ESSAYS IN ECONOMIC THEORY

by

Jean Tirole

Ingénieur, Ecole Polytechnique
(1976)

Ingénieur, Ecole Nationale des Ponts et Chaussées
(1978)

Docteur de 3 ième cycle, Mathématiques de la Décision
Université Paris IX
(1978)

Submitted to the Department of Economics
In Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

June 1981

Signature of Author ____
Department of Economics
May 10, 1981

Certified by ______________________________
Eric Maskin
Thesis Supervisor

Accepted by ______________________________
Richard Eckaus
Chairman, Departmental Committee for Graduate Students

ESSAYS IN ECONOMIC THEORY

by

Jean Tirole

Ingénieur, Ecole Polytechnique
(1976)

Ingénieur, Ecole Nationale des Ponts et Chaussées
(1978)

Docteur de 3 ième cycle, Mathématiques de la Décision
Université Paris IX
(1978)

Submitted to the Department of Economics
In Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

June 1981

Signature of Author Jean Tirole
Department of Economics
May 10, 1981

Certified by Eric Maskin
Eric Maskin
Thesis Supervisor

Accepted by Richard S. Eckaus
Richard Eckaus
Chairman, Departmental Committee for Graduate Students

ESSAYS IN ECONOMIC THEORY

by

JEAN TIROLE

Submitted to the Department of Economics
on May 15, 1981 in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

ABSTRACT

Essay I

This Essay studies the properties of fixed-price equilibrium and related concepts. Fixed-price allocations possess two basic properties: implementability (they are decentralized by a set of quantity constraints) and order (only one side of a given market is constrained). Their characterization as social Nash optima stresses the lack of coordination between markets, and furthermore provides easy proof of existence and study of dimensionality. Then optimality properties are examined. Constrained Pareto optima are generically not implementable, and thus are not K-equilibria. On the other hand, even a K-equilibrium which is not dominated by any other K-equilibrium need not be optimal in the class of implementable allocations, and moreover an implementable Pareto optimum need not be orderly.

Essay II

This Essay analyzes how an early entrant in a market can exploit its headstart by strategic investment. The answer depends crucially on the solution used. We argue that "perfect equilibrium" is the most appropriate concept for the study of dynamic rivalry.

Our analysis is based on Spence's (1979) paper, "Investment Strategy and Growth in a New Market". We establish the existence of the set of perfect equilibria in the no-discounting case, and suggest that one particular equilibrium is most reasonable. This equilibrium, also valid with discounting, involves the follower firm being forever deterred from investing to its steady-state reaction curve, in contrast to Spence's proposed solution. Finally, we consider entry deterrence.

Essay III

This Essay considers the possibility of static and dynamic speculation when traders have rational expectations. Its central theme is that, unless traders have different priors or are able to obtain insurance in the market, speculation relies on inconsistent plans, and thus is ruled out by rational expectations. Static speculative markets with and without insurance motives are characterized. Then a sequential asset market with a finite number of traders and differential information is described in order to study the speculation created by potential capital gains. Price bubbles and their martingale properties are examined. It is argued that price bubbles rely on the myopia of traders and that they disappear if traders adopt a truly dynamic maximizing behavior.

Thesis Supervisor: Eric Maskin

Title: Professor of Economics

Acknowlegments

The Department of Economics at MIT is a very congenial place to write a thesis. During the three years I have spent here I have benefited from numerous interactions with the faculty members and my fellow students. In particular, Drew Fudenberg and Eric Maskin's help was invaluable, and I am delighted to include joint papers with each of them in this thesis. It is difficult to name all those who contributed to my formation here, but I would like to thank Peter Diamond, my second reader, Rudiger Dornbusch , Joe Farrell, Stan Fischer, Tim Kehoe, Bob Solow, and among the students, David Levine and Carl Shapiro. Eric Maskin proved to be an exceptional thesis advisor, never selfish with his time. My special thanks go to him at this time.

Table of Contents

Essay I

On the Efficiency of Fixed Price Equilibria

(with Eric Maskin)

Chapter 1

Introduction

Theorems characterizing equilibrium in economies that fail to satisfy some of the strictures of the Arrow-Debreu model have recently abounded. In particular, papers by Grossman [1977], Grossman and Hart [1979], and Hahn [1979] have studied the efficiency properties of equilibrium with incomplete market structures and have established analogues of the two principal theorems of welfare economics. In this paper we undertake a study of the efficiency of a model with a different kind of imperfection, fixed prices. More precisely, we start by showing (Proposition 2) that Grandmont's [1977] notion of K-equilibrium (Keynesian equilibrium) in fixed-price models, which embraces both the Drèze [1975] and Benassy[1] [1975] equilibrium concepts is equivalent to a kind of social Nash optimum[2], in which optimization is incompletely coordinated across markets and where the control variables are quantity constraints. Viewing K-equilibria as social Nash optima, we believe, permits a better understanding of the structure of the set of equilibria (see the existence theorem (Proposition 3) and the study of local dimension (Chapter 6)).

K-equilibria possess two important properties: order (the requirement that at most one side of the market can be quantity-constrained) and voluntary exchange (no one trades more of any good than he wants to). We examine the inter-relations among K-equilibria, order, and voluntary exchange (sometimes called implementability) and their connection with the two most natural concepts of optimality in a fixed-price economy: constrained Pareto optimality (optimality relative to trades that are feasible at the fixed-prices[3]) and implementable Pareto optimality (optimality relative to feasible trades that satisfy voluntary exchange). We show first (Proposition 5) that a common definition of order (c.f., Grandmont [1977]), in fact, implies that exchange is voluntary (assuming that preferences are convex and differentiable). We, therefore, consider a less demanding notion of order, weak order[4], which is distinct from voluntary exchange. By analogy with weak order, we introduce a weaker form of voluntary exchange; viz., weak implementability[5]. We prove (Proposition 7) that, with convexity and differentiability, order is equivalent to the conjunction of weak implementability (weak voluntary exchange) and weak order. We then demonstrate (Proposition 8) that constrained Pareto optima, although weakly orderly, are, except by accident, non-implementable and, hence, non-orderly. Then Proposition 9 shows that a Pareto maximal element in the set of implementable allocations - i.e., an implementable Pareto optimum - need not be weakly orderly, and thus need not be a K-equilibrium. These last two results mean that whether the economy is centralized or decentralized (i.e., whether or not traders are compelled to make trades or are free to make them), it may be efficient to "constrain" both sides of the market.

As we have said, K-equilibria are equivalent to social Nash optima. However, they are not optimal in the familiar sense. In particular, it is quite possible for one K-equilibrium to Pareto-dominate another. Proposition 10 demonstrates, moreover, that even non-dominated K-equilibria (i.e., Keynesian Pareto optima) need not be implementable Pareto optima. Although Proposition 9 and 10 apply when the only restriction on preferencesare convexity and differentiability, there is a subclass of preferences - that of "no spillover effects" - for which the KPOs and IPOs coincide (Proposition 11).

Chapter 2

Notation and Definitions

Consider an economy of m + 1 goods indexed by h(h = 0,1,...,m), whose price vector p is fixed ($p_o = 1$), and n traders indexed by i(i = 1,...,n) where trader i has a feasible net trade set $X^i \subseteq R^{m+1}$. We assume that X^i is convex and contains the origin (so that trading nothing is possible) and that trader i's preferences (denoted by $\succsim^i$) are continuous and strictly convex on this set. We will at times require preferences to be differentiable as well. Following Grandmont [1977], we define an equilibrium for such an economy as follows:

Definition 1: A K-equilibrium is a vector of net trades $(t^1,...,t^n)$ associated with the vector of quantity constraints $((\underline{Z}^1,\overline{Z}^1),...,(\underline{Z}^n,\overline{Z}^n))$ (with $\underline{Z}^i \leq 0$, $\overline{Z}^i \geq 0$, $\underline{Z}^i_o = -\infty$ and $\overline{Z}^i_0 = +\infty$) such that, for all i,

(a) t^i is feasible at prices p: $t^i \in \tilde{X}^i = X^i \cap \{t^i | p.t^i = 0\}$

(b) quantity constraints are observed: $\underline{Z}^i \leq t^i \leq \overline{Z}^i$

(c) exchange is voluntary: t^i is the $\succsim^i$ - maximal element among net trades satisfying (a) and (b).

(d) exchange is orderly: if, for some commodity h, some agents i and j,

$(\tilde{t}^i,\tilde{t}^j) \in \gamma^i_h(\underline{Z},\overline{Z}) \times \gamma^j_h(\underline{Z},\overline{Z})$, $\tilde{t}^i \succ^i t^i$ and $\tilde{t}^j \succ^j t^j$, then $(t^i_h - \overline{Z}^i_h)(t^j_h - \underline{Z}^j_h) \geq 0$, where

$\gamma^i_h(\underline{Z},\overline{Z}) = \{\tilde{t}^i \in \tilde{X}^i | \underline{Z}^i_k \leq \tilde{t}^i_k \leq Z^i_k \quad \forall k \neq 0,h\}$

(e) aggregate feasibility: $\sum_i t^i = 0$

For any trade t^i by trader i, we may confine our attention to the <u>"canonical" rations</u> $\underline{Z}(t^i)$ and $\overline{Z}(t^i)$, associated with that trade: for $h \neq 0$, if $t_h^i \geq 0$, then $\overline{Z}_h(t^i) = t_h^i$ and $\underline{Z}_h(t^i) = 0$; if $t_h^i \leq 0$, then $\overline{Z}_h(t^i) = 0$ and $\underline{Z}_h(t^i) = t_h^i$.

Voluntary exchange implies that agents are not forced to trade more of any good than they want to. An allocation characterized by voluntary exchange is said to be implementable. Formally, we have:

<u>Definition 2</u>: An <u>implementable allocation</u>[6] is a vector of net trades $(t^1,\ldots,t^n)$ satisfying conditions (a), (b), (c), and (e) for the canonical rations associated with these trades.

A market is orderly if buyers and sellers are not both constrained on that market. The next two definitions represent alternative attempts to capture the idea of order. First we introduce property (d'), which is equivalent to (see Proposition 1), but somewhat easier to work with, than (d). (d'): A vector of net trades $(t^1,\ldots,t^n)$ satisfies property (d') if, for all markets h, there exists no alternative vector $(\tilde{t}^1,\ldots,\tilde{t}^n) \varepsilon \prod_i \gamma_h^i(\underline{Z}(t^i), \overline{Z}(t^i))$ such that $\tilde{t}^i \gtrsim^i t^i$ (with at least one strict preference) and $\sum_i \tilde{t}_h^i = 0$.[7]

The following definition is equivalent to one in Grandmont [1977].

Definition 3: An orderly allocation is a vector of net trades $(t^1,\ldots,t^n)$ satisfying (a), (b), (d'), and (e) for the canonical rations associated with those trades.

The problem with the above definition of an orderly allocation, if one is attempting to distinguish between the notions of order and voluntary exchange, is that it itself embodies elements of voluntary exchange. Indeed, we will show below (Proposition 4) that, with differentiability, the above concept of order implies voluntary exchange. Heuristically, this is because, under the definition of order, the trade $\tilde{t}^i$ in $\gamma_h^i(t^i)$ could be preferred to t^i simply because t^i involves forced trading on a market $k \neq h$ and not because $\tilde{t}^i$ relaxes a constraint on market h. Therefore, we define an alternative notion of order (due to Younès [1975]) that is free from the taint of voluntary exchange. We first define property (d'):

(d'): A vector of net trades $(t^1,\ldots,t^n)$ satisfies property (d') if, for all markets h, there exists no alternative vector $(\tilde{t}^1,\ldots,\tilde{t}^n) \in \prod_{i=1}^{n} \overline{\gamma}_h^i(t^i)$ such that, for each i, $\tilde{t}^i \geq_i t^i$ (with at least one strict preference) and $\sum_i \tilde{t}_h^i = 0$, where $\overline{\gamma}_h^i(t^i) = \{\overline{t}^i \in \tilde{X}^i \mid \overline{t}_k^i = t_k^i,\ k \neq o,h\}$
Notice that properties (d') and (d") are identical except that the latter requires that alternative net trade vectors be identical to the original trades in all markets other than h and 0.

Definition 4: A weakly orderly allocation is a vector of net trades $(t^1,\ldots,t^n)$ satisfying property (a), (b), (d"), and (e) for the canonical rations associated with the trades.

An orderly allocation is obviously weakly orderly. By analogy with weak order, we may define a concept of weak implementability. We first introduce a weaker version of property (c):

(c'): A vector of net trades $(t^1,\ldots,t^n)$ satisfies property (c') if, for all markets h, there do not exist i and $\tilde{t}^i \in \bar{\gamma}^i_h(t^i)$ such that $\tilde{t}^i \gtrsim {}^i t^i$ (with at least one strict preference) and $\underline{Z}_h\ (t^i) \leq \tilde{t}^i_h \leq \bar{Z}_h(t^i)$. We now have:

Definition 5: A weakly implementable allocation is a vector of net trades $(t^1,\ldots,t^n)$ satisfying conditions (a), (b), (c') and (e) for the canonical rations associated with these trades.

Below we shall be interested in the Pareto-maximal elements in the sets of K-equilibria, implementable allocations, orderly allocations and weakly orderly allocations, which will be called K-Pareto optima (KPO), implementable Pareto optima (IPO), orderly Pareto optima (OPO), and weakly orderly Pareto optima (WPO), respectively. An ostensibly still stronger notion of optimality, selecting Pareto-maximal elements in the set of all feasible allocations, is constrained Pareto optimality:

Definition 6: A constrained Pareto optimum (CPO) is a Pareto optimum of the economy for feasible consumption sets $\tilde{X}^i = X^i \cap \{t^i | p.t^i = 0\}$, (i.e., it solves the program:

$$(*) \max \sum_{i=1}^{n} \lambda^i u^i(t^i) \text{ subject to } t^i \in \tilde{X}^i \text{ and } \sum_i t^i = 0,$$

for some choice of non-negative λ^i's, where the u^i's are utility functions representing preferences over net trades.

Chapter 3

Characterization and Existence of K-Equilibrium

We first check the consistency of the definitions:

Proposition 1: {K-equilibrium allocations} = {Implementable allocations} $\cap$ {Orderly allocations}.

Proof: We need just check that (d) and (d') are equivalent[8,9]. If (d') is not satisfied for $(t^1,\ldots,t^n)$, there exist h and $(\tilde{t}^1,\ldots,\tilde{t}^n) \in \prod_{i=1}^{n} \gamma_h^i(\underline{Z}(t^i), \overline{Z}(t^i))$ such that $(\tilde{t}^1,\ldots,\tilde{t}^n)$ Pareto-dominates $(t^1,\ldots,t^n)$. Because $(\tilde{t}^1,\ldots,\tilde{t}^n)$ maintains equilibrium on market h, we can infer that at least one agent is demand-constrained and one supply-constrained in $(t^1,\ldots,t^n)$, which contradicts (d). If, on the other hand, (d) is not satisfied by $(t^1,\ldots,t^n)$, there exist h, i, j, and

$$(\tilde{t}^i,\tilde{t}^j) \in \gamma_h^i(\underline{Z}(t^i), \overline{Z}(t^i)) \times \gamma_h^j(\underline{Z}(t^j), \overline{Z}(t^j))$$

such that $(\tilde{t}^i,\tilde{t}^j)$ Pareto-dominates (t^i,t^j) and $(\tilde{t}_h^i - t_h^i) \times (\tilde{t}_h^j - t_h^j) < 0$. Therefore, if the constraints on market h are relaxed by the amount min $\{|\tilde{t}_h^i - t_h^i|, |\tilde{t}_h^j - t_h^j|\}$, property (d') is contradicted. Q.E.D.

We now turn to the characterization of K-equilibria in terms of social Nash optima. As indicated above, the idea behind a social Nash optimum is to consider m uncoordinated planners, one for each market h ($h \neq 0$), who choose quantity constraints to maximize a weighted sum of consumers' utilities, subject to keeping equilibrium on their own market and given the rations chosen by the other planners. A social Nash optimum is then defined as a Nash equilibrium of that "game". Formally:

Definition 7: For each i, let $\overset{\circ}{t}^i(p,\underline{Z}^i,\overline{Z}^i)$ solve trader i's preference maximization problem, given prices p and rations $\underline{Z}^i$ and $\overline{Z}^i$. Define the indirect utility function $v^i(p,\underline{Z}^i,\overline{Z}^i) = u^i(\overset{\circ}{t}^i(p,\underline{Z}^i,\overline{Z}^i))$, where u^i is a utility function representing i's preferences. Suppose that, for fixed positive weights $\{\lambda_h^i\}$, the manager of market h chooses $\underline{Z}_h^i$ and $\overline{Z}_h^i$ for each i so as to maximize $\sum_{i=1}^{n} \lambda_h^i v^i(p,\underline{Z}^i,\overline{Z}^i)$ subject only to the constraint $\sum_{i=1}^{n} \overset{\circ}{t}_h^i(p,\underline{Z}^i,\overline{Z}^i) = 0$ and taking as given the rations $\underline{Z}_k^i$ and $\overline{Z}_k^i$ in each market $k \neq h$, 0. The allocation corresponding to the equilibrium of such a "game" is a social Nash optimum.

By definition, a social Nash optimum is an implementable allocation. From the equivalence between (d) and and (d') it is also orderly. Conversely, a K-equilibrium is a social Nash optimum. We have thus characterized the set of K-equilibria:

Proposition 2: {K-equilibria} = {Social Nash optima}

The set of weakly orderly and orderly allocations can be characterized similarly. In particular, a weakly orderly allocation is equivalent to a social Nash optimum where the instruments of planner h are the trades $\{t_h^i\}$ on his own market. The characterization of orderly allocations, although straightforward, is less natural because of the hybrid nature of these allocations: an orderly allocation is a social Nash optimum where each planner chooses trades on his own market given the canonical rations associated with the allocations chosen by the other planners. In other words, each planner assumes that the others have the power only to choose rations, whereas, in fact, they choose the actual trades. Under differentiability, and with at least three goods, Corollary 6 below guarantees that this kind of social Nash optimum is identical to that of Definition 7.

As a by-product of the characterization of K-equilibria as social Nash optima, we obtain a straightforward proof of the existence of a K-equilibrium at prices p based on the Social Equilibrium Existence Theorem of Debreu [1952]:

Proposition 3: Under the above assumptions, for any vector of positive prices p and any choice of positive weights $\{\lambda_h^i\}$, there exists a social Nash optimum and, hence, a K-equilibrium, associated with those prices and weights.

Proof: We show that each planner faces a concave, continuous objective function and that his feasible strategy space is a convex, compact valued and continuous correspondence of the strategies of the other planners.

To see the concavity of the objective function, consider two alternative choices of constraints, $((\underline{Z}^1,\overline{Z}^1),\ldots,(\underline{Z}^n,\overline{Z}^n))$ and $((\tilde{\underline{Z}}^1,\tilde{\overline{Z}}^1),\ldots,(\tilde{\underline{Z}}^n,\tilde{\overline{Z}}^n))$. For any α, $0 \leq \alpha \leq 1$, the trade $\alpha \overset{\circ}{t}{}^i(p,\underline{Z}^i,\overline{Z}^i) + (1-\alpha)\overset{\circ}{t}{}^i(p,\tilde{\underline{Z}}^i,\tilde{\overline{Z}}^i)$ is feasible for the constraints

$(\alpha \underline{Z}^i + (1-\alpha)\ \tilde{\underline{Z}}^i,\ \alpha\overline{Z}^i + (1-\alpha)\tilde{\overline{Z}}^i)$. The concavity of the utility functions then implies the concavity of the planner's objective function. Continuity follows immediately from the continuity of preferences. The set of feasible strategies for planner h is defined by:

$$\{((\underline{Z}_h^1, \overline{Z}_h^1), \ldots, (\underline{Z}_h^n, \overline{Z}_h^n)) \mid \sum_{i=1}^{n} \overset{\circ}{t}{}_h^i\ (p, \underline{Z}_h^i, \underline{Z}^i)_{h(}, \overline{Z}_h^i, \overline{Z}^i)_{h(}) = 0\}^{10}$$

and is denoted by $\Gamma_h(\underline{Z})_{h(}, (\overline{Z})_{h(})$. Without loss of generality, we can restrict $\Gamma_h(\underline{Z})_{h(}, \overline{Z})_{h(})$ to canonical quantity constraints. $\Gamma_h(\underline{Z})_{h(}, \overline{Z})_{h(})$ is not empty because it contains (0,0). It is bounded because if t^i is the preferred vector in $\gamma_h^i(\underline{Z})_{h(}, \overline{Z})_{h(})$, $\min \{t_h^i, 0\} \leq \underline{Z}_h^i \leq 0$ and $0 \leq \overline{Z}_h^i \leq \max \{t_h^i, 0\}$. It is closed because of the continuity of the $\overset{\circ}{t}{}^i$'s. To see that $\Gamma_h(\underline{Z})_{h(}, \overline{Z})_{h(})$ is convex, choose $(\underline{Z}_h, \overline{Z}_h) = ((\underline{Z}_h^1, \overline{Z}_h^1)$ $, \ldots, (\underline{Z}_h^n, \overline{Z}_h^n))$ and $(\tilde{\underline{Z}}_h, \tilde{\overline{Z}}_h)$ in $\Gamma_h(\underline{Z})_{h(}, \overline{Z})_{h(})$ and consider $\alpha(\underline{Z}_h, \overline{Z}_h) + (1-\alpha)\ (\tilde{\underline{Z}}_h, \tilde{\overline{Z}}_h)$ for $0 \leq \alpha \leq 1$. If, for example, $\overset{\circ}{t}{}_h^i\ (p, \underline{Z}_h^i, \underline{Z}^i)_{h(}, \overline{Z}_h^i, \overline{Z}^i)_{h(}) = \underline{Z}_h^i$, then, because constraints are canonical, $\overset{\circ}{t}{}_h^i\ (p, \underline{Z}_h^i, \underline{Z}^i)_{h(}, \overline{Z}_h^i, \overline{Z}^i)_{h(}) = \underline{Z}_h^i$ and $\overset{\circ}{t}{}_h^i\ (p,\ \alpha\underline{Z}_h^i + (1-\alpha)\underline{Z}_h^i, \underline{Z}^i)_{h(}, \alpha\overline{Z}_h^i + (1-\alpha)\ \overline{Z}_h^i,\ \overline{Z}^i)_{h(}) = \alpha\underline{Z}_h^i +$ $(1-\alpha)\ \underline{Z}_h^i$. Similarly, for the upper constraints. That the correspondence Γ_h is upper semi-continuous follows from the continuity and convexity of preferences. It is also immediate that Γ_h is lower hemi-continuous. Finally, we can restrict the domain of Γ_h to only those rations $(\underline{Z})_{h(}, \overline{Z})_{h(})$ which could ever be canonical constraints. This domain is obviously closed, bounded and convex. We can thus apply the Social Equilibrium Existence Theorem to conclude that a social Nash optimum exists. Q.E.D.

Chapter 4

Order and Voluntary Exchange

We next observe that under our hypotheses of convexity, and differentiability, weak implementability is actually no weaker than implementability.

Proposition 4: If preferences are differentiable (we assume strict convexity throughout the paper), a weak implementable allocation is implementable.

Proof: If $(t^1,\ldots,t^n)$ is weakly implementable, then, for each i and $h \neq 0$, the partial derivative of i's utility function with respect to good h is equal to: $\lambda p_h + \mu_h^i - \nu_h^i$ where $\lambda(p.t^i) = 0$, $\mu_h^i(\bar{Z}_h^i - t_h^i) = 0$, $\nu_h^i(t_h^i - \underline{Z}_h^i) = 0$. Thus, the first order conditions for a utility maximum (subject to the budget constraint, and the quantity constraints) are satisfied. But, from convexity, the first order conditions are sufficient for a maximum. Hence, $(t^1,\ldots,t^n)$ is implementable. Q.E.D.

We can now demonstrate that if preferences are differentiable and there are at least three markets, order implies voluntary exchange.

Proposition 5: If preferences are differentiable and $m \geq 2$, an orderly allocation is implementable.

Proof: Consider an orderly allocation $(\bar{t}^1,\ldots,\bar{t}^n)$. If this allocation is not weakly implementable, then there exist i, h, and $\tilde{t}^i \epsilon \bar{\gamma}_h^i(\bar{t}^i)$ such that $\tilde{t}^i >_i \bar{t}^i$ and $\underline{Z}_h(\bar{t}^i) \leq \tilde{t}_h^i \leq \bar{Z}_h(\bar{t}^i)$, and, because $\tilde{t}_{h'}^i = \bar{t}_{h'}^i$, $\sum_{j \neq i} \bar{t}_{h'}^j + \tilde{t}_{h'}^i = 0$. Thus,

$(\bar{t}^i,\ldots,\tilde{t}^i,\ldots,\bar{t}^n)$ contradicts the order of $(\bar{t}^1,\ldots,\bar{t}^n)$, and so we conclude that $(\bar{t}^i,\ldots,\bar{t}^n)$ must be weakly implementable after all. From Proposition 4, $(\bar{t}^1,\ldots,\bar{t}^n)$ is thus implementable. Q.E.D.

