Theoretical Economics 2 (2007), 299–354 1555-7561/20070299

Efficiency in repeated trade with hidden valuations

SUSAN ATHEY
Department of Economics, Harvard University

DAVID A. MILLER
Department of Economics, University of California, San Diego

We analyze the extent to which efficient trade is possible in an ongoing relationship between impatient agents with hidden valuations (i.i.d. over time), restricting attention to equilibria that satisfy ex post incentive constraints in each period. With *ex ante* budget balance, efficient trade can be supported in each period if the discount factor is at least one half. In contrast, when the budget must balance *ex post*, efficiency is not attainable, and furthermore for a wide range of probability distributions over their valuations, the traders can do no better than employing a posted price mechanism in each period. Between these extremes, we consider a "bank" that allows the traders to accumulate budget imbalances over time, but only within a bounded range. We construct non-stationary equilibria that allow traders to receive payoffs that approach efficiency as their discount factor approaches one, while the bank earns exactly zero expected profits. For some probability distributions there exist equilibria that yield exactly efficient payoffs for the players and zero profits for the bank, but such equilibria require high discount factors.

KEYWORDS. Repeated trade, Myerson–Satterthwaite Theorem, repeated games, private information, dynamic mechanism design, ex post incentive compatibility, budget balance.

JEL CLASSIFICATION. C72, C73, D82, L14.

1. INTRODUCTION

The problem of trade concerns two people, one of whom possesses a perishable good that both of them like. The basic question is: If their valuations are private information,

Susan Athey: athey@fas.harvard.edu
David A. Miller: d9miller@ucsd.edu

We are indebted to Sean Meyn, Jeffrey Ely, and two referees for valuable suggestions. We are also grateful for discussions with David Goldbaum, Andrzej Skrzypacz, and Yixiao Sun, as well as comments from seminar participants at Stanford, UCSD, the 2005 NBER/NSF Decentralization Conference, NYU, and Pennsylvania. Ryoji Hiraguchi provided outstanding research assistance. Athey gratefully acknowledges support from the National Science Foundation, grants SES–9983820 and SES–0351500. An earlier version of this paper appeared as a chapter in Miller's Ph.D. dissertation at Stanford University. He thanks Susan Athey, Ilya Segal, and Steven Tadelis, his dissertation committee, for their valuable guidance. He also thanks Jonathan Levin and Andrzej Skrzypacz for participating in his dissertation defense. He gratefully acknowledges support from the Stanford Institute for Economic Policy Research via a dissertation fellowship.

can they design monetary payments that give both of them the incentive to allocate the good to the one with the highest valuation? Trade in a static setting has been well studied (following Myerson and Satterthwaite 1983). We consider the extent to which trade can be supported in repeated environments in which valuations are identically and independently distributed (i.i.d.) over time.

We study equilibria that satisfy *ex post incentive compatibility* (IC) and *ex post individual rationality* (IR) in each period.[1] These ex post constraints guarantee that our equilibria are robust to certain types of model mis-specification. Firstly, in our model traders' private information in any period concerns only their own valuations for that period, but our equilibria would not be disrupted if the model were expanded to allow for arbitrary higher order beliefs over their valuations in that period. That is, our equilibria are robust to the introduction of any payoff-irrelevant signals, even if they are correlated with the traders' current valuations (although they cannot be correlated with future valuations). Secondly, many of the equilibria we construct are stationary, and a stationary equilibrium is not disrupted even if the distribution of valuations is perturbed, as long as the agents' future utilities do not decrease to the point that ex post IR is violated.

Our main results, briefly, are that equilibria under our ex post constraints have the following properties.

1. If unlimited, actuarially fair insurance against budget imbalances is available, then efficiency is attainable for discount factors of at least one half.

2. If the institutional environment cannot absorb any budget imbalances, then efficiency is unattainable, and furthermore for many settings the optimal equilibrium in the repeated game is simply to repeat the same inefficient posted price mechanism that is optimal in the one-shot game (for all discount factors).

3. If the traders have access to a bounded, collateralized credit line, and they have sufficient collateral, then (a) attainable payoffs approach efficiency as the discount factor approaches one, and furthermore (b) in some settings efficiency is attainable for sufficiently high discount factors.

Result (1) can be viewed as favorable to trade, since it demonstrates that efficiency can be supported in an ongoing trading relationship even under constraints much stronger than those under which the Myerson–Satterthwaite theorem ruled out efficiency in one-shot settings.

On the negative side, efficient trade in result (1) relies on unbounded insurance, which is an unreasonably strong condition. The problem is that in order to support efficiency in each period, the traders must allow for budget imbalances after many realizations. The accumulated imbalances can be conceptualized as the balance of a credit line account, which is provided to the traders by a budget-breaking institution, such as a bank, at a zero rate of interest. The problem is that if the players are to trade efficiently in

[1] I.e., ex post equilibrium in every stage game; see Chung and Ely (2002).

every period, then the balance must follow a random walk, so the traders will eventually either exhaust their wealth by depositing it all into the account, or exhaust the bank's resources by withdrawing them all from the account.

Given that result (1) cannot yield exact efficiency under realistic institutions, the next question to ask is what fraction of efficiency is attainable in the absence of a budget-breaking institution. Result (2) shows that efficiency cannot be approximated, no matter how patient are the traders. For many settings, the traders optimally sacrifice a large fraction of the possible gains from efficient trade by using a simple "posted price mechanism," whereby the buyer and seller accept or reject the chance to trade at a single, fixed price. The advantage of a posted price mechanism is that the only transfer required is the price itself, and the price is paid if and only if trade occurs; the disadvantage, of course, is its inefficiency.

Taken together, results (1) and (2) might seem to imply that there is little hope for traders to attain payoffs close to efficiency using realistic institutions. Fortunately, results (3a) and (3b) show that this is not at all the case. For result (3a), we construct a class of mechanisms that supports efficient trade in every period, yet assures that the accumulated budget imbalances stay within a compact interval, and that the institution that provides the credit line account earns zero expected profits. These mechanisms operate by fine-tuning the payments between the players so that when the account balance is low they always make a deposit, but when the account balance is high they always make a withdrawal. Whether this fine-tuning is possible depends on the probability distribution over their valuations.

For result (3b), we construct a more broadly applicable class of mechanisms such that, as in result (3a), the accumulated imbalances are confined to a compact interval and the institution that provides the credit line account earns zero expected profits. Given a sufficiently high discount factor, the fraction of periods in which trade is efficient can be made arbitrarily close to one, and the traders' expected payoffs can be made arbitrarily close to the efficient payoffs. The way that the players trade and make payments in each period depends on the balance in the account at the start of the period. When the balance is close to the lower bound, the players trade inefficiently so as to make certain that the balance does not decrease, and increases in expectation. When the balance is close to the upper bound, the players do not trade at all and simply make a withdrawal from the account. When the balance is not close to either bound, the players trade efficiently.

A numerical example (using the uniform distribution for the traders' valuations) indicates that our near-efficient mechanisms can be relevant even when the traders' discount factor is not close to one. In the example, the traders can capture 90.1% of the gains from efficient trade even when their discount factor is just 0.67. When their discount factor is 0.95, they capture 98.9% of the gains from efficient trade. For comparison, the optimal posted price mechanism for the example captures only 75% of the gains from efficient trade, while the efficient mechanism from result (3a) captures all the gains from trade but requires a discount factor of at least 0.93.

Before introducing the model, we first provide an institutional interpretation of our key assumptions. Then in Section 2 we outline the model, the mechanism design approach, and some preliminary results. Section 3 proves result (1), and Section 4 proves result (2). Section 5 expands the model to allow mechanisms to depend on the account balance, and proves result (3). Section 6 considers some of the related literature, and Section 7 concludes.

1.1 *Institutions and assumptions*

The institutional settings we consider are represented by the restrictions that they impose on equilibria in the repeated relationship. The institutions and assumptions we consider fall into three categories: incentive compatibility (IC), budget balance (BB), and individual rationality (IR). At the outset, we should point out that IC imposes restrictions on equilibria in the game, while the BB and IR assumptions are better thought of as conditions on the structure of the game itself. We are concerned with two possible ways of mis-specificying the model: failing to account for payoff-irrelevant information that is correlated with the players' valuations for the object, and failing to specify the correct distribution of valuations.

Incentive compatibility Our equilibrium concept in the repeated game is *ex post perfect public equilibrium* (EPPPE; Miller 2007b). This concept requires that the mechanism in each stage game be implemented in ex post equilibrium (Chung and Ely 2002), which corresponds to ex post IC and ex post IR.[2] Although ordinary perfect public equilibrium is more commonly used in the literature, it requires that the players have common knowledge of the distribution over their valuations in every period. And even if the players start the period with common knowledge of the distribution, a perfect public equilibrium will generally not survive if any player subsequently learns any additional information that is correlated with the valuations. For example, if a player were able to delay his announcement or spy on the other player's information, he might be tempted to make an untruthful announcement once learning about the other player's type. EPPPE, in contrast, does not require common knowledge of the distribution, because a player's truthful announcement is incentive compatible regardless of any information he may learn that is not relevant to his own payoff.

Our equilibria do require the players to have common knowledge that each others' incentives satisfy ex post IC. Many of the equilibria we construct are stationary—that is, they do not employ changes in promised future utility to provide incentives. In stationary equilibria, there is common knowledge of ex post IC if it is common knowledge that the distribution over valuations does not change over time, even if the distribution itself is not common knowledge.[3] We construct also some non-stationary equilibria; in

[2]Since the parties in our model have private valuations, ex post implementation is equivalent to dominant strategies implementation.

[3]If the players do not have common knowledge of the distribution, it may be difficult for them to reach agreement on which equilibria are optimal. But this is a problem of equilibrium selection, not equilibrium robustness.

these cases it is sufficient, although perhaps not necessary, that the players have common knowledge of the distribution in each period and that no new information about a particular period's distribution arises until the previous period is completed. To ease the exposition, we maintain the assumption that players have common knowledge of the distribution (which allows us to avoid defining the universal type space).

It is difficult to identify institutions that could support the assumption that players learn nothing other than their own valuations during the period, and thereby justify imposing interim IC rather than ex post IC. At a minimum, such institutions would need to enforce simultaneous communication and prevent spying; these requirements suggest an institutional environment in which the traders interact anonymously. But when the traders are anonymous, it is hard to justify common knowledge in the first place, since they will not be familiar with each others' tastes and opportunity costs.

Our concept of optimality requires the distribution of valuations to be known. However, the equilibria we construct in Sections 3 and 4 are stationary (i.e, they use the same mechanism in every period along the equilibrium path). Since a stationary equilibrium does not use changes in future utility to provide incentives, a change in the underlying distribution has the effect of a lump sum transfer, and so does not disrupt ex post IC. So if we design an optimal stationary equilibrium using the wrong distribution, it still satisfies ex post IC even though it may not be optimal.

Finally, we note that for environments in which interim IC is reasonable, the appropriate equilibrium concept in the repeated game is ordinary perfect public equilibrium, and the folk theorem applies. That is, if the players are sufficiently patient, they can support exact efficiency even under ex post budget balance (see Miller 2007a). In a separate note (Athey and Miller 2006), we show how to compute the lowest discount factor that can support efficient trade in an ordinary perfect public equilibrium.

Budget balance We consider several alternative restrictions on the balance of payments between the parties. At one extreme, an institutional environment with unbounded, actuarially fair insurance requires only that the payments between the parties balance in expectation; this is ex ante BB. It is important to note that an institution that provides unbounded insurance must be truly willing to absorb arbitrarily large accumulated imbalances over time, since we show that the absolute value of the accumulated imbalances eventually exceeds any finite bound with probability one when trade is efficient in every period. This requirement casts doubt on the feasibility of such an institution. At the other extreme, a lack of institutions for insurance—even self-insurance—may require that the payments between the parties balance exactly in each period; this is ex post BB. Between these two extremes is an environment in which insurance is available but limited. For example, if the trading partners have access to a credit line account with finite bounds on the balance, we say that any mechanism must satisfy *bounded budget account* (BBA).

There are several ways to conceptualize BBA. It can be considered a form of self insurance, by which traders use their own finite wealth to absorb the budget imbalances incurred in the course of trading. That is, when the mechanism requires them to make a "deposit," they must set aside those funds to be reserved until a future period in which the mechanism allows them to take a "withdrawal." Our preferred heuristic for thinking about BBA is that the traders have access to a credit line account with a zero interest rate,[4] and they can borrow from this account up to the amount of collateral that they hold. (They can also engage in "negative borrowing" by putting in extra funds, but only up to a finite bound.) In this sense, the credit limit on the account depends on their material wealth.[5]

We note that ex ante BB and BBA are not robust to changes in the distribution of valuations, since they depend on the details of the distribution.[6] Ex post BB is robust, of course, since it requires budget balance for every realization. Since budget balance properties in general do not depend on higher order beliefs (except through the possible failure of incentive compatibility), when ex post IC holds they are robust to the introduction of any payoff-irrelevant information.

Individual rationality Individual rationality is a constraint that compares payoffs attained within the mechanism to some outside options that are not chosen in equilibrium. Ex post IR allows the players to make this comparison after they have learned the realizations of all players' types. The outside option we consider is autarky, in which the players do not trade. We can interpret "no trade" literally, but an alternative interpretation is that the expected value of trade with other agents is normalized to zero (under the assumption that the trading partners have higher potential surplus trading with one another than with alternative partners). Autarky not only is a stage game equilibrium, but also yields the minimax utilities for both players, and hence yields the lowest expected utility possible in any sequential equilibrium. Selecting autarky as the outside option is essentially equivalent to allowing the players to commit not to renegotiate their equilibrium selection.[7]

Ex post IR, when it is satisfied with strict inequality, is robust to small changes in the distribution of valuations, but can fail if there is a large change. Ex post IR, naturally, is robust to any payoff-irrelevant information because it applies after players have learned what their payoffs will be.

---

[4]Non-zero interest rates could also be considered at the cost of additional complexity; we discuss this in footnote 17, below.

[5]BBA can be motivated in a general equilibrium context as follows. Consider an economy with three agents. They each have an initial endowment of a uniform durable consumption good. One agent is the bank, one is the buyer, and one is the seller. Each period the seller is endowed with a random perishable object (the object of trade), and each trader is endowed with a large, finite quantity of a uniform perishable consumption good. Each trader's utility function is quasilinear in the sum of her uniform durable and uniform perishable consumption goods, but the bank values only the durable good. Since the traders cannot put more in the bank than they have durable consumption goods, and the bank cannot pay out more than the amount of durable consumption good that it started with, this economy requires BBA.

[6]Under BBA, the traders are asked to make payments that depend on their expectations of future payoffs.

[7]If players could not so commit, then the outside options would need to be adjusted endogenously to take into account the extent to which the players could renegotiate off the equilibrium path. Since this would affect the set of equilibrium paths that could be supported, investigating the efficiency of trade with renegotiation is a substantial and potentially interesting topic. We do not pursue it here.

2. The model

2.1 *Stage mechanisms*

There are two players—a buyer (player b) and a seller (player s). At the outset of each stage, the seller is endowed with a perishable object, and each player $i \in \{b, s\}$ realizes a private type $\theta_i \in [\underline{\theta}, \overline{\theta}] \subseteq \mathbb{R}_+$ that indicates his or her valuation for the object.[8] Let $\boldsymbol{\theta} = (\theta_b, \theta_s)$ and $\Theta = [\underline{\theta}, \overline{\theta}]^2$. We write $\tilde{\boldsymbol{\theta}}$ for the random variable whose realization is $\boldsymbol{\theta}$. Each player's type is independently distributed according to an atomless cumulative distribution function F_i with full support on Θ_i, with a probability density function that is uniformly bounded away from zero.[9]

The timing in each period is as follows. First, players learn their private types. Second, they simultaneously make public announcements. Third, they transact according to the stage mechanism specified for that period, simultaneously trading the object (if called for) and receiving their appropriate monetary payments. Finally, the owner of the object (after trade) consumes it.

Each player has quasilinear period utility of the form $u_i(q_i, t_i, \theta_i) = \theta_i q_i + t_i$, where $q_i \in [0, 1]$ gives the probability that player i will own the object and $t_i \in \mathbb{R}$ represents a monetary payment to player i.

By the revelation principle, it is without loss of generality to restrict attention to direct revelation mechanisms. In a direct revelation mechanism, after learning their types, the players simultaneously announce their types as $\hat{\boldsymbol{\theta}} = (\hat{\theta}_b, \hat{\theta}_s) \in \Theta$, and then the stage mechanism specifies a probability distribution over which player receives the object, a monetary award for each player, and a continuation reward for each player that summarizes the utilities he or she will receive in all future stages. Let $\mathscr{Q}$ be the set of functions mapping Θ to $\{(x, y) \in [0, 1]^2 : x + y = 1\}$, let $\mathscr{T}$ be the set of functions mapping Θ to $\mathbb{R}^2$, and let $\mathscr{W}$ be the set of functions mapping Θ to $\mathbb{R}^2$. Then a stage mechanism g is written as $g = \langle q, t, w \rangle \in \mathscr{Q} \times \mathscr{T} \times \mathscr{W}$; that is, q is the "allocation rule," t is the "monetary payment function," and w is the "continuation reward function."

The efficient allocation rule in this context is essentially unique; i.e., whenever the traders do not value the object equally, the rule must assign the object to the trader with the higher valuation. For the purposes of this paper the allocation in case of a tie is unimportant, so we use $q_b^*(\boldsymbol{\theta}) \equiv \mathbf{1}\{\theta_b \geq \theta_s\}$ as "the" efficient allocation rule (where $\mathbf{1}\{\cdot\}$ is the indicator function).

We use the following shorthand notation to represent the ex post gains from efficient trade:

$$v^{\mathrm{GET}}(\boldsymbol{\theta}) \equiv (\theta_b - \theta_s) q_b^*(\boldsymbol{\theta}).$$

Another useful quantity is the expectation of joint utility under an efficient allocation:

$$v^{\mathrm{ETU}} \equiv \mathbb{E}[\max_i \tilde{\theta}_i].$$

[8]When referring to one of these extreme valuations as the realization of player i's valuation, we use the notation $\overline{\theta}_i \equiv \overline{\theta}$ or $\underline{\theta}_i \equiv \underline{\theta}$ to avoid confusion about which player realized the extreme valuation.

[9]We assume throughout that all relevant functions are measurable with respect to the measure associated with this distribution.

Note that $v^{\mathrm{GET}}(\boldsymbol{\theta})$ is defined at the ex post phase as a function of $\boldsymbol{\theta}$, while v^{ETU} is defined at the ex ante phase as a constant.

2.2 *Recursive mechanisms*

In the infinitely repeated setting the players share a common discount factor $\delta \in (0,1)$. Their valuations are i.i.d. from one period to the next. Each trader's recursive utility is the average discounted value of his utility in the stage game combined with the continuation reward:

$$U_i(\boldsymbol{\theta}, g) \equiv (1-\delta)\, u_i(q(\boldsymbol{\theta}), t(\boldsymbol{\theta}), \theta_i) + \delta\, w_i(\boldsymbol{\theta}).$$

Let v_b and v_s give the expected average continuation utilities at the start of some period; then quasilinearity and feasibility imply that $v_b + v_s \leq v^{\mathrm{ETU}}$. Individual rationality (defined below) further requires that $v_b \geq 0$ and $v_s \geq \mathbb{E}[\tilde{\theta}_s]$, so we write the set of feasible and individually rational expected continuation rewards as

$$\mathscr{V} = \left\{ \mathbf{v} = (v_s, v_b) \in \mathbb{R}^2_+ : v_b + v_s \leq v^{\mathrm{ETU}} \text{ and } \mathbb{E}[\tilde{\theta}_s] \leq v_s \right\}.$$

We focus attention on the class of pure strategy perfect public equilibria. A recursive mechanism can be used to specify the equilibrium path of a perfect public equilibrium (PPE), while an off-equilibrium threat of a trigger punishment supports the path. Miller (2007a) provides a general proof that any recursive mechanism satisfying interim IC, ex post IR, and a notion of budget balance appropriate to the game corresponds to a PPE; and any ceiling on the payoffs of recursive mechanisms also applies to PPE.[10] A recursive mechanism specifies which stage mechanism should be used in a given period as a function of the "promised utility" from the previous period.

DEFINITION 1. Given δ, a *recursive mechanism* is a triplet $\langle V, \gamma, \mathbf{v}^0 \rangle$ such that $V \subset \mathscr{V}$, $\gamma : V \to \mathscr{Q} \times \mathscr{T} \times \mathscr{W}$ with $\gamma(\mathbf{v}) = \langle q(\cdot;\mathbf{v}), t(\cdot;\mathbf{v}), w(\cdot;\mathbf{v}) \rangle$, $\mathbf{v}^0 \in V$, and the following conditions are satisfied.

(i) *Promise keeping*: For all i and all $\mathbf{v}$, $\mathbb{E}\left[U_i(\tilde{\boldsymbol{\theta}}, \gamma(\mathbf{v}))\right] = v_i$.

(ii) *Coherence*: For all $\boldsymbol{\theta} \in \Theta$ and all $\mathbf{v} \in V$, $w(\boldsymbol{\theta}; \mathbf{v}) \in V$.

The set V contains all payoffs that can be attained in the mechanism, and $\mathbf{v}^0$ is the initial payoff.[11] Since V and $\mathbf{v}^0$ are often implicit in the definition of γ, we refer to γ also as a recursive mechanism.

---

[10]Although restricting attention to perfect public equilibria is not always without loss of generality with respect to equilibrium payoffs, Fudenberg and Levine (1994, Theorem 5.2) implies that in our setting (in particular, because θ_b and θ_s are independent), any sequential equilibrium payoff can be supported by a perfect public equilibrium. Even if θ_b and θ_s are not independent, there is no known way for equilibria in non-public strategies to outperform perfect public equilibria. In particular, the critique of Kandori and Obara (2006) does not apply because no player, by altering his own announcement, can change the probability distribution of the other player's announcement.

[11]There are a couple of important differences between V and the equilibrium set defined by Abreu, Pearce, and Stacchetti (1990). First, we have not yet imposed incentive constraints; thus, without further qualifications, the payoffs in V need not be supportable in equilibrium. In addition, when we do impose

Promise keeping implies that the stage mechanism in each period delivers, in expectation, utility equal to the continuation rewards promised at the end of the previous period. A recursive mechanism must specify coherent continuation rewards. Given the promised utility $\mathbf{v}$, the function $w(\cdot;\mathbf{v})$ gives the expected average continuation rewards for the subsequent period as a function of the current period's reported types.

In a recursive mechanism of this kind, there are two means of transferring utility: immediate monetary payments and changes in continuation rewards. Changes in continuation rewards can be delivered by adjusting the stage mechanisms that are used in future periods as a function of what happens in the current period. Though some of the cases we examine do not employ changes in continuation rewards, in general it is the availability of anticipated future utility that enables us to design recursive mechanisms that satisfy not only the relatively weak properties of an ordinary PPE, but also some stronger properties that are desirable in weaker institutional settings. The properties we employ are interpreted in Section 1.1; here, we formally define them.

