Czech Economic Review 8 (2014) 33–54
Acta Universitatis Carolinae Oeconomica

Expectations of Bailout and Collective Moral Hazard

Milan Horniaček[*]

Received 30 August 2012; Accepted 29 July 2014

Abstract We analyze an infinite horizon stochastic game with discounting of future single period payoffs that models interaction between Cournot duopolists producing differentiated products. There are two states of demand for an output of each firm. The probabilities of their occurrence depend on random factors and on investing or not investing into product innovation. If the low state of demand occurs for an output of each firm, then the government not observing firms' investments into product innovation bails out each firm with a positive output. A weakly renegotiation-proof perfect public equilibrium is the solution concept that we apply. It is a perfect public equilibrium in which no two replaceable continuation equilibrium payoff vectors are strictly Pareto ranked. Taking into account expected bailout, there exists a weakly renegotiation-proof perfect public equilibrium with a strictly Pareto efficient equilibrium payoff vector, in which, along the equilibrium path, the firms collude on not investing into product innovation.

Keywords Bailout, innovation, perfect public equilibrium, weakly renegotiation-proof equilibrium, stochastic game
JEL classification C73, D43

1. Introduction

The current economic crisis prompted governments to various forms of support of firms (e.g., in the form of scrapping subsidy aimed at support of car makers—see Miťková (2009) for the analysis of its impact). Although in some cases (like the scrapping subsidy) such a support was an indirect one, we will use the term "bailout" for it. (During the crisis, even the European Commission is willing to approve government aid that it would ban in more prosperous times.) Bailouts can be expected especially when all or most of the firms in an industry are in trouble. In such a case, it can be argued that the firms themselves are not responsible for their bad results. Nevertheless, expectations of bailout can reduce the endeavors of firms to help themselves, especially through product or technological innovation. In particular, if none of the firms in the industry invests in research and development (henceforth, R&D), expected long term profit (expressed, for example, as expected average discounted profit) of each of them can be thanks to bailouts higher than if all of them invested in R&D. In such a case, it is in the interest of all firms to cooperate in suppressing R&D. If one of them deviates, invests in R&D and its expenditures lead to product (or technological) innovation, the other firms can do the same, but in the way that punishes the deviator (e.g., by a signi-

[*] Comenius University in Bratislava, Faculty of Social and Economic Sciences, Mlynské luhy č. 4, 821 05 Bratislava, the Slovak Republic. Phone: +421 220 66 98 61, E-mail: milan.horniacek@fses.uniba.sk.

ficant increase in the output of innovated product). That is, a deviation does not lead to competition in innovation, but to innovation by the other firms aimed at punishing the deviator. In the present paper, we illustrate this behavior by a simplified game theoretic model.

The model has the form of an infinite horizon discrete time stochastic game with discounting of future single period payoffs. There are just two players in it—the Cournot duopolists producing differentiated products. (This simplification, as well as the other simplifications of the model described below, enables us to concentrate on the analysis of the impact of expected bailouts on R&D activity of strategically behaving firms. We need not to distract a reader by complicated assumptions and formulae that would be unavoidable in the analysis of a more general game.) An expected average discounted profit is the payoff of each firm. We do not view the government as a player in the analyzed game. It does not behave strategically. Its decision on bailing out the firms follows a fixed simple rule—the bailout is granted to a firm if and only if it produces a positive output and both firms face the low state of demand for their product.

A weakly renegotiation-proof perfect public equilibrium (henceforth, WRPPPE) is the solution concept that we apply to the analyzed game. It is a combination of a perfect public equilibrium used by Fudenberg et al. (1994) and a weakly renegotiation-proof equilibrium developed by Farrell and Maskin (1989) (that coincides with an internally consistent equilibrium of Bernheim and Ray 1989). Thus, it is a subgame perfect equilibrium in public strategies, in which no continuation equilibrium vector of expected average discounted profits is strictly Pareto dominated by another continuation equilibrium vector of expected average discounted profits prescribed for the subgame with the same initial state of demands. WRPPPE does not require immunity of an equilibrium to renegotiation by the grand coalition to any perfect public equilibrium in any subgame. (A renegotiation-proof equilibrium of Maskin and Tirole (1988), who restrict attention to Markov strategies, requires immunity of an equilibrium to renegotiation by the grand coalition to any Markov perfect equilibrium in each subgame.) Moreover, it does not require immunity of an equilibrium to renegotiation by the grand coalition to a continuation equilibrium of another WRPPPE in any subgame. (Strongly renegotiation-proof equilibrium of Farrell and Maskin (1989), which coincides with externally consistent equilibrium of Bernheim and Ray (1989), is based on such a requirement.) Nevertheless, WRPPPE addresses the main problem of collective consistency of an equilibrium. It is "forgiveness-free"—it rules out renegotiations by the grand coalition that would consist in forgiving a deviator and returning to the play prescribed when nobody is punished. Such renegotiations are not only the most tempting ones, but they are also the most easily planned ones. They do not require any computation that has not been already done in evaluation of the original equilibrium.

We base a WRPPPE on a perfect public equilibrium and restriction to pure public strategies. Nevertheless, each perfect public equilibrium is a subgame perfect equilibrium, i.e., it is immune to unilateral deviations to (pure or behavioral) non-public strategies.

The paper is organized as follows. In the following section, we describe the analyzed model and define a WRPPPE. In Section 3, we prove a sufficient condition for the existence of a WRPPPE for discount factors close enough to one, in which the equilibrium vector of expected average discounted profits is strictly Pareto efficient and, along the equilibrium path, the firms collude on not investing into product innovation.

2. Model

Throughout the paper, $\mathbb{N}$ denotes the set of positive integers and $\mathbb{R}$ denotes the set of real numbers. We endow each finite dimensional real vector space with the Euclidean topology and $\mathbb{R}^\infty$ with the product topology.

When a vector is an argument of a function, we use only one pair of brackets (e.g., we write $y^*(\omega_{11},\omega_{21})$ instead of $y^*((\omega_{11},\omega_{21}))$).

The time horizon of the analyzed game is $\mathbb{N}$. There are two players in it. They are Cournot duopolists that produce differentiated products. The goods produced by them are substitutes.

Whenever we refer, in the same sentence or formula, to firms j and i, we assume that $j \in \{1,2\}$ and $i \in \{1,2\}\setminus\{j\}$.

For firm $j \in \{1,2\}$, the set of possible states of demand for its product is $\Omega_j = \{\omega_{j1},\omega_{j2}\}$. We set $\Omega = \Omega_1 \times \Omega_2$ and use the term "state of demands" for $\omega \in \Omega$. For each $j \in \{1,2\}$, the inverse demand function $P_j : [0,\infty)^2 \times \Omega_j \to [0, P_j(0,0,\omega_{j2})]$ assigns to each vector of outputs $y = (y_j, y_i)$ of duopolists and each state of demand $\omega_j \in \Omega_j$ the unit price at which demand for j's product equals y_j.

Assumption 1. *For each $j \in \{1,2\}$:*

(i) *P_j is continuous in y at each $(y,\omega_j) \in [0,\infty)^2 \times \Omega_j$;*

(ii) *P_j is nonincreasing in y_1 and y_2 at each $(y,\omega_j) \in [0,\infty)^2 \times \Omega_j$ and strictly decreasing in y_1 and y_2 at each $(y,\omega_j) \in [0,\infty)^2 \times \Omega_j$ with $P_j(y,\omega_j) > 0$;*

(iii) *P_j is concave in y_j at each $(y,\omega_j) \in [0,\infty)^2 \times \Omega_j$ with $P_j(y,\omega_j) > 0$;*

(iv) *P_j is twice continuously differentiable with respect to y_1 and y_2 at each $(y,\omega_j) \in [0,\infty)^2 \times \Omega_j$ with $P_j(y,\omega_j) > 0$;*

(v) *$P_j(y,\omega_{j1}) < P_j(y,\omega_{j2})$ for each $y \in [0,\infty)^2$ with $P_j(y,\omega_{j2}) > 0$;*

(vi) *$\frac{\partial P_j(y,\omega_{j1})}{\partial y_j} \le \frac{\partial P_j(y,\omega_{j2})}{\partial y_j}$ for each $y \in [0,\infty)^2$ with $P_j(y,\omega_{j1}) > 0$;*

(vii) *for each $\omega_j \in \Omega_j$, there exists $y_j^{\max}(\omega_j) > 0$ such that $P_j(y,\omega_j) = 0$ for each $y \in [0,\infty)^2$ with $y_j \ge y_j^{\max}(\omega_j)$, $P_j(y_j,0,\omega_j) > 0$ for each $y_j \in [0, y_j^{\max}(\omega_j))$, and $y_j^{\max}(\omega_{j2}) > y_j^{\max}(\omega_{j1}) > 0$.*

Parts (i)–(iv) of Assumption 1 are standard in the analysis of a Cournot oligopoly producing substitutes. Part (v) just states that ω_{j2} is the high and ω_{j1} is the low state of demand for product of firm $j \in \{1,2\}$. Part (vi) says that, for a fixed output of the competitor, the inverse demand curve at the high state of demand is not steeper than at the low state of demand. Part (vii) ensures that for each firm, at each state of demand, the set of outputs that it can sell for a positive price has a finite positive supremum (i.e., an output equal to or exceeding the supremum cannot be sold for a positive price even if the firm is alone in the market). We assume that each firm $j \in \{1,2\}$, in each period $t \in \mathbb{N}$, at each state of demand $\omega_j \in \Omega_j$, chooses its output from the interval $[0, y_j^{\max}(\omega_j)]$.