That there be at least three markets and that preferences be differentiable are hypotheses essential for the validity of the preceeding proposition. Consider, for example, a two-market economy as represented in the Edgeworth box in Figure 1. Point A represents the initial endowment;

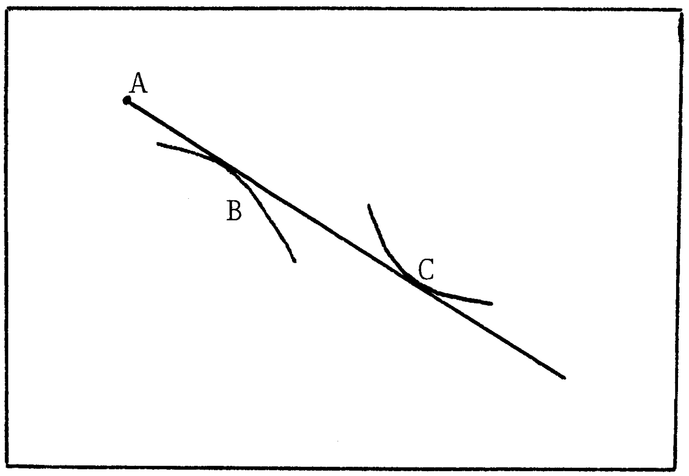


Figure 1

the line through A, prices; and the curves tangent to the line, indifference curves. Any allocation between B and C is clearly orderly, but not implementable since it involves forced trading by the agent whose indifference curve is tangent at B. To see that differentiability is crucial, consider a two-person three-good economy where agents have preferences of the form $\log \min \{x_1,x_2\} + \log x_0$. Given the preferences, we can treat goods 1 and 2 together as a composite commodity, since traders will always hold goods 1 and 2 in equal amounts. Thus the economy is, in effect, reduced to two goods, and so Figure 1 again becomes applicable. A trivial corollary of Proposition 5 is: Corollary 6: If preferences are differentiable and $m \geq 2$, {Orderly allocations} = {K - equilibrium allocations}.

We finally recall a result of Younès [1975] (also proved by Sylvestre [1978]),which a special case of corollary 6 when $m \geq 2$.

Proposition 7: (Younès) If preferences are differentiable: {Implementable allocations} $\cap$ {Weakly orderly allocations} = {K-equilibrium allocations}.

Proof: By definition, a K-equilibrium is implementable and weakly orderly. Conversely consider an allocation which is implementable and weakly orderly. From implementability, differentiability and convexity of preferences, the first order conditions are necessary and sufficient for maximization. It is then easy to check that the restrictions on the shadow prices of rations imposed by weak order and order are the same (see Younès).

Chapter 5

Optimality

We next turn to constrained Pareto optimality. We show that although a constrained Pareto optimal allocation is weakly orderly, it is ordinarily neither implementable nor orderly, at least when preferences are differentiable.

Proposition 8: A constrained Pareto optimum (CPO) is (i) weakly orderly (which implies that {Constrained Pareto optimal allocations}= {weakly orderly Pareto optimal allocations}) and (ii) with differentiable preferences, neither implementable nor (when $m \geq 2$) orderly, if it is not a Walrasian equilibrium allocation[11], if each trader is assigned a strictly positive weight in the program (*), and if there is some (i.e., non-zero) trade on every market.

Proof: Let $(t^1,...,t^n)$ be a CPO. If it were not weakly orderly, then trades could be altered on some market h, leaving trades on other markets undisturbed, in a Pareto-improving way, a contradiction of optimality. Therefore, (i) is established.

Suppose that $(t^1,...,t^n)$ is not a Walrasian equilibrium allocation, that preferences are differentiable, that there is non-zero trade on every market, and that all traders have positive weight in program (*). We will establish that $(t^1,...,t^n)$ is not implementable. Because it is not Walrasian, there exists at least one market h and one agent i who would prefer a trade different from t_h^i, given his trades on other markets $k \neq 0,h$. If, say, trader i is a net buyer of h, either he would like to buy more or to buy less of good h. If less, the non-implementability of $(t^1,\ldots,t^n)$ follows immediately. Assume, therefore, that he would like to buy more. Because, by assumption, there is non-zero trade on market h, there are traders who sell positive quantities of good h. If among these traders, there exists an agent j who would like to sell less of good h, the proof is, again, complete. If there exists j who would like to sell more than $-t_h^j$ units of good h (given his trades on markets other than 0 and h), i and j can arrange a mutually beneficial trade at prices p, contradicting constrained Pareto optimality. Therefore, assume that all sellers on market h are unconstrained. From differentiability, forcing them to sell a bit more of good h does not change their utility to the first order but does increase i's utility. Therefore, if the allocation assigns positive weight to i in (*), it involves forced trading. Thus $(t^1,\ldots,t^n)$ is not implementable. If $m \geq 2$, Proposition 5 implies it is not orderly. Q.E.D.

The hypothesis of differentiability in Proposition 8 is, as in previous results, essential. Crucial too is the assumption that all traders have positive weight in the program (*). To see this, refer again to Figure 1. Point B is both constrained Pareto optimal and implementable. However, the trader whose indifference curve is tangent to C has zero weight. (Note, incidentally, that all the other CPO's - which constitute the line segment between B and C - are non-implementable). Finally, the hypothesis of non-zero trade on each market is necessary. Refer, for example to the Edgeworth box economy in Figure 2. Initial endowments

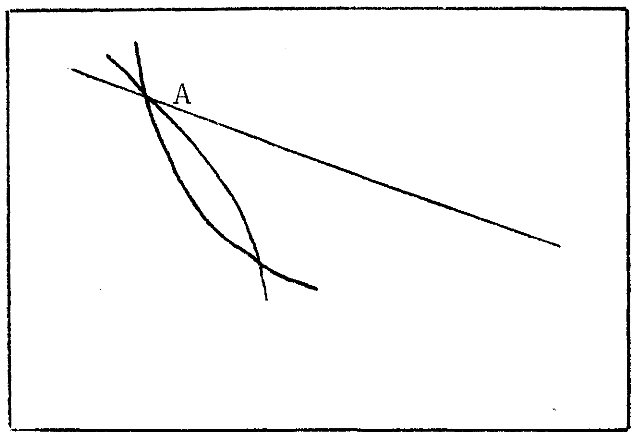


Figure 2

are given by A, which is also a constrained Pareto optimum relative to the price line drawn. Although A does not involve forced trading, it does not violate the Proposition, as it involves no trade at all.

Although differentiability is a restrictive assumption, the non-zero weight and trade assumptions rule out only negligibly many CPO's. On the basis of Proposition 8, we may conclude that, with differentiability, CPO's are generically non-implementable and non-orderly.

We now consider the set of Pareto optima among implementable allocations: the implementable Pareto optima. As opposed to a social Nash optimum where uncoordinated planners choose the rations on their own market according to their own weights, an IPO is a Pareto optimum for a unique planner choosing all the rations. Obvious questions are whether IPO's are necessarily orderly or even weakly orderly. The following proposition demonstrates that this is not the case.

Proposition 9: Implementable Pareto optima need not be weakly orderly (nor, *a fortiori*, orderly).

Proof: The proof takes the form of an example. Consider a two-trader, three-good economy in which trader A derives utility only from good 0 and has an endowment of one unit of each of goods 1 and 2. Trader B has a utility function of the form

$$U(x_0, x_1, x_2) = \frac{15}{8} x_1 + \frac{3}{2} x_2 - 3x_1x_2 - 3x_1^2 - x_2^2 + x_0$$

where x_i is consumption of good i, and an endowment of one unit of good 0. All prices are fixed at 1. It can be verified that trader B's unconstrained demands for goods 1 and 2 at these prices are $\frac{1}{12}$ and $\frac{1}{8}$, respectively. This is an IPO in which all the weight is assigned to trader B. In this IPO, trader A is constrained on both markets, and buys $\frac{1}{12} + \frac{1}{8} = \frac{5}{24}$ units of good 0. Now consider an IPO in which trader A buys $\frac{11}{48}$ units of good 0. If such an IPO exists, trader B must be constrained either on market 1 or 2. If the constraint is on market 2, we have:

(1) $x_1^B + x_2^B = \frac{11}{48}$ (because trader A buys $\frac{11}{48}$ units of 0)

and

(2) $6x_1^B + 3x_2^B = \frac{7}{8}$ (from maximization of utility with respect to good 1)

Solving equations (1) and (2), we find $x_2^B = \frac{1}{6}$, which is greater than B's unconstrained demand, $\frac{1}{8}$. Thus, if the IPO exists, trader B must be constrained on market 1. Now, if trader B is constrained from buying more than $\frac{1}{24}$ units of good 1, demand for good 2 is $\frac{3}{16}$. Notice that $\frac{3}{16} + \frac{1}{24} = \frac{11}{48}$. Thus, if trader B is so constrained and trader A is constrained from selling more than $\frac{1}{24}$ units of good 1 and $\frac{3}{16}$ units of good 2, the resulting allocation is an IPO. However, it is not weakly orderly, because given a purchase of $\frac{3}{16}$ units of good 2, trader

B would like to buy $\frac{5}{96}$ units of good 1. Since $\frac{5}{96} > \frac{1}{24}$, both traders A and B are constrained on market 1. Q.E.D.

K-equilibria do not have the welfare properties associated with Walrasian equilibria. In particular, it is possible for one K-equilibrium to Pareto dominate another.[12] Nonetheless, one might expect the Keynesian Pareto optima - the Pareto maximal allocations within the class of K-equilibria - to have "good" welfare properties. For instance, one might conjecture that they are IPO's. That this need not be so is demonstrated by the following:

Proposition 10: A KPO need not be an IPO.

Proof: The proof is again by example. Consider an economy similar to that of the proof of Proposition 9 but with two additional goods. Specifically take:

$$U^A = x_o + \frac{15}{8} x_3 + \frac{3}{2} x_4 - 3x_3x_4 - 3x_3^2 - x_4^2$$

$$U^B = x_o + \frac{15}{8} x_1 + \frac{3}{2} x_2 - 3x_1x_2 - 3x_1^2 - x_2^2.$$

Suppose that trader A has endowments of $\frac{19}{24}$, $\frac{1}{3}$, and $\frac{1}{3}$ units of goods 0,1, and 2, respectively, whereas B's endowments consist of $\frac{19}{24}$, $\frac{1}{3}$, and $\frac{1}{3}$ units of good 0, 3, and 4, respectively. All prices are fixed at 1. It can be verified that if unconstrained on markets 3 and 4, trader A demands $\frac{1}{12}$ and $\frac{1}{8}$ units, respectively, independent of constraints he faces on other markets. Similarly, trader B demands $\frac{1}{12}$ and $\frac{1}{8}$ units, respectively, of goods 1 and 2 if unconstrained on those markets. Thus, the unconstrained demand on all four markets are less than the unconstrained supplies: $\frac{1}{3}$ units in each case. Consequently, from order, the only possible K-equilibrium is one in which demand is unconstrained on every market. Trader A's equilibrium net trades vector is therefore $(0, -\frac{1}{12}, -\frac{1}{8}, \frac{1}{12}, \frac{1}{8})$. The two traders enjoy utilities of $\frac{205}{192}$ each. Because this is the unique K-equilibrium it is a KPO. Now suppose that trader B is constrained from buying more than $\frac{1}{24}$ units of good 1 and that A is constrained from buying more than $\frac{1}{24}$ units of good 3. It is easily checked that B will then demand $\frac{3}{16}$ units of good 2 and A, $\frac{3}{16}$ units of good 4. Thus, we obtain an implementable allocation in which tracer A's net trade vector is $(0, -\frac{1}{24}, -\frac{3}{16}, \frac{1}{24}, \frac{3}{16})$ and B's is $(0, \frac{1}{24}, \frac{3}{16}, -\frac{1}{24}, -\frac{3}{16})$. But these net trades generate utilities of $\frac{835}{768}$ for each trader. Because $\frac{835}{768} > \frac{205}{192}$, this implies that the KPO is not an IPO. Q.E.D.

We can summarize the results (with differentiability) so far in a schematic diagram (Figure 3).

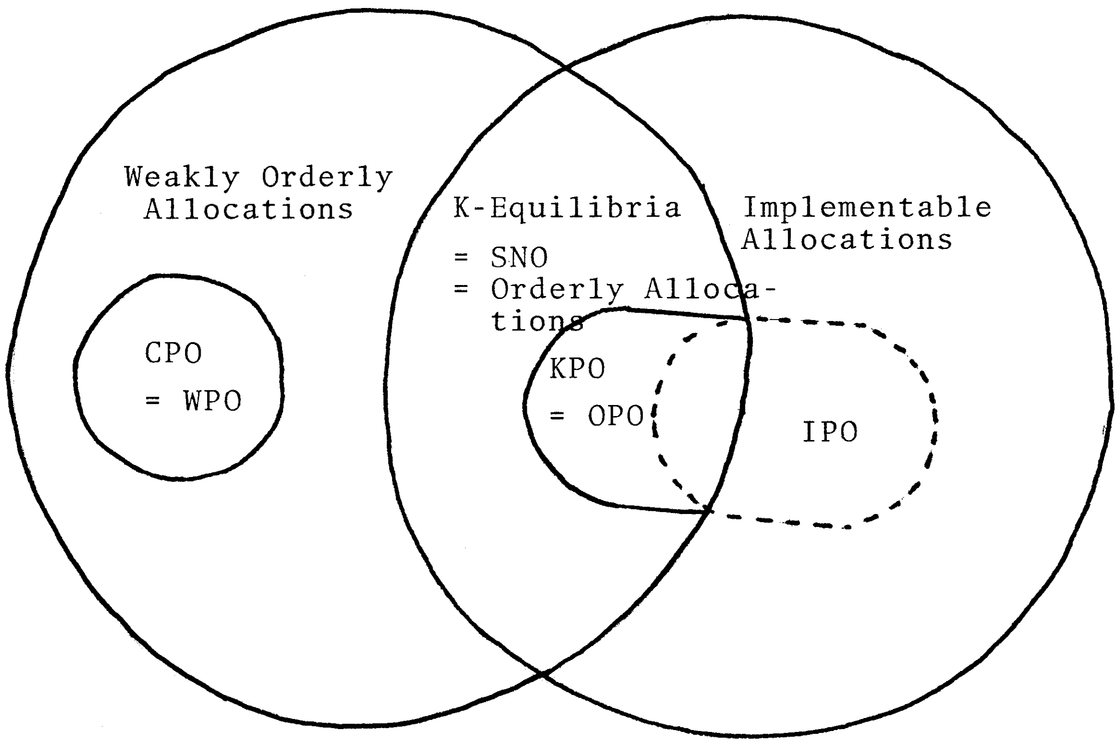


Figure 3

. The No Spillover Case

One "unappealing" feature of Figure 3 is that the set of IPO's is neither completely within or without the set of weekly orderly allocations, and, more specifically, the set of KPO's. However, with an additional hypothesis, this unaesthetic property disappears.

By the absence of spillovers, we mean that a change in a constraint on a market does not alter net trades in any of the other markets, except the unconstrained market. A sufficient condition to obtain no spillovers is that the traders' utility functions take the form $u^i = t_0^i + \sum_{h=1}^{n} \Phi_h^i(t_h^i)$. In the no spillover case, the only change in Figure 3 is that the IPO set shrinks to coincide with the KPO and OPO sets. We have:

<u>Proposition 11</u>: In the case of no spillovers, {IPO} = {KPO} = {OPO}.

<u>Proof</u>: An IPO must be orderly. Otherwise, slightly relaxing the constraints in market h for one demand-constrained and one supply-constrained agent would be implementable (since it would not disturb the other markets) and Pareto improving. Q.E.D.

Chapter 6

Structure of K-Equilibria, Constrained Pareto Optima and Keynesian Pareto Optima Sets.

In this section we undertake a study of the local dimension of the different sets.

K-Equilibria

Assume that at a K-equilibrium the m markets (h=1,...,m) comprise at least one binding constraint (a constraint is binding if the derivative of the indirect utility function with respect to the constraint is different from zero). Let us show that, on market h, the number of degrees of freedom is equal to the number of binding constraints (b^h), minus one;[13] for that let us remember that a K-equilibrium is a social Nash optimum. The first order condition for a social Nash optimum with a binding ration r_h^i yields $\lambda_h^i \frac{\partial v^i}{\partial r_h^i} = \mu_h$, where μ_h is the multiplier associated with market h's equilibrium constraint. Let

$$\tilde{\lambda}_h^i = \frac{\lambda_h^i}{\mu_h} \text{ and } F = \begin{bmatrix} \tilde{\lambda}_h^i \frac{\partial v^i}{\partial r_h^i} - 1 \\ \\ \sum_i \overset{o}{t}{}_h^i \,(p, \underline{Z}^i, \overline{Z}^i) \end{bmatrix}$$

Let $\tilde{\lambda}_h$ be the b_h-dimensional vector of modified weights $\tilde{\lambda}_h^i$. Consider the equation: $F(\tilde{\lambda}_h, r_h) = 0$, where r_h is the b_h-dimensional vector of binding rations on market h. The Jacobian of F is:

$$DF = \left[\begin{array}{ccc|ccc} \ddots & & 0 & \ddots & & 0 \\ & \frac{\partial v^i}{\partial r^i_h} & & & \tilde{\lambda}^i_h \frac{\partial^2 v^i}{\partial r^{i2}_h} & \\ 0 & & \ddots & 0 & & \ddots \\ \hline 0 & \cdots & 0 & 1 & \cdots & 1 \end{array}\right]$$

DF is of rank $(b^h + 1)$. Because F is a function of $(2b^h)$ variables, the inverse image $F^{-1}(0)$ is a manifold of dimension $(b^h - 1)$. The next step consists in projecting this manifold into the b_h-dimensional space of rations. It is easy to check that the projection has the same dimension as $F^{-1}(0)$. The local dimension of the set of K-equilibria is thus $\sum_h (b^h - 1) = b - m$, where b is the total number of binding constraints.

Constrained Pareto Optima

A constrained Pareto Optimum is a Pareto Optimum of the economy with consumption sets $\tilde{X}^i$ and induced preferences. Since we have assumed differentiable convexity, the local dimension of the set of constrained Pareto optima in the space of feasible allocations is (n-1), where n is the number of traders (A rigorous proof of this fact would follow the line of Smale (1976)).

Keynesian Pareto Optima

The only (direct) way to change the weight between two traders is to change their rations for a market on which they are both constrained. Call $T = \{(i,j) | i \text{ and } j \text{ are both on at least one market}\}$. $\tilde{T}$ is obtained from T by eliminating the redundant pairs; more precisely, in $\tilde{T}$, starting from i there can be at most one sequence of pairs: (i,j), $(j,k),\ldots,(\ell,i)$ leading back to i. The local dimension of the set of Keynesian Pareto Optima is then:

$\text{Min}\ [|\tilde{T}|,\ n - 1]$.

Footnotes

1. Benassy's equilibrium is a K-equilibrium if preferences are convex.

2. The term is due to Grossman (1977). Grossman's SNO, however, is related only in spirit to our own.

3. Younès (1975) calls this concept p-optimality.

4. The concept of weak order is due to Younès (1975) and Malinvaud-Younès (1977).

5. This concept was suggested to us by J.P. Benassy.

6. Younès (1977) calls this concept a p-equilibrium.

7. Grandmont, Laroque, and Younès (1978) call property (d') market by market efficiency. We shall sometimes use the shorter term:

   $\gamma_h^i\ (t^i)$ for $\gamma_h^i(\underline{Z}(t^i),\ \overline{Z}(t^i))$.

8. This equivalence is demonstrated by Grandmont, Laroque, and Younès (1978).

9. Note that, by analogy with (d), one can define a concept of weak order involving only two agents. It is easy to check that this definition and (d") (the analog of definition (d')) are equivalent.

10. If x is an m-dimensional vector, the notation $x_{)h(}$ denotes the vector $(x_1,\ldots,x_{h-1},\ x_{h+1},\ldots,x_n)$, i.e., the vector obtained by deleting component h.

11. With differentiable preferences, a Walrasian allocation is simply an allocation such that for each agent i and each good h, i's marginal rate of subtitution between h and the numéraire is equal to p_h (for details on the definitions of a Walrasian allocation under non-differentiability, see Silvestre (1978).

12. To see that this is so, recall that Hahn (1978) has shown that at Walrasian prices, a non-Walrasian K-equilibrium can exist. But this equilibrium must be Pareto dominated by the Walrasian equilibrium.

13. A constraint r_h^i is binding if $\frac{\partial v^i}{\partial r_h^i} \neq 0$

References

Benassy, J.P., (1975). "Neo-Keynesian Disequilibrium Theory in a Monetary Economy", Review of Economic Studies, 42.

Debreu, G., (1952) "A Social Equilibrium Existence Theorem", Proceedings of the National Academy of Sciences of the U.S.A., 38.

Drèze, J., (1975) "Existence of an Equilibrium Under Price Rigidity and Quantity Rationing", International Economic Review, 46.

Grandmont, J.M., (1977) "The Logic of the Fixed-Price Method", Scandinavian Journal of Economics.

Grandmont, J.M., Laroque, G., and Younès, Y., (1978) "Disequilibrium Allocations and Recontracting", Journal of Economic Theory, pp. 86-102.

Grossman, S., (1977) "A Characterization of the Optimality of Equilibrium in Incomplete Markets", Journal of Economic Theory, 15, 1-15.

Grossman, S., and Hart, O., (1979) "A Theory of Competitive Equilibrium in Stock Market Economies", Econometrica , 42, 293-330.

Hahn, F., (1978) "On Non-Walrasian Equilibria" , Review of Economic Studies, 45, 1-18.

Hahn, F., (1979) Comment on Grossman-Hart, Mimeo.

Malivaud, E., and Younès, Y., (1977) "Une Nouvelle Formulation Générale pour l'Etude des Fondements Microéconomiques de la Macroéconomie", Cahiers du Seminaire d'Econometrie, CNRS, Paris.

Silvestre , J., (1978a) "Fix Price Analysis: A Synopsis of Three Solution Concepts", Working Paper IP 270, Center for Research on Management Science, University of California, Berkeley.

Silvestre , J., (1978b) "Fix Price Analysis: The Classification of Disequilibrium Regimes", Working Paper IP271, Center for Research in Management Science, University of California, Berkeley.

Younès, Y., (1975) "On the Role of Money in the Process of Exchange and The Existence of a Non-Walrasian Equilibrium", Review of Economic Studies, 132, 489-502.

Essay II

Capital as a Commitment:
Strategic Investment in Continuous Time

(with Drew Fudenberg)

Chapter 1

Introduction

This paper analyzes the extent to which an early entrant in a new market can exploit its head start by investing more than it would, were all firms to choose their steady-state capital levels simultaneously. In the absence of depreciation, such "over-investment" by the lead firm is locked in, forcing its rivals to make their subsequent decisions taking the leader's capital stock as given. Their response, when anticipated by the leader, encourages it to invest to reduce the investment of other firms. This idea is due to Spence (1979), on whose model we base our analysis.

The effects of this temporal asymmetry on the solution to the investment game depend crucially on the solution concept used. One of the main purposes of this paper is to illuminate this point. We emphasize the shortcomings of the Nash solution concept, which permits firms to make empty threats; that is, threats which would not in fact be in the firm's interest to carry out, were its bluff to be called. In our model, firms have a quite powerful, although self-destructive, threat which can be used to enforce many seemingly implausible investment paths as Nash solutions. Thus, we base our analysis on Selten's (1965) concept of (subgame) "perfect equilibrium", which rules out empty threats. In extending Selten's discrete-time formulation to continuous-time games, we notice that solution concepts even more restrictive than our extension of perfectness may be desirable.

We relate our work to the entry-deterrence literature in Chapter 2. After laying out the model (Chapter 3) and explaining perfect equilibrium (Chapter 4), we consider the special case of no discounting (Chapter 5). We characterize the set of perfect equilibria, which includes, but is not confined to, Spence's solution. Chapter 6 singles out one particular equilibrium as the most reasonable solution. In the special case in which the two firms are identical, and enter the market at the same time, the outcome of the equilibrium we pick on the grounds of purely non-cooperative considerations is nothing but the dynamic joint profit maximization path. This solution, though appealing, is here justified by an ad-hoc argument which certainly merits a more detailed treatment, but the problem of refining the perfect equilibrium concept for infinite-horizon games is beyond the scope of this paper. We then (Chapter 7) analyze the discounting case. With discounting, Spence's solution is not in general Nash, and *a fortiori* not perfect. We conclude with a few remarks about entry deterrence.

In contrast to Spence's work, our results suggest that the steady state of the game will usually be on neither firm's steady-state reaction curve. The "follower" firm may be permanently restrained at some capital level below its reaction curve by the leader's credible threat to respond to investment by investing itself.

Chapter 2

Background

The study of entry deterrence and strategic interactions between firms with investments as the control variables has been motivated in large part by the realization that previous models using prices or quantities as controls were unrealistic. In traditional limit-pricing models, the monopolist's price (or quantity) before entry was taken by prospective entrants as a signal, or threat, of what the price (or quantity) would be post-entry, although no reason was given that the post-entry levels would indeed be the same. This lacuna prompted the criticism that the perceptive entrant should not be misled by the pre-entry level. The monopolist's threat was not credible, because it could and would deviate from the threat in the event of entry.

Capital levels, in contrast, do have commitment value. That is, given a large pre-entry capital stock, the "threat" to have a large post-entry stock is credible. However, for this threat to matter, there must be a reason that increasing the established firm's capital stock would decrease the payoff to entry.

Spence (1977), in a seminal paper, analyzed pre-emptive investment in a two-period model. In the first period, existing firms optimized their capital levels, considering the response of their second-period payoff to their first-period decision. If entry had occurred, established firms were assumed to either produce to capacity or to charge their marginal variable cost in the second period; however, given entry, neither reaction was likely to be profit-maximizing. While the assumption ensured that larger first-period capital stocks would decrease the payoff to entry, positing an ad-hoc rule for post-entry behavior finessed the question of why capital levels should be threats. Dixit (1980) considered the same two-period problem, using the more reasonable assumption of Cournot-Nash behavior in the second period. In his model, the established firm's first-period capital stock was a meaningful threat because it lowered the second-period marginal cost of the established firm, thus inducing it to have higher Nash equilibrium output in the second period.

Spence (1979) examined the problem in continuous time. The continuous-time framework offers a way to explicitly analyze post-entry behavior. The equilibrium payoffs thus computed can be used as "second-period" payoffs to consider the decision to enter and strategic pre-entry activity by the established firm. The continuous time specification also allows for varying degrees of headstart by the leader, while in the two-period models established firms could build up to any capital level before entry.

