Essays on Altruism and Health Care Markets

STOCKHOLM SCHOOL OF ECONOMI
EFI, THE ECONOMIC RESEARCH INSTITUT

EFI Mission

EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution which works independently of economic, political and sectional interests. It conducts theoretical and empirical research in management and economic sciences, including selected related disciplines. The Institute encourages and assists in the publication and distribution of its research findings and is also involved in the doctoral education at the Stockholm School of Economics.

EFI selects its projects based on the need for theoretical or practical development of a research domain, on methodological interests, and on the generality of a problem.

Research Organization

The research activities are organized in nineteen Research Centers within eight Research Areas. Center Directors are professors at the Stockholm School of Economics.

ORGANIZATION AND MANAGEMENT	
Management and Organisation; (A)	Prof Sven-Erik Sjöstrand
Center for Ethics and Economics; (CEE)	Adj Prof Hans de Geer
Public Management; (F)	Prof Nils Brunsson
Information Management; (I)	Prof Mats Lundeberg
Center for People and Organization (PMO)	Acting Prof Jan Löwstedt
Center for Innovation and Operations Management; (T)	Prof Christer Karlsson
ECONOMIC PSYCHOLOGY	
Center for Risk Research; (CFR)	Prof Lennart Sjöberg
Economic Psychology; (P)	Prof Lennart Sjöberg
MARKETING	
Center for Information and Communication Research; (CIC)	Adj Prof Bertil Thorngren
Center for Consumer Marketing; (CCM)	Associate Prof Magnus Söderlund
Marketing, Distribution and Industrial Dynamics; (D)	Prof Lars-Gunnar Mattsson
ACCOUNTING, CONTROL AND CORPORATE FINANCE	
Accounting and Managerial Finance; (B)	Prof Lars Östman
Managerial Economics; (C)	Prof Peter Jennergren
FINANCE	
Finance; (FI)	Prof Clas Bergström
ECONOMICS	
Center for Health Economics; (CHE)	Prof Bengt Jönsson
International Economics and Geography; (IEG)	Prof Mats Lundahl
Economics; (S)	Prof Lars Bergman
ECONOMICS STATISTICS	
Economic Statistics; (ES)	Prof Anders Westlund
LAW	
Law; (RV)	Prof Erik Nerep

Chairman of the Board: Prof Sven-Erik Sjöstrand. *Director:* Associate Prof Bo Sellstedt,

Adress

EFI, Box 6501, S-113 83 Stockholm, Sweden • Internet: www.hhs.se/efi/
Telephone: +46(0)8-736 90 00 • Fax: +46(0)8-31 62 70 • E-mail efi@hhs.se

ESSAYS ON ALTRUISM AND HEALTH CARE MARKETS

Björn Persson


STOCKHOLM SCHOOL OF ECONOMICS
EFI, THE ECONOMIC RESEARCH INSTITUTE

Dissertation for the Degree of Doctor of Philosophy, Ph.D.
Stockholm School of Economics, 2001

ISBN NR 91-7258-563-3

Keywords:
Altruism, innovation, insurance, learning, pharmaceutical industry, pricing strategy, product differentiation, regulation.

Distributed by:
EFI, Stockholm School of Economics,
Box 6501, S-113 83 Stockholm, Sweden

*Printed by:*HHS PrintCenter, Stockholm 2001

Contents

Acknowledgements

During the writing of this thesis, I have received advice and encouragement from many. First, I would like to thank my thesis committee: Magnus Johannesson, Bengt Jönsson, and Jörgen Weibull for their guidance and support. I have benefited greatly from their expertise in a wide range of economic issues, and I highly value the strong personal commitment that they have shown to my work. I owe special thanks to Magnus Johannesson for his unfailing support and helpfulness, which have been essential to the completion of this thesis.

In this thesis, I was fortunate to share authorship with Mats Ekelund, Magnus Johannesson, and Jörgen Weibull. All of these collaborations were both intellectually rewarding and enjoyable.

Many others at the Stockholm School of Economics have contributed to the helpful and cordial atmosphere during my years in graduate school. I am very grateful to Tore Ellingsen, Ulf Gerdtham, and Douglas Lundin who made valuable comments and suggestions that improved this thesis. Erik Berglöf, Annie Maccoby, Martin Hill, Sophie Langenskiöld, Chloé Le Coq, and María Sáez-Martí have my warm thanks for their strong support and friendship.

I have appreciated the kind help and administrative assistance I have received from Carin Blanksvärd, Britt-Marie Eisler, Pirjo Furtenbach, Ritva Kiviharju, Kerstin Niklasson, and Ingrid Nilsson over the years at the SSE.

I wish to express my gratitude to the National Association of Swedish Pharmacies (*Stiftelsen Apoteksbolagets Fond för forskning och studier i hälsoekonomi och socialfarmaci*) for financial support during the writing of this thesis. I also wish to thank Läkemedelsindustriföreningen (LIF) for its generous provision of data.

I thank my parents Ulla and Bengt for their constant encouragement. Finally, I am grateful to Tina, Jens, Emil, and, in particular, to Jody.

Stockholm in April 2001

Björn Persson

Summary of essays

Summary of essays

This thesis consists of two parts. The first part includes two papers that deal with the Swedish pharmaceutical market, and one paper that looks at strategic incentives that arise in optimal treatment involving untried drugs. The second part, consisting of two papers, examines some implications of altruism.

I. Pharmaceutical pricing in a regulated market.

The first essay compares pharmaceutical pricing in the price regulated Swedish market with pricing in the US market, where no such regulation exists. We consider all New Chemical Entities (NCEs) introduced in Sweden between the years 1987 and 1997 and compare our results with the findings of a previous study performed on US data (see Lu and Comanor [1998][1]). We examine the development of launch prices (relative to prices of existing substitutes) and price dynamics using the same explanatory variables as in Lu and Comanor. Specifically, we consider the influence on pricing of the degree of therapeutic innovation and the presence of similar brand name pharmaceuticals. Furthermore, in line with Lu and Comanor we also consider whether generic competition has any effect on pricing. Similar to Lu and Comanor, we find that the relative introductory prices are positively correlated with the degree of therapeutic innovation. The average relative launch prices are, however, higher in Sweden. Contrary to the findings on the US market, real prices for NCEs fall substantially over time for all classes of therapeutic innovation. Moreover, unlike Lu and Comanor, we find no effect of brand substitutes or from generic competition on either introduction prices or price dynamics. We conclude that the price regulation effectively inhibits the use of penetration strategies, and that price competition between pharmaceuticals is less pronounced in Sweden than in the US.

II. Innovativeness and market shares in the pharmaceutical industry.

In the second essay, we use the same NCE data to examine the extent to which newly introduced pharmaceuticals gain market shares in Sweden. We assume that new pharmaceuticals differ from existing products in two dimensions. The first dimension represents the pharmaceutical's (medical) efficacy. The second dimension represents other relevant characteristics that the consumer values, e.g., the extent

[1]Lu, Z. J., and W. S. Comanor [1998] "Strategic Pricing of New Pharmaceuticals", *Review of Economics and Statistics*, 80 (1), 108-118.

and nature of side effects or the administration form. Hence, the former dimension is a general quality aspect over which consumers have homogeneous preferences, whereas the latter represents specific quality aspects over which consumers have heterogeneous preferences. We therefore employ a simplified model of horizontal and vertical product differentiation as a tool to derive the main implications these assumptions entail. We use the degree of therapeutic innovation as a proxy for medical quality, or efficacy. The extent of brand substitutes is assumed to reflect the horizontal dimension in the market. The empirical findings correspond to the predictions of the theoretical model. In particular, we find that vertically differentiated pharmaceuticals gain larger market shares on average and command higher prices than horizontally differentiated pharmaceuticals. Moreover, as a general rule competing substitutes have less influence on the former than on the latter.

III. A simple model of strategic experimentation.

The third essay also relates to health markets, in a somewhat more abstract formulation. We consider a simple model of strategic interaction in which two individuals acting under uncertainty must learn about a payoff relevant parameter. Specifically, we assume that the individuals are physicians who may invest in knowledge about the efficacy of an unknown treatment (e.g., a pharmaceutical drug). Two treatments are available, both of which produce binary outcomes of either success or failure. One of the treatments has a known success rate and represents an established drug, whereas the other is a new treatment with unknown probability of success. At the outset, both physicians have identical prior assessments of the new (unknown) treatment's success rate. Only by prescribing the treatment will the physicians gain information about its characteristics. In each of two periods, the physicians allocate the treatments to their patients. The problem consists of choosing whether to prescribe the new or the established treatment at an early stage. If the physicians prescribe the new treatment, they will gain information that can be used to treat subsequent patients more effectively. On the other hand, prescribing the new treatment also involves a gamble. Therefore, the physicians face a trade-off whether to sacrifice payoff at an early stage for increased future payoff.

This paper considers two information scenarios. In the first, we suppose that the physicians cannot observe each other's prescription choices. The problem is then a simple variant of the one-armed bandit problem, extensively discussed in the statistics literature (see e.g., Gittins [1989][2]). In the other scenario, we suppose that they can observe the result of each other's treatments as well as their own. This assumption introduces a possibility for the physicians to learn about the new treatment without having to administer it to their own patients.

[2]Gittins, J. C. [1989] *Multi-armed bandit allocation indices*, Wiley, Chichester.

In the pure equilibria for the interesting parameter values, one of the physicians prescribes the new treatment at an early stage, and the other prescribes the established treatment. This outcome also produces the highest combined expected payoff. However, strong free riding effects are likely to arise since the physicians both prefer to let the other prescribe the new treatment: it is better to play safe and use the established treatment, while gaining the freely available information generated by the other physician's experimentation with the new treatment. Therefore, failure to coordinate the experimentation activity between the physicians is likely to lead to an out-of-equilibrium outcome in which none of the physicians prescribes the new treatment. In this case, no new information is gained about the new treatment, and this outcome is the worst in terms of expected payoffs.

IV. Altruism as a cause of insurance market failure.

This essay examines the behavior on insurance markets when individuals have altruistic concerns for others' welfare. We consider a large population of equally altruistic individuals who all run a risk of losing their health with some exogenous probability. Against a premium, the individuals can obtain insurance which fully restores their health. Due to the altruistic preferences, the individuals may consider transferring funds to those who are in worse health states than themselves. This potentially has distorting effects on the individuals' decisions to purchase insurance: it is now possible to forego insurance and exploit the good will of others if an adverse health outcome arises. When only a few individuals interact, the presence of altruism can potentially lead to severe savings and insurance distortions. In contrast to most of the previous literature, we consider how altruism may create incentives to forego insurance when there is a large number of individuals. In addition to the free riding effects that arise when receivers exploit the donors, free riding may now arise where donors exploit other donors. Voluntary donations are a public good, and therefore risk being under-provided in equilibrium. That is, each individual may consider their personal contribution to be so marginal in relation to the total that giving seems superfluous. Consequently, there is a possibility that no one makes any donations in equilibrium.

We find, however, that there will be a positive total donation even when the population size approaches infinity. In the case where all the individuals have the same probability of falling ill, we find that even moderate levels of altruism can cause complete market failure. In that case, no individual purchases insurance and those who remain healthy will help those who fall ill. However, it is simultaneously an equilibrium that all individuals purchase insurance and that no transfers are made. When we introduce some heterogeneity between the individuals, this result is modified and altruism must now be rather substantial to cause any severe distortions on the

insurance market.

We also consider how altruism relates to the information structure that prevails between the insurance providers and the individuals who purchase insurance. In the absence of altruism, asymmetric information about the risk of falling ill may give rise to well-known adverse selection effects when individuals with low health risks to prefer not to insure. A natural question is then whether altruism reinforces or counteracts these effects. We find that altruism strengthens the adverse selection effects and thus causes more individuals to forego insurance.

V. Non-reciprocal altruism in dictator games.

This essay also looks at behavior consistent with preferences for others' utility. We consider how subjects allocate funds between others and themselves in an experimental setting. In particular, we let subjects play a dictator game in which one individual divides a sum of money between himself and a recipient who must accept this allocation. The standard game theoretic prediction of the outcome in this game is that the dictator keeps the money and shares nothing with the recipient. Since deviations from this outcome have been observed in numerous experiments, it has been suggested that the standard assumption of (solely) payoff maximizing individuals is misguided. Instead, individuals are assumed to have altruistic preferences, or a taste for fairness. This explanation has, however, also met with some dispute. The reason why individuals make positive contributions, it is argued, is that they expect others to reciprocate in future interactions. This argument is found in e.g., Hoffman et al. [1996] [3] who hypothesized that increasing social distance would lead to decreasing donations, where social distance is a measure of anonymity between subjects (and the experimenter).

In this essay, we test this hypothesis with a large degree of social distance, randomly drawing the recipients from the general Swedish population. They are unaware that they participate in the experiment. This design removes any remaining reciprocity from the double blind procedure used by Hoffman et al. In line with their hypothesis, we should therefore expect the average donation to decrease to zero. We contrast this design with a replication of the standard double blind procedure used in Hoffman et al., and test the null hypothesis that the donation distributions do not differ in the two experimental treatments. We found that subjects made positive donations in both designs. Although the average donation was slightly lower in the standard double blind design, we could not reject the null hypothesis that the two treatments produced the same donations. This result was interpreted in favor of individuals having altruistic concerns.

[3] Hoffman, E., McCabe, K., and V. Smith [1996] "Social distance and other-regarding behavior in dictator games", *American Economic Review* 86, 653-660.

Essay I: Pharmaceutical pricing in a regulated market

Pharmaceutical pricing in a regulated market

MATS EKELUND AND BJÖRN PERSSON*
STOCKHOLM SCHOOL OF ECONOMICS

ABSTRACT. We compare how new pharmaceuticals are priced in the price regulated Swedish market with the pricing of new pharmaceuticals in the US market, as studied by Lu and Comanor [1998]. We collect a data set consisting of all New Chemical Entities (NCEs) launched in Sweden between 1987 and 1997, and test the same models as Lu and Comanor [1998]. In line with their results, we find that introductory prices depend on the degree of therapeutic innovation. The average relative launch prices are, however, higher in Sweden.

Further, contrary to the results from the US market, Swedish real prices for NCEs fall substantially over time for all classes of therapeutic innovation. Also contrary to the findings of Lu and Comanor [1998] we find no effect of the presence of branded substitutes on either introduction prices or price dynamics. Our results indicate that the price regulation discourages the use of penetration strategies and decreases price competition between brand name drugs.

1. INTRODUCTION

The rapid increase in pharmaceutical expenditures in many countries (see e.g. Abbott [1995] and Besley and Gouveia [1994]), has generated much interest in the pricing of pharmaceuticals. It is sometimes argued that the presence of patent protection and third-party financing could lead to excessive prices, and therefore many nations have some form of expenditure regulation. One example is price caps, which are currently gaining in popularity with regulatory agencies and government institutions. However, the use of price caps as a means to curb pharmaceutical expenditures is the subject of some controversy. Opponents of price (cap) regulation argue that this may adversely affect incentives to develop new and better products, since producers are not adequately reimbursed for the massive investments needed to bring new drugs to the market. This standpoint is taken by many in the United States where no such regulation exists, (see e.g. Danzon [1997]). It has also been maintained that price caps are not even effective in curbing pharmaceutical outlays (Abbott [1995] and Danzon and Chao [2000]).

*We thank Magnus Johannesson for his advice and comments. We have also benefited from comments and discussions with Tore Ellingsen, Ulf Gerdtham, and Bengt Jönsson.

In order to evaluate the effects of price regulations and to resolve the controversies surrounding pharmaceutical prices, we believe that further empirical work is desirable. In this paper, we take a step in this direction, performing a basic comparison of pricing patterns in a market with regulated prices with a market where prices are not regulated. More specifically, we consider how the pricing strategies of pharmaceuticals in Sweden relate to the pharmaceutical pricing in the US market using identical explanatory variables.

There have been a number of empirical studies on the pricing of new patented drugs in the US market. Reekie [1978], in a seminal study, examined the introductory prices of new chemical entities (NCEs) launched in the US between 1958 and 1975. He found that the introductory price (relative to existing substitutes) depended on the degree of therapeutic advance, i.e. the price premium was larger for drugs offering important therapeutic gains compared to drugs offering minor therapeutic gains. Reekie [1978] also found that prices tended to increase faster for drugs introduced at lower prices compared to drugs with higher launch prices.

The pricing pattern observed by Reekie is consistent with the observation made by Dean [1969], who distinguished between two pricing strategies for new products: skimming and penetration. A skimming strategy involves setting a relatively high introductory price to skim off the highest willingness to pay, and then lowering the price. Penetration is the opposite strategy: lower introductory prices followed by increased prices as demand picks up. Dean [1969] argued that skimming strategies are more often used for products offering major advantages over existing products, and that penetration strategies are used for products that offer only marginal improvements over existing products.

Recently, Lu and Comanor [1998] (henceforth LC) performed a study of 144 NCEs introduced in the US between 1977 and 1987.[1] The principal motive of their paper was to explore and quantify the demand side determinants for NCE prices. In their view, therapeutic value and market structure are the main explanatory variables for NCE pricing. This is in contrast to the more established supply side arguments often used in the debate over pharmaceutical prices.

LC classified the new pharmaceuticals into three different classes of therapeutic gain, depending on whether the drugs represented a major, a minor, or no improvement over existing drugs. Introductory prices for drugs representing important therapeutic gains were on average about three times higher than those for existing substitutes, whereas new drugs that offered little therapeutic gains were launched at about the same price as existing substitutes. The real prices for drugs representing important therapeutic gains were relatively stable over time. However, for drugs rep-

[1] In their study, the term New Molecular Entities (NME) is used.

resenting small improvements the real prices increased substantially. These results are thus consistent with Dean's [1969] observation of skimming and penetration strategies. LC also found that the presence of branded substitutes had a clear negative effect on both introductory prices and price changes over time.

LC concluded that competitive forces play an important role in the US pharmaceutical market, and hypothesized that an introduction of price regulations may discourage penetration strategies.

The purpose of the present study is to compare the outcome in a (price) regulated market with the results in the US market by performing the same tests as in LC. To do so, we examine a data set consisting of all NCEs launched in Sweden between 1987 and 1997, and estimate the same models as in LC. In line with their results, we find that introductory prices reflect the degree of therapeutic innovation. Contrary to the results in the US market, the real prices for NCEs fall substantially on average over time for all classes of therapeutic gain, so the price cap regulation rules out the use of penetration strategies. More interestingly, we find no evidence of branded substitutes having any effect on either introductory prices or price dynamics. It therefore appears that price competition between pharmaceutical brands is indeed less pronounced under price cap regulation.