Definition 2. A recursive mechanism γ satisfies *ex post incentive compatibility (IC)* if

$$\theta_i \in \arg\max_{\hat{\theta}_i}(1-\delta)\, u_i(q(\hat{\theta}_i,\theta_{-i};\mathbf{v}), t(\hat{\theta}_i,\theta_{-i};\mathbf{v}),\theta_i)+\delta\, w_i(\hat{\theta}_i,\theta_{-i};\mathbf{v})$$

for all $\boldsymbol{\theta}\in\Theta$, all $i\in\{b,s\}$, and all $\mathbf{v}\in V$.

Definition 3. A recursive mechanism γ satisfies *ex post individual rationality (IR)* if

$$U_b(\boldsymbol{\theta},\gamma(\mathbf{v}))\geq 0$$
$$U_s(\boldsymbol{\theta},\gamma(\mathbf{v}))\geq (1-\delta)\,\theta_s+\delta\mathbb{E}[\tilde{\theta}_s],$$

for all $\boldsymbol{\theta}\in\Theta$ and all $\mathbf{v}\in V(\delta,\gamma)$.

Definition 4. A recursive mechanism γ satisfies *ex post budget balance (BB)* if $t_b(\boldsymbol{\theta};\mathbf{v}) + t_s(\boldsymbol{\theta};\mathbf{v})=0$ for all $\boldsymbol{\theta}\in\Theta$ and all $\mathbf{v}\in V(\delta,\gamma)$. It satisfies *ex ante BB* if $\mathbb{E}[t_b(\tilde{\boldsymbol{\theta}};\mathbf{v})+t_s(\boldsymbol{\theta};\mathbf{v})]=0$ for all $\mathbf{v}\in V(\delta,\gamma)$.

In Sections 3 and 4, we consider mechanisms that do not employ changes in continuation rewards, which we call "stationary."

Definition 5. A recursive mechanism γ is *stationary* if $V(\delta,\gamma)$ is a singleton.

2.3 *Consequences of incentive compatibility*

This subsection presents some standard consequences of ex post and interim IC. These consequences are used throughout the paper. The following lemma characterizes mechanisms with deterministic, nondecreasing allocation rules, including the efficient mechanism.[12]

incentive constraints, we include individual rationality constraints that are specified directly in terms of agents' "outside options," which in our model are the payoffs if the relationship ends. We do not include these outside options in V, since in a mechanism that respects individual rationality constraints, the outside options are never realized.

[12]The results extend to stochastic mechanisms (i.e., mechanisms that allow $q\in(0,1)$) in a natural way.

LEMMA 1. *Consider a stage mechanism $g = \langle q, t, w\rangle \in \mathscr{Q} \times \mathscr{T} \times \mathscr{W}$, where there exists a non-decreasing function $\lambda : \mathbb{R} \to \mathbb{R}$ such that $q_b(\boldsymbol{\theta}) = 1$ if and only if $\theta_b \geq \lambda(\theta_s)$, and $q_b(\boldsymbol{\theta}) = 0$ otherwise. (Since the inverse of λ may not be a function when λ is not strictly increasing, define $\lambda^{-1}(\theta_b) \equiv \sup\{\theta_s : \theta_b \geq \lambda(\theta_s)\}$.) Then, g satisfies ex post IC if and only if there exist functions $h_b : \Theta_s \to \mathbb{R}$ and $h_s : \Theta_b \to \mathbb{R}$ such that, for all $\boldsymbol{\theta}$,*

$$t_b(\boldsymbol{\theta}) + \frac{\delta}{1-\delta} w_b(\boldsymbol{\theta}) = -q_b(\boldsymbol{\theta}) \cdot \lambda(\theta_s) + h_b(\theta_s) \tag{1}$$

$$t_s(\boldsymbol{\theta}) + \frac{\delta}{1-\delta} w_s(\boldsymbol{\theta}) = q_b(\boldsymbol{\theta}) \cdot \lambda^{-1}(\theta_b) + h_s(\theta_b). \tag{2}$$

PROOF. Let us begin with ex post IC. As is standard, the envelope theorem implies that if g satisfies ex post IC, then wherever the derivative exists (which is almost everywhere),

$$\begin{aligned}\frac{\partial U_i(\boldsymbol{\theta}, g)}{\partial \theta_i} &= \frac{\partial}{\partial \theta_i}\left[(1-\delta)\left(u_i(q(\hat{\theta}_i, \theta_{-i}), t(\hat{\theta}_i, \theta_{-i}), \theta_i)\right) + \delta w_i(\hat{\theta}_i, \theta_{-i})\right]\Big|_{\hat{\theta}_i = \theta_i} \\ &= (1-\delta) q_i(\boldsymbol{\theta}).\end{aligned}$$

This in turn implies, by the fundamental theorem of calculus, that

$$\begin{aligned}\frac{1}{1-\delta} U_i(\boldsymbol{\theta}, g) &\equiv q_i(\boldsymbol{\theta}) \cdot \theta_i + t_i(\boldsymbol{\theta}) + \frac{\delta}{1-\delta} w_i(\boldsymbol{\theta}) \\ &= \frac{1}{1-\delta} U_i(\underline{\theta}_i, \theta_{-i}, g) + \int_{\underline{\theta}_i}^{\theta_i} q_i(s, \theta_{-i})\, ds.\end{aligned} \tag{3}$$

For each trader, substituting $h_i(\theta_{-i})$ for $U_i(\underline{\theta}_i, \theta_{-i}, g)/(1-\delta)$ and substituting in for the form of q_i given in the proposition yields

$$\frac{1}{1-\delta} U_b(\boldsymbol{\theta}, g) = h_b(\theta_s) + \int_{\underline{\theta}_b}^{\theta_b} \mathbf{1}\{b \geq \lambda(\theta_s)\}\, db$$

$$\frac{1}{1-\delta} U_s(\boldsymbol{\theta}, g) = h_s(\theta_b) + \int_{\underline{\theta}_s}^{\theta_s} \mathbf{1}\{s > \lambda^{-1}(\theta_b)\}\, ds,$$

which simplify to (1) and (2). It is straightforward to verify that any mechanism that satisfies (3) satisfies ex post IC. □

An efficient mechanism that satisfies (1) and (2) is known as a Vickrey–Clarke–Groves (VCG; Clarke 1971; Groves 1973; Vickrey 1961) mechanism. A VCG mechanism thus implements the payments $t_b(\boldsymbol{\theta}) = -\theta_s q_b^*(\boldsymbol{\theta}) + h_b(\theta_s)$ and $t_s(\boldsymbol{\theta}) = \theta_b q_b^*(\boldsymbol{\theta}) + h_s(\theta_b)$, where h_i is an arbitrary function that depends only on θ_{-i}. In words, each trader receives the full social surplus of the transaction, plus a fixed amount that depends only on the other trader's type.

3. Equilibria with ex ante budget balance

Ex ante BB corresponds to a situation in which unbounded insurance is available to absorb budget imbalances in the trading relationship. Given such insurance, we construct a stationary recursive mechanism that achieves efficiency with ex post IR and ex post IC for $\delta \geq \frac{1}{2}$. The proof adapts the logic of the Myerson–Satterthwaite theorem to a situation in which extra surplus is available because of the prospect of a future trading relationship. Intuitively, the parties are willing to pay additional fees that balance the budget in expectation, because by failing to pay they would lose the future surplus.

Proposition 1. *There exists a stationary recursive mechanism that satisfies efficiency, ex ante BB, ex post IR, and ex post IC if and only if $\delta \geq \frac{1}{2}$.*

Proof. For a stationary, efficient mechanism, $V(\delta,\gamma)$ is a singleton—call it (v_b, v_s)—which satisfies $v_b + v_s = v^{\text{ETU}}$, and there is a single stage mechanism $\langle q, t, w\rangle$ used in every period. Furthermore, $w_b(\boldsymbol{\theta}) = v_b$ and $w_s(\boldsymbol{\theta}) = v_s$ for all $\boldsymbol{\theta} \in \Theta$.

Consider first sufficiency. We construct a stage mechanism as follows.

1. The allocation is efficient, e.g. $q_b \equiv q_b^*$.

2. The expected future gains from trade are split equally: for all $\boldsymbol{\theta} \in \Theta$,

$$w_b(\boldsymbol{\theta}) = v_b = \frac{1}{2}\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})]$$
$$w_s(\boldsymbol{\theta}) = v_s = \frac{1}{2}\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})] + \mathbb{E}[\tilde{\theta}_s].$$

3. Monetary payments are as follows:

$$t_b(\boldsymbol{\theta}) = -\theta_s q_b^*(\boldsymbol{\theta}) - \frac{1}{2}\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})]$$
$$t_s(\boldsymbol{\theta}) = \theta_b q_b^*(\boldsymbol{\theta}) - \frac{1}{2}\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})].$$

For each player, the first term in the monetary payment is his or her externality on the other player, and the second term is fixed with respect to his or her announcement. Thus the total transfers (monetary payments plus future surplus, which is fixed) take the form of a VCG mechanism, so efficiency and ex post IC are satisfied by construction.

Utilities in each period can be written as follows:

$$\begin{aligned} u_b(q^*(\boldsymbol{\theta}), t(\boldsymbol{\theta}), \theta_b) &= (\theta_b - \theta_s) q_b^*(\boldsymbol{\theta}) - \frac{1}{2}\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})] \\ &= v^{\text{GET}}(\boldsymbol{\theta}) - \frac{1}{2}\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})] \\ u_s(q^*(\boldsymbol{\theta}), t(\boldsymbol{\theta}), \theta_s) &= (1 - q_b^*(\boldsymbol{\theta}))\theta_s + \theta_b q_b^*(\boldsymbol{\theta}) - \frac{1}{2}\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})] \\ &= \theta_s + v^{\text{GET}}(\boldsymbol{\theta}) - \frac{1}{2}\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})]. \end{aligned}$$

Thus the promise-keeping constraint is satisfied, because in expectation each player receives $\frac{1}{2}\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})]$ plus his or her reservation utility, which is exactly v_i.

Ex ante BB is satisfied because for each $\boldsymbol{\theta}$ the externality terms of the monetary payments sum to the ex post gains from trade $v^{\text{GET}}(\boldsymbol{\theta})$; ex ante these are balanced in expectation by the fixed terms of the monetary payments. In other words, the gains from trade are given out twice through the externality payments and efficient ownership of the object, while they are paid back once through the fixed fees to balance the budget.

Finally, ex post IR requires

$$\frac{U_b(\boldsymbol{\theta},g)}{1-\delta}=(\theta_b-\theta_s)q_b^*(\boldsymbol{\theta})+\frac{1}{2}\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})]\left(\frac{2\delta-1}{1-\delta}\right)\geq 0$$
$$\frac{U_s(\boldsymbol{\theta},g)}{1-\delta}=\theta_s+(\theta_b-\theta_s)q_b^*(\boldsymbol{\theta})+\frac{1}{2}\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})]\left(\frac{2\delta-1}{1-\delta}\right)+\frac{\delta}{1-\delta}\mathbb{E}[\tilde{\theta}_s]\geq\theta_s+\frac{\delta}{1-\delta}\mathbb{E}[\tilde{\theta}_s]$$

for all $\boldsymbol{\theta}$. The worst case realization of types for both constraints is $(\underline{\theta}_b,\overline{\theta}_s)$, in which case the constraints are satisfied if and only if $\delta\geq\frac{1}{2}$.

Now consider whether there exist alternative transfers satisfying ex ante BB, ex post IR and ex post IC for $\delta<\frac{1}{2}$. By Lemma 1, any such alternative can be obtained from the proposed transfers by adding a function $h_b(\theta_s)$ to $t_b(\boldsymbol{\theta})$ and adding a function $h_s(\theta_b)$ to $t_s(\boldsymbol{\theta})$. For ex ante BB, if such $h_b(\theta_s)$ is ever strictly positive for a non-zero measure subset of Θ_s, then, to compensate, either $h_b(\theta_s)$ or $h_s(\theta_b)$ must be strictly negative for a nonzero measure subset of Θ_s or Θ_b (respectively). However, if $h_b(\theta_s)<0$ is ever realized, the ex post IR constraint for $\underline{\theta}_b$ will be violated for all $\delta\leq\frac{1}{2}$, and similarly if $h_s(\theta_b)<0$ is ever realized the ex post IR constraint for $\overline{\theta}_s$ will be violated for all $\delta\leq\frac{1}{2}$. □

4. Equilibria with ex post budget balance

In this section, we analyze second-best efficiency for a trading environment with the weakest institutions, under which the players cannot insure against budget imbalances. We show that the optimal mechanism under these circumstances is usually simple, stationary, and highly inefficient—regardless of the discount factor. But first, Miller (2007b) establishes an impossibility result for efficient trade with ex post BB and ex post IC.

Proposition 2. *There exists $\varepsilon>0$ such that $v_s+v_b<v^{ETU}-\varepsilon$ for all $\delta\in[0,1)$, for every recursive mechanism γ that satisfies ex post BB and ex post IC, and for all $\mathbf{v}\in V(\delta,\gamma)$.*

To understand this result, note that ex post BB requires that current monetary payments balance ex post, and efficiency in future periods requires that transfers of future surplus must balance ex post as well.[13] However, ex post IC means that an efficient mechanism must give each player a transfer (including future surplus) that is sensitive to his or her announcement only through the term that represents his or her realized externality on the other player. Hence the sum of transfers includes a term that is not

[13]The result holds also if ex post BB is relaxed to allow for free disposal. Free disposal of money has an effect similar to non-stationarity, which we consider below. Free disposal of the object is not helpful, since to obtain near-efficiency the object would need to be retained arbitrarily often.

additively separable in types (the gains from trade) as well as terms that are additively separable (recall the expressions for transfers under ex post IC, (1–2)). Thus an inconsistency arises, and efficient trade is impossible.

Hagerty and Rogerson (1987) show that in a static trade setting, ex post BB, ex post IR, and ex post IC can be attained only by "posted price mechanisms"—mechanisms of the form $q(\boldsymbol{\theta}) = \mathbf{1}\{\theta_b \geq \lambda(\theta_s)\}$, where[14]

$$\lambda(\theta_s) = \begin{cases} \overline{p} & \text{if } \theta_s \leq \overline{p} \\ \infty & \text{otherwise.} \end{cases}$$

"Posted price mechanisms" are so named because they can be implemented indirectly by the following procedure: a price $\overline{p}$ is posted prior to the players learning their types, and then after they learn their types privately the buyer and seller each makes an independent decision about whether to trade at that price. Only if both desire to trade does trade occur; if trade occurs, the buyer pays the seller $\overline{p}$, and otherwise there is no payment. Thus, payments are budget-balanced ex post, and it is not necessary to use variation in future continuation rewards to provide incentives. Posted price mechanisms satisfy ex post IR for all $\delta \geq 0$, because the buyer never pays more than his valuation if he receives the object, while the seller never receives less than her valuation if she gives it up, and neither makes any payment when the object is not traded. Hence under ex post BB the optimal *stationary* recursive mechanism is no better than the optimal static mechanism.

However, if we allow continuation rewards to vary, it may be possible for a nonstationary recursive mechanism to dominate the optimal stationary mechanism by offering greater efficiency in the first period while promising lower continuation rewards following some realizations. Since monetary payments and changes in continuation rewards are perfect substitutes in terms of providing incentives (see Lemma 1), the problem of designing the first period of an optimal mechanism under ex post BB boils down to designing an optimal static mechanism under free disposal:

$$\max_{\langle q,t,w\rangle} \mathbb{E}\left[q_b(\tilde{\boldsymbol{\theta}})(\tilde{\theta}_b - \tilde{\theta}_s) + \frac{\delta}{1-\delta}\left(w_b(\tilde{\boldsymbol{\theta}}) + w_s(\tilde{\boldsymbol{\theta}})\right)\right] \tag{4}$$

s.t. ex post IR, ex post IC, and

$$t_b(\boldsymbol{\theta}) + t_s(\boldsymbol{\theta}) + \frac{\delta}{1-\delta}\left(w_b(\boldsymbol{\theta}) + w_s(\boldsymbol{\theta})\right) \leq \frac{\delta}{1-\delta} \max_{\boldsymbol{\theta}'} \left[w_b(\boldsymbol{\theta}') + w_s(\boldsymbol{\theta}')\right] \quad \forall\, \boldsymbol{\theta}. \tag{5}$$

This problem can be solved using linear programming. For a more detailed exposition and proof, see Miller (2007b).

The tradeoff between inefficiency in the current allocation and inefficient continuation rewards is difficult to resolve analytically, because when considering allocation rules q and continuation rewards there is a great deal of flexibility in selecting the corresponding functions h_b and h_s. We have not been able to derive a general formula

[14] Hagerty and Rogerson allow the price to be randomized, but if a randomized price is ever optimal then there must exist also an optimal deterministic price.

for how alternative choices of the allocation rule map into expected loss of future surplus. In Athey and Miller (2007), we present the results of applying numerical linear programming to solve (4–5), for a variety of probability distributions. Our main numerical finding is quite striking: the optimal mechanism is often a stationary, posted-price mechanism. Further, even when a posted price mechanism is not optimal, the optimal non-stationary mechanism improves on the best posted price mechanism by very little.[15]

Since posted-price mechanisms satisfy ex post IR for all δ, when a posted price mechanism is optimal, even as $\delta \to 1$ the optimal mechanism in this institutional environment is the same as the optimal mechanism in a one-shot trading relationship with the same institutions. This result provides some justification for the use of simple posted price schemes in practice, even in ongoing trading relationships. The following example illustrates the properties of an optimal mechanism.

EXAMPLE 1. For the case where each player's value is uniformly distributed on the unit interval, the optimal scheme is stationary, and the optimal posted price is $\overline{p} = \frac{1}{2}$. This provides stationary values of $v_b^{\mathrm{SB}} = \mathbb{E}[U_b(\tilde{\boldsymbol{\theta}}, g^{\mathrm{SB}})] = \frac{1}{16}$ and $v_s^{\mathrm{SB}} = \mathbb{E}[U_s(\tilde{\boldsymbol{\theta}}, g^{\mathrm{SB}})] = \frac{9}{16}$, implying that $w_b^{\mathrm{SB}}(\boldsymbol{\theta}) = \frac{1}{16}$ and $w_s^{\mathrm{SB}}(\boldsymbol{\theta}) = \frac{9}{16}$ for all $\boldsymbol{\theta}$. In contrast, the expected social surplus from efficient trade is $\frac{2}{3}$. ◊

In sum, we have illustrated that under the weakest institutional environment (one that ensures only simultaneous exchange of money and the object at the ex post phase), the optimal equilibrium is inefficient. However, it often takes a familiar, simple form that is commonly observed in practice.

5. EQUILIBRIA WITH BOUNDED BUDGET ACCOUNT

In the case of ex ante BB, the only role played by the future repetitions of the stage game is to provide enough extra surplus to relax the IR constraints, while in the case of ex post BB, future repetitions optimally played no role at all in many examples. Complementing these results, Proposition 2 shows that no efficient recursive mechanism, stationary or otherwise, can achieve ex post BB and ex post IC. Even so, in this section we show that the capabilities of nonstationary recursive mechanisms occupy a useful territory that extends beyond the limits of what stationary recursive mechanisms can achieve. In this territory, a nonstationary recursive mechanism can achieve ex post IC, ex post IR, and either exact or approximate efficiency (depending on the probability distribution over $\tilde{\boldsymbol{\theta}}$ and the discount factor), if the traders jointly have access to a collateralized credit account and possess sufficient collateral.

[15]We do not have a general characterization of the conditions under which a posted price mechanism is optimal. From examining our numerical results, it appears that nonstationary schemes become optimal when some regions of types are relatively unlikely, and it is not too costly to bear some future surplus loss in those regions in order to provide incentives for efficient trade in regions that are more likely. It should be noted that for cases where we find that a non-stationary mechanism is optimal, additional steps are required to derive the full recursive mechanism.

The institutional environment that motivates our modeling of the joint account can be described as follows. The traders hold durable consumption goods that can be used as collateral. An institution (the "bank") offers them a credit line account from which they can jointly withdraw (i.e., borrow) up to the value of their collateral; the bank shares the same discount factor as the players.[16] The traders can also deposit funds into the account, up to some finite limit. The interest rate on the account is zero, regardless of the balance.[17] For simplicity of exposition, we normalize the balance to zero when the players have withdrawn the maximum amount. Using this normalization, if $\bar{A}$ is the value of the collateral, a balance of $\bar{A}$ corresponds to the initial balance of the credit line prior to any withdrawals or deposits.[18] We focus on mechanisms that yield zero expected profits to the bank.

It is worth highlighting that we rely heavily on the assumption of non-cooperative play. In particular, we rule out joint deviations by the players. They cannot collude against the bank in order to take the maximum possible withdrawal in each period, and keep the account balance permanently at its lowest level. Vulnerability to collusion by agents is a pervasive feature of the mechanism design literature, and renegotiation among agents is commonly assumed away. However, it is important to remember that these assumptions may be strong in practice.

To model the credit line account we introduce a new state variable, $A \in \mathbb{R}$, which tracks the current cash balance of the account. Any imbalances, positive or negative, in the sum of payments to the players result in (respectively) a deposit into or a withdrawal from the account. We also enlarge the set of feasible and ex ante IR promised utility vectors, since the traders can increase their joint promised utility beyond v^{ETU} by withdrawing funds from the account:

$$\mathscr{V}^{+} = \left\{(v_s, v_b) \in \mathbb{R}^2_+ : v_s \geq \mathbb{E}[\tilde{\theta}_s]\right\}.$$

We now define a recursive mechanism that takes both v and A as state variables.

Definition 6. Given δ, an *account-recursive mechanism* is a triplet $\langle V_a, \gamma_a, \mathbf{v}^0 \rangle$, such that $V_a \subset \mathbb{R} \times \mathscr{V}^+$, $\gamma_a : \mathbb{R} \times \mathscr{V}^+ \to \mathscr{Q} \times \mathscr{T} \times \mathscr{W}$ with $\gamma_a(A, \mathbf{v}) = \langle q(\cdot; A, \mathbf{v}), t(\cdot; A, \mathbf{v}), w(\cdot; A, \mathbf{v}) \rangle$, $\mathbf{v}^0 \in V_a$, and the following are satisfied.

[16]Although this assumption is not used directly in our analysis, we state it to avoid some unattractive possibilities in the interpretation of the mechanism. For example, with different discount factors the bank and the agents might like to make intertemporal trades.

[17]Our analysis could be extended to allow for non-zero interest rates as well, and the modifications would be fairly straightforward for a relatively low interest rate that is the same for balances above and below the initial balance. The mechanisms we construct below have regions of (high) account balances where players withdraw money on average and (low) regions where they deposit money on average; with a positive interest rate, the players would increase these deposits and withdrawals by particular constants. The discount factors required to support the mechanisms would generally be higher, and the required range of account balances larger.

[18]The bank should never confiscate the collateral, since no payments are required on a zero-interest credit line. Our discussion of collateral is a matter of interpretation; if somehow the bank could observe that the trading relationship ended, the collateral would protect the bank. Formally, in our model collateral pins down the "initial balance" of the account.