Spence found that the equilibrium was for the lead firm to invest as quickly as possible to some capital level and then stop. As this level was chosen knowing the follower's response, the result was much like the equilibrium in a static Stackelberg game.

While Spence used Nash equilibrium as the formal solution concept, the idea of perfect equilibrium underlies his paper and the historical development of the entry-deterrence literature[1]. The Nash concept in its usual formulation is most appropriate in static games. Taking one's opponent's strategies as given is reasonable when those strategies are only their current moves, but taking as given strategies which specify future moves, either as functions of time only, ("open-loop" strategies) or as functions of time and some state variables, ("closed-loop" strategies) allows empty threats. Thus the Nash concept is ill-suited to the investigation of commitment and credibility. The more restrictive perfectness concept is a refinement of Nash equilibrium which allows only credible threats. This restriction greatly reduces the size of the set of equilibria.

Chapter 3

The Model

A. Description

Our model is that of Spence (1979). We consider a market with two firms, firm one and firm two, each with an associated capital stock $K^i(t)$ $(i = 1,2)$. At time zero, firm two has just entered the market and has no capital, $K^2(0) = 0$; while firm one has an exogenously given capital stock $K^1(0) = K^1_0$. This is the "post-entry" game. Later, in Chapter 8, we will discuss firm one's behavior before firm two's entry[2].

We assume that given the capital stocks $K^i(t)$, there is at each time t an instantaneous equilibrium in the product market, with associated net revenues (i.e., total revenues minus operating costs) $\Pi^i\ (K^1(t),\ K^2(t))$. The Π^i could, for example, be the result of a Nash equilibrium in quantities given short-run costs at time t. This assumption implies that the choice of quantities at time t has no effect on the game later. Were it costly to change output levels, or were firms to rely on self-financing, so that by producing more now a firm could impair its rival's ability to invest, this assumption would be inappropriate. We make the assumption, as Spence did, to permit us to ignore quantity decisions and to focus on commitment via investment [3].

Capital is measured so that the (constant) price of investment (new capital) is one. Thus, each firm's instantaneous profit stream is $[\Pi^i(K^1(t),\ K^2(t)) - I^i(t)]$, where I^i

is firm i's investment at time t. Each firm tries to maximize its net present value of profits.

$$V^i \equiv \int_0^\infty [\Pi^i(K^1(t), K^2(t)) - I^i(t)]e^{-r^i t}\, dt$$

where r^i is i's discount rate. We assume that Π^i is differentiable, concave in K^i, and that $\Pi^i_{ij} < 0$. (See Appendix 1 for some examples). There is no depreciation so $\dot{K}^i(t) = I^i(t)$. Finally, no firm may disinvest, and each firm has a constant upper bound on the amount of its investment:

$$\forall t,\ I^i \varepsilon [0, \bar{I}^i]$$

We next define the functions $R_1(K^2)$ and $R_2(K^1)$ by $\frac{\partial \Pi^1}{\partial K^1}(R_1(K^2), K^2) = r^1$, and $\frac{\partial \Pi^2}{\partial K^2}(K^1, R_2(K^1)) = r^2$ (these are indeed functions by the concavity of the Π^i). They are the "steady-state reaction functions"; that is, they are the reaction functions of the following game:

Firm i chooses K^i to maximize its steady-state value, $SSV^i \equiv \frac{\Pi^i(K^1, K^2)}{r^i} - K^i$. This is a one-move game in which the firms simultaneously choose capital levels which they are required to maintain forever. In a slight abuse of notation, we will let R_j denote both the function $R_j(K^i)$ and its graph. We assume that R_1 and R_2 have a unique intersection which will be the Nash equilibrium of the steady-state game. We will denote this intersection as $N = (K^1(N), K^2(N))$ (note that we will write both $K^i(t)$ and $K^i(P)$, where P is a specified point). We further assume that at N the absolute value of the slope of R_1 is greater than that of R_2 so that the steady-state equilbrium is stable under the usual myopic adjustment process. We denote by S_1 the Stackelberg equilibrium of the steady-state game with firm one as the leader, which will be unique if, as we shall assume, $\Pi^1(K^1, R_2(K^1))$ is concave. $K^1(S_1)$ satisfies $\frac{d}{dK^1} \Pi^1(K^1(S_1), R_2(K^1(S_1))) = r^1$. We make the symmetric assumption for firm two.

We shall also wish to discuss the special case in which firms do not discount their payoffs. In this "no-discounting" case, firm i is assumed to maximize $\lim_{T\to\infty} \{1/T \int_0^T [\Pi^i(K^1(t), K^2(t)) - I^i(t)]\, dt\}$. We assume the monopoly profit is finite so this limit is bounded above. The corresponding steady-state game has objective functions $SSV^i = \Pi^i(K^1, K^2)$, so the corresponding reaction functions are defined by:[4]

$$\frac{\partial}{\partial K^i} \Pi^i(R_i(K^j), K^j) = 0.$$

We shall model the firms' investment paths as solutions to a non-cooperative game. Firm i's strategy is a function $I^i(K^1, K^2, t)$, specifying firm i's investment as a function of the state (K^1, K^2) and time. Note that firm i's investment is assumed to depend on the history of the game only through the current state.

Define the "Industry Growth Path" (IGP) from a given initial position as the locus on which each firm invests as quickly as possible. When the IGP from a given point goes to the right of N, we shall say that firm one is the leader at that point. (See Figure 1).

As there is no depreciation, from a point $P = (K^1(P), K^2(P))$ the state can only move to points Q with $K^1(Q) \geq K^1(P)$ and $K^2(Q) \geq K^2(P)$. We shall say that such Q are "accessible form", "after", or "above" P, and that P is "before" or "below" Q. We note lastly that the investment game is stationary.

B. Nash Equilibria

As a motivation for introducing perfect equilibrium, we consider three elements of the set of Nash equilibria in the special case in which firms do not discount their payoffs. Assume that in the initial position, $K^1(0) < K^1(N)$ and that firm one is the leader.

First Equilibrium: Firm i's strategy is "invest as quickly as possible until $K^i = K^i(N)$, then stop". Neither firm can gain by deviating, given the other's strategy, for if T_2 is the (fixed) time at which firm two stops investing, firm one is maximizing

$$\lim_{T\to\infty} \{1/T \, [\int_0^{T_2} [\Pi^1(K^1(t), K^2(t)) - I^1(t)]dt$$

$$+ \int_{T_2}^{T} [\Pi^1(K^1(t), K^2(t)) - I^1(t)] \, dt]\} =$$

$$0 + \lim_{T\to\infty} 1/T \int_{T_2}^{T} [\Pi^1(K^1(t), K^2(N)) - I^1(t)] \, dt$$

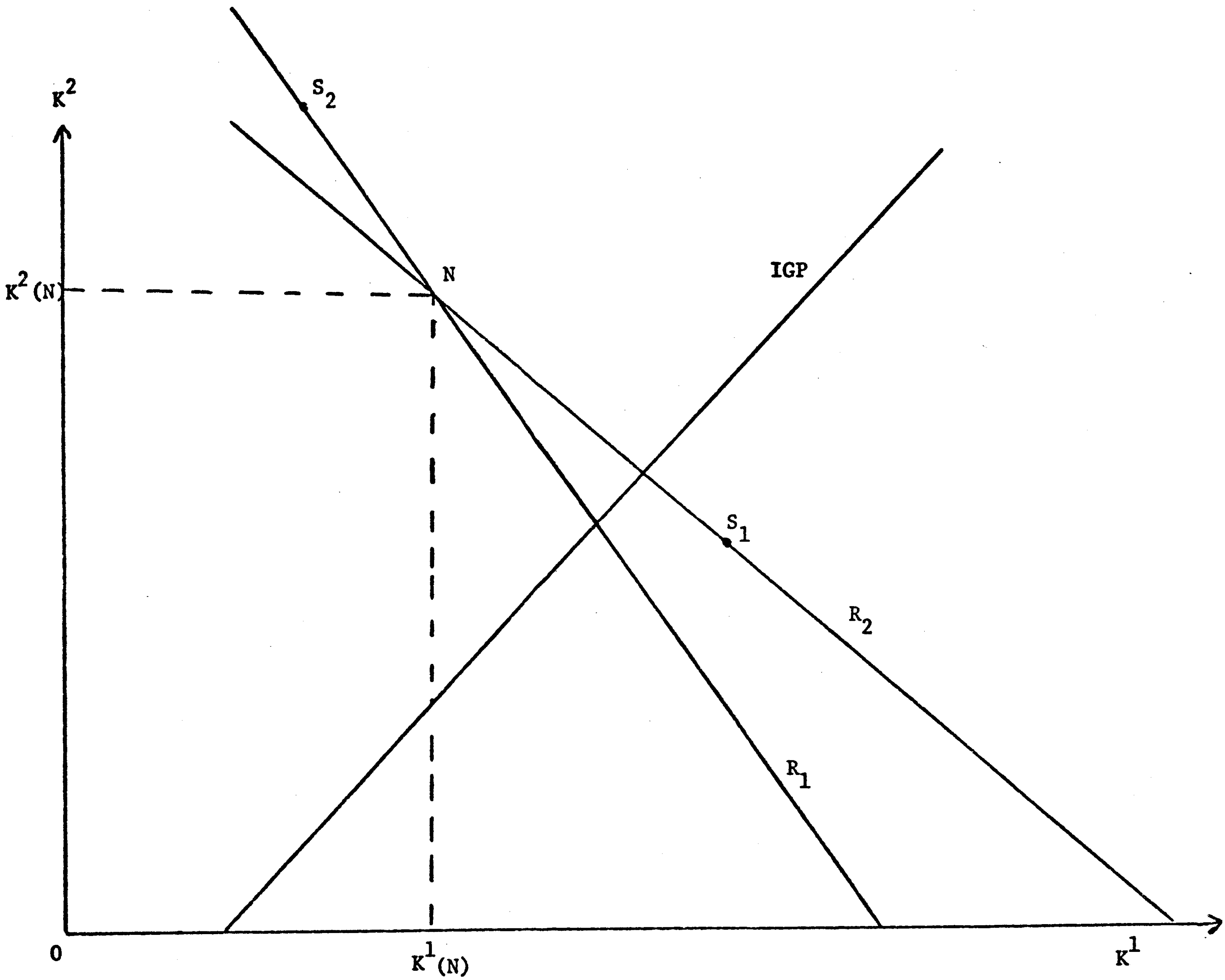


Figure 1

which is maximized if $\Pi_1^1 \; (K^1(t),\, K^1(N)) = 0$ (where the subscript indicates partial differentiation) for all t sufficiently large, that is if firm one invests to $K^1(N)$; and similarly for firm two. Note that in this example the speed of investment does not influence a firm's payoff (only the final state matters); in particular investing as quickly as possible to $K^1(N)$ is a best response for firm one to the given strategy of firm two.

Second Equilibrium: Firm one announces that if K^2 is ever greater than $\tilde{K}^2 = K^2(N) - \Delta$ $(\Delta > 0)$, it will invest until K^1 equals some $\tilde{K}^1$ such that $\max_{K^2} \Pi^2(\tilde{K}^1, K^2) < \Pi^2(R_1(\tilde{K}^2), \tilde{K}^2)$, otherwise firm one stops investing when K^1 reaches $R_1(\tilde{K}^2)$. Note that if Δ is sufficiently small, $\tilde{K}^1$ will exist. Let firm two's strategy be to invest until $K^2(t) = \tilde{K}^2$. Observe that this is a best response to firm one's strategy, and that, given firm two's strategy, firm one maximizes its payoff by stopping at $K^1 = R_1(\tilde{K}^2)$, as its strategy prescribes. Again, both firms are actually indifferent about the speed of investment.

Third Equilibrium: Firm two's strategy is "invest as fast as possible until the state hits (the graph of) R_2, and then stop". Now define $B(K^1(t),\, K^2(t)) = (K^1(B),\, K^2(B))$ to be the point on R_2 which maximizes firm one's payoff over the subset of R_2 which is attainable from $(K^1(t),\, K^2(t))$ given that firm two invests as quickly as possible to R_2. B is firm one's preferred point on R_2 between $(K^1(t),\, R_2(K^1(t)))$

and the intersection of the IGP through $(K^1(t), K^2(t))$ with R_2. If S_1, the Stackelberg point defined earlier, is attainable, $B = S_1$. Let firm one's strategy be "invest as quickly as possible until $K^1(t) = K^1(B)$, then stop".

Let us verify that this is indeed a Nash equilibrium: Given firm two's strategy, firm one maximizes its payoff by investing to $K^1(B)$, and it may as well do so as quickly as possible. Given firm one's strategy, if firm two invests as quickly as possible to R_2 the steady state will be at $B(K^1(0), K^2(0))$. Now observe that if firm two deviates by investing less quickly at some point below R_2, firm one will recalculate B as the attainable set will now include points on R_2 with lower K^2. As, however, the "old" B is still attainable, such a deviation by two cannot lead to a steady-state with $K^1 < K^1(B(K^1(0), K^2(0)))$, and thus cannot increase firm two's payoff. In this equilibrium, the speed with which firm two invests does matter. If it slowed and if S_1 were to the south-east of the attainable set on R_2, then firm one would have both the opportunity and the incentive (from the concavity of $\Pi^1(K^1,R_2(K^1)))$ to drive the steady-state to the south-east along R_2.

All three of the above equilibria are Nash, but equilibria one and two rely on empty threats. That the latter does we trust is obvious. The first equilibrium rests on an empty threat in the following way. If firm one did not stop at $K^1(N)$ but at $\bar{K}^1(N) = K^1(N) + \varepsilon$, the state would still be below R_2 when firm one stops. Were firm two to optimize from this point it would invest only to $R_2(\bar{K}^1)$, not to $K^2(N)$. Firm two threatens to invest to $K^2(N)$ regardless of what firm one does: but this threat is empty because firm one could anticipate the threat would not be carried out were firm one to present firm two with the fait accompli of investment to $\bar{K}^1$.

The first equilibrium is an "open-loop" equilibrium, that is an equilibrium for the game in which the players precommit themselves to time-paths of investment at the start. The second equilibrium is a "closed-loop" equilibrium in that the strategies prescribe moves which depend on the state, and not just on time. Note that open-loop equilibria can be viewed as (trivial) closed-loop equilibria in which the move at time t is a constant function of the state. The third equilibrium is also closed-loop in that the strategies depend on the state; moreover, it has the appealing property of not being enforced by empty threats, as we will show.

Chapter 4

Perfect Equilibrium

To make this idea of "no empty threats" more precise, we use Selten's (1975) concept of (subgame) perfect equilibrium. Before giving the definition, we discuss Selten's example. Figure 2 describes a game in extensive form. Player one has a choice of up or down; if he plays down, player two can choose up or down. The payoffs to each player are indicated at the ends of the tree.

There are two Nash equilibria in this game. In the first, player one plays up, and player two plays "if you go down, I'll go up", in the other, player one plays down, and two plays down. The first equilibrium is like the path to N in our investment game in that it relies on a threat by player two to do something he actually wouldn't do. Were player one to play down, player two would play down. Perfect equilibrium requires that players predict their opponents' future moves on the basis of their knowledge of the game, instead of being "deceived" by announcements.

Definition: A set of strategies for a game in extensive form is a perfect equilibrium if the strategies yield a Nash equilibrium for every subgame; where for every node of the game tree, the associated subgame is the game beginning at that node.

In Selten's example, the only Nash equilibrium at the lower node is "player two plays down", since the Nash

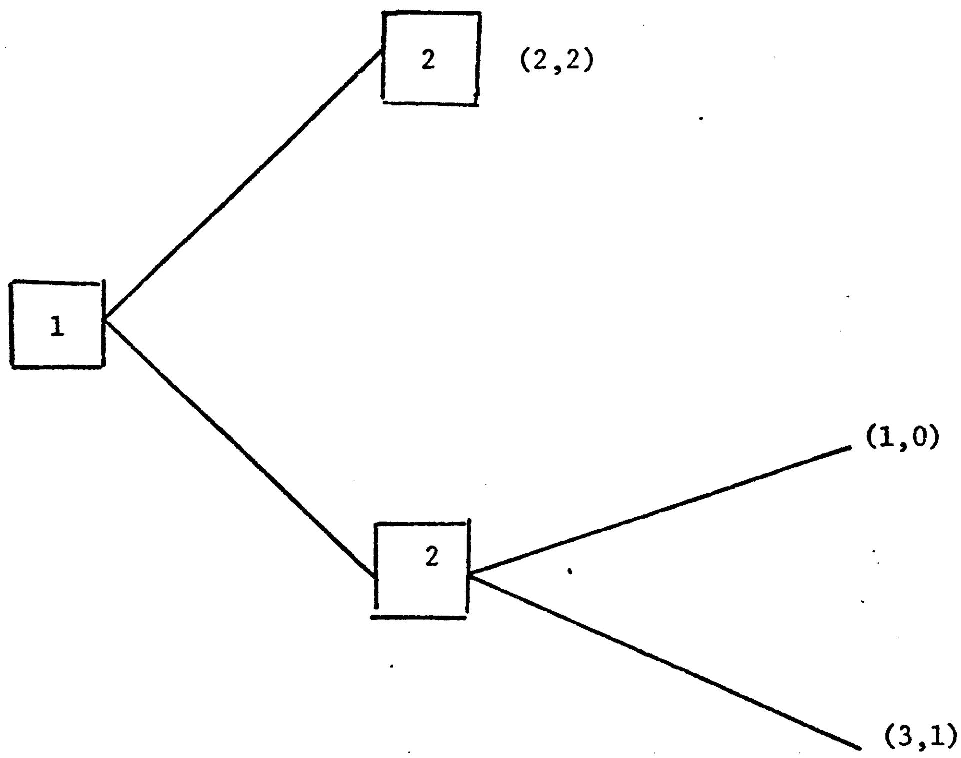


Figure 2

equilibrium of a one-player game is equivalent to that player making an optimal choice. Thus, only the "good" Nash equilibrium, (Down, Down), is perfect. Note that in finite-horizon discrete-time games of perfect information, such as the example, perfect equilibria are obtained by backwards induction from the terminal nodes. In games of perfect information this process will result in a simple choice by one player at each node.

Note also that if we change the (1,0) payoff to (1,1) there will be two Nash equilibria at the lower node, and thus two different perfect equilibria, depending on which Nash equilibrium we specify in the second period. In finite games of perfect information one expects the perfect equilibrium path will be unique, as the one-player games at each node will typically have unique solutions, except on small sets of singularities[5].

We extend the definition of perfect equilibrium in a natural way to continuous-time by identifying nodes of the game with points in state space. Thus, we require that the strategies yield a Nash equilibrium for the sub-game starting from any point (K^1,K^2) in the state space (because of the stationarity of the game, we will, after the discussion of the previous equilibria, exclude time from the state space).

The investment game is simultaneous-move and thus is <u>not</u> a game of perfect information. When choosing a move at time t, neither player knows the other's move at that instant. Moreover, the investment game does not have a fixed terminal time. For these reasons, we should not expect the game to have a unique perfect equilibrium. Indeed, we shall demonstrate the existence of an infinity of perfect equilibria in the no-discounting case[6].

Chapter 5

Analysis of the No-Discounting Case

We now return to the no-discounting case, to see if any of our proposed strategy pairs are perfect equilibria. The strategies which gave the path to N are not perfect: if firm one deviated by investing a bit past $K^1(N)$, to $\bar{K}^1$, firm two's only Nash response would be to invest only up to $R_2(\bar{K}^1)$, not up to $K^2(N)$ as announced. Thus the strategies are not Nash at the point $(\bar{K}^1, R_2(\bar{K}^1))$, where firm two's Nash response would be to stop, but its strategy says to invest. In the second equilibrium, firm one induced firm two to stop at $\tilde{K}^2$ by threatening to punish any further investment by investing so much that firm two could not possibly gain. The moves prescribed by the strategies at $(R_1(\tilde{K}_2), (\tilde{K}_2 + \varepsilon))$ are "firm one invests, firm two does not". This investment by firm one drives the state to the east, away from R_1, and can only lower firm one's payoff. That is, "firm one invests to $\tilde{K}^1$, firm two does not invest" is not a Nash equilibrium for the subgame starting from $(R_1(\tilde{K}^2), \tilde{K}^2 + \varepsilon)$. So the second equilibrium is not perfect.

To show that the path of the third equilibrium _is_ the result of perfect equilibrium strategies, we must extend the strategies given to cover the rest of the state-space.

We divide the state-space in four regions: (1) the set not below either reaction curve; (II) the set below R_2 and to the right of the IGP through N; (III) the IGP through N, below the reaction curves; and (IV) the set below R_1 to the left of the IGP through N. (Figure 3). Next we specify strategies for each region:

<u>Region I</u>: Neither firm invests.

<u>Region II</u>: Firm one invests as quickly as possible until K^1 reaches the level associated with the steady-state Stackelberg point, S_1; that is, $I^1(t) = \bar{I}^1$ if $K^1(t) \leq K^1(S_1)$, and zero otherwise. Firm two invests as quickly as possible.

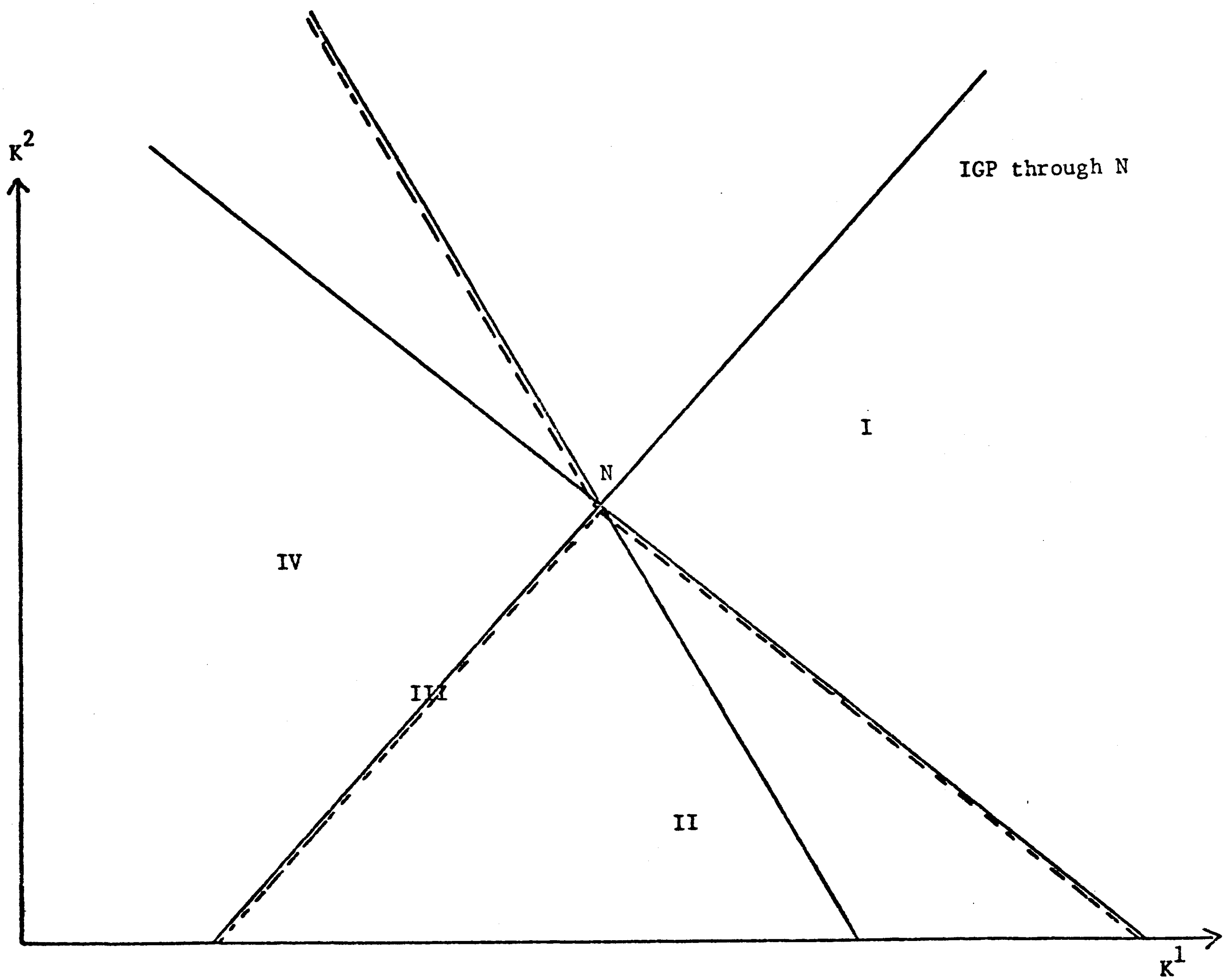
<u>Region III</u>: Both firms invest as quickly as possible (i.e., the state moves along the IGP).

<u>Region IV</u>: Symmetric with Region II.

With these strategies each player is either investing as fast as possible, or not investing. Thus we can characterize the direction of movement at a point as one of four vectors: ↑, →, ↗ , or ·. The proposed strategies are shown in Figure 4.

<u>Proposition 1</u>: These strategies form a perfect equilibrium.

Although the proof is somewhat lengthy, we include it in the text to help readers unfamiliar with perfect equilibrium with the style of reasoning involved; subsequent proofs are given in appendices. Readers familiar with extensive form games may easily convince themselves of Proposition 1 by looking at Figure 4, and skip the proof.