Each firm $j \in \{1,2\}$, in each period $t \in \mathbb{N}$, decides whether to invest into R&D aimed at product innovation. In order to keep the model as simple as possible, we distinguish only two decisions on R&D investment: 0 (not investing into R&D) and 1 (investing into R&D). The cost of investing into R&D for firm $j \in \{1,2\}$ is $b_j > 0$. The outcome of R&D investment in each firm $j \in \{1,2\}$ is stochastic. This stochastic outcome combines random factors affecting success or failure of R&D and random disturbances of demand. If it is successful, it leads to a higher state of the demand. It shifts the graph of the inverse demand function for j's product (for a fixed competitor's output) upwards. It can but need not change also its slope. According to part (vi) of Assumption 1, if it changes the slope of the inverse demand curve, it does not increase the absolute value of the latter. In other words, a price of the innovated product is not more sensitive to the quantity brought to the market than a price of the original product.

We denote state of demand for product of firm $j \in \{1,2\}$ in period $t \in \mathbb{N}$ by $\omega_j(t)$ and let $\omega(t) = (\omega_1(t), \omega_2(t))$, $\omega(1)$ is given. For each $j \in \{1,2\}$, function $\mu_j : \Omega_j^2 \times \{0,1\}^2 \to [0,1]$ assigns to each $(\omega_j(t), \omega_j(t+1), I_j(t), I_i(t)) \in \Omega_j^2 \times \{0,1\}^2$ the probability of occurrence of state of demand $\omega_j(t+1)$ for j's product in period $t+1$ when state of demand for j's product in period t is $\omega_j(t)$, j's investment into R&D in period t is $I_j(t)$, and the investment into R&D by its competitor in period t is $I_i(t)$. The probability distributions specified by μ_1 and μ_2 are independent.

Assumption 2. *For each $j \in \{1,2\}$:*

(i) $1 > \mu_j(\omega_{j1}, \omega_{j2}, 1, 0) > \mu_j(\omega_{j1}, \omega_{j2}, 1, 1) > \mu_j(\omega_{j1}, \omega_{j2}, 0, 0)$
$= \mu_j(\omega_{j1}, \omega_{j2}, 0, 1) = 0,$

(ii) $1 > \mu_j(\omega_{j2}, \omega_{j2}, 1, 0) > \mu_j(\omega_{j2}, \omega_{j2}, 1, 1) > \mu_j(\omega_{j2}, \omega_{j2}, 0, 0)$
$= \mu_j(\omega_{j2}, \omega_{j2}, 0, 1) = 0,$

(iii) $\mu_j(\omega_{j2}, \omega_{j2}, 1, I_i) > \mu_j(\omega_{j1}, \omega_{j2}, 1, I_i)$ *for each* $I_i \in \{0,1\}$.

Part (i) of Assumption 2 states that the probability of transition from the low to the high state of demand for a firm's product is positive if and only if the firm invests into R&D and it decreases in the case of the competitor's investment into R&D. Part (ii) expresses analogous relationships for keeping the high state of demand. Part (iii) states that, when a firm invests into R&D, the probability of keeping the high state of demand exceeds the probability of achieving it from the low state of demand.

Firm $j \in \{1,2\}$ has cost function $c_j : [0, y_j^{\max}(\omega_{j2})] \to [0, c_j(y_j^{\max}(\omega_{j2}))]$.

Assumption 3. *For each* $j \in \{1,2\}$*,* c_j *is* (i) *twice continuously differentiable,* (ii) *strictly increasing,* (iii) *convex, and* (iv) $c_j(0) = 0$.

We can justify part (iv) of Assumption 3 by normalization of single period payoff of each firm by subtracting its fixed costs.

For each $j \in \{1,2\}$, function $\eta_j : [0, y_j^{\max}(\omega_{j2})] \times [0, y_i^{\max}(\omega_{i2})] \times \Omega_j \to \mathbb{R}$ expresses firm j's single period profit from sale of its output (without taking into account the government's subsidy and cost of investment into R&D). It is defined by $\eta_j(y_j, y_i, \omega_j) = P_j(y_j, y_i, \omega_j) y_j - c_j(y_j)$.

For each $j \in \{1,2\}$, function $r_j : [0, y_i^{\max}(\omega_{i2})] \times \Omega_j \to [0, y_j^{\max}(\omega_{j2})]$ is firm j's reaction function. It assigns to each feasible output of firm i and each state of demand for j's output $\omega_j \in \Omega_j$, j's output that maximizes its profit from sale of its output. It is defined by

$$r_j(y_i, \omega_j) = \arg\max \left\{ \eta_j(y_j, y_i, \omega_j) \mid y_j \in \left[0, y_j^{\max}(\omega_j)\right] \right\}.$$

It follows from Assumptions 1 and 3 that firms' reaction functions are well-defined.

Both firms observe the state of demand for product of each of them in each period before making decisions on output and R&D investment. Each firm observes ex post the competitors output (i.e., each $j \in \{1,2\}$, at the beginning of each period $t \in \mathbb{N}$, knows the output of firm i in all preceding periods). The government also observes the state of demand for product of each firm at the beginning of each period. Neither the competitor nor the government observes a firm's investment into R&D. In each period $t \in \mathbb{N}$, in which $\omega_j(t) = \omega_{j1}$ for each $j \in \{1,2\}$, the government pays to each firm with a positive output subsidy $M > 0$. If $\omega_j(t) = \omega_{j2}$ for at least one $j \in \{1,2\}$, the government pays no subsidy to any firm. Thus, the government pays subsidy only if both firms face low state of demand, assuming that such a situation is a result of an external shock and the firms do not bear responsibility for it. On the other hand, it pays subsidy only to a firm that (despite an external shock) continues in its activity—i.e., it produces a positive output.

In order to simplify the formulae, we define function $\phi : \Omega \to \{0,1\}$ by $\phi(\omega) = 1$ if and only if $\omega_1 = \omega_{11}$ and $\omega_2 = \omega_{21}$. Thus, the government pays subsidy to each firm with a positive output if and only if state of demands ω satisfies $\phi(\omega) = 1$.

Remark 1. In order to avoid misunderstanding, we stress that the government's subsidy is not an instrument for stimulating R&D in the firms. Its aim is to help the industry in which all firms face low state of demand. In such situation, it cannot be excluded that the whole industry has been hit by unfavorable random factors. By paying the subsidy, the government mitigates social consequences of low state of demand for the whole industry—it enables both firms to continue to produce and saves their employees from unemployment and their suppliers from a plummeting demand. If only one firm faces low state of demand, it does not get the subsidy because social consequences of such situation are milder. Some employees of the failing firm can find job in the more successful firm. Suppliers of the failing firm can increase sales to the more successful firm. Moreover, by refusal to pay the subsidy to the only failing firm, the government forces the firms to care about their future sales.

It is true that firms have to publish annual reports and the government can find the expenditures on R&D from them. Nevertheless, expenditures on many activities other than product innovation (e.g. expenditures on innovation of administrative processes in a firm aimed at increasing the use of electronic documents) can he hidden in R&D cost. “Creative accounting” can hide even some operational cost among R&D expenditures (e.g. people from the R&D department can be assigned to perform operational tasks). Moreover, even specialized grant agencies have problems with uncovering of embezzlement of R&D funds (see, for example, Kouzbit 2013; Caroll 2009). Therefore, (taking into account the number of firms in the economy) it is not realistic to believe that the government could be able to find out the actual use of expenditures reported in the annual report as R&D expenditures.

Both firms discount their future single period profits by the common discount factor $\delta \in (0,1)$. (Our qualitative results would continue to hold if the firms used different discount factors.) $\Gamma(\delta)$ denotes the analyzed game with discount factor δ.

We denote the set of non-terminal (i.e., finite) histories in $\Gamma(\delta)$ by H and the set of terminal (i.e., infinite) histories by Z. $H = \bigcup_{t\in\mathbb{N}} H(t)$, where $H(t)$ is the set of histories leading to period t. $H(1)$ contains only the history $(\omega_1(1), \omega_2(1))$. For each $t \in \mathbb{N}\setminus\{1\}$, each $h \in H(t)$ has the form

$$h = \left\{(\omega_1(1), \omega_2(1)), \{y_1(\tau), y_2(\tau), I_1(\tau), I_2(\tau), \omega_1(\tau+1), \omega_2(\tau+1)\}_{\tau=1}^{t-1}\right\},$$

where, for each $j \in \{1,2\}$ and each $\tau \in \{1,\ldots,t-1\}$, $y_j(\tau)$ is firm j's output in period τ, $I_j(\tau)$ is its investment into R&D in period τ, and $\omega_j(\tau+1)$ is state of demand for j's product in period $\tau+1$. (Thus, a history leading to period t contains also the state of demands in the latter period.)

A public history contains only observable actions and observable consequences of actions, i.e., states of demands and firms' outputs. It does not contain firms' investments into R&D (which are not publicly observable). (See Fudenberg et al. (1994) for the characterization of public histories and strategies.) We denote the set of non-terminal public histories by H_p. $H_p = \bigcup_{t\in\mathbb{N}} H_p(t)$, where $H_p(t)$ is the set of public histories leading to period t. $H_p(1) = H(1)$. For each $t \in N\setminus\{1\}$, each $h_p \in H_p(t)$ has the form

$$h_p = \left\{(\omega_1(1), \omega_2(1)), \{, y_1(\tau), y_2(\tau), \omega_1(\tau+1), \omega_2(\tau+1)\}_{\tau=1}^{t-1}\right\}.$$

We restrict attention to pure public strategies in $\Gamma(\delta)$. (Our results would continue to hold if we allowed behavioral strategies, including behavioral strategies that condition actions on signals of a publicly observable randomizing device.) A pure public strategy of firm $j \in \{1,2\}$ is a function that assigns to each $t \in \mathbb{N}$ and each $h \in H_p(t)$ a pair $(y_j(t), I_j(t)) \in [0, y_j^{\max}(\omega_j(t))] \times \{0,1\}$. The firms make decisions on output and investment into R&D simultaneously. When a firm makes a decision on its output and its investment into R&D in period t, it knows the public history leading to it. We denote the set of pure public strategies of firm $j \in \{1,2\}$ in $\Gamma(\delta)$ by S_j and set $S = S_1 \times S_2$. For each $j \in \{1,2\}$, $\pi_j : S \to \mathbb{R}$ is firm j's payoff function in $\Gamma(\delta)$ defined on the set of profiles of pure public strategies. It assigns to each $s \in S$, j's expected average discounted profit (that takes into account also government subsidies and cost of investments into R&D) generated by it. (In order to keep the notations as simple as possible, we do not explicitly indicate the dependence of the expected average discounted profits on the discount factor in the symbols for firms' payoff functions in $\Gamma(\delta)$.) We define function $\pi : S \to \mathbb{R}^2$ by $\pi(s) = (\pi_1(s), \pi_2(s))$.