(i) *Promise keeping*: $\mathbb{E}[U_i(\tilde{\boldsymbol{\theta}}, \gamma_a(A, \mathbf{v}))] = v_i$ for all i and all $(A, \mathbf{v}) \in V_a$.

(ii) *Coherence*: Letting

$$\alpha(\boldsymbol{\theta}; A, \mathbf{v}) \equiv A - \left(t_s(\boldsymbol{\theta}; A, \mathbf{v}) + t_b(\boldsymbol{\theta}; A, \mathbf{v})\right),$$

$\left(\alpha(\boldsymbol{\theta}; A, \mathbf{v}), w(\boldsymbol{\theta}; A, \mathbf{v})\right) \in V_a$ for all $(A, \mathbf{v}) \in V_a$ and all $\boldsymbol{\theta} \in \Theta$.

(iii) *Account keeping*: For all $(A, \mathbf{v}) \in V_a$,

$$v_b + v_s = (1-\delta)\mathbb{E}\left[\sum_{\tau=1}^{\infty} \delta^{\tau-1}\left(\sum_{i=b,s} q_i\left(\tilde{\boldsymbol{\theta}}^{(\tau)}; \tilde{A}^{(\tau)}, \tilde{\mathbf{v}}^{(\tau)}\right)\tilde{\theta}_i^{(\tau)} + \tilde{A}^{(\tau)} - \tilde{A}^{(\tau+1)}\right)\right],$$

where $\tilde{\boldsymbol{\theta}}^{(\tau)}$ is the instance of the random variable $\tilde{\boldsymbol{\theta}}$ in the τth period, $\tilde{A}^{(1)} \equiv A$, $\tilde{\mathbf{v}}^{(1)} \equiv \mathbf{v}$, $\tilde{A}^{(\tau+1)} \equiv \alpha(\tilde{\boldsymbol{\theta}}^{(\tau)}; \tilde{A}^{(\tau)}, \tilde{\mathbf{v}}^{(\tau)})$, $\tilde{\mathbf{v}}^{(\tau+1)} \equiv w(\tilde{\boldsymbol{\theta}}^{(\tau)}; \tilde{A}^{(\tau)}, \tilde{\mathbf{v}}^{(\tau)})$, and the expectation is taken at time zero over all $\tilde{\boldsymbol{\theta}}^{(1)}, \tilde{\boldsymbol{\theta}}^{(2)}, \ldots$.

For shorthand, we refer to the function γ_a as an account-recursive mechanism, since V_a is often implicit. We address the selection of $\mathbf{v}^0$ later. Note that since there is no upper bound on $\mathscr{V}^+$, a condition such as account keeping is needed to ensure that the mechanism does not promise ever-increasing levels of utility without eventually providing that utility through either allocative payoffs or withdrawals from the account.

DEFINITION 7. An account-recursive mechanism γ_a satisfies *bounded budget account* (BBA) if there exists $B \in \mathbb{R}_+$ such that $\{A \in \mathbb{R} : (A, \mathbf{v}) \in V_a\} \subseteq [0, B]$.

The following helpful lemma allows us to avoid demonstrating that account keeping holds for BBA mechanisms that promise bounded per-period utility. The proof is in Section A.1 (in the Appendix).

LEMMA 2. *Suppose a triplet* $\langle V_a, \gamma_a, \mathbf{v}^0\rangle$ *satisfies promise keeping, coherence, and BBA. If* $\sup\{v_b + v_s : \exists\, A \text{ s.t. } (A, (v_b, v_s)) \in V_a\} < \infty$, *then* γ_a *is an account-recursive mechanism.*

For the remainder of this section, we often refer to account-recursive mechanisms simply as "mechanisms," and we use the term "value correspondence" for the correspondence that associates each account balance with a set of promised utility vectors.

5.1 *Efficient account-recursive mechanisms*

An initial, benchmark result shows that it is impossible to design a *stationary* mechanism (an account-recursive mechanism that uses the same stage mechanism in every period) that satisfies efficiency, ex post IC, ex post IR, and BBA, because there exist series of realizations of types that either deplete the account or send it over the upper bound.

PROPOSITION 3. *There does not exist a stationary account-recursive mechanism that satisfies BBA, ex post IR, and ex post IC, and allocates efficiently in every period.*

PROOF. We established above (using (1–2)) that any mechanism satisfying ex post IC and efficiency must have either imbalances of monetary payments or changes in future joint surplus, but a stationary mechanism cannot have changes in future joint surplus. Any stage mechanism thus has a realization of $\tilde{\boldsymbol{\theta}}$ for which the imbalance of monetary payments is either strictly positive or strictly negative. Since the mechanism is stationary, starting from any account balance, a long enough series repeating this same realization sends the account balance past either the upper or lower bound. □

A natural follow-on question concerns whether it is possible to design a nonstationary mechanism that has the desired properties. We begin by considering a class of mechanisms that allocates efficiently in every period, and where the sum of player promised utilities takes a low value for low account balances, and a high value for high account balances, with a discontinuous jump at some intermediate account balance. Restricting the sum of promised utilities to be constant except at discontinuous jump points greatly simplifies the construction of a mechanism, making it possible to specify the form of the mechanism analytically. We show, however, that our mechanism requires restrictions on the probability distribution over $\tilde{\boldsymbol{\theta}}$ and relatively high patience for many such distributions. Although more complex schemes could probably be constructed to achieve efficiency for a wider range of distributions, our mechanism illustrates the qualitative features such schemes should have, and it further highlights the important role of patience in achieving efficiency.

To begin, we state a condition that places restrictions on the discount factor jointly with the probability distribution over $\tilde{\boldsymbol{\theta}}$.

CONDITION 1. There exist real-valued functions h_b and h_s such that

$$h_b(\theta_s) \geq 0 \text{ for all } \theta_s \in \Theta_s,\ h_s(\theta_b) \geq \frac{\delta}{1-\delta}\mathbb{E}\big[\tilde{\theta}_s\big] \text{ for all } \theta_b \in \Theta_b \tag{6}$$

$$v^{GFT}(\boldsymbol{\theta}) + h_b(\theta_s) + h_s(\theta_b) \leq \delta\mathbb{E}\big[\tilde{\theta}_s + 2v^{GFT}(\tilde{\boldsymbol{\theta}}) + h_b(\tilde{\theta}_s) + h_s(\tilde{\theta}_b)\big] \text{ for all } \boldsymbol{\theta} \in \Theta. \tag{7}$$

Lemma 4, in Section A.2, shows that Condition 1 is satisfied if and only if it is possible to find a stage mechanism with the following properties: (i) continuation rewards are equal to the promised utility from the stage mechanism for all type realizations, (ii) ex post IR and IC hold, and (iii) the traders can make a deposit but not take a withdrawal. We see that the condition can be satisfied only if δ is sufficiently high. Furthermore, for some distributions over $\tilde{\boldsymbol{\theta}}$, the condition fails for all δ, while for others the condition holds for sufficiently high δ. We delay further interpretation of Condition 1 until after we describe a class of efficient mechanisms that exists if Condition 1 holds.

PROPOSITION 4. *If Condition 1 holds, there exists an account-recursive mechanism that satisfies ex post IR, ex post IC, and BBA, and such that efficient trade is implemented in every period.*

The proof, in Section A.2, constructs a mechanism that is characterized by two regimes, each of which implements efficient trade. Which regime is active depends on the account balance. For $A \in [0, A_{dp}]$, the "deposit" regime is active, and for $A \in (A_{dp}, A_{wd}]$, the "withdrawal" regime is active. The numbers A_{dp} and A_{wd} are chosen such that the account balance is always within the interval $[0, A_{wd}]$. In the deposit regime, the monetary payments satisfy $t_b(\boldsymbol{\theta}) + t_s(\boldsymbol{\theta}) \leq 0$, so that the players never take withdrawals from the account. In the withdrawal regime, the monetary payments satisfy satisfy $t_b(\boldsymbol{\theta}) + t_s(\boldsymbol{\theta}) \geq 0$, so that the players never make deposits into the account. Promised utility in the deposit regime is $\mathbf{v}^{dp}$, where $v_b^{dp} + v_s^{dp} < v^{\mathrm{ETU}}$, while promised utility in the withdrawal regime is $\mathbf{v}^{wd}$, where $v_b^{wd} + v_s^{wd} > v^{\mathrm{ETU}}$.

Within each regime, the players' promised utility does not vary with the account balance, but it does differ across regimes. In particular, promised utility is higher in the withdrawal regime than in the deposit regime. As a result, when the traders switch from the withdrawal regime to the deposit regime, they see a potentially large loss in the present value of their expected future utility, since even a small decrease in promised (per-period) utility can be large in present value terms if they are patient. In order to induce truthful revelation in states of the world where the mechanism would specify switching to the deposit regime, the traders must be compensated through large payments in states that lead to a switch. That in turn implies that a large withdrawal will be taken from the joint account, just as the traders switch to the deposit regime. But to prevent running out of funds in such a case, the critical value of the account, A_{dp}, at which the agents switch to the deposit regime must be large enough that the account contains enough funds to compensate them for the switch. This bound on A_{dp} may increase with agents' patience, so it may take a long time to raise enough funds to increase the balance above A_{dp} again, and a correspondingly long time before they can take withdrawals again.

The mechanism of Proposition 4 does not necessarily deliver zero profits to the bank, and it does not necessarily provide the players efficient expected payoffs (i.e., $\mathbf{v}^0$ has not been determined). In particular, if the traders started from an account balance of zero, they would have to make deposits that they would not expect to recover in the long run. In that case, the bank would earn a corresponding expected profit. However, if the account is interpreted as a collateralized credit line, the players need not expect to lose money to the bank. Let $\bar{A}$ be the average account balance for the ergodic distribution over account balances (which exists and is unique, by Lemma 5, in Section A.4). If the players have collateral worth at least $\bar{A}$, then they can borrow up to this amount from the bank. We then assure the players the full expected benefits of their relationship—and assure the bank zero expected profits—by means of a randomization that takes place before the first period of the mechanism.

COROLLARY 1. *Suppose that Condition 1 holds. Then there exists a recursive mechanism γ_a satisfying the properties in Proposition 4, an initial account balance $\bar{A} \in (0, A_{wd})$, and an initial randomization over deposits and withdrawals (which occurs prior to the first period of the mechanism), such that, taking expectations prior to the first period over both the initial randomization and all possible realizations of play under the mechanism,*

(i) *the expected account balance for each period is exactly $\bar{A}$, and*

(ii) *the expected average joint payoffs are v^{ETU}.*

The proof is in Section A.3. The result follows from choosing the initial account balance equal to the expected account balance at the ergodic distribution, $\bar{A}$, and choosing the initial randomized deposit or withdrawal so as to jump immediately to the ergodic distribution of the mechanism.[19] From a perspective prior to the initial randomization, if the first period account balance has a distribution equal to the ergodic distribution, then all future account balances are also distributed according to the ergodic distribution. Therefore by definition the expected account balance along the path of play is exactly $\bar{A}$ in every period. We take $\bar{A}$ to be the traders' starting balance prior to the initial randomization, so that the bank expects, on average, no change in the account balance in any future period. When the account is interpreted as a collateralized line of credit, this means the players should hold collateral valued at no less than $\bar{A}$.

Since the bank earns zero expected profits, the traders similarly must earn zero expected profits through changes in their account balance, and therefore when computing their expected payoffs in the mechanism we can restrict attention to the utility that they earn directly from trade, which by construction is exactly their expected total profit from efficient trade, v^{ETU}.

Now we return to discuss Condition 1. It is an open question whether Condition 1 is necessary for an efficient BBA mechanism to exist. But we argue that even if Condition 1 is not necessary for an efficient BBA mechanism to exist, BBA is still difficult to satisfy when Condition 1 fails.

In particular, note that Condition 1 is necessary for the existence of an efficient BBA mechanism for which the sum of the traders' values is constant in A over a sufficiently large interval containing $A = 0$, by Lemma 4. So when Condition 1 fails, if an efficient BBA mechanism exists then the sum of the traders' values must increase with the account balance near $A = 0$. That way, in expectation players would be better off in the future than at $A = 0$, and so they would be willing to pay more into the account. Of course, these higher promised values would themselves need to be implemented through stage mechanisms that allowed players to keep more of the gains from trade, but keeping more of the gains from trade in a given period implies that expected increases in the account balance must be smaller (and some withdrawals must occur) so that the expected path would lead away from $A = 0$ more slowly. This limits the slope of the sum of the traders' values.

There are also other important restrictions on the slope of the sum of the traders' values. If a vector of types is realized that requires the traders to take a withdrawal from the account, then they anticipate both the increase in their current payoffs due to the withdrawal and the decrease in their future payoffs due to the reduction in promised utilities associated with low account balances. If a certain transfer of utility is required to induce truthful reporting of this particular realization of types, then the withdrawal

[19] That is, $\mathbf{v}^0$ is determined as an outcome of the initial randomization. To define the mechanism formally to include the initial randomization, we would need to treat $\mathbf{v}^0$ as a random variable.

must be larger than the desired transfer of utility, in order to compensate for the decrease in continuation values. Indeed, if the sum of the traders' promised utilities increases too fast with respect to the account balance throughout an interval near $A=0$, then any withdrawal from the account has a net negative impact on the players in that interval.

For these reasons, we expect that in general the existence of efficient mechanisms requires high discount factors. It also seems likely that there are some probability distributions over $\tilde{\boldsymbol{\theta}}$ (in particular, those for which the maximum gains from trade are large relative to the average gains from trade) for which efficient mechanisms do not exist for any discount factor, because the required payments into the account at $A=0$ are too large. We have not been able to formally prove this conjecture.

One challenge in resolving this question is that the stage mechanisms that satisfy the desired properties for particular account balances depend on the details of the probability distribution over $\tilde{\boldsymbol{\theta}}$. In addition, the set of possible mechanisms is so large and complex that it is difficult to derive general characterizations of the properties of the value set V_a as a function of the mechanism γ_a. Numerical analysis might be able to shed further light on this issue. In the next subsection, however, we take a different approach: we specify a mechanism that applies for a wide range of discount factors and probability distributions, but one that achieves only approximate efficiency, with the level of efficiency depending on the discount factor.

In summary, this subsection generates several conclusions. First, for some distributions of $\tilde{\boldsymbol{\theta}}$, a non-stationary mechanism with two regimes can achieve efficiency in every period. Second, when this mechanism is augmented with an appropriately chosen randomization prior to the first period, the agents expect to receive all the gains from trade, and ex ante expected deposits into and out of the account are equal to zero. Third, the specific mechanism we construct satisfies ex post IC and ex post IR only if the discount factor is sufficiently high, and if the average gains from trade are large relative to the maximum gains from trade. If these features fail, a different type of mechanism (one that delivers increasing utility as a function of account balances near the lower bound of the account) would be necessary, although it is not clear whether such a mechanism exists. Finally, we conjecture that for some distributions over $\tilde{\boldsymbol{\theta}}$ there is no discount factor for which it is possible to implement efficient trade in every period.

5.2 *Almost-efficient account-recursive mechanisms*

We now turn to analyze the existence of mechanisms that deliver approximately efficient trade. Subject to a lower bound on δ that is comparatively loose for many distributions, the mechanism we construct satisfies BBA, ex post IR, and ex post IC. In addition, as δ approaches 1, the fraction of periods in which efficient allocation is used approaches 1, and the sum of the players' ex ante expected per-period utilities approaches the efficient level, v^{ETU}. For every probability distribution over $\tilde{\boldsymbol{\theta}}$, there exists a δ close enough to 1 such that either an efficient mechanism of the form described in Section 5.1 exists, or an "almost-efficient" mechanism of the form described in this section exists, or both exist.

The mechanisms we construct employ inefficient allocation only when A nears one of the bounds. When the account balance nears zero, the mechanism makes use of a "revenue regime," which is guaranteed not to withdraw money from the account.

We construct an inefficient stage mechanism, $g(\cdot;A,\mathbf{v})=\langle q^r(\cdot),t^r(\cdot;A),w^r(\cdot;A)\rangle$, that utilizes misallocation in order to increase the account balance. The allocation function for the revenue regime sets $q_b^r(\boldsymbol{\theta})=\mathbf{1}\{\theta_b \geq \lambda_r(\theta_s)\}$, where λ_r is a strictly increasing function. In addition, we impose the following restriction on λ_r, which we show implies that the revenue regime always (weakly) adds cash to the account:

$$\lambda_r^{-1}(\overline{\theta}_b)-\lambda_r(\underline{\theta}_s)\leq 0 \tag{8}$$

for all $\boldsymbol{\theta}\in\Theta$. Let Λ_r be the space of all such functions. One subclass of Λ_r is

$$\lambda_r(\theta_s)=\frac{\overline{\theta}-\underline{\theta}}{2}\left(\frac{2(\theta_s-\underline{\theta})}{\overline{\theta}-\underline{\theta}}\right)^n+\frac{\overline{\theta}-\underline{\theta}}{2}+\underline{\theta}, \tag{9}$$

where $n>0$. For $n=1$ this simplifies to $\lambda_r(\theta_s)=\theta_s+\frac{1}{2}(\overline{\theta}-\underline{\theta})$. As $n\to\infty$, this class converges to a posted price mechanism with a price of $\frac{1}{2}(\overline{\theta}-\underline{\theta})$. (Note that the posted price mechanism itself raises zero revenue, but is not an element of Λ_r since it employs a λ function that is not strictly increasing.)

We make use of notation that represents the net gains from the revenue mechanism, accounting for both the benefits of trade and a portion of the sum of payments in each period, given $\lambda_r\in\Lambda_r$:

$$v^{\text{NGR}}(\boldsymbol{\theta})=q_b^r(\boldsymbol{\theta})\cdot\left(\theta_b-\theta_s+\lambda_r^{-1}(\theta_b)-\lambda_r(\theta_s)\right).$$

Note that $v^{\text{NGR}}(\boldsymbol{\theta})\geq 0$ for all $\boldsymbol{\theta}$ and $\mathbb{E}[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]>0$ by the definition of q_b^r. This is true despite the fact that $\lambda_r^{-1}(\theta_b)-\lambda_r(\theta_s)\leq 0$ by (8), so that $v^{\text{NGR}}(\boldsymbol{\theta})$ is always less than the benefits of trade in the revenue regime. Note further that $v^{\text{NGR}}(\boldsymbol{\theta})\leq v^{\text{GET}}(\boldsymbol{\theta})$ for all $\boldsymbol{\theta}$, recalling that $v^{\text{GET}}(\boldsymbol{\theta})$ represents the gains from trade under efficiency, since the efficient allocation rule $q_b^*(\boldsymbol{\theta})$ maximizes the value of $q(\boldsymbol{\theta})(\theta_b-\theta_s)$.

We rely on a condition on primitives that, for sufficiently high $\delta<1$, is satisfied whenever Condition 1 fails. It is under this condition that we construct almost-efficient mechanisms.

Condition 2. There exist $\delta<1$ and $\lambda_r\in\Lambda_r$ such that

$$\mathbb{E}[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]+\min\left\{2\delta\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})],\overline{\theta}-\underline{\theta}\right\}\geq 2\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})]. \tag{10}$$

Lemma 3. *There exists $\delta<1$ sufficiently high that either Condition 1 or Condition 2 is satisfied.*

Proof. First, since $\mathbb{E}[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]>0$ and Λ_r is nonempty, there exists $\delta<1$ such that $\mathbb{E}[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]+2\delta\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})]\geq 2\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})]$, regardless of Condition 1. Fix such a δ. Second, if for some $\lambda_r\in\Lambda_r$, (10) fails, then we have $\mathbb{E}[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]+2\delta\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})]>\mathbb{E}[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]+\overline{\theta}-\underline{\theta}$, which implies Condition 1. □

A number of subtleties arise in constructing mechanisms with the desired properties. Perhaps most salient is that (just as in the efficient mechanism of the last section) when agents need to switch between regimes that promise different levels of utility, they must be compensated by a potentially large withdrawal or deposit, since ex post IC would be disrupted if a large change in promised utility were not offset by a compensating monetary payment. Note that the revenue regime must promise relatively low joint utility, since not only is their allocation inefficient, but the traders must also expect to make costly deposits to raise the account balance. So if the regime adjacent to the revenue regime offered higher promised joint utility, then in states of the world where the mechanism would specify switching to the revenue regime, the traders would have to take large compensatory withdrawals in order to support ex post IC. But to avoid running out of funds in such a case, the critical value of the account (denoted A_r) at which the agents should switch to the revenue regime would need to be high enough that the account contained enough funds to accommodate the compensatory withdrawals. Since the magnitude of such withdrawals would increase with the agents' patience, it could take a long time to raise enough funds to increase the balance above A_r again, and a correspondingly long time before allocations became efficient again.

In order to keep the region of inefficient allocation small, we construct a mechanism in which changes in the sum of promised utilities can occur only at account balances far from the bounds. The mechanism we propose has the feature that A_r is small and does not vary with δ. In addition, there are two different regimes where efficient trade is implemented. For a region of moderately low account balances, the players use an "upward drift" regime, where trade is efficient and on average the agents deposit money into the account. For a region of higher account balances, the players use a "downward drift" regime, where trade is efficient and on average the agents withdraw money from the account. Finally, for a small region of very high account balances, the agents use a "payout" regime: trade shuts down for one period and the agents withdraw money from the account, shifting them back into a regime with efficient trade while maintaining an upper bound on the account size. In this way, the scheme ensures that the agents implement efficient trade most of the time, and the account balance usually stays in a moderate range.

As patience grows, large payments may be required when transitioning between the two efficiency regimes (the upward drift and downward drift regimes), but the range of account balances where those regimes are used also grows with patience. Since the sum of promised utilities does not change between the revenue regime and the upward drift regime, and since it also does not change between the downward drift regime and the payout regime, large monetary payments are not required at these transitions.

Proposition 5. *For any $\varepsilon > 0$, there exists $\underline{\delta} < 1$ such that, for all $\delta \in (\underline{\delta}, 1)$, there exists an account-recursive mechanism γ_a such that*

(i) ex post IR, ex post IC, and BBA are satisfied

(ii) starting from any initial account balance within the account bounds, the expected fraction of periods in which the object is allocated efficiently is at least $1-\varepsilon$.

The proof is in Section A.4. We focus our discussion on "almost-efficient" mechanisms that apply when Condition 2 holds. By Lemma 3, in all other situations Condition 1 must hold, in which case the result follows directly from Proposition 4. In the almost-efficient mechanism there are four regimes, for which the transition points $0 < A_r < A_u < A_d < A_p < \infty$ are chosen so that

$$V_a = \big([0, A_r] \times \{\mathbf{v}^r\}\big) \cup \big((A_r, A_u] \times \{\mathbf{v}^u\}\big) \cup \big((A_u, A_d] \times \{\mathbf{v}^d\}\big) \cup \big((A_d, A_p] \times \{\mathbf{v}^p\}\big) \tag{11}$$

and the promised utilities are chosen so that

$$v_b^r + v_s^r = v_b^u + v_s^u = \mathbb{E}\big[\tilde{\theta}_s + v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})\big] \tag{12}$$

$$v_b^d + v_s^d = v_b^p + v_s^p = \mathbb{E}\big[\tilde{\theta}_s + 2v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})\big]. \tag{13}$$

Furthermore, A_r and $A_p - A_d$ do not vary with δ, while $A_d - A_u$ and $A_u - A_r$ grow without bound as $\delta \to 1$. The regimes are characterized as follows.