<u>Figure 3</u>

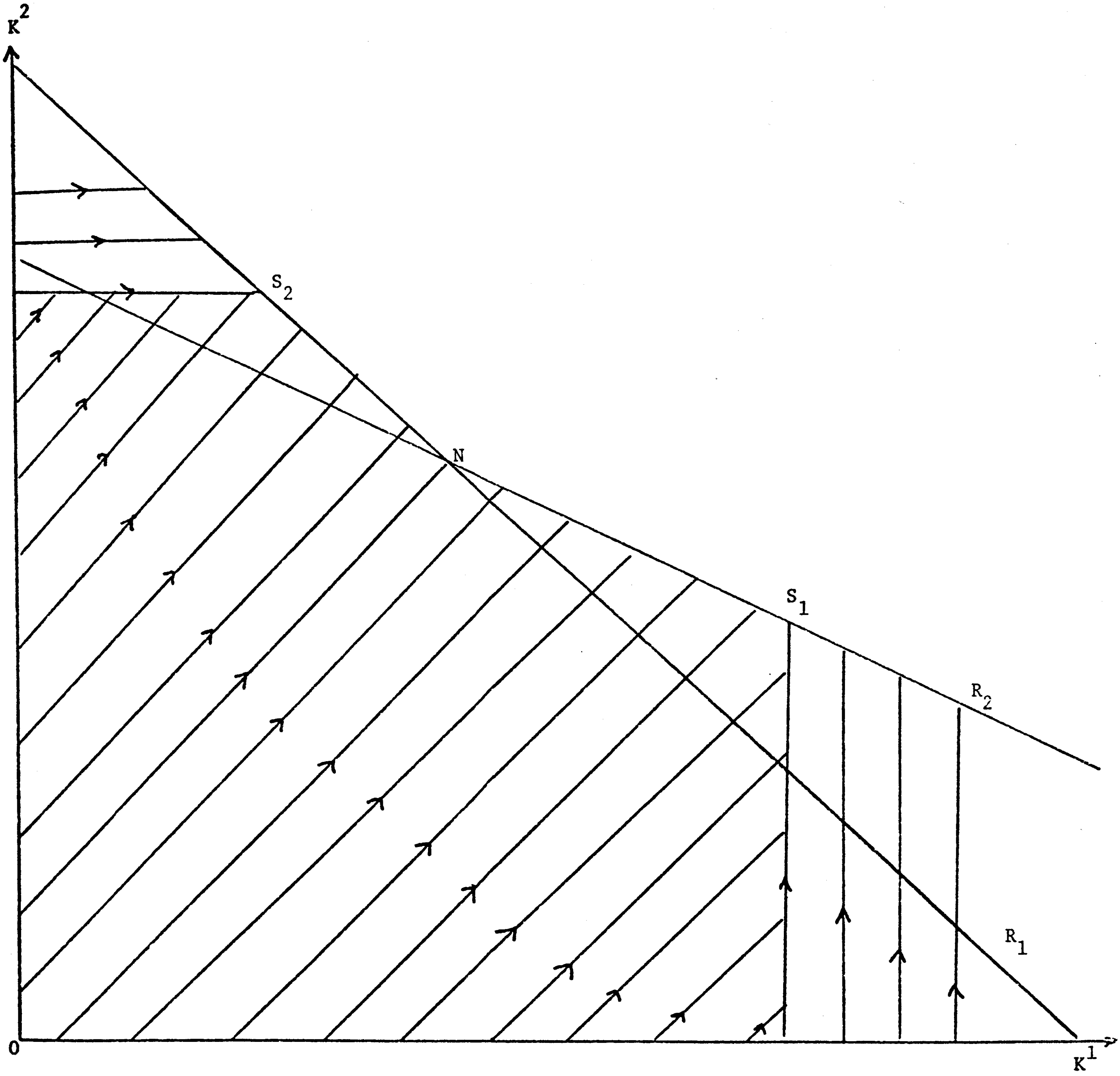


Figure 4

Proof: We must check that the strategies yield a Nash equilibrium at every point. We first examine the strategies in Region I. Once the state is in this region, it can never leave. Each firm takes its rival's strategy as given, and so assumes its rival will never resume investment. Each firm will thus maximize its payoff given its rival's fixed capital stock. As the Π^i were assumed to be concave in K^i, each firm wishes to be as close to its reaction curve as possible, which, in Region I, requires that the firm not invest. So the strategies yield a Nash equilibrium in Region I.

Next we consider starting at an arbitrary point in Region II. Is firm two's strategy optimal from any initial point in the Region, given firm one's strategy? First we will show that, given firm one's strategy, the best firm two can do, starting from a point in the Region, is to choose an investment path which leads to a steady-state on R_2. No investment path of firm two could lead the state into Regions III or IV. Moreover, starting at any point in Region II, firm one will stop investing in finite time, for any investment path of firm two. Then, because when firm one stops the state will be on or below R_2, the best firm two can do is to choose a path leading to a steady-state on R_2, as we already know that once the state reaches R_2 the best firm two can do is to stop.

Firm two's proposed strategy in Region II is to invest as quickly as possible, so it can deviate only by investing less quickly. But we observe that, given firm one's strategy, such a deviation can never lead to a steady-state preferred by firm two. Therefore, firm two cannot gain by deviating.

Is firm one's strategy optimal from any initial point in Region II, given firm two's strategy? Assume for the moment that firm one is constrained to investment paths such that the state does not enter Regions III or IV. [We will later show that this constraint is not binding]. This means that whatever firm one does, given firm two's strategy, the state will go to R_2, where, as we have seen, it will remain forever. Consider the following maximization problem for firm one: choose a steady-state Q* in the set of points on R_2 which are attainable from the initial point, given that firm two invests as quickly as possible. From our assumption that $\Pi^1(K^1,R_2(K^1))$ is concave, Q* is unique and is either S_1 or the attainable point nearest it. The proposed strategy for firm one in Region II "invest as quickly as possible until $K^1 = K^1(S_1)$" is equivalent to "invest as quickly as possible until $K^1(t) = K^1(Q^*)$" and is thus optimal for firm one, given firm two's strategy and given that the state does not enter Regions III and IV.

We now check that firm one could not gain from choosing an investment path which allowed the state to enter Regions III or IV. By an argument analogous to that for firm two in Region II, the best firm one can do in Regions III or IV is to arrive at a steady-state on R_1. However, this steady-state can not be better for firm one than the one at N, which is attainable from the initial point in Region II whenever firm one can send the state in Regions III or IV. So firm one's strategy in Region II is a best response to that of firm two, and we have shown that the strategies yield a Nash equilibrium from any initial point in Region II.

In Region III, a deviation by either firm makes the other the leader, and we have seen that firm one will not do this. The argument for firm two here, and for both firms in Region IV, are exactly symmetric. Q.E.D.

So we do have a perfect equilibrium. This equilibrium is the solution proposed by Spence. The lead firm invests past its steady state reaction curve to deter investment by its rival. When firm one is able to invest to $K^1(S_1)$ before the state reaches R_2, the steady state Stackelberg point is the outcome. When this is not possible, the firms are in a race to R_2.

While the solution is asymmetric, it does not rely, as does Stackelberg equilibrium in quantities on the assumption of asymmetric behavior. Instead, the asymmetry is derived from asymmetric initial conditions. Of course, the Stackelberg equilibrium quantities can be viewed as a (perfect) Nash equilibrium in which the leader chooses its quantity first. A frequent criticism of the Stackelberg equilibrium in quantities is given that both firms would like to move first, why should one firm let another be the leader? Both firms would prefer to be the leader in the investment game as well. The point is that it is easier to believe that a firm might start investing first, than to believe that it might always manage to produce first.

We have demonstrated one perfect equilibrium for the no-discounting case; it is far from unique. Moreover, as we shall show in Chapter 7, it will not in general be an equilibrium with discounting. We now establish the existence of a set of perfect equilibria, and indicate which equilibrium in the set we consider to be the most reasonable; this equilibrium moreover has the advantage of carrying over to the discounting case.

Notice that, along an IGP in Region II, there exists a point P such that firm two prefers that the state remain constant forever to the state moving along the IGP to R_2 and stopping there forever; and that, past R_1, firm one always prefers that all investment cease (see Figure 5). Assume that P is to the north-west of the IGP going to S_1. Then, as we have seen, the strategies "invest as quickly as possible to R_2" are a Nash equilibrium for the subgame starting at P. Given the strategies at all points above P ("stop, "stop") at $(K^1(P), K^2(P))$ is also a Nash Equilibrium at P. We will use this observation to construct a set of perfect equilibria, each with this "early stopping" property.

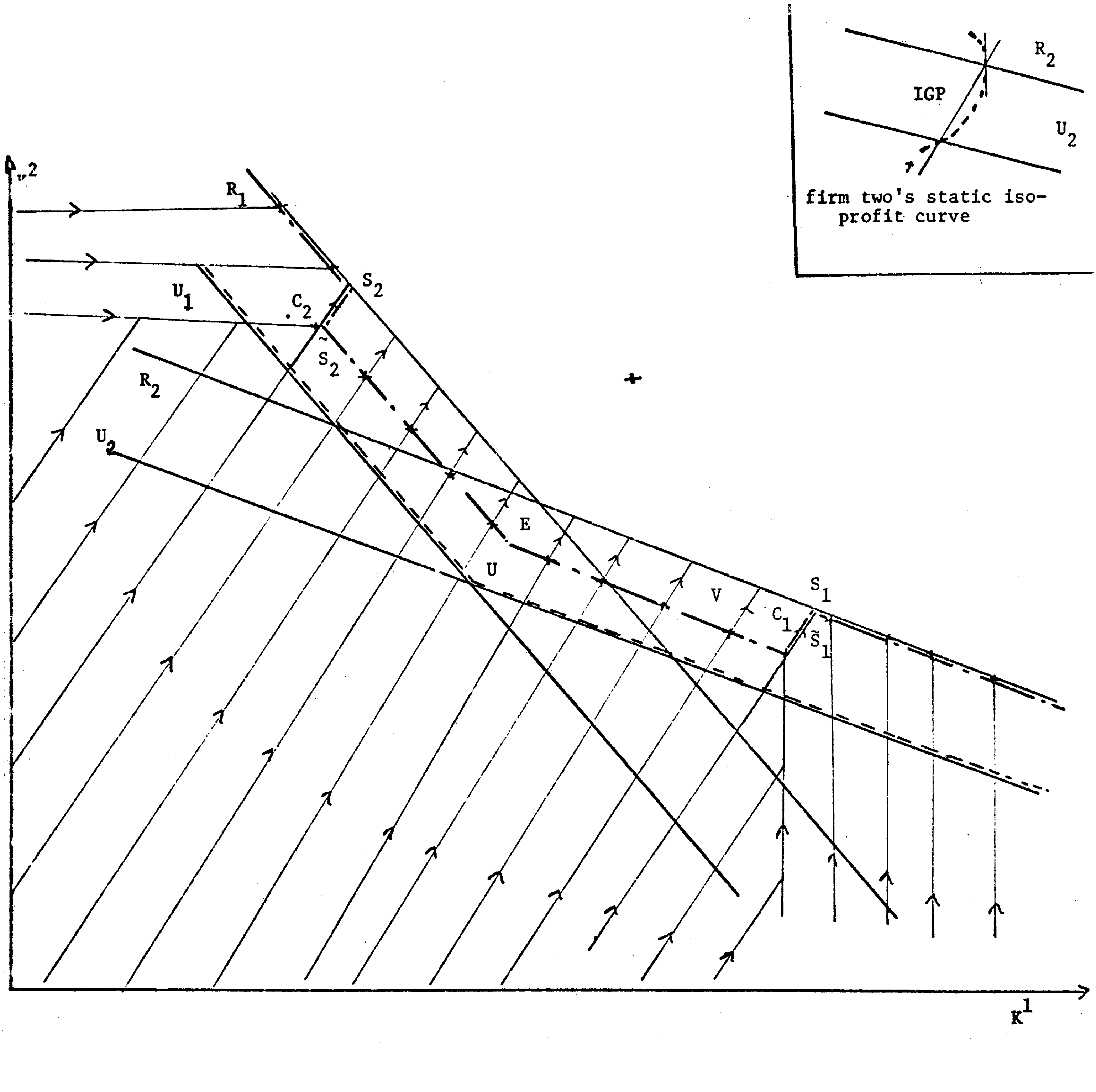

R2
IGP
U2
firm two's static iso-
profit curve
R1
S2
U1
C2
S2
R2
U2
E
U
V
S1
C1
S1
K1

Figure 5

To do so, we define U_2 as the set of points at which firm two is just indifferent between the state remaining constant forever, and the state moving along the IGP to R_2 and then staying there; define U_1 similarly and let U be the upper envelope of U_1 and U_2. Let C_1 (respectively C_2) for "credible", be the IGP from U_2 (U_1) to S_1 (S_2). At all points southeast of C_1 "stop" is firm one's Nash response to "invest to R_2", so firm one cannot credibly threaten to invest to R_2, and there cannot be an early stopping equilibrium. To the northwest of C_1, "invest to R_2" is firm one's best reply to "invest to R_2", so there firm one does have a credible threat to invest.

We now state:

Proposition 2: Take any downward sloping line E connecting C_1 and C_2, above U and below the upper envelope of the reaction curves. Call T - for terminal surface - the line formed by E, C_1 and C_2 between E and the reaction curves, R_1 to the left of S_2 and R_2 to the right of S_1 (one such T is depicted in Figure 5). One can construct perfect equilibrium strategies such that the equilibrium path stops on T. Moreover, the terminal point will be on E if $\underset{\sim}{K}^1(0) < K^1(\tilde{S}_1)$ and $K^2(0) < K^2(\tilde{S}_2)$, where $\tilde{S}_i$ is the intersection of E and C_i.

Proof: The proof, essentially identical to the proof of Proposition 1, is given in Appendix Two.

Note that all the steady states between U_2 and R_2 are Pareto-superior to the steady state on R_2 where the state would end up if both players invested as fast as possible. Once the state crosses R_1, firm one's only incentive for investing is to reduce firm two's investment, so it definitely prefers that they both stop sooner.

Above U_2, while firm two would prefer to invest were firm one to uncontingently stop, firm two is deterred from investment by firm one's credible threat to invest up to R_2. This threat is not credible to the right of C_1, because here firm one prefers either the point on R_2 directly above (which would be the steady state were firm two alone to invest), or S_1, to the point where the IGP intersects R_2. In this region firm one cannot deter two's investment.

In summary, while the concept of perfect equilibrium allowed us to discard Nash equilibria based on empty threats, it does not yield a unique solution to the no-discounting game. For this reason, we proceed in the next section to further restrict the equilibrium set.

Chapter 6

Restricting the Set of Perfect Equilibria

We feel that not all of the perfect equilibria described above are equally reasonable. Unfortunately, no refinement of the perfect equilibrium concept which yields unique solutions to infinite-horizon games is yet well-established. For this reason, we shall content ourselves with a very informal argument which we hope will convey our intuition.

We begin by adopting Harsanyi's ([1964], pp. 678-79) view that "the essential difference between cooperative and non-cooperative games consists only in the fact that in the latter the players are unable to cooperate in achieving a payoff vector *outside* the set of equilibrium *points* - however desirable this may be for all of them - but there is no reason whey they should not cooperate *within* the set of equilibrium points". Thus, we shall speak of the firms coordinating their strategies while insisting that we are still discussing a non-cooperative game. The issue is *which* equilibrium the players should expect. The search for a unique equilibrium can be seen as an extension of the idea of perfectness: What should the players expect to happen in each subgame? Perfect equilibrium says that the only reasonable thing to expect is a Nash equilibrium, but does not suggest which of several Nash equilibria to expect. We shall therefore base our argument for a particular equilibrium on a discussion of the coordination of the players' expectations.

As a means of representing this coordination, we introduce the idea of contracts <u>between</u> the players (see Aumann): at any time, either player can propose a contract which specifies a particular perfect equilibrium from that point on. Let us emphasize that contracts are not binding - they simply are a means of coordinating on a perfect equilibrium. Each firm compares its payoff from accepting the contract to what it can obtain by not accepting and in making this comparison considers the possibility of contracting later on. We impose perfectness on the contracting process in the sense that no firm can threaten to later refuse a contract which it would in fact accept. These contracts will enable us to rule out a great many perfect equilibria, and to say that the equilibrium should involve "early stopping", but we will be forced to introduce a second argument to obtain a unique equilibrium.

We first rule out any perfect equilibrium which involves investment above the reaction curves. Such equilibria do exist, with each firm investing only to limit the steady state capital stock of the other. These equilibria can be likened to two people beating each other the head, each attacking only to induce the other to stop sooner. Both firms have an incentive to propose ("stop", "stop") as the equilibrium above the reaction curves, because ("stop", "stop") strictly Pareto dominates all other outcomes of subgames starting in this region.

Having pinned down what happens in Region I, we now turn to the game below R_2 and to the right of the IGP going to N. Past R_1, firm one is investing only to reduce firm two's capital accumulation. Thus, firm two might be able to anticipate that, past R_1, firm one would always be willing to stop if firm two did; that is "stop" is firm one's only Nash response to "stop". Roughly speaking firm two could become the leader, in the sense that firm two could choose its investment path knowing that when it stopped, firm one would too.

We would therefore like to introduce the locus I_2, which is defined as those points at which firm two's steady state isoprofit curve is tangent to the IGP. I_2 is below R_2; its computation is performed for a simple case in Appendix 3. The terminal surface constructed from I_2,I_1 (the analog of I_1 with firm two as the leader), C_1,C_2 and the reaction curves will be called E* (see Figures 6 and 7). Firm two prefers the state to stop at any point on or above I_2 rather than moving along the IGP through that point and then stopping. Thus on I_2, firm two would prefer to stop investing knowing that firm one will follow suit, to investing and provoking strategic investment by firm one. Were firm two faced with this choice at E*, either to stop or to provoke firm one's investment, then E* would be a

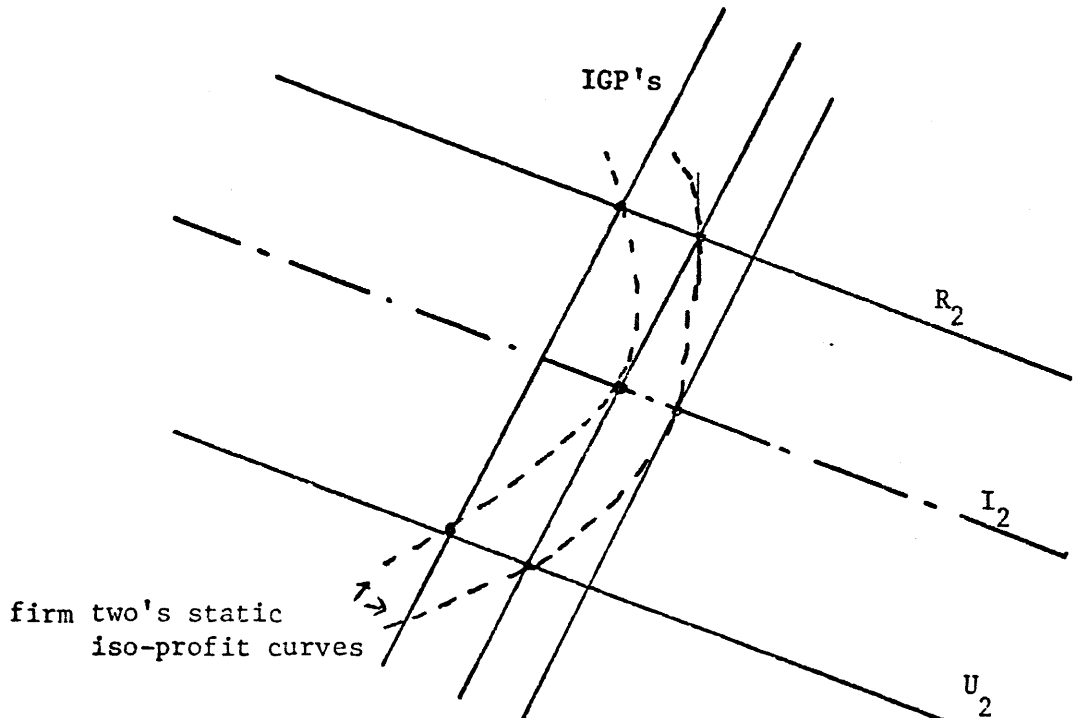
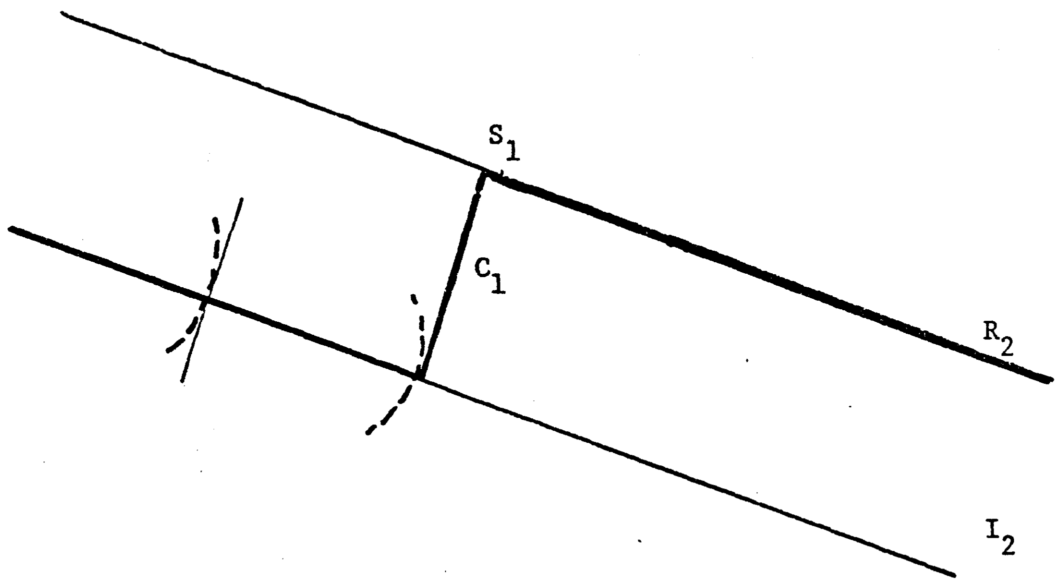


Figure 6



<u>Figure 7</u>: construction of E* in the no discounting case

"natural" terminal surface, with firm two's (dominant) strategy below E* being "invest".

We claim that equilibria which yield stopping below E* are not reasonable. Whatever firm one's behavior below I_2, firm two gains by investing up to I_2, given that firm two can arrange for the ("stop, "stop") equilibrium to occur at I_2. We now use the idea of mutually beneficial contracts to argue that, at I_2, firm two can indeed arrange for the ("stop", "stop") equilibrium: This equilibrium is the best possible outcome for firm one for the subgame starting at I_2, and firm one would therefore always accept a contract which specified it. Note that while firm one <u>below</u> I_2 might like to <u>claim</u> that it would not agree to stop at I_2, we are imposing perfectness on the game with contracts: regardless of firm one's announcements, firm two correctly anticipates that firm one would indeed agree to stop at I_2 (or for that matter, anywhere above R_1).

Next we examine the game starting from any point A on I_2 (see Figure 8). First assume that the firms coordinate at A, stipulating that they both stop on some terminal surface. What are the reasonable stopping points? Introduce firm two's isoprofit curve through A; any agreed upon steady state has to be to the left of this curve, since firm two can always guarantee itself the payoff corresponding to the steady state at A. For a similar reason, the steady state has to be below firm one's isoprofit curve through the intersection of R_2 and the IGP through A. We now have the set

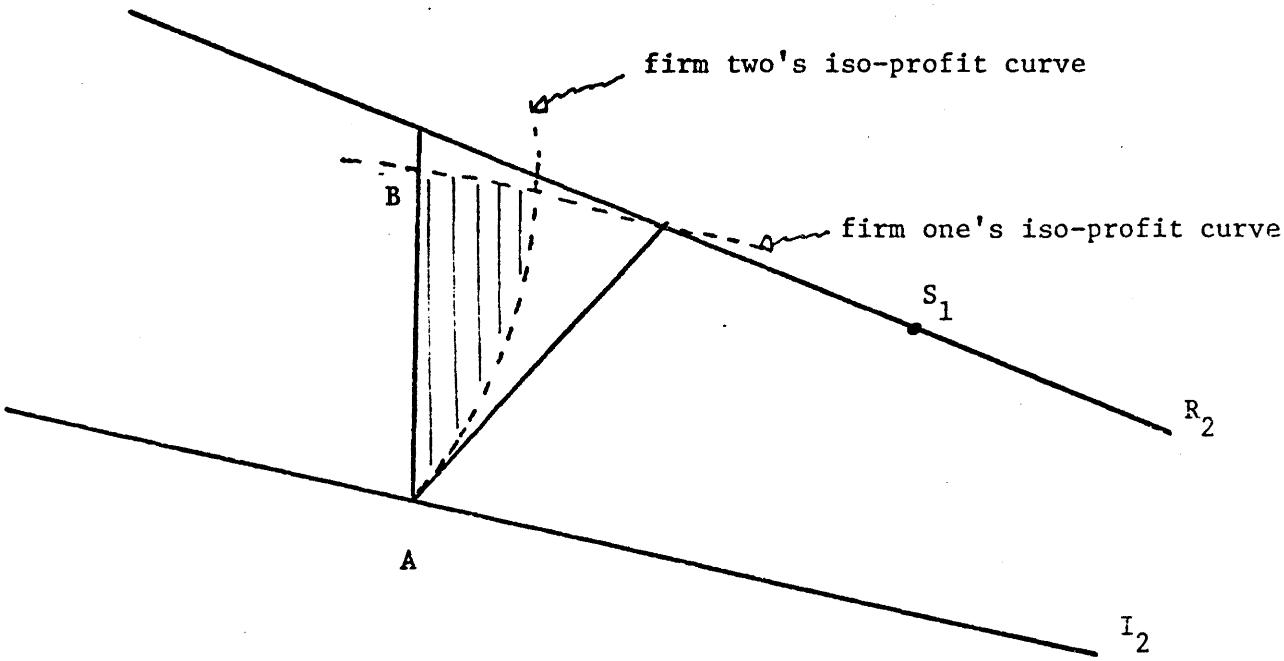


Figure 8: The "contract zone"

of potential cooperation steady states (see Figure 8) from which we conclude that <u>any equilibrium in the game with contracts must be an "early-stopping" equilibrium</u>, as were the players to expect the equilibrium to go to R_2 they could find a contract which they both preferred. The "Pareto-optimal" steady states in this set are the steady states vertically above A, between A and B; they all are the steady states of some perfect equilibrium starting from A (take horizontal terminal surfaces through the steady state).