The remainder of this paper is structured as follows. We first give a brief description of the pharmaceutical market in Sweden. Then we informally discuss some of the hypotheses that the regulatory constraint provides and put these in relation to the findings in LC. We continue with a description of the estimated models and the data, then with a presentation of the results. The final section offers a concluding discussion.

2. The pharmaceutical market in Sweden

In Sweden, a substantial fraction of the pharmaceutical consumption is subsidized by the government. For pharmaceutical products that are included in the public insurance program, the state will fully reimburse individual consumption exceeding some fixed amount per annum.[2] If pharmaceutical producers wish to have their products reimbursed through this system, prices must be set by the appropriate government agency. Prior to 1993, prices were determined by the National Corporation of Swedish Pharmacies (NCSP) after negotiations with the producers. The NCSP is government-owned and monopolizes the retailing of all pharmaceuticals in Sweden. When the NCSP was in charge of the price negotiations, firms were not allowed to sell their products until the prices were agreed upon. If the NCSP and a firm were

[2]Presently, this amount is SEK 1800 (approx. USD 180). Although the co-payment schedule has changed somewhat over the time for our sample (1987-1997), the state financed part has always been large.

unable to reach an agreement about the price, the firm could apply to engage the Medical Products Agency (MPA)[3] as a mediator. In 1993, the National Social Insurance Board (NSIB) replaced the NCSP as the government agency responsible for administering the drug benefit scheme. Since that time, participation in the public insurance program has been optional, so producers are allowed to market their drugs outside the program and set prices freely. However, since consumption outside the benefit program must be financed completely out-of-pocket, producers have rarely exercised this option. The objective of the NSIB is to determine pharmaceutical prices so as to:

- ensure a consistently high level of public health;
- ensure that a sufficient range of pharmaceutical products are available at reasonable prices;
- support an efficient production of pharmaceuticals;
- encourage the research and development of new pharmaceuticals.[4]

Furthermore, when setting the launch price for a new product, the regulating agency should consider the medical merits and the health economic value of the product, the price in comparable countries,[5] and the price (and reimbursement cost) for related treatments (pharmaceutical or other). The health economic value represents the societal benefit the new drug brings in terms of savings for the total reimbursement bill (e.g., through effects on other products comprised in the benefit schedule), savings on social insurance due to reduced sick leave, health care savings in the form of a reduced number of physician visits, and also savings due to a reduced amount of care outside the direct health care sector (e.g. home care).

The manufacturer is responsible for providing the regulating agency with the relevant information about the drug.[6] This information should also contain, apart from the above, forecasts of sales in Sweden the first two years after launch, and predicted sales at a steady state level.

Once an introductory price has been set, it cannot be raised within the two first years after launch. Thereafter, applications for price increases for the whole product line, but not individual products, are considered once per year. As a general rule,

[3]The MPA is roughly the Swedish equivalent of the FDA in the US.

[4]These points are in accordance with the 1990 Transparency Directive of the European Union.

[5]Most notably other northern European countries.

[6]The regulating agency has no means of conducting investigations of its own, but must rely on the information supplied by the applicant or outside sources.

the producer may increase prices within some specified margin and is granted some leeway for individual products in the basket. For instance, price increases on certain products may be allowed if the producer lowers prices on other products in the basket.

Since 1993, the reference pricing system has been used in Sweden. This system applies when the patents for brand name pharmaceuticals have expired and there are generic substances in the market. If generic alternatives are available, the reimbursement scheme allows full coverage for products that cost no more than 1.1 times the lowest priced generic pharmaceutical. If consumers want brand drugs priced above this level, they have to pay the differential out-of-pocket.

3. Preliminary discussion

3.1. The US and Swedish markets. LC argue that demand factors explain well the determinants for NCE prices in the US, and their findings are in line with what standard economic theory predicts. In particular, they find that therapeutic advantage has a positive effect on relative launch prices since willingness to pay increases in quality. Also, pharmaceuticals that are mainly intended for chronic conditions command on average lower relative launch prices than drugs primarily intended for acute conditions. Moreover, the former display a pattern of increasing prices over time, so there is evidence that penetration pricing is being used for chronic drugs. This is expected since chronic drugs are to a large extent repeat purchase items, and the appropriate strategy for pricing such goods is by a penetration scheme: low initial prices in order to attract a consumer base whose loyalty can be exploited when the producer increases prices later.

Furthermore, LC find substantial evidence that the presence of brand name substitutes has a negative effect on relative introduction prices. This is also expected: even though brand name products are all patented they may still have similar therapeutic properties and functions. Therefore, prescribing physicians have more products to choose from in treating a given condition.

Finally, LC find that on average the presence of generic competition among the brand name substitutes has a *positive* effect on relative launch prices for NCEs. LC report conflicting evidence on the effect of generic entry on NCE prices from earlier studies, i.e. Caves et al. [1991], Frank and Salkever [1995], and Grabowski and Vernon [1992]. One explanation for the positive price effects offered both in Frank and Salkever [1995] and in LC, is that it may be more profitable for a brand name producer to focus on satisfying the inelastic segment of demand when generic drugs enter the market. In this instance, other brand name producers may follow suit and increase their prices too. In short, LC conclude that the demand factor variables they consider explain both launch prices and price dynamics in the American pharmaceutical market.

We expect some of the effects that LC describe to be present also in the Swedish data. However, given the broad regulatory policy objectives stated above, it is not obvious how the two markets compare given the same explanatory variables. For example, the price regulation in Sweden makes it difficult for firms to raise the price of one of their products without simultaneously lowering prices on other products. Consequently, we do not expect to observe penetration strategies. Instead, we are more likely to observe constant or falling real prices over the whole range of NCEs in the sample, although it is conceivable that the prices of less therapeutically advanced drugs fall less than those of more innovative drugs.

As far as launch prices are concerned, it is reasonable to expect that therapeutically advanced drugs are rewarded with premia by the regulator as compensation for high R&D outlays. Although the regulator could argue that, given that these costs are sunk, it would be rational to offer the producer marginal cost coverage only. Such behavior, however, is not in line with the regulator's objective to "encourage the research and development of new pharmaceuticals". Also, producers would be hesitant to supply the Swedish market with their products in case they were not sufficiently compensated.[7] Hence, we expect similar patterns as in LC. However, we do not expect any major differences between drugs intended for acute and chronic drugs since penetration pricing is not really an option for the producers.

LC argue that the market structure, or, more specifically, the nature of competition, is a strong candidate in explaining both launch prices and price dynamics. In particular, the existence of more branded substitutes has depressing effects on NCE launch prices. In the regulated environment, however, the market structure may not reflect price setting to the same extent so the relationship between prices and the number of substitutes is less clear-cut. On the one hand, the presence of branded substitutes provides the regulator with more bargaining power to keep prices low by exercising "yardstick regulation". On the other hand, in the interest of the public health the regulating agency has a commitment to ensure the availability of a variety of horizontally differentiated products. For any given level of demand, each product then risks obtaining a smaller market share so the regulator compensates them by allowing higher prices. Consequently, the strong connection between NCE prices and brand substitutes observed in the American data may be less pronounced in the Swedish data.

As mentioned above, there is mixed empirical evidence of the relation between

[7]It can be said that the Swedish market is very small in an international perspective so given the sunk costs, any price above marginal cost should be accepted by the producers. However, there are additional costs associated with the marketing of a new drug, e.g. training of pharmaceutical consultants, promotion campaigns, etc. Therefore, the option of not supplying the drug may be seriously weighed.

prices on new pharmaceuticals and the presence of generic competition among the branded substitutes. Here, we have no reason to believe that the presence of generics among the brand name substitutes should affect the NCE prices in any systematic way. The ambiguous effects from branded competition are expected to be present here also. Furthermore, brand name substitutes with generic competitors tend to be older drugs located further away from the NCE in product space than newer brands without generic competition. Therefore we expect a very weak link between NCE prices and generic competition in our sample.

4. Empirical analysis

4.1. Equations and variables. We now proceed with the estimation of the LC model on the Swedish data. We consider the following two models:

$$(1) \qquad LWRIS = \alpha_0 + \alpha_1 A + \alpha_2 B + \alpha_3 ACUTE + \alpha_4 LNS + \alpha_5 \left\{ \begin{array}{c} DG \\ LG \end{array} \right\} + \varepsilon,$$

and:

$$\begin{aligned} (2) \qquad LRATIO \;=\; & \beta_0 + \beta_1 A + \beta_2 B + \beta_3 ACUTE + \beta_4 DALP + \\ & \beta_5 DBLP + + \beta_6 DCLP + \beta_7 LNS + \beta_8 \left\{ \begin{array}{c} BG \\ LGS \end{array} \right\} + \delta, \end{aligned}$$

where:

- $LWRIS$ denotes the logarithm of the ratio of the launch price of an NCE to the weighted average price of existing brand name substitutes;
- $LRATIO$ is the logarithm of the ratio of the real price (CPI deflated) of a new drug four years after launch to its real introductory price;
- A and B are dummy variables that take the value 1 if an NCE is $A(B)$ classified and zero otherwise;
- $ACUTE$ indicates whether an NCE is used primarily for an acute condition;
- LNS denotes the logarithm of 1 plus the number of existing brand name substitutes at the time of launch;

- *DG* indicates whether an NCE had generic competition among its brand name substitutes at the time of launch;
- *LG* denotes the logarithm of 1 plus the percentage of substitutes that had generic rivals at the time of launch;
- *BG* indicates whether or not an NCE has a generic rival four years after launch, when there was none at launch;
- *LGS* is the logarithm of the ratio of the percentage of brand name substitutes that had a generic rival after four years to the same percentage at the time of NCE launch, (expressed as whole numbers plus 1);
- *DALP*, *DBLP*, and *DCLP* are interactive variables equaling the product of *LWRIS* and the variables indicating whether the NCE is classified as *A*, *B*, or *C*, respectively.

These models are in effect identical to the ones estimated in LC, and therefore provides a framework for comparing the results in the two different markets.

4.2. The data. The data we use consist of all NCEs approved by the Medical Products Agency (MPA) in Sweden between 1987 and 1997. In total the MPA approved 335 NCEs in this time period. Of these NCEs, we excluded 89 from the present analysis either because they were never marketed in Sweden, or because a price had not yet been set on the NCE at the time of this analysis. Our data set thus includes a total of 246 NCEs.

No official rating of therapeutic advance is available for pharmaceuticals in Sweden. As a part of this research project, a rating was carried out by pharmacologists connected to the MPA, (see Beermann and Rosén [1999]). Beermann and Rosén had access to all relevant NCE information from the MPA and collaborated with the agency to determine the rating. The purpose was to provide a rating of the therapeutic advance for all NCEs approved between 1987 and 1997, according to how the NCE was judged by the MPA at the time of approval. They classified the NCEs according to the FDA rating system that was used also in the study by LC. These classes are:

- *A*- Important therapeutic gain: Drug may provide effective therapy (by virtue of greatly increased efficacy or safety) for a disease not adequately treated or diagnosed by any marketed drug, or provide markedly improved treatment of a disease through improved efficacy or safety (including decreased abuse potential),

- *B*- Modest therapeutic gain: Drug has a modest but real advantage over other available marketed drugs; for example, somewhat greater effectiveness, decreased adverse reactions, more convenient route of administration, etc., and

- *C*- Little or no therapeutic gain: Essentially duplicates in medical importance and therapy for one or more existing drugs.

In addition to rating the new NCEs according to therapeutic advance Beermann and Rosén [1999] also classified the drugs according to type; that is, whether a drug was indicated primarily for treatment of acute illnesses or chronic conditions.

Beermann and Rosén used the WHO Anatomic Therapeutic Classification (ATC) system to define substitute drugs (ATC Index with DDDs (Defined Daily Dosage), [1997]). The ATC system is divided into 14 anatomical groups (e.g. Heart and Circulation). Each anatomical group is divided into main therapeutic groups, and then into therapeutic subgroups that are further divided into chemical/therapeutic subgroups (e.g. HMG-Co A reductase inhibitors). The chemical/therapeutic subgroups are finally divided into groups according to chemical substance. For example, a chemical substance is the cholesterol lowering drug Simvastatin that belongs to the HMG-CoA reductase inhibitors chemical/therapeutic subgroup.

We defined close substitutes as drugs in the same chemical/therapeutic subgroup that share the same indication.

This is similar to the definition used by LC, and for most of the new NCEs substitutes were defined in this way. In the event that an NCE did not have any substitute in the same chemical drug class, the substitute was defined as the drug that was most commonly prescribed for that indication in the year before the introduction of the NCE.

For Class *A* drugs it is often difficult to define substitute drugs, since many of these compounds are intended for indications not previously treated by any marketed drug (see the above definition of class *A* drugs). In the ranking of therapeutic advance according to the FDA classification scheme, Beermann and Rosén [1999] therefore divided the class *A* drugs into:

(*i*) drugs that lacked a substitute drug at the time of introduction (drugs for indications not previously treated by drugs), and

(*ii*) drugs with substitute drugs at the time of introduction (drugs providing markedly improved efficacy or safety over existing drugs).

Out of the total 246 NCEs in this study, there were a total of 8 class *A* drugs with substitutes and there were 28 class *A* drugs without substitutes. Naturally, the drugs without substitutes could not be included in the analysis of relative introductory prices since their relative prices could not be defined.

In order to estimate relative prices between NCEs and substitutes comparable dosages need to be defined. For this purpose we used the Defined Daily Doses (DDD) system recommended by the WHO for studies of drug use (ATC Index with DDDs). A DDD is defined as the average daily dose of a drug used by an adult for treatment of the main medical indication of the drug. Official DDDs were available for about two-thirds of the drugs in our sample and were used in these cases. For the remaining drugs daily dosages were based on the recommended average daily dosages in FASS (Pharmaceutical Specialities in Sweden), a widely used medical reference book that contains instructions/recommendations on the appropriate use of all drugs in Sweden. When recommendations regarding dosages were not sufficiently specific or not available in FASS, we consulted clinical pharmacologists (Beermann and Rosén) for the appropriate daily dosages. When doses were given in amount per square meter (e.g. for creams and lotions) or amount per kg body weight, the daily doses were based on one square meter and 70 kg, respectively.

Some pharmaceuticals are available in more than one form of administration. For example, the same drug can be taken as a pill or a fluid. As a rule we calculate the price for tablets when these were available. If there was no tablet, we chose the administration form of the package containing the largest number of daily dosages.

We based all prices on the official retail prices of drugs in Sweden.[8] These prices are set once a year and are reported in FASS. We estimated a price per daily dose for all NCEs and all substitute drugs. We estimated the average price for existing substitutes as the weighted average of the price of all substitutes, with the market shares in the year before the introduction of the NCE as weights. We collected data on substitute quantities from SDM (Swedish Drug Market (1987-1997), a database containing the quantities of all drugs sold on the Swedish market graciously provided by the Swedish Association of the Pharmaceutical Industry. We then adjusted all prices to 1997 prices using the official consumer price index produced by the Statistical Yearbook of Sweden [1999].

[8]In Sweden the price is independent of the buyer (which is not the case in e.g. the US).

5. Results

5.1. Summary statistics. Table 1 summarizes some statistics on relative introductory prices for 218 NCEs introduced between 1987 and 1997 (excluding the 28 A class drugs without substitutes).

Table 1: Relative Introdutory Prices

Class	Relative Introductory Price					Substitutes[a)]	
	N	Median	Mean	Std	Range	I	II
A							
Acute	5	1.43	1.59	0.56	1.10 – 2.55	1	0
Chronic	3	10.97	9.40	4.88	3.93 – 13.30	1	0
Combined	8	2.04	4.52	4.83	1.10 – 13.30	1	0
B							
Acute	30	1.64	3.75	4.96	0.17 – 19.28	4.11	0.03
Chronic	42	2.55	3.92	4.77	0.09 – 20.32	3.80	0.17
Combined	72	2.25	3.86	4.85	0.09 – 20.32	3.93	0.11
C							
Acute	63	1.26	2.16	3.09	0.04 – 16.70	4.40	0.12
Chronic	75	1.13	2.18	3.54	0.12 – 26.17	4.49	0.18
Combined	138	1.16	2.17	3.33	0.04 – 26.17	4.45	0.15

[a)]Subheadings I and II denote the average number of branded substitutes, and the fraction of substitutes with generic competition, respectively.

On average, relative launch prices increase with the degree of innovation. The average relative launch price is 2.2 for C drugs, 3.9 for B drugs and 4.5 for A drugs. The median relative launch prices are substantially lower than the mean, indicating that prices are highly skewed with a few relatively high priced NCEs. The median relative launch prices reveal patterns similar to the average prices, however. The only exception is that the median launch price decreases somewhat for A drugs compared to B drugs. This result should be interpreted with some care given that there are only eight A drugs (with substitutes) in the sample. For drugs in classes B and C, the average relative launch prices are about the same for acute and chronic drugs. For A drugs the relative launch prices are lower for acute drugs than for chronic drugs, but again this comparison is highly limited by the small number of A drugs.

The number of branded substitutes at introduction decreases with the degree of therapeutic innovation. The average number of branded substitutes at introduction is 1 for A drugs, 3.7 for B drugs and 4.5 for C drugs. The fraction of branded substitutes with generic competition at introduction of the NCE also decreases with

the degree of therapeutic innovation. For A drugs, none of the branded substitutes face generic competition whereas 15 percent of the branded substitutes for C drugs face generic competition.

In Table 2 we show some summary statistics about the inflation adjusted change in the price of NCEs four years after introduction. This comparison includes all the NCEs in our database with a follow-up time of at least four years (N=149).