For each $j \in \{1,2\}$, each $t \in \mathbb{N}$, and each $h \in H_p(t)$, the prescription of $s_j \in S_j$ after public history h has the form $s_j(h) = (y_j(t), I_j(t))$, where $y_j(t) \in [0, y_j^{\max}(\omega_j(t))]$ is firm j's output and $I_j(t) \in \{0,1\}$ is its investment into R&D in period t.

For each $h \in H$, we denote by $\Gamma_{(h)}(\delta)$ the subgame of $\Gamma(\delta)$ following the non-terminal history h. For each $h \in H_p$, we denote by $\Gamma_{(h)}(\delta)$ the public subgame of $\Gamma(\delta)$ following the non-terminal public history h. It is the union of subgames following the histories whose public components coincide with h. We indicate restrictions of sets and functions defined for $\Gamma(\delta)$ to public subgame $\Gamma_{(h)}(\delta)$ by subscript (h).

When $s^* \in S$ is an equilibrium of $\Gamma(\delta)$ and $h \in H_p$, then $s^*_{(h)}$ is the continuation equilibrium of $\Gamma_{(h)}(\delta)$ and $\pi_{(h)}(s^*_{(h)})$ is the continuation equilibrium payoff vector. We say that continuation equilibria $s^*_{(h)}$ and $s^*_{(h')}$ are replaceable if and only if the states of demands at the beginning of public subgames $\Gamma_{(h)}(\delta)$ and $\Gamma_{(h')}(\delta)$ coincide. In such a case, the sets of strategy profiles and their payoff consequences in these two subgames coincide (i.e., $S_{(h)} = S_{(h')}$ and $\pi_{(h)}(s) = \pi_{(h')}(s)$ for each $s \in S_{(h)}$). Thus, the firms can use $s^*_{(h)}$ in $\Gamma_{(h')}(\delta)$ and $s^*_{(h')}$ in $\Gamma_{(h)}(\delta)$. If the latter condition is not satisfied, the sets of feasible output vectors in the first period of the two subgames are different and consequences of investments into R&D in the first period (probability distributions on the set of states of demands in the second period) are also different. These differences generate differences in the following periods.

A WRPPPE is the solution concept that we apply to $\Gamma(\delta)$. We obtain it by applying the concept of a weakly renegotiation-proof equilibrium developed by Farrell and Maskin (1989) to perfect public equilibrium (i.e., to a subgame perfect equilibrium in public strategies). That is, we supplement the definition of a perfect public equilibrium by the requirement that there do not exist two replaceable continuation equilibria with the continuation equilibrium payoff vectors that are strictly Pareto ranked.

It is worth noting that each perfect public equilibrium is a subgame perfect equilibrium. That is, no firm in no subgame can increase its expected average discounted profit by a unilateral deviation to a non-public (pure or behavioral) strategy (see Fudenberg et al. 1994, p. 1002).

We end this section by giving the formal definition of a WRPPPE tailored to $\Gamma(\delta)$.

Definition 1. *A strategy profile $s^* \in S$ is a WRPPPE of $\Gamma(\delta)$ if*

(i) *there do not exist $h \in H_p$, $j \in \{1,2\}$, and $s_j \in S_{j(h)}$ such that*

$$\pi_{j(h)}(s_j, s^*_{i(h)}) > \pi_{j(h)}(s^*_{(h)}),$$

(ii) *and there do not exist $h \in H_p$ and $h' \in H_p$ such that $s^*_{(h)}$ and $s^*_{(h')}$ are replaceable and $\pi_{j(h)}(s^*_{(h)}) > \pi_{j(h')}(s^*_{(h')})$ for each $j \in \{1,2\}$.*

3. Existence of a strictly Pareto efficient WRPPPE with moral hazard

In this section, we give a sufficient condition for the existence of a strictly Pareto efficient WRPPPE of $\Gamma(\delta)$ with moral hazard, i.e. a WRPPPE, in which the equilibrium vector of expected average discounted profits is strictly Pareto efficient and (along the equilibrium path) no firm invests into R&D but both firms rely on bailout by the government in the low states of demand for their products.

Our sufficient condition for the existence of a strictly Pareto efficient WRPPPE with moral hazard is based on the following two assumptions.

Assumption 4. *There exists $y^* \in \prod_{j=1}^2 \left(0, y_j^{\max}\left(\omega_{j1}\right)\right)$ with the following properties:*

(i) *setting $v_j^* = \eta_j\left(y_j^*, y_i^*, \omega_{j1}\right) + M$, $j \in \{1,2\}$, we have*

$$v_1^* + v_2^* = \max\left\{\sum_{j=1}^2 \left(\eta_j\left(y_j, y_i, \omega_{j1}\right) + M\right) \mid y \in \prod_{j=1}^2 \left[0, y_j^{\max}\left(\omega_{j1}\right)\right]\right\};$$

(ii) *there do not exist $\left(I_1(\omega), I_2(\omega)\right) \in \{0,1\}^2 \setminus \{(0,0)\}$ and $y(\omega) \in [0, y_1^{\max}(\omega_1)] \times [0, y_2^{\max}(\omega_2)]$, $\omega \in \Omega$, such that for some $\omega \in \Omega$*

$$\begin{aligned}&\sum_{j=1}^2 \Big[\sum_{\omega' \in \Omega} \left(\prod_{k=1}^2 \mu_k\left(\omega_k, \omega_k', I_k(\omega), I_i(\omega)\right)\right) \times \\ &\times \left(\eta_j\left(y_j(\omega'), y_i(\omega'), \omega_j'\right) + \phi(\omega') M \times \operatorname{sgn}\left(y_j(\omega')\right)\right) \\ &- b_j I_j(\omega)\Big] \geq v_1^* + v_2^*\,; \end{aligned} \tag{1}$$

(iii) *for each $i \in \{1,2\}$, $y_j \in \left[0, y_j^{\max}\left(\omega_{j1}\right)\right]$ defined by*

$$y_j > r_j\left(0, \omega_{j1}\right) \;\&\; \eta_j\left(y_j, 0, \omega_{j1}\right) = v_j^* - M \tag{2}$$

satisfies

$$\begin{aligned}&\mu_i\left(\omega_{i1}, \omega_{i2}, 1, 0\right) \eta_i\left(r_i\left(y_j, \omega_{i2}\right), y_j, \omega_{i2}\right) + \\ &+ \left(1 - \mu_i\left(\omega_{i1}, \omega_{i2}, 1, 0\right)\right)\left(\eta_i\left(r_i\left(y_j, \omega_{i1}\right), y_j, \omega_{i1}\right) + \operatorname{sgn}\left(r_i\left(y_j, \omega_{i1}\right)\right) M\right) \\ &- b_i > v_i^*. \end{aligned} \tag{3}$$

Assumption 5. *For each $j \in \{1,2\}$,*

$$\begin{aligned} v_j^+ &= \mu_j\left(\omega_{j1}, \omega_{j2}, 1, 0\right) \eta_j\left(r_j\left(0, \omega_{j2}\right), 0, \omega_{j2}\right) \\ &\quad + \left(1 - \mu_j\left(\omega_{j1}, \omega_{j2}, 1, 0\right)\right)\left[\eta_j\left(r_j\left(0, \omega_{j1}\right), 0, \omega_{j1}\right) + M\right] - b_j \\ &> \eta_j\left(r_j\left(0, \omega_{j1}\right), 0, \omega_{j1}\right) + M, \end{aligned} \tag{4}$$

$$\eta_i\left(r_i\left(r_j\left(0,\omega_{j1}\right),\omega_{i1}\right),r_j\left(0,\omega_{j1}\right),\omega_{i1}\right)+M<v_i^*, \tag{5}$$

$$\begin{aligned}&\mu_j\left(\omega_{j1},\omega_{j2},1,1\right)\eta_i\left(r_i\left(r_j\left(0,\omega_{j2}\right),\omega_{i2}\right),r_j\left(0,\omega_{j2}\right),\omega_{i2}\right)\\&+\left(1-\mu_j\left(\omega_{j1},\omega_{j2},1,1\right)\right)\eta_i\left(r_i\left(r_j\left(0,\omega_{j1}\right),\omega_{i2}\right),r_j\left(0,\omega_{j1}\right),\omega_{i2}\right)-b_i<v_i^*,\end{aligned} \tag{6}$$

$$\eta_i\left(r_i\left(r_j\left(0,\omega_{j2}\right),\omega_{i2}\right),r_j\left(0,\omega_{j2}\right),\omega_{i2}\right)<v_i^*. \tag{7}$$

Assumptions 4 and 5 may seem to be complicated. Nevertheless, they have simple verbal interpretation. They are based on comparison of firms' expected profits computed by summing the expected profit from sale of the output (taking into account the probabilities of different states of demands for a given current state of demands) with the government's subsidy and subtracting the costs of investments into R&D (that will lead to the probabilities of states of demands considered in computing the expected profit).