Should firm one agree to coordinate on any "Pareto-optimal" perfect equilibrium other than the one stopping at A (on I_2)? Although we have no formal argument to offer, we feel that the answer is no. Our intuition is based on the fact that advance contracts allow firm two to commit itself to stop later and thus to prevent firm one's strategic investment. It is not clear why firm one would help firm two to so commit itself by accepting such contracts. Assume that firm two proposes to coordinate on a perfect equilibrium going to a steady state C above A; should firm one believe firm two and stop investing until the state reaches C? If it did, once at C, firm two could again propose an equilibrium with a vertical path, say to C' (see Figure 9), and so on, until the state reaches R_2 so that, <u>a posteriori</u>, firm one would have actually preferred to invest until R_2. Loosely speaking this situation may be compared to a sequential blackmail game in which one player

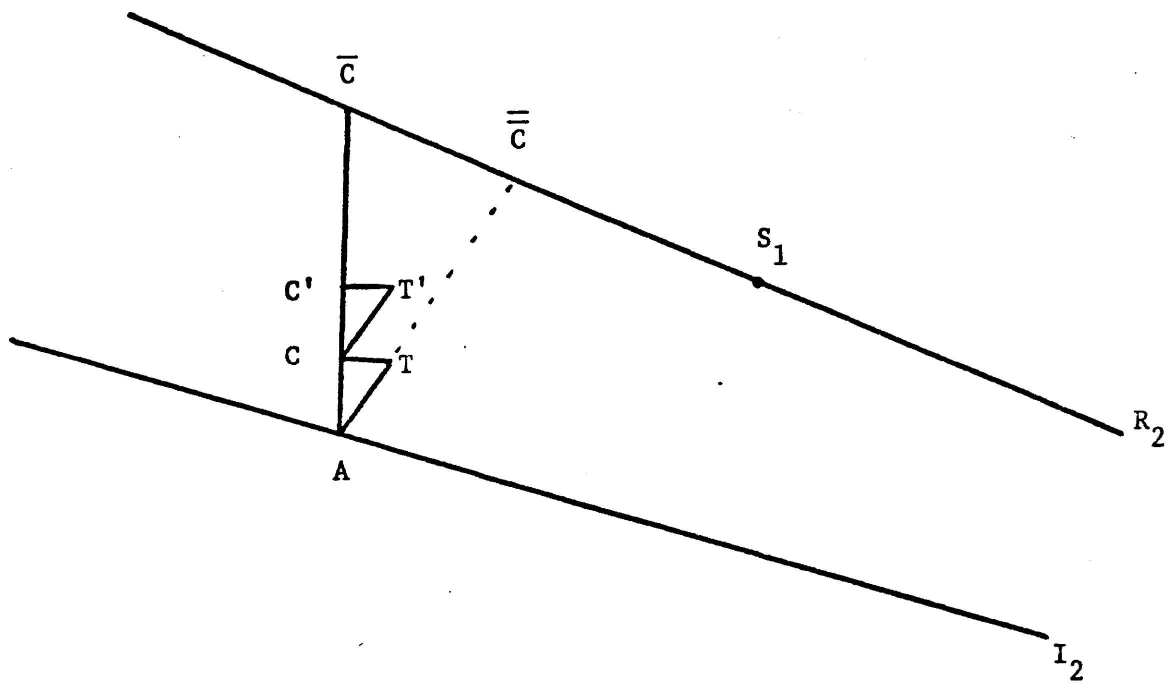


Figure 9:

At A on I_2, one perfect equilibrium has for terminal surface the horizontal segment T. The two firms, if they coordinate on this perfect equilibrium go to C. The difficulty is that at C, firm two has an incentive to repropose an equilibrium with another horizontal terminal surface, say T', leading to C', and so on, up to $\bar{C}$. Thus a posteriori firm one would have preferred to go to $\bar{\bar{C}}$.

(firm two) obtains a concession from the other (firm one) on the pretext that the blackmain will then end; the problem is that the blackmailer has no reason to stop, so that it is rational for the victim to refuse the first demand. What happens under both reaction curves? At a point on the upper envelope of I_1 and I_2, the Pareto frontier of the contract zone is composed of two segments , one horizontal and the other vertical. In this sense power is more equally distributed among the two firms, and the bargaining argument must be formulated both ways. Now, if the upper envelope of I_1 and I_2 is a terminal surface, it is clear that the state will never stop under it, since it is a dominant strategy for at least one firm to invest up to the terminal surface.

Alternatively one can analyze the game in which only "immediate-stopping" contracts are considered (immediate stopping contracts specify that both firms stop investing at the state where the contract is concluded). Note that in the region above the lower envelope of the reaction curves (i.e., the region where at most one firm would like to go on), this game is equivalent to the game where no contracts are considered, but firms can blow up their investment plant. In this game, firm two is the only player able to call a halt above R_1 and below R_2 (by proposing a contract to stop immediately, which will be accepted by firm one, or by blowing up its investment plant). This power is actually self-destructive in the sense that, if firm two does not propose an immediate contract or destroy its investment facilities, firm one should presume that firm two wants to invest more.

Our preferred perfect equilibrium has an interesting property in the special case where the IGP through the original point goes through the intersection I of I_1 and I_2. The steady state of the non-cooperative game is I (see the proof of proposition 2 for the description of the associated strategies). At this point, the isoprofit curves of firm one and firm two are tangent, since they are both tangent to the IGP. This means that each firm's profit is maximized given the other firm's profit level. In other words, the non-cooperative outcome is constrained efficient from the point of view of the two firms, given that monetary transfers are prohibited. Moreover, if the maximum investment speeds are equal, so that the slope of the IGP is equal to one, the non-cooperative outcome is nothing but the joint-profit maximization point. This in particular will be true in the completely symmetric case, where both firms are identical and enter the market at the same date (see Figure 10). Deviations from joint-profit maximization are then due to the eagerness of one of the firms to take advantage of the temporal asymmetry. Even in this case purely non-cooperative behavior leads to early stopping outcomes, i.e., to restricted competition.

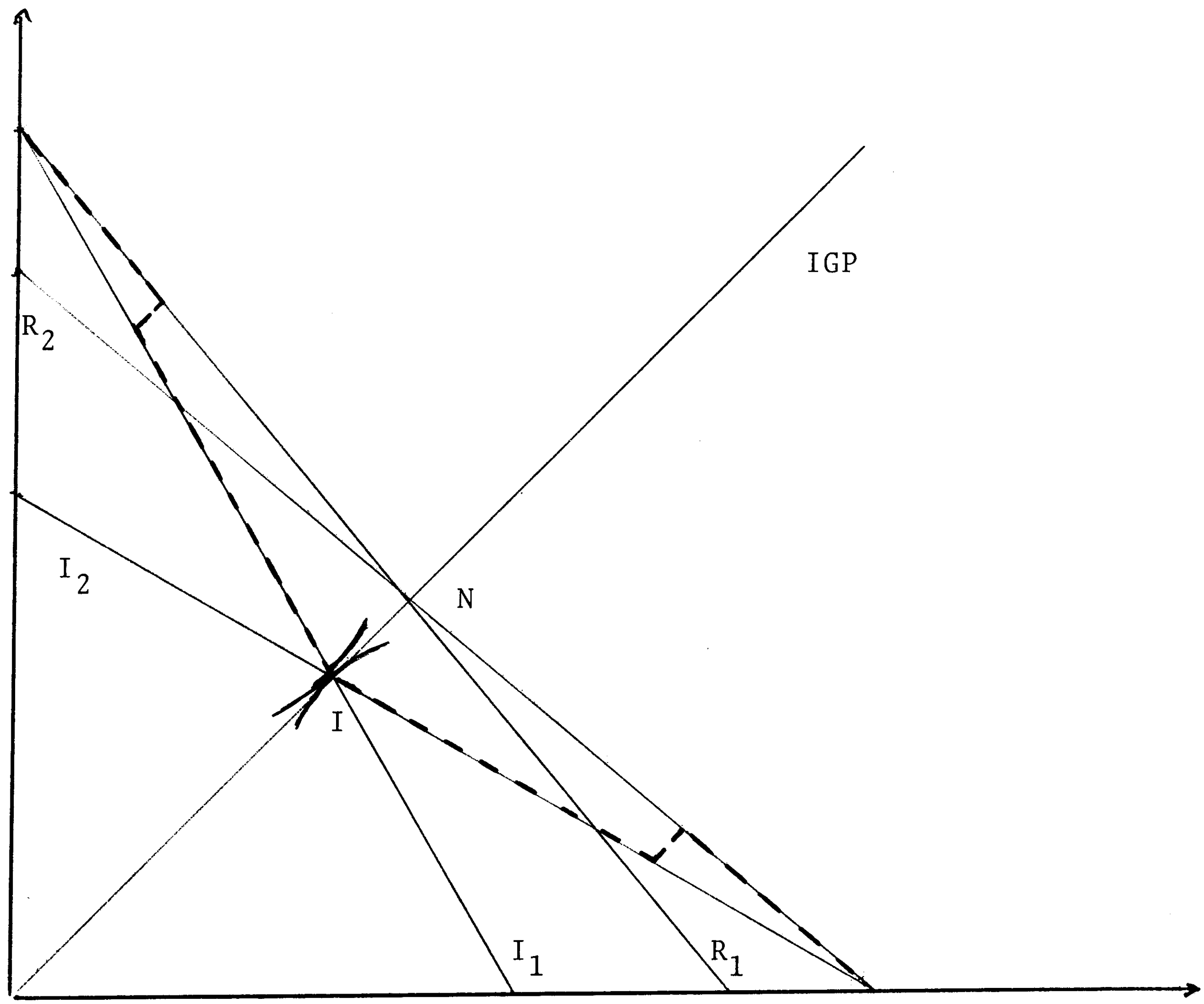


Figure 10*: Symmetric Case: The Non-Cooperative Outcome is the Joint-Profit Maximization Solution.

* This picture is drawn for the capacity constraint case with linear demand and constant marginal cost (up to the capacity constraint).

We conclude that the only reasonable equilibrium on which the firms could coordinate is the one which yields stopping at E*. We recognize, however, that this coordination process might break down. In this case, it is difficult to say what would happen. The observed moves might not correspond to any pair of perfect equilibrium strategies if the two firms expect different equilibria to prevail. In any case, a coordination equilibrium should not involve investing to R_2 (early stopping property), and most reasonably should stop on E*.

Chapter 7

The Discounting Case

Now that we have explored the model, and perfect equilibrium, in the no-discounting case, we proceed to the more realistic situation in which firms do discount future payoffs. Once time matters, firms must consider not only the eventual steady state, but also the approach to it.

Recall the first perfect equilibrium discussed, in which firm one acted as a Stackelberg leader, stopping below its preferred attainable point on R_2. This was Spence's solution. It is not generally correct with discounting, because such an investment path frequently will not be optimal for firm one, as shown below. The optimal control for firm one, when firm two invests as quickly as possible, will typically be a "two-switchpoint path" on which firm one invests, stops, and then invests again. Nor will a "modified Spence solution" with a two-switch point path going to R_2 usually be an equilibrium: if the second switchpoint occurs above the curve U_2, which is defined as before, then firm two will not want to invest to R_2, but will prefer to stop at a capital level just less than that at which firm one resumes investment.

On the other hand, the "stop at I_2" coordination equilibrium will be shown to carry over to the discounting case. Firm one's use of a two-switchpoint path will not induce firm two to want to stop earlier, as by definition below I_2 firm two prefers moving along the IGP to stopping.

We now discuss firm one's choice of an optimal path, assuming that firm two invests as fast as possible, in order to derive necessary (but not sufficient) conditions for a perfect equilibrium in which firm two invests as quickly as possible up to R_2. Spence proved that the optimal path for firm one to a point on R_1 is to invest as quickly as possible to the capital level associated with that point, and then stop. Intuitively, the firm is trying to minimize the horizontal distance betweeen the state and its reaction function, $K^1(t) - R_1\ (K^2(t))$, and a "most-rapid approach" path is at each instant no further away than any other feasible path. Spence claimed that such a path was optimal for reaching an arbitrary point. Were this true, then the solution he proposed would have been a perfect equilibrium with discounting. However, if the terminal point is above R_1, firm one will again try to stay as near its reaction curve as possible. Below R_1, this requires the most rapid approach path, and by symmetry we expect that above R_1 it will require a "least rapid approach" path.

To verify this intuition, we state firm one's control problem (i.e., for a given $\bar{K}^2\ (t),\ \forall t$)

$$\begin{cases} \displaystyle\max_{0 \leq I^1(t) \leq \bar{I}^1} \int_0^{\bar{T}} [\Pi^1(K^1(t),\ \bar{K}^2(t)) - I^i(t)]e^{-r^1 t} \\ \text{s.t.}\quad \dot{K}^1(t) = I^1(t),\ K^1(0) = \bar{K}^1 \end{cases}$$

$(\bar{K}^2(t) = \bar{I}^2 t,\ \bar{T}$ given by the "target" on $R_2)$

The Hamiltonian is:

$$H = e^{-r^1 t}[\Pi^1 - I^1] + \lambda I^1$$

First order conditions:

$$I^1 \text{ maximizes } H \rightarrow I^1 = \bar{I}^1 \text{ if } \lambda > e^{-r^1 t}$$

$$I^1 \varepsilon [0, \bar{I}^1] \text{ if } \lambda = e^{-r^1 t}$$

$$I^1 = 0 \text{ if } \lambda < e^{-r^1 t}$$

and $\dot{\lambda} = -\frac{\partial H}{\partial K^1} = -e^{-r^1 t}\Pi_1^1 \; (K^1, K^2)$

We see immediately that as the Hamiltonian is linear in investment, the optimal control is bang-bang. The path cannot involve an interval of singular control for if it did $\lambda = e^{-r^1 t}$, $\dot{\lambda} = -r^1 e^{r^1 t} \rightarrow \Pi_1^1 = r^1$ and the interval would be along R_1. As this is impossible in the absence of depreciation, we conclude there can be no interval of indeterminate investment.

Next we argue that there can be at most two switchpoints, and that the switch from investing to not investing must occur below R_1. Examine the switching equation, $\Sigma = \lambda - e^{-r^1 t}$. Below R_1, $\dot{\Sigma} < 0$, above R_1, $\dot{\Sigma} > 0$. Thus a switch from on to off could not occur above R_1, as the on-off switch requires that Σ be decreasing. Likewise, the off-on switch cannot occur below R_1. As switches must alternate in type, there can be at most one switchpoint in each region, and so no more than two in all. As a corrolary, any vertical segment of the path must intersect R_1.

We will call one switch-point curves with switchpoints below R_1 (i.e., "on-off" switchpoints) "J-curves" and two switch-point curves "S-curves". There can also be "upside-down J-curves" with switchpoints above R_1 (i.e., "off-on" switchpoints). The impossibility of a J-curve with switch-point above R_1 is easy to see in a diagram (Figure 11). In FIgure 11, the dotted S-curve clearly dominates the J-curve for firm one.

Now that we know how best for firm one to get to a given point, we can ask which point on R_2 firm one would choose, again assuming that firm two invests as quickly as possible to R_2. In particular, we might want to know when firm one's best reply involves choosing a J-curve. While we can not answer this question in general, we can partially characterize firm one's choice, and can find examples in which an S-curve is optimal.[7]

This completes our discussion of firm one's control problem. Let us now see the implication of S-curves for perfect equilibrium in the game.

Consider, for example, the perfect equilibrium path going to R_2. If the second switchpoint is above U_2, then firm two will prefer to stop below the switchpoint to forestall firm one's resuming investment; that is, "invest as fast as possible to R_2" will not be a Nash response for two. It was this observation which prompted our exploration of "early-stopping" equilibria.

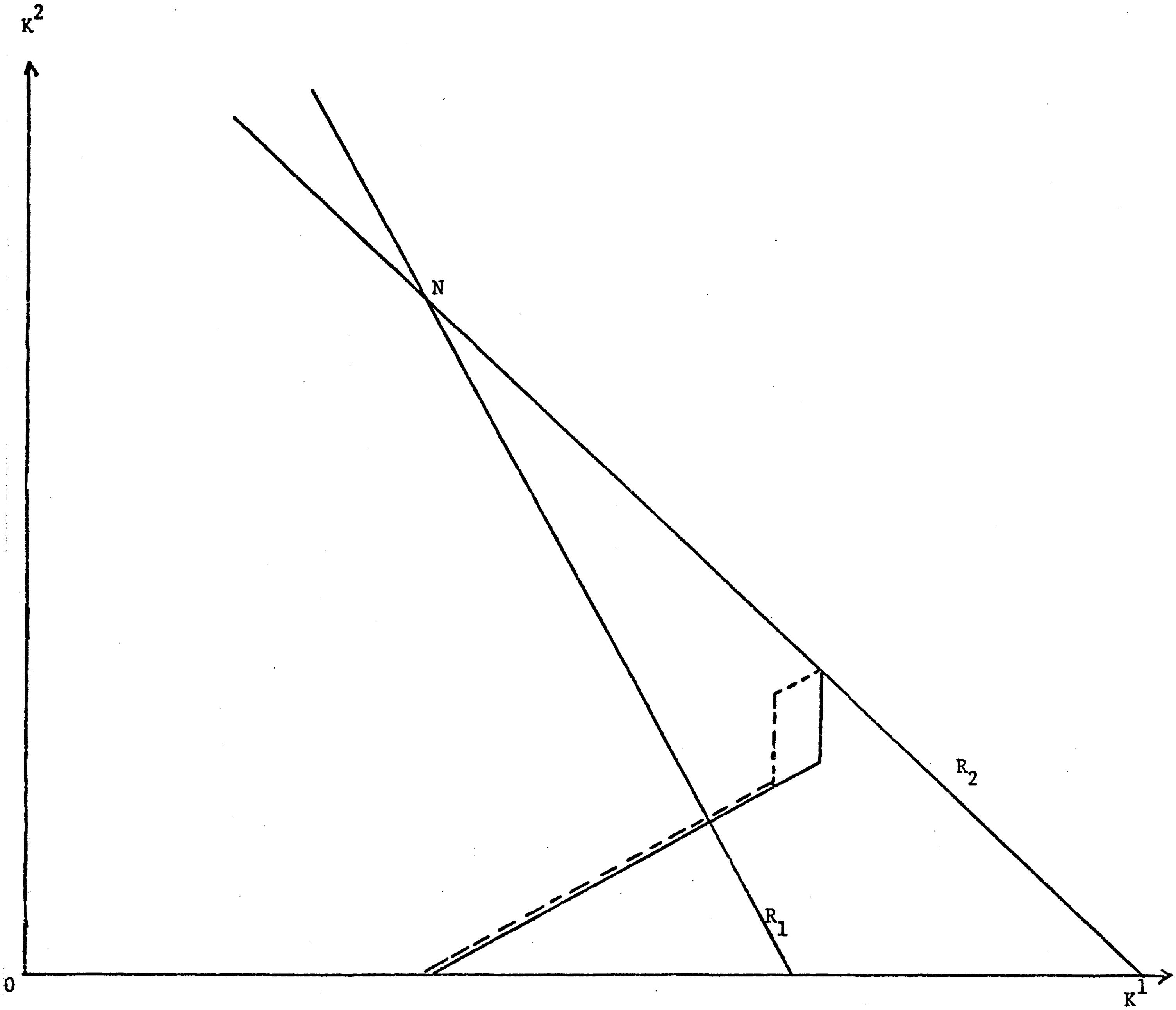


Figure 11

Let us now define the analogs of the terms developed for the no-discounting case. First the Stackelberg point S_1 on R_2 is such that just under that point, firm one is indifferent between letting firm two invest up to R_2, and limiting firm two's investment by investing (i.e., $(K^1(S_1), K^2(S_1))$ is given by the system of equations:

$$K^2 = R_2(K^1) \text{ and } -1 + \frac{\Pi^1_1(K^1,K^2) + |\Pi^1_2(K^1,K^2)| \frac{dR_2(K^1)}{dK^1}}{r^1} = 0).$$

The set of points above which firm one has a credible threat to invest in response to firm two's investing, C_1, is defined as the set of points where firm one would resume investment given that firm two invests up to R_2. In the no-discounting case, C_1 was part of the IGP going to S_1; whenever the state was above C_1, firm one preferred moving along the IGP to moving vertically, assuming that firm two invested up to R_2. In the discounting case, it is easily shown that S_1 belongs to C_1, that C_1 is to the northwest of the IGP to S_1, and that its slope is always less than the slope of the IGP through the point.

We define I_2 as before, to be the locus on which firm two is indifferent between having the state remain constant forever, and having it move a bit along the IGP, and stop forever (i.e., the equation of I_2 is: $\Pi^2_2(K^1,K^2)+\Pi^2_1(K^1,K^2)\frac{\bar{I}^1}{\bar{I}^2} = r^2$).

As in the no-discounting case, the following defines a perfect equilibrium. Let the terminal surface E* be the upper envelope of I_1 and I_2 between C_2 and C_1, C_1 and C_2 between the upper envelope of I_1 and I_2 and the reaction curves, and the reaction curves (R_2 to the southeast of C_1, and R_1 to the northwest of C_2), where I_1, C_2 are defined symmetrically to I_2, C_1. Above the reaction curves, neither firm invests; between E* and the reaction curves, both firms invest; and since E* is a terminal surface, no firm invests on E*. Finally, when firm one is the "leader", it solves a control problem knowing that firm two invests up to the terminal surface, and firm two invests as quickly as possible and symmetrically when firm two is the "leader". See Figure 12.

The arguments in favor of this equilibrium are the same as the ones in the no-discounting case. First, firm two would never agree to coordinate on a perfect equilibrium stopping under E*, since it can at worst propose a contract stopping all investment once on E*. The equilibria stopping between E* and the reaction curves do not seem reasonable for the same reason as in Chapter 6.

We conclude that our preferred "early stopping" equilibrium in the no-discounting case <u>carries over to the discounting case</u>. It can be checked that in the completely symmetric case where firms are identical and enter the market at the same date, our preferred perfect equilibrium leads to a path identical to the <u>dynamic joint-profit maximizing path</u>, i.e., the path that the firms would follow could they sign a binding contract and transfer money between them.

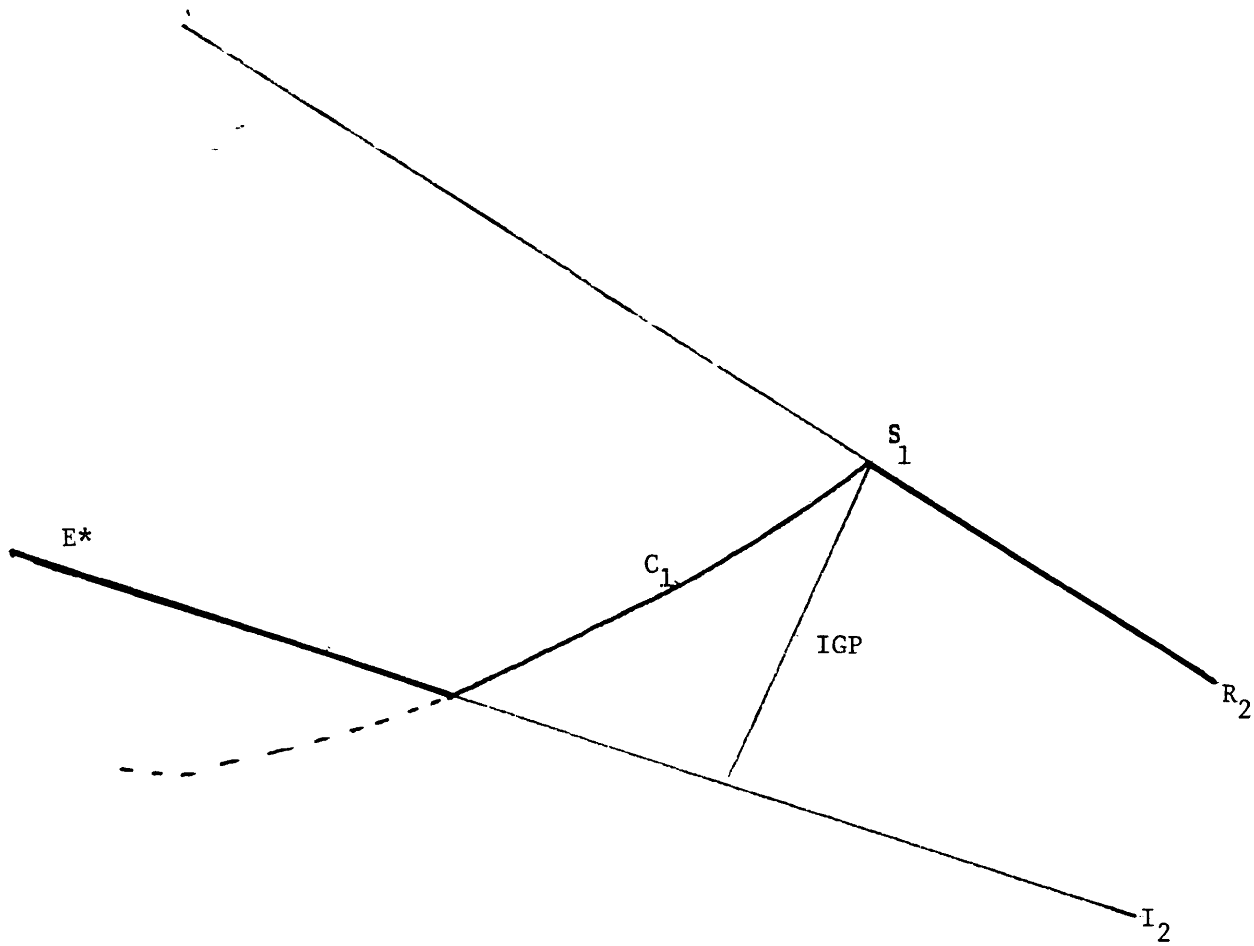


Figure 12: Construction of E* in the discounting case

Chapter 8

Entry

Having analyzed the "post entry" game that corresponds to the second period of two-period models, we turn to the question of entry deterrence. The answers to this question are a by-product of our previous analysis. Note that the reduction of the equilibrium set to one element (Chapters 6 and 7), is necessary in order to have a well defined trajectory from a given state point and therefore to discuss entry. The established firm starts investing at time 0, and firm two starts at time t_2 (i.e., t_2 is time zero in the post-entry game).