- The *revenue regime*, which is employed when $A \in [0, A_r]$, uses the allocation rule $q_b(\boldsymbol{\theta}) = \mathbf{1}\{\theta_b \geq \lambda_r(\theta_s)\}$, so as to ensure that the traders do not withdraw from the account, and make a strictly positive expected deposit. Thus the balance never declines in this regime, and the balance almost surely rises above A_r within a finite number of periods. The traders' joint promised utility in this regime is $v_b^r + v_s^r = \mathbb{E}\big[\tilde{\theta}_s + v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})\big]$, which is the same promised utility they would obtain if they were to use the revenue regime mechanism in every period.

- The *upward drift regime*, which is employed when $A \in (A_r, A_u]$, uses an efficient allocation rule. The payment rule is set so that the players make a strictly positive expected deposit (this is the "upward drift"). The expected deposit is calibrated so that traders' joint promised utility is the same as in the revenue regime. This means that transitions between the revenue and upward drift regimes do not require any special withdrawals or deposits to counterbalance changes in promised utility. So if the account balance is in the upward drift regime in one period, then the traders' deposit or withdrawal is determined only by the realization of $\tilde{\boldsymbol{\theta}}$, as long as that deposit or withdrawal yields a new balance that is less than A_u. The transition point A_r is chosen high enough that, if the account balance is in the upward drift regime in one period, then the account balance in the following period is at least zero.

- The *downward drift regime*, which is employed when $A \in (A_u, A_d]$, also uses an efficient allocation rule. The payment rule is set to ensure a downward drift—that is, the traders make a strictly positive expected withdrawal. Somewhat arbitrarily, the expected withdrawal is calibrated so that efficient payoffs lie exactly halfway between $\mathbf{v}^d$ and $\mathbf{v}^u$; i.e., $\frac{1}{2}\big(v_b^u + v_s^u + v_b^d + v_s^d\big) = v^{\mathrm{ETU}}$, which can be verified using (12–13).

  Since the upward and downward drift regimes use different promised utilities but the same allocation rule, to preserve incentive compatibility the traders must be compensated with special deposits or withdrawals when they transition between these two regimes. Consider what happens when, starting from the downward drift regime, the traders make a withdrawal that brings the account balance into the upward drift regime. While in the downward drift regime, they were promised high joint utility, $\mathbf{v}^d$, but starting in the next period their promised joint utility will be low, $\mathbf{v}^u$. To compensate them for this change, they should receive an additional joint payment equal to the present value of the change, $\frac{\delta}{1-\delta}(v_b^d + v_s^d - (v_b^u + v_s^u))$. This is implemented by making an additional withdrawal for this amount. For δ close to one, this withdrawal is large. Similarly, when transitioning from the upward drift regime to the downward drift regime, they must make an additional large deposit. The result is that transitions between these two regimes always involve large changes in the account balance. The transition points A_u and A_d are chosen so that such transitions always yield balances in the range (A_r, A_d).

- The *payout regime*, which is employed when $A \in (A_d, A_p]$, always assigns the object to the seller, so as to avoid trade. In this regime, the traders always take a lump sum withdrawal from the account, so that the account balance always decreases in this regime, and the balance falls below A_d in a fixed number of periods. The withdrawal is calibrated so that the joint promised utility in the payout regime is the same as in the downward drift regime, so that transitions between these two regimes do not involve any special withdrawals or deposits to counterbalance changes in promised utility. Finally, the upper bound A_d is chosen high enough that, starting from the downward drift regime, the maximum possible account balance in the following period is no more than A_d.

  We note that the payout regime is not an essential feature of the mechanism. Instead, the downward drift regime could prohibit deposits. Then withdrawals would be larger on average, corresponding to a higher per-period promised joint utility in the downward drift regime. This would require larger withdrawals and deposits when transitioning between the upward and downward drift regimes, and thus increase the size of the requisite bounds on the account balance.

Notice that the sizes of the revenue and payout regimes (i.e., the magnitudes of A_r and $A_p - A_d$) are determined by properties that do not vary with δ—namely, the largest possible transitions into these regimes from adjacent regimes that share the same promised utility. Therefore these regimes remain the same size regardless of δ. On the other hand, the sizes of the upward and downward drift regimes (i.e., the magnitudes of $A_u - A_r$ and $A_d - A_u$) must increase on the order of $1/(1-\delta)$. These, of course, are the regions in which trade is efficient. In the proof, we demonstrate that every mechanism constructed in this way has a unique, atomless ergodic distribution. The drift properties of the different regimes ensure that—at this ergodic distribution—the proportion of periods spent in the revenue and payout regimes converges to zero as $\delta \to 1$.

As in the previous subsection, we ensure that the players receive their expected gains from trade and that the bank receives zero expected profits by means of a randomization that takes place before the first period of the mechanism.

COROLLARY 2. *For any $\varepsilon > 0$, there exists $\underline{\delta} < 1$ such that, for all $\delta \in (\underline{\delta}, 1)$, there exists a recursive mechanism γ_a satisfying the properties in Proposition 5, an initial account balance $\bar{A} \in (0, A_p)$, and an initial randomization over deposits and withdrawals (which occurs prior to the first period of the mechanism), such that, taking expectations over both the initial randomization and all possible realizations of play under the mechanism,*

(i) the expected account balance for each period is exactly $\bar{A}$, and

(ii) the expected joint utility is at least $v^{ETU} - \varepsilon$.

As before, the initial randomization ensures that the bank earns zero profits in expectation. The traders similarly must earn zero expected profits through changes in their account balance, and therefore when computing their expected payoffs in the mechanism we can restrict attention to the utility that they earn directly from trade. Since they start at the steady state, their expected joint average utility from trade is at least $\omega(\delta)v^{\mathrm{ETU}}$, where $\omega(\delta)$ is the proportion of periods spent in the upward and downward drift regimes given δ. Proposition 5 demonstrates that $\lim_{\delta \to 1} \omega(\delta) = 1$, and hence the traders' expected payoffs can be made arbitrarily close to v^{ETU} by a high enough choice of δ.

5.3 *Numerical examples of account-recursive mechanisms*

This section presents numerical examples of the exact-efficient and almost-efficient mechanisms constructed in Sections 5.1 and 5.2. For the examples in this section, we assume that θ_s and θ_b are independently and uniformly distributed on the unit interval. For simplicity of notation, we treat numerical approximations as if they were exact.

For the almost-efficient mechanisms considered in Section 5.2, we choose λ_r according to (9) with $n = 3$, and $C_u = C_d = 0.03$ (these are arbitrary parameters used in the construction of Proposition 5; see (29–30)). For $n = 3$, our almost-efficient mechanism is valid for $\delta \geq 0.67$. (Higher values of n can be used to support any $\delta > \frac{5}{8}$.) An exact-efficient mechanism of the form considered in Section 5.1 requires $\delta \geq 0.93$.

Table 1 displays steady state characteristics of almost-efficient, exact-efficient, and optimal posted price mechanisms for this probability distribution over $\tilde{\boldsymbol{\theta}}$. For the almost-efficient and exact-efficient mechanisms, some characteristics were approximated by Monte Carlo simulation.[20] Almost-efficient mechanisms are exhibited for $\delta \in \{0.67, 0.80, 0.90, 0.95, 0.99\}$; the exact-efficient mechanism is exhibited for $\delta \in \{0.95, 0.99\}$, since it is not valid for $\delta \leq 0.92$.

Table 1 demonstrates that, compared to the optimal posted price mechanism, the almost-efficient mechanisms we construct perform quite well even for low discount factors. The optimal posted price mechanism captures only 75% of the gains from efficient trade, but when $\delta = 0.67$ our almost-efficient mechanism captures 90% of the gains from

[20]In each case, a starting balance was chosen randomly according to the uniform distribution on $[0, A_p]$, and the mechanism was simulated for $10^5 + 10^6$ periods. A new starting balance was chosen randomly from among the last 10^6 of these periods, and was then adopted as the starting balance for an additional 10^7 periods. The ergodic distributions and steady state statistics were computed from these last 10^7 periods.

Mechanism	δ	$\mathbb{E}[v_b+v_s]$	%GFT	$\frac{1}{1-\delta}\mathbb{E}[v_b+v_s]$	$\bar{A}$	$\max A$
Almost efficient	0.67	0.650	90.1	1.97	1.70	2.65
	0.80	0.655	93.0	3.27	2.02	3.21
	0.90	0.661	96.7	6.61	2.81	4.65
	0.95	0.665	98.9	13.30	4.30	7.52
	0.99	0.667	99.998	66.67	15.88	30.51
Exact efficient	0.95	0.667	100	13.33	18.76	29.69
	0.99	0.667	100	66.67	105.34	162.94
Optimal posted price	any δ	0.625	75.0		0	0

Table 1: Characteristics of sample mechanisms. Approximate characteristics of sample mechanisms: almost-efficient and exact-efficient mechanisms as described in Section 5.3, and the optimal posted price mechanism as described in Section 4. Estimates are based on the ergodic distribution. Almost-efficient mechanisms are valid for $\delta \geq 0.67$, exact-efficient mechanisms are valid for $\delta \geq 0.93$, and the optimal posted price mechanism is valid for any δ. $\mathbb{E}[v_s+v_b]$ is the expected average joint utility. %GFT is the percent of the gains from efficient trade captured by each mechanism. $\frac{1}{1-\delta}\mathbb{E}[v_b+v_s]$ is the expected total joint utility, for comparison with $\bar{A}$, which is the average account balance as well as the amount of collateral the traders should hold. Finally, $\max A$ is the least upper bound on the account balance.

efficient trade. Even at this low discount factor, the total value of the traders' relationship is greater than the collateral they must hold. And as $\delta \to 1$, our almost-efficient mechanisms converge to efficiency, and the total value of the relationship grows faster than the collateral requirements.

Table 2 displays additional characteristics of the almost-efficient mechanisms. We see that in the revenue and upward drift regimes, agents receive lower per-period payoffs than they would in the posted price mechanism. This follows naturally, since the dynamic mechanism in those regions provides the same payoffs as a stationary mechanism where money is always deposited into the account (e.g., there is money burning); following the analysis of the previous section, for this probability distribution over $\tilde{\boldsymbol{\theta}}$, the posted price mechanism is optimal within the class of stationary mechanisms when no withdrawals are allowed but it is possible to burn money. We see also that when $\delta \geq .90$, for all possible account balances the joint value from continuing the relationship is greater than the value of the account; Table 3 shows that our exact-efficient mechanisms do not have this property.

The efficient mechanisms we construct, of course, attain efficiency exactly, but require $\delta \geq 0.93$. Table 1 shows that they also require significantly larger bounds on the account balance, compared to our almost-efficient mechanisms at the same discount factors. Table 3 displays additional characteristics of the exact-efficient mechanisms.

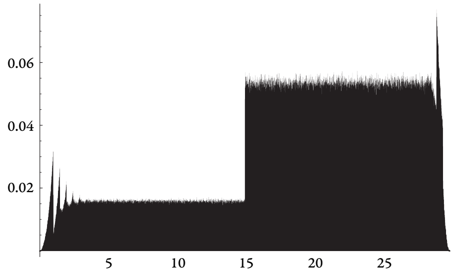
Figure 1 displays estimated ergodic densities for the exact-efficient mechanisms. The densities are sharply discontinuous at A_{dp}, and mostly flat otherwise. A large

δ	$\frac{v_b^u+v_s^u}{1-\delta}$	$\frac{v_b^p+v_s^p}{1-\delta}$	A_r	A_u	A_d	A_p	f_r	f_u	f_d	f_p
0.67	1.85	2.19	0.77	1.99	2.54	2.65	0.153	0.449	0.352	0.046
0.80	3.05	3.62	0.77	2.27	3.10	3.21	0.115	0.456	0.398	0.031
0.90	6.10	7.24	0.77	2.99	4.54	4.65	0.061	0.471	0.456	0.013
0.95	12.20	14.47	0.77	4.43	7.41	7.52	0.022	0.487	0.488	0.003
0.99	60.98	72.35	0.77	15.92	30.40	30.51	7×10^{-5}	0.499	0.501	6×10^{-7}

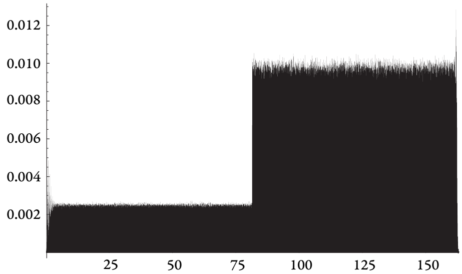
Table 2: Characteristics of almost-efficient mechanisms. Approximate characteristics of almost-efficient account-recursive mechanisms, as described in Section 5.3. Estimates are based on the ergodic distribution. A_m, for $m \in \{r,u,d,p\}$, is the maximum account balance in regime m; A_p is the least upper bound on the account. f_m is the fraction of periods spent in regime m. For all these mechanisms, $v_b^r + v_s^r = v_b^u + v_s^u = 0.610$ and $v_b^d + v_s^d = v_b^p + v_s^p = 0.724$.

δ	$v_b^{dp}+v_s^{dp}$	$v_b^{wd}+v_s^{wd}$	$\frac{v_b^{dp}+v_s^{dp}}{1-\delta}$	$\frac{v_b^{wd}+v_s^{wd}}{1-\delta}$	A_{dp}	A_{wd}	f_{dp}	f_{wd}
0.95	0.104	0.833	2.083	16.667	14.854	29.688	0.229	0.771
0.99	0.021	0.833	2.083	83.333	81.438	162.937	0.205	0.795

Table 3: Characteristics of exact-efficient mechanisms. Approximate characteristics of exact-efficient account-recursive mechanisms, as described in Section 5.3. Estimates are based on the ergodic distribution. A_m, for $m \in \{dp, wd\}$, is the maximum account balance in regime m; A_{wd} is the least upper bound on the account. f_m is the fraction of periods spent in regime m.



(a) $\delta = 0.95$



(b) $\delta = 0.99$

Figure 1: Ergodic distributions for sample exact-efficient mechanisms. Approximate ergodic distributions of the account balance for two exact-efficient mechanisms, as described in Section 5.3, for two values of δ. In each graph, the horizontal axis displays the account balance, from 0 to A_{wd}. The vertical axis displays the probability density.

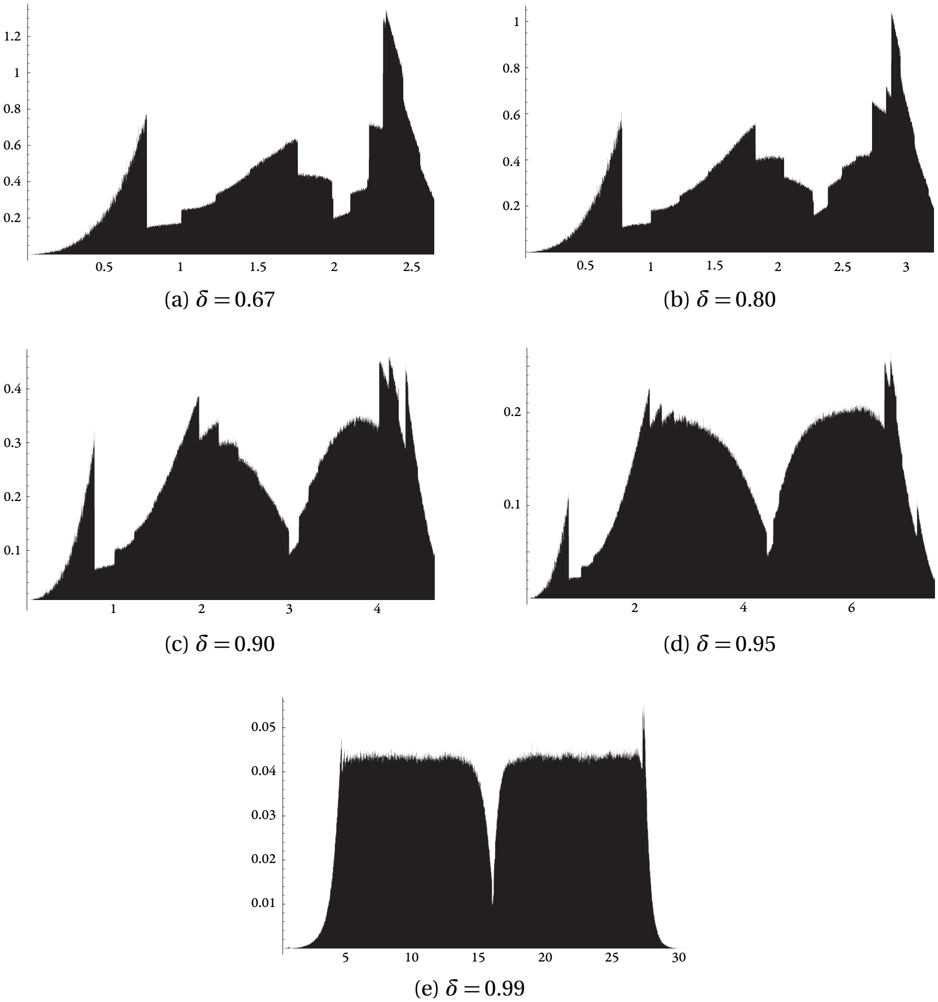


Figure 2: Ergodic distributions for sample almost-efficient mechanisms. Approximate ergodic distributions of the account balance for five almost-efficient mechanisms, as described in Section 5.3, each for a different value of δ. In each graph, the horizontal axis displays the account balance, from 0 to A_p. The vertical axis displays the probability density.

interval of account balances below A_{dp} is reached only from lower balances, while a large interval above A_{dp} is reached only from higher balances. The spikes at the extreme account balances correspond to large changes that occur in conjunction with regime shifts.

Figure 2 displays estimated ergodic densities for the almost-efficient mechanisms. For very high discount factors there are spikes in the distribution that appear to be atoms at low magnification. At higher magnification these regions appear to be atomless, as expected. For high discount factors, the ergodic density of each almost-efficient mechanism features a distinctive “canyon” at A_u, since probability mass is drawn away from

A_u by the large compensatory withdrawals and deposits that are likely in that vicinity. Also for high discount factors, the density drops off quickly near the account bounds, so that the probability mass in the revenue regime and the payout regime is very small.

6. Related literature

In their seminal paper, Myerson and Satterthwaite (1983) analyze the static problem for the case where players' valuations are private and independently distributed. They show that efficient trade is impossible under any mechanism that satisfies interim IR, interim ("Bayesian") IC, and ex ante BB. McAfee and Reny (1992) demonstrate a mechanism that overturns this result when valuations are correlated, but Chung and Ely (2002) show that imposing "ex post equilibrium" (i.e., ex post IC and ex post IR) restores Myerson and Satterthwaite's negative conclusion. The proof of the Myerson–Satterthwaite Theorem makes clear that efficiency could be restored if the IR constraints could be relaxed, so that players would be willing to make payments in excess of their gains from trade, as is the case in an ongoing relationship.

Our mechanism design approach employs dynamic programming in the spirit of Abreu, Pearce, and Stacchetti (1986, 1990), who show that the incentives provided by future play can be summarized in a continuation reward function that maps today's outcomes into future equilibrium payoffs. This allows us to apply standard tools of mechanism design to the dynamic program, exploiting the equivalence between incentives offered by monetary payments in the present and changes in promised future utility (Athey and Bagwell 2001; Athey, Bagwell, and Sanchirico 2004). We wish to emphasize that the mechanism design approach does not rely on the presence of an independent mechanism designer in the model; instead, the traders mutually agree on a self-enforcing mechanism that they themselves design, where the "mechanism" itself can be thought of as a metaphor for an unstructured trading relationship in the presence of certain institutions. In relational contracting problems, this approach can be used to show that productive relationships can be sustained even when contract enforcement is incomplete. Furthermore, efficient mechanisms can often be stationary due to the players' ability to "settle up" in the present by paying money rather than trading future utility (Levin 2003; Rayo forthcoming).

Our approach is related also to that of Fudenberg, Levine, and Maskin (1994), who prove a folk theorem for repeated games with hidden actions or information. Although they focus on the case in which players cannot make monetary payments, their framework can be extended to accommodate monetary payments (indeed, this would greatly simplify their analysis). Furthermore, although their result is for finite type spaces, Miller (2007a) extends it to the continuous type spaces that we consider here. When applied to the repeated trade problem, these results can be used to show that as the discount factor approaches unity, the best equilibrium approaches efficiency, under a mechanism that satisfies interim IC and ex post BB.[21] Additionally,

[21] An alternative way to achieve efficiency in a perfect public equilibrium with interim IC is based on the ideas of Cramton, Gibbons, and Klemperer (1987). When players are patient, the seller can potentially be induced to sell rights to half of the object to the agent at a fixed price at the beginning of each period. If the

Jackson and Sonnenschein (2004) consider long but finite time horizons, and construct interim IC and ex post BB mechanisms that approximate efficiency in a repeated trade setting, by linking the outcomes across periods. Their mechanisms yield approximate efficiency for sufficiently long time horizons and sufficiently patient traders. Our approach is distinguished by a focus on weaker institutions, corresponding to ex post IC.

Our results employ the concept of ex post perfect public equilibrium (EPPPE) in repeated games with hidden information. EPPPE was introduced by Miller (2007b), who proves that an ex post IC, ex post BB mechanism cannot attain or approximate efficiency, even as the discount factor approaches one, in a class of games that includes the repeated trade game as a special case.

Our result that ex post IC (together with ex post IR and ex post BB) leads to a substantial loss of efficiency can be related to recent findings by other authors showing that inefficiency may be optimal in dynamic games of hidden information. In particular, in studies considering very different stage games, Athey, Bagwell, and Sanchirico (2004), Athey, Atkeson, and Kehoe (2005), and Amador, Werning, and Angeletos (2006) all establish conditions (typically involving restrictions on the hazard rate of the distribution of hidden information) under which optimal mechanisms do not make use of agents' hidden information, and inefficiency results. Interestingly, all of these papers use interim IC, while our results are driven by the imposition of ex post IC; the similarity with our work is that, due to various features of the problems they study, in all of these papers the instruments available to provide incentives are associated with a reduction of social surplus.

A number of other authors have employed alternative approaches to the problem of efficiency with a large number of traders. Rustichini, Satterthwaite, and Williams (1994), Swinkels (2001), Satterthwaite and Williams (2002), and Tatur (2005) consider double auctions with large numbers of traders; their results indicate that such settings can yield asymptotic efficiency, but generally not budget balance. Satterthwaite and Shneyerov (2007) show that asymptotically efficient trade is possible in a dynamic, decentralized matching market with a large number of traders. In a continuous time general equilibrium setting, Taub (1994) shows that an optimal contract among a finite number of asymmetrically informed traders of a divisible good is inefficient. Our work complements these approaches by expanding the time horizon of the analysis rather than the number of traders.