Assumption 4 states that there exists a collusive outcome in which the firms refrain from investing into R&D and choose an output vector that gives both of them positive profit from sale of their output and has the following three properties. First, it maximizes the sum of their profits in the low states of demand for their products. Second, it gives the sum of profits that is higher than the sum of expected profits generated by any other choice of investments into R&D and output vectors. Third, each firm i can increase its expected profit by a unilateral deviation to investing into R&D and choosing the optimal reaction to the competitor's output for each state of demand for its product unless the competitor produces (irrespective of the state of demand for i's product) an output that gives him lower profit from sale of his product than the collusive outcome.

Assumption 5 states that each firm j, when, starting in the low state of demand for its product, it invests into R&D and produces its monopoly output in each state of demand for its product (while the competitor produces nothing), earns expected profit exceeding its profit from its monopoly output in the low state of demand for its product (when the state of demand for the competitor's product is also low; see (4)). Obviously, the latter profit exceeds its profit in the collusive outcome. Taking into account part (iii) of Assumption 2, the same conclusion holds when firm j starts in the high state of demand for its product. In addition, a deviation without investment into R&D gives the competitor lower profit than the collusive outcome (see (5)). Also, a deviation with investment into R&D gives the competitor lower expected profit than the collusive outcome. (This follows from (6) and (5).) The last conclusion holds irrespective of the state of demand in which a firm starts, as well as irrespective of the state of demand in which the competitor starts. (This follows from part (iii) of Assumption 2 and parts (ii), (v), and (vi) of Assumption 1. Parts (v) and (vi) of Assumption 1 imply that, for any firm, its monopoly output in the high state of demand is higher than its monopoly output in the low state of demand for its product.) Finally, the best response in the high state of demand for its own product to the competitor's monopoly output in the high state of demand for its product gives each firm lower profit than the collusive outcome (see (7)).

In the Appendix we give an example of the game that satisfies Assumptions 1–5. Assumptions 1–5 together form a sufficient condition for the existence of a WRPPPE with moral hazard and strictly Pareto efficient vector of expected average discounted profits.

Remark 2. As noted in the Section 1, a public perfect equilibrium defined by Fudenberg et al. (1994) is one of the two solution concepts for dynamic games that are combined in WRPPPE. Nevertheless, our game is quite different from the one analyzed by Fudenberg et al. (1994). They analyze a countable infinitely repeated game. Players' actions in the current period affect neither the sets of their feasible actions in the following period nor their payoff consequences. We analyze a stochastic game in which players' actions in the current period affect payoff consequences of their actions in the following period by affecting the distribution of states of demands (that are payoff relevant in the sense described in Maskin and Tirole (1988)) in the following period. These differences between the analyzed games are reflected in the fact that our Assumptions 4 and 5 have no relation to Conditions 6.1 and 6.2 in Fudenberg et al. (1994, p. 1020–1021). The latter deal with the distribution of public outcomes induced by players' deviations from the actions prescribed by their strategies. Our Assumptions 4 and 5 deal with payoffs induced by profiles of players' actions.

Proposition 1. *Let Assumptions 1–5 hold. Then there exist $\underline{\delta} \in (0,1)$ such that, for each $\delta \in [\underline{\delta},1)$, $\Gamma(\delta)$ has a WRPPPE in which the vector of expected average discounted profits is strictly Pareto efficient on the set $\{\pi(s) \mid s \in S\}$ and, along the equilibrium path, no firm invests into R&D.*

Proof. Proof with mathematical details is deferred to the Appendix.

Now we give the outline of the proof of Proposition 1. In the first period both firms refrain from investment into R&D and produce outputs that maximize the sum of their profits in the state of demands at the beginning of the game. In the following periods, unless a deviation occurs, they again refrain from investment into R&D and produce positive outputs that maximize the sum of their profits when the state of demand for the product of each of them is low. Refraining from investment into R&D leads to persistence of low states of demand. If firm i unilaterally deviates when no punishment takes place, it is punished by exclusion from the market (producing nothing and refraining from investment into R&D) for T_{i1} periods. The competitor invests into R&D in each period of punishment except for the last one. In each period of punishment of firm i, he produces his monopoly output in the state of demand for his product that occurs. (If the punisher invests into R&D in the last period of punishment, he is punished in the same way as for a unilateral deviation when no punishment takes place.) In each period of punishment, the punishing firm earns single period profit that is greater than its single period collusive profit (i.e., its single period profit when none of the firms is punished). The duration of punishment T_{i1} is long enough to ensure that the punishment wipes out any gain of deviating firm i from a single period deviation in output.

Now suppose that a unilateral deviation of firm i in output in period t is coupled with its investment into R&D in period t and a subsequent deviation in output in period $t+1$ in each state of demand for i's product. In this case, the punisher can react by its investment into R&D only since period $t+1$. As a consequence of this, for discount factor close to one, i's gain from investment into R&D in period t and deviation in output in period $t+1$ (i.e., the difference between its expected profit in period $t+1$ discounted to period t and its investment cost incurred in period t) exceeds its single period collusive profit. (This follows from part (iii) of Assumption 4.) Thus, the latter gain exceeds i's average discounted profit in the subgame at the beginning of which its punishment for T_{i1} periods is triggered. Therefore, such a deviation cannot be punished by restarting of punishment lasting for T_{i1} periods. Instead, such a deviation (as well as any new deviation by firm i during its punishment planned to last T_{i1} periods) triggers a new punishment lasting for $T_{i2} > T_{i1}$ periods. During this longer punishment (unless a deviation from it occurs) the firms behave in the same way as during the shorter one. The expected gain from a unilateral deviation of the punished firm in output or from its investment into R&D with a subsequent deviation in output during the longer punishment is lower than its single period collusive profit. (The reason for this is that now it faces the punisher who invested into R&D.) Thus, for discount factor close enough to one, this gain does not exceed the punished firm's average discounted profit in the subgame at the beginning of which the longer punishment is triggered. Therefore, restarting of the longer punishment wipes out this gain.

The strategies outlined above depend only on past output vectors, on the past states of demands, and on the current state of demands, i.e., only on elements of public histories. (When a firm deviates by investing into R&D the competitor reacts to this deviation only if it leads to the high state of demand for the deviator's product.) Hence, they are public strategies.

A renegotiation from a continuation equilibrium, in which t periods of punishment of firm i remain, to a continuation equilibrium, in which $\tau < t$ periods of i's punishment remain, would decrease the expected average discounted profit of the punisher - firm j. The same holds for a renegotiation to a continuation equilibrium in which none of the firms is punished or firm j is punished. A renegotiation from a continuation equilibrium, in which t periods of punishment of firm i remain, to a continuation equilibrium, in which $\tau > t$ periods of i's punishment remain, would decrease the average discounted profit of firm i. A renegotiation from a continuation equilibrium in which none of the firms is punished to a continuation equilibrium in which firm i is punished would decrease average discounted profit of firm i. Thus, the outlined strategy profile is a WRPPPE.

The outlined strategy profile prescribes (along the equilibrium path) refraining from investment into R&D by both firms and the output vector that maximizes the sum of profits in the current state of demands in each period. It follows from part (ii) of Assumption 4 that the firms cannot increase the sum of their expected average discounted profits by investing into R&D and modifying their output plans. Thus, the outlined strategy profile maximizes the sum of firms' expected average discounted profits. Therefore, the vector of expected average discounted profits generated by it is strictly Pareto efficient.

Remark 3. Our game is quite different also from the one analyzed by Chen (1995). He deals with a countable infinitely repeated game. Players' actions in the current period affect neither the sets of their feasible actions in the following period nor their payoff consequences. Thus, all non-terminal histories lead to the same payoff relevant state.

In our stochastic game, players' actions in the current period affect the distribution of payoff relevant states of demands in the following period. Moreover, there is no stage game in a stochastic game. Hence, Chen's result on the stage game best response (Chen 1995, Proposition 1, p. 603) cannot be applied to our game. Differences between the analyzed games are reflected also in the different definitions of a weakly renegotiation-proof equilibrium. Chen requires that no two continuation equilibrium payoff vectors are strictly Pareto ranked (see Chen 1995, p. 602). We require that there do not exist two replaceable continuation equilibria (i.e., two continuation equilibria with the same initial state of demands) with the continuation equilibrium payoff vectors that are strictly Pareto ranked. Therefore, our result neither contradicts to nor extends that of Chen (1995).

Remark 4. Our aim in this paper was not to show that a weakly renegotiation-proof equilibrium in public strategies with strictly Pareto efficient payoff vector can exist in a stochastic game with (some) unobservable actions of the players. The aim was to show that a too generous subsidy to firms in a failing industry can lead to persistent collective moral hazardous behavior that is even sustainable in a WRPPPE and gives a strictly Pareto efficient payoff vector. This aim motivated the formulation of Assumptions 4 and 5. These assumptions ensure that an agreement between firms not to invest into R&D (and rely on the government's subsidy) is sustainable in a WRPPPE with strictly Pareto efficient payoff vector. They do not rule out existence of another WRPPPE with strictly Pareto efficient payoff vector. The set of such additional equilibria depends on specific functional forms and parameters of the model.