Table 2: Ratios of Inflation Adjusted Prices 4 Years after Introduction

Class	Ratios of Inflation Adjusted Prices 4 Years after Introduction					Substitutes[a)]	
	N	Median	Mean	Std	Range	I	II
A							
Acute	16	0.78	0.83	0.28	0.52 – 1.65	0.31	0
Chronic	12	0.75	0.73	0.12	0.44 – 0.91	1.57	0
Combined	28	0.77	0.78	0.21	0.44 – 1.65	0.90	0
B							
Acute	20	0.93	0.90	0.33	0.34 – 2.05	2.75	0.10
Chronic	26	0.82	0.84	0.13	0.64 – 1.10	2.46	0.12
Combined	46	0.86	0.86	0.24	0.34 – 2.05	2.60	0.11
C							
Acute	42	0.89	0.89	0.29	0.40 – 2.51	4.58	0.13
Chronic	33	0.79	0.81	0.10	0.64 – 1.07	4.03	0.14
Combined	75	0.84	0.85	0.23	0.40 – 2.51	4.33	0.14

[a)]Subheadings I and II denote the average number of branded substitutes, and the fraction of substitutes with generic competition, respectively.

The average real price decreases over time for all classes of drugs: for A drugs the average real price decline is 22 percent, for B drugs it is 14 percent, and for C drugs the average real price decline is 15 percent. There is thus a tendency for a larger price decline for A drugs compared to B and C drugs. There is also a tendency for a greater decline in prices for chronic drugs compared to acute drugs. The range of price changes for the individual drugs is wide, with some drugs more than doubling in price and some decreasing to a third of the introductory price.

5.2. Regression results. Table 3 shows the regression results of the determinants of relative introductory prices (Equation 1).

Table 3: Regression Analysis of Determinants of Relative Introductory Prices

Variable	Regression Equations[a)]		
Constant	0.485*** (2.895)	0.486*** (2.886)	0.486*** (2.889)
A	0.920** (2.446)	0.920** (2.483)	0.921** (2.442)
B	0.463*** (2.988)	0.463*** (2.980)	0.462*** (2.978)
Acute	−0.197 (−1.371)	−0.197 (−1.351)	−0.200 (−1.375)
LNS	−0.124 (−1.239)	−0.124 (−1.138)	−0.119 (−1.125)
DG		0.004 (0.019)	
LG			−0.077 (−0.153)
N	218	218	218
R2	0.082	0.083	0.083
R2ADJ	0.065	0.061	0.061
F	4.794***	3.817***	3.822***

[a)]Numbers in parentheses are t-values. *Significance at the 10 % level. **Significance at the 5 % level. ***Significance at the 1 % level.

The variable for class A drugs is statistically significant at the 5 percent level and the variable for class B drugs is significant at the 1 percent level. Moreover, both have the expected signs. This strongly suggests that A and B drugs are launched at higher relative prices than C drugs. The coefficient for A drugs is also, as expected, higher than for B drugs, but this difference between A and B drugs is not significant. The variable for acute drugs is negative but not significant. The variable for the number of branded substitutes is also negative and insignificant. The extent of generic competition on the market does not show a significant effect either. The explanatory power of the regression equations for introductory prices is low with an R-square value of 8 percent.

Table 4 reports the results for the determinants of the real price change four years after introduction.

Table 4: Regression Analysis of Rates of Change of Real Prices

Variable	Determinants of Real Price Changes[a)]					
Constant	−0.202*** (−4.813)	−0.203*** (−4.811)	−0.201*** (−4.782)	−0.175*** (−3.039)	−0.174*** (−3.000)	−0.168*** (−2.900)
A	−0.138*** (−2.624)	−0.138*** (−2.625)	−0.140*** (−2.671)	−0.126 (−0.855)	−0.130 (−0.880)	−0.137 (−0.934)
B	0.032 (0.726)	0.033 (0.750)	0.031 (0.722)	0.003 (0.040)	0.004 (0.050)	0.005 (0.059)
Acute	0.069* (1.856)	0.068* (1.840)	0.064* (1.739)	0.076* (1.932)	0.077* (1.943)	0.075* (1.900)
DALP				−0.008 (−0.330)	−0.008 (−0.320)	−0.008 (−0.314)
DBLP				−0.001 (−0.039)	−0.001 (−0.081)	−0.003 (−0.196)
CDLP				−0.009 (−0.762)	−0.009 (−0.761)	−0.010 (−0.871)
LNS	−0.020 (−0.872)	−0.016 (−0.621)	−0.011 (−0.476)	−0.008 (−0.090)	−0.014 (−0.539)	−0.010 (−0.395)
BG		−0.023 (−0.356)			−0.032 (−0.468)	
LGS			−0.280 (−1.175)			−0.313 (−1.280)
N	128	128	128	149	149	149
R2	0.093	0.096	0.104	0.094	0.095	0.104
R2ADJ	0.064	0.059	0.067	0.049	0.043	0.053
F	3.567***	2.862**	3.138***	2.089**	1.845*	2.041**

[a)]Numbers in parentheses are t-values. *Significance at the 10 % level. **Significance at the 5 % level. ***Significance at the 1 % level.

We estimated the equations both with and without the interaction terms between relative introduction prices and the drug classification. The reason for this is that we may include the A class drugs that had no substitute drug at the time of introduction.

The variable for class A drugs is negative and highly significant, whereas the variable for class B drugs is positive but insignificant. This indicates that prices for A drugs decline faster over time than prices for B and C drugs, but that there is no significant difference between B and C drugs. The variable for acute drugs has a positive coefficient, but fails to be significant at the 5 percent level. The variables for generic competition and brand name competition have no significant effect on the change in prices over time. Likewise, the interaction variables between drug classification and relative introductory price are insignificant. The explanatory power of the regression equations is low with an R-square value of about 10 percent.

6. Concluding remarks

The main purpose of this paper was to compare NCE pricing outcomes in a regulated market (Sweden) with those reported from a market without price regulation (US). In accordance with LC we found that introductory prices depend on the degree of therapeutic innovation. The patterns were similar in both studies, but the average relative introductory prices were higher for all classes of drugs in the regulated market.

The inflation-adjusted price change over time differed substantially between the studies. In the present analysis, real prices declined over time for all classes of drugs. This result is in contrast to the substantial price increase that was reported for C drugs in the US.

Compared to the US market, the Swedish market is thus characterized by higher relative launch prices and falling real prices over time. This pattern of high relative launch prices and declining real prices over time is in line with what one might expect from the price cap regulation in Sweden. Although penetration pricing was not observed for any of the classes of drugs, there was a small tendency for the prices of class A drugs to fall more than the prices of class B and class C drugs. This was the only similarity with the US market in pricing patterns over time.

Contrary to the results in LC there was also a tendency for real prices of acute drugs to fall less over time than those of chronic drugs. LC also found that branded substitutes had a substantial negative effect on both introductory prices and price changes over time, and that the extent of generic competition tended to increase the relative launch prices. We found no significant effect of the number of substitutes, branded or generic, neither on the relative launch prices nor on the price changes over time.

We note that the NCE samples used in the Swedish and American studies are taken in different time periods. The samples are roughly of the same length but the

present sample starts some ten years after the one used in LC. This discrepancy may have some effect on the validity of the comparison.

The differences in price behavior between the US and the Swedish markets suggest that the competitive forces at work in the two regimes are different. Pricing decisions seem to play a major role in the US, whereas it is difficult to see from the Swedish data that prices are as important strategic variables in the regulated market. Clearly, price caps limit the possibilities to use certain pricing strategies which then naturally do not appear in the Swedish data. It generally appears that some of the pricing constraints imposed by the regulator offset some disciplinary effects that market forces bring. The high variance in introductory prices and the lack of price effects from branded competition in the Swedish market are examples of this.

References

[1] Abbott III, T.A. [1995] "Price Regulation in the Pharmaceutical Industry: Prescription or Placebo?", *Journal of Health Economics* 14, 551-565.

[2] Besley, T., and M. Gouveia [1994] "Alternative Systems of Health Care Provision", *Economic Policy* 9(19), 199-258.

[3] Beermann, B., and A. Rosén [1999] "Rating innovative therapeutic benefits of medicines licensed in Sweden 1987 - 1997", *International Journal of Pharmaceutical Medicine* 13, 123-126.

[4] Caves, R. E., M. D. Whinston, and M. A. Hurwitz [1991] "Patent Expiration, Entry, and Competition in the US Pharmaceutical Industry", *Brookings Papers on Economic Activity: Microeconomics,* 3-65.

[5] Danzon, P. M. [1997] *"Pharmaceutical Price Regulation: National Policies versus Global Interests"*, The AEI Press, Washington, D.C.

[6] Danzon, P. M., and L.- W. Chao [2000] "Cross-National Price Differences for Pharmaceuticals: How Large, and Why?", *Journal of Health Economics* 19, 151-165.

[7] Dean, J. [1969] "Pricing of Pioneering Products", *Journal of Industrial Economics* 17, 165-179.

[8] *FASS (Pharmaceutical Specialities in Sweden).* Läkemedelsinformation AB, Elanders, Kungsbacka, 1988 – 1998.

[9] Frank, R. E., and D.S. Salkever [1997] "Generic Entry and the Pricing of Pharmaceuticals", *Journal of Economics and Management Strategy* 6 (1), 75-90.

[10] Grabowski, H. G., and J. M. Vernon [1992] "Brand Loyalty, Entry, and Price Competition in Pharmaceuticals after the 1984 Drug Act", *Journal of Law and Economics* 35, 331-350.

[11] Laffont, J.-J., and J. Tirole [1993] *"A Theory of Incentives in Procurement and Regulation"*, MIT Press, Cambridge and London.

[12] Lu, Z. J., and W. S. Comanor [1998] "Strategic Pricing of New Pharmaceuticals", *Review of Economics and Statistics* 80(1), 108-118.

[13] Reekie, W.D. [1978] "Price and Quality Competition in the United States Drug Industry", *Journal of Industrial Economics* 26, 223-237.

[14] Reekie, W.D. [1981] "Price and Quality Competition in the Drug Market: Evidence from the US and the Netherlands", in Helms, R. B. (ed.), *"Drugs and Health"*, Washington, DC: Enterprice Institute, 123-39.

[15] *Statistical Yearbook of Sweden* 1999, Statistics Sweden, Stockholm.

[16] *SDM, (Swedish Drug Market)*, Statistics compiled by The Swedish Association of the Pharmaceutical Industry, 1987 – 1997.

[17] Schweitzer, S.O. [1997] *"Pharmaceutical Economics and Policy"*, Oxford University Press, New York.

[18] WHO Collaborating Centre for Drug Statistics Methodology, *"ATC Index With DDD's 1997"*, Oslo.

Essay II: Innovativeness and market shares in the pharmaceutical industry

Innovativeness and market shares in the pharmaceutical industry

MATS EKELUND AND BJÖRN PERSSON*
STOCKHOLM SCHOOL OF ECONOMICS

ABSTRACT. We analyze the pharmaceutical market with a model of horizontal and vertical product differentiation. The implications from the model are tested on data from the Swedish pharmaceutical market. We find that vertically differentiated drugs gain larger market shares, command higher prices, and are less sensitive to substitutes than drugs that are only horizontally differentiated.

1. INTRODUCTION

Pharmaceutical producers are bringing more and more sophisticated drugs to the market, contributing to the quality of life of millions of patients. This, however, is occurring at rising cost, as reflected in the shares of national GDPs devoted to pharmaceutical expenditures. In most countries, there is extensive government involvement in public reimbursement schemes reflecting a concern that pharmaceutical products be efficiently provided to citizens. The associated price regulations stemming from these schemes are an example of how decision makers try to control the rising costs caused by increasing consumption of medical drugs. Meanwhile, there is a general awareness of the trade-off between controlling costs and providing proper incentives for efficient consumption of drugs on the one hand, and the need to reward medical R&D on the other. Given the welfare gains at stake, it is worthwhile to explore how incentives for R&D are provided on the pharmaceutical market. In this paper, we ask how the introduction of new innovative drugs is rewarded through market shares and prices in the Swedish market.

Earlier research on the economics of the pharmaceutical market has focused mainly on three topics: behavior at generic entry after patent expiration, pricing policies and price sensitivity of pharmaceuticals, and the effectiveness of promotion outlays. In the first category the literature has addressed questions concerning various barriers to entry, price competition, and factors that determine entry. Examples in this category are Grabowski and Vernon [1992], Frank and Salkever [1997] on pricing and barriers to entry, and Scott Morton [1999] on factors that determine generic entry. Reekie [1978] and Lu and Comanor [1998] analyzed pricing strategies as a function of product

*We thank Magnus Johannesson for his advice and comments. We have also benefited from comments and discussions with Tore Ellingsen, Ulf Gerdtham, and Bengt Jönsson.

innovativeness. Ellison et al. [1997] studied the price sensitivity of generic and therapeutic substitutes, and Mortimer [1997] considered the influence of insurance on prescription patterns. The effectiveness and importance of advertising and promotion was analyzed by Berndt et al. [1995], and earlier by Hurwitz and Caves [1988].

Recent work has offered new perspectives on the pharmaceutical market. For instance, Berndt et al. [1999] investigated the existence of network effects in the antiulcer market, and Stern and Trajtenberg [1998] used micro data to study physicians' prescription behavior. Lichtenberg [1998] studied creative destruction in the pharmaceutical market, measuring the effect of new drug introductions on the sales of old drugs. Lichtenberg's results suggest that the most innovative drugs increase aggregate sales one-to-one, while less innovative new drugs decrease the sales of old drugs more than they increase aggregate sales.

Unlike earlier work that was limited to generic pharmaceuticals or particular submarkets, this paper considers the direct interplay between innovativeness, prices, and market shares for new innovative pharmaceuticals in many submarkets. We hope to shed some light on the most important variables which determine how a new innovative pharmaceutical is rewarded in the market. We focus on the differentiation aspect of the good, a factor which we believe is fundamental for understanding how new innovative drugs capture market shares. More specifically, we assume that new innovative drugs are differentiated from each other both in vertical and horizontal dimensions. Along the vertical dimension, pharmaceuticals differ in their general medical efficacy, whereas the horizontal dimension measures how the drugs cater to idiosyncratic tastes or needs. We provide an analysis of this in a model and test its implications on Swedish quarterly data. We consider New Chemical Entities (NCEs) in 55 therapeutic categories with observations ranging over a twelve-year period. Our main result is that the more innovative and vertically differentiated drugs capture larger market shares and higher prices compared with the less innovative drugs that are only horizontally differentiated. The market shares of the vertically differentiated drugs also seem less sensitive to the number of substitutes. We also conjecture that more innovative drugs capture larger market shares when consumers are insensitive to prices. If prices are regulated in such an environment, relative prices must be set differently in order to maintain the same incentives as in the unregulated setting.

The paper is organized as follows. After a discussion of the pharmaceutical market, we present an address model of product differentiation in the spirit of Hotelling. We proceed with a description of the data and an empirical analysis of some of the implications we derive from the model. Finally, we conclude by summarizing the main results.

2. The pharmaceutical market

It is sometimes argued that while there is nothing fundamentally unique about the pharmaceutical market, many features present in other industries are combined and exaggerated.

At the very heart of the industry is a heavy focus on research and development. Bringing a new therapeutic drug to the market is a lengthy and costly process: the average time period between the scientific conception and the marketing of a new drug is reported to be about ten years in the US, and the associated average cost is nearly 300 million USD.[1] The primary reasons are the stochastic nature of new drug development and the rigorous safety regulations imposed on the industry. For every new pharmaceutical brought to the market, tens of thousands of molecular structures have been screened. An overwhelming majority of the candidate compounds will fail and be discarded on the way, leaving only a tiny subset for further testing first on animals and then on humans.[2] Once the required testing is complete, and the developing firm has a marketable end product, the manufacturer must obtain approval from the regulatory authorities to start selling the drug. If the drug is approved the firm usually initiates an intense marketing to boost demand. The pharmaceutical industry spends, on average, about the same share of its revenues on marketing as it does on research (Comanor and Schweitzer [1995]). This figure is matched only in certain markets for sophisticated electronic components. Hence, the industry is extreme also in this sense.

The demand side is also atypical. The final consumer (the patient) does not usually choose the specific product, or even treatment. Instead, a better informed agent (the physician) makes the choice for the consumer. In addition, the physician may be constrained in her choice of products due to third party payment and other regulations. The high degree of insurer involvement is particular to the pharmaceutical market, and to the health industry in general.

Pharmaceuticals are fairly specialized products, and individuals respond idiosyncratically to different drugs. Therefore individual consumers may not agree on the ranking of a given pharmaceutical in relation to others, although there is consensus over the general efficacy or therapeutic benefit of the drug. The individual reaction to a certain brand (e.g. with regard to the nature and intensity of its side effects) is often hard to predict ex ante. For this reason, pharmaceuticals are, to a high degree, experience goods whose benefits must be judged after consumption.

It is natural to assume that the process leading to a good match between an in-

[1] *The Economist* [1998], Schweitzer [1997].

[2] On average, only 1 out of 10 000 molecules that enter the initial screening process is actually developed into a drug. (*The Economist* [1998])

dividual consumer and a drug is perceived as a cost by the patient. Individuals are adverse to exposing themselves to unwelcome side effects that may result from consuming new pharmaceuticals. The consumer (and in some sense also the prescribing physician) is thought of as making a "consumption investment" in a given drug which will increase the differentiation relative to the ex ante.

Therefore we suggest that if an individual has found a good or even acceptable match, he will be reluctant to try a new drug even if the new drug is expected to be slightly better. In other words, an individual who switches between brands will incur an additional cost.

As mentioned above, we have chosen to concentrate on the product differentiation aspect of pharmaceutical drugs. We contend that therapeutic drugs are well suited for the traditional story of horizontal and vertical product differentiation. This literature stems from the seminal contributions of Launhardt and Hotelling and has been elaborated by numerous others. The basic idea is that pharmaceuticals provide some degree of therapeutic advantage or general medical efficacy, which constitutes the vertical component of the products. All individuals (patients and/or physicians) agree on the ranking of the products in this respect. That is, all individuals prefer a higher medical efficacy product to a lower efficacy product, all else being equal, implying that they would choose the former over the latter. In addition, there is a horizontal dimension representing the individuals' idiosyncratic tastes for products. In our framework we assume that various pharmaceuticals may be ranked differently by individual consumers, even when they have the same medical potency to treat a certain condition. The reason is that some individuals may experience more severe side effects from certain pharmaceuticals than other individuals, or simply that they prefer a certain profile of side effects over another. This horizontal component therefore suggests that they would not necessarily prefer the same drug if prices were equal.