We consider two cases: "deterministic entry" in which the date of firm two's earlier possible investment, t_2, is known; and "stochastic entry" in which firm one has an exogenous subjective probability distribution over t_2. In both cases, firm two will enter if and only if the net present value of doing so is positive; this net present value $V_2(K^1(t_2))$ is nothing but the present discounted value at t_2 of firm two's profits along the perfect equilibrium path from $(K^1(t_2),0)$ to the stopping curve E^*. Similarly one can define $V_1(K^1(t_2))$ as the present discounted value at t_2 of firm one's profits if firm two enters - let K^{1d} be the entry deterring level: $V_2(K^{1d}) = 0$. It is easy to see that firm two will enter if and only if $K^1(t_2) < K^{1d}$. Define.

$$\bar{V}_1(K^1(t_2)) = \begin{cases} V_1(K^1(t_2)) & \text{if } K^1(t_2) < K^{1d} \\ V_m(K^1(t_2)) & \text{if } K^1(t_2)) \geq K^{1d} \end{cases}$$

where $V_m(K^1(t_2))$ is the present discounted value at t_2 of firm one's profits when it starts as $K^1(t_2)$ and acts as a monopolist. $\bar{V}_1(K^1(t_2))$ is discountinuous at K^{1d} (it jumps up).

1. Deterministic Entry: The date of potential entry t_2 is known. Firm one's maximization problem is:

$$\text{(A)} \quad \begin{cases} \text{Max} & \{\int_0^{t_2} [\Pi^1(K^1) - I^1]e^{-r^1 t}\, dt + e^{-r^1 t_2}\, \bar{V}_1(K^1(t_2))\} \\ \text{s.t.} & \dot{K}^1 = I^1 \end{cases}$$

where $\Pi^1(K^1)$ is the net (monopolistic) revenue of firm one. Depending on the solution of this problem, firm one will either deter entry by reaching (or overreaching) K^{1d}, or choose to let two enter. If it decides to overshoot K^{1d} at t_2, we shall say that entry is blockaded. This happens when the entry deterring level K^{1d} is less than the level of capital given by the intersection of firm one's reaction curve and the horizontal axis (and t_2 is "high enough") - this of course is not the interesting case - In general, the entry deterrence problem will be solved by computing the solution of (A).

2. Stochastic Entry: Let us assume that firm one has some subjective probability distribution about t_2 (for example firm two's entry is conditioned by the acquisition of know-how, whose date is uncertain). Let $F(t_2)$ be the cumulative distribution function of dates of entry. Firm one, assumed risk

neutral, solves the following stochastic control problem:

$$\text{(B)} \quad \begin{cases} \text{Max} \quad \left\{\int_0^\infty \left[\int_0^{t_2} (\Pi^1(K^1)-I^1)e^{-r^1 t}\, dt + e^{-r^1 t} \bar{V}(K^1(t_2))\right] dF(t_2)\right\} \\ \text{s.t.} \quad \dot{K}^1 = I^1 \end{cases}$$

The solution of (B) may involve different kinds of strategies; a simple case to analyze has the conditional probability of potential entry between t_2 and $(t_2 + dt_2)$ independent of time t_2 (exponential distribution case). Firm one will then use one of the following two strategies. It may either accumulate capital as quickly as possible up the deterring level K^{1d} (if firm two does not enter in between), and possibly overshoot; or accumulate capital as quickly as possible up to a given (non-deterring) level (if firm two does not enter in between) and then wait for firm two's entry. Again after firm two's entry the path followed is the perfect equilibrium path from the point of entry. Firm one's choice is then reduced to picking the level of capital ("target") which maximizes its expected payoff.

Chapter 9

Discussion

In closing, we would like to speculate on the extension of our results to models with depreciation. Intuitively, with high depreciation rates, the leader's capital stock is not at all locked in, and we expect it will thus lack an incentive to strategically overinvest. Were this intuition to carry over to low depreciation rates, then strategic investment would seem unlikely.

While we have been unable to find a perfect equilibrium for the game with depreciation, we have found a family of Nash equilibria with the following properties: for all positive depreciation rates the only steady state is at N; for low depreciation the path stays a long time above R_1; and for sufficiently high depreciation the path never goes above R_1. This equilibrium thus supports our intuition about the high-deprecation case, and yet exhibits strategic investment for low depreciation, perhaps justifying some optimism as to the reasonableness of such behavior.[8]

We conclude that the temporal asymmetries do give the lead firm an advantage, in that it may invest past its steady state Nash level to deter investment by others. The lead firm chooses its investment path considering the follower's eventual choice of steady-state; however, past its reaction curve the lead firm would always prefer both firms stopping to any alternative, so the follower does not always invest to R_2, but decides where to stop knowing that the leader will follow suit. The steady state will typically be an "early stopping" equilibrium below R_2. Of course, below its reaction curve, the follower would prefer to invest, ceteris paribus, but the leader's credible threat precludes further investment. This purely non-cooperative restriction of competition between the firms is reinforced by the observation that in the completely symmetric case, the outcome of the most reasonable perfect equilibrium is nothing but the dynamic joint-profit maximization path.

We argued that the study of entry deterrence, commitment, and credible threats should proceed by restricting the Nash equilibrium set to perfect equilibria. This restriction is is both simple and compelling. Our work suggests that further restrictions on the equilibrium set are desirable. While we used an ad-hoc argument for singling out a particular "early stopping" equilibrium, we hope that future research will develop a rigorous and general approach useful for the many infinite horizon, imperfect-information games that arise naturally in economic modelling.

Footnotes

1. Nick Papadopoulos has studied Spence's model using a differential game approach. While the Starr-Ho solution concept he used is essentially that of perfectness (with piecewise differentiable valuation function), he based his analysis on the assumptions that the state stopped on the reaction curves, and that the valuation function was continuous below them, which effectively imposed investing to the reaction curves as the only possible solution.

2. We will not, however, discuss the determination of the time period between the two entries. One could imagine that firms engage in search for new investment opportunities, with the probability of "discovery" depending on search expenditures, observed profits of other firms, and on potential profits for the searcher. The last case would present no difficulties for our analysis: One could "fold back" the post-discovery game, replacing it with the associated payoffs. Were firm one's conduct to affect the probability distribution of firm two's date of entry, then firm one would need to consider new issues such as "lying low" to avoid discovery, and our analysis would not be applicable.

3. In another paper, "Learning by Doing and Market Performance". we analyze a game in which current quantity decisions influence future costs and thus do have commitment value.

4. Note that for this game to be interesting, and for the reaction functions to be well-defined, there must be a flow cost of capital, or firms would choose arbitrarily large capital levels.

5. This claim can be made precise as follows: Consider an n-player finite game of perfect information in extensive form. Then associated with each path is a n-tuple of payoffs for the n players; as the game is finite, there are finitely many such paths; denote the number of paths by I. Then, holding the game tree fixed, we can identify games with specifications of payoffs for each path; that is, with elements of $[R^n]^I$. The claim is then that the complement of the closure of the set of elements of $[R^n]^I$ for which the associated game has multiple perfect equilibria is of full measure.

6. We believe that the multiplicity of equilibria in the investment game (see below) should be attributed to the infinite horizon, and not to the imperfect information, because multiplicity is also present in a discrete-time sequential move analog of the game.

7. We have shown:

   1) If the optimal J-curve goes to the left of S_1, it is dominated by an S-curve.

   2) On the other hand, if $\partial \Pi^1/\partial K^2$ is linear in K^2, then if the optimal J-curve ends to the right of S_1, it is globally optimal.

   3) The optimal S-curve always ends to the left of S_1.

   4) With either S or J curves, the terminal point chosen goes to S_1, when possible, as r^1 goes to zero.

   If we had incorporated a financial constraint on investment (such as requiring all investment to be financed by retained earnings) as in Spence (1979), then J-curves would be more often optimal. Such as integral constraint would have greatly complicated the analysis.

8. This path is the one assumed by Eaton and Lipsey (1980) who analyze entry-deterrence with depreciation in the case of unbounded investment speed and arbitrarily small difference in dates of entry.

References

Dixit, A. "A Model of Duopoly Suggesting a Theory of Entry Barriers". Bell Journal of Economics, Vol. 10, No. 1 (Spring 1979), pp. 20-32.

Dixit, A., "The Role of Investment in Entry Deterrence". Economic Journal 90 (March 1980), pp. 95-106.

Eaton, B.D. and Lipsey, R.G. "Capital, Commitment, and Entry Equilibrium". Queen's University Discussion Paper No. 265 (July 1980).

Fudenberg, D. and Tirole, J. "Learning by Doing and Market Performance" in preparation.

Harsanyi, J.C. "A General Solution for Finite Non-Cooperative Games, Based on Risk-Dominance". Advances in Game Theory, M. Dreshler et al., eds., Annals of Mathematics, Study 52, Princeton 1964, pp. 627-650.

Kydland, F.E. "Equilibrium Solutions in Dynamic Dominant Player Models" Journal of Economic Theory, Vol. 15, No. 2 (August 1977), pp. 307-324.

Maskin, E., and Newberry, D. "Rational Expectations and Market Power: The Paradox of the Disadvantageous Tariff on Oil" M.I.T. Working Paper No. 227, November 1978.

Papadopolous, N. "Growth in a New Market". Harvard University, March 1980.

Schelling, T.C. The Strategy of Conflict. Cambridge, Harvard University Press, 1980.

Schmalensee, R. "Economies of Scale and Barriers to Entry". Sloan School Working Paper, M.I.T., June 1980, No. 1130-80.

Selten. R. "Spieltheoretische Behandlung eines Oligopolmodells mit Nachfragerträheit" Zeitschrift für die gesamte Staatswissenschaft, Vol. 12, 1965.

Selten, R. "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games". International Journal of Game Theory, No. 1, 1975. pp. 25-55.

Spence, A.M. "Entry, Capacity, Investment, and Oligopolistic Pricing". Bell Journal of Economics, Vol. 8, No. 2 (Autumn 1977). pp. 534-44.

Spence, A.M. "Investment Strategy and Growth in a New Market", Bell Journal of Economics, Vol. 10, No. 1, 1979. pp. 1-19.

Starr, A.W., and Ho, Y.L. "Further Properties of Nonzero-Sum-Differential Games". *Journal of Optimization Theory and Applications*, Vol. 3, No. 4 (1969). pp. 207-219.

Appendix 1

Examples of Instantaneous Payoff Functions

In this appendix we compute the Π^i in two special cases and verify that our assumptions hold for capital levels at which strategic investment would occur.

The simplest example one can think of is that the instantaneous equilibrium be Cournot (i.e., Nash in quantities q^i) given instantaneous costs, which are taken to be $C^i(q^i,K^i) = c^iK^i$ for $q^i \leq K^i$, with capital functioning solely as a capacity constraint. Then, at each instant, firm i chooses q^i to maximize $q^i(p(q^i + q^j) - c^i)$ with $q^i \leq K^i$. When both capacity constraints are binding, our assumptions on the Π^i are nothing but the usual assumptions in the static Cournot model that: the profit function is concave ($\Pi^i_{ii} < 0$); that the reaction functions are downward sloping ($\Pi^i_{ij} < 0$) and have a unique intersection, that the Cournot equilibrium is stable under the usual dynamics and that there is a unique Stackelberg point on each reaction curve. If the state (K^1,K^2) is such that firm i has excess capacity, then the state can never move so as to make firm i's constraint binding. Therefore, nothing firm i does can influence firm j's investment, and firm i will not invest. Knowing this, firm j will simply invest until its capital level is optimal given firm i's. So we can solve directly the trivial game when at least one firm has excess capacity, and need impose assumptions only in the complementary case, which we discuss in the paper.

Appendix 2

Proof of Proposition Two

Pick any point on E under both reaction curves. By analogy with Chapter 5, the IGP going to this point will be called Region III. Region II (IV) will be the region to the right (left) of Region III and under T, and Region I is defined as before as the region above both reaction curves. Another area deliminated by E, C_1, C_2 and the upper envelope of R_1 and R_2, must still be named: Region V (see Figure 5).

In Region I, the strategies are the same as before (no one invests). In Regions II, III and IV, the strategies are defined as before except that the terminal surface is taken to be T rather than the upper envelope of R_1 and R_2. On T, both firms stop. Lastly in Region V, both firms invest.

We have to check that for each firm and for any initial state, the firm's strategy as described above is an optimal one, given the other firm's strategy. By our previous analysis we know that this is true in Regions I and V. Since E is located above U , no firm has an incentive to invest once on E, since it would lead to the steady state on the upper envelope of R_1 and R_2 on the IGP through the point on E. We thus have to check that the strategies are Nash in Regions II, III and IV. For that notice that firm two's (one's) (steady state) payoff increases when the state moves towards bigger $K^1(K^2)$ along T on the boarder of Region II (Region IV).

This, as shown in Chapter 5, implies that firm two's (one's) best response to firm one's (two's) strategy at any point of region II (Region IV) is to invest. Lastly firm one's (two's) optimization problem at any point in Region II (Region IV), given the other firm's strategy, induces a Nash strategy at any point of Region II (Region IV). Thus the strategies are in perfect equilibrium. Q.E.D.

Notice also that if for example $K^1(0) \geq K^1(\tilde{S}_1)$, the outcome of the investment game is a steady state on T just above $(K^1(0),0)$ (i.e., on C_1 or on the part of R_2 to the southeast of S_1)

Appendix 3

Example of the Computation of the Curve I_2

We consider the "capacity model" with $C^i(q^i,K^i) = c^iK^i$ for $q^i \le K^i$, $= \infty$ otherwise, linear demand, and no discounting. Then when the capacity constraint is binding $\pi^1(K^1,K^2) = K^1(1-K^1-K^2) - c^1K^1$, $R_2(K^1) = \frac{1-K^1-c^2}{2}$ and $S_1 = (\frac{1+c^2-2c^1}{2}, \frac{1+2c^1-3c^2}{4})$ just as in the static Cournot case.

We now compute I_2: it is given by

$$\frac{\partial \pi^2}{\partial K^2} = \frac{-\bar{I}^1}{\bar{I}^2}\frac{\partial \pi^2}{\partial K^1} = -i\,\frac{\partial \pi^2}{\partial K^1} \quad (\text{where } i = \frac{\bar{I}^1}{\bar{I}^2}) \text{ or}$$

$K^2 = \frac{1-K^1-c^2}{2+i}$. Firm one's preferred point on I_2 is

$K^1 = \frac{1-c^1(\frac{2+i}{1+i}) + c^2(\frac{1}{1+i})}{2}$, which can be greater than or less than $K^1(S_1)$.

Essay III

On the Possibility of Speculation Under Rational Expectations

Chapter 1

Introduction

Speculation is generally defined as a process for transferring price risks. Given this (admittedly vague) definition, there is considerable disagreement about the conditions which allow a speculative market to arise. The Working theory (see Hirshleifer [1975,1977], Feiger [1976]) makes differences in beliefs the key to speculative behavior: in particular the degree of traders' risk aversion affects only the size of their gamble. Associated with this theory and (as we shall see below) potentially at the root of its internal inconsistency is the idea that better informed traders are able to make money on the average. On the other hand, the Keynes-Hicks theory of speculation emphasizes not differences in beliefs, but differences in willingness to take risk or in initial positions as the foundation of a speculative market. The social function of speculation is thus to shift price risks from more to less risk averse traders or from traders with riskier positions to those with less risky positions. In other words, speculation in the Keynes-Hicks tradition is a substitute for insurance markets.

For markets with sequential trading (e.g., a stock market), one can give a more precise definition of speculation: Harrison and Kreps [1978], following Kaldor and Keynes, say that "investors exhibit speculative behavior if the right to resell [an] asset makes them willing to pay more for it than they would pay if obliged to hold it forever."

In this paper, we investigate the possibility of speculative behavior when traders have rational expectations.

The idea behind a rational expectations equilibrium (<u>REE</u>) is that each trader is able to make inferences from the market price about the profitability of his trade. Traders know the statistical relationship between the market price and the realized value of their trade (the "forecast function") and use the information conveyed by the price as well as their private information to choose their demands. As such the definition is still ambiguous, and this ambiguity has led to some misinterpretation of the concept of REE, particularly concerning speculation. If an agent's utility function is not strictly concave, his demand for an asset based on its price and his information (including the information conveyed by the price) may be multivalued. An auctioneer may then have to choose a trade from the agent's demand correspondence which depends on the other agents' trades to clear the market. But because the agent's trade then depends on those of others, it is not measurable with respect to the price and his information. That is, as Kreps [1977] recognized, the agent may then infer information not only from the price, but also paradoxically from the quantity he trades. This may seem to be merely a technical point, but it is important for our purpose, since many models of speculation assume risk neutrality. To see how easily the issue of measurability arises, let us describe an extremely simplified version of Feiger's [1978] futures market model. Consider a one-shot market for an asset the unit value of which is a random variable equal to +1 with probability 1/2, and -1 with probability 1/2. There are two classes of risk neutral traders. The "informed traders" know with certainty the realization of the random variable (i.e., they receive a signal perfectly correlated with it). The "uninformed traders" do not have any information other than the prior probability distribution; hence their expectation before looking at the market price is 0. We assume, as Feiger does, that traders cannot buy or sell more than their collateral constraint, taken to be their wealth. Moreover, the wealth of the uninformed traders exceeds that of the informed traders, which implies that the uninformed traders "make the price" (if the price were to differ from the expectation of the uninformed traders, no matter how formed, the market would stay unbalanced). In this model, there is a trivial (non-informative) self-fulfilling forecast function: Whatever the signal received by the informed traders, the market price is 0. Given this forecast function, uninformed traders are indifferent between buying and selling. Informed traders buy (sell) when their signal tells them that the value of the asset is +1 (-1). Hence, whatever the realization of the random variable, the uninformed traders, as a group, lose an amount of money equal to the wealth of the informed traders. Traders maximize their expected pay-off using their own information and the self-fulfilling forecast function, but they ignore the information conveyed by the magnitude of their trade.[2]

The concept of REE has been interpreted in two ways. In the first interpretation the forecast function - i.e., the statistical relationship between the market price and the profitability of the trades - is learned over time by the traders (see e.g., Bray [1980b]). The interpretation of course does not apply to markets which open only infrequently nor to markets which open repeatedly, but involve frequent structural shifts. Alternatively, the forecast function may be justified by an argument of the same nature as the foundation of some other equilibrium concepts in economic theory (Nash equilibrium, Bayesian equilibrium ...): vaguely speaking, there exists some common knowledge about the structure of the market, or more generally the game, and an equilibrium arises when each trader presumes that the other traders use their equilibrium strategies. For a market with heterogeneous information, assume that it is common knowledge (in Aumann [1976]'s sense) that a) trader i has utility function u^i; b) all the traders have the same prior[3] and are rational; c) the market clears. Then the traders can compute the forecast function associated with a REE. The assumption that the preferences are common knowledge is, of course, very strong, and one might want to substitute a') for a): a') trader i's utility function u^i belongs to set U^i (e.g., U^i is the set of concave utility functions). Then a REE has the property that no trader can refute the forecast function, in the following sense: Each trader i can construct a utility function in U^j, and a trade depending on the price, j's information for each trader j different from i, such that, for any set of signals the traders receive, all the traders maximize given their information and the market clears. The true utility functions and the REE trades satisfy these requirements, by definition.[4]

The object of this paper it to explore some implications of the concept of REE for the theory of speculation. We consider one-shot markets ("static speculation" - Chapter 3, as well as markets with sequential trading ("dynamic speculation" - Chapter 4.)

First we argue in Chapter 3a that, contrary to the Working-Hirshleifer - Feiger view, rational and risk averse traders never trade solely on the basis of differences in information. Risk neutral traders may trade, but do not expect any gain from their trade. The understanding of this no-betting result is facilitated by the second interpretation of a REE, where traders have some non-statistical knowledge about the market. Consider a purely speculative market (i.e., a market where the aggregate monetary gain in zero and insurance plays no role). Assume that it is common knowledge that traders are risk averse, rational, have the same prior and that the market clears. Then it is also common knowledge that a trader's expected monetary gain given his information must be positive in order for him to be willing to trade. The market clearing condition then imposes that no trader expects a monetary gain from his trade. This process can be illustrated by the following elementary example: At the beginning of a seminar the speaker states a proposition. Suppose that the validity of the proposition is in question; and that each member of the audience but the speaker either has no information about its validity or else has some counter-example in mind. In the first case, the member will not be willing to bet with the speaker, who, after all, having worked on the topic before the seminar, is endowed with superior information. In the second case, he will be willing to bet that the proposition is incorrect. The speaker can therefore deduce that only members of the audience having a counter-example in mind will be willing to bet with him.[5] Consequently, the speaker will not be willing to bet at all.

The idea behind this result is due to Stiglitz [1974] and a formal treatment can be found in Kreps [1977]. Milgrom and Stokey [1980] deduce a similar result using an interesting approach.

The motivation for recalling this theorem is that it provides insight on the rest of the paper, and moreover has important consequences for the theory of speculation. In particular, it definitely contradicts the Working theory for markets with rational traders having the same prior (but differential information). Indeed one might ask what is needed in order to observe speculative behavior; in Chapter 3a, we give four conditions giving rise to static speculation.

In Chapter 3b, we point out how the kind of reasoning used to show the impossibility of pure speculation can be extended to derive characterizations of "non-zero-sum games", e.g., of markets where trading is justified by insurance. Consider a one-shot speculative market ("futures market") with two classes of traders: risk-averse traders ("farmers") holding an initially risky position ("crop") and risk-averse traders endowed with information ("speculators"). Even in the case where all the relevant information is revealed by the futures price, trade occurs in agreement with the Keynes-Hicks theory, and the speculators receive a risk premium for insuring farmers. Introduce now into the market competitive, risk neutral traders ("insurance companies") who have the same information and rationality as the farmers. It is shown that, in a REE, speculators are deprived of their raison d'être (insurance) and do not trade any more. Indeed, the equilibrium exhibits a very peculiar feature: Some or all of the information retained by the speculators is revealed by the price, in spite of the fact that they do not trade! The usual justifications given for REE seem very weak in such a context and certainly demand more analysis.

While Chapter 3 is concerned with the characterization of one period speculative markets, Chapter 4 considers a sequential market in order to focus on the second definition of speculation - that investors exhibit a speculative behavior if the right to resell an asset makes them willing to pay more for it than they would pay if obliged to hold it forever. To this purpose we describe a stock market as a sequence of rational expectations equilibria. The dividends of a given firm $(d_0, d_1, \ldots, d_t \ldots)$ are assumed to follow an exogenously given stochastic process. Each trader, who is assumed to be risk neutral, will have in each period some information (signal) about the process. This information differs among traders. In a sequential speculative market, the concept of REE can be further refined. Often in markets with homogeneous information, traders are assumed to base their behavior on the comparison between the current price and (the probability distribution of) next period's price; the corresponding REE for a stock market with heterogeneous information will be named "myopic REE". We show that, for any given period, even if short sales are prohibited, a trader will not expect a gain from his trade, regardless of what information he may possess (of course, the price expectation is taken relative to the trader's own information and the information he can infer from the market). This does not mean that the price of the stock has to be equal to any market fundamental (i.e., the expected present discounted value of dividends). The right to resell the asset in general makes traders willing to pay more for it than they would pay if obliged to hold it forever, i.e., more than their market fundamental. Moreover, in an equilibrium of a stock market with infinite horizon, the market fundamental will in general differ for all traders. On the contrary, in a finite horizon stock market, it is shown that the price is equal to the market fundamental of any active trader (of any trader if short sales are allowed). Next define for each active trader a price bubble as the difference between the market price and his market fundamental. The martingale properties of those price bubbles are exhibited. Lastly, it is shown that, in the special case where everyone has the same information in every period, the uniqueness of the market fundamental leads to a price bubble identical for all traders, and this price bubble follows a (discounted) martingale.

One may nevertheless dislike the concept of myopic REE, especially in an economy with a finite number of traders. Indeed, a sequence of myopic REE does not necessarily lead to a well defined (i.e., converging) expected gain function for each trader. In Chapter 4a we exhibit an elementary example of myopic REE where any optimal strategy (i.e., maximizing a trader's expected pay-off over the whole time horizon) requires the trader to realize his profits in finite time (i.e., quit the market). This is inconsistent, as the set of traders is finite. We are thus led to define a fully dynamic REE as a sequence of self-fulfilling forecast functions such that there exists for each agent a sequence of (information contingent) stock holdings, called a "strategy", satisfying the following properties: (1) in each period t and for any information a trader i may have at time t, the corresponding strategy maximizes i's expected present discounted gain from t on (i's posterior being computed from the common prior and i's information-acquired individually and inferred from the market price); (2) the market clears in each period and for any information traders have in this period.

As one might expect, the definition of a fully dynamic REE puts very strong restrictions on the type of price and expectation functions which can arise in equilibrium.

In fact, Chapter 4b shows that in a fully dynamic REE, price bubbles disappear and every trader's market fundamental equals the price of the stock, regardless of whether short sales are allowed or not. This implies that a speculative behavior in the Kaldor-Keynes-Harrison-Kreps sense cannot be observed in a fully dynamic REE.