4. Conclusions

We have identified a sufficient condition for the existence of a WRPPPE of $\Gamma(\delta)$ for the discount factor close enough to one, in which the firms collude on refraining from investments into R&D and relying on the government's subsidy. Such a behavior is tantamount to collusive collective moral hazard. The government pays the subsidy to both firms because it does not know that collusive refraining from investments into R&D is the main culprit for low state of demand for products of both of them. Investments into R&D play a peculiar—we could say, even perverse—role in the WRPPPE with moral hazard (constructed in the proof of Proposition 1). They are not an instrument of competition between firms but an instrument of punishing the competitor's deviation from the collusive arrangement based on refraining from them. When a firm deviates, invests into R&D, and its investments lead to the high state of demand for its product, the competitor punishes it by investing into R&D and producing its monopoly output (in both states of demand for his product) for a finite number of periods. Of course, the government's rule for granting subsidies to firms, which does not address the fact that investments into R&D are not observable, is responsible for such a perverse use of the latter investments.

From the game theoretic point of view, the present paper is (to our best knowledge) the first one that applies the principle of weak renegotiation-proofness to a game with several payoff relevant states, in which both players take an action in each period.

The model of Maskin and Tirole (1988) is the one with alternating moves. One firm takes an action in odd periods and the other in even periods. Thus, each firm can react to a change in the payoff relevant state caused by an action of the competitor before the latter moves again. In our model, a deviator who invests into R&D can reap benefits of the deviation in the immediately following period, before the punisher can react by investing into R&D. This makes weakly renegotiation-proof punishments more complicated than in the case of countable infinitely repeated games (like in Farrell and Maskin 1989) and games with alternating moves.

Acknowledgment The research reported in the present paper was supported by the grant VEGA 1/0181/12 from the Slovak Ministry of Education, Science, Research, and Sport.

References

Bernheim, B. D. and Ray, D. (1989). Collective Dynamic Consistency in Repeated Games. *Games and Economic Behavior*, 1(4), 295–326.

Caroll, J. (2009). Korean Scientist Convicted in R&D Embezzlement Scandal. http://www.fiercebiotechresearch.com/story/korean-scientist-convicted-r-d-embezzlement-scandal/2009-10-26 [Retrieved on February 5, 2014]

Chen, K.-P. (1995). On Renegotiation-Proof Equilibrium under Imperfect Monitoring. *Journal of Economic Theory*, 65(2), 600–610.

Farrell, J. and Maskin, E. (1989). Renegotiation in Repeated Games. *Games and Economic Behavior*, 1(4), 327–360.

Fudenberg, D., Levine, D. and Maskin, E. (1994). The Folk Theorem with Imperfect Public Information. *Econometrica*, 62(5), 997–1039.

Kouzbit, O. (2013). More Embezzlement Seen at Glonass – $ 213 Million Total Over Three Years. http://www.ewdn.com/2013/05/28/more-embezzlement-seen-at-glonass-213-million-total-over-three-years/ [Retrieved on February 5, 2014]

Maskin, E. and Tirole, J. (1988). A Theory of Dynamic Oligopoly II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles. *Econometrica*, 56(3), 571–599.

Miťková, V. (2009). Modeling the Crisis in Automobile Industry – The Applied CGE Study. In Brožová, H. and Kvasnička, R. (eds.), Proceedings of the Mathematical Methods in Economics 2009. Prague, Czech University of Life Sciences, Faculty of Economics and Management, p. 221–226.

Appendix

First, we give an example of a game satisfying Assumptions 1–5.

Example 1. For each $j \in \{1,2\}$,

$$\begin{aligned}
P_j\left(y_j, y_i, \omega_{j2}\right) &= \max\left\{20 - 2y_j - y_i, 0\right\},\\
P_j\left(y_j, y_i, \omega_{j1}\right) &= \max\left\{10 - 2y_j - y_i, 0\right\},\\
c_j\left(y_j\right) &= 5y_j,\\
b_j &= 2,\\
\mu_j\left(\omega_{j1}, \omega_{j2}, 1, 0\right) &= 0.8,\\
\mu_j\left(\omega_{j1}, \omega_{j2}, 1, 1\right) &= 0.6,
\end{aligned}$$

other values of function μ_j are either directly specified in Assumption 2 or (in order to satisfy Assumptions 4 and 5) can have any values satisfying Assumption 2. We set $M = 15$.

It is easy to verify that this example satisfies Assumptions 1–3. The maximization of the sum of profits in the low state of demand for product of both firm gives each firm profit (gross of the government subsidy) 2.0833. Therefore, $v^* = (17.0833, 17.0833)$. The maximization of the sum of profits in the high state of demand for product of both firms gives each firm profit (gross of cost of investment into R&D) 18.75. We have $18.75 - 2 = 16.75 < v_1^* = v_2^*$. The maximization of the sum of profits when the state of demand for its product is high for one firm and low for the other firm gives sum of profits (gross of cost of investment into R&D) equal to 27.111. We have $27.111 - 2 = 25.111 < v_1^* + v_2^*$. Thus, v^* satisfies parts (i) and (ii) of Assumption 4. To see that part (iii) of Assumption 4 is satisfied, it is enough to note that the left hand side of (1) is less than

$$\begin{aligned}
&\sum\nolimits_{j=1}^{2}\Big[\sum\nolimits_{\omega'\in\Omega}\Big(\prod\nolimits_{k=1}^{2}\mu_k\left(\omega_k, \omega_k', I_k(\omega), I_i(\omega)\right)\Big)\\
&\times\left(\eta_j\left(y_j\left(\omega'\right), y_i\left(\omega'\right), \omega_j'\right) + \phi\left(\omega'\right) M \times \operatorname{sgn}\left(y_j\left(\omega'\right)\right) - b_j\zeta_j\left(\omega_j'\right)\right)\Big],
\end{aligned}\tag{A1}$$

where $\zeta_j\left(\omega_{j1}\right) = 0$ and $\zeta_j\left(\omega_{j2}\right) = 1$ for each $j \in \{1,2\}$. Using the results on the maximal sums of profits in various states of demands given above, the expression (A1) is (for specification of functions μ_1 and μ_2 satisfying Assumption 2) less than $v_1^* + v_2^*$.

For each $i \in \{1,2\}$, the output y_j defined in (2) equals 1.9717. Substituting this into the left hand side of (3), we get $18.203 > 17.0833 = v_1^* = v_2^*$. Thus, part (iii) of Assumption 4 is satisfied.

It is easy to verify by direct computations that Assumption 5 is satisfied.

Proof of Proposition 1. *Preliminaries.*
For $j \in \{1,2\}$, let $v_i^{\max} = \max\{\eta_i\left(r_i\left(0, \omega_{i2}\right), 0, \omega_{i2}\right), \eta_i\left(r_i\left(0, \omega_{i1}\right), 0, \omega_{i1}\right) + M\}$. It is the highest single period payoff that firm i can earn. For $i \in \{1,2\}$ and $T \in \mathbb{N}$ consider inequality

$$(1-\delta)\, v_i^{\max} + \delta^{T_{i1}+1} v_i^* - \delta v_i^* \le 0. \tag{A2}$$

For the limit case $\delta = 1$, (A2) is satisfied as equality. The derivative of its left hand side with respect to δ evaluated at $\delta = 1$ is positive if and only if

$$T_{i1} > \frac{v_i^{\max}}{v_i^*}.$$

We set

$$T_{i1} = \min\left\{n \in \mathbb{N} \mid n > \frac{v_i^{\max}}{v_i^*}\right\},\ i = 1,2.$$

Then there exists $\delta_1 \in (0,1)$ such that (A2) holds for each $i \in \{1,2\}$ and each $\delta \in [\delta_1, 1)$. Next, for each $i \in \{1,2\}$ and $T_{i2} \in \mathbb{N}$, consider inequalities

$$(1-\delta)\, v_i^{\max} + \delta^{T_{i2}+1} v_i^* - \delta^{T_{i1}} v_i^* \le 0, \quad \text{(A3)}$$

$$(1-\delta)\,\eta_j\left(r_j\left(0,\omega_{j1}\right), r_i\left(r_j\left(0,\omega_{j1}\right),\omega_{i2}\right),\omega_{j1}\right) + \delta\left(1-\delta^{T_{i2}}\right) v_j^+ + \delta^{T_{i2}+1} v_j^* \ge v_j^*. \quad \text{(A4)}$$

For the limit case $\delta = 1$, (A3) and (A4) are satisfied as equalities. The derivative of the left hand side of (A3) with respect to δ evaluated at $\delta = 1$ is positive if and only if

$$T_{i2} > \frac{v_i^{\max} - v_i^* + T_{i1} v_i^*}{v_i^*}.$$

The derivative of the left hand side of (A4) with respect to δ evaluated at $\delta = 1$ is negative if and only if

$$T_{i2} > \frac{v_j^* - \eta_j\left(r_j\left(0,\omega_{j1}\right), r_i\left(r_j\left(0,\omega_{j1}\right),\omega_{i2}\right),\omega_{j1}\right)}{v_j^+ - v_j^*}.$$

We set

$$T_{i2} = \min\left\{\begin{array}{c} n \in \mathbb{N} \mid \\ n > \max\left\{\frac{v_i^{\max} - v_i^* + T_{i1} v_i^*}{v_i^*}, \frac{v_j^* - \eta_j\left(r_j\left(0,\omega_{j1}\right), r_i\left(r_j\left(0,\omega_{j1}\right),\omega_{i2}\right),\omega_{j1}\right)}{v_j^+ - v_j^*}\right\}\end{array}\right\},$$
$$i = 1,2.$$