The market for anti-depressants is a pharmaceutical submarket with both horizontal and vertical product differentiation, and here serves as an example. In this market, there are two predominant types of pharmaceuticals, the tricyclic drugs and the Serotonin inhibitors, SSRIs. The tricyclic drugs belong to an older class of anti-depressants. These drugs increase the production of neurotransmitters of which low levels are associated with depression. The SSRIs belong to a newer generation of drugs. Instead of increasing the production of some neurotransmitters, these inhibit the uptake of a certain neurotransmitter, serotonin.

It has not been scientifically ascertained that the SSRIs are on average more effective in providing relief or treating depression than the older tricyclics. The main contribution of the SSRIs is the reduced probability of adverse side effects and gen-

erally less severe side effects. The first of the SSRI drugs (Prozac®) was therefore not classified as an *A* drug (important therapeutic gain) according to the FDA classification system, but as a *B* drug (modest therapeutic gain). Due to the reduced probabilities of side-effects, the SSRIs are vertically differentiated from the tricyclics. There is also inter-brand horizontal differentiation (within the SSRI group and within the tricyclics group), and intra-brand differentiation (between the two groups). For example, Prozac® is horizontally differentiated from the other SSRIs in that it has the longest half-life. This characteristic is attractive for individuals who sometimes forget to take their medication, but it is a disadvantage to those who occasionally experience adverse reactions. SSRIs are horizontally differentiated from tricyclics in that they produce certain side effects (such as gastrointestinal disease) with higher probability than the tricyclics, but there is a (much) lower probability for a wide range of other symptoms. For this reason, some patients may have a nonpecuniary reason to prefer a tricyclic drug to an SSRI drug, even though SSRIs are generally perceived as more advanced and innovative.[3]

The Swedish pharmaceutical market. In the Swedish pharmaceutical market consumers are largely subsidized, and prices are determined in a regulated environment. Until 1993, pharmaceutical prices were determined in negotiations between producers and the National Corporation of Swedish Pharmacies (NCSP). The NCSP is a state-owned company that monopolizes the retailing of all drugs in Sweden. In 1993 a government agency, the National Social Insurance Board (NSIB), assumed the role of the NCSP in the negotiations with the pharmaceutical industry. The regulatory rules also changed slightly, so that from 1993 a producer need only negotiate a price with the NSIB if it wants consumers to be reimbursed for its products within the national pharmaceutical benefit scheme.[4] Prices are set annually, but both producers and the agency may initiate re-negotiations in cases where there are wide discrepancies between the anticipated and the actual performance of their products. In the present sample there are few cases of drastic price changes, however. In most cases real prices decline slightly over time. We think of the outcome as the result of a one-shot bargaining process between two equally potent parties. Both parties consider the price levels in comparable countries and seem reluctant to deviate from the norm in the European Union, not least because of the potential effects of parallel import. Once the initial price is agreed upon, the price path is normally corrected only so that the price falls just below the CPI, unless the drug performs very unexpectedly.

[3] An excellent review on the medical properties of anti-depressants is provided in the paper "Pharmaceutical Innovations and Market Dynamics: Tracking Effects on Price Indexes for Anti-depressant Drugs" by Berndt et al. [1996].

[4] A producer may opt not to do this and set prices independently.

Pharmaceutical consumption in Sweden is financed in three different ways. The cost of prescription drugs is shared between the NSIB and the patient. The patient pays a decreasing fraction of her consumption out-of-pocket up to a certain level, above which pharmaceuticals are entirely covered by the NSIB via federal taxes. This scheme is re-initiated annually so that once the consumers have reached the level where co-payment stops, they pay nothing for additional pharmaceutical consumption for the remainder of that year.[5] Pharmaceuticals that are sold to and consumed in hospitals are financed by the regional authorities, which organize and provide most of the health care in Sweden. Finally, prescription-free, over-the-counter pharmaceuticals are paid for by consumers alone.

Physicians in general have little direct incentive to prescribe the most cost-effective pharmaceuticals. However, they may have indirect incentives to take account of their prescription costs. For example, some hospitals have pharmaceutical boards supervising the prescriptions and the use of drugs. These boards try to influence physicians to make more cost-effective decisions. Physicians may also be concerned about the costs borne by the patient as a result of the prescription, or about the general costs for taxpayers. Whether physicians' incentives to prescribe cost-effective drugs are strong enough to have an important impact on prescription patterns is still largely an open question. (For a study on the prescription behavior of Swedish physicians, see Lundin [2000]).

3. The theoretical framework

3.1. Assumptions. In order to examine the effect that innovativeness has on market shares we adopt a simple address model of product differentiation. In this class of models, stemming from the classic work of Hotelling [1929], it is assumed that there are J products located in some characteristics space W and that consumers have preferences defined on this set. The utility for an individual $i \in I$ is then a function $u_i(\mathbf{w}, \mathbf{p})$ of his address in relation to the products and the prices $\mathbf{p} \in \mathbb{R}^J$ charged by the suppliers. The individuals are distributed over the set W according to some distribution Φ. The consumers are often assumed to have inelastic demand for exactly one product, so a strategy for a consumer i is a decision vector $\mathbf{d}_i : W \times \mathbb{R}^J \to \{0,1\}^J$ with $\sum_{j \in J} d_{ij} = 1$. All firms[6] $j \in J$ are assumed to maximize profits $\pi_j(\mathbf{w}, \mathbf{p})$ by choosing location and price. A firm strategy would then be a pair (w_j, p_j) in some prescribed order. A Nash equilibrium for this game, if it exists, will then give rise to an expression for the market share for all firms $j \in J$,

$$M_j = M_j(\mathbf{w}^*, \mathbf{p}^*, \Phi(\mathbf{w}^*), J) \ , \tag{1}$$

[5]This level is SEK 1800, or approximately USD 180.

[6]Each firm produces exactly one product, so j denotes both product and producer here.

which will be in the focus of our attention in the empirical analysis.

We will here consider a simplified version of a setup used by Economides [1993]. In his model, firms play a multi-stage game by choosing whether or not to enter and then choosing location, quality, and price. Here we simply focus on the price game, taking all other variables as given. That is, we consider J firms exogenously located on $W = H \times V$ where H is a circle of unit circumference, and where $V = \{0, v\}$ with $v > 0$. A firm location is then a pair $(h, v) \in W$ which has been determined outside the game. The first component, $h \in H$, represents the *variety* and the second component, $v \in V$, represents product *quality*. We think of H as the horizontal dimension over which individuals have idiosyncratic preferences, and V denotes the vertical dimension over which the individuals have the same preference ranking. In our framework a pharmaceutical product is then identified by these two dimensions, where the variety caters to individual specific needs, and quality is simply the overall medical potency or efficacy of the drug. That is, at given prices, more quality is preferred to less by all individuals, but some individuals may prefer one certain variety over others. The firms maximize profits by choosing a price $p_j \in \mathbb{R}_+, \forall j \in J$ given a vector $\mathbf{w}$ of locations. We assume that all firms have identical marginal costs, normalized to zero. The consumers are distributed on H rectangularly so Φ is simply the Lebesgue measure on the unit interval. The utility function for an individual i, located at x_i, purchasing from supplier j, located at $w_j = (h_j, v_j)$ is given by

$$u_i(w_j, p_j, \alpha, t) = v_j - t|h_j - x_i| - \alpha p_j \,, \tag{2}$$

where t is a transport cost and where α is a parameter measuring consumer price sensitivity. The transport cost is, as usual, interpreted as the utility loss an individual incurs by consuming a non-ideal product, and it is the same across all individuals. The parameter α measures the weight the consumer puts on prices in their choice of pharmaceutical. We first consider the case where consumers attach some importance to prices, that is when $\alpha > 0$. For reasons of comparison we also consider the case where $\alpha = 0$, that is when the price has no influence whatsoever on consumer choice. This is intended to reflect the case where pharmaceutical expenditures are completely tax financed and when prices are set by a regulator. In this setting the consumers only choose an optimal action, whereas the firms take no action. We begin by describing the equilibrium of the price game.

3.2. Equilibrium. In line with Economides we restrict attention to localized competition strategies which leave each firm with positive demand. A firm j then faces demand $D_j(\mathbf{p}, \mathbf{w}, J)$:

$$D_j = \frac{\alpha(p_{j+1} + p_{j-1} - 2p_j) + 2v_j - v_{j+1} - v_{j-1} + t(h_{j+1} - h_{j-1})}{2t} \,, \tag{3}$$

where z_j is the location of a consumer indifferent between suppliers j and $j+1$. Each firm $j \in J$ then solves the maximization program:

$$\max_{p_j} \left[\pi_j \left(\mathbf{w}, \mathbf{p}\right) = p_j D_j \left(\mathbf{p}, \mathbf{w}, J\right)\right] \ .$$

This gives rise to the J equilibrium equations

$$\mathbf{A}\mathbf{p}^* = \mathbf{y} + \mathbf{B}\mathbf{v} \ , \tag{4}$$

where

$$\begin{aligned} \mathbf{A}_j &= (0, .., 0, -1/4, 1, -1/4, 0, .., 0) \\ y_j \left(\mathbf{h}\right) &= t \left(x_{j+1} - x_{j-1}\right) / 4\alpha \\ \mathbf{B}_j &= (0, .., 0, -1/4\alpha, 1/2\alpha, -1/4\alpha, 0, .., 0) \end{aligned} \tag{5}$$

are the jth rows of $\mathbf{A}$, $\mathbf{y}$, and $\mathbf{B}$, respectively. As Economides points out, the matrix $\mathbf{A}$ is circulant[7] and may be inverted as long as its column sum is different from zero. The equilibrium price is then given by

$$\mathbf{p}^* = \mathbf{A}^{-1} \left[\mathbf{y} + \mathbf{B}\mathbf{v}\right] \ . \tag{6}$$

If we make the simplifying assumptions that (1): the firms are placed equidistantly on H and (2): that there is only one producer of quality $v > 0$, say j, we have that

$$\begin{aligned} p_j^* &= \tfrac{t}{2J\alpha} \textstyle\sum_{j \in J} a_{jk}^{-1} + \tfrac{a_{jd}^{-1}}{2\alpha} v \\ p_{j\pm1}^* &= \tfrac{t}{2J\alpha} \textstyle\sum_{k \in J} a_{j+1k}^{-1} - \tfrac{a_{j+1d}^{-1}}{4\alpha} v \\ p_{j\pm n}^* &= \tfrac{t}{2J\alpha} \textstyle\sum_{k \in J} a_{j+nk}^{-1}, \text{ for } n > 1 \ . \end{aligned} \tag{7}$$

where a_{jk}^{-1} is the kth element in sum of the jth row in $\mathbf{A}^{-1}$, and where a_d^{-1} is the diagonal element in $\mathbf{A}^{-1}$. The assumption of equidistantly located producers provides a clearer link between the idea of horizontal product differentiation and the number of substitutes in the market. This simplification enables us to let the number of firms J serve as a measure of variety competition on the circle H. In the empirical analysis the number of substitutes will serve as a variable for the degree of horizontal differentiation.

[7] A matrix is circulant if each row is equal to the preceding upper row moved one step to the right.

Since the matrix $\mathbf{A}$ is circulant, $\mathbf{A}^{-1}$ has the properties that the row sum is the same across rows and that the diagonal elements are the same. Let $\beta = \sum_{k \in J} a_{jk}^{-1}$ for all j and let γ be the main diagonal entry of $\mathbf{A}^{-1}$. The equilibrium prices are then given by:

$$\begin{aligned} p_j^* &= \tfrac{1}{2\alpha}\left(t\beta/J + \gamma v\right) \\ p_{j\pm 1}^* &= \tfrac{1}{2\alpha}\left(t\beta/J - \gamma v/2\right) \\ p_{j\pm n}^* &= \tfrac{1}{2\alpha} t\beta/J, \text{ for } n > 1 \ . \end{aligned} \tag{8}$$

From these expressions we see how location affects optimal prices for the different firms. The high quality producer j sets the highest equilibrium price, as expected. Also, prices are decreasing in the number of firms J. Furthermore, the effects of the (localized) competition pushes down the equilibrium price for a low quality producer that is located next to a high quality producer. This is so since the lower quality producers that are placed immediately next to the high quality producer do not have the same variety advantage as competitors farther away on the circle. The scope for variety competition is more limited and therefore firm $j's$ immediate neighbors must price more aggressively than other lower quality suppliers.

The equilibrium market shares are given by:

$$\begin{aligned} M_j &= \tfrac{1}{J} + \tfrac{2v(4-3\gamma(J))}{8t} \\ M_{j\pm 1} &= \tfrac{1}{J} + \tfrac{4v(\gamma(J)-1)}{8t} \\ M_{j\pm 2} &= \tfrac{1}{J} - \tfrac{v\gamma(J)}{8t} \ , \text{ for } n > 1 \ , \\ M_{j\pm n} &= \tfrac{1}{J}, \text{ for } n > 2 \ , \end{aligned} \tag{9}$$

where we must impose that $v < 8t/J\gamma$, so the quality premium cannot be too large.[8] Since the rows in $\mathbf{A}$ sum to less than unity, γ is decreasing in J (and approaching 1 from above). The condition that γ be less than $6/5$ is also met for all J, implying that the high quality producer will enjoy the largest market share of all producers.

Interestingly, the firms with the smallest market shares are not the immediate neighbors of a high quality producer, but instead the neighbors $j+2$ and $j-2$, one slot farther away. The producers closest to the high quality supplier must set a more competitive price than the other lower quality producers. Hence, the neighboring producer closest to him in turn, $j+2$, will yield some market share for this reason.

[8] This is to ensure that no firm gets zero demand in the restricted set of equilibria that we consider.

In fact, due to the more intense price competition between firms j and $j+1$, the immediate neighbors of a high quality producer will gain a larger market share than all other firms (except the high quality supplier). Also, we note that a high quality producer loses less market share relative to a lower quality producer when the number of firms increases. This would imply that the high quality producer is less sensitive to competitors than lower quality producers are. If we compare the profits accruing to the different producers we find that

$$\pi_j(\mathbf{p}^*) \geq \pi_{j+n}(\mathbf{p}^*) \geq \pi_{j+1}(\mathbf{p}^*) \geq \pi_{j+2}(\mathbf{p}^*) \ , \tag{10}$$

indicating that the high quality producer enjoys the highest profit, whereas the least preferred slot is the one next to the immediate neighbor of the high quality producer.[9]

3.3. Case where $\alpha = 0$. In the case where consumers are indifferent to prices, the only relevant variable is product location. Here we assume that firms take no action and simply offer their products given their locations in W. The consumers now choose the pharmaceutical that maximizes gross benefit given their location. If we assume rectangular Φ on H as above, the following market shares will result:

$$\begin{aligned} M_j &= \tfrac{1}{J} + \tfrac{v}{t} \\ M_{j\pm 1} &= \tfrac{1}{J} - \tfrac{v}{2t} \\ M_{j\pm n} &= \tfrac{1}{J} \ , \text{ for } n > 1 \ . \end{aligned} \tag{11}$$

The distribution of market shares is slightly different than in the former case. As before, the high quality producer j will enjoy the largest market share but this share is larger than when firms set prices optimally. Also, producer $j's$ larger market share comes at the expense of the immediate neighbors', $j+1$ and $j-1$, market shares. All other $J-3$ producers get an equal market share of $1/J$. Clearly, the absence of price as a strategic variable will cause some consumers closer to $j+1$ and $j-1$ to now prefer product j, which they did not prefer when the net benefit was the decision criterion. Since they cannot respond by lowering prices, producers $j+1$ and $j-1$ will lose market shares compared to the above. The firms $j+2$ and $j-2$ are not affected by the competitive pricing of firms $j+1$ and $j-1$, respectively, and therefore now enjoy the same market shares as the other lower quality producers have. Of course, we cannot determine which of the two "regimes" $\alpha > 0$ or $\alpha = 0$, is preferred by the different producers since we had assumed an exogenous regulatory price.

[9] The middle inequality is valid as long as $t \leq \frac{\beta-(\gamma-1)\left(v^2J^2\gamma-2\beta vJ\right)+2vJ}{4\beta}$, which is likely to hold, especially for large J.

3.4. Remark. In view of the admittedly highly stylized model above, we expect certain features to be present in the subsequent empirical analysis. Firstly, we expect that innovative drugs will command higher prices on the average and gain larger market shares than do the lower quality producers. Also, the idea that most low quality producers will be relatively more sensitive to an increased number of competitors than will the high quality producer, is consistent with the theory.

This outcome is expected to be more pronounced in the case where consumers are fully reimbursed for their pharmaceutical expenditures.

4. Empirical analysis

Data. The data consist of quantities sold and price quotes for a set of NCEs introduced in Sweden between 1987 and 1997. Each NCE has been labeled with an indicator B and C representing the therapeutic advantage that the NCE offers over existing substitutes.

Quantities. We have defined quantity as the number of daily dosages sold during a given period of time. We use the quantities for daily dosage in accordance to the Defined Daily Doses (DDD) system recommended by the WHO for studies of drug use (ATC Index with DDDs, [1997]). A Defined Daily Dose denotes the average daily dose of a drug used by an adult for the treatment of the main medical indication of the drug. Official DDDs were available for about two-thirds of the drugs in our NCE sample, and we used them in these cases. For the remaining drugs, we based quantities for daily dosages on the recommended average daily dosages in FASS (Pharmaceutical Specialities in Sweden 1988-1998), a reference book widely used by physicians in Sweden. When recommendations regarding dosages were not sufficiently specific or not available in FASS, we consulted pharmacologists for the appropriate quantities. In some cases however, it was not possible to establish an average daily dose since the use of the drug is too varied.

NSIB provided the data on quantities sold (measured as DDD per substance, whenever an official DDD was available) over the period 1996-1999:2. In order to obtain the relevant data for the period 1987-1995, however, we estimated the price per DDD for each NCE and approximated the number of daily dosages sold per quarter with the ratio between the sales revenue and the estimated price per daily dose. The prices of pharmaceuticals sold in Sweden are reported in FASS and the prices for pharmaceuticals sold by the NSIB are the same regardless of purchaser.[10] As a rule, we have chosen to calculate the price for tablets when these were available; if they were not, we chose the administration form of the package containing the largest

[10]We do not face the problems caused by multi-tariffs or price discrimination common in the US that make it difficult to determine the price of a pharmaceutical.

number of daily dosages. NSIB provided us with data on the sales of substances 1987-1999:2 defined as revenues per quarter. It should be noted that of the B and C drugs introduced on the Swedish market between 1987-1997, we were not able to analyze all empirically. In order to include an NCE drug in the analysis we needed data on prices per DDD or sales in DDDs for all substitutes. Since the data on sales and quantities were on the ATC code (Anatomic Therapeutic Classification, see below) level, we had to exclude all markets where at least one substitute drug had generic competition. However, there are three markets where we were able to obtain prices and quantities for generic competitors for some drugs in 1998 and 1999. For this reason there are three drugs with observations only in the second and third year. (This is the reason why the number of observations reported in tables 2 and 3 is larger for the second year than for the first year).