7. Conclusion

In this paper, we have shown that the institutional environment is important in determining the structure and performance of the optimal trading relationship. When the institutional environment provides unlimited insurance against imbalances, efficiency is attainable in a stationary mechanism for any discount factor of at least one half, as we show in Section 3. When there is no source of insurance, so that ex post budget balance is appropriate, the optimal mechanism for trade entails a sacrifice of efficiency.

seller refuses, then trade breaks down. Subsequently, an auction can be used to allocate the object among the two players. This scheme would satisfy interim IC but not ex post IC within a period.

Strikingly, for a wide range of distributions over $\tilde{\boldsymbol{\theta}}$, a simple posted price mechanism is optimal for all discount factors, so that the repeated nature of the relationship plays no role at all.

Section 5 shows that players may benefit from employing a joint credit line account, so that they can self-insure against budget imbalances. For patient players, if they hold sufficient collateral then such an account can sometimes enable them to trade efficiently, and more generally can allow them to approximate efficient trade. However, schemes that keep the size of the account within bounds can be quite complex to design. The problem is that when agents anticipate making systematic deposits to the account (as when the account balance gets small) or systematic withdrawals (when the account balance gets large), incentives for truthful revelation in the current period are affected, as players recognize that their announcements affect the future account activity. Counteracting those incentives requires the agents to adjust their payments in the current period. Since these adjustments might be large, they could require large withdrawals from the account just when its balance is running low. We construct two types of mechanisms—one exactly efficient and one approximately efficient—that keep the account within specified bounds. The exact efficient mechanism relies on a condition that some probability distributions over $\tilde{\boldsymbol{\theta}}$ violate.

We show that for all distributions, if the discount factor is sufficiently high then either the exactly efficient mechanism can be used, or else the approximately efficient mechanism is available (or both). Both mechanisms subtly calibrate the players' incentives, so that players sometimes (when the balance is moderately low) implement efficient trade and, on average, deposit money into the account; and where at other times (when the balance is moderately high) players implement efficient trade and, on average, withdraw money from the account. In the approximately efficient mechanism, when the balance gets very low or very high, players use inefficient trade to raise revenue or withdraw money so that the account stays within fixed bounds.

These mechanisms are not stationary, but they are ergodic. By starting their trading relationship with a randomization that replicates the ergodic distribution over account balances, the traders can jump immediately to the steady state. This allows us to demonstrate that the mechanisms give all the expected gains from trade to the traders without extracting any resources from the insuring institution.

Several extensions might be interesting to consider in future research. In this paper we focus on the case of independent private valuations; a natural extension concerns interdependent and correlated valuations. All our results can be easily extended to correlated private valuations, with small changes in the details; ex post IC provides the necessary robustness. Our analytic results also extend naturally to the case of interdependent valuations when the requisite allocation rules are ex post implementable. That is, if the efficient allocation rule is ex post implementable, then Propositions 1–4 extend with minor modifications, and if there is an ex post implementable allocation rule for the revenue regime then Proposition 5 extends as well.

A second set of extensions involves placing restrictions on the magnitude of payments that are possible. If we assumed that players did not have access to cash outside of the surplus generated in the period's trade, then non-stationary schemes would be required to provide players sufficient incentives. In an extreme case, we might imagine that players could not use monetary payments at all, which might be the case if multi-directional trade takes place within a firm or among politicians in a legislature. In that case, non-stationary schemes would use future rewards and punishments to provide incentives, as in Fudenberg, Levine, and Maskin (1994) and Athey and Bagwell (2001). A third set of extensions considers situations in which players cannot fully commit not to renegotiate off the equilibrium path. Notions of renegotiation proofness or costly renegotiation would impose new constraints on the utilities attainable in equilibrium, and lead to new conclusions about the forms of optimal equilibria.

Appendix: Proofs not included in the text

A.1 *Proof of Lemma 2 (page 314)*

By promise keeping and coherence, for each i and each $(A,\mathbf{v})\in V_a$,

$$v_i=(1-\delta)\mathbb{E}\big[\tilde{\theta}_i q_i(\tilde{\boldsymbol{\theta}};A,\mathbf{v})+t_i(\tilde{\boldsymbol{\theta}};A,\mathbf{v})\big]+\delta\mathbb{E}\big[w_i(\tilde{\boldsymbol{\theta}};A,\mathbf{v})\big].$$

Therefore

$$\begin{aligned}\sum_i v_i&=(1-\delta)\mathbb{E}\Big[\sum_i\tilde{\theta}_i q_i(\tilde{\boldsymbol{\theta}};A,\mathbf{v})+A-\alpha(\tilde{\boldsymbol{\theta}};A,\mathbf{v})\Big]+\delta\mathbb{E}\Big[\sum_i w_i(\tilde{\boldsymbol{\theta}};A,\mathbf{v})\Big]\\&=(1-\delta)\mathbb{E}\Big[\sum_i\tilde{\theta}_i^{(1)}q_i(\tilde{\boldsymbol{\theta}}^{(1)};\tilde{A}^{(1)},\tilde{\mathbf{v}}^{(1)})+\tilde{A}^{(1)}-\tilde{A}^{(2)}\Big]+\delta\mathbb{E}\big[\sum_i\tilde{v}_i^{(2)}\big]\\&=(1-\delta)\sum_{\tau=1}^{T}\delta^{\tau-1}\mathbb{E}\Big[\sum_i\tilde{\theta}_i^{(\tau)}q_i(\tilde{\boldsymbol{\theta}}^{(\tau)};\tilde{A}^{(\tau)},\tilde{\mathbf{v}}^{(\tau)})+\tilde{A}^{(\tau)}-\tilde{A}^{(\tau+1)}\Big]+\delta^T\mathbb{E}\Big[\sum_i\tilde{v}_i^{(T+1)}\Big]\end{aligned}$$

for all T, where expectations are taken at time zero over all $\tilde{\boldsymbol{\theta}}^{(1)},\tilde{\boldsymbol{\theta}}^{(2)},\ldots$. Hence account keeping is satisfied if

$$\lim_{T\to\infty}\delta^T\mathbb{E}\Big[\sum_i\tilde{v}_i^{(T+1)}\Big]=\lim_{T\to\infty}(1-\delta)\sum_{\tau=T+1}^{\infty}\delta^{\tau-1}\mathbb{E}\Big[\sum_i\tilde{\theta}_i^{(\tau)}q_i(\tilde{\boldsymbol{\theta}}^{(\tau)};\tilde{A}^{(\tau)},\tilde{\mathbf{v}}^{(\tau)})+\tilde{A}^{(\tau)}-\tilde{A}^{(\tau+1)}\Big].$$

Since $\tilde{A}$, $\tilde{\mathbf{v}}$, and $\tilde{\theta}$ are all uniformly bounded random variables, both sides of this equation are exactly zero.

A.2 *Proof of Proposition 4 (page 315)*

We begin with a lemma that relates Condition 1 to the problem of finding a stage mechanism that satisfies ex post IC, ex post IR, has no withdrawals from the account, and has continuation rewards equal to the promised utility of the stage mechanism. The stage mechanism used when $A=0$ has these features.

Let Problem (SP) be the problem of choosing a stage mechanism $g=\langle q,t,w\rangle$ to maximize

$$\mathbb{E}[U_b(\tilde{\boldsymbol{\theta}},g)]+\mathbb{E}[U_s(\tilde{\boldsymbol{\theta}},g)]$$

subject to efficiency ($q=q^*$), ex post IR, ex post IC, $t_b(\boldsymbol{\theta})+t_s(\boldsymbol{\theta})\le 0$ for all $\boldsymbol{\theta}$, and

$$w(\boldsymbol{\theta})=\big(\mathbb{E}[U_b(\tilde{\boldsymbol{\theta}},g)],\mathbb{E}[U_s(\tilde{\boldsymbol{\theta}},g)]\big) \text{ for all } \boldsymbol{\theta}.$$

Let Problem (SP′) be the problem of choosing real-valued functions h_b, h_s to maximize $\mathbb{E}[h_b(\tilde{\theta}_s)+h_s(\tilde{\theta}_b)]$ subject to (6–7).

LEMMA 4. *For a given δ, if Condition 1 holds then there exist $\breve{h}_b$ and $\breve{h}_s$ in the constraint set of Problem (SP′), and $\breve{g}=\langle q^*,\breve{t},\breve{w}\rangle$ is in the constraint set of Problem (SP), where, for all $\boldsymbol{\theta}$,*

$$\breve{t}_b(\boldsymbol{\theta})=-q_b^*(\boldsymbol{\theta})\cdot\theta_s+\breve{h}_b(\theta_s)-\frac{\delta}{1-\delta}\breve{v}_b$$
$$\breve{t}_s(\boldsymbol{\theta})=q_b^*(\boldsymbol{\theta})\cdot\theta_b+\breve{h}_s(\theta_b)-\frac{\delta}{1-\delta}\breve{v}_s$$
$$\breve{w}_b(\boldsymbol{\theta})=\breve{v}_b=(1-\delta)\mathbb{E}\left[\breve{h}_b(\tilde{\theta}_s)+\int_{\underline{\theta}_b}^{\tilde{\theta}_b} q_b^*(b,\tilde{\theta}_s)\,db\right]=\mathbb{E}[U_b(\tilde{\boldsymbol{\theta}},\breve{g})] \tag{14}$$
$$\breve{w}_s(\boldsymbol{\theta})=\breve{v}_s=(1-\delta)\mathbb{E}\left[\breve{h}_s(\tilde{\theta}_b)-\int_{\tilde{\theta}_s}^{\overline{\theta}_s}\big(1-q_b^*(\tilde{\theta}_b,s)\big)\,ds+\overline{\theta}_s\right]=\mathbb{E}[U_s(\tilde{\boldsymbol{\theta}},\breve{g})]. \tag{15}$$

If instead Condition 1 fails, then there is no solution to Problem (SP).

PROOF. First, consider Problem (SP). For any g for which (a) continuation rewards are constant at $w(\boldsymbol{\theta})=(v_b,v_s)$ and (b) ex post IC holds, the envelope theorem implies that

$$\frac{\mathbb{E}[U_b(\tilde{\boldsymbol{\theta}},g)]}{1-\delta}=\mathbb{E}\left[\underline{\theta}_b q_b^*(\underline{\theta}_b,\tilde{\theta}_s)+t_b(\underline{\theta}_b,\tilde{\theta}_s)+\frac{\delta}{1-\delta}v_b+\int_{\underline{\theta}_b}^{\tilde{\theta}_b} q_b^*(b,\tilde{\theta}_s)\,db\right],$$
$$\frac{\mathbb{E}[U_s(\tilde{\boldsymbol{\theta}},g)]}{1-\delta}=\mathbb{E}\left[\overline{\theta}_s(1-q_b^*(\tilde{\theta}_b,\overline{\theta}_s))+t_s(\tilde{\theta}_b,\overline{\theta}_s)+\frac{\delta}{1-\delta}v_s-\int_{\tilde{\theta}_s}^{\overline{\theta}_s}\big(1-q_b^*(\tilde{\theta}_b,s)\big)\,ds\right].$$

Ex post IC implies that monetary payments take the form, for some h_b and h_s,

$$t_b(\boldsymbol{\theta})=-\theta_s q_b^*(\boldsymbol{\theta})+h_b(\theta_s)-\frac{\delta}{1-\delta}v_b \tag{16}$$
$$t_s(\boldsymbol{\theta})=\theta_b q_b^*(\boldsymbol{\theta})+h_s(\theta_b)-\frac{\delta}{1-\delta}v_s. \tag{17}$$

Thus,

$$t_b(\underline{\theta}_b,\theta_s)=h_b(\theta_s)-\frac{\delta}{1-\delta}v_b$$
$$t_s(\theta_b,\overline{\theta}_s)=h_s(\theta_b)-\frac{\delta}{1-\delta}v_s,$$

since $q_b^*(\boldsymbol{\theta})=0$ for these cases.

Let

$$\begin{aligned} v^{\mathrm{EIR}} &\equiv \mathbb{E}\left[\overline{\theta}_s + \int_{\underline{\theta}_b}^{\tilde{\theta}_b} q_b^*(b,\tilde{\theta}_s)\,db - \int_{\tilde{\theta}_s}^{\overline{\theta}_s} \left(1-q_b^*(\tilde{\theta}_b,s)\right)ds\right] \\ &= \mathbb{E}\left[\overline{\theta}_s + q_b^*(\tilde{\boldsymbol{\theta}})(\tilde{\theta}_b-\tilde{\theta}_s) - \left(1-q_b^*(\tilde{\boldsymbol{\theta}})\right)(\overline{\theta}_s-\tilde{\theta}_s) - q_b^*(\tilde{\boldsymbol{\theta}})(\overline{\theta}_s-\tilde{\theta}_b)\right] \\ &= \mathbb{E}\left[\tilde{\theta}_s + 2q_b^*(\tilde{\boldsymbol{\theta}})\tilde{\theta}_b - 2q_b^*(\tilde{\boldsymbol{\theta}})\tilde{\theta}_s\right] = \mathbb{E}[\tilde{\theta}_s] + 2\mathbb{E}[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}})]. \end{aligned}$$

It then follows that

$$\frac{\mathbb{E}[U_b(\tilde{\boldsymbol{\theta}},g)]}{1-\delta} + \frac{\mathbb{E}[U_s(\tilde{\boldsymbol{\theta}},g)]}{1-\delta} = v^{\mathrm{EIR}} + \mathbb{E}\left[h_b(\tilde{\theta}_s)+h_s(\tilde{\theta}_b)\right]. \tag{18}$$

Thus, given (16–17) and the specified q and w, maximizing $\mathbb{E}[h_b(\theta_s)+h_s(\theta_b)]$ is equivalent to maximizing $\mathbb{E}[U_b(\boldsymbol{\theta},g)]+\mathbb{E}[U_s(\boldsymbol{\theta},g)]$.

We construct a mechanism to solve Problem (SP) using the monetary payments in (16–17). Ex post IR for the buyer requires that

$$\theta_b q_b^*(\boldsymbol{\theta}) + t_b(\boldsymbol{\theta}) + \frac{\delta}{1-\delta} v_b \geq 0,$$

for all $\boldsymbol{\theta}$—which, using (16), holds if and only if $h_b(\theta_s)\geq 0$ for all θ_s. For the seller,

$$\theta_s\left(1-q_b^*(\boldsymbol{\theta})\right) + t_s(\boldsymbol{\theta}) + \frac{\delta}{1-\delta} v_s \geq \theta_s + \frac{\delta}{1-\delta}\mathbb{E}[\tilde{\theta}_s],$$

which, using (17), holds for all $\boldsymbol{\theta}$ if and only if $h_s(\theta_b)\geq \frac{\delta}{1-\delta}\mathbb{E}[\tilde{\theta}_s]$ for all θ_b. Thus, (6) is equivalent to ex post IR under (16–17). Now consider the constraint that $t_b(\boldsymbol{\theta})+t_s(\boldsymbol{\theta})\leq 0$ for all $\boldsymbol{\theta}$. Using (16–17), and substituting in the requirement that $(v_b,v_s)=\left(\mathbb{E}[U_b(\tilde{\boldsymbol{\theta}},g)],\mathbb{E}[U_s(\tilde{\boldsymbol{\theta}},g)]\right)$ using (18), the no-withdrawal constraint becomes (7). It then follows that Problems (SP′) and (SP) are equivalent in the sense given in the statement of the lemma, and that Condition 1 is necessary and sufficient for each problem to have a non-empty constraint set. □

We now construct the mechanism that proves Proposition 4. The value set is defined using A_{dp}, A_{wd}, $\mathbf{v}^{dp}$, and $\mathbf{v}^{wd}$ with $0<A_{dp}<A_{wd}<\infty$, as $V_a=\left([0,A_{dp}]\times\{\mathbf{v}^{dp}\}\right)\cup\left((A_{dp},A_{wd}]\times\{\mathbf{v}^{wd}\}\right)$. Note that promised utilities are constant within each of the regimes.

For both regimes, we construct monetary payments in terms of continuation rewards. (We define the continuation rewards later.) For an arbitrary regime $m\in\{dp,wd\}$, let

$$t_b^m(\boldsymbol{\theta};A) = -q_b^*(\boldsymbol{\theta})\cdot\theta_s + h_b^m(\theta_s) - \frac{\delta}{1-\delta} w_b^m(\boldsymbol{\theta};A) \tag{19}$$

$$t_s^m(\boldsymbol{\theta};A) = q_b^*(\boldsymbol{\theta})\cdot\theta_b + h_s^m(\theta_b) - \frac{\delta}{1-\delta} w_s^m(\boldsymbol{\theta};A). \tag{20}$$

Observe that the last term on each line cancels out the utility that each trader receives from his or her continuation reward. Therefore the total transfers (monetary payments plus continuation rewards) are exactly as specified in the VCG mechanism (Lemma 1), establishing that ex post IC is satisfied. Furthermore, promised utilities in regime m are exactly

$$v_b^m = (1-\delta)\mathbb{E}\left[v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) + h_b^m(\tilde{\theta}_s)\right] \tag{21}$$

$$v_s^m = (1-\delta)\mathbb{E}\left[\tilde{\theta}_s + v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) + h_s^m(\tilde{\theta}_b)\right]. \tag{22}$$

The deposit regime The deposit regime is characterized by two conditions, in addition to the usual ex post IR and IC.

1. The traders must not withdraw money from the account (i.e., $t_b^m(\boldsymbol{\theta};A)+t_s^m(\boldsymbol{\theta};A) \leq 0$ for all $\boldsymbol{\theta} \in \Theta$ and all $A \in [0, A_{dp}]$).

2. Subject to the first constraint, $v_s^{dp} + v_b^{dp}$ is maximized (or approximately maximized).

By Lemma 4, these constraints are satisfied by choosing $h_b^{dp} = \breve{h}_b$ and $h_s^{dp} = \breve{h}_s$ (where $\breve{h}_b$ and $\breve{h}_s$ solve Problem (SP′) in Lemma 4).[22] Thus $\mathbf{v}^{dp} = \breve{\mathbf{v}}$ as given in (14–15). Further, Lemma 4 establishes that any stage mechanism in the deposit regime respects ex post IR.

We need to specify the continuation reward function, w^{dp}. If a transition remains in the deposit regime, continuation rewards are fixed at $w^{dp}(\boldsymbol{\theta};A) = \mathbf{v}^{dp}$. Starting from an account balance of $A \leq A_{dp}$, in order for a transition to remain in the deposit regime, the new account balance must be less than A_{dp}. Using (18) (which applies because continuation values are constant in this region) and (19–20), this is the case if and only if

$$\begin{aligned} A &- \left(t_b^{dp}(\boldsymbol{\theta};A) + t_b^{dp}(\boldsymbol{\theta};A)\right) \\ &= A - \left(v^{\text{GET}}(\boldsymbol{\theta}) + h_b^{dp}(\theta_s) + h_s^{dp}(\theta_b) - \frac{\delta}{1-\delta}\left(v_b^{dp} + v_s^{dp}\right)\right) \\ &= A - \left(v^{\text{GET}}(\boldsymbol{\theta}) + h_b^{dp}(\theta_s) + h_s^{dp}(\theta_b) - \delta\left(v^{\text{EIR}} + \mathbb{E}[h_b^{dp}(\tilde{\theta}_s) + h_s^{dp}(\tilde{\theta}_b)]\right)\right) \\ &\leq A_{dp}. \end{aligned} \tag{23}$$

Otherwise, the transition is to switch to the withdrawal regime. Thus, we define $w^{dp}(\boldsymbol{\theta};A) = \mathbf{v}^{dp}$ if (23) holds, and let $w^{dp}(\boldsymbol{\theta};A) = \mathbf{v}^{wd}$ otherwise.

The withdrawal regime The withdrawal regime is characterized by an analogous pair of constraints, in addition to ex post IR and IC.

1. The traders must not deposit money into the account (i.e., $t_b^m(\boldsymbol{\theta};A)+t_s^m(\boldsymbol{\theta};A) \geq 0$ for all $\boldsymbol{\theta} \in \Theta$ and all $A \in (A_{dp}, A_{wd}]$).

[22] Since under Condition 1 the constraint set for Problem (SP′) is non-empty, either an exact solution exists or an approximate solution (satisfying the constraint set) exists. If an exact solution does not exist, choose $\breve{h}_b$, $\breve{h}_s$, and $\breve{g}$ to be an approximate solution.

2. Subject to the first constraint, $v_s^{wd}+v_b^{wd}$ is minimized (or approximately minimized).

To satisfy these constraints, we set $h_b^{wd}(\theta_s)=0$ for all θ_s and $h_s^{wd}(\theta_b)=\delta v^{\text{EIR}}/(1-\delta)$ for all θ_b. That the first constraint is satisfied is evident from (21–22). The second constraint is satisfied as well, since the traders' joint withdrawal from the account is zero whenever there is no trade. They cannot be induced to withdraw any less when trade does occur, because any such inducement would need to operate through either h_b^{wd} or h_s^{wd}, neither of which can be conditioned on trade, so that any decrease in $h_i^{wd}(\theta_{-i})$ for some particular value of θ_{-i} would lead to a net deposit into the account for realizations of θ_i such that trade occurs. Note that these definitions of h_b^{wd} and h_s^{wd} imply also that ex post IR is satisfied given the monetary payments specified in (19–20).

Note that we can now use (21–22) to solve for $\mathbf{v}^{wd}$:

$$\begin{aligned}v_b^{wd}&=(1-\delta)\mathbb{E}\left[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})\right]\\ v_s^{wd}&=(1-\delta)\mathbb{E}\left[\tilde{\theta}_s+v^{\text{GET}}(\tilde{\boldsymbol{\theta}})+\frac{\delta}{1-\delta}v^{\text{EIR}}\right].\end{aligned}$$

This implies that $v_b^{wd}+v_s^{wd}=v^{\text{EIR}}$, which in turn implies that $t_b^{wd}(\boldsymbol{\theta};A)+t_s^{wd}(\boldsymbol{\theta};A)=v^{\text{GET}}(\boldsymbol{\theta})$ as long as the new balance $\alpha(\boldsymbol{\theta},A,\mathbf{v}^{wd})$ is still within the withdrawal regime—that is, if

$$A-v^{\text{GET}}(\boldsymbol{\theta})\geq A_{dp}. \tag{24}$$

As long as (24) holds, we set $w^{wd}(\boldsymbol{\theta};A)=\mathbf{v}^{wd}$, while otherwise we let $w^{dp}(\boldsymbol{\theta};A)=\mathbf{v}^{dp}$.

Respecting account bounds Because promised utilities are lower in the deposit regime than in the withdrawal regime (i.e., $v_b^{dp}+v_s^{dp}<v_b^{wd}+v_s^{wd}$), by our definitions the traders must make a potentially large deposit into the account when transitioning from the deposit regime to the withdrawal regime, and take a potentially large withdrawal from the account when transitioning from the withdrawal regime to the deposit regime. Intuitively, this is necessary because otherwise the players would be tempted to misreport $\boldsymbol{\theta}$ when doing so would enable them to switch to a regime with higher promised utility, or avoid switching to a regime with lower promised utility. The more patient are the traders, the more they value a change in continuation rewards, and hence the larger the deposit or withdrawal must be. Mechanically, this operates through (19–20), where by our definitions a slight change in $\boldsymbol{\theta}$ (to induce a regime switch) leads to a large change in $w^m(\boldsymbol{\theta};A)$.