Then there exists $\delta_2 \in (0,1)$ such that, for each $i \in \{1,2\}$, (A3) and (A4) hold for each $\delta \in [\delta_2, 1)$. Further, it follows from (5) that there exists $\delta_3 \in (0,1)$ such that

$$\eta_i\left(r_i\left(r_j\left(0,\omega_{j1}\right),\omega_{i1}\right), r_j\left(0,\omega_{j1}\right),\omega_{i1}\right) + M \le \delta^{T_{i2}} v_i^*,\ i = 1,2 \quad \text{(A5)}$$

for each $\delta \in [\delta_3, 1)$. It follows from (6) that there exists $\delta_4 \in (0,1)$ such that

$$\begin{aligned}&\delta\mu_j\left(\omega_{j1},\omega_{j2},1,1\right)\eta_i\left(r_i\left(r_j\left(0,\omega_{j2}\right),\omega_{i2}\right), r_j\left(0,\omega_{j2}\right),\omega_{i2}\right)\\ &+\delta\left(1-\mu_j\left(\omega_{j1},\omega_{j2},1,1\right)\right)\eta_i\left(r_i\left(r_j\left(0,\omega_{j1}\right),\omega_{i2}\right), r_j\left(0,\omega_{j1}\right),\omega_{i2}\right) - b_i \le \delta^{T_{i2}+1} v_i^*,\\ &i = 1,2, \qquad \text{(A6)}\end{aligned}$$

for each $\delta \in [\delta_4, 1)$. It follows from (4) that there exists $\delta_5 \in (0,1)$ such that

$$\begin{aligned}
&\delta \mu_j (\omega_{j1}, \omega_{j2}, 1, 0)\, \eta_j \left(r_j \left(0, \omega_{j2}\right), 0, \omega_{j2}\right) \\
&+ \delta \left(1 - \mu_j (\omega_{j1}, \omega_{j2}, 1, 0)\right) \left[\eta_j \left(r_j \left(0, \omega_{j1}\right), 0, \omega_{j1}\right) + M\right] - b_j \\
&\geq \delta \left[\eta_j \left(r_j \left(0, \omega_{j1}\right), 0, \omega_{j1}\right) + M\right],\ j = 1,2
\end{aligned} \tag{A7}$$

for each $\delta \in [\delta_5, 1)$. It follows from (7) that there exists $\delta_6 \in (0,1)$ such that

$$\eta_i \left(r_i \left(r_j \left(0, \omega_{j2}\right), \omega_{i2}\right), r_j \left(0, \omega_{j2}\right), \omega_{i2}\right) \leq \delta^{T_{i2}} v_i^*,\ i = 1,2 \tag{A8}$$

for each $\delta \in [\delta_6, 1)$.

We set $\underline{\delta} = \max\{\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6\}$. For each $\omega \in \Omega$ fix

$$y^* (\omega) \in \arg\max \left\{ \begin{array}{l} \eta_1 (y_1, y_2, \omega_1) + \eta_2 (y_1, y_2, \omega_2) \mid \\ y \in [0, y_1^{\max} (\omega_1)] \times [0, y_2^{\max} (\omega_2)] \end{array} \right\}. \tag{A9}$$

Thus, $y^* = y^* (\omega_{11}, \omega_{21})$.

For each $t \in \mathbb{N} \setminus \{1\}$ and each $h \in H_p(t)$, we denote by h^- the subhistory of h leading to period $t-1$. It contains the same information as h except for firms' outputs in period $t-1$ and state of demands in period t.

Description of strategy profile s^*. For each $i \in \{1,2\}$ function $d_i : H_p \to \{0, 1, \ldots, T_{i2} + 1\}$ assigns to each $h \in H_p$ the number of periods in which firm i has to be punished in public subgame $\Gamma_{(h)}(\delta)$ for a deviation prior to the beginning of $\Gamma_h(\delta)$. It is defined recursively as follows. We set $d_1 (\omega_1 (1), \omega_2 (1)) = d_2 (\omega_1 (1), \omega_2 (1)) = 0$. Take (arbitrary) $t \in N \setminus \{1\}$ such that functions d_1 and d_2 are already defined for each $h \in H_p(t-1)$. Consider (arbitrary) $h \in H_p(t)$. If $d_1(h^-) = d_2(h^-) = 0, \omega_1(t) = \omega_{11}$, $\omega_2(t) = \omega_{21}$, and $y(t-1) = y^*(\omega_1(t-1), \omega_2(t-1))$, then $d_1(h) = d_2(h) = 0$. If $d_1(h^-) = d_2(h^-) = 0$ and either $\omega_1(t) = \omega_{12}$ or $y_1(t-1) \neq y_1^*(\omega_1(t-1), \omega_2(t-1))$, then $d_1(h) = T_{11}$ and $d_2(h) = 0$. If $d_1(h^-) = d_2(h^-) = 0$, $\omega_1(t) = \omega_{11}$, $y_1(t-1) = y_1^*(\omega_1(t-1), \omega_2(t-1))$, and either $\omega_2(t) = \omega_{22}$ or

$$y_2(t-1) \neq y_2^*(\omega_1(t-1), \omega_2(t-1)),$$

then $d_2(h) = T_{21}$ and $d_1(h) = 0$. If $1 < d_i(h^-) \leq T_{i2}$, $\omega_i(t) = \omega_{i1}$, and $y_i(t-1) = 0$, then $d_i(h) = d_i(h^-) - 1$ and $d_j(h) = 0$. If $d_i(h^-) = 1$, $\omega_i(t) = \omega_{i1}$, $\omega_j(t) = \omega_{j1}$, and $y_i(t-1) = 0$, then $d_i(h) = d_j(h) = 0$. If $d_i(h^-) = T_{i2} + 1$, $\omega_i(t) = \omega_{i1}$, and $y_j(t-1) = r_j\left(0, \omega_{j1}\right)$, then $d_i(h) = T_{i2}$ and $d_j(h) = 0$. If $d_i(h^-) > 0$, $\omega_i(t) = \omega_{i1}$, $\omega_i(t-1) = \omega_{i1}$, and $y_i(t-1) > 0$, then $d_i(h) = T_{i2}$ and $d_j(h) = 0$. If $d_i(h^-) > 0$, $\omega_i(t) = \omega_{i2}$, and $\omega_j(t) = \omega_{j2}$, then $d_i(h) = T_{i2}$ and $d_j(h) = 0$. If $d_i(h^-) > 0$, $\omega_i(t) = \omega_{i2}$, and $\omega_j(t) = \omega_{j1}$, then $d_i(h) = T_{i2} + 1$ and $d_j(h) = 0$. If $d_i(h^-) = T_{i2} + 1$, $\omega_i(t) = \omega_{i1}$, and $y_j(t-1) \neq r_j\left(0, \omega_{j1}\right)$, then $d_j(h) = T_{j1}$ and $d_i(h) = 0$. If $d_i(h^-) = 1$, $\omega_i(t) = \omega_{i1}$, $\omega_j(t) = \omega_{j2}$, and $y_i(t-1) = 0$, then $d_j(h) = T_{j1}$ and $d_i(h) = 0$. Note that we have $d_j(h) = 0$ for each $h \in H_p$ with $d_i(h) > 0$. Therefore, in the following text, whenever we deal with the case $d_i(h) > 0$, we do not specify that $d_j(h) = 0$.

Now we define $s^* \in S$. For $h \in H_p(1)$, $s_i^*(h) = (y_i^*(\omega_1(1), \omega_2(1)), 0)$ for each

$i \in \{1,2\}$. Next consider (arbitrary) $t \in \mathbb{N} \setminus \{1\}$ and $h \in H_p$. If $d_1(h) = d_2(h) = 0$, then $s_i^*(h) = (y_i^*, 0)$ for each $i \in \{1,2\}$. If $d_i(h) = 1$, then $s_i^*(h) = (0,0)$ and $s_j^*(h) = (r_j(0, \omega_j(t)), 0)$. If $1 < d_i(h) \leq T_{i2}$, then $s_i^*(h) = (0,0)$ and $s_j^*(h) = (r_j(0, \omega_j(t)), 1)$. If $d_i(h) = T_{i2} + 1$, then $s_i^*(h) = \left(r_i\left(r_j\left(0, \omega_{j1}\right), \omega_{i2}\right), 0\right)$ and $s_j^*(h) = \left(r_j\left(0, \omega_{j1}\right), 1\right)$.

Clearly, when both firms follow s^* (i.e., when $d_1(h) = d_2(h) = 0$ for each $h \in H_p$ that occurs), no firm invests into R&D.

Strategy profile s^* is a perfect public equilibrium. $\Gamma(\delta)$ is a game continuous at infinity. (Firms' future single period profits are discounted. There is a common upper bound on the absolute value of profit of any firm in any period. This follows from the fact that there are only two states of demand for product of each firm, the interval of feasible outputs of each firm in each state of demand for its product is compact, and inverse demand and cost functions are continuous.) Therefore, in order to show that there do not exist a public subgame and a firm that can increase its expected average discounted profit in it by a unilateral deviation, it is enough to consider unilateral single period deviations in output and unilateral single period deviations in investment into R&D followed by a deviation in output in the immediately following period.