Definition of markets. We have defined the therapeutic markets using the World Health Organization (WHO) Anatomic Therapeutic Classification (ATC) system. The ATC system is divided into 14 anatomical groups (e.g. Heart and Circulation). Each anatomical group is divided into main therapeutic groups, and then into therapeutic subgroups that are further divided into chemical/therapeutic subgroups (e.g. HMG-Co A reductase inhibitors). The chemical/therapeutic subgroups are further divided into groups according to chemical substance. An example of a chemical substance is the cholesterol-lowering Simvastatin that belongs to the HMG-CoA reductase inhibitors chemical/therapeutic subgroup.

We define a competitor or substitute to an NCE as a pharmaceutical within the same chemical therapeutic subgroup sharing the main indication. In some cases there are no competitors in the same therapeutic group as the NCE at the time of market entry. In some of these cases, pharmacologists have indicated certain pharmaceuticals outside the chemical therapeutic subgroup which may be considered as substitutes. We define the market share of an NCE as the number of daily doses sold during a given period of time divided by the total number of daily doses sold of the NCE and its competitors combined.

There are several caveats with this definition of a market. First, pharmaceuticals outside the market as we have defined it, may be prescribed to certain patients for the same indication for which the NCE is intended. Therefore, this definition will underestimate the actual market. Second, drugs that are within the same chemical therapeutic subgroup and that share the same indication may be prescribed to some patients with other indications. This will overestimate the actual market size in our definition. For these reasons, the market shares as we have calculated them are necessarily approximations of the true market shares.

Therapeutic classes. No official rating on therapeutic gain for pharmaceuticals is available in Sweden. We have therefore used the rating system constructed by pharmacologists (see Beermann and Rosén [1999]). Beermann and Rosén [1999] classified all NCEs introduced in Sweden between 1987 and 1997 according to the FDA rating system used in the US. This system contains three different classes of therapeutic advantage.

- *Class A:* Important therapeutic gain: Drug may provide effective therapy (by virtue of greatly increased efficacy or safety) for a disease not adequately treated or diagnosed by any marketed drug, or provide markedly improved treatment of a disease through improved efficacy or safety (including decreased abuse potential).
- *Class B*: Modest therapeutic gain: Drug has a modest but real advantage over other available marketed drugs; for example, somewhat greater effectiveness, decreased adverse reactions, more convenient route of administration, etc.
- *Class C*: Little or no therapeutic gain: Essentially duplicates in medical importance and therapy for one or more already existing drugs.

We have excluded class A drugs in our empirical study since they rarely have well-defined substitutes.

Variables. The following variables are included in the subsequent analysis.

- M_t is the market share of an NCE obtained at time t where $t \in \{1, 2, 3\}$ indicates the first, second, and the third years after launch.
- B indicates whether the NCE is a class B drug.
- C indicates whether the NCE is a class C drug.
- SUB_t is the number of substitutes at period t. If a substitute has been in the market for less than four of the quarters during the period, its contribution to the number of substitutes is calculated as the number of quarters divided by four.
- $PRATIO$ is the price of the NCE divided by the weighted average price of the number of substitutes at the time of introduction.

Econometric model. In our model pharmaceuticals are differentiated by price, the vertical quality dimension indicated by the classification of the drug (B or C), and the presence of a horizontal dimension approximated by the number of therapeutic substitutes in the market.

To see whether the location in product space will influence the relative price of a new drug, we first regress the price ratio on the B indicator variable, the number of substitutes and the number of substitutes multiplied with a dummy for a B classified drug respectively. The latter variable is included in order to detect systematic differences between B and C drugs with respect to how they are affected by the number of substitutes. The regression equation is

$$(1) \qquad \log PRATIO = \alpha_0 + \alpha_1 B + \alpha_2 \log SUB_1 + \alpha_3 B \log SUB_1 + \varepsilon \ .$$

We then regress the variable for the market shares on the number of substitutes and the indicator variable for a B drug. Since prices are determined by the location in product space in our model, we exclude this variable in the second regression. We expect that the price and market shares are determined by the same factors. The price sensitivity of physicians may also be low due to a lack of pecuniary incentives.[11] The second regression equation therefore provides information on how the new pharmaceutical's exogenous location in product space determines market success. We expect a vertically differentiated (B) drug to be more successful in terms of gained market share than a C drug. This difference is influenced by the degree of horizontal differentiation. If the vertical quality is valued more by consumers than the horizontal aspect we should expect B classified drugs to take large market shares irrespective of the number of substitutes. The pharmaceuticals in class C, on the other hand, would be more sensitive to horizontal differentiation.

To test the degree of product differentiation we regress the market shares for the first four quarters for each $t \in \{1, 2, 3\}$ combined on the number of substitutes and the dummy for B classified drugs.

$$(2) \qquad \log M_t = \beta_0 + \beta_1 B + \beta_2 \log SUB_t + \upsilon_t \ .$$

In order to detect if there are any systematic differences between B drugs and C drugs with respect to how they are affected by the number of substitutes, we also regress the equation above with an added variable for the B dummy multiplied with the number of substitutes.

$$(3) \qquad \log M_t = \delta_0 + \delta_1 B + \delta_2 \log SUB_t + \delta_3 B \log SUB_t + \xi_t \ .$$

[11] The price ratio is far from significant at any conventional levels when it is included in the regression.

5. Results

When we run regression (1) we find that the B indicator variable is positively correlated with the log of the price ratio ($\log PRATIO$), and that the variable for the number of substitutes ($\log SUB$) is negatively correlated with the log of the price ratio. The results are summarized in table 1 below.

Table 1: Regression 1: Price ratio and product differentiation.

	logPRATIO[a)]
Constant	0.430** (2.068)
B	0.503* (1.701)
logSUB1	−0.373* (−1.697)
BlogSUB1	−0.009 (0.002)
N	83[b)]
F-test	3.910***
R2	0.129
R2adj	0.096
Breusch-Pagan	6.619*
RESET	0.656

[a)]Results in the regression are corrected for heteroscedasticity. Numbers in parentheses are the t-values of the coefficients. The RESET-test is conducted with the predicted value raised to two, three and four. *Significance at the 10% level, ** significance at the 5% level and *** significance at the 1% level. [b)]There are two fewer observations in the price regression (1) compared to the market shares regression (2). This is because two drugs lacked well-defined relative prices per DDD, but their market shares were zero.

Both the B indicator variable and the interaction variable for the number of substitutes are significant at the 10 percent level. The coefficient of the interaction variable for the number of substitutes of B drugs is quite insignificant and indicates that there is no difference between the way the prices of B drugs and C drugs are affected by the number of substitutes. Even though the degree of significance is low, the results are consistent with the hypothesis that product differentiation has an influence on prices. A vertically differentiated product is expected to command a higher price when it enters the market, and we also expect that a larger number

of substitutes will affect prices adversely. We cannot tell whether the correlations between differentiation and prices are due to a market pricing strategy by the firms, whether the differentiation has affected the bargaining position of the firms vis-à-vis the regulating authorities, or both. The fact that the price variable is far from significant when it is included in regression (3) may indicate that prices play a minor role in determining the market success of the drug. If this is the case, the price is not likely to be an important variable in the market strategy but rather a result of the bargaining game between the firm and the regulating authorities. A higher quality and a smaller number of substitutes may strengthen the threat of the firm's withholding a drug from the market.

Table 2: Regression 2: Market shares, therapeutical rating, and number of substitutes.

	logM1[a)]	logM2	logM3
Constant	−2.605***	−1.315***	−1.587***
	(−8.947)	(−3.743)	(−6.125)
B	1.392***	2.051***	1.10***
	(3.410)	(3.716)	(3.087)
logSUB	−1.780***	−2.017***	−1.420***
	(−4.113)	(−4.118)	(−3.217)
N	85	88	78
F-test	17.94***	18.30***	10.72***
R2	0.304	0.301	0.222
R2adj	0.287	0.285	0.202
Breusch-Pagan	9.04**	6.14**	24.15***
RESET	2.62	1.36	1.40

[a)]Results in the regressions are corrected for heteroscedasticity. Numbers in parentheses are the t-values of the coefficients. The RESET-test is conducted with the predicted value raised to two, three and four. *Significance at the 10% level, ** significance at the 5% level and *** significance at the 1% level.

In the second set of equations, where we regress the log of the market share ($\log M$) on the number of substitutes and the B indicator variable for the first three periods, we find that all variables are significant at the 1 percent level or lower. The B indicator variable has the expected sign, and vertically differentiated class B pharmaceuticals seem on average to gain larger market shares than the class C pharmaceuticals. This is further supported by the descriptive statistics presented in

Table 3: Regression 3: Testing the difference in effect from the number of substitutes on the market share of B drugs and C drugs, respectively.

	logM1[a)]	logM2	logM3
Constant	−2.309*** (−6.781)	−1.04** (−2.597)	−1.311*** (−4.316)
B	0.714 (1.532)	1.241* (1.892)	0.239 (0.570)
logSUB	−2.262*** (−4.023)	−2.408*** (−3.984)	−1.802*** (−3.214)
BlogSUB	1.557* (1.983)	1.508 (1.563)	1.441** (2.002)
N	85	88	78
F-test	13.30***	12.91***	8.15***
R2	0.330	0.316	0.248
R2adj	0.305	0.291	0.218
Breusch-Pagan	9.185**	6.854*	29.439***
RESET	1.304	0.982	0.680

[a)]Results in the regressions are corrected for heteroscedasticity. Numbers in parentheses are the t-values of the coefficients. The RESET-test is conducted with the predicted value raised to two, three and four. *Significance at the 10% level, ** significance at the 5% level and *** significance at the 1% level.

table 4.

The first year in the market, the average market share of a B drug is 33 percent compared with 10 percent for a C drug. The corresponding figures for the second and third years are 44 percent and 46 percent for class B pharmaceuticals, and 18 percent and 22 percent for class C pharmaceuticals, respectively. The number of substitutes is slightly smaller for B drugs than for C drugs, 1.6, 1.8, and 1.9 compared with 2.1, 2.3, and 2.3, for the first, second, and third years, respectively. The variable for the number of substitutes is negative, as expected, and lends support to the hypothesis that a larger degree of horizontal differentiation obstructs market penetration. The intercept is larger in the regression equations for the second and third years compared to the first year. The average market shares increase between the first and second years and seem to stabilize in the third year. In the last period both the B indicator variable and the number of substitutes have a smaller impact on the variable for average market share. This is indicated in the smaller coefficients of the B and

Table 4: Average market shares and number of substitutes.

	Year 1[a)]	Year 2	Year 3
Market share			
B drugs	0.33 (0.31)	0.44 (0.31)	0.46 (0.34)
C drugs	0.10 (0.16)	0.18 (0.22)	0.22 (0.25)
Number of substitutes			
B drugs	1.6	1.8	1.9
C drugs	2.1	2.3	2.3

[a)]Numbers in parentheses are standard deviations for market shares.

$\log SUB$ coefficients. Also, the R-square value decreases from 0.30 in the second regression to 0.22 in the third regression. A possible explanation for this is that consumer learning about the new drug and the producers' marketing efforts become more important over time relative the initial product space location.

Finally we add the variable with the B dummy multiplied with the number of substitutes in order to determine whether there are any systematic differences in how B and C drugs are affected by the number of substitutes. For all three periods, the coefficient is of nearly the same magnitude as that of the coefficient of the $\log SUB$ variable, but with the opposite sign. In the first period the coefficient of the interaction variable is very close to being significant at the 5 percent level. During the second period it falls short of being significant at the 10 percent level, and for the third period it is significant at the 5 percent level. It thus seems that there are differences in the way the market share of a B drug is influenced by the number of substitutes compared with a C drug, and that a class B drug seems to be much less affected.

The data is afflicted with heteroscedasticity. For this reason t-values are estimated based on White's heteroscedasticity consistent covariance matrix whenever the Breusch-Pagan test is significant.

6. Discussion

In the preceding analysis we focus on pharmaceuticals' location in product space as the major determinant of pricing and market success. We assume that each therapeutic market consists of a unit circumference circle where each consumer demands one variety. We interpret the number of substitutes as a measure of competition

in the sense that more substitutes improve the match between a consumer and a certain pharmaceutical. This implies that it is more difficult for a new brand to establish itself the larger the number of substitutes. This, however, presupposes that the potential scope for differentiation (or the possibility to locate farther away from the competitors) is the same across markets. This is clearly a simplification. Some markets have a larger potential for differentiation than others, and therefore a larger number of substitutes. The strength of the local monopoly enjoyed by each variety naturally depends on the number of potential niches. Certain illnesses affect homogeneous groups of patients while others affect more heterogeneous parts of the population. For example, while prostate cancer is a disease afflicting elderly men only, depression may strike a person of any age and of either sex. It might be the case that elderly men react more homogeneously to side effects than the more diverse part of the population that suffers from depression. Therefore, the potential to gain substantial market power may be lower for products intended for the former condition, given the same number of substitutes. If there were a way to control for this heterogeneity, we would get a clearer picture of how the number of substitutes affects the degree of competition.

Another caveat concerns the assumption of exogenous product location. Even though it is difficult for a firm to determine exactly where a product will be located, it is nevertheless true that some markets may be more prone to differentiation than others. If the scope for product differentiation increases the opportunities to make profits within a market, which is likely, this will increase the R&D efforts to locate on these markets. Product differentiation may also be endogenous in another and more subtle way. Although some drugs are very similar with respect to efficacy and side effects, their attributes may be emphasized ex post in the marketing of the drug. Consumers may therefore perceive differences between pharmaceuticals that are founded in factors beyond the physical product. If product differentiation is endogenous in some of these ways, the number of substitutes will not correspond to the actual competition in therapeutic markets. However, given the complexities of the pharmaceutical market we believe that the simplifying assumption of exogenous product differentiation is necessary.

7. Summary and policy implications

The location in product space seems to influence both prices and the market success enjoyed by new innovative pharmaceuticals. Vertically differentiated class B drugs are priced higher and capture a larger market share than C classified drugs, which are only horizontally differentiated. There is some evidence that the number of substitutes existing in the market affects the price of a drug relative to the existing substitutes for both B and C drugs. A plausible explanation is that the bargaining position of

the firm vis-à-vis the regulating authorities, and thereby the price level, is influenced by the perceived innovativeness of the drug. However, the price relative to existing substitutes does not seem to have any important effect on the captured market share. This might lend support to the hypothesis that the prescribing physicians in Sweden are rather insensitive to prices. We find that the degree of horizontal differentiation interpreted in terms of the number of existing substitutes adversely influences the market share captured by a new pharmaceutical. However, this effect seems to be more pronounced for C drugs than for B drugs.

Our results indicate that the more innovative B classified drugs are better rewarded than the less innovative C drugs. Further, there is no evidence that horizontally differentiated C drugs have any important effect on the market shares that the more innovative B drugs capture.

The empirical findings in this paper basically conform with the predictions of our model of product differentiation, suggesting that this approach might be valuable for further research in this area.

Two cases of our model with product differentiation were discussed: One with price sensitive consumers and price as a strategic variable, and one without price sensitive consumers. In the latter case the more innovative drugs tended to capture larger market shares. This observation should be taken into consideration by regulators who set prices administratively. In a market with price-insensitive consumers, innovative drugs can have lower prices but still be as profitable as if they were priced freely in a market with price-sensitive consumers. Less innovative drugs lose out the most in markets with regulated prices and price insensitive consumers.

Expenditures may also be affected by the price sensitivity of consumers. Since more innovative drugs take larger market shares in the absence of price sensitivity and usually receive higher prices than less innovative drugs, expenditures in price-insensitive markets may be more strongly affected by new innovations.

References

[1] Andersson G.F., and J.P. Poullier [1999] "Health Spending, Access, and Outcomes: Trends in Industrialized Countries", *Health Affairs* 18 179-192.

[2] Berndt, E.R., Pindyck, R.S., and P. Azoulay [1999] "Network Effects and Diffusion in Pharmaceutical Markets: Antiulcer Drugs", *NBER Working Paper* 7024.

[3] Berndt, E.R., I.M. Cockburn, and Z. Griliches [1996] "Pharmaceutical Innovations and Market Dynamics: Tracking Effects on Price Indexes for Antidepressant Drugs", *Brookings Papers on Economic Activity Microeconomics*, 133-188.

[4] Berndt, E.R., L. Bui, D.R. Reily, and G.L. Urban [1995] "Information, Marketing, and Pricing in the U.S. Antiulcer Drug Market", *AEA Papers and Proceedings* 85 (2), 100-105.

[5] Comanor, W.S., and S.O. Schweitzer, [1995] "The Pharmaceutical Industry", in Adams, W., and Brock, J.W. (eds.) *The structure of American Industry* 9th edition. Englewood Cliffs, N.J., Prentice Hall.

[6] Dos Santos Ferreira, R., and J.-F. Thisse [1996] "Horizontal and Vertical Differentiation: The Launhardt Model", *International Journal of Industrial Organization* 14, 485-506.

[7] *The Economist* February 21, 1998.

[8] Ellison, S.F., I. Cockburn, Z. Griliches, and J. Hausman [1997] "Characteristics of Demand for Pharmaceutical Products: an Examination of Four Cephalosporins", *RAND Journal of Economics* 28 (3), 426-446.

[9] *FASS*, 1988 – 1998, (Pharmaceutical Specialities in Sweden), Apoteksbolaget.

[10] Frank, R.G., and D.S Salkever [1997] "Generic Entry and the Pricing of Pharmaceuticals", *Journal of Economics & Management Strategy* 6 (1), 75-90.

[11] Grabowski, H.G., and J.M. Vernon [1992] "Brand Loyalty, Entry, and Price Competition in Pharmaceuticals After the 1984 Drug Act", *Journal of Law and Economics* XXXV, 331-333.

[12] Hotelling, H. [1929] "Stability in Competition", *Economic Journal* 39, 41-57.