The paper is organized as follows: Chapter 2 provides notation and defines a REE. Chapter 3 deals with static speculation; Chapter 3a demonstrates the impossibility of pure speculation under rational expectations and risk aversion ("zero-sum game") and Chapter 3b considers the link between insurance and information conveyed by prices ("non-zero sum game"). Chapter 4 develops a model of REE in a sequential market with differential information, and answers the question of whether dynamic speculation in the Kaldor-Keynes-Harrison-Kreps sense is consistent with rational expectations. Lastly the conclusion summarizes our results and considers their implications for real world asset markets.

Chapter 2

Notation and Definition of a REE

The definition given in this section is intentionally vague and its content will be made precise in every particular context. Consider a market with I risk averse or risk neutral traders: $i = 1,\ldots,I$. Their net transactions $\{x^i\}$ are effected at price p. The market clears when: $\sum_i x^i = 0$. Let E be the set of pay-off relevant environments (for example E may be the set of potential spot prices next period). Each trader receives a private signal s^i belonging to a set S^i. The vector of all signals is: $s = (\ldots,s^i,\ldots)$ belonging to a set S (contained in $\times_i S^i$). Then $\Omega = E \times S$ is the set of states of nature, and we assume that all the traders have the <u>same prior</u> ν on Ω. Let T be a set contained in S; we shall denote by $\nu^i(s^i|T)$ the marginal probability of signal s^i conditional on $\{s \in T\}$. $\nu^i(s^i)$ will denote the prior probability of signal s^i.

Definition 1

A <u>REE</u> is a forecast function Φ which associates with each set of signals s a price $p = \Phi(s)$, <u>and</u> a set of trades $x^i(p,s^i,S(p))$ for each agent i, relative to information s^i and $s \in S(p) \equiv \Phi^{-1}(p)$, such that:

1. $x^i(p, s^i, S(p))$ maximizes i's expected utility conditional on i's private information s^i, and the information conveyed by the price S(p).

2. The market clears: $\sum_i x^i(p, s^i, S(p)) = 0$

Note that by writing $x^i(p, s^i, S(p))$, we impose the measurability requirement on trader i's demand, i.e., his demand depends only on his information and not on the other traders' signals or trades. Sometimes we shall use the shorthand $x^i(p, s^i)$ for $x^i(p, s^i, S(p))$.

The application of the concept of REE to the characterization of speculative markets is the central theme of the rest of the paper. Since the diversity of the models necessitated by the study of the different facets of speculation may be disagreeable to a reader interested in a particular aspect, the different parts are written so as to be self-contained, and thus can be read more or less independently.

Chapter 3

Static Speculation

a. The Impossibility of Pure Speculation: In this sub-section, we formalize the notion that pure speculation - or participation in a zero-sum game - is inconsistent with risk aversion and rational expectations. We shall say that a market is purely speculative if the total monetary gain is non positive and the participants' initial positions (corresponding to no trade on the market) are uncorrelated with the return of the asset.[6]

Trader i ($i = 1,\dots,I$) buys (sells), at price p, x^i claims which entitle (force) him to receive (give up) ($\tilde{p}\, x^i$) once the value of the random price $\tilde{p}$ is known. Every trader has a concave utility function, and his initial position is uncorrelated with $\tilde{p}$. i's ex post ("realized") gain is: $G^i = (\tilde{p} - p)x^i$. From the market clearing condition: $\sum_i G^i = 0$.

Trader i's individual information is a signal s^i in S^i. We shall assume that all signals have a positive probability:

$$\forall i,\ \forall s^i \in S^i: \nu^i(s^i) > 0$$

A REE of the purely speculative market is a forecast function $p = \Phi(s) \Longleftrightarrow s \in S(p) = \Phi^{-1}(p)$ such that:

$$\sum_i x^i\ (p, s^i, S(p)) = 0$$

Since trader i has a concave utility function, and has the option not to trade, he must expect a non-negative gain:

$$E(G^i | s^i, S(p)) \geq 0 \qquad (1)$$

This has to be true for any signal s^i belonging to the projection $S^i(p)$ of $S(p)$ on S^i. This implies:

$$E(G^i | S(p)) = \sum_{s^i \in S^i(p)} E(G^i | s^i, S(p)) \; \nu^i(s^i | S(p)) \geq 0 \qquad (2)$$

Because the market is purely speculative:

$$\sum_i G^i = 0 \rightarrow \sum_i E(G^i | S(p)) = 0 \qquad (3)$$

This implies in turn that:

$$\forall i: E(G^i | S(p)) = 0$$
$$\rightarrow \forall i: E(G^i | s^i, S(p)) = 0 \qquad (4)$$

In other words, no trader can expect a gain in a REE. We can now state:

<u>Proposition</u> 1: In a REE of a purely speculative market with risk averse or risk neutral traders, risk averse traders do not trade; risk neutral traders may trade, but they do not expect any gain from their trade.

Note also that a self-fulfilling equilibrium need not satisfy Proposition 1, as exemplified by the model based on Feiger [1978] presented in the introduction (equation (1) does not hold if the gain function G^i is not measurable with respect to p, s^i and S(p)). In this model, the uninformed traders expect a zero gain for any given trade, and the informed traders a gain equal to their wealth. The problem is that the trade of the uninformed traders depends on the signal received by the informed traders, and that the former are not allowed to make any inference from the quantity they trade.

Proposition 1 shows that one must relax at least one of the previous assumptions in order to observe static speculation:

(a) One may introduce risk loving traders.

(b) One may depart from the strict Bayesian assumption that priors are identical for everybody and that differences in beliefs are simply the result of differences in information.

(c) One other way of transforming the market into a "positive-sum game" from the point of view of the set of rational agents is to introduce a non rational agent: A related method consists in introducing traders whose (possibly stochastic) demand or supply is independent of the market price (see Grossman [1976], [1977] and

Grossman-Stiglitz [1976]), although one must be cautious and give a more complete description of the model before calling these traders irrational. The set of all rational players is then able to take advantage of this type of players, who, roughly speaking, face an unfair bet.

(d) The absence of correlation between the initial position of the traders and the market outcome (and the corresponding impossibility for anyone to use the market to hedge) is a central condition for the non-existence of a "pure betting market". If this condition fails to hold, the market can be seen as a means of supplying insurance to traders with risky positions. This view vindicates the Keynes-Hicks position and is the essence of Danthine [1978]'s model of a futures market (see Chapter 3b.)

The distinction between (c) and (d) is not as clear-cut as it might seem, if one considers the examples of REE which can be found in the literature. Consider, for example, Grossman [1976]'s one-period stock market; there is a fixed supply $\bar{x}$ of the stock. If, following Grossman, one assumes that traders have constant absolute risk aversion utility functions, the demands are independent of wealth and thus one does not have to specify who owns the initial stock in order to compute the equilibrium price. However, the stock market equilibrium may be interpreted in terms of (c) if the holders of the initial stock $\bar{x}$ sell the whole stock to the set of rational buyers whatever the price or in terms of (d) if the rational traders also own the initial stock and thus try to hedge (or speculate) on the market.

Let us now examine where the previous argument breaks down when one of the assumptions is relaxed. First if a trader either is risk loving or has an initially risky position on the market (cases (a) and (d)), he may in equilibrium expect a negative gain: Thus (1) does not hold. (1) also fails to hold when one introduces irrational agents or fixed supplies or demands into the market (case (c)); to illustrate this simply, assume that there is a fixed supply $\bar{x}$ of a risky asset, so that the market clearing condition is: $\sum_i x^i = \bar{x}$. Assume further that all traders have the same information and the same constant absolute risk aversion utility function, and that the distribution of the future price of the asset is normal. It is well known[7] that the demand of the rational traders is proportional to $(E(\tilde{p}) - p)$ where $E(\tilde{p})$ denotes the expectation of the price relative to the common information. Thus in equilibrium: $k(E(\tilde{p}) - p) = \bar{x}$ $(k > 0)$. The aggregate expected gain of the rational traders is then: $\frac{\bar{x}^2}{k} > 0$, whereas the traders with the fixed supplies expect an aggregate loss $(\frac{\bar{x}^2}{-k})$ relative to their not selling the asset. Finally, if traders have different priors[8] (case (b)), the sum of the expected gains may well be strictly positive: Since the posteriors have to be computed from different priors, (3) does not hold.

b. Insurance and Information: Links between information and insurance have been much studied in information theory. In particular, a recurrent theme is that information, by revealing to some traders their advantageous position, will in general restrict trade, and thus destroy the possibility of insurance. In this section we consider the effect of the introduction of competitive, risk-neutral traders, called "insurance companies", on a market where transfers of information take place. We consider a futures market analogous to that of Danthine [1978][9] where risk averse farmers, who hold an initially risky position, sell in order to obtain insurance, and risk averse speculators, who are endowed with some information about the future spot price, buy. This is *not* a "zero-sum game", since farmers are willing to pay a risk premium in order to get insurance; thus Proposition 1 of Chapter 3a does not hold. The futures price may reveal part of or all the relevant information held by the speculators, but, even in the fully revealing case, trade takes place in agreement with the Keynes-Hicks theory. Imagine now that competitive, risk neutral traders enter the futures market. They are supposed to have the same information as the farmers - i.e., the one conveyed by the futures price. Clearly they are more efficient in insuring farmers than the speculators; they deprive them of their raison d'être. The speculators' only motive to trade would be their superiority in information, or, in other words, the insurance companies have reduced the game between farmers and speculators to a pure speculation game, which, we know from Chapter 3a is impossible. Thus in a REE, the speculators do not trade. Let us now give a sketch of the model (for more details, see Danthine).

Risk averse <u>farmers</u> ($i = 1,\dots,I$) plant a crop in period 1 and harvest it in period 2. The cost of planting is c^i, and the quantity harvested $q^i(c^i)$. The price of the harvest in period 2 is some random variable $\tilde{p}$; to simplify we assume that $\tilde{p}$ is independent of the harvest, i.e., the consumers have an infinitely elastic demand curve. Every trader has the same prior on $\tilde{p}$; to make things concrete, let $\tilde{p} = \bar{p} + \eta$, where $\eta \sim N(o,\sigma_\eta^2)$. Farmer i, who has an initially risky position sells f^i units of futures contracts at price p at time 1.

The other participants in the futures market are competitive, risk averse <u>"speculators"</u> ($j = 1,\dots,J$). Each speculator j has some information (signal) $p^j \equiv \tilde{p} + s^j$ about the future spot price $\tilde{p}$, where $s^j \sim N(o,\sigma_s^2)$ (again the distribution is not essential for our purpose). Speculator j buys b^j in the futures market. As is usual in a REE, the market price p conveys some or all the relevant information about the set of signals $(s^1, \ldots, s^j)$.

Farmers and speculators have Von Neumann-Morgenstern utility functions, so that their optimization problems are:

Farmer i:

$$\underset{\{c^i, f^i\}}{\text{Max}} \quad E\{u^i(\tilde{p}\ q^i(c^i) + (p - \tilde{p})\ f^i - c^i) \mid p\}$$

Speculator j:

$$\underset{\{b^j\}}{\text{Max}} \quad E\{u^j((\tilde{p} - p)b^j) \mid s^j, p)\}$$

The futures market is in equilibrium when:

$$\sum_i f^i(p) = \sum_j b^j\ (s^j, p)$$

The derivations are given in Danthine; all we have to observe is that even in the case where speculators have no informational superiority (the price reveals the sufficient statistic $\frac{\sum_j s^j}{|J|}$), trade takes place and the speculators make a profit by providing insurance to the farmers.

Introduce now <u>competitive risk neutral traders</u> (k = 1,...,K), called insurance companies. Insurance company k buys b^k in the futures market on the basis of the information revealed by the market price p. Equilibrium in the futures market is now achieved when:

$$\sum_i f^i(p) = \sum_j b^j\ (s^j, p) + \sum_k b^k(p)$$

We now show that the presence of insurance companies implies that in a REE speculators do not trade.

First observe that the risk neutrality of the insurance companies implies that their price expectation: $E(\tilde{p}|p)$ (and thus the one of the farmers) is equal to the futures price:

$$p = E(\tilde{p}|p) \tag{5}$$

In passing, this implies that, whatever their optimal crop, farmers buy complete insurance: $\forall i$: $f^i = q^i(c^i)$. To conform to the notation of the introduction, let:

$$x^i \equiv -f^i(\forall i),\ x^j \equiv b^j(\forall j), x^k \equiv b^k(\forall k) \text{ and}$$
$$G^l \equiv (\tilde{p} - p)\, x^l$$

for $l = i,j,k$.

Market clearing requires that:

$$\sum_i G^i + \sum_j G^j + \sum_k G^k = 0 \tag{6}$$

Assume now that for some signals $s = (\ldots,s^j,\ldots)$ and associated price $p = \Phi(s)$ (where Φ is the forecast function), there exists a speculator j_o that: $x^{j_o}(s^{j_o},p) \neq 0$. This implies that:

$$E(G^{j_o}|s^{j_o},p) > 0$$

Since speculators are always free not to trade, we know that for all signals and for each speculator j:

$$E(G^j | s^j, p) \geq 0$$

By integration with respect to s^j (consistent with p):

$$E(G^j | p) \geq 0 \quad (7)$$

and for j_o:

$$E(G^{j_o} | p) > 0 \quad (8)$$[10]

(6), (7) and (8) imply that:

$$\sum_i E(G^i | p) + \sum_k E(G^k | p) < 0$$

which contradicts (5).

We can now state:

Proposition 2: Consider a one-shot speculative market with three classes of traders - risk averse traders with initially risky positions ("farmers"), - risk averse traders with initially riskless positions ("speculators") and - risk neutral traders ("insurance companies"). In a REE, speculators will not trade, regardless of their and others' information.[11]

That the speculators do not trade any longer in equilibrium should not surprise us given that the justification of their trade before the introduction of insurance companies was insurance and not their superiority in information. Note that, apparently paradoxically, if a REE exists,[12] some or all the information of the speculators is revealed by the price although they do not trade. This is straining the credibility of a REE to the limit and it is difficult to see how such an equilibrium could arise. The lack of theory about how a REE is reached is of course a general drawback of the REE analysis, but seems particularly problematic in our context.

Also we have not found a definite answer to the welfare implications of Proposition 2. Note that Proposition 2 does not say that the speculators' information (which has a social value) is destroyed, i.e., is not transmitted by the market price and thus remains unused. It is clear that in the presence of insurance companies, the speculators are worse off (whereas the insurance companies neither gain nor lose). The farmers are better insured since they do not pay a premium relative to their expectation. If the transmission of information by the price does not "deteriorate" with the introduction of insurance companies (i.e., $\forall\, s \in S$, $S^{ins}(s) \subseteq S(s)$ where $S(s)$, $S^{ins}(s)$ are subsets of S representing the information conveyed by the price without and with insurance companies), then the farmers are unambiguously better off.

Chapter 4

Dynamic Speculation: A Model of a Stock Market with Heterogeneous Information

This section is particularly concerned with the Kaldor-Keynes-Harrison-Kreps definition of speculation, according to which investors exhibit speculative behavior if the right to resell an asset makes them willing to pay more for it than they would pay if obliged to hold it forever. To this intent we describe the market for a given stock as a sequence of REE.

The stock may be traded at dates $t = 0, 1, 2, \ldots$. The (non-negative) dividend d_t is declared immediately prior to time t trading, and paid to traders who hold the stock at (t-1). As in Harrison and Kreps [1978], we assume that the sequence of dividends $\{d_0, d_1, \ldots, d_t, \ldots\}$ is an exogenously given stochastic process (for example driven by the demand in the market of the firm's output,...). At time t, the stock is traded at price p_t.

There is a finite set of traders $i = 1,\ldots,I$. Trader i is assumed to be risk neutral and to discount the future with the discount factor γ. His holding of the stock at time t is x_t^i [13] and given an aggregate stock $\bar{x}$, the market clearing condition is $\sum_i x_t^i = \bar{x}$ $(= \sum_i x_{t-1}^i)$. If short sales are prohibited, we impose that $x_t^i \geq 0$. In this case we shall say that trader i is active at time t if: either

$x_t^i \neq x_{t-1}^i$ or $0 < x_t^i = x_{t-1}^i < \bar{x}$. If short sales are allowed, the convention will be that every trader is active at every period. The motivation for this definition will become clear later. The market is active at time t if some traders are active. Let us now describe the information available to traders at each date, and then define selling strategies and sequences of REE of the stock market.

Information

Let E be the set of pay-off relevant environments. E is here taken to be the set of potential processes governing the sequence of dividends. At each period t, trader i has some private information about the underlying stochastic process; this may include past dividends, past prices, market studies, tips, etc. We represent trader i's information at time t as an element (event) s_t^i of a partition F_t^i of a set S^i. It is natural to assume that the partition F_t^i becomes finer and finer over time: $F_t^i \subseteq F_{t+1}^i$. The vector of all signals at time t is $s_t = (\ldots, s_t^i, \ldots)$. s_t is subset of a set S contained in $\underset{i}{X}\, S^i$. Let $\Omega \equiv E \times S$ denote the set of states of nature. We shall assume that all traders have the same prior ν on Ω. Let ν^i denote the marginal probability distribution on S^i; for simplicity we assume that all signals have positive measure: $\forall i$, $\forall t$, $\forall s_t^i \in F_t^i$: $\nu^i (s_t^i) > 0$ (this assumption can easily be relaxed).

At each time t, trader i can derive some information in addition to his private signal s_t^i simply by observing the current price p_t. Anticipating ourselves a bit, a REE at time t is characterized by a forecast function Φ_t, which associates with any set of signals s_t a price $p_t = \Phi_t(s_t)$. Conversely, the observation of price p_t indicates that s_t belongs to $S_t(p_t) \equiv \Phi_t^{-1}(p_t)$. ($S_t(p_t)$ is an element of $\{\underset{i}{X} F_t^i\}$). For notational simplicity, we shall often use the shorthand S_t for $S_t(p_t)$. To summarize, at time t trader i has information (s_t^i, S_t) based on his private signal (s_t^i) and the information conveyed by the price (S_t).

Consider a trader i having at time t information (s_t^i, S_t). This information can be regarded as a probability distribution on S (which takes zero values except on $s_t^i \times S_t^{-i}$, where S_t^{-i} denotes the projection of S_t on $\{\underset{j\neq i}{X} S^j\}$) and thus on Ω. This in turn induces a conditional probability distribution on $\{\underset{j}{X} F_{t+\tau}^j\}$, $\forall\tau \geq 1$: Trader i assigns a probability to any set of signals $s_{t+\tau} = (\ldots, s_{t+\tau}^j, \ldots)$ in $\{\underset{j}{X} F_{t+\tau}^j\}$. With a set of signals $s_{t+\tau}$ received by the traders at time $(t+\tau)$, there will be associated a price $p_{t+\tau} = \Phi_{t+\tau}(s_{t+\tau})$, so that trader i will have information $(s_{t+\tau}^i, S_{t+\tau} = \Phi_{t+\tau}^{-1}(p_{t+\tau}))$. To summarize, with each information (s_t^i, S_t) at time t, trader i associates a probability of having at time $(t+\tau)$ information $(s_{t+\tau}^i, S_{t+\tau})$ and facing price $p_{t+\tau}$.

In the following we shall define two types of REE: myopic REE (4a) and fully dynamic REE (4b). For each definition we shall characterize the equilibrium. (The propositions which are proven only in the case where short sales are prohibited also hold in the simpler case where they are allowed, following the convention that all traders are active).

a. Myopic REE: It is often assumed in the literature on sequential trading that traders choose their trades on the basis of short run considerations;[14] more precisely, in each period they compare their current trading opportunities with the expected trading opportunities in the following period. The application of this concept to a stock market with heterogeneous information leads to the following definition:

Definition

A myopic REE is a sequence of self-fulfilling forecast functions $s_t = (\ldots, s_t^i, \ldots) \to p_t = \Phi_t(s_t) \Longleftrightarrow s_t \in S_t(p_t) \equiv \Phi_t^{-1}(p_t)$, such that there exists a sequence of associated stock holdings $\{x_t^i(s_t^i, p_t)\}$ for each trader, satisfying:

1) Market Clearing: $\forall t,\ \forall s_t: \sum_i x_t^i(s_t^i, p_t) = \bar{x}$

2) Short Run Optimizing Behavior:

i) If short sales are allowed:

$$\forall t,\ \forall s_t,\ \forall i: p_t = E[\gamma d_{t+1} + \gamma p_{t+1} \mid s_t^i, S_t]$$

ii) If short sales are prohibited:

If $p_t = E[\gamma d_{t+1} + \gamma p_{t+1} \mid s_t^i, S_t]$, then $x_t^i(s_t^i, p_t) \in [0, \bar{x}]$

- - > - , - - = 0

- - < - , - - = $\bar{x}$

The interpretation of 2) is that each trader maximizes his expected short-run gain.

We now prove that even if short sales are prohibited, the price p_t must be equal to the expectation of the sum of the discounted dividend and the discounted next period price for any trader active at time t - i.e., no trader expects a short-run gain from his trade.

<u>Proposition 3</u>: Even is short sales are prohibited, for any trader i active at time t:

$$p_t = E(\gamma d_{t+1} + \gamma p_{t+1} | s_t^i, S_t)$$

<u>Proof</u>: Let $g_t^i \equiv -p_t \Delta x_t^i$ and ${}_t g_{t+1}^i \equiv [p_{t+1} + d_{t+1}] \Delta x_t^i$ denote the changes in i's cash flows at t and (t+1) resulting from his trade $\Delta x_t^i \equiv x_t^i - x_{t-1}^i$ at time t.

From the market clearing condition at time t: $\forall s_{t+1}$:

$$\sum_i g_t^i = 0 \text{ and } \sum_i {}_t g_{t+1}^i = 0$$

This implies:

$$\sum_i [g_t^i + \gamma \; {}_t g_{t+1}^i] = 0$$

Taking the expectation relative to the set of signals consistant with p_t:

$$\sum_i E[g_t^i + \gamma \; {}_t g_{t+1}^i | S_t] = 0$$

From the maximizing behavior of agent i:

$$\forall s_t \in S_t: E[g_t^i + \gamma_t g_{t+1}^i | s_t^i, S_t] \geq 0$$

Thus if for some information s_t and some trader i_o, we had:

$$E(g_t^{i_o} + \gamma \; {}_t g_{t+1}^{i_o} | s_t^{i_o}, S_t) \geq 0$$

we would conclude that:

$$\sum_i E(g_t^i + \gamma_t g_{t+1}^i | S_t) > 0$$

a contradiction.

Q.E.D.

It is often thought that the price of an asset in a speculative market may reflect speculative attributes as well as the asset's basic value, that is the price of an asset is the sum of its fundamental value and its speculative value ("price bubble"). Sargent and Wallace [1973] and Flood and Garber [1980] show in a monetary model with homogeneous information that price bubbles are not inconsistent with rational expectations; they are not even inconsistent with a positive probability in any period that the bubble "bursts" and the market "crashes down" to the market fundamental (see Blanchard [1979]; now the price has to grow faster during the duration of the bubble than in the previous case in order for the asset holders to be compensated for the probability of a crash). We study those price bubbles in our stock market with heterogeneous information.

Given information (s_t^i, S_t), one can define a <u>market fundamental</u> as the expectation of the present discounted value of future dividends:

$$F(s_t^i, S_t) \equiv E(\sum_{\tau=1}^{\infty} \gamma^{\tau} d_{t+\tau} | s_t^i, S_t)$$

For the price p_t consistent with S_t, the <u>price bubble</u> as seen by an individual with information (s_t^i, S_t) is defined by:

$$B(s_t^i, p_t) \equiv p_t - F(s_t^i, S_t)$$

Note that the price bubble depends on the information, and thus generally differs among individuals. Note also that the Kaldor-Keynes definition of speculative behavior of trader i amounts to:

$$B(s_t^i, p_t) > 0$$

We now formalize the idea that in a finite horizon asset market, price bubbles can not exist, and thus that no speculative behavior (in the previous sense) should be observed.

<u>Proposition 4</u>: In a stock market with finite horizon $\bar{T}$, whether short sales are allowed or not, the price bubbles are all equal to zero for the traders active in the market. Thus a market fundamental can be uniquely defined as the common market fundamental of all active traders, and is equal to the price:

$$\forall t, \forall i \text{ active at } t: \quad p_t = E(\sum_{u=t+1}^{\bar{T}} \gamma^{u-t} d_u | s_t^i, S_t)$$

Proof: The price of the stock at $\bar{T}$ is 0. Consider a trader i who is active at $(\bar{T}-1)$. Proposition 3 implies that:

$$p_{\bar{T}-1} = E(\gamma d_{\bar{T}-1} | s^i_{\bar{T}-1}, S_{\bar{T}-1})$$

This means that at $(\bar{T}-1)$ an active trader is indifferent between selling and holding the stock until the end period $\bar{T}$. Consider now an active trader i at time $(\bar{T}-2)$. According to his information at $(\bar{T}-1)$, he will hold the stock until $\bar{T}$ or trade at $(\bar{T}-1)$. But we saw that if he is active, he is indifferent between trading and holding. Thus: $\forall i$, $\forall s_{T-2}$ such that i is active at $(\bar{T}-2)$:

$$p_{\bar{T}-2} = E(\gamma d_{\bar{T}-1} + \gamma^2 d_{\bar{T}} | s^i_{\bar{T}-2}, S_{\bar{T}-2})$$

Proposition 4 is then proved by induction. Q.E.D.

Thus in a finite horizon stock market, backward induction from the final "crash" leads to the absence of price bubbles. The picture changes dramatically in the infinite horizon case.

Let us now assume that the horizon is infinite , and investigate the martingale properties of the price bubbles.

Proposition 5: (a) If short sales are allowed, then price bubbles are (discounted) martingales: $\forall i$, $\forall (s_t^i, S_t)$, $\forall T \geq 1$:

$$B(s_t^i, p_t) = \gamma^T E(B(s_{t+T}^i, p_{t+T}) | s_t^i, S_t)$$

(b) If short sales are prohibited, the price bubble of trader i endowed with information (s_t^i, S_t) satisfies the preceeding martingale property between t and (t+T) if, conditionally on his information at t, trader i is active in each period t, t+1,...,t+T-1.