First we show that there do not exist $h \in H_p$ and firm $i \in \{1,2\}$ that can increase its expected average discounted profit in $\Gamma_{(h)}(\delta)$ by a unilateral deviation in output in the first period of $\Gamma_{(h)}(\delta)$. If s^* does not prescribe any punishment in the first period of $\Gamma_{(h)}(\delta)$, the claim follows from (A2). (Note that, for $\omega(1) \in \Omega \setminus \{(\omega_{11}, \omega_{21})\}$, $\eta_i(y_i^*(\omega(1)), y_j^*(\omega(1)), \omega_i(1)) \geq 0$. Otherwise, taking into account part (iv) of Assumption 3, we would have $y_i^*(\omega(1)) > 0$. Thus, we could increase the value of the objective function in problem (A9) by decreasing i's output to 0. By an analogous reasoning, using part (i) of Assumption 4, $\eta_j(y^*) > 0$ for each $j \in \{1,2\}$.) If s^* prescribes the beginning of the punishment of firm $i \in \{1,2\}$ lasting for T_{i1} periods in the first period of $\Gamma_{(h)}(\delta)$ and i deviates in output in the latter period, then the claim follows from (A3). If s^* prescribes continuation of punishment of $i \in \{1,2\}$ or starting of punishment of i lasting for T_{i2} periods in the first period of $\Gamma_{(h)}(\delta)$, then the deviation gives i single period profit no higher than $\delta^{T_{i2}} v_i^*$ (using (A5) and (A8)) and then the punishment lasting for T_{i2} periods is restarted. Thus, i's average discounted profit in $\Gamma_{(h)}(\delta)$ is at most $\delta^{T_{i2}} v_i^*$. Without a deviation it is at least $\delta^{T_{i2}} v_i^*$. Therefore, i does not gain by a deviation. If s^* prescribes starting of punishment of $i \in \{1,2\}$ lasting for $T_{i2} + 1$ periods in the first period of $\Gamma_{(h)}(\delta)$, then $s_i^*(h) = \left(r_i\left(r_j\left(0, \omega_{j1}\right), \omega_{i2}\right), 0\right)$ and $s_j^*(h) = \left(r_j\left(0, \omega_{j1}\right), 1\right)$, so i cannot increase its single period profit by a deviation in output in the first period of $\Gamma_{(h)}(\delta)$.

If s^* prescribes starting of punishment of $i \in \{1,2\}$ lasting for $T_{i2} + 1$ periods in the first period of $\Gamma_{(h)}(\delta)$ and firm j deviates in output in the latter period, it is punished for T_{j1} periods by zero single period profit. Thus, using (A2), (A4), and part (iii) of Assumption 2, j does not gain by such a deviation. In all other cases during i's punishments j produces its monopoly output, so it cannot increase its single period profit by a deviation in output.

Next we show that there do not exist $h \in H_p$ and firm $i \in \{1,2\}$ that can increase its expected average discounted profit in $\Gamma_{(h)}(\delta)$ by a unilateral deviation in investments into R&D followed by a deviation in output in the immediately following period. (In

(A10)–(A13), we work with the upper bound $v_i^{\max}$ on i's expected single period profit in the second period of $\Gamma_{(h)}(\delta)$. Clearly, in the cases considered in (A10)–(A20), i's investment into R&D without subsequent deviation in output decreases its expected average discounted profit in $\Gamma_{(h)}(\delta)$.) If $i \in \{1,2\}$ invests into R&D in the first period of $\Gamma_{(h)}(\delta)$ with $h \notin H_p(1)$, in which s^* does not prescribe any punishment, without deviation in output and deviates in output in the second period of $\Gamma_{(h)}(\delta)$, then, using (A2) and the fact that $T_{i2} > T_{i1}$, its expected average discounted profit in $\Gamma_{(h)}(\delta)$ is not higher than

$$\begin{aligned}
&(1-\delta)(v_i^* - b_i) + (1-\delta)\delta v_i^{\max} + \delta^{T_{i2}+2} v_i^* \\
&< (1-\delta) v_i^* + \delta\left[(1-\delta) v_i^{\max} + \delta^{T_{i2}+1} v_i^*\right] \\
&< (1-\delta) v_i^* + (1-\delta) v_i^{\max} + \delta^{T_{i1}+1} v_i^* \leq v_i^*.
\end{aligned} \tag{A10}$$

Thus, i does not gain by the deviation. If $h \in H_p(1), (A10)$ is replaced by

$$\begin{aligned}
&(1-\delta)(\eta_i(y^*(\omega(1))) - b_i) + (1-\delta)\delta v_i^{\max} + \delta^{T_{i2}+2} v_i^* \\
&< (1-\delta)\eta_i(y^*(\omega(1))) + \delta\left[(1-\delta) v_i^{\max} + \delta^{T_{i2}+1} v_i^*\right] \\
&< (1-\delta)\eta_i(y^*(\omega(1))) + (1-\delta) v_i^{\max} + \delta^{T_{i1}+1} v_i^* \\
&\leq (1-\delta)\eta_i(y^*(\omega(1))) + \delta v_i^*.
\end{aligned} \tag{A11}$$

Thus, again, i does not gain by the deviation. If i's investment into R&D in the first period of $\Gamma_{(h)}(\delta)$ is coupled with its deviation in output, then (A10) is replaced by

$$\begin{aligned}
&(1-\delta)(v_i^{\max} - b_i) + (1-\delta)\delta v_i^{\max} + \delta^{T_{i2}+2} v_i^* \\
&< (1-\delta) v_i^{\max} + \delta\left[(1-\delta) v_i^{\max} + \delta^{T_{i2}+1} v_i^*\right] \\
&\leq (1-\delta) v_i^{\max} + \delta^{T_{i1}+1} v_i^* \leq \delta v_i^* < v_i^*,
\end{aligned} \tag{A12}$$

and (A11) is replaced by

$$\begin{aligned}
&(1-\delta)(v_i^{\max} - b_i) + (1-\delta)\delta v_i^{\max} + \delta^{Ti+2} v_i^* \\
&< (1-\delta) v_i^{\max} + \delta\left[(1-\delta) v_i^{\max} + \delta^{T_{i2}+1} v_i^*\right] \\
&\leq (1-\delta) v_i^{\max} + \delta^{T_{i1}+1} v_i^* \leq \delta v_i^* \\
&\leq (1-\delta)\eta_i(y^*(\omega(1))) + \delta v_i^*,
\end{aligned} \tag{A13}$$

where we used (A3) and then (A2). Thus, i does not gain by the deviation.

Next consider $\Gamma_{(h)}(\delta)$, in the first period of which s^* prescribes the beginning or continuation of i's punishment (i.e., $d_i(h) > 0$). First, suppose that i's punishment prescribed by s^* has to last for at most T_{i2} periods (i.e., $d_i(h) \leq T_{i2}$; note that this implies that $s_i^*(h) = (0,0)$), i deviates by investing into R&D without deviation in output in the first period of $\Gamma_{(h)}(\delta)$ and it deviates in output in the second period of $\Gamma_{(h)}(\delta)$ unless the state of demand for its product is high and the state of demand for the competitors product is low in the second period of $\Gamma_{(h)}(\delta)$. If i's investment into

R&D leads to the high state of demand for its product in the second period of $\Gamma_{(h)}(\delta)$, then, using (A6), its expected average discounted profit in $\Gamma_{(h)}(\delta)$ is not higher than

$$(1-\delta)\,\delta^{T_{i2}+1}v_i^* + \delta^{T_{i2}+2}v_i^* < (1-\delta)\,\delta^{T_{i2}}v_i^* + \delta^{T_{i2}+1}v_i^* = \delta^{T_{i2}}v_i^*. \tag{A14}$$

If i's investment into R&D does not lead to the high state of demand for its product in the second period of $\Gamma_{(h)}(\delta)$, then, using (A5) and the fact that the competitor's monopoly output is higher in the high state than in the low state of demand for his product, i's expected average discounted profit in $\Gamma_{(h)}(\delta)$ is not higher than

$$(1-\delta)\left(\delta^{T_{i2}+1}v_i^* - b_i\right) + \delta^{T_{i2}+2}v_i^* < \delta^{T_{i2}}v_i^*. \tag{A15}$$

Thus, i does not gain by the deviation. If i's punishment prescribed by s^* has to last for $T_{i2}+1$ periods (i.e., $d_i(h) = T_{i2}+1$; note that this implies that $s_i^*(h) = \left(r_i\left(r_j\left(0,\omega_{j1}\right),\omega_{i2}\right),0\right)$), $A14$) is replaced by

$$\begin{aligned}&(1-\delta)\,\eta_i\left(r_i\left(r_j\left(0,\omega_{j1}\right),\omega_{i2}\right),r_j\left(0,\omega_{j1}\right),\omega_{i2}\right) + (1-\delta)\,\delta^{T_{i2}+1}v_i^* + \delta^{T_{i2}+2}v_i^*\\ &= (1-\delta)\,\eta_i\left(r_i\left(r_j\left(0,\omega_{j1}\right),\omega_{i2}\right),r_j\left(0,\omega_{j1}\right),\omega_{i2}\right) + \delta^{T_{i2}+1}v_i^*\end{aligned} \tag{A16}$$

and (A15) is replaced by

$$\begin{aligned}&(1-\delta)\left(\eta_i\left(r_i\left(r_j\left(0,\omega_{j1}\right),\omega_{i2}\right),r_j\left(0,\omega_{j1}\right),\omega_{i2}\right) - b_i\right) + (1-\delta)\,\delta^{T_{i2}+1}v_i^* + \delta^{T_{i2}+2}v_i^*\\ &< (1-\delta)\,\eta_i\left(r_i\left(r_j\left(0,\omega_{j1}\right),\omega_{i2}\right),r_j\left(0,\omega_{j1}\right),\omega_{i2}\right) + \delta^{T_{i2}+1}v_i^*.\end{aligned} \tag{A17}$$

Thus, again, i does not gain by the deviation. If i's punishment prescribed by s^* has to last for at most T_{i2} periods and its investment into R&D is coupled with a positive output in the first period of $\Gamma_{(h)}(\delta)$, (A14) is replaced by

$$\begin{aligned}&(1-\delta)\,\delta^{T_{i2}}v_i^* + (1-\delta)\,\delta^{T_{i2}+1}v_i^* + \delta^{T_{i2}+2}v_i^*\\ &= \delta^{T_{i2}}v_i^*\left[1-\delta+(1-\delta)\,\delta+\delta^2\right] = \delta^{T_{i2}}v_i^*\end{aligned} \tag{A18}$$

and (A15) is replaced by

$$\begin{aligned}&(1-\delta)\left(\delta^{T_{i2}}v_i^* - b_i\right) + (1-\delta)\,\delta^{T_{i2}+1}v_i^* + \delta^{T_{i2}+2}v_i^*\\ &< (1-\delta)\,\delta^{T_{i2}}v_i^* + (1-\delta)\,\delta^{T_{i2}+1}v_i^* + \delta^{T_{i2}+2}v_i^* = \delta^{T_{i2}}v_i^*,\end{aligned} \tag{A19}$$

where we used (A5) and (A8). Thus, i does not gain by the deviation.