[13] Hurwitz, M.A., and R.E Caves [1988] "Persuasion or Information? Promotion and the Shares of Brand Name and Generic Pharmaceuticals", *Journal of Law and Economics* XXXI, 299-320.

[14] Lichtenberg, F.R. [1998] "*Pharmaceutical Innovation as a Process of Creative Destruction*", Columbia University and NBER.

[15] Lu, J.Z., and W.S. Comanor [1998] "Strategic Pricing of New Pharmaceuticals", *Review of Economics and Statistics* 80(1), 108-118.

[16] Lundin, D. [2000] "Moral Hazard in Physician Prescription Behavior", *Journal of Health Economics* 19, 639-662.

[17] Mortimer, O.M. [1997] "*Demand for Prescription Drugs: The Effects of Managed care Pharmacy Benefits*", Ph.D. Thesis, University of California, Berkely.

[18] Norman, G., and J.-F. Thisse (eds) [1994] *The Economics of Product Differentiation, vols I and II,* Elgar Reference Collection, London.

[19] Reekie, D. [1978] "Price and Quality Competition in the United States Drug Industry", *Journal of Industrial Economics* 26 (3), 223-237.

[20] Rosén, A., and B. Beermann [1999] "Rating Innovative Therapeutic Benefits of Medicines Licensed in Sweden 1987-1997", *International Jounal of Pharmaceutical Medicine* 13, 123-126.

[21] Salop, S. [1979] "Monopolistic Competition with Outside Goods", *Bell Journal of Economics* 10, 141-156.

[22] Scott Morton, F.M. [1999] "Entry Decisions in the Generic Pharmaceutical Industry", *RAND Journal of Economics* 30 (3), 421-440.

[23] Schweitzer, S.O. [1997] *Pharmaceutical Economics and Policy,* Oxford University Press, New York.

[24] Stern, S., and M. Trajtenberg [1998] "Empirical Implications of Physician Authority in Pharmaceutical Decisionmaking", *NBER Working Paper* 6851.

[25] WHO [1997] Collaborating Centre for Drug Statistics Methodology "*ATC Index with DDDs*", Oslo, Norway.

Essay III: A simple model of strategic experimentation

A simple model of strategic experimentation

BJÖRN PERSSON*
STOCKHOLM SCHOOL OF ECONOMICS

ABSTRACT. This paper offers a simple framework for studying the strategic interaction between two agents who can observe each others' actions. There is uncertainty over the payoff distribution for a certain action, which is common to both agents. The agents can gain information either by experimenting and taking the uncertain action themselves, or by observing the other agent's outcome from taking the uncertain action.

The main finding is that although the combined expected payoff is maximized in the pure subgame perfect equilibria, this outcome is unlikely to arise due to conflicting preferences with respect to the equilibria. Strong incentives to free ride on the other agent's experimentation activities are therefore likely to lead to the worst possible outcome.

1. INTRODUCTION

In cases where individuals must make decisions with uncertain consequences, they can gain information through experimentation and/or through observations of others in similar situations. The widespread existence of scenarios in which learning from others seems to take place is reflected in a large body of literature in the social sciences. A main issue in this line of research concerns the socially efficient level of experimentation. In many instances it is natural to assume that there is a cost associated with acquiring payoff-relevant information. Given that uncertainty relates to a parameter that is common to all individuals, there might be scope for strategic interaction in information gathering. In the event that outcomes from experimentation are publicly observable, problems associated with free riding arise. If individuals may acquire information from others' experimentation at no cost, we might suspect that outcomes will prevail in which less experimentation takes place than would be socially optimal.

The purpose of this paper is to study a learning problem in which the individuals make repeated decisions, and the outcomes from their decisions are publicly observable. It attempts to do this in the simplest possible setup. There are two individuals who both live for two periods only, and there are two possible actions to choose from in each period. The only uncertainty present is that neither individual knows with certainty the payoff distribution of one of the actions.

*I am grateful to Magnus Johannesson and Jörgen W. Weibull for advice and suggestions.

The fact that the decision makers need to take action in two consecutive periods adds a feature that has not often been considered in the literature on social learning. Most of the literature considers myopic individuals and the learning dynamics pertaining to society as a whole, rather than the individual agents. The herd or information cascade literature, initiated independently by Banerjee [1992] and Bikhchandani, Hirshleifer, and Welch [1992], deals with one-shot decisions, and with how agents may rationally discard their private signal in lieu of a more powerful public signal. The usual result entails action convergence, and some of the subsequent literature has discussed the robustness of these results, e.g. Smith and Sorensen [2000].

In the framework of this paper, the individuals must consider also how their behavior in the first period influences their expected payoff in the second period. As a concrete example, we let the decision makers be health care providers who choose between two treatments to administer to their patients. Although the general set-up of the problem has a wide range of applications, this particular interpretation comes very naturally to mind. The situation where a doctor must choose repeatedly between treatments that have unknown characteristics is also often used as a motivating example in the literature on statistical learning, specifically in bandit problems. Choosing a treatment whose characteristics are uncertain involves a gamble, but it also provides information that can be used subsequently to treat patients more effectively in the future. The example in this paper also constitutes a bandit problem, albeit with a two-period horizon only.

Bandit problems have been studied since the 1930s, and there is by now a very substantial statistics literature on this topic. The basic multi-armed bandit problem can be described as follows. A decision maker has to select one of many statistically independent reward processes at all decision epochs. The chosen process yields an immediate reward when sampled, and then changes state. The processes that were not selected do not change state and yield no reward. By sampling from the different processes the decision maker generates a sequence of immediate rewards, and the objective is to find the process that generates the highest expected discounted present value. The fundamental characteristic of a bandit problem, then, is how immediate gain is contrasted against future gain. Selecting a certain alternative may yield a superior immediate payoff whereas selecting another alternative (and thereby sacrificing some immediate payoff) may prove more beneficial in the long run. Translated to our problem, there is a conflict between treating the patients arriving at an early stage and treating those who arrive at a later stage when more is known about the uncertain treatments. The fact that the maximization problem cannot be intertemporally separated is what complicates the bandit problem.

A well-known result in bandit theory is that optimal learning may not always imply perfect learning. If the decision maker becomes sufficiently pessimistic about

the payoff potential from a certain alternative (after a series of poor outcomes) he may rationally abandon this alternative forever, even if it might in fact be the optimal process. One of the first papers using bandit theory in economics, Rotschild [1974], made precisely this point. He considered a monopolist choosing prices over an infinite horizon and showed that the rational policy entails charging a suboptimal price with positive probability.

Given the obvious generality of bandit problems these have, however, found rather few other applications in the economics literature.[1] An exception is in the literature on job search, see e.g., Jovanovic [1979] and Viscusi [1979].

The literature on multi-player bandits is scarcer still, and we know of only two papers in the economics literature. Aoyagi [1998] considers the issue of action convergence in a two-armed bandit setup. More specifically, he addresses an earlier conjecture made by Rotschild [1974] in relation to the results found in the more recent literature on common knowledge. Rotschild suggested that when there are two players, they may end up settling on different arms forever even if they may observe each other's actions (but not outcomes). The reason would be that both players think that the other has had a series of bad experiences with the other arm. This reasoning conflicts with the notion that it is impossible for the players to agree to disagree about which arm is superior. Aoyagi finds, under certain weak assumptions about the support of the payoff distributions, that in any Nash equilibrium the players will settle on the same arm with probability one.

Aoyagi focuses exclusively on the topic of action convergence, in the limit of an infinite time horizon, and the issue of whether agents learn the true values of the parameters which govern the reward processes. He does not consider explicit strategies or their relation to issues concerning social efficiency.

A paper that is concerned with social efficiency is Bolton and Harris [1999], and the present formulation is close to theirs in spirit. They consider a one-armed bandit problem in continuous time extended to a game where the players can observe each others' outcomes, not only each others' actions as in Aoyagi [1998]. They identify and characterize the set of stationary Markov perfect equilibria, and how these relate to the number of players. In particular, they find a unique symmetric equilibrium where the players devote any given time to both a safe and a risky action. Bolton and Harris do not consider pure equilibria in their model, except for noting that no such symmetric equilibria exist. They find that the equilibrium level of experimentation is less than efficient due to incentives to free ride, but they also identify an encouragement effect which offsets the free riding effect somewhat. If an individual

[1] "[a bandit problem] embodies in essential form a conflict evident in all human action", Whittle [1982, p 210]. (Also cited in Berry and Fristedt, [1985, p 5]).

expects others to experiment in the future, she can be stimulated to increase her own experimentation if this reduces the time until the information resulting from others' experimentation becomes available. Hence, in their setup there is a flavor of both intertemporal substitutability and intertemporal complementarity that work in opposite directions. They conclude that the former effect typically dominates the latter, so the equilibrium outcomes are generically inefficient.

The present paper considers a particularly simple set up, focussing on pure strategies which we solve for explicitly. In particular, we contrast the scenario where the decision makers can observe the outcomes of each others' actions with the scenario where they cannot. We find that the pure *equilibria* in the two different scenarios are the same in terms of learning and efficiency. However, we argue that the actual *outcome* in the scenario with mutual observability is likely to be an out-of-equilibrium outcome due to coordination failures. The coordination failures that arise due to conflicting preferences over the outcomes are detrimental to learning and efficiency. In this framework, the outcomes from the experimentation activities are equally useful for both, regardless of the decision maker produces it. Both agents then prefer to let the other decision maker acquire the information, since experimenting has a cost in terms of foregone expected payoff. The ensuing free riding effects are potentially strong enough to completely thwart incentives to learn about the uncertain action, even in an environment in which the decision makers would benefit from this knowledge at a future stage.

The remainder of the paper is devoted to a thorough description of a basic model where both decision makers can each take one observation of the unknown action. It then offers a slight extension to a setting where the decision makers each have the possibility to make more than one observation of the unknown action. After a brief discussion of the results, a final section concludes.

2. The basic model

2.1. The problem. Consider the following simple two-period formulation. There are two health care providers (e.g. physicians) who are visited by one patient each in both periods. The physicians must choose a health care technology (a treatment, e.g. a pharmaceutical drug) to administer to the patients. We suppose that there are two such technologies available, and they both produce a binary stochastic outcome of either success or failure. A successful treatment results in a positive payoff to the physician, whereas the physician gets zero payoff if the treatment fails. The problem is that the physicians do not know with certainty which of the two treatments is superior.[2] For simplicity, we assume that one of the technologies has a known success

[2]The patients are assumed to be homogeneous, so if a certain health care technology proves better for one patient, then it is better for all patients (in probabilistic terms). This (unrealistic)

propensity. The other technology, however, has an unknown success propensity and the physicians share a common prior belief over its success probability. In the first period, the physicians simultaneously choose one of the two treatments to prescribe to their respective patients. Then they observe the outcome from the treatment, receive the payoff, and update their beliefs about the unknown treatment according to Bayes' theorem. Armed with reassessed beliefs, they enter into the last period and make another simultaneous prescription choice for a new set of patients. Finally, the outcomes in the second period are observed, and payoffs are collected.

In this problem, it is natural to consider the first period as a potential experimental stage where the physicians decide whether or not to make an information investment. Since the characteristics of one of the treatments are not completely known, the first period might serve as an information gathering stage where the physicians improve their knowledge of the unknown treatment's performance. The knowledge so obtained is then used to treat the patient in period two as effectively as possible. This interpretation is more natural when there is a large set of patients rather than just one or two in each period. In this instance we may think of the first stage as a clinical trial and the second period as the actual treatment phase.

As mentioned above, we find it useful to distinguish between two alternative formulations for comparative reasons. In the first (A), we assume that the outcomes of the treatments are private information held only by the treating physician. In the second (B), we assume that the outcomes of the treatments are publicly available. Consequently, in scenario A each physician observes only how their own patient reacts to the chosen treatment. The only means for the physicians to acquire information about the unknown treatment is therefore to prescribe it and record the outcome. (Prescribing the known treatment will add no knowledge about the unknown treatment's characteristics since the treatments are independent of each other.) In scenario B, however, each physician may, in addition to her own prescription outcomes, observe the outcomes resulting from the other physician's prescriptions. This assumption clearly introduces an additional possibility to obtain knowledge about the unknown treatment's efficacy. In particular, it now becomes possible to acquire information about the unknown treatment even without trying it. Since the physicians may learn from observing each other, we could expect that incentives to prescribe the unknown treatment are different in the two scenarios. This is the focus of our analysis.

2.2. A game theoretic description. This section presents the problem described above in standard game theoretic terms. The two scenarios A and B constitute two dynamic games of complete information, and we refer to them as model A

assumption of independence between patients and treatments is made for simplicity.

and model B, respectively. The two models are identical in every respect (except for the information the players may hold at the end of each of the subgames), and we use the following notation for both A and B. Formally, let $a_{it} \in \mathcal{A} = \{1,2\}$ denote the treatment that physician $i \in I = \{1,2\}$ prescribes in period $t \in T = \{1,2\}$. When physician i administers a treatment in period t an outcome $Z_{it} \in \{1,0\}$ is produced, representing success and failure, respectively. The sequence of outcomes observed by the physicians then form independent Bernoulli distributions with success frequencies given by the parameter $\theta = (\theta_1, \theta_2) = \Theta \subseteq (0,1)^2$. We let μ be a probability measure on θ representing the prior beliefs the physicians have over the success rates. We assume that treatment 2 is known, so both physicians believe that $\theta_2 = y \in [0,1]$ with $\mu-$ probability one. It is useful to let $X_t \in \{1,0\}$ denote the outcome from treatment 1 in period t, and let $Y_t = y$ denote the outcome from treatment 2. Hence, we can summarize the observation in t as $Z_t \in \{1,0,y\}$, and this is also the payoff the physicians receive in period t. We let $\mathcal{H}_t^A$ and $\mathcal{H}_t^B$ denote the set of possible histories observable (immediately) after stage t and $\mathcal{H}^{A(B)} = \mathcal{H}_1^{A(B)} \cup \mathcal{H}_2^{A(B)}$ is the set of all possible observations during the course of play. The null history is denoted by $\mathcal{H}_0^{A(B)}$. A (pure) strategy for player $i \in I$ in A (B) is a function $s_i : \mathcal{H}^{A(B)} \rightarrow \mathcal{A}$, namely a rule that assigns a treatment to any given observed history of events. We denote by S_i the set of pure strategies available to player i, and $S = S_1 \times S_2$ is the set of strategy profiles. Let Λ (with generic element λ) be the product set of the parameter support set Θ and the set of sample histories $\mathcal{H}$, that is $\Lambda = \Theta \times \mathcal{H}$. The prior distribution μ over θ, along with a strategy profile $s \in S$, will induce a probability distribution on the space Λ, and we denote this distribution by μ_s. Physician i attempts to maximize the expected discounted sum of successful observations, that is, the expression:

$$V_i(s_i, s_j) = \int_\Lambda [Z_{i1} + \delta_i Z_{i2}] \, \mu_s(d\lambda) = \mathbb{E}_{\mu_s}[Z_{i1} + \delta_i Z_{i2}] \ , \tag{1}$$

by choosing a pure strategy $s_i \in S_i$, and where $\delta_i \in [0,1]$ is a discount factor. In order to make the analysis as transparent as possible, we will assume that the prior belief μ over the unknown treatment's success propensity is rectangularly distributed on the unit interval. This choice is purely for analytical convenience. Since we make a special point of obtaining explicit expressions this is a convenient choice of distribution, although it may not be the most realistic case. The flat distribution is a special case of the beta distribution, which is a natural choice when modelling

beliefs over proportions.[3] This is partly due to the fact that it has the convenient property of being closed under repeated sampling, (see e.g., DeGroot [1970]). We will restrict our attention to the case where both physicians have the same discount factor $\delta_i = \delta_j = \delta$.

We now proceed by characterizing the subgame perfect equilibria in the two models A and B.

2.3. Version A. As mentioned above, a strategy for player i is a sequence of treatment choices in periods one and two, $s_i = (a_{i1}, a_{i2})$. A pure strategy equilibrium is a profile $s^* = (s_1^*, s_2^*)$ where s_i^* maximizes $V(s_i, s_j)$ for $i, j \in I$ and $i \neq j$. In this game all actions and outcomes are private information, so the optimal strategy of physician i is independent of the strategy used by physician j. Therefore, model A corresponds to a particularly simple version of the classical one-armed bandit problem. We apply backward induction to find a pure subgame perfect equilibrium profile s^* (if one exists). In the last period, physician $i \in I$ will choose the action $a_{i2} \in \{1, 2\}$ that maximizes her expected immediate payoff, $\mathbb{E}_{\mu_s}[Z_{i2}]$. Her optimal action a_{i2}^* is therefore given by:

$$a_{i2}^* = \begin{cases} 1 \text{ if } \mathbb{E}\left[X_2 \mid \mathcal{H}_1^A\right] > y \\ \\ 2 \text{ otherwise}^5 \end{cases},$$

where $\mathcal{H}_1^A$ denotes the information generated by the choice in the first stage. Due to the dynamic structure of the problem, the decision maker's choice of treatment in period one will have an impact on the decision situation also in period two. By Bellman's principle, optimality prescribes that the best initial action be taken given that the decision maker continues optimally. With a slight abuse of notation, we let $V(a_{i1}, a_{i2}^*) \in \{V(1), V(2)\}$ denote the expected value to physician i when she prescribes treatment $a_{i1} \in \{1, 2\}$ in stage one, and then continues optimally in stage two. An optimal strategy for physician i must then solve:

$$\sup_{a_i \in \{1,2\}} \left[V(a_{i1}, a_{i2}^*)\right] .$$

[3] A random variable X is distributed beta with parameters a and b, (with $a, b > 0$), if X has an absolutely continuous distribution whose density $f(x)$ is:

$$f(x \mid a, b) = \begin{cases} = \frac{x^{a-1}(1-x)^{b-1}}{\int_{[0,1]} x^{a-1}(1-x)^{b-1} dx} \text{ if } 0 < x < 1 , \\ 0 \text{ otherwise} . \end{cases}$$

[7] We assume that the physicians choose the known treatment in period two whenever there is a tie.