If the mechanism is to respect the bounds on the account, the range of balances in each regime must be at least as large as the largest possible change in the balance that can be caused by switching into that regime.

The smallest account balance that can result from switching from the withdrawal regime to the deposit regime occurs starting from A just above A_{dp}, after $(\overline{\theta}_b,\underline{\theta}_s)$ is realized. Then the lower bound of the account is respected if

$$\begin{aligned}
&A_{dp} - \max_{\boldsymbol{\theta}'}\left(t_s^{wd}(\boldsymbol{\theta}';A_{dp}) + t_b^{wd}(\boldsymbol{\theta}';A_{dp})\right) \\
&\quad = A_{dp} - \left(\overline{\theta} - \underline{\theta} + h_b^{wd}(\underline{\theta}) - \frac{\delta}{1-\delta} v_b^{dp} + h_s^{wd}(\overline{\theta}) - \frac{\delta}{1-\delta} v_s^{dp}\right) \\
&\quad = A_{dp} - \left(\overline{\theta} - \underline{\theta} + \frac{\delta}{1-\delta} v^{\mathrm{EIR}}\right) + \delta\left(v^{\mathrm{EIR}} + \mathbb{E}[h_b^{dp}(\tilde{\theta}_s) + h_s^{dp}(\tilde{\theta}_b)]\right) \\
&\quad = A_{dp} - \left(\overline{\theta} - \underline{\theta} + \frac{\delta^2}{1-\delta} v^{\mathrm{EIR}} - \delta\mathbb{E}[h_b^{dp}(\tilde{\theta}_s) + h_s^{dp}(\tilde{\theta}_b)]\right) \geq 0.
\end{aligned}$$

Let $\boldsymbol{\theta}^{dp}$ be the value of $\boldsymbol{\theta}$ that minimizes $v^{\mathrm{GET}}(\boldsymbol{\theta}) + h_b^{dp}(\theta_s) + h_s^{dp}(\theta_b)$, the variable part of monetary payments in the deposit regime. Then the largest account balance that can result when switching from the deposit regime to the withdrawal regime occurs starting from $A = A_{dp}$, after $\boldsymbol{\theta} = \boldsymbol{\theta}^{dp}$ is realized. Then the upper bound on the account balance, A_{wd}, is respected if

$$A_{wd} \geq A_{dp} - \left(v^{\mathrm{GET}}(\boldsymbol{\theta}^{dp}) + h_b^{dp}(\theta_s^{dp}) + h_s^{dp}(\theta_b^{dp}) - \frac{\delta}{1-\delta}\left(v_b^{wd} + v_s^{wd}\right)\right).$$

Summary The requirements on A_{wd} and A_{dp} as a function of h_b^{dp} and h_s^{dp} are

$$A_{dp} > \overline{\theta} - \underline{\theta} + \frac{\delta^2}{1-\delta} v^{\mathrm{EIR}} - \delta\mathbb{E}[h_b^{dp}(\tilde{\theta}_s) + h_s^{dp}(\tilde{\theta}_b)] \tag{25}$$

$$A_{wd} > A_{dp} - \left(v^{\mathrm{GET}}(\boldsymbol{\theta}^{dp}) + h_b^{dp}(\theta_s^{dp}) + h_s^{dp}(\theta_b^{dp}) - \frac{\delta}{1-\delta} v^{\mathrm{EIR}}\right). \tag{26}$$

Then, for any A_{dp} and A_{wd} that satisfy (25–26), there exists a mechanism with monetary payments and continuation reward functions defined above, that satisfies efficiency in every period, ex post IR, ex post IC, and BBA (as verified by the fact that A_{wd} is finite, by our analysis of transitions between the two regimes, and by the fact that only deposits occur in the deposit regime and only withdrawals occur in the withdrawal regime). Furthermore, the promise keeping and coherence requirements are satisfied by construction, while account keeping follows from Lemma 2.

A.3 *Proof of Corollary 1*

Since Condition 1 holds, there exists a recursive mechanism γ_a satisfying the properties in Proposition 4; by Remark 1 (in Section A.5), one of these properties is the existence of a unique ergodic distribution over account balances, denoted Ψ. Set the initial account balance at the ergodic expected account balance, $\bar{A} \equiv \int_{[0,A_p]} A\, d\Psi$. Choose an initial randomization over deposits and withdrawals prior to the first period of the mechanism according to Ψ. Since Ψ is the ergodic distribution, the initial account balance, the expected immediate post-randomization account balance, and the expected account balance in every subsequent period are all equal to $\bar{A}$ (where expectations are taken over both the initial randomization and all possible realizations of play under the mechanism). By account keeping, the traders' expected average joint payoffs thus satisfy $v_s + v_b = v^{\mathrm{ETU}}$, completing the proof.

A.4 *Proof of Proposition 5*

As discussed in the text, when Condition 1 holds for some $\delta < 1$, the result follows directly from Proposition 4; when Condition 1 fails for all $\delta < 1$, Condition 2 must hold for some $\delta < 1$. For the remainder of the proof we thus assume that Condition 2 holds.

This proof proceeds in several steps. We first describe the mechanism in terms of primitives, and then prove that it satisfies the basic properties required: ex post IC, ex post IR, BBA, coherence, promise keeping, and account keeping. Because there are four regimes, describing transitions among regimes is somewhat complicated, and a number of cases must be checked to establish that the mechanism is feasible.

Our next result establishes ergodicity, a result that is complicated by the fact that the distribution over new account balances starting at a particular point has atoms, and further the transition probabilities are discontinuous as a function of the initial account balances at the points where the regimes switch.

The most complex step, then, is to establish approximate efficiency. To do so we proceed as follows. First, we subdivide the regions of potential account balances, so that transitions occur only within the same region or to neighboring regions. Next, we establish some bounds on the relative likelihood of various regions. Finally, we show that the expected fraction of periods spent in the revenue regime and the payout regime can be made arbitrarily small by choosing $\delta < 1$ sufficiently high.

The mechanism The recursive account mechanism we construct consists of four regimes, $m \in \{r, u, d, p\}$, as follows: the revenue regime ($m = r$, used when $0 \le A \le A_r$), the efficiency regime with upward drift ($m = u$, used when $A_r < A \le A_u$), the efficiency regime with downward drift ($m = d$, used when $A_u < A \le A_d$), and the payout regime ($m = p$, used when $A_d < A \le A_p$). The revenue regime entails inefficient trade and always raises money for the joint account, while the payout regime entails no trade and always reduces the balance. The two efficiency regimes entail efficient trade, but differ in the fixed components of the payment functions, so that the ex ante expected sum of payments is either positive (in the upward drift regime) or negative (in the downward drift regime).

Each regime can be described by a stage mechanism, where, if m is the appropriate regime for A, $g(\cdot; A, \mathbf{v}) = \langle q^m(\cdot), t^m(\cdot; A), w^m(\cdot; A)\rangle$. The allocation function takes the form $q_b^m(\boldsymbol{\theta}) = \mathbf{1}\{\theta_b \ge \lambda_m(\theta_s)\}$, where λ_m is a nondecreasing function, as analyzed in Lemma 1. In the efficiency regimes, $\lambda_u(\theta_s) = \lambda_d(\theta_s) = \theta_s$, while in the payout regime, $\lambda_p(\theta_s) = \theta_s + \overline{\theta} - \underline{\theta}$ (so that trade never occurs). For the revenue regime, we require that λ_r be strictly increasing and satisfy (8). Note that $v^{\mathrm{NGR}}(\boldsymbol{\theta}) \ge 0$ by definition of $q_b^r(\boldsymbol{\theta})$, which in turn implies that

$$q_b^r(\boldsymbol{\theta}) \cdot \left(\lambda_r^{-1}(\theta_b) - \lambda_r(\theta_s)\right) \ge -(\overline{\theta} - \underline{\theta}), \tag{27}$$

a property that we use below.

The critical levels of the balance that determine which regime is used are:

$$A_r = \overline{\theta} - \underline{\theta} - \mathbb{E}[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}})] - (\mathbb{E}[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})]) \tag{28}$$

$$A_u = A_r + \overline{\theta} - \underline{\theta} + \max\left\{0,\ \frac{1}{1-\delta}\mathbb{E}[2\delta v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - (\delta+1)v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})]\right\} + \frac{C_u}{1-\delta} \tag{29}$$

$$A_d = A_u + \mathbb{E}[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}})] + \frac{1+\delta}{1-\delta}\mathbb{E}[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})] + \frac{C_d}{1-\delta} \tag{30}$$

$$A_p = A_d + \mathbb{E}[v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})]$$

for some positive constants C_u and C_d. Note that $A_r > 0$ since (10) holds, and $A_p > A_d > A_u > A_r$.

For the continuation reward functions, where $\mathbf{v}^m \equiv (v_b^m, v_s^m)$, let

$$\begin{aligned}
v_b^r &= (1-\delta)\mathbb{E}[q_b^r(\tilde{\boldsymbol{\theta}})(\tilde{\theta}_b - \lambda_r(\tilde{\theta}_s))]\\
v_s^r &= \mathbb{E}[\tilde{\theta}_s] + \mathbb{E}[q_b^r(\tilde{\boldsymbol{\theta}})(\lambda_r^{-1}(\tilde{\theta}_b) - \tilde{\theta}_s)] + \delta\,\mathbb{E}[q_b^r(\tilde{\boldsymbol{\theta}})(\tilde{\theta}_b - \lambda_r(\tilde{\theta}_s))]\\
v_b^u &= (1-\delta)\mathbb{E}[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}})]\\
v_s^u &= \mathbb{E}[\tilde{\theta}_s] + \mathbb{E}[v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})] - (1-\delta)\mathbb{E}[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}})]\\
v_b^d &= (1-\delta)\mathbb{E}[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}})]\\
v_s^d &= \mathbb{E}[\tilde{\theta}_s] + \delta\mathbb{E}[v^{\mathrm{GFT}}(\tilde{\boldsymbol{\theta}})] + \mathbb{E}[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})]\\
v_b^p &= (1-\delta)\delta\mathbb{E}[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}})]\\
v_s^p &= \mathbb{E}[\tilde{\theta}_s] + (1-\delta+\delta^2)\mathbb{E}[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}})] + \mathbb{E}[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})].
\end{aligned}$$

Note that $v_b^r + v_s^r = v_b^u + v_s^u$ and $v_b^p + v_s^p = v_b^d + v_s^d$. The equilibrium set V_a is given by (11).

We next define payments in terms of the continuation reward functions; later, when we check feasibility and bounded budget account, we define the relevant continuation reward functions. For $m \in \{r, u, d\}$, let

$$\begin{aligned}
t_b^m(\boldsymbol{\theta}; A) &= -q_b^m(\boldsymbol{\theta}) \cdot \lambda_m(\theta_s) - \frac{\delta}{1-\delta} w_b^m(\boldsymbol{\theta}; A)\\
t_s^m(\boldsymbol{\theta}; A) &= q_b^m(\boldsymbol{\theta}) \cdot \lambda_m^{-1}(\theta_b) + \frac{\delta}{1-\delta}\mathbb{E}[\tilde{\theta}_s] - \frac{\delta}{1-\delta} w_s^m(\boldsymbol{\theta}; A)\\
&\quad - \mathbf{1}\{m = u\}(\mathbb{E}[2v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})]) + \mathbf{1}\{m \in \{r, u\}\}\frac{\delta}{1-\delta}\mathbb{E}[v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})]\\
&\quad + \mathbf{1}\{m = d\}\frac{1}{1-\delta}\mathbb{E}[2\delta v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})].
\end{aligned}$$

It is also useful to separate out the part of the payment functions that does not depend on the continuation reward functions. Thus, for $i \in \{b, s\}$, $m \in \{r, u, d\}$, let

$$\hat{t}_i^m(\boldsymbol{\theta}; A) = t_i^m(\boldsymbol{\theta}; A) + \frac{\delta}{1-\delta} w_i^m(\boldsymbol{\theta}; A).$$

For $m = p$, let $t_b^p(\boldsymbol{\theta}; A) = 0$ and $t_s^p(\boldsymbol{\theta}; A) = \mathbb{E}[2v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})]$.

Incentive compatibility and individual rationality Note that given the proposed allocation rules, regardless of the continuation reward functions, by Lemma 1 these payments guarantee that ex post IC is satisfied. It is possible also to verify ex post IR at this point. For the buyer, this requires that (using definitions)

$$\frac{1}{1-\delta}U_b(\boldsymbol{\theta}, g^m(\cdot;A)) = q_b^m(\boldsymbol{\theta})(\theta_b - \lambda_m(\theta_s)) \geq 0. \tag{31}$$

By the definitions of $q_b^m(\boldsymbol{\theta})$ for each m, this always holds.

The seller's ex post IR constraint requires that for each m,

$$\frac{1}{1-\delta}U_s(\boldsymbol{\theta}, g^m(\cdot;A)) \geq \theta_s + \frac{\delta}{1-\delta}\mathbb{E}[\tilde{\theta}_s].$$

Substituting in definitions, for each $m \in \{r, u, d\}$ there is a K_m such that

$$\frac{1}{1-\delta}U_s(\boldsymbol{\theta}, g^m(\cdot;A)) = \theta_s + q_b^m(\boldsymbol{\theta}) \cdot (\lambda_m^{-1}(\theta_b) - \theta_s) + \frac{\delta}{1-\delta}\mathbb{E}[\tilde{\theta}_s] + K_m, \tag{32}$$

which is greater than $\theta_s + \frac{\delta}{1-\delta}\mathbb{E}[\tilde{\theta}_s]$ for all $\boldsymbol{\theta}$ if and only if $K_m \geq 0$. Substituting in the definitions of t_s^m for $m \in \{r, u, d\}$ reveals that

$$\begin{aligned} K_r &= \frac{\delta}{1-\delta}\mathbb{E}\left[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right] \\ K_u &= -\mathbb{E}\left[2v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right] + \frac{\delta}{1-\delta}\mathbb{E}\left[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right] \\ K_d &= \frac{1}{1-\delta}\mathbb{E}\left[2\delta v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right]. \end{aligned}$$

Thus K_r is always non-negative. When

$$\delta \geq 1 - \frac{\mathbb{E}\left[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right]}{2\mathbb{E}\left[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})\right]}, \tag{33}$$

$K_u \geq 0$ is guaranteed, and since (33) also requires $\delta > \frac{1}{2}$, $K_d \geq 0$ as well. Finally, consider the seller's ex post IR constraint for $m = p$. Below, we assign $w_s^p(\boldsymbol{\theta}; A) = v_s^d$. Then,

$$\frac{1}{1-\delta}U_s(\boldsymbol{\theta}, g^p(\cdot;A)) = \theta_s + \mathbb{E}[2v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})] + \frac{\delta}{1-\delta}v_s^d. \tag{34}$$

Since $v_s^d \geq \mathbb{E}[\tilde{\theta}_s]$, the ex post IR constraints are satisfied.

Promise keeping and account keeping Using the expressions for payoffs in (31), (32), and (34) and taking expectations, it is also straightforward to verify that the value set V_a is given by (11), satisfying the promise keeping requirement. Account keeping follows from Lemma 2.

Coherence and bounded budget account We now consider the relationship between continuation rewards and the account balance, taking care to ensure that transitions from one regime to another are consistent with the coherence condition. With some abuse of notation, we write A_{m-1} or A_{m+1} to indicate $A_{m'}$, where m' is the regime corresponding to the interval of balances just below or just above the interval for regime m, respectively. Let $A_{r-1}=0$. When we say that starting from m, given $(\boldsymbol{\theta};A)$ "a transition occurs to regime m'," we mean that we assign $w_b^m(\boldsymbol{\theta};A)=v_b^{m'}$ and $w_s^m(\boldsymbol{\theta};A)=v_s^{m'}$. Consistency of the definitions requires that when we assign a transition to state m',

$$\begin{aligned}A_{m'-1} < \alpha(\boldsymbol{\theta};A,\mathbf{v}^m) &= A-(t_b^m(\boldsymbol{\theta};A)+t_s^m(\boldsymbol{\theta};A))\\ &= A-(\hat{t}_b^m(\boldsymbol{\theta};A)+\hat{t}_s^m(\boldsymbol{\theta};A))+\frac{\delta}{1-\delta}(v_b^{m'}+v_s^{m'})\le A_{m'}.\end{aligned}\tag{35}$$

In such a case, we say that "for this $(\boldsymbol{\theta};A)$, a transition to m' is consistent."

Notice that there may be multiple regimes m' where such a consistency requirement is satisfied, given that values differ across regimes. Thus, we specify a hierarchy on the transitions. Formally, we assign continuation rewards as follows: for $m=p$, $w_b^p(\boldsymbol{\theta};A)=v_b^d$ and $w_s^p(\boldsymbol{\theta};A)=v_s^d$. For $m\in\{r,u,d\}$, starting from $(\boldsymbol{\theta};A)$ where $A_{m-1}<A\le A_m$ (so that we start from regime m), if a transition to m is consistent, then $w_b^m(\boldsymbol{\theta};A)=v_b^m$ and $w_s^m(\boldsymbol{\theta};A)=v_s^m$; otherwise, if a transition to $m+1$ is consistent, then $w_b^m(\boldsymbol{\theta};A)=v_b^{m+1}$ and $w_s^m(\boldsymbol{\theta};A)=v_s^{m+1}$; otherwise, if $m>r$ and a transition to $m-1$ is consistent, then $w_b^m(\boldsymbol{\theta};A)=v_b^{m-1}$ and $w_s^m(\boldsymbol{\theta};A)=v_s^{m-1}$. Note that by construction, any continuation rewards assigned using this hierarchy lie in the set V_a defined in (11), and so feasibility holds.

It remains to check that this hierarchy assigns unique continuation rewards for all $(\boldsymbol{\theta};A)\in V_a$. Our analysis of the transitions makes use of the fact that starting from a given regime, the realizations of payments differ only in the term $\frac{\delta}{1-\delta}(v_b^{m'}+v_s^{m'})$, which (using definitions) is higher for higher m'.

Transitions starting from the revenue regime: $A\le A_r$ Anticipating a transition to either the revenue regime or the upward-drift regime (the only two allowed by our definitions) the sum of payments can be written as

$$t_b^r(\boldsymbol{\theta};A)+t_s^r(\boldsymbol{\theta};A)=q_b^r(\boldsymbol{\theta})\cdot(\lambda_r^{-1}(\theta_b)-\lambda_r(\theta_s)).$$

Our restriction (8) implies that this sum of payments is always non-positive. Our assignment rule implies that the transition is to remain in the revenue regime if

$$0\le A-q_b^r(\boldsymbol{\theta})\cdot(\lambda_r^{-1}(\theta_b)-\lambda_r(\theta_s))\le A_r,\tag{36}$$

and otherwise the transition is to the upward-drift regime. Using (27), the minimum possible sum of payments (corresponding to the maximum possible deposit in the account) is $-(\overline{\theta}-\underline{\theta})$; thus, starting from $A\le A_r$, the largest possible subsequent account balance is $A_r+(\overline{\theta}-\underline{\theta})$, which is less than A_u. Thus, for all $(\boldsymbol{\theta};A)$ such that $A\le A_r$, we have assigned consistent continuation rewards.

Transitions starting from the upward-drift regime: $A_r < A \le A_u$ In the upward-drift regime,

$$\hat{t}_b^u(\boldsymbol{\theta};A)+\hat{t}_s^u(\boldsymbol{\theta};A)=v^{\text{GET}}(\boldsymbol{\theta})-\mathbb{E}[2v^{\text{GET}}(\tilde{\boldsymbol{\theta}})-v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]+\frac{\delta}{1-\delta}\mathbb{E}[\tilde{\theta}_s+v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})].$$

The transition is to remain in the upward-drift regime if, using (35),

$$\begin{aligned}A_r &< A-\big(\hat{t}_b^u(\boldsymbol{\theta};A)+\hat{t}_s^u(\boldsymbol{\theta};A)\big)+\frac{\delta\big(v_b^u+v_s^u\big)}{1-\delta}\\ &=A-\big(v^{\text{GET}}(\boldsymbol{\theta})-\mathbb{E}[2v^{\text{GET}}(\tilde{\boldsymbol{\theta}})-v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]\big)\le A_u.\end{aligned}\tag{37}$$

If the latter condition fails, we next consider a transition to the downward-drift regime, which happens if

$$A_u < A-\Big(v^{\text{GET}}(\boldsymbol{\theta})-\mathbb{E}[2v^{\text{GET}}(\tilde{\boldsymbol{\theta}})-v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]-\frac{2\delta}{1-\delta}\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})-v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]\Big)\le A_d.\tag{38}$$

Notice that the middle term of (38) is maximized when $v^{\text{GFT}}(\boldsymbol{\theta})=0$, so that the maximum balance attainable starting from $A \le A_u$ is

$$A_u+\frac{1}{1-\delta}\mathbb{E}[2v^{\text{GET}}(\tilde{\boldsymbol{\theta}})-(1+\delta)v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]\le A_d,$$

which can be directly verified using the definition of A_d. Thus, if the first inequality in (37) holds, then either the last inequality in (37) holds as well, or else (using the fact that $v_b^d+v_s^d > v_b^u+v_s^u$) both inequalities in (38) hold. Therefore, it remains only to verify transitions when the first inequality in (37) fails. A transition occurs to the revenue regime if

$$0\le A-\big(\hat{t}_b^u(\boldsymbol{\theta};A)+\hat{t}_s^u(\boldsymbol{\theta};A)\big)+\frac{\delta}{1-\delta}\big(v_b^r+v_s^r\big)\le A_r.\tag{39}$$

Since $v_b^r+v_s^r=v_b^u+v_s^u$, the second inequality in (39) holds if and only if the first inequality in (37) fails. We can verify that the first inequality in (39) holds, so that the account balance stays positive, by noting that at worst,

$$\begin{aligned}A-(t_b^u(\boldsymbol{\theta};A)+t_s^u(\boldsymbol{\theta};A)) &= A-\big(\overline{\theta}-\underline{\theta}-\mathbb{E}[2v^{\text{GET}}(\tilde{\boldsymbol{\theta}})-v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]\big)\\ &=A-A_r,\end{aligned}$$

which is non-negative since $A\ge A_r$ in the upward-drift regime.