It follows from (A7) and part (iii) of Assumption 2 that firm j cannot increase its expected average discounted profit in the subgame, in which s^* prescribes start or continuation of a punishment of firm i in any of its period except the last one, by refraining from investment into R&D in the first period of the subgame.

Finally, consider investment into R&D by firm j in the first period of $\Gamma_{(h)}(\delta)$, in which the punishment of firm i ends and the state of demand for j's output is $\widetilde{\omega}_j(1)$, with j's deviation in output in the second period of $\Gamma_{(h)}(\delta)$. Using (A2) and the fact that $T_{j2} > T_{j1}$, j's expected average discounted profit in $\Gamma_{(h)}(\delta)$ is not higher than

$$\begin{aligned}
&(1-\delta)(\eta_j(r_j(0,\widetilde{\omega}_j(1)),0,\widetilde{\omega}_j(1))-b_j)+(1-\delta)\delta v_j^{\max}+\delta^{T_{j2}+2}v_j^* \\
&<(1-\delta)\eta_j(r_j(0,\widetilde{\omega}_j(1)),0,\widetilde{\omega}_j(1))+\delta\left[(1-\delta)v_j^{\max}+\delta^{T_{j2}+1}v_j^*\right] \\
&<(1-\delta)\eta_j(r_j(0,\widetilde{\omega}_j(1)),0,\widetilde{\omega}_j(1))+(1-\delta)v_j^{\max}+\delta^{T_{j1}+1}v_j^* \\
&\leq(1-\delta)\eta_j(r_j(0,\widetilde{\omega}_j(1)),0,\widetilde{\omega}_j(1))+\delta v_j^*. \qquad \text{(A20)}
\end{aligned}$$

Thus, j does not gain by the deviation.

Strategy profile s^* is weakly renegotiation-proof. We have to show that for each pair of subgames $\Gamma_{(h)}(\delta)$ and $\Gamma_{(h')}(\delta)$ with the same initial state of demands $\widetilde{\omega}(1)\in\Omega$, there is a firm $k\in\{1,2\}$ such that

$$\pi_{k(h)}\left(s_{(h)}^*\right)\geq\pi_{k(h')}\left(s_{(h')}^*\right). \qquad \text{(A21)}$$

We start with case $d_i(h)\in\{1,\dots,T_{i2}\}$ and $d_i(h')<d_i(h)$. (Note that this excludes the case $\widetilde{\omega}_j(1)=\omega_{j1}$ and $\widetilde{\omega}_i(1)=\omega_{i2}$.) If $d_i(h)>1$,

$$\begin{aligned}
\pi_{j(h)}\left(s_{(h)}^*\right)&>(1-\delta)\eta_j(r_j(0,\widetilde{\omega}_j(1)),0,\widetilde{\omega}_j(1))+\delta\left(1-\delta^{d_i(h)-1}\right)v_j^*+\delta^{d_i(h)}v_j^* \\
&>v_j^*, \qquad \text{(A22)}
\end{aligned}$$

where we used (A7), part (iii) of Assumption 2, and the fact that

$$\eta_j\left(r_j\left(0,\omega_{j1}\right),0,\omega_{j1}\right)>\eta_i\left(y_j^*,y_i^*,\omega_{j1}\right)$$

and (using (4) and the fact that $\mu_j\left(\omega_{j1},\omega_{j2},1,0\right)\in(0,1)$)

$$\eta_j\left(r_j\left(0,\omega_{j2}\right),0,\omega_{j2}\right)>\eta_j\left(r_j\left(0,\omega_{j1}\right),0,\omega_{j1}\right)+M.$$

If $d_i(h)=1$, j's average discounted profit in $\Gamma_{(h)}(\delta)$ equals

$$\pi_{j(h)}\left(s_{(h)}^*\right)=(1-\delta)\eta_j(r_j(0,\widetilde{\omega}_j(1)),0,\widetilde{\omega}_j(1))+\delta v_j^*>v_j^*. \qquad \text{(A23)}$$

Clearly, if $d_i(h')=0$, then $\pi_{j(h')}\left(s_{(h')}^*\right)\leq v_j^*$. (If $d_j(h')=T_{j2}+1$, $\widetilde{\omega}_i(1)=\omega_{i1}$, and $\widetilde{\omega}_j(1)=\omega_{j2}$, the latter inequality follows from (A3).) Thus, in this case, taking into account (A22) and (A23), (A21) holds for $k=j$. Next consider subcase $d_i(h')>0$. Denote by w_j the sum of expected discounted profits of firm j in the first $d_i(h')$ periods of $\Gamma_{(h')}(\delta)$ when both firms follow s^*. Then $\pi_{j(h')}\left(s_{(h')}^*\right)=(1-\delta)w_j+\delta^{d_i(h')}v_j^*$ and

$$\begin{aligned}
\pi_{j(h)}\left(s_{(h)}^*\right)&>(1-\delta)w_j+\delta^{d_i(h')}\left(1-\delta^{d_i(h)-d_i(h')}\right)v_j^*+\delta^{d_i(h)}v_j^* \\
&=(1-\delta)w_j+\delta^{d_i(h')}v_j^*,
\end{aligned}$$

where we used analogous reasoning as in the derivation of (A22). Thus, if $d_i(h') > 0$, then (A21) holds for $k = j$.

Next consider case $d_i(h') \in \{2, \dots, T_{i2}\}$ and $0 < d_i(h) < d_i(h')$. Then

$$\pi_{i(h)}\left(s^*_{(h)}\right) = \delta^{d_i(h)} v_i^* > \delta^{d_i(h')} v_i^* = \pi_{i(h')}\left(s^*_{(h')}\right).$$

Thus, (A21) holds for $k = i$.

Next consider case $d_i(h) = T_{i2} + 1$, $\widetilde{\omega}_i(1) = \omega_{i2}$, $\widetilde{\omega}_j(1) = \omega_{j1}$, and $d_i(h') = 0$. Then, using (A4),

$$\pi_{j(h)}\left(s^*_{(h)}\right) \geq v_j^* > \pi_{j(h')}\left(s^*_{(h')}\right).$$

Thus, (A21) holds for $k = j$.

Finally, consider case $d_1(h) = d_2(h) = 0$ and $d_i(h') > 0$. (Note that this implies that $\widetilde{\omega}_1(1) = \omega_{11}$ and $\widetilde{\omega}_2(1) = \omega_{21}$.) Then

$$\pi_{i(h)}\left(s^*_{(h)}\right) = v_i^* > \pi_{i(h')}\left(s^*_{(h')}\right).$$

Thus, (A21) holds for $k = i$.

Equilibrium payoff vector is strictly Pareto efficient. We have

$$\pi(s^*) = \begin{pmatrix} (1-\delta)\,\eta_1\left(y_1^*(\omega(1)), y_2^*(\omega(1)), \omega_1(1)\right) + \delta v_1^* \\ (1-\delta)\,\eta_2\left(y_2^*(\omega(1)), y_1^*(\omega(1)), \omega_2(1)\right) + \delta v_2^* \end{pmatrix}.$$

In order to show that $\pi(s^*)$ is strictly Pareto efficient on the set $\{\pi(s) \mid s \in S\}$, it is enough to show that there does not exist $s \in S$ with

$$\pi_1(s) + \pi_2(s) > \pi_1(s^*) + \pi_2(s^*). \tag{A24}$$

Strategy profile s^* does not prescribe investment into R&D in any period by any firm and it prescribes an output vector maximizing the sum of firms' single period profits from sale of their outputs in state of demands $\omega(1)$ in the first period and in state of demands $(\omega_{11}, \omega_{21})$ in the other periods. Therefore, if $s \in S$ satisfies (A24), it has to prescribe investment into R&D by some firm in some period. Since $\Gamma(\delta)$ is a game continuous at infinity and the sum of expected average discounted profits is a scalar, it is enough to consider investments into R&D in one period and the choice of output vector in the immediately following period, in any state of demands that occurs with a positive probability in it. That is, (since functions μ_1 and μ_2 have only current states of demand for firms' products and their current investments into R&D as their arguments) it is enough to show that there do not exist $\omega \in \Omega$, $(I_1(\omega), I_2(\omega)) \in \{0,1\}^2 \setminus \{(0,0)\}$, and

$$y(\omega') \in [0, y_1^{\max}(\omega_1)] \times [0, y_2^{\max}(\omega_2)],\ \omega' \in \Omega,$$

such that

$$\sum_{j=1}^{2}\Big[\sum_{\omega'\in\Omega}\Big(\prod_{k=1}^{2}\mu_k\left(\omega_k,\omega_k',I_k(\omega),I_i(\omega)\right)\Big)$$
$$\times\,\delta\left(\eta_j\left(y_j\left(\omega'\right),y_i\left(\omega'\right),\omega_j'\right)+\phi\left(\omega'\right)M\times\operatorname{sgn}\left(y_j\left(\omega'\right)\right)\right)-b_jI_j(\omega)\Big]>\delta\left(v_1^*+v_2^*\right). \tag{A25}$$

Suppose that there exist $\omega\in\Omega$, and $y(\omega')$, $\omega'\in\Omega$, such that (A25) is satisfied. Then (1) is satisfied for ω. This contradiction with part (ii) of Assumption 4 completes the proof of strict Pareto efficiency of $\pi(s^*)$. $\square$