The expressions for the value functions $V(a_{i1}, a_{i2}^*)$ are:

$$V(a_{i1}, a_{i2}^*) = \begin{cases} \mathbb{E}\left[X_1 + \delta \max\left\{\mathbb{E}\left[X_2 \mid \mathcal{H}_1^A\right], y\right\} \mid \mu\right] \text{ if } a_{i1} = 1 \\ y + \delta \mathbb{E}\left[\max\left\{\mathbb{E}\left[X_2 \mid \mathcal{H}_1^A\right], y\right\} \mid \mu\right] \text{ if } a_{i1} = 2 \end{cases}, \tag{2}$$

Hence, the solution to physician $i's$ problem is given by $s_i^* = (a_{i1}^*, a_{i2}^*)$ as specified above. Physician $j \neq i$ solves an identical maximization program, and combining the two gives the subgame perfect equilibrium profile $s^* \in (s_1^*, s_2^*)$. In our case where μ is Lebesgue measure we have that:

$$V(1) = \begin{cases} \frac{1+2\delta y}{2} \text{ if } y \geq \frac{2}{3} \\ \frac{3+\delta(2+3y)}{6} \text{ if } \frac{1}{3} \leq y < \frac{2}{3} \\ \frac{(1+\delta)}{2} \text{ if } y < \frac{1}{3} \end{cases}, \tag{3}$$

and

$$V(2) = \begin{cases} (1+\delta)\, y \text{ if } y \geq \frac{1}{2} \\ \frac{2y+\delta}{2} \text{ if } y < \frac{1}{2} \end{cases}. \tag{4}$$

Physician $j \neq i$ faces an identical problem, so the symmetric subgame perfect Nash equilibrium is given by:

$$s_i^* = \begin{cases} (2,2) \text{ if } y \geq \frac{2\delta+3}{3(\delta+2)} \\ (1,2) \text{ if } \frac{1}{3} \leq y < \frac{2\delta+3}{3(\delta+2)} \text{ and } X_{i1} = 0 \\ (1,1) \text{ if } \frac{1}{3} \leq y < \frac{2\delta+3}{3(\delta+2)} \text{ and } X_{i1} = 1, \text{ or if } y < \frac{1}{3} \end{cases}, \tag{5}$$

for $i, j \in I$.

The results are standard and the trade-off between immediate and future payoffs is clearly illustrated here. Let y^c denote the cutoff value for y where the physicians are indifferent (disregarding the tie) between the two treatments in the first stage, given that they continue optimally in period two. Using the notation above, $y^c = \{y \in (0,1) : V(1) = V(2)\}$. In case the physicians are myopic, we have that $y^c = E[X \mid \mu] = 1/2$, that is, the myopic rule is to prescribe the known treatment if $y \geq$

1/2. The optimal choice for any positive discount factor δ, however, is to prescribe the unknown treatment in the first stage if $y < (2\delta + 3)/(3\delta + 6)$, which is larger than 1/2 for all $\delta > 0$. That is, the greater the weight the physicians put on the second stage, the stronger the incentives are to prescribe the unknown treatment (as measured by y). As δ grows without bound, the cutoff value y^c approaches 2/3 indicating that it is indeed optimal to sacrifice some immediate payoff in hope of obtaining a larger future payoff. In fact, it is always better to take the unknown treatment than the known treatment, given that they have the same expected success rate. This result is also a characteristic feature of bandit problems, as discussed in e.g., Berry and Fristedt [1985].

Another point worth noting here is that it is never optimal to choose the uncertain treatment in stage two, unless it was chosen in the first stage. This is a consequence of (a version of) an optimal stopping theorem applied to this problem: if it ever becomes optimal to choose the safe option, it must also be optimal to stick with the safe option forever after, since no new information can arise from playing it. Clearly, this does no longer have to be true if the physicians can observe each other.

We say that *experimentation* occurs whenever $y^c > 1/2$; that is when optimal behavior prescribes that the unknown treatment be administered in stage one, even though its current expected success rate is lower than that of the known treatment. Before turning to version B of our problem, we summarize the main result in the following proposition.

Proposition 1. *Let* $Y^\delta = \left[\frac{1}{2}, \frac{2\delta+3}{3\delta+6}\right)$. *If* $y \in Y^\delta$, *the unique subgame perfect equilibrium is for both physicians* $i \in I$ *to experiment.*

2.4. Version B. We now consider the formulation where the physicians can observe each other's treatment choices and the associated outcomes. As above, the problem for physician i is to maximize the expression $V(s_i, s_j)$ where:

$$V(s_i, s_j) = \begin{cases} \mathbb{E}\left[X_1 + \delta \max\left\{\mathbb{E}\left[X_2 \mid \mathcal{H}_1^B\right], y\right\} \mid \mu\right] & \text{if } a_{i1} = 1 \\ y + \delta \mathbb{E}\left(\max\left\{\mathbb{E}\left[X_2 \mid \mathcal{H}_1^B\right], y\right\} \mid \mu\right) & \text{if } a_{i1} = 2 \end{cases}; \tag{6}$$

where $\mathcal{H}_1^B$ now denotes all the possible histories generated by the actions taken by *both* physicians in period one.

First we note that if physician j chooses the known treatment in the first period, the decision situation for physician $i \neq j$ is the same as in version A above. Evidently, if physician j prescribes the known treatment to her first patient she will contribute no information about the characteristics of the unknown treatment. The decision

situation for physician i changes if physician j does prescribe the unknown treatment to her first patient. Given this behavior of j, physician i will know just as much about the unknown treatment as if she had prescribed it herself, without having to take the gamble with her first patient. This argument clearly has a public good flavor, and we may therefore expect that the equilibrium entails less experimentation than in version A.

Again abusing notation slightly, let $V(a_{i1}, a_{j1})$ denote the expected payoff to physician i from choosing treatment $a \in \{1,2\}$ in period one, given that physician $j \neq i$ prescribes treatment $a \in \{1,2\}$ in period one, and that i continues optimally in period two. If we restrict attention to the strategically non-trivial case where $y \in Y^{\delta}$, the respective values become:

$$V(a_{i1}, a_{j1}) = \begin{cases} \frac{6+\delta(3+8y)}{12} \text{ if } (a_{i1}, a_{j1}) = (1,1) \\ \frac{3+\delta(2+3y)}{6} \text{ if } (a_{i1}, a_{j1}) = (1,2) \\ \frac{6y+\delta(2+3y)}{6} \text{ if } (a_{i1}, a_{j1}) = (2,1) \\ (1+\delta)\, y \text{ if } (a_{i1}, a_{j1}) = (2,2) \end{cases} .$$

The ranking of the expected payoffs is

$$V(2,1) \geq V(1,1) \geq V(1,2) \geq V(2,2) \ ,$$

(with strict inequalities if $y \neq 1/2$), for both physicians. Given that one of the physicians (i, say) prescribes the known treatment to her first patient, the other physician would rather prescribe the unknown treatment to her first patient. This is natural in view of the results in version A : physician j acts as if in isolation. Given that i prescribes the unknown treatment, j would prefer to free ride on the information produced by physician i while prescribing the known treatment to her own first patient. For $y \in Y^{\delta}$, the physicians initially believe that the known treatment is superior to the unknown treatment. Therefore, the known treatment is preferable in the first period, given that the same information set is reached prior to stage two. The payoff structure suggests, however, that these equilibria may be difficult to attain since the physicians have conflicting preferences over the asymmetric equilibria. Consequently, the physicians face a coordination problem: both physicians prefer to be the one *not* experimenting. Therefore, it can be argued that the likely outcome in this one shot game is that both choose to prescribe the known treatment to their first period patients. This is in fact the worst possible outcome. The payoff structure is in fact

highly reminiscent of that in a Hawk-Dove game, in which coordination failure has long been recognized to cause Pareto dominated outcomes. The main result from the analysis above is thus as follows.

Proposition 2. *Given that $y \in Y^{\delta}$, there are two asymmetric pure strategy subgame perfect equilibria where one physician experiments and the other does not.*[8] *However, coordination failure will potentially cause a Pareto dominated out-of-equilibrium outcome in which none of the physicians experiments.*

We note that an equal amount of information about the unknown treatment is produced in the equilibria in versions A and B. However, the expected payoffs are not the same in the two versions, since the information generated by the experimenting physician also benefits the other physician. The highest individual expected payoff accrues in version B to the physician who does not experiment (when the other does). Moreover, the highest expected *combined* payoff is achieved in the asymmetric equilibria in model B.

If *both* physicians experiment under mutual observability, as in version B, the combined payoff would be less: the benefit to the physicians (as a group) from taking the safe action in period one outweighs the benefit derived from the extra information produced when both physicians experiment. Although not an equilibrium, this outcome produces the second highest combined expected payoffs and therefore dominates the equilibrium outcome in version A, where both individuals experiment in isolation. Furthermore, this outcome produces the most information about the unknown treatment's success frequency since two (joint) observations are made in stage one. This also suggests that there are values of the known treatment's success frequency y for which *cooperative* experimentation would occur, when individual experimentation does not. If both physicians were to experiment and treat their respective patients "as one" (that is, treated jointly) in stage one, optimality would prescribe that they experiment for all y values in the set Y^{δ}_{team}, where

$$Y^{\delta}_{team} = \left[\frac{1}{2}, \frac{3(\delta+2)}{4(\delta+3)}\right) .$$

Clearly, $Y^{\delta} \subset Y^{\delta}_{team}$, implying that joint experimentation should optimally occur for a larger set of parameter values of the known treatment. However, although this behavior improves the combined expected payoff and increases knowledge about the unknown treatment, the expected payoff is still inferior to the asymmetric equilibrium outcomes in version B : in a sense, too much information is produced in this outcome.

[8] There is also an inefficient mixed strategy equilibrium where both physicians experiment with probability $p = [6y - 3 + \delta(3y - 2)]/\delta(8y - 5)$.

Consequently, in a cooperative solution (or a social planner solution), only one of the physicians should experiment with the unknown treatment. The other physician should prescribe the known treatment in stage one, and subsequently use the information produced by the experimenting physician. In order to arrive at any of the two equilibrium outcomes, it is required that one physician "yield" to the other and experiment in stage one. This is, as argued above, perhaps unlikely to occur since the physicians are identical in every respect. Rather, there is a good possibility that the worst outcome arises in which none of them experiments.

A natural question is how the analysis above changes if we consider a larger population of patients in each period. In the following section we see that the results above remain robust in this case.

3. Extended patient population

Suppose now that each of the two physicians is visited by more patients who should be treated as effectively as possible. There are a few different, although related, scenarios to consider here. First, we assume that the physicians have a binary experimentation choice in both periods: either they experiment and prescribe the unknown treatment to all n patients, or they do not and let all n patients have the known treatment. This perhaps somewhat unnatural assumption constitutes a first step, since it is the closest to the basic model considered above.

A more natural formulation would make the physicians able to prescribe both treatments in both periods. That is, we may let both physicians $i \in I$ choose a subset of $k_{i1} \in \{0, 1..., n\}$ patients to whom they assign the unknown treatment in period one. Then they observe the outcomes, and based on the results from the observations they choose a new subset k_{i2} from a new set of N patients.

We consider also the following third case. Both physicians each face a total (known) number of n patients to treat over both periods. The problem for physician i is to choose an optimal number k_i individuals for whom to assign the unknown treatment in period one (the test stage). Given the outcomes, physician i then prescribes the treatment she believes to be superior to the remaining $n - k_i$ patients in period two (the treatment stage). Her objective is still to maximize the overall number of successful treatments in both periods. The difference between this formulation and the two preceding ones is that here, the physicians do not get a new batch of patients in period two. Although this framework is different from the setup in the basic model considered in the previous section, it can be argued that it is perhaps the most natural formulation. Moreover, it has been considered in recent studies of clinical trial sizes.[9] It is therefore worthwhile to consider its extension to the B version with

[9]See e.g., Gittins and Pezeshk [2000], and Su, Cheng, and Berry [1999].

mutual observability. We will refer to the formulations above simply as cases I, II, and III.

In order to facilitate comparisons with the basic model, we consider the various cases in versions A and B separately. Moreover, we assume that $\delta_i = 1$ for $i, j \in I$ throughout this section for analytical convenience.

3.1. Case I. Here we suppose that the physicians have a binary experimentation choice in both periods with $k_{i1} = \{n, 0\}$ patients. Suppose first that the physicians have no possibility to observe each other, that is, we consider version A. The problem is the same as the one in the basic case, with the second period value exchanged to $\mathbb{E}[\max\{\mathbb{E}[X_2|k_{i1}], y\}]$, where $\mathbb{E}[X_2|k_{i1}]$ denotes the expectation of X_2 conditional on the history generated by k_{i1} observations. For ease of exposition, we let the function $F(n)$ be the expected value of the maximum of the estimated unknown treatment's success probability given n trials, $\mathbb{E}[X_2|n]$, and the known treatment's success frequency, y. That is, we define

$$F(n) = \mathbb{E}[\max\{\mathbb{E}[X_2|n], y\}] \ . \tag{7}$$

Intuitively, it seems natural that this value should increase as information accrues (otherwise information would not be valuable). This is indeed the case: the function $F(n)$ is non-decreasing in n, provided that the support of the distribution has at least two interior points (see Su, Cheng, and Berry [1998]). In the case where the prior belief over the unknown treatment's success propensity is rectangular, it will in fact be strictly increasing in n. The function $F(n)$ is then given by:

$$F(n) = \frac{1}{n+1} \sum_{k=0}^{n} \max\left\{\frac{k+1}{n+2}, y\right\} \ . \tag{8}$$

This expression is conveniently approximated by the function $\Phi(n)$:

$$\Phi(n) = \frac{1}{n+1}\left[\sum_{k=0}^{y(n+2)-1} y + \sum_{k=y(n+2)}^{n} \frac{k+1}{n+2}\right] \approx \frac{n + y^2(n+2)}{2(n+1)} \ , \tag{9}$$

a concave function which is strictly increasing in n.[10]

As in the basic case, we calculate the critical value y_n^c of the known treatment's success parameter below which the physicians should prescribe the unknown treatment. In line with the section above, we can define an experimentation set $Y_n^1 =$

[10] From here on, all calculations are based on this approximation.

$Y_n = \left[\frac{1}{2}, y_n^c\right)$, which depends on the number of observations n, where it is optimal to prescribe the unknown treatment. The critical level y_n^c is (approximately) given by:

$$y_n^c = \frac{2n + 2 - [2n^2 + 3n + 2]^{1/2}}{n + 2} .$$

The critical value y_n^c is increasing in n, implying that when the number of potential observations available increases, the unknown treatment will be chosen for larger values of y in model A. Hence, in this sense experimentation can be said to occur more easily if there are more patients: this point also relates to the "joint" experimentation case mentioned in the previous section. Also, since $\Phi(n)$ is increasing in n, the expected payoff increases in n in model A.

The situation in the model B version of this extension will qualitatively remain the same as in the basic model B case. While the upper bound of experimentation set Y_n increases with the number of patients n, the free riding effects that the mutual observability assumption produces are still equally powerful here. In other words, the two asymmetric equilibria where one physician experiments and the other does not are now sustained for higher values of y^c, but this outcome is still not very likely since the incentives to free ride on the other physician's experimentation are the same as in the basic case.

More specifically, consider an experimenting physician i in the asymmetric equilibrium. Given that the other physician (j) does not experiment, i's incentives *not* to deviate increase with the number of patients n, by virtue of the results from version A just above. Similarly, physician j will have stronger incentives *not* to deviate from her choice of not experimenting as n increases, since she also benefits from the knowledge produced by i : although her benefit increases in n regardless of whether she chooses the unknown or the known treatment, she benefits *more* from not prescribing the unknown treatment. This again is a consequence of decreasing marginal returns to information. Hence, in this sense, the asymmetric equilibrium is sustainable for larger values of y when n gets larger. The requirement on y in order for the experimenting physician not to deviate is $y < (E[X_1] + \Phi(n))/2$, which is increasing in n. Similarly, for j not to deviate we have that $y \geq E[X_1] + \Phi(2n) - \Phi(n)$ must hold, which is decreasing in n, but bounded from below by $1/2$ since $\lim_{n\to\infty} \Phi(2n) - \Phi(n) = 0$.

We summarize the results as follows.

Proposition 3. *If n denotes the number of outcomes resulting from the unknown treatment in period one, then we have that $Y_n \subset Y_{n+1}$ (for given y) in model A. Although the set Y where an asymmetric equilibrium is sustained increases in n, the coordination failure is just as likely to occur as in the basic model B.*

3.2. Case II. In this case, we suppose that the physicians can make an (approximately) continuous choice of k patients from all n patients in both periods. In version A, where the physicians can only observe the outcomes from their own treatments, both physicians $i \in I$ choose k_{i1} and k_{i2} in order to maximize the following expression:

$$V(k_{i1}, k_{i2}) = \tfrac{1}{n}\left[k_{i1}\mathbb{E}[X_1] + (n - k_{i1})\, y + k_{i2}\mathbb{E}[X_2 \,|k_{i1}] + (n - k_{i2})\, y\right] , \tag{10}$$

subject to $k_{i1}, k_{i2} \in [0, n]$. Clearly, it must be optimal to prescribe the treatment currently believed to be superior to all n patients in the last period, irrespective of what happens in the first period. The problem is therefore to choose k_{i1} as to maximize:

$$V(k_{i1}, k_{i2}) = \frac{1}{n}\left[k_{i1}\mathbb{E}[X_1] + (n - k_{i1})\, y + n\Phi(k_{i1})\right] . \tag{11}$$

The solution to this problem is given by:

$$k_{i1}^{\star} \approx \begin{cases} \left[\frac{n(1-y^2)}{2y-1}\right]^{1/2} & \text{if } y > 1/2 \\ n \text{ otherwise} \end{cases} , \tag{12}$$

for both $i, j \in I$. This result will be modified when we consider version B where the physicians have access to each others' prescription outcomes. In this case, the best reply function for physician i is given by:

$$k_{i1}(k_{j1}) = \left[\frac{n(1-y^2)}{2y-1}\right]^{1/2} - k_{j1} . \tag{13}$$

The best reply functions for the physicians coincide, and we have a continuum of equilibria. We note that if physician j treats one additional patient with the unknown treatment in stage one, physician i will optimally reduce her prescription of the unknown treatment by one. A symmetric equilibrium has both physicians sharing the experimentation activities:

$$k_{i1}^{\star} = k_{j1}^{\star} = \frac{1}{2}\left[\frac{n(1-y^2)}{2y-1}\right]^{1/2} . \tag{14}$$

However, the same reasoning that applies in the basic model is also valid here. Although none of the physicians has an incentive to deviate unilaterally there is still an incentive to free ride on the other. This is no surprise in light of the results obtained earlier. Since experimentation produces the same information, irrespective of who conducts it, and both physicians can acquire it from each other with out cost, it is better to play it safe and let the other experiment. The preferred equilibrium by

both physicians is therefore to choose no experimentation in the first period, and let the other physician experiment with her patients. As in the basic case of version B in section 2, there is thus a strong incentive not to experiment that potentially leads to an adverse outcome with no experimentation. In sum, we have the following proposition.