Transitions starting from the downward-drift regime: $A_r < A \le A_u$ Using definitions,

$$\hat{t}_b^d(\boldsymbol{\theta};A)+\hat{t}_s^d(\boldsymbol{\theta};A)=v^{\text{GET}}(\boldsymbol{\theta})-\frac{1}{1-\delta}\mathbb{E}[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]+\frac{\delta}{1-\delta}\mathbb{E}[\tilde{\theta}_s+2v^{\text{GET}}(\tilde{\boldsymbol{\theta}})].$$

A transition to remain in the downward-drift regime occurs if

$$A_u < A-(\hat{t}_b^d(\boldsymbol{\theta};A)+\hat{t}_s^d(\boldsymbol{\theta};A))+\frac{\delta\big(v_b^d+v_s^d\big)}{1-\delta}=A-\big(v^{\text{GET}}(\tilde{\boldsymbol{\theta}})-\mathbb{E}[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]\big)\le A_d.\tag{40}$$

If the second inequality in (40) fails, using the fact that $v_b^d + v_s^d = v_b^p + v_s^p$, transition to the payout regime occurs if

$$A_d < A - \left(v^{\text{GET}}(\boldsymbol{\theta}) - \mathbb{E}\left[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right]\right) \leq A_p. \tag{41}$$

Note that by the definition of A_p, the second inequality in (41) never fails. Thus, if (41) fails, the first inequality in (40) must then fail; if so, given that $v_b^u + v_s^u < v_b^d + v_s^d$, when the first inequality in (40) fails,

$$A - (\hat{t}_b^d(\boldsymbol{\theta}; A) + \hat{t}_s^d(\boldsymbol{\theta}; A)) + \frac{\delta}{1-\delta}\left(v_b^u + v_s^u\right) \leq A_u.$$

Thus, to verify consistency when (40) and (41) fail, whereby a transition to the upward-drift regime is called for, it only remains to show that

$$\begin{aligned} A_r &< A - (\hat{t}_b^d(\boldsymbol{\theta}; A) + \hat{t}_s^d(\boldsymbol{\theta}; A)) + \frac{\delta}{1-\delta}\left(v_b^u + v_s^u\right) \\ &= A - \left(v^{\text{GET}}(\boldsymbol{\theta}) + \frac{1}{1-\delta}\mathbb{E}\left[2\delta v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - (\delta+1)v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right]\right). \end{aligned} \tag{42}$$

This is most difficult to satisfy when $v^{\text{GET}}(\boldsymbol{\theta}) = \overline{\theta} - \underline{\theta}$ and $A = A_u$. The definition of A_u guarantees that (42) holds.

Transitions from the payout regime: $A_d < A \leq A_p$ Given such an A, for all $\boldsymbol{\theta} \in \Theta$ the sum of payments is

$$t_b^p(\boldsymbol{\theta}; A) + t_s^p(\boldsymbol{\theta}; A) = \mathbb{E}[2v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})].$$

Thus, even in the most challenging case where $A = A_p = A_d + \mathbb{E}[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]$,

$$\begin{aligned} A - (t_b^p(\boldsymbol{\theta}; A) + t_s^p(\boldsymbol{\theta}; A)) &= A_d + \mathbb{E}\left[v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right] - \mathbb{E}\left[2v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right] \\ &= A_d - 2\mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})] < A_d. \end{aligned}$$

In addition, we can guarantee that $A_u \leq A - (t_b^p(\boldsymbol{\theta}; A) + t_s^p(\boldsymbol{\theta}; A))$, since in the most challenging case, where $A = A_d$,

$$A_u \leq A_d - \mathbb{E}\left[2v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right],$$

which follows since

$$\begin{aligned} &A_d - A_u - \mathbb{E}\left[2v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right] \\ &\quad = \mathbb{E}\left[v^{\text{GET}}(\tilde{\boldsymbol{\theta}})\right] + \frac{1+\delta}{1-\delta}\mathbb{E}\left[v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right] - \mathbb{E}[2v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})] + \frac{C_d}{1-\delta} \\ &\quad = \frac{2\delta}{1-\delta}\mathbb{E}\left[v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})\right] + \frac{C_d}{1-\delta} > 0. \end{aligned}$$

Thus, we have verified that a consistent continuation reward is assigned for all $(\boldsymbol{\theta}; A)$, so coherence is satisfied. Since the account balance is always in the interval $[0, A_p]$, BBA is satisfied as well.

Ergodicity Given α and $(A,\mathbf{v}^m)\in V_a$, let

$$G(\hat{A};A)\equiv\int_{\boldsymbol{\theta}\in\Theta}\mathbf{1}\{\alpha(\boldsymbol{\theta};A,\mathbf{v}^m)\leq\hat{A}\}\,dF(\boldsymbol{\theta})$$

be the probability that $\alpha(\boldsymbol{\theta};A,\mathbf{v}^m)$ is less than $\hat{A}$ in the period following an account balance of A. (Note that this depends on δ; we leave δ out of the notation for simplicity.)

We will demonstrate that the Markov process on $A\in[0,A_p]$ that is induced by our mechanism is uniformly ergodic, and hence a unique ergodic probability measure Ψ exists and the probability measure over the account balance in period t converges to Ψ uniformly at a geometric rate as $t\to\infty$.

Lemma 5. *The Markov process on* $[0,A_p]$ *induced by our mechanism is uniformly ergodic, and* Ψ *is atomless with full support on* $[0,A_p]$.

The proof is in Section A.5. A somewhat complicated approach is necessary to establish this claim, since the presence of discontinuities and atoms in the transition process implies that the process violates properties that are often used to show existence of an invariant distribution, such as the Feller property.[23] To this end, we invoke the powerful Theorem 16.0.2 from Meyn and Tweedie (1993), hereafter M&T.[24] We show also that Ψ is atomless and has full support on $[0,A_p]$, by employing M&T Theorem 10.4.9.

Near-efficiency The next step in the proof is to show that as $\delta\to 1$, efficient allocation is used arbitrarily often. We begin by defining two subintervals of the upward drift regime: U_0 and $U_1(\delta)$, where U_0 is adjacent to the revenue regime and $U_1(\delta)$ is adjacent to U_0. Next we show that the ergodic probability of the revenue regime is bounded by a multiple of the ergodic probability of U_0. We then characterize the ergodic density in U_0 and $U_1(\delta)$, and identify a lower bound on this density, as an increasing function of the minimum density in U_0. Since $U_1(\delta)$ grows without bound as $\delta\to 1$, this lower bound requires the minimal density in U_0 to converge to zero. Finally, we demonstrate that if the minimal density in U_0 converges to zero, then the ergodic probability of U_0 must converge to zero as well, and hence the ergodic probability of the revenue regime converges to zero. The argument for the payout regime is analogous and omitted.

Subdividing the upward drift regime First we construct a subinterval U_0 of the upward drift regime, such that any transition from the revenue regime to the upward drift regime is a transition to U_0, and any transition from the upward drift regime to the revenue regime is a transition from U_0.

Within the upward-drift regime, the sum of payments when the traders transition back into the regime is

$$v^{\mathrm{GET}}(\boldsymbol{\theta})-\mathbb{E}[2v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}})-v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})]. \tag{43}$$

[23]See Meyn and Tweedie (1993, Section 6.1).

[24]We are indebted to Sean Meyn for suggesting a productive approach for applying results from M&T to this problem.

Thus, a transition out of the upward drift regime into the revenue regime is possible only if the starting balance A satisfies

$$A_r < A \leq A_r + \overline{\theta} - \underline{\theta} - \mathbb{E}\left[2v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})\right] < A_r + \overline{\theta} - \underline{\theta}.$$

When transitioning from the revenue regime to the upward-drift regime, the increase in the account balance is always less than or equal to $\overline{\theta} - \underline{\theta}$. Accordingly, we can choose $\varepsilon \in \left(0, \frac{1}{1-\delta} C_u\right)$ small, and let $U_0 \equiv \left(A_r, A_r + \overline{\theta} - \underline{\theta} + \varepsilon\right) \equiv \left(A_r, \overline{U}_0\right)$, so that any transition from the revenue regime into the upward drift regime is a transition into U_0, and any transition from the upward drift regime into the revenue regime is a transition from U_0.

We now construct a subinterval $U_1(\delta)$ of the upward drift regime, such that any transition from U_0 back into the upward drift regime regime is a transition to $U_0 \cup U_1(\delta)$, and such that no transition from the downward drift regime is ever a transition into $U_0 \cup U_1(\delta)$.

Recall that C_u is a parameter of the mechanism which, along with δ, determines the length of the upward-drift regime. Define $U_1(\delta) \equiv \left[\overline{U}_0, A_r + \frac{1}{1-\delta} C_u\right) \equiv \left[\overline{U}_0, \overline{U}_1(\delta)\right)$. Fix $C_u > 0$ sufficiently large that any transition from U_0 back into the upward drift regime is a transition to $U_0 \cup U_1(\delta)$:

$$\overline{U}_0 + \overline{\theta} - \underline{\theta} - \mathbb{E}\left[2v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})\right] \leq A_r + \frac{1}{1-\delta} C_u.$$

Given the definition of A_u in (29), no transition from the downward drift regime can occur into $U_1(\delta)$, by (42). Note that by the construction of U_0, above, no transition from $U_1(\delta)$ can occur into the revenue regime.

Bounding the probability of the revenue regime Let $R \equiv [0, A_r]$ be the region of account balances in the revenue regime. Since transitions from R go only back to R or to U_0,

$$\Psi(U_0) > \Psi(R) - \int_{A \in R} G(A_r; A)\, d\Psi(A),$$

where the last term represents the probability of remaining in R having started there. This implies that $\Psi(U_0) > \Psi(R)\left(1 - \sup_{A \in R} G(A_r; A)\right)$ and thus the ratio $\Psi(U_0)/\Psi(R)$ is bounded below by $\Psi(U_0)/\Psi(R) \geq 1 - \sup_{A \in R} G(A_r; A) \equiv \sigma_0 > 0$ for all δ, since $\sup_{A \in R} G(A_r; A) < 1$ (see (28) and (36)). By the definitions of the monetary payment functions, $G(A_r; A)$ does not depend on δ for $A \in R$.

Characterizing the ergodic density in the upward drift regime This section derives a simple expression for the ergodic density in the upward drift regime (45, below).

Consider the region $U_1(\delta)$, and let $G^*(\cdot) \equiv G(\cdot + \hat{A}; \hat{A})$ represent the probability distribution over changes in the account balance in this region; G^* is well-defined since $G(\cdot + \hat{A}; \hat{A})$ does not vary with $\hat{A}$ in this region (nor does it vary with δ).

Note that G^* is the cumulative distribution function of the random change in the account balance on $U_1(\delta)$. Using our definitions of monetary payment functions, this random variable is a function of $\boldsymbol{\theta}$, and can be written as $y \equiv -v^{\mathrm{GET}}(\boldsymbol{\theta}) + \mathbb{E}\left[2v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) -\right.$

$v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})]$. We write the support of G^* as $[\underline{y}, \overline{y}]$, with $\overline{y} \equiv \mathbb{E}[2v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})]$ and $\underline{y} \equiv \overline{y} - (\overline{\theta} - \underline{\theta})$. It also follows from our definitions that G^* can be represented as a mixture of a density component on the interval $[\underline{y}, \overline{y})$ (corresponding to the realizations of $\boldsymbol{\theta}$ for which trade occurs) with an atom at $\overline{y}$ (corresponding to the realizations of $\boldsymbol{\theta}$ for which no trade occurs). Note that the definition of G^* implies that

$$\int_{y \in [\underline{y}, \overline{y}]} y \, dG^*(y) = \mathbb{E}\big[\mathbb{E}[2v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})] - v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}})\big] = \mathbb{E}\big[v^{\mathrm{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\mathrm{NGR}}(\tilde{\boldsymbol{\theta}})\big] > 0. \tag{44}$$

At the ergodic distribution, transitions across any boundary are equalized, so that

$$\begin{aligned}\int_{\hat{A} \in (A, A - \underline{y}]} G(A; \hat{A}) \, d\Psi(\hat{A}) &= \int_{\hat{A} \in (A - \overline{y}, A]} (1 - G(A; \hat{A})) \, d\Psi(\hat{A}) = \int_{\hat{A} \in (A - \overline{y}, A]} d\Psi(\hat{A}) - \int_{\hat{A} \in (A - \overline{y}, A]} G(A; \hat{A}) \, d\Psi(\hat{A}) \\ &= \Psi\big((A - \overline{y}, A]\big) - \int_{\hat{A} \in (A - \overline{y}, A]} G(A; \hat{A}) \, d\Psi(\hat{A}),\end{aligned}$$

and, consequently,

$$\Psi\big((A - \overline{y}, A]\big) = \int_{\hat{A} \in (A - \overline{y}, A - \underline{y}]} G^*(A - \hat{A}) \, d\Psi(\hat{A}).$$

Then differentiating with respect to A yields

$$\psi(A) - \psi(A - \overline{y}) = G^*(\underline{y}) \psi(A - \underline{y}) - G^*(\overline{y}) \psi(A - \overline{y}) + \int_{y \in [\underline{y}, \overline{y})} \psi(A - y) \, dG^*(y),$$

where $\psi \equiv d\Psi([0, A])/dA$ and dG^* is the differential for the probability measure implied by G^*.[25] Since $G^*(\underline{y}) = 0$ and $G^*(\overline{y}) = 1$, this simplifies to

$$\psi(A) = \int_{y \in [\underline{y}, \overline{y})} \psi(A - y) \, dG^*(y). \tag{45}$$

Bounding the density in the upward drift regime We now construct a lower bound on the ergodic density ψ in the region $U(\delta) \equiv U_0 \cup U_1(\delta)$; this is the part of the upward drift regime that can be transitioned into only from the upward drift and revenue regimes. The ergodic density of account balances on $U(\delta)$ must lie above a lower bound that decreases linearly as the account balance increases. The idea is that the ergodic density cannot decrease with the account balance too fast in regions where transitions in expectation shift probability mass toward higher account balances.

Recall that Ψ and ψ depend on δ; now we write Ψ_δ and ψ_δ to make this clear. Recall also that $\overline{U}_1(\delta)$ is the upper bound of $U_1(\delta)$.

[25] The density representation ψ exists by the Radon–Nikodym theorem, since Ψ is equivalent (in the terminology of M&T) to μ. (Recall that μ is the uniform distribution on $[0, A_p]$. The result follows from the fact, previously established, that the ergodic distribution is atomless with full support.)

LEMMA 6. *For all $A \in U(\delta)$,*

$$\psi_\delta(A) > \frac{\overline{U}_1(\delta) - A}{\overline{U}_1(\delta) - A_r} \inf_{A \in U_0} \{\psi_\delta(A)\}.$$

PROOF. Suppose to the contrary that there exists $A \in U(\delta)$ such that

$$\psi_\delta(A) \leq \frac{\overline{U}_1(\delta) - A}{\overline{U}_1(\delta) - A_r} \inf_{A \in U_0} \{\psi_\delta(A)\}.$$

Note that, for all $\delta < 1$, $\inf_{A \in U(\delta)}\{\psi_\delta(A)\} > 0$: since ψ is atomless with full support, and since G^* has a density component, starting from any small open interval in $U(\delta)$ with positive Ψ-measure, the probability of reaching any other open interval in $U(\delta)$ must be positive given a sufficiently large number of steps. (A more detailed version of this argument is used below in Lemma 8.)

Let

$$p^* = \sup\left\{ p : \frac{\overline{U}_1(\delta) - A}{\overline{U}_1(\delta) - A_r} p \leq \psi(A) \text{ for all } A \in U(\delta) \right\}.$$

Now, define the density[26] $\underline{\psi}_\delta$ as follows.

(i) $\underline{\psi}_\delta(A) = \psi_\delta(A)$ for all $A < A_r$.

(ii) $\underline{\psi}_\delta(A_r) = \inf_{A \in U_0} \psi_\delta(A)$.

(iii) $\underline{\psi}_\delta(A) = 0$ for all $A \geq \overline{U}_1(\delta)$.

(iv) For A in the interior of $U(\delta)$,

$$\underline{\psi}_\delta(A) = \frac{\overline{U}_1(\delta) - A}{\overline{U}_1(\delta) - A_r} p^*.$$

Note that by construction, $\underline{\psi}_\delta(A) \leq \psi_\delta(A)$ for all $A \in U(\delta)$, and, for any $\varepsilon > 0$, there exists $A^* \in U_1(\delta)$ such that

$$\underline{\psi}_\delta(A^*) + \varepsilon > \psi_\delta(A^*). \tag{46}$$

Choose $0 < \varepsilon < p^* \mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})] / (\overline{U}_1(\delta) - A_r)$. Since $A^* - \overline{y} > A_r$, we have the following contradiction:

$$\begin{aligned}
\psi_\delta(A^*) &= \int_{y \in (\underline{y}, \overline{y}]} \psi_\delta(A^* - y)\, dG^*(y) \geq \int_{y \in (\underline{y}, \overline{y}]} \underline{\psi}_\delta(A^* - y)\, dG^*(y) \geq \underline{\psi}_\delta\left(\int_{y \in (\underline{y}, \overline{y}]} (A^* - y)\, dG^* \right) \\
&= \underline{\psi}_\delta\big(A^* - \mathbb{E}[v^{\text{GET}}(\tilde{\boldsymbol{\theta}}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]\big) > \underline{\psi}_\delta(A^*) + \varepsilon > \psi_\delta(A^*),
\end{aligned}$$

where the first relation is by (45), the second relation is by construction of $\underline{\psi}_\delta$, the third relation is by the convexity of $\underline{\psi}_\delta$ on $[A_r, A_p]$, the fourth relation is by (44), the fifth relation follows from the definitions of $\underline{\psi}_\delta$ and of ε, and the final relation follows from $\varepsilon > 0$ and (46). □

[26] This is not a probability density, since it integrates to less than 1.

This lemma implies that $\Psi_\delta(U(\delta)) > \frac{1}{2}\big(\overline{U}_1(\delta) - A_r\big)\inf_{A\in U_0}\{\psi_\delta(A)\}$, since there is a linear lower bound on $\psi_\delta(A)$. Since $\overline{U}_1(\delta) \to \infty$ as $\delta \to 1$, it must be that $\inf_{A\in U_0}\{\psi_\delta(A)\} \to 0$ as $\delta \to 1$. Hence, for any $\varepsilon > 0$, there exists $\underline{\delta} < 1$ such that $\inf_{A\in U_0}\{\psi_\delta(A)\} < \frac{1}{2}\varepsilon$ for all $\delta > \underline{\delta}$, and hence for all $\delta > \underline{\delta}$ there exists $\tilde{A}_\delta \in U_0$ such that $\psi_\delta(\tilde{A}) < \varepsilon$. Fix such an ε.

Since one-step transitions from U_0 to U_0 follow the same rule as any transitions from $U(\delta)$ to $U(\delta)$, and the support of G^* is $\big[\underline{y}, \overline{y}\big]$, we have (using (45)),

$$\varepsilon > \psi_\delta(A) \geq \int_{\{y\in(\underline{y},\overline{y}):A-y\in U_0\}} \psi_\delta(A-y)\, dG^*(y). \tag{47}$$

The following two lemmas let us use (47) to show that $\Psi_\delta(U_0)$ cannot exceed a multiple of $\inf_{A\in U_0}\{\psi_\delta(A)\}$, and therefore $\Psi_\delta(U_0) \to 0$ as $\delta \to 1$.

The first lemma demonstrates that G^* is uniformly bounded away from zero on part of its support. This lemma is purely technical, and unfortunately somewhat cumbersome to notate. The rest of the argument operates in essentially the same way as it would if G^* were uniformly bounded away from zero on all of its support.

Lemma 7. *There exists $\eta > 0$ such that, for all $y \in \big(\frac{1}{2}\underline{y}, \overline{y}\big)$, $\frac{d}{dy}G^*(y) \geq \eta$.*

Proof. From (43), $G^*(y) = \Pr\big\{y \leq \overline{y} - v^{\mathrm{GET}}(\boldsymbol{\theta})\big\}$. Let $\upsilon > 0$ be the square of the uniform lower bound on F_s and F_b. Then, for any $y, y' \in (\frac{1}{2}\underline{y}, \overline{y})$ with $y' > y$,

$$\begin{aligned}\Pr\{y \leq \overline{y} - v^{\mathrm{GET}}(\boldsymbol{\theta}) < y'\} &= \Pr\{\overline{y} - y \geq v^{\mathrm{GET}}(\boldsymbol{\theta}) > \overline{y} - y'\} \\ &\geq \left(\frac{(\overline{\theta} - \underline{\theta} - \overline{y} + y')^2}{2} - \frac{(\overline{\theta} - \underline{\theta} - \overline{y} + y)^2}{2}\right)\upsilon = \frac{1}{2}(y' - y)(y + y' - 2\underline{y})\upsilon,\end{aligned}$$

where the inequality is from the uniform lower bound on the integral of $dF_s\, dF_b$ over the region of Θ specified in the first line, and the final equality is because $\overline{\theta} - \underline{\theta} = \overline{y} - \underline{y}$ by construction. Then

$$\frac{d}{dy}G^*(y) = \lim_{y' \to y} \frac{\Pr\big(y \leq \overline{y} - v^{\mathrm{GET}}(\boldsymbol{\theta}) < y'\big)}{y' - y} \geq (y - \underline{y})\upsilon,$$

and hence $\frac{d}{dy}G^*(y) \geq -\frac{1}{2}\underline{y}\,\upsilon \equiv \eta > 0$ for all $y \in \big(\frac{1}{2}\underline{y}, \overline{y}\big)$. □

Let $\big(G^*\big)^n$ be the cumulative distribution function for n-step transitions within the $U(\delta)$ region. The following lemma demonstrates that the density of $\big(G^*\big)^n$ is uniformly bounded away from zero on part of its support, and that this part of the support grows linearly in n. It makes use of Lemma 7. Like Lemma 7, it is purely technical, and the rest of the argument operates in the same way as it would if $\big(G^*\big)^n$ were uniformly bounded away from zero on all its support.

LEMMA 8. *For any* $n \geq 1$, *there exists* $\eta_n^* > 0$ *such that, for any* $A \in U(\delta)$ *and any* $A' \in [A + \frac{n-1}{2}\underline{y}, A + (n-1)\overline{y}]$, $\frac{d}{dA}(G^*)^n(A') \geq \eta_n^*$.

PROOF. Choose $\eta > 0$ to satisfy the statement of Lemma 7. Define the cumulative distribution function $\hat{G}^*$ by $\hat{G}^*(\frac{1}{2}\underline{y}) = 0$, $\frac{d}{dy}\hat{G}^*(y) = \eta$ for $y \in [\frac{1}{2}\underline{y}, \overline{y}]$, and $\frac{d}{dy}\hat{G}^*(y) = 0$ otherwise, so that (by Lemma 7) $\frac{d}{dy}\hat{G}^*$ is a subdensity of $\frac{d}{dy}G^*$; i.e., $\frac{d}{dy}G^*(y) \geq \frac{d}{dy}\hat{G}^*(y)$ for all y.

Consider the "sub-Markov" process associated with $\hat{G}^*$ (it is just like a Markov process except that its transition density $\frac{d}{dy}\hat{G}^*$ integrates to less than one). Starting from $A \in U(\delta)$, suppose that $A + \frac{n}{2}\underline{y} \in U(\delta)$ as well. Then the process can reach the region $[A + \frac{n}{2}\underline{y}, A + \frac{n-1}{2}\underline{y})$ (for $n \geq 1$) in exactly n steps only if it has reached the region $[A + \frac{n-1}{2}\underline{y}, A + \frac{n-2}{2}\underline{y})$ in $n-1$ steps. Thus the n-step sub-cumulative distribution function $(\hat{G}^*)^n$ must satisfy

$$\begin{aligned}(\hat{G}^*)^n(\tfrac{n}{2}\underline{y} + x) &= \int_{\frac{n-1}{2}\underline{y}}^{\frac{n-1}{2}\underline{y}+x} \left(\int_{\frac{1}{2}\underline{y}}^{\frac{n}{2}\underline{y}+x-y} d\hat{G}^*(\hat{y}) \right) d(\hat{G}^*)^{n-1}(y) \\ &= \int_{\frac{n-1}{2}\underline{y}}^{\frac{n-1}{2}\underline{y}+x} (\tfrac{n-1}{2}\underline{y} + x - y)\eta \, d(\hat{G}^*)^{n-1}(y)\end{aligned}$$

for any $x \in (0, -\frac{1}{2}\underline{y})$. Therefore

$$\frac{d}{dx}(\hat{G}^*)^n(\tfrac{n}{2}\underline{y} + x) = \int_0^x \eta \frac{d}{d\hat{x}}(\hat{G}^*)^{n-1}(\tfrac{n-1}{2}\underline{y} + \hat{x})\, d\hat{x}. \tag{48}$$

Applying (48) inductively to $\frac{d}{dx}\hat{G}^*(\frac{1}{2}\underline{y} + x) = \eta$ (from Lemma 7) yields

$$\frac{d}{dx}(\hat{G}^*)^n(\tfrac{n}{2}\underline{y} + x) = \frac{\eta^n x^{n-1}}{(n-1)!}$$

for any $x \in (0, -\frac{1}{2}\underline{y})$. By similar reasoning, for $x \in (0, \overline{y})$,

$$-\frac{d}{dx}(\hat{G}^*)^n(n\overline{y} - x) = \frac{\eta^n x^{n-1}}{(n-1)!}.$$

Furthermore, since $\frac{d}{dy}\hat{G}^*$ is log-concave on its support, $\frac{d}{dy}(\hat{G}^*)^n$ is also log-concave and thus single peaked[27] and therefore

$$\frac{d}{dy}(\hat{G}^*)^n(y) \geq \frac{\eta^n}{(n-1)!}\min\{\overline{y}^{n-1}, (-\tfrac{1}{2}\underline{y})^{n-1}\} \equiv \eta_n^* > 0.$$

for all $y \in [\frac{n-1}{2}\underline{y}, (n-1)\overline{y}]$. $\square$

[27] The convolution of two log-concave densities is log-concave; see Karlin (1968, pp. 332–333).

Choose an integer n large enough that $U_0 \subset \left(\overline{U}_0 + \frac{n-1}{2}\underline{y}, A_r + (n-1)\overline{y}\right)$. Note that n may be chosen independently of δ. Then, since ψ_δ is the ergodic distribution associated with ξ^n for any $n \geq 1$, we can also write (47) as

$$\varepsilon > \psi_\delta(A) \geq \int_{\left\{y \in \left(\frac{n-1}{2}\underline{y}, (n-1)\overline{y}\right) : A - y \in U_0\right\}} \psi_\delta(A-y)\, d\left(G^*\right)^n, \tag{49}$$

where (recall from (47)) for any $\varepsilon > 0$ there exists $\underline{\delta} < 1$ such that (49) is satisfied for all $\delta > \underline{\delta}$. Since $\frac{d}{dy}\left(G^*\right)^n(y)$ exceeds η_n^* (from Lemma 8) on the entire region of integration,

$$\begin{aligned}
\varepsilon > \psi_\delta(A) &\geq \int_{\left\{y \in \left(\frac{n-1}{2}\underline{y}, (n-1)\overline{y}\right) : A - y \in U_0\right\}} \psi_\delta(A-y)\, d\left(G^*\right)^n \\
&\geq \eta_n^* \left(\int_{\left\{y \in \left(\frac{n-1}{2}\underline{y}, (n-1)\overline{y}\right) : \tilde{A} - y \in U_0\right\}} \psi_\delta(A-y)\, dy \right) = \eta_n^* \left(\int_{A' \in U_0} \psi_\delta(A)\, dA' \right) = \eta_n^* \Psi_\delta(U_0).
\end{aligned}$$

Since we can choose $\underline{\delta} < 1$ high enough to satisfy this condition for any $\varepsilon > 0$ independently of n (and η_n^*), it must be that $\lim_{\delta \to 1} \Psi_\delta(U_0) \to 0$.

Bounding the probability of inefficient allocation This completes the proof, since $\Psi_\delta(R) \leq \frac{1}{\sigma_0}\Psi_\delta(U_0)$ converges to 0 as $\delta \to 1$; a similar but omitted argument implies that $\Psi_\delta(P) \to 0$ as $\delta \to 1$. Thus the probability of inefficient allocation at the stationary distribution Ψ_δ can be made arbitrarily small by sufficiently high choice of $\delta < 1$.

A.5 *Proof of Lemma 5*

Let $\xi(\cdot; A)$ be the probability measure associated with $G(\cdot; A)$; then ξ is the transition function of the Markov process on $A \in [0, A_p]$ that is induced by the mechanism we have defined. We call this Markov process Ξ.

In what follows, let ξ^n be the n-step transition function of Ξ, let μ be the uniform distribution on $[0, A_p]$, let supp$(\cdot)$ be the support operator on measures, and let int$(\cdot)$ be the interior operator on subsets of $[0, A_p]$.

Definition 8 (M&T, p. 106). For an integer $n > 0$, a set $\mathscr{S} \subseteq [0, A_p]$ is ν_n*-small* if there exists a non-trivial Borel measure ν_n such that $\xi^n(\mathscr{B}; A) \geq \nu_n(\mathscr{B})$ for all $A \in \mathscr{S}$ and any Borel set $\mathscr{B}$. $\mathscr{S}$ is *small* if $\mathscr{S}$ is ν_n-small for some n.

Definition 9 (M&T Proposition 4.2.1). A measure $\hat{\mu}$ is an *irreducibility measure* for Ξ if for all $A \in [0, A_p]$ and for any $\mathscr{B} \subseteq [0, A_p]$ with $\hat{\mu}(\mathscr{B}) > 0$, there exists an $n < \infty$ such that $\xi^n(\mathscr{B}; A) > 0$.

Lemma 9. *Ξ is uniformly ergodic if and only if $[0, A_p]$ is small.*

This lemma is part of the statement of M&T Theorem 16.0.2 and is highlighted here without proof for reference. We establish that $[0, A_p]$ is small in several steps.

1. We claim that μ is a *maximal irreducibility measure* for Ξ (see M&T Section 4.2). By M&T Proposition 4.2.2, there are four relevant requirements for this.

   (a) μ must be an irreducibility measure for Ξ, which follows by the definitions of our payment functions. In particular, the definitions of our monetary payment functions imply that for any two account balances A and A', there exists an n-period sequence $\{\boldsymbol{\theta}^t\}_{t=1}^n$ such that, starting from A, A' is reached in the nth period. Then, for any open set $\mathscr{U} \supset \{A'\}$ there must exist a positive probability set of n-period and $(n+1)$-period sequences of realizations of $\boldsymbol{\theta}$ such that $\mathscr{U}$ is reached from A.

   (b) For any set $\mathscr{U} \subset [0, A_p]$ with zero μ-measure, the set of points from which $\mathscr{U}$ is reached in finite time with positive probability must have zero μ-measure as well. By the definition of payments in the four regimes, a point $A \in [0, A_p]$ is reached with positive probability from another point A' in one period with positive probability only if A is reached from A' if no trade occurs in that period. Conditional on A' being located in a particular one of the four regimes, by construction of the monetary payment rules A' is a translation of A by a distance that is fixed for all δ. Hence for any zero μ-measure set $\mathscr{U} \subset [0, A_p]$, the set of points from which $\mathscr{U}$ is reached in a single period with positive probability has zero μ-measure. Applying this fact inductively yields the conclusion that the set of points from which $\mathscr{U}$ is reached in finite time with positive probability is a countable union of sets of zero μ-measure, which in turn has zero μ-measure.

   (c) It must be that any other measure μ' is an irreducibility measure for Ξ if and only if it is absolutely continuous with respect to μ. The "if" part of the statement is implied by requirement (a), above; here we demonstrate the "only if" part. Suppose to the contrary that μ' is an irreducibility measure but it assigns positive measure to some zero μ-measure set $\mathscr{U}$, so that it is not absolutely continuous with respect to μ. Then by definition we need that for all $A \in [0, A_p]$, there exists $n < \infty$ such that $\xi^n(\mathscr{U}; A) > 0$. In particular, choose $A \in [0, A_r]$ such that $A \notin \mathscr{U}$. Then, for any $n < \infty$, the definition of the payments in the revenue regime imply that $\xi^n(\cdot; A)$ has exactly one atom, located at A, corresponding to the event that no trade occurs in any of the n periods, while the event that trade occurs yields a distribution over subsequent account balances that is absolutely continuous with respect to μ. Hence there cannot exist $n < \infty$ with $\xi^n(\mathscr{U}; A) > 0$. But this contradicts the supposition that $\mu'(\mathscr{U}) > 0$ and μ' is an irreducibility measure for Ξ.

   (d) μ must be "equivalent"[28] to

$$\int_{[0,A_p]} K(\cdot; A)\mu'(dA)$$

[28]Two measures are "equivalent" if they share the same null sets; i.e., they are mutually absolutely continuous.

for any irreducibility measure μ', where

$$K(\cdot;A)=\sum_{n=0}^{\infty}\xi^n(\cdot;A)2^{-(n+1)}.$$

As M&T show in the proof of their Proposition 4.2.2, it suffices to show that μ is equivalent to

$$\int_{[0,A_p]} K(\cdot;A)\mu(dA).$$

M&T Proposition 4.2.1 states if μ is an irreducibility measure for Ξ, it follows that for any $A\in[0,A_p]$, $K(\mathscr{B};A)>0$ whenever $\mu(\mathscr{B})>0$. In turn, this implies that μ is equivalent to $\int_{[0,A_p]} K(\cdot;A)\mu(dA)$, since both are atomless with full support.

2. We claim that there exists $\varepsilon>0$ such that any interval $[A,A+\varepsilon]\subseteq[0,A_d]$ is small. Since the probability density function of F_i is uniformly bounded away from zero on Θ, from the construction of the payment functions it is assured that $\xi(\cdot;A)$ has a density component with interval support for all $A\in[0,A_d]$, and furthermore the density component of $\xi(\cdot;A)$ is uniformly bounded away from zero on its support.[29] Also from the construction of the payment functions, whenever $A+\varepsilon\in\operatorname{int}\big(\operatorname{supp}\big(\xi(\cdot;A)\big)\big)$, there exists an open interval contained in $\bigcap_{A'\in[A,A+\varepsilon]}\operatorname{supp}\big(\xi(\cdot;A')\big)$. Choose $\varepsilon>0$ such that $A+\varepsilon\in\operatorname{int}\big(\operatorname{supp}\big(\xi(\cdot;A)\big)\big)$ for all $A\in[0,A_d]$ (this is possible by the definitions of the monetary payment functions), and for any such A choose $\mathscr{C}_A$ to be an open interval in $\bigcap_{A'\in[A,A+\varepsilon]}\operatorname{supp}\big(\xi(\cdot;A')\big)$. Note that $\xi(\cdot;\hat{A})=\int_{A'\in(\cdot)}\frac{d}{dA'}G(A';\hat{A})dA'$, where $\frac{d}{dA'}G(A';\hat{A})$ is a density function that is well-defined almost everywhere. Let

$$\nu_A(\cdot)=\int_{A'\in(\cdot\cap\mathscr{C}_A)}\inf_{\hat{A}\in[A,A+\varepsilon]}\frac{d}{dA'}G(A';\hat{A})dA',$$

where the infimum in the integrand is strictly positive for almost all $A'\in\mathscr{C}_A$ since for each A' the density function is uniformly bounded above zero by the same bound. Hence ν_A is a non-trivial Borel measure. Therefore, for all $A'\in[A,A+\varepsilon]$ and any Borel set $\mathscr{B}$, $\xi(\mathscr{B};A')\geq\nu_A(\mathscr{B})$, so $[A,A+\varepsilon]$ is small.

3. We claim that Ξ is aperiodic. If Ξ were periodic then there would exist $n>1$ and a "cycle" of disjoint Borel sets $\mathscr{D}_1,\ldots,\mathscr{D}_n$ such that for $i=1,\ldots,n-1$, $A\in\mathscr{D}_i$ implies $\xi(\mathscr{D}_{i+1};A)=1$, and $A\in\mathscr{D}_n$ implies $\xi(\mathscr{D}_1;A)=1$. (The latter claim is implied by M&T Theorem 5.4.4, since we have already shown that Ξ is μ-irreducible). It is

[29] Actually, the density is zero at the endpoint of the support that would result from the realization $(\overline{\theta},\underline{\theta})$, but removing a small open neighborhood of this endpoint from the support would leave a positive uniform lower bound on the density over the remainder of the support. Since this caveat has no effect on the analysis, for simplicity of exposition we continue for now as if the density were uniformly bounded away from zero on its support. In Lemma 7, we treat this issue more formally.

easy to see that in the region $[0, A_d]$ there are no cycles with $n > 1$, since for any A in this region we have that $A \in \text{supp}(\xi(\cdot; A))$. The payout regime, however, is more challenging, since if $A \in (A_d, A_p]$ then $\text{supp}(\xi(\cdot; A)) = \{A - 2\mathbb{E}[v^{\text{GET}}(\tilde{\theta}) - v^{\text{NGR}}(\tilde{\boldsymbol{\theta}})]\}$, and so for any set $\mathscr{D}_1 \subseteq (A_d, A_p]$ there exists $n > 1$ and a collection of disjoint Borel sets $\mathscr{D}_2, \ldots, \mathscr{D}_n$ such that $A \in \mathscr{D}_i$ and $\xi(\mathscr{D}_{i+1}; A) = 1$ for $i = 1, \ldots, n-1$. However, the cycle cannot be completed: for any such collection, $\mathscr{D}_i < \mathscr{D}_{i+1}$, and for a finite n, for $A \in \mathscr{D}_{n-1}$, $\text{supp}(\xi(\cdot; A)) < A_d$. But then for this A, our transition rule in the downward drift regime guarantees that $A' \in \text{supp}(\xi(\cdot; A))$ implies $A' \in \text{supp}(\xi(\cdot; A'))$. Hence there can be no cycle with $n > 1$.

4. We claim that $[0, A_d]$ is small. By M&T Proposition 5.5.3, every small set is "petite" (see M&T Section 5.5.2). By M&T Proposition 5.5.5, since Ξ is μ-irreducible, every finite union of petite sets is petite; since $[0, A_d]$ is covered by a finite collection of sets $\{[A, A + \varepsilon] \subseteq [0, A_d]\}$, $[0, A_d]$ is petite. Finally, by M&T Theorem 5.5.7, since Ξ is μ-irreducible and aperiodic, every petite set is small.

5. We claim that $[0, A_p]$ is small. When the account balance enters P it then falls deterministically in steps of equal length until it returns to $[0, A_d]$. Let m be the maximum number of steps needed to exit P. Since $\xi([0, A_d]; A)$ is uniformly bounded above zero for all $A \in [0, A_d]$, there exists $\eta > 0$ such that $\xi^m([0, A_d]; A) \geq \eta$ for all $A \in [0, A_p]$. Then M&T Proposition 5.2.4(i) yields the claim.

As previewed above, since the state space $[0, A_p]$ is small, M&T Theorem 16.0.2 implies that Ξ is uniformly ergodic, which means that a unique ergodic measure Ψ exists, and the probability measure over the account balance in period t converges to Ψ uniformly at a geometric rate as $t \to \infty$. Since Ξ is uniformly ergodic and μ-irreducible, by M&T Proposition 10.1.1 it is recurrent. Then M&T Theorem 10.4.9 implies that Ψ is equivalent to μ, and therefore is atomless and has full support on $[0, A_p]$.

Remark 1. The exact-efficient mechanism (see Section A.2) shares the same properties: it has a unique, atomless ergodic distribution with full support on $[0, A_{wd}]$, to which the account balance in period t converges at a geometric rate as $t \to \infty$. The proof operates as immediately above, with some simplifications and changes. First, in step 1(c), choose A in the withdrawal regime, which has the same relevant property as the revenue regime. Second, step 3 simplifies because an exact-efficient mechanism has no payout regime. For the same reason, step 5 is unnecessary.

A.6 *Proof of Corollary 2*

Given δ sufficiently close to 1, there exists a recursive mechanism γ_a satisfying the properties in Proposition 5; by Lemma 5, one of these properties is the existence of a unique ergodic distribution over account balances, denoted Ψ. Set the initial account balance at the ergodic expected account balance, $\bar{A} \equiv \int_{[0, A_p]} A \, d\Psi$. Choose an initial randomization over deposits and withdrawals prior to the first period of the mechanism according to Ψ. Since Ψ is the ergodic distribution, the initial account balance, the expected immediate post-randomization account balance, and the expected account balance in every subsequent period are all equal to $\bar{A}$ (where expectations are taken over both the initial randomization and all possible realizations of play under the mechanism). By account keeping, the traders' expected average joint payoffs are thus

$$\begin{aligned} v_b + v_s &= (1-\delta)\mathbb{E}\left[\sum_{\tau=0}^{\infty}\delta^{\tau}\left(\int_{A^{(\tau)}\in[0,A_p]}\big(q_b(\boldsymbol{\theta}^{(\tau)};A^{(\tau)},\mathbf{v})\theta_b^{(\tau)} + q_s(\boldsymbol{\theta}^{(\tau)};A^{(\tau)},\mathbf{v})\theta_s^{(\tau)}\big)\,d\Psi + \bar{A} - \bar{A}\right)\right] \\ &= \int_{A\in[0,A_p]} \mathbb{E}\big[\big(q_b(\tilde{\boldsymbol{\theta}};A,\mathbf{v})\tilde{\theta}_b + q_s(\tilde{\boldsymbol{\theta}};A,\mathbf{v})\tilde{\theta}_s\big)\big]\,d\Psi \\ &\geq \Psi\big((A_r,A_d]\big)\mathbb{E}\big[\big(q_b^*(\tilde{\boldsymbol{\theta}})\tilde{\theta}_b + q_s^*(\tilde{\boldsymbol{\theta}})\tilde{\theta}_s\big)\big], \end{aligned}$$

where the inequality is by the efficient allocation in the upward and downward drift regimes. By Proposition 5, $\Psi\big((A_r,A_d]\big) \to 1$ as $\delta \to 1$, completing the proof.

References

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1986), "Optimal cartel equilibria with imperfect monitoring." *Journal of Economic Theory*, 39, 251–269. [327]

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1990), "Toward a theory of discounted repeated games with imperfect monitoring." *Econometrica*, 58, 1041–1063. [306, 327]

Amador, Manuel, Iván Werning, and George-Marios Angeletos (2006), "Commitment vs. flexibility." *Econometrica*, 74, 365–396. [328]

Athey, Susan, Andrew Atkeson, and Patrick J. Kehoe (2005), "The optimal degree of discretion in monetary policy." *Econometrica*, 73, 1431–1475. [328]

Athey, Susan and Kyle Bagwell (2001), "Optimal collusion with private information." *RAND Journal of Economics*, 32, 428–465. [327, 330]

Athey, Susan, Kyle Bagwell, and Chris Sanchirico (2004), "Collusion and price rigidity." *Review of Economic Studies*, 71, 317–349. [327, 328]

Athey, Susan and David A. Miller (2006), "Efficiency in repeated games with private information and very impatient players." Working paper. [303]

Athey, Susan and David A. Miller (2007), "Posted price mechanisms in repeated trade." Working paper. [312]

Chung, Kim-Sau and Jeffrey C. Ely (2002), "Ex-post incentive compatible mechanism design." Unpublished paper, Department of Economics, Northwestern University. [300, 302, 327]

Clarke, Edward H. (1971), "Multipart pricing of public goods." *Public Choice*, 11, 17–33. [308]

Cramton, Peter, Robert Gibbons, and Paul Klemperer (1987), "Dissolving a partnership efficiently." *Econometrica*, 55, 615–632. [327]

Fudenberg, Drew and David K. Levine (1994), "Efficiency and observability with long-run and short-run players." *Journal of Economic Theory*, 62, 103–135. [306]

Fudenberg, Drew, David K. Levine, and Eric Maskin (1994), "The folk theorem with imperfect public information." *Econometrica*, 62, 997–1039. [327, 330]

Groves, Theodore (1973), "Incentives in teams." *Econometrica*, 41, 617–631. [308]

Hagerty, Kathleen M. and William P. Rogerson (1987), "Robust trading mechanisms." *Journal of Economic Theory*, 42, 94–107. [311]

Jackson, Matthew O. and Hugo F. Sonnenschein (2004), "Overcoming incentive constraints by linking decisions (extended version)." Available at `http://www.stanford.edu/~jacksonm/linkinglong.pdf`. [328]

Kandori, Michihiro and Ichiro Obara (2006), "Efficiency in repeated games revisited: The role of private strategies." *Econometrica*, 74, 499–519. [306]

Karlin, Samuel (1968), *Total Positivity*. Stanford University Press, Stanford. [347]

Levin, Jonathan (2003), "Relational incentive contracts." *American Economic Review*, 93, 835–857. [327]

McAfee, R. Preston and Philip J. Reny (1992), "Correlated information and mechanism design." *Econometrica*, 60, 395–421. [327]

Meyn, Sean P. and Richard L. Tweedie (1993), *Markov Chains and Stochastic Stability*. Communications and Control Engineering Series, Springer-Verlag, London. [342]

Miller, David A. (2007a), "Attainable payoffs in repeated games with interdependent private information." Unpublished paper, Department of Economics, University of California, San Diego. [303, 306, 327]

Miller, David A. (2007b), "The dynamic cost of ex post incentive compatibility in repeated games of private information." Unpublished paper, Department of Economics, University of California, San Diego. [302, 310, 311, 328]

Myerson, Roger B. and Mark A. Satterthwaite (1983), "Efficient mechanisms for bilateral trading." *Journal of Economic Theory*, 29, 265–281. [300, 327]

Rayo, Luis (forthcoming), "Relational incentives and moral hazard in teams." *Review of Economic Studies*. [327]

Rustichini, Aldo, Mark A. Satterthwaite, and Steven R. Williams (1994), "Convergence to efficiency in a simple market with incomplete information." *Econometrica*, 62, 1041–1063. [328]

Satterthwaite, Mark A. and Artyom Shneyerov (2007), "Dynamic matching, two-sided incomplete information, and participation costs: existence and convergence to perfect competition." *Econometrica*, 75, 155–200. [328]

Satterthwaite, Mark A. and Steven R. Williams (2002), "The optimality of a simple market mechanism." *Econometrica*, 70, 1841–1863. [328]

Swinkels, Jeroen M. (2001), "Efficiency of large private value auctions." *Econometrica*, 69, 37–68. [328]

Tatur, Tymon (2005), "On the trade off between deficit and inefficiency and the double auction with a fixed transaction fee." *Econometrica*, 73, 517–570. [328]

Taub, Bart (1994), "Currency and credit are equivalent mechanisms." *International Economic Review*, 35, 921–956. [328]

Vickrey, William (1961), "Counterspeculation, auctions, and competitive sealed tenders." *Journal of Finance*, 16, 8–37. [308]

Submitted 2005-10-17. Final version accepted 2007-8-3. Available online 2007-8-3.