Proof: The proof is a simple application of the law of iterated projections.

a) By definition of an equilibrium: $\forall t$, $\forall s_t$, $\forall i$:

$$p_t = E(\gamma d_{t+1} + \gamma p_{t+1} | s_t^i, S_t)$$

$$= E[\gamma d_{t+1} + \gamma [E[\gamma d_{t+2} + \gamma p_{t+2} | s_{t+1}^i, S_{t+1}]] | s_t^i, S_t]$$

$$= E[\gamma d_{t+1} + \gamma^2 d_{t+2} + \gamma^2 p_{t+2} | s_t^i, S_t]$$

By induction:

$$p_t = E(\sum_{\tau=1}^{T} \gamma^\tau d_{t+T} + \gamma^T p_{t+T} | s_t^i, S_t) \quad (9)$$

$$= E(\sum_{\tau=1}^{T} \gamma^\tau d_{t+\tau} | s_t^i, S_t) + \gamma^T E(E(\sum_{\tau=1}^{\infty} \gamma^\tau d_{t+T+\tau} | s_{t+T}^i, S_{t+T}) | s_t^i, S_t)$$

$$+ \gamma^T E(B(s_{t+T}^i, p_{t+T}) | s_t^i, S_t) \quad (10)$$

Using the law of iterated projections:

$$p_t = E(\sum_{\tau=1}^{\infty} \gamma^\tau d_{t+\tau} | s_t^i, S_t) + \gamma^T E(B(s_{t+T}^i, p_{t+T}) | s_t^i, S_t)$$

or:

$$B(s_t^i, p_t) = \gamma^T E(B(s_{t+T}^i, p_{t+T}) | s_t^i, S_t) \quad (11)$$

b) It is clear that the proof still holds without short sales, if trader i is active in every intermediate period for any state of information which can occur, given that i's information at t is (s_t^i, S_t). Q.E.D.

Special Case: Homogeneous Information (myopic REE version of Radner [1972]'s equilibrium of plans, prices and price expectations). Assume that all traders have at each period the same information, i.e., receive the same signal $s_t \varepsilon F_t$. The price p_t conveys no extra information, and traders base their expectation on s_t. The following proposition is trivial:

Proposition 6: In a stock market with homogeneous information, whether short sales are allowed or not, the price bubble is the same for every trader, and has the martingale property.

Note that if a heterogeneous information REE is fully revealing, i.e., $S_t(p_t)$ is a "sufficient statistic" for the set of signals in each period, we are in a situation analogous to the special case.

. As explained in the introduction, a myopic REE exhibits some rather unattractive features. This can be illustrated by a simple stock market with no uncertainty. Assume there is one unit of a stock, whose price at time t is p_t. A constant dividend $d_t = 1$ $(t \geq 1)$ is distributed just before trading. If traders have a discount factor $\frac{1}{2}$, the market fundamental is $(\frac{1}{2} + \frac{1}{4} + \ldots) \times 1 = 1$.

A myopic REE is simply a price function p_t such that:

$$p_t = \frac{1}{2}(1 + p_{t+1})$$

The general solution is: $p_t = 1 + \alpha 2^t$, where $\alpha 2^t$ represents a price bubble. Assume there are two individuals (or two types) A and B, and consider the following sequence of traders (trader A is the initial owner of the stock):

At time 0, trader A sells the stock to trader B at price 2

......1...... B .... A 3

......2...... A .... B 5

......3...... B .... A 9

etc.

This is a myopic REE. The first thing to observe is that, if we try to compute the discounted gains of the traders, they do not converge:

$$G^A = 2 - \frac{1}{2}(3) + \frac{1}{4}(5+1) - \frac{1}{8}(9) + \frac{1}{16}(17+1) \ldots\ldots - \frac{1}{2^{2k-1}}(1+2^{2k-1}) + \frac{1}{2^{2k}}(1+2^{2k}+1) \ldots.$$

$$G^B = -2 + \frac{1}{2}(3+1) - \frac{1}{4}(5) + \frac{1}{8}(9+1) - \frac{1}{16}(17) + \ldots. + \frac{1}{2^{2k-1}}(1+2^{2k-1}+1) - \frac{1}{2^{2k}}(1+2^{2k})\ldots.$$

Thus, it is not possible to define present discounted gains associated with the myopic REE strategies. Nevertheless, we may observe that A (resp. B) can always guarantee himself 2 (resp. 0) by leaving the market just after selling. In fact, if a trader wants to maximize his present discounted gain, he has to "realize his profits" by refusing to repurchase the stock at some date; this strategy can also be viewed as a dominant strategy in that the trader avoids running the risk of getting stuck with a devalued stock if the other trader switches to a "finite time strategy". Thus it would be natural to assume that A's payoff is 2 and B's payoff is 0. But those payoffs are inconsistent since they must add up to the market fundamental which is 1.

To summarize, in a myopic REE, each trader must 1) believe that he will be able to sell the asset, 2) realize his profits in finite time. These two conditions are inconsistent with the assumption that the number of traders is finite.[16]

b. Fully Dynamic REE

Requiring that the strategy of each trader maximizes his expected present discounted gain leads to the definition of a fully dynamic REE.

Definition

A fully dynamic REE is a sequence of self-fulfilling forecast functions: $s_t = (\ldots, s_t^i, \ldots) \rightarrow p_t = \Phi_t(s_t) \iff s_t \in S_t(p_t) \equiv \Phi_t^{-1}(p_t)$, such that there exists a sequence of (information contingent) stock holdings (strategies) $x_t^i(s_t^i, p_t)$ satisfying:

(1) Market Clearing: $\forall t, \forall s_t: \sum_i x_t^i(s_t^i, p_t) = \bar{x}$.

(2) Maximizing Behavior: At each time t, and for any information (s_t^i, S_t) trader i may possess, i's strategy (restricted to the information sets reachable from (s_t^i, S_t) at t), maximizes i's expected present discounted gain from t on - i's posterior being computed from the common prior and i's information (s_t^i, S_t).

As the following proposition shows, long-run maximizing behavior considerably restricts the eligible set of price and forecast functions:

Proposition 7: Whether short sales are allowed or not, price bubbles do not exist in a fully dynamic REE:

$\forall t, \forall s_t, \forall i: F(s_t^i, S_t) = p_t$, i.e.,

$B(s_t^i, p_t) = 0$.

Proof: We prove proposition 7 in the case where short sales are prohibited.

Let $(x_t^i(s_t^i, p_t))$ be a set of optimal strategies.

Let $G_t^i \equiv \sum_{\tau=1}^{\infty} \gamma^\tau d_{t+\tau} x_{t+\tau-1}^i + \sum_{\tau=1}^{\infty} \gamma^\tau p_{t+\tau} (x_{t+\tau-1}^i - x_{t+\tau}^i)$

be the discounted sum of realized dividends and capital gains associated with i's optimal strategy.

Clearly the G_t^i's add up to the market fundamental times the quantity of the stock:[17]

$$\sum_i G_t^i = (\sum_{\tau=1}^{\infty} \gamma^\tau d_{t+\tau})\, \bar{x} = f_t \bar{x} \qquad (12)$$

where f_t denotes the "realized market fundamental", i.e., the discounted sum of the realized dividends per unit of stock from t on. The proof uses the following lemmas:

Lemma 1: The market fundamental relative to the market information exceeds the price: $\forall s_t$: $F(S_t) \geq p_t$

Proof of Lemma 1: Since trader i optimizes, he can not gain by selling x_t^i and leaving the market at time t:

$$E(G_t^i | s_t^i, S_t) \geq x_t^i\, p_t$$

Thus:

$$E(G_t^i | S_t) = \sum_{s_t^i \varepsilon S_t^i} E(G_t^i | s_t^i, S_t)\ \nu^i(s_t^i | S_t)$$

$$\geq p_t (\sum_{s_t^i \varepsilon S_t^i} x_t^i\ (s_t^i, p_t)\ \nu^i(s_t^i | S_t))$$

where S_t^i denotes the projection of S_t on F_t^i. The last expression in brackets is nothing but the statistical average of i's stockholding at price p_t. This implies

$$\sum_i E(G_t^i | S_t) \geq p_t\, \bar{x}$$

or

$$E(f_t\, \bar{x} | S_t) \geq p_t\, \bar{x} \Rightarrow F(S_t) \geq p_t \qquad \text{Q.E.D.}$$

Lemma 2: No trader expects a gain from his trade at time t:

$\forall t, \forall s_t, \forall i$: $E(G_t^i | s_t^i, S_t) = E(G_t^i(x_{t-1}^i) | s_t^i, S_t)$
where $G_t^i(x_{t-1}^i)$ is defined as G_t^i except that at t, i holds x_{t-1}^i instead of $x_t^i(s_t^i, p_t)$ (the holding strategies being unchanged after t).

Proof of Lemma 2: From (12), the trading game at t, given the holding strategies beyond t is a zero-sum game:

$$\forall t,\ \forall\{s_t\}: \sum_i (G_t^i - G_t^i(x_{t-1}^i)) = 0$$
$$\Rightarrow \sum_i E(G_t^i - G_t^i(x_{t-1}^i) \mid S_t) = 0$$

The optimizing trader i cannot improve upon $\{x_t^i\}$ by holding x_{t-1}^i at t, and following the same strategy beyond t:

$$\forall t,\ \forall s_t,\ \forall i:\ E(G_t^i - G_t^i(x_{t-1}^i) \mid s_t^i, S_t) \geq 0$$

Now one can apply the same argument as in the proofs of propositions 1,2, and 3 to the functions $\{G_t^i - G_t^i(x_{t-1}^i)\}$ Q.E.D.

Using Lemma 1: $\forall t,\ \forall s_t,\ \forall i: \sum_{s_t^i \in S_t^i} F(s_t^i, S_t)\ \nu^i(s_t^i \mid S_t) \geq p_t$

Imagine now that the market fundamental of some agent i_o who does not hold the whole stock at the start of the period $(x_{t-1}^{i_o} < \bar{x})$ were to strictly exceed the price: $F(s_t^{i_o}, S_t) > p_t$. Then i_o could buy and make a strictly positive expected profit, contradicting Lemma 2. Thus: $\forall i$ such that $x_{t-1}^i \neq \bar{x}$, $\forall s_t$: $F(s_t^i, S_t) = p_t$. Integrating the previous equality gives: $F(S_t) = p_t$. Now if i holds the whole stock at the beginning of the period, his market fundamental can not be lower than p_t without contradicting Lemma 2. Thus:

$\forall s_t^i$: $F(s_t^i, S_t) \geq p_t$ - But then: $F(s_t^i, S_t) = p_t$ Q.E.D.

Remark 1: Proposition 7- the law of equalization of market fundamentals - does not imply that the price p_t fully reveals the complete signal s_t. Consider the following example: There are two traders A and B and two Bernoulli processes independent and uncorrelated over time:

$$s_t^A = \begin{cases} 0 \text{ with probability } 1/2 \\ 1 \text{ with probability } 1/2 \end{cases} \qquad s_t^B = \begin{cases} 0 \text{ with probability } 1/2 \\ 1 \text{ with probability } 1/2 \end{cases}$$

Assume that the dividend depends on the signals in the following way:

$$d_{t+1} = (s_t^A + s_t^B) \bmod 2$$

(i.e., $d_{t+1}(0,0) = d_{t+1}(1,1) = 0$, $d_{t+1}(0,1) = d_{t+1}(1,0) = 1$)

With a discount factor 1/2, the market fundamental corresponding to the absence of information is 1/2. It is easy to see that the following non informative price function is a fully dynamic REE.

s_t^A \ s_t^B	0	1
0	1/2	1/2
1	1/2	1/2

Remark 2: Harrison and Kreps (1978) have shown that in a stock market in which priors differ, yet are common knowledge, and are never updated, the market price strictly exceeds the market fundamental of the traders. Thus the right

to resell the stock gives traders the incentive to pay more for it than if they were obliged to hold it forever.

Their result may still hold with identical priors, differential information and updating, if one takes a <u>self-fulfilling equilibrium</u>. Consider the following example (due to David Kreps): The model is the same as in the previous remark, except for the dividend process: $d_{t+1} = s_t^A + s_t^B$.

The following stationary price function leads to a <u>self-fulfilling equilibrium</u>:

s_t^A \ s_t^B	0	1
0	$\frac{9}{16}$	$\frac{21}{16}$
1	$\frac{21}{16}$	$\frac{21}{16}$

For example, when $s_t^B = 1$, B believes that the next dividend will be 1 with probability 1/2 and 2 with probability 1/2, since he cannot infer anything from the price. Thus he is willing to pay:

$$\frac{1}{2}\left(\left(\frac{1}{2} \times 1 + \frac{1}{2} \times 2\right) + \left(\frac{1}{4} \times \frac{9}{16} + \frac{3}{4} \times \frac{21}{16}\right)\right) = \frac{21}{16}$$

Now assume that $(s_t^A, s_t^B) = (0,1)$. A is fully informed (and is willing to pay $\frac{17}{16}$); thus he does not want to hold the stock (short sales are assumed to be prohibited). B, who holds the stock, has for a market fundamental:

$$\frac{1}{2}\left(\frac{1}{2} \times 1 + \frac{1}{2} \times 2\right) + \frac{1}{4}\left(\frac{1}{2} \times 1 + \frac{1}{4} \times 0 + \frac{1}{4} \times 2\right) + \ldots = \frac{20}{16} < \frac{21}{16} = p_t$$

thus the Harrison-Kreps result holds.

The explanation of this result is the following: B anticipates that he will be able to sell when information is (0,1); since his market fundamental $(\frac{16}{16})$ will then be lower than the price $(\frac{21}{16})$, he is now willing to pay more $(\frac{21}{16})$ than his market fundamental $(\frac{20}{16})$.

This should remind us of the Feiger model: When B observes $s_t^B = 1$, he ought to realize that he is playing against a better informed trader. Of course if the quantity x_t^B were "measurable" (i.e., depended only on B's information), B would be willing to trade even if he realized that A is better informed, but then equilibrium would be destroyed. Even if B's information makes him indifferent to all feasible trades, it makes a difference whether B lets an auctionner (or the market) pick his trade, or if he chooses it himself![18]

Note that the kind of price bubble arising in the previous example cannot be observed in a myopic REE (and of course not in a fully dynamic REE where no price bubble exists). Proposition 5 tells us that in a myopic REE prices have to grow "on average" if price bubbles exist. On the contrary the price function of the previous example is constant over time; so are the price bubbles.[19] These price bubbles come from the fact that traders never infer any information from the quantity they trade. This lasting "mistake" is permanently embodied in the price in addition to the market fundamental (note that one might add any exponentially growing price bubble of the usual kind to the self-fulfilling equilibrium).

To conclude, the Harrison-Kreps result still holds with myopic REE, but the price bubbles are of the exponentially growing kind. However, in a fully dynamic REE, no speculation in the Kaldor-Keynes sense can be observed.

Chapter 5

Conclusions

We first summarize our observations:

1. The Working theory of speculation - that speculation is created by differences in expectations - relies on irrationality of at least some traders or on differences in priors. Differences in expectations, contrary to differences in risk aversion or in riskiness of initial positions, do not give rise to any gain from trade if the priors are identical; it is then intuitive that the Working theory is inconsistent with rationality of the traders.

2. Insurance motives create meaningful gains from trade (non-zero-sum game) and may give rise to speculation. This is demonstrated by a futures market with risk averse traders with initially risky positions (farmers) and risk averse traders with initially riskless positions and endowed with information (speculators). To show that the *raison d'être* of the trade between farmers and speculators is insurance and not differences in information, we introduce risk neutral traders (insurance companies) into the market. Then speculators do not trade any more. We also indicate that such an equilibrium strains the credibility of REE to the limit, since some or all of the information detained by the speculators is conveyed by the price.

3. Dynamic speculation can be characterized by the same methods as static speculation. A handy way to represent

a sequential asset market is to describe it as a sequence of REE. We distinguished two kinds of REE's: Myopic REE's where price bubbles may develop, but possess martingale properties, and fully dynamic REE's where price bubbles do not exist at all. The concept of REE rules out price bubbles arising from differences in information whereas the intertemporal maximizing behavior and the equilibrium concept rule out bubbles of the exponentially growing kind. We indicated that the concept of fully dynamic REE was most reasonable if one is willing to posit rationality of the traders. This certainly does not deny the relevance of (and the need for) positive models trying to describe actual price bubbles by non-rational behaviors.[20]

What are the implications of our results for a real world asset market? It is clear that the assumptions underlying a REE are very strong: in particular the existence of a self-fulfilling forecast function is by itself difficult to justify. In some simple cases, such as the one of the speaker in a seminar given in the introduction, one may be able to appeal to the existence of some common knowledge about the market; in particular, in the real world, people should certainly worry about the quality of their information relative to the other traders'.

Behind the mathematics lie two basic principles. First, one should not count on differences in information in order to achieve a speculative gain. This result is best understood by using the second interpretation of a REE as a forecast function which cannot be refuted on the basis of the common knowledge about the market described in the introduction, and by observing that not everyone can possess "better than average" information. Of course, in a market where some other traders do rely on the belief that they have superior information, it might pay to do so as well. We then face a recursive problem. The question is: Can rational traders expect in equilibrium a speculative gain based on their allegedly superior information or their information concerning the other traders' behavior? The common knowledge version of a REE would require the answer to be no. Second, in a dynamic framework with a finite number of agents, a rational trader will not enter a market where a bubble has already grown, since some traders have already realized their gains and left a negative-sum game to the other traders. Again, if one is able to find a "sucker", it may pay to participate. The point is that in an equilibrium with a finite number of traders, it is not possible for everyone to find a buyer and avoid "getting stuck with a hot potato". This is not to deny the positive relevance of Keynes' "Castles in the Air" theory, which undoubtedly explains a number of speculative phenomenons: in fact, more research should be devoted to the manipulability and controllability of speculative markets. But section IV certainly vindicates the "Firm Foundation" asset pricing theory as a normative concept; moreover, the views developed above have some counterparts in the investment literature (see, e.g., Malkiel [1975]).[21]

References

Aumann, R. (1976) "Agreeing to Disagree" The Annals of Statistics, 1236-1239.

Blanchard, O. (1979) "Speculative Bubbles, Crashes and Rational Expectations" Economic Letters, 3, 387-389.

Bray, M. (1980a) "Futures Trading, Rational Expectations and the Efficient Market Hypothesis", forthcoming in Econometrica.

Bray, M. (1980b) "Learning, Estimation, and the Stability of Rational Expectations", Research Paper #550. Graduate School of Business, Stanford.

Danthine, J.P. (1978) "Information, Futures Prices and Stabilizing Speculation" Journal of Economic Theory, 17, 79-98.

De Groot, M. (1970) Optimal Statistical Decisions, New York. McGraw-Hill.

Feiger, G. (1976) "What is Speculation?" Quarterly Journal of Economics XC 677-687.

Feiger, G. (1978) "Divergent Rational Expectations Equilibrium in a Dynamic Model of a Futures Market" Journal of Economic Theory 17, 164-178.

Flood, R., Garber P. (1979) "Market Fundamentals vs. Price Bubbles: The First Tests", Journal of Political Economy, forthcoming.

Grossman, S. (1976) "On the Efficiency of Competitive Stock Markets where Traders have Diverse Information" Journal of Finance 31, 573-585.

Grossman, S. (1977) "On the Existence of Futures Markets, Noisy Rational Expectations, and Informational Externalities" Review of Economic Studies, 64, 431-449.

Grossman, S. (1979) "Rational Expectations and the Allocation of Resources Under Asymmetric Information: A Survey" Working Paper #79-18, University of Pennsylvania.

Grossman, S., J. Stiglitz (1976) "Information and Competitive Price Systems" American Economic Review Papers and Proceedings, 246-252.

Harrison, J.M., D.M. Kreps (1978) "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations" Quarterly Journal of Economics, 323-336.

Hirshleifer, J. (1975) "Speculation and Equilibrium: Information, Risks and Markets" Quarterly Journal of Economics, LXXXIX, 519-42.

Hirshleifer, J. (1977) "The Theory of Speculation Under Alternative Regimes of Markets" Journal of Finance 32, 975-1000.

Kreps, D.M. (1977) "A Note on 'Fulfilled Expectations' Equilibria" Journal of Economic Theory, pp. 32-43.

Levine, D. (1980) "Theory of Price Bubbles" Mimeo M.I.T.

Malkiel, B.G. (1975) "A Random Walk Down Wall Street" Norton & Company Inc., second edition.

Milgrom, P., N. Stokey (1980) "Information, Trade and Common Knowledge", Discussion Paper #377, Northwestern University.

Miller, E. (1977) "Risk, Uncertainty and Divergence of Opinion" Journal of Finance 32, 1151-1168.

Radner, R. (1972) "Existence of Equilibrium in Plans, Prices and Price Expectations in a Sequence of Markets" Econometrica, 40, 289-303.

Radner, R. (1979) "Rational Expectations Equilibrium: Generic Existence and the Information Revealed by Price" Econometrica, 47, 655-678.

Samuelson, P. (1973) "Proof that Properly Discounted Present Values of Assets Vibrate Randomly" Bell Journal of Economics, 369-374.

Sargent, T.J., N. Wallace (1973) "The Stability of Models of Money and Growth with Perfect Foresight" Econometrica.

Stiglitz, J.E. (1974) "Information and Capital Markets", Oxford University, Mimeo.

Footnotes

1. This work arose from discussions I had with Drew Fudenberg and Eric Maskin on speculation. I am also very grateful to them as well as to Peter Diamond, David Levine and David Kreps for helpful comments. This paper has benefited from discussions in seminars at MIT, Cambridge University and the Roy seminar in Paris. Eric Maskin provided very helpful comments on the current version.

2. Perhaps a better name than REE is "fulfilled expectations equilibrium" (see Kreps (1977)). In this paper, we require the demand of a trader to depend only on the price, the trader's private information and the information conveyed by the price.

3. We shall assume throughout the paper that traders have a common prior: Different beliefs are due solely to differences in information (signals).

4. The first version of this paper used the definition of a REE as a forecast function which cannot be refuted by any trader on the basis of the above common knowledge about the market. Eric Maskin pointed out to me that this definition was equivalent to the usual definition of a REE when one imposes measurability. Hence the results derived in this paper do not depend on any knowledge of the traders about the market other than the price and the relationship between signals and prices.

5. This point may remind the reader familiar with the literature on auctions of the winner's curse.

6. "Uncorrelated" is relative to the information of the trader. This definition is more stringent than the condition that the initial position be uncorrelated with the return of the asset and the signal received by the trader. Kreps (1977) observes that the information conveyed by the price may introduce some correlation between the initial position and the return of the asset and thus create an insurance motive for speculation. Hence, we assume that the initial positions of all traders are uncorrelated with the return of the asset and the set of signals.

7. See, for example, Grossman (1976, 1977).

8. See, for example, Harrison-Kreps (1978), Hirshleifer (1975, 1977) and Miller (1977).

9. For a more complete model of a futures market, see Bray (1980a).

10. In Chapters 3a and 4 where it is assumed that all signals have a positive probability, the derivation of (8) is straightforward. Here we assumed that s^{jo} is drawn from a continuous distribution; one must then be cautious and invoke a continuity argument in order to derive (8).

11. Remember that "riskless" must not be taken too literally: it simply means that the speculators' initial positions are uncorrelated with p. Also Proposition 2 is stated in slightly more general terms than would be allowed by the proof. It is easy to check that even if farmers and insurance companies have private information (signals), speculators do not trade in a REE.

12. For example if all speculators have the same information (signal), it is easy to see that a REE exists and is unique. This is the fully revealing one: $p = \bar{p} + s$.

13. Traders do not have a budget constraint at each period. They can borrow and lend at the rate $(1/\gamma - 1)$.

14. See, for example, Sargent-Wallace (1973), Flood-Garber (1979), Blanchard (1979) as well as some of the literature on growth with heterogeneous capital goods.

15. Proposition 3 can be seen as a generalization to heterogeneous information of Samuelson (1973)'s theorem 3 (note also that Samuelson hypothesizes that a stock's present price is set at the expected discounted value of its future dividends, which we do not impose here). It also shows that, even with differential information, no trader can gain from his trade, which restricts the validity of Samuelson's conjecture (p. 373) that "there is no incompatibility in principle between behavior of stock prices that behave like random walk at the same time that there exists subsets of investors who can do systematically better than the average investors" to subsets of traders with "measure zero".

16. This is not true with infinite number of traders. For example, in an overlapping generation model, a price bubble is consistent with each generation leaving the market after realizing its profit.

17. With an infinite number of traders, the adding up in (12) may make no sense. Consider the perfect information stock market described at the end of Chapter 4a. Assume now that there exists a countable number of infinitely-lived traders $(A_0, A_1, \ldots, A_t, \ldots)$. Consider the following

sequence of trades (A_0 holds the stock initially):

At time 0, trader A_0 sells the stock to trader A_1 at price 2

.... 1, ... A_1 .... A_2 3
.... t, ... A_t .... A_{t+1}.......$(1+2^t)$

then the present discounted pay-off for all traders but A_0 is 0; for A_0, it is 2. But the market fundamental is 1. This may remind the reader of the familiar paradoxes of infinity. Note that this example does not depend on non-maximizing behavior of the traders.

18. This point has already been recognized by David Kreps ((1977), section on "information from quantity").

19. The price bubbles, for trader A for example, are $\frac{1}{16}$, $\frac{5}{16}$, $\frac{1}{16}$, when information is {(0,0), (0,1), (1,0) or (1,1)}

20. For an example of a behavioral theory of price bubbles, see Levine (1980).

21. Does this imply that throwing darts on a financial page of the Wall Street Journal is an optimal strategy for portfolio selection? Not really, as evidenced by the example of a player with a random demand in a model à la Grossman (see end of Chapter 3a).