Proposition 4. *The optimal experimentation size in a population of n patients in period one is of order $n^{1/2}$ in model A. In model B, there are many equilibria which involve experimentation from both physicians that are all Pareto efficient. However, a likely outcome is that neither experiments due to the information externality.*

3.3. Case III. Finally, we consider the third formulation where the decision concerns the optimal clinical trial size $k \leq n$, rather than simply whether or not to experiment in period one. As mentioned above, this setup is different from the one in the basic model. However, we find the formulation interesting because it has a natural appeal and, moreover, it appears elsewhere in the literature on clinical trials.

Beginning again with model A, the maximization problem facing physician $i \in I$ is now to choose k_i as to maximize the expression:

$$V_i(k_i) = \frac{1}{n}\left[k_i \mathbb{E}[X_1] + (n - k_i)\,\mathbb{E}\left[\max\left\{\mathbb{E}\left[X_2 \middle| k_i\right], y\right\}\right]\right] , \tag{15}$$

subject to $0 \leq k_i \leq n$. The solution to this program shows that the optimal clinical trial size k_i^* is of order $n^{1/2}$:

$$k_i^* \approx \frac{1}{y}\left[n\left(1 - y^2\right)\right]^{1/2} . \tag{16}$$

This result is in line with the result offered in Su, Cheng, and Berry [2000] : for non-extreme values of y and for large n, the optimal clinical trial size should be substantially smaller than the patient population. Using this as a benchmark, we now turn to version B. Not unexpectedly, it turns out that the equilibrium clinical trial size is smaller than in version A. The best reply functions are:

$$k_i(k_j) \approx \frac{1}{y}\left[\left(n + k_j\right)\left(1 + y^2\right)\right]^{1/2} - k_j , \tag{17}$$

for $i, j \in I$, and the unique pure equilibrium is symmetric, given by:

$$k^* \approx \frac{1}{2y}\left[n\left(1 - y^2\right)\right]^{1/2} . \tag{18}$$

The physicians will reduce their trial sizes by about a half compared with the result in model A. The flavor of this result is similar to the one in the cases above: the

experimental activities decrease when they no longer act in isolation. It is important to note, however, that this is framework is slightly different from the basic model and the first extension. There is only one decision node here, so the physicians cannot free ride on each other's experimentation to the same extent as in the basic model. Therefore, the effect of the coordination failure is not nearly as severe as in that case. Nevertheless, in equilibrium each expects the other to contribute with observations from the unknown treatment, so the physicians assign fewer patients to the unknown treatment in version B than they do in version A.

4. Concluding discussion

The objective of this paper has been to establish a simple framework for analyzing the strategic behavior that might arise in situations in which individuals can observe each others' actions. We considered a two period scenario in which physicians choose between two treatments, of which one has unknown characteristics. Learning then becomes relevant since the objective is to treat patients as effectively as possible. In particular, we compared two different information structures. In the first, version A, the physicians could only learn about the unknown treatment's efficacy by prescribing it to their patients. In the second, version B, the physicians could also without cost observe the outcomes from each others' prescriptions as well as from their own. This opened up another possibility to learn about the unknown treatment since the outcomes were assumed to be independent of the physicians and patients. We defined *experimentation* as the act of prescribing the unknown treatment in the first period, even though the prior belief was that the other treatment was superior. Within the confines of a special example, we solved explicitly for the equilibrium strategies in the two versions, and characterized the possible outcomes.

For the relevant parameter values, our results suggest that experimentation is less likely to occur when the physicians can observe each other than when they cannot. The reason for this is that the information produced as a result of prescribing the unknown treatment is a public good. Therefore, there is a strong incentive for both physicians not to experiment and thereby increase their expected payoff by choosing the "safe" treatment in the first period. For the interesting values of the known treatment's success propensity y, we typically have asymmetric equilibria where one of the physicians experiments with the unknown treatment, and the other prescribes the known treatment. Since the physicians are symmetric in all respects, there is no particular reason to believe that they will end up in an asymmetric equilibrium where one of them "yields" to the other. Therefore, a likely outcome is that both choose the known treatment in both periods and thereby forego the expected payoff that they otherwise would get.

It is interesting to compare how the different outcomes would be ranked from

a societal perspective. A social planner concerned with the sum of the physicians' welfare prefers the asymmetric equilibria since they provide the highest expected combined payoff. The optimal amount of experimentation is the same in versions A and B in terms of the knowledge gained. In the latter case, it is enough that one of the physicians experiments and that the other free rides on this information. If both individuals were to experiment under mutual observability, the expected payoffs decrease. This outcome is, however, the second most preferred by a social planner. Hence, to ensure the highest expected combined payoff, the planner would promote optimal experimentation by assigning this activity to some physicians, and dictate that this information be available to other physicians.

The general result remains valid also in the extended version, in which more observations are available. The efficiency loss due to coordination failures will increase when the number of patients is larger: although more information is generated in the asymmetric equilibria, this information is lost in the out-of-equilibrium outcome, just as in the basic case.

Thus, this analysis has found that there is an increased risk that knowledge about an unknown treatment is not advanced in a setting where the decision makers can observe each other, in contrast to a situation in which they cannot. In short, this setup that allows for observational learning is highly afflicted by the social inefficiencies that are usually associated with public good analysis.

References

[1] Aoyagi, M. [1998] "Mutual Observability and the Convergence of Actions in a Multi-Person Two-Armed Bandit Model", *Journal of Economic Theory* 82, 405-424.

[2] Banerjee, A. [1992] "A Simple Model of Herd Behavior", *Quarterly Journal of Economics* 107, 797-818.

[3] Bikhchandani, S., Hirshleifer, D., and I. Welch [1992] "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades", *Journal of Political Economy* 100 (5), 992-1026.

[4] Berry, D.A. [1972] "A Bernoulli Two-Armed Bandit", *Annals of Mathematical Statistics* 43, 871-897.

[5] Berry, D.A., and B. Fristedt [1985] *Bandit Problems, Sequential Allocation of Experiments*, Chapman and Hall, London.

[6] Bolton, P., and C. Harris [1999] "Strategic Experimentation", *Econometrica* 67, 349-374.

[7] Cheng, Y., F. Su, and D. A. Berry [1998] "Group Sequential Strategies in Two-Armed Bandit Problems: Implications for Clinical Trial Sample Sizes", unpublished manuscript, Department of Statistics, Duke University.

[8] DeGroot, M. [1970] *Optimal Statistical Decisions*, McGraw-Hill, New York.

[9] Gittins, J.C. [1979] "Bandit Processes and Dynamic Allocation Indices", *Journal of the Royal Statistical Society, Series B (Methodological)* 41 (2), 148-177.

[10] Gittins, J.C., and H. Pezeshk [2000] "How Large Should a Clinical Trial Be?", *Journal of the Royal Statistical Society, Series D (The Statistician)* 49 (2), 177-187.

[11] Jovanovic, B. [1979] "Job Matching and the Theory of Turnover", *Journal of Political Economy* 87 (5), 972-990.

[12] Rotschild, M. [1974] "A Two-Armed Bandit Theory of Market Pricing", *Journal of Economic Theory* 9, 185-202.

[13] Smith, L. and P. Sorensen [2000] "Pathological Outcomes from Observational Learning", *Econometrica* 68 (2), 371-398.

[14] Viscusi, W. [1979] "Job Hazards and Worker Quit Rates: An Analysis of Adaptive Worker Behavior", *International Economic Review* 20 (1), 29-58.

[15] Whittle, P. [1982] *Optimization Over Time: Dynamic Programming and Stochastic Control, Vol I*, Wiley, New York.

Essay IV: Altruism as a cause of insurance market failure

Altruism as a cause of insurance market failure

BJÖRN PERSSON AND JÖRGEN W. WEIBULL*
STOCKHOLM SCHOOL OF ECONOMICS

ABSTRACT. We here develop a model in order to study the effect of altruism on individuals' insurance decisions. The main question is if the potential strategic incentive to free ride on others' altruism can lead to insurance market failure. One relevant concern is whether there will be any transfers at all in a large economy, even with altruistic individuals, since the incentive to give may be weakened as the population grows - altruists free-riding on each other. Our main results indicate that altruism can have a detrimental effect on insurance markets in a setting where the individuals are identical. If, however, individuals differ much as to the health risk, and these individual risks are observed by the insurers, then we find that the degree of altruism must be (perhaps unrealistically) high in order for altruism to cause insurance market failure. A more complex pattern is found in the case of asymmetric information: low levels of altruism increases the number of equilibria, as compared with the case without altruism, while high levels of altruism knocks out all insurance equilibria.

1. INTRODUCTION

Every year in the US, where the extent of private donations to charity is the largest in the world, more than two-thirds of the population makes voluntary contributions to a wide range of charities, interest groups, and non-profit organizations. A recent article (Rose-Ackerman [1996]) reports that the value of the total monetary donations in the US exceeded $120 billion in 1990, nine-tenths of which were personal contributions. The fact that individuals donate money to strangers could be interpreted as evidence that people have altruistic concerns for others' welfare.[1]

Altruism can, however, potentially cause socially suboptimal outcomes by creating strategic incentives for individuals to exploit the benevolence of others. The reason is that an individual with altruistic preferences cannot commit not to help another individual in need, and knowing this the receiving individual may for example

*Department of Economics, Stockholm School of Economics, SE - 113 83 Stockholm Sweden, and the Research Institute of Industrial Economics, SE - 114 85 Stockholm, Sweden.

[1]Giving up consumption for the benefit of others may also be consistent with selfish behavior. Individuals may be motivated by guilt or a sense of duty, or have preferences for fairness (see e.g., Andreoni [1988]). Furthermore, individuals may want not to see poor people in the streets, or be concerned with reputation effects and social status, or they may hope to be helped in the future if they help others now (cooperative egoism).

over-consume at an early stage in expectation of help at a later stage. The altruist recognizes this, and rationally over-saves in order to provide support to the other. Therefore, an inefficient allocation of resources may result. Interestingly, both the altruist and the exploiting individual could be better off if the altruist could commit to his or her future support. This is sometimes referred to as the Samaritan's Dilemma, after Buchanan [1975] who identified the incentive problem.

In this paper, we examine how incentives to purchase insurance are influenced by the presence of altruism. We consider individuals who in addition to their own utility from consumption care also about other's utility from consumption.[2] All individuals are assumed to be equally altruistic toward each other. The individuals consume one good only, which we refer to as *health* for the sake of concreteness. We believe that health serves as a good motivation for our purpose, although it puts no restriction on our treatment as such: in the analysis, health is tantamount to income and can be traded among the individuals. The cost incurred for improving another individual's health equals a loss of one's own health to exactly the same degree. We imagine that initially each individual is endowed with one unit of health and faces an exogenous risk of falling ill (losing health). Prior to the realization of such a health loss, individuals can purchase health insurance that fully protects them from this adverse outcome. Due to the presence of altruistic preferences, however, individuals may want to make health transfers to those who are worse off than themselves - for example to uninsured individuals who fall ill. Therefore, some individuals may opt to forego insurance and instead rely on the altruism of others. Our main task, then, is to investigate how the possibility of receiving charity affects individuals' decisions to buy insurance in the market. In particular, we ask under what circumstances, and to what extent, altruism can cause the insurance market to partially fail or collapse completely.

Most of the previous literature on altruism and efficiency has considered the consequences that arise from interaction between only a few (usually two or three) individuals. A paper closely related to ours is Lindbeck and Weibull [1988], which provides a game-theoretic analysis of the incentive conflict considered in Buchanan's informal account mentioned above. That study focuses on the case of two equally or unequally altruistic individuals who make intertemporal consumption decisions. As a special case, one individual may be altruistic toward the other who is purely selfish. In subgame perfect equilibrium, the selfish individual will squander at an early stage, thus leaving the altruist to provide for him in the next stage. The altruist cannot credibly commit to not help the other individual, and such equilibrium behavior may be Pareto inefficient, an inefficiency that can prevail even if both individuals are equally

[2]This is sometimes called *pure* altruism. An *impure* altruist also derives utility directly from the act of giving *per se*, a so called warm glow effect (see Andreoni [1989]).

altruistic. It is argued in that study that the externality caused by altruism may justify mandatory pension schemes from an efficiency viewpoint, since these provide a form of binding commitments. The conclusion is at variance with Becker's [1974], [1976] so called rotten-kid theorem.[3]

The role for mandatory insurance as a means to alleviate distortions due to altruism is further elaborated in Coate [1995], who considers an economy with two altruists and one selfish individual. The selfish individual is risk-averse and subject to a possible loss of income, but can obtain insurance against this loss. The two altruists are wealthy, and this is known by the selfish individual. The latter will under-insure in anticipation of help. In Coate's model, one source of inefficiency comes from the fact that the poor individual, by relying on the altruists, does not choose a socially optimal level of protection. Another inefficiency arises from the fact that the altruists' giving may not be optimal from their own viewpoint. The reason is that charity is a public good to both altruists, and is therefore afflicted with well-known free-riding problems: each altruists has an incentive to free-ride on the other's altruism, and therefore to under-provide charity. Consequently, there is a "double" rationale for the public provision of in-kind transfers to the poor from a government acting on behalf of the altruists. Coate also notes that the public insurance solution may not be unambiguously (Pareto) optimal in case the altruists also derive "warm-glow" (Andreoni [1989]) from giving charity. He suggests that the charitable behavior of others is key to explaining why many (low-income) citizens in the US lack basic health insurance and disaster insurance.

In the scenarios described above, altruism can potentially cause severe social inefficiencies. It is not obvious, however, to what extent these results hold when one considers a large population. In particular, if there are many altruistic individuals, the free riding problem associated with giving will naturally be intensified. Each potential contributor may consider his or her donation to be so marginal that giving does not seem worthwhile. A major concern we address, then, is whether there will be any transfers at all in a large economy. Free riding among contributors is considered in Andreoni [1988], who argues that the public good view of altruism (pure altruism) cannot possibly account for the extensive charity donations observed empirically. As the population grows, only the wealthiest individuals will make contributions, and average donations will converge to zero.[4] In light of this result, we might expect that the effects of altruism are too small to cause insurance markets to fail: if would-be recipients expect only diminutive charity contributions, they clearly have an incentive

[3]A treatment of the Samaritan's Dilemma and its relation to Becker's rotten-kid theorem is also found in Bruce and Waldman [1990].

[4]Andreoni concludes that motives other than altruism are needed to explain actual donation patterns.

to self-protect efficiently.

Another question we address here relates altruism to the information structure that prevails between insurers and their clients. The adverse selection mechanism that normally kicks in when there is asymmetric information about the propensity of loss is a well-known cause for insurance failure. Little is known, however, how the adverse selection effect interacts with altruism. Specifically, we ask whether altruism will reinforce or counteract the adverse selection effect. It is possible that high-risk individuals will have weaker incentives to insure, compared to a situation without altruists. This could cause insurance premia to fall and thus bring more low-risk individuals into the insurance market, and thus alleviate adverse selection. On the other hand, it is possible that altruism may provide a "cushion" that weakens the incentive to buy insurance. We therefore examine the traditional case in which the insurers know only the average risk level in the population while each individual knows his or her individual risk.

Our main results indicate that altruism can have a detrimental effect on insurance markets in a setting where the individuals are *ex ante* identical. If, however, individuals differ much as to the health risk, and these individual risks are observed by the insurers, then we find that the degree of altruism must be (perhaps unrealistically) high in order for altruism to cause insurance market failure. A similar, but somewhat more complex pattern is found in the case of asymmetric information: although altruism's net effect is to reinforce the adverse selection effects, this distorting influence becomes strong only at high levels of altruism. At low levels, it causes multiplicity of equilibria (one additional equilibrium, on top of the two equilibria in the absence of altruism), while at high levels of altruism it knocks out the socially least inefficient equilibrium, and the only remaining equilibrium is complete collapse of the insurance market. Hence, there is a cause for public health insurance programs, but perhaps not as strong (at realistically low levels of altruism) as one might have thought.

This paper is structured as follows. First we develop a base model in which all individuals are *ex ante* identical in every respect, and there is no informational asymmetry. Then we extend the model to allow for individuals with differing health risks; there is distribution of health risks in the population, and this distribution is known by everyone in the economy. We consider both the case of complete information about each individual's health risk, as well as the case when each individual's health risk is his or her private information. The results, and some directions for further research are discussed in the final section. An appendix establishes a connection between altruistic concerns for others' utility from health and altruistic concerns for others' welfare including others' altruistic concerns for others welfare etc. *ad infinitum*.

2. The Base Model

Consider an economy consisting of n *ex ante* identical individuals. Each individual is in each period in some (health) state, $h \geq 0$. We will call an individual in state $h = 1$ "healthy" and an individual in state $h = 0$ "ill." All individuals live in (the same) two periods, $t = 0$ and $t = 1$. In period 1, every individual is healthy with probability $1 - p$ and ill with probability p, where $p \in [0, 1]$. In period 0 there exists an insurance market against this risk. The cost of bringing someone who has fallen ill back to good health is 1, and the premium for this insurance is actuarially fair, $w = 1 - p$. Hence, the expected profit to an insurer is zero, and an individual who buys insurance is certain to be in health state $h = 1 - p$ in period 1. We assume in the base model that all individuals and insurers know the risk p of falling ill.

Let x denote the number of individuals who choose to buy the insurance in period 0. In period 1, the total population in general consists of three groups: (A) those who did not buy the insurance in period 1 and did not fall ill, and hence are in state $h = 1$, (B) those who bought the insurance, and thus are in state $1 - p$, and (C) those who did not buy the insurance and fell ill, and accordingly are in state $h = 0$. Although each uninsured individual's health state is a random variable, we assume that the size of each population group is deterministic.[5] Every individual being exposed to the same risk of falling ill, their sizes are accordingly given by

$$n_A = (1 - p)(n - x), \; n_B = x, \text{ and } n_C = p(n - