending and the Smoothing of Uninsurable Income." in Prescott, Edward C. and Neil Wallace, ed.. **Contractual Arrange-**

ments for Intertemporal Trade, Minnesota Studies in Macroeconomics series, vol. 1 Minneapolis: University of Minnesota Press, 3-25.

King-Casas, B., D. Tomlin, C. Anen, C. F. Camerer, S. R Quartz and P. R. Montague (2005), "Getting to Know You: Reputation and Trust in a Two-Person Economic Exchange," *Science*, April 1, Vol. 308, 78-83.

Mobius, M. (2001), "Trading Favors," mimeo.

Pinker, S. (2002), **The Blank Slate: The Modern Denial of Human Nature**, Penguin Press Science.

Ridley, M. (1997), **The Origins of Virtue: Human Instincts and the Evolution of Cooperation**, Viking Press.

Seabright, P. (2004), **The Company of Strangers: A Natural History of Economic Life**, Princeton: Princeton University Press.

Sethi, R. and E. Somanathan (2003), "Understanding Reciprocity," *Journal of Economic Behavior and Organization*, Vol. 50, Issue 1, 1-27.

Skryzpacz, A. and H. Hopenhayn (2004), "Tacit Collusion in Repeated Auctions," *Journal of Economic Theory*, 114.1, January, 153-69.

Wang, C. (1995), "Dynamic Insurance with Private Information and Balanced Budgets," *Review of Economic Studies*; 62.4, October, 577-95.

Watson, J. (1999), "Starting Small and Renegotiation," *Journal of Economic Theory*, 85, 52-90.

Watson, J. (2002), "Starting Small and Commitment," *Games and Economic Behavior*, 38, 176-99.

9. Appendix

9.1. Proof of Proposition 3.15

Consider the widest HSSGL. Let λ be the set of points on this HSSGL for which there exist multiple implementations. Suppose to the contrary that $\lambda \neq \emptyset$. Then it is straightforward to show that λ is convex and symmetric around the 45-degree line; therefore, λ contains $(\widetilde{u}, \widetilde{u})$, the middle point of this HSSGL. We will show that $(\widetilde{u}, \widetilde{u})$ is uniquely implemented, which then establishes that $\lambda = \emptyset$.

Consider a point (u, v) on the widest HSSGL and an implementation of it, $i = \{x, y, r, s, u_{i\theta}, v_{i\theta}, u_o, v_o\}$. Following the proof of Lemma 3.9, given any implementation, we can find an alternative implementation such that (2.4) and (2.5) bind, $r = s = 0$ and $u_{i0} = u_{i1} \equiv u_i$, with all other variables remaining the same. For such an implementation, suppose that (2.2) is slack; that is, suppose $\beta(u_a - u_o) > x \geq 0$. Then, for small $\varepsilon > 0$, if we decrease u_a by ε, increase u_o by $\frac{p}{1-2p}\varepsilon$, and change nothing else, the resulting implementation is feasible and implements (u, v). The same argument applies to a slack (2.3) as well. Therefore, given any implementation, we can find another implementation with the same values for x and y and with (2.2) and (2.3) binding.

We know that $\widetilde{i} = \{\widetilde{x} = \widetilde{y} = \frac{\beta}{\beta+\beta^*}, \widetilde{r} = \widetilde{s} = 0, \widetilde{u}_{a\theta} = \widetilde{v}_{b\theta} = \overline{u}, \widetilde{u}_{b\theta} = \widetilde{v}_{a\theta} = \underline{u}, \widetilde{u}_o = \widetilde{v}_o = \widetilde{u}\}$ implements $(\widetilde{u}, \widetilde{u})$.

We now argue that $(\widetilde{u}, \widetilde{u})$ is uniquely implemented. Suppose to the contrary that there exists another implementation $i = \{x, y, r, s, u_{i\theta}, v_{i\theta}, u_o, v_o\}$ of $(\widetilde{u}, \widetilde{u})$. As established above, we can focus on an implementation i such that (2.4), (2.5),(2.2) and (2.3) bind, $r = s = 0$, and $u_{i0} = u_{i1} \equiv u_i$. Define the following: $\Delta x \equiv x - \widetilde{x}$, $\Delta y \equiv y - \widetilde{y}$, $\Delta u_\pi \equiv u_\pi - \widetilde{u}_\pi$, $\Delta v_\pi \equiv v_\pi - \widetilde{v}_\pi$ for $\pi \in \{a, b, o\}$. Then $\Delta x = -\Delta y$, since $x + y = \widetilde{x} + \widetilde{y}$.

First, suppose that $x \neq \widetilde{x}$. As (2.2) and (2.3) bind under both $\widetilde{i}$ and i, we have that $\Delta x = \beta(\Delta u_a - \Delta u_o)$ and $\Delta y = \beta(\Delta v_b - \Delta v_o) = -\beta(\Delta u_b - \Delta u_o)$. Further, the promise-keeping constraint, (2.6), must hold under both $\widetilde{i}$ and i. Thus,

$$\begin{aligned} 0 &= -p\Delta x + pqk\Delta y + \beta[p\Delta u_a + p\Delta u_b + (1-2p)\Delta u_o] \\ &= -p\Delta x + pqk\Delta y + \beta[p(\Delta u_a - \Delta u_o) + p(\Delta u_b - \Delta u_o) + \Delta u_o], \end{aligned}$$

which implies

$$\Delta u_o = \frac{p(qk-1)}{\beta}\Delta x.$$

Since $\Delta x \neq 0$ and $\frac{p(qk-1)}{\beta} > 0$, Δu_o and Δx have the same sign. Recall that $\widetilde{u}_{a\theta} \equiv \widetilde{u}_a = \overline{u}$. Thus, $\Delta u_a \leq 0$. Now, if $\Delta x = \beta(\Delta u_a - \Delta u_o) > 0$, then $\Delta u_o < 0$, which is a contradiction. So, $\Delta x \leq 0$ must hold. Using a similar argument, we can show that $\Delta y \leq 0$ must hold as well. Then $\Delta y = -\Delta x$ implies $\Delta x \geq 0$, so that $\Delta x = \Delta y = 0$.

Second, suppose that $\Delta u_o \neq 0$ and $\Delta x = \Delta y = 0$. As (2.2) and (2.3) bind under both $\widetilde{i}$ and i, we have that $\Delta u_a = \Delta u_o$ and $\Delta u_b = \Delta u_o$. As just argued, $\Delta u_a \leq 0$. Similarly, with $\widetilde{u}_{b\theta} \equiv \widetilde{u}_b = \underline{u}$, $\Delta u_b \geq 0$. Thus, it must be that $\Delta u_o = 0$.

We conclude that $(\widetilde{u}, \widetilde{u})$ is uniquely implemented. Thus, $\lambda = \emptyset$. That is, every point (u, v) on the widest HSSGL is implemented uniquely. ■

9.2. Strongly Symmetric Equilibria (SSE)

We provide here proofs concerning strongly symmetric equilibria (SSE).Given any $\psi_s = [u_{aut}, u]$, following APS, define

$$
\begin{aligned}
B^{ss}(\psi_s) = \{ & v : \exists x \in [0,1],\ r \in [0, kx],\ v_o, v_{10}, v_{11} \in \psi_s \text{ such that} \\
& IC_x : 1 - x + q(r + \beta v_{11}) + (1-q)\beta v_{10} \geq 1 + \beta v_o \\
& IC_\theta : kx - r + \beta v_{11} \geq kx + \beta v_{10} \\
& PK : v = p[1 - x + q(r + \beta v_{11}) + (1-q)\beta v_{10}] \\
& \qquad + p[q(kx - r + \beta v_{11}) + (1-q)\beta v_{10}] \\
& \qquad + (1-2p)\beta v_o \qquad \qquad \qquad \qquad \}
\end{aligned}
$$

Let $\psi_s^* = [u_{aut}, u_{\max}]$ be the maximal fixed point of B^{ss}. That is, if $[u_l, u_h]$ is a fixed point of B^{ss}, then $[u_l, u_h] \subset [u_{aut}, u_{\max}]$.

Refer a pair (q, p) as an information structure. Consider the set $I = \{(q,p) : q \in (\frac{1}{k}, 1],\ p \in (0, \frac{1}{2}]\}$, which is the set of all feasible information structures.

9.3. Solving for ψ_s^*

Start with a very large u^1. Given u^n, by slightly abusing the notation, define u^{n+1} as follows:

$$
\begin{aligned}
u^{n+1} = B^{ss}(u^n) = \max\ v =& p[1 - x + q(r + \beta v_{11}) + (1-q)\beta v_{10}] \\
& + p[q(kx - r + \beta v_{11}) + (1-q)\beta v_{10}] \\
& + (1-2p)\beta v_o \\
=& p[1 + (qk-1)x + 2\beta(q v_{11} + (1-q) v_{10})] \\
& + (1-2p)\beta v_o
\end{aligned}
$$

subject to

$$x \in [0,1],\ r \in [0, kx],\ v_o, v_{10}, v_{11} \in [u_{aut}, u^n]$$
$$IC_x : 1 - x + q(r + \beta v_{11}) + (1-q)\beta v_{10} \geq 1 + \beta v_o$$
$$IC_\theta : kx - r + \beta v_{11} \geq kx + \beta v_{10}$$

We will employ APS to solve for $u_{\max}$. Accordingly, if $u^1 > u_{\max}$, then $B^{ss}(u^n) < u^n$ and $\lim_{n\to\infty} u^n = u_{\max} = B^{ss}(u_{\max})$. Let $u_{eff} = \frac{pqk}{1-\beta}$ be the average utility of the first best solution, i.e. investing $x = 1$ every period when some agent receives positive income. Then $u_{\max} \leq u_{eff}$; therefore, it would suffice to start with $u^1 = u_{eff.}$

Proposition: For any $u \geq u_{aut}$, IC_x and IC_θ bind at the solution of $B^{ss}(u)$.

Proof:

The proof proceeds via three claims.

Claim 1: $v_{11} = u$.

Proof: If $v_{11} < u$, then increasing v_{11} increases the objective without violating IC_x and IC_θ. Contradiction.□

Claim 2: IC_θ is binding.

Proof: Suppose in contrary that IC_θ is slack. Then $u - v_{10} > \frac{r}{\beta} \geq 0$, i.e. $u > v_{10}$. Now increase v_{10} by $\varepsilon > 0$. IC_x becomes slack, IC_θ continues to hold if ε is small enough. The objective increases. Contradiction.□

Claim 3: IC_x is binding.

Proof: Suppose in contrary that IC_x is slack. Then $v_o = u$ and $x = 1$. To see this, check the following: If $v_o < u$, then increase v_o by $\varepsilon > 0$. IC_x is not violated if ε is small enough; IC_θ is not affected; and the objective increases. Contradiction. If $x < 1$, then increase x by $\varepsilon > 0$. IC_x is not violated if ε is small enough; IC_θ is not affected; and the objective increases since $qk > 1$. Contradiction.

Substituting $v_o = u$ and $x = 1$, IC_x becomes $q(r+\beta u) + (1-q)\beta v_{10} > 1 + \beta u$, equivalently $qr > 1 + (1-q)\beta(u - v_{10})$. Binding IC_θ yields $r = \beta(u - v_{10})$. These together imply $(2q-1)r > 1$. Thus, a contradiction is immediate unless $2q-1 > 0$. In that event, $r > \frac{1}{2q-1} > 0$, and so $u > v_{10}$. We can thus increase v_{10} by $\varepsilon > 0$, and decrease r by $\beta\varepsilon$, and IC_θ continues to hold. Then the total change on the left hand side of IC_x can be computed as $(1-2q)\beta\varepsilon$. Since IC_x is slack by supposition, IC_x continues to hold if ε is small. The total change in the objective can be computed as $2p(1-q)\beta\varepsilon > 0$, so the objective increases. Contradiction.□

This completes the proof of the proposition.

Now, binding IC_x and binding IC_θ imply $v_o = 2qu + (1-2q)v_{10} - \frac{x}{\beta}$. Substituting $v_{11} = u$, $v_o = 2qu + (1-2q)v_{10} - \frac{x}{\beta}$, and $r = \beta(u - v_{10})$ yields

$$B^{ss}(u) = \max\ p + 2q(1-p)\beta u + (p(qk+1)-1)x + (1-2q(1-p))\beta v_{10}$$

subject to

$$v_{10}, v_o = 2qu + (1-2q)v_{10} - \frac{x}{\beta} \in [u_{aut}, u] \quad (9.1)$$

$$0 \leq x \leq 1 \quad (9.2)$$

$$\beta(u - v_{10}) \leq kx \quad (9.3)$$

The following three curves will be crucial in characterizing the optimal strongly symmetric equilibrium:

$$I : p = \frac{1}{qk+1} \text{ equivalently } p(qk+1) - 1 = 0$$

$$II : p = \frac{2q-1}{2q} \text{ equivalently } 1 - 2q(1-p) = 0$$

$$III : p = \frac{k(2q-1)-1}{qk-1}$$

Curve I is convex and decreasing in q. Curves II and III are both concave and increasing in q. Furthermore, all three curves intersect at $q^* = \frac{k+\sqrt{k^2+8k}}{4k} \in (\frac{1}{2}, 1)$. For $q < q^*$, curve I lies above II, which lies above III. For $q > q^*$, curve III lies above II, which lies above I. The three curves partition the set of information structures into six subsets. See Figure 3. We drop the superscipt of u^n to simplify the notation.

Region 1: $p(qk+1) - 1 \leq 0$, $1 - 2q(1-p) \geq 0$, $p - \frac{k(2q-1)-1}{qk-1} \geq 0$

The coefficient of x and v_{10} are nonpositive and nonnegative, respectively, in the objective of $B^{ss}(u)$. Therefore, the objective function is nonincreasing in x and nondecreasing in v_{10}. Setting $x = 0$, $v_{10} = u$, check that $v_o = u$ and $r = 0$ so that all the constraints are satisfied. This implies $B^{ss}(u) = p + \beta u$ for all u. Then $B^{ss}(u) < u$ as long as $u > u_{aut}$. Therefore, $u^\infty = \lim_{n\to\infty} u^n = u_{aut}$. Hence, $u_{\max} = u_{aut}$ in this region.

Region 2 and 3: $p(qk+1) - 1 > 0$, $1 - 2q(1-p) \geq 0$

The coefficient of x is positive and the coefficient of v_{10} is nonnegative in the objective of $B^{ss}(u)$. Check whether $x = 1$ and $v_{10} = u$ is a solution for $B^{ss}(u)$. Substituting $x = 1$ and $v_{10} = u$, we obtain $B^{ss}(u|x = 1, v_{10} = u) = p+\beta u+p(qk+1) - 1$. Also, $B^{ss}(u|x = 1, v_{10} = u) < u$ if and only if $u > u_{aut} + \frac{p(qk+1)-1}{1-\beta} = \gamma$. Therefore, starting with $u^1 = u_{eff} > \gamma$, we obtain a decreasing sequence of $\{u^n\}$ with $u^\infty = \lim_{n\to\infty} u^n = \gamma$.

Now check feasibility of the solution $x = 1$ in the limit: $v_o = u^\infty - \frac{1}{\beta} \geq u_{aut}$ if and only if $\beta \geq \frac{1}{p(qk+1)}$. Since $p(qk + 1) - 1 > 0$ i.e. $\frac{1}{p(qk+1)} < 1$ in these regions, there exists $\beta \geq \frac{1}{p(qk+1)}$. Then, for all $u \geq u_{\max}$, all the constraints are satisfied when $x = 1$, $v_{10} = u$. Therefore, $u_{\max} = u_{aut} + \frac{p(qk+1)-1}{1-\beta}$. Furthermore, check that $\lim_{p\to 1/2} u_{\max} = \lim_{p\to 1/2} u_{eff}$.

To address the possibility raised in Proposition 5.2, we now further suppose that $q \leq 1/2$ and $\beta < \frac{1}{p(qk+1)}$. Suppose that both $x < 1$ and $v_{10} < u_{\max}$ is satisfied in the solution of $B^{ss}(u_{\max})$. Then, we may increase v_{10} by ε and increase x by $(1 - 2q)\beta\varepsilon$. If $\varepsilon > 0$ is small enough, $x < 1$ and $v_{10} < u_{\max}$ continue to hold. Furthermore, v_o remains the same as defined by (9.1). Thus, all constraints hold and the objective of $B^{ss}(u_{\max})$ increases, which is a contradiction. As a result, either $x = 1$ or $v_{10} = u_{\max}$ in the solution of $B^{ss}(u_{\max})$. Suppose $x = 1$. Given $q \leq 1/2$, we may use (9.1) to find that

$$v_o = 2qu_{\max} + (1 - 2q)v_{10} - \frac{1}{\beta} \leq 2qu_{\max} + (1 - 2q)u_{\max} - \frac{1}{\beta} = u_{\max} - \frac{1}{\beta}.$$

Arguing as in the previous paragraph, we may now use $\beta < \frac{1}{p(qk+1)}$ and conclude that it is not possible that $x = 1$ in the solution of $B^{ss}(u_{\max})$. Thus, it can only be that $x < 1$ and $v_{10} = u_{\max}$ hold in the solution of $B^{ss}(u_{\max})$. Note that (9.1) is the only constraint that causes $x < 1$. Therefore, (9.1) is binding from below in the solution of $B^{ss}(u_{\max})$. This yields $v_o = 2qu_{\max} + (1 - 2q)u_{\max} - \frac{x}{\beta} = u_{aut}$, so that $x = \beta(u_{\max} - u_{aut})$. Substituting $v_{10} = u_{\max}$ and $x = \beta(u_{\max} - u_{aut})$ into the objective of $B^{ss}(u_{\max})$, we obtain

$$u_{\max} = B^{ss}(u_{\max}) = p + \beta u_{\max} + (p(qk + 1) - 1)\beta(u_{\max} - u_{aut}).$$

Simplifying and using $\beta < \frac{1}{p(qk+1)}$, we obtain $u_{\max} = u_{aut}$.

Regions 4,5,6: $1 - 2q(1 - p) < 0$

The coefficient of v_{10} is negative in the objective function. Also $1-2q(1-p) \leq 0$ implies $q \geq \frac{1}{2(1-p)} > \frac{1}{2}$, so the coefficient of v_{10} in (9.1) is negative as well.

Regions 5,6: $p(qk+1)-1<0$ i.e. $p<\frac{1}{qk+1}$

The coefficient of x is negative in the objective function. q varies between $\frac{1}{2}$ and 1. We will consider the following three subregions:

Subregion 1: $q<\frac{1+k}{2k}\Leftrightarrow 2q-\frac{1}{k}<1$

This subregion contains a part of Region 6. Note that $2q-\frac{1}{k}\in(0,1)$. So, $(2q-\frac{1}{k})u+(1-2q+\frac{1}{k})v_{10}\in[u_{aut},u]$ if $v_{10}\in[u_{aut},u]$. Also, if (9.3) binds, $v_o=(2q-\frac{1}{k})u+(1-2q+\frac{1}{k})v_{10}$. Now, suppose that (9.3) is slack at the optimal solution. Then decrease x so that (9.3) binds. Then $v_o\in[u_{aut},u]$ holds because of the previous argument, and the objective increases. A contradiction. Therefore, (9.3) is binding at the optimal solution.

Now consider $x>0$, $v_{10}<u$, and a decrease in x by $\varepsilon>0$. In order to satisfy $\beta(u-v_{10})=kx$, increase βv_{10} by $k\varepsilon$. This changes the objective by $\Delta=-(p(qk+1)-1)\varepsilon+(1-2q(1-p))k\varepsilon$. Check that $\Delta>0\Leftrightarrow p>\frac{k(2q-1)-1}{qk-1}$. The last inequality holds in Region 6. Therefore, check $x=0$ and $v_{10}=u$. All the constraints are satisfied when $x=0$ and $v_{10}=u$. So, $x=0$ and $v_{10}=u$ hold at the optimal solution for all $u>u_{\max}$. Then $B^{ss}(u)=p+\beta u$, and $B^{ss}(u)<u\Leftrightarrow u>u_{aut}$, so that we have $u_{\max}=u_{aut}$.

Subregion 2: $\frac{1+k}{2k}\leq q<q^*$

Parts of Region 5 and Region 6 are included.

Suppose that (9.3) binds at the optimal solution. Then $v_o=(2q-\frac{1}{k})u+(1-2q+\frac{1}{k})v_{10}$ is (weakly) decreasing in v_{10}, and $v_o=u$ when $v_{10}=u$. Suppose $\frac{k+1}{2k}<q$. Then $v_o\leq u$ implies that $v_{10}=v_o=u$, which implies $x=0$. Alternatively, suppose $\frac{k+1}{2k}=q$. Then $v_o=u$ for all v_{10}. If $x>0$ and $v_{10}<u$, we can follow the argument above (for subregion 1), and decrease x by $\varepsilon>0$ and increase βv_{10} by $k\varepsilon$. We then satisfy (9.3) and induce $\Delta>0$, since $p>\frac{k(2q-1)-1}{qk-1}=0$. Thus, $v_{10}=v_o=u$ and $x=0$ again follows. So, in either case, $B^{ss}(u)=p+\beta u$.

Now suppose that (9.3) is slack. If $v_o=2qu+(1-2q)v_{10}-\frac{x}{\beta}<u$, we can increase the objective by decreasing x. So, $v_o=u$ must hold. Then $x=\beta(2q-1)(u-v_{10})$. Substituting x in $B^{ss}(u)$, and taking its partial derivative with respect to v_{10}, we obtain $\frac{\partial B^{ss}(u)}{\partial v_{10}}=p\beta[-2kq^2+qk+1]>0$ since $q<q^*$. Also check that $v_{10}=u$ implies $x=0$ and $v_o=u$. That is, all the constraints are satisfied. Therefore, $v_{10}=v_o=u$ and $x=0$ hold in the solution of $B^{ss}(u)$. Again, $B^{ss}(u)=p+\beta u$.

We obtain $B^{ss}(u)=p+\beta u$ in both cases. Hence, by taking the limit, we obtain $u_{\max}=u_{aut}$ in this subregion.

Subregion 3: $q\geq q^*$

Only a part of Region 5 is included.

Suppose that (9.3) binds at the optimal solution of $B^{ss}(u_{\max})$. The same argument in Subregion 2 applies: $v_o = (2q - \frac{1}{k})u + (1 - 2q + \frac{1}{k})v_{10}$ is decreasing in v_{10}, and $v_o = u$ when $v_{10} = u$. Then $v_o \leq u$ implies that $v_{10} = v_o = u$, which implies $x = 0$. So, $B^{ss}(u_{\max}) = p + \beta u_{\max}$, which yields $u_{\max} = u_{aut}$. We will rule out this possibility next.

Now suppose that (9.3) is slack. Then, by the same reasoning in Subregion 2, $v_o = u$ and $x = \beta(2q-1)(u - v_{10})$. Substituting these in $B^{ss}(u)$, we obtain $\frac{\partial B^{ss}(u)}{\partial v_{10}} = p\beta[-2kq^2 + qk + 1] < 0$ since $q \geq q^*$. Therefore choose, $v_o = u$, $x = \beta(2q-1)(u - v_{10})$, and v_{10} as small as possible subject to $x \leq 1$ and $v_{10} \geq u_{aut}$. Check that $\beta(u - v_{10}) \leq kx$ is equivalent to $q \geq \frac{1+k}{2k}$, which is satisfied in this subregion. So, either (i) $v_{10} = u_{aut}$ and $x = \beta(2q-1)(u - u_{aut}) \leq 1$, or (ii) $x = 1$ and $v_{10} = u - \frac{1}{\beta(2q-1)} \geq u_{aut}$ holds in the solution. As we start with a large u, $x = \beta(2q-1)(u - u_{aut})$ will exceed 1, therefore case (ii) will hold for large u.

Now check if case (ii) holds in the limit. In case (ii), we have

$$\begin{aligned} B^{ss}(u) &= p + 2q(1-p)\beta u + (p(qk+1) - 1) + (1 - 2q(1-p))\beta(u - \frac{1}{\beta(2q-1)}) \\ &= p + \beta u + \lambda \end{aligned}$$

where $\lambda = (p(qk+1) - 1) - \frac{1-2q(1-p)}{2q-1} = \frac{p}{2q-1}(2kq^2 - qk - 1) \geq 0$ since $q \geq q^*$.

In the limit, we obtain $u^\infty = u_{aut} + \frac{\lambda}{1-\beta}$. So, $v_{10} = u^\infty - \frac{1}{\beta(2q-1)} \geq u_{aut}$ is equivalent to $\beta \geq \hat{\beta} = \frac{1}{1+p(2kq^2-qk-1)}$. So, for $\beta \geq \hat{\beta}$, we obtain $u_{\max} = u_{aut} + \frac{\lambda}{1-\beta}$ and $\lim\limits_{q\to 1} u_{\max} = \lim\limits_{q\to 1} u_{eff}$. This also rules out binding (9.3). Note that $\hat{\beta} < 1 \Leftrightarrow q > q^*$

Region 4: $p \geq \frac{1}{qk+1}$.

In this region, the coefficients of x and v_{10} in the objective of $B^{ss}(u)$ are nonnegative and negative, respectively. So, the objective function is nonincreasing in x and decreasing in v_{10}.

Obviously, all the constraints cannot be slack at the optimal solution. Consider $x < 1$ and $v_{10} > u_{aut}$. Decrease v_{10} by ε and increase x by $\beta(2q-1)\varepsilon$. Then v_o remains unchanged. The left hand side of (9.3) increases by $\beta\varepsilon$. The right hand side of (9.3) increases by $k(2q-1)\beta\varepsilon$. Note that $q > \frac{1+k}{2k}$ i.e. $k(2q-1) > 1$ in this region. Therefore, (9.3) becomes slack and the objective increases. So (9.3) is slack in the optimal solution.

The same argument also implies that (i) if $x < 1$ then $v_{10} = u_{aut}$, and (ii) if $v_{10} > u_{aut}$ then $x = 1$. Otherwise it would be possible to increase the objective as above.

In case (i), $x < 1$ would also imply $v_o = u_{aut}$. Otherwise, a small increase in x would increase the objective without violating any constraint. Similarly, in case (ii), $v_{10} > u_{aut}$ would also imply $v_o = u$. Otherwise, a small decrease in v_{10} would increase the objective without violating any constraint, since (9.3) is slack.

In case (i), solving x from $v_o = 2qu+(1-2q)v_{10}-\frac{x}{\beta}$, we obtain $x = 2q\beta(u-u_{aut})$. Then $x < 1$ is equivalent to $u - u_{aut} < \frac{1}{2q\beta}$. In case (ii), solving v_{10} from $v_o = 2qu + (1-2q)v_{10} - \frac{x}{\beta} \in [u_{aut}, u]$, we obtain $v_{10} = u - \frac{1}{\beta(2q-1)}$. Then $v_{10} > u_{aut}$ is equivalent to $u - u_{aut} > \frac{1}{\beta(2q-1)}$. Since $\frac{1}{2q\beta} < \frac{1}{\beta(2q-1)}$, $u - u_{aut} < \frac{1}{2q\beta}$ and $u-u_{aut} > \frac{1}{\beta(2q-1)}$ cannot hold simultaneously. For large u, $u-u_{aut} > \frac{1}{\beta(2q-1)}$ holds. Therefore, for large u, by setting $x = 1$, $v_o = u$ and $v_{10} = u - \frac{1}{\beta(2q-1)}$, we obtain $B^{ss}(u) = p+\beta u+\lambda$ and $u^{\infty} = u_{aut}+\frac{\lambda}{1-\beta}$ as above. Also, $v_{10} = u^{\infty}-\frac{1}{\beta(2q-1)} \geq u_{aut}$ is equivalent to $\beta \geq \hat{\beta}$ as above. So, for $\beta \geq \hat{\beta}$, we obtain $u_{\max} = u_{aut} + \frac{\lambda}{1-\beta}$ and $\lim_{q\to 1} u_{\max} = \lim_{q\to 1} u_{eff}$. ■

Figure 1: The Partition of the Information Structure

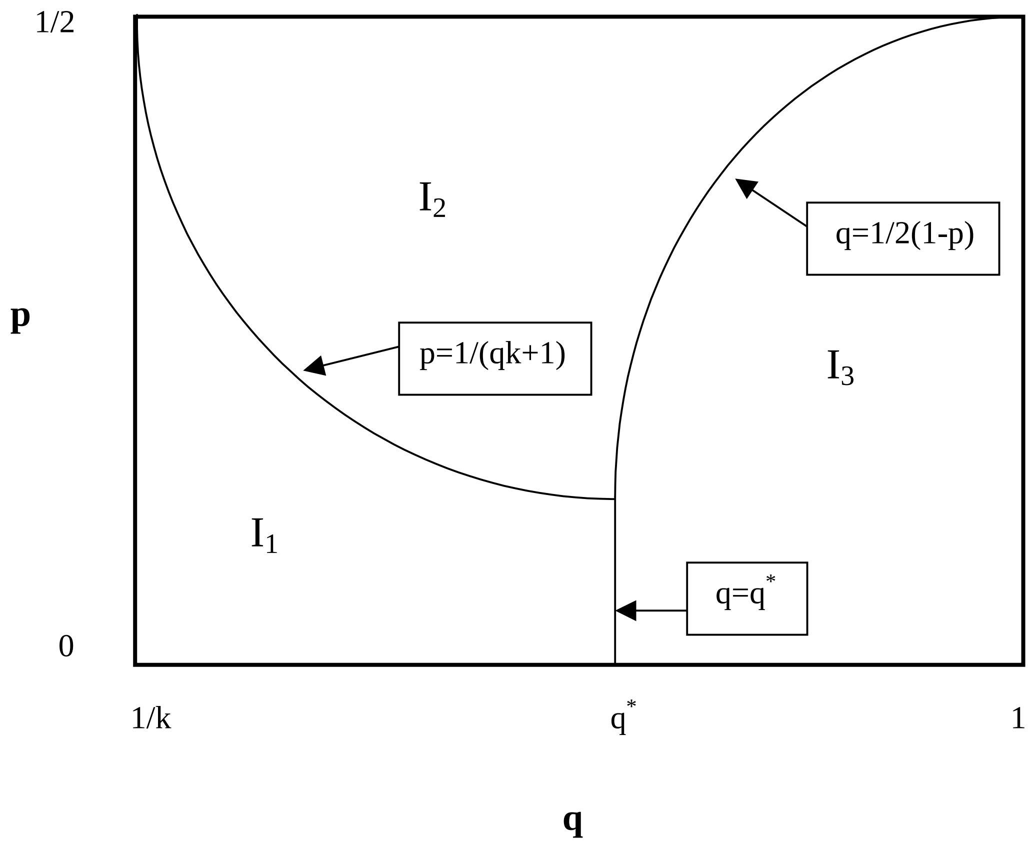

Figure 2

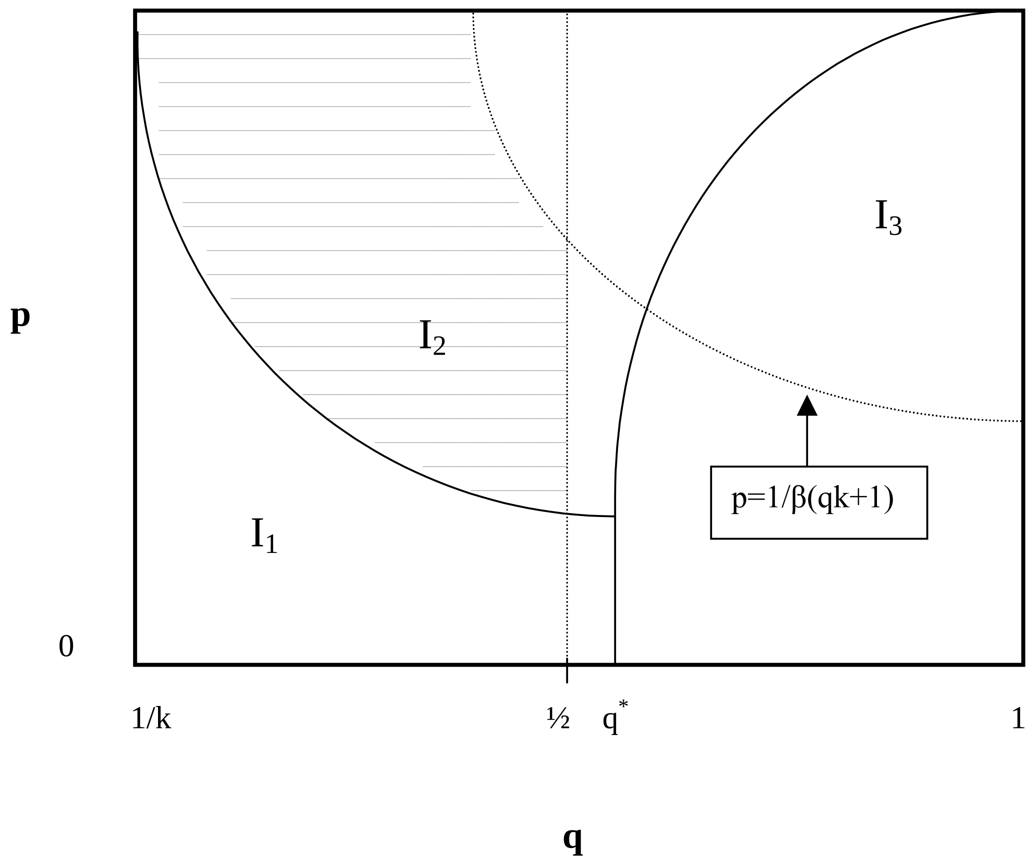

Figure 3: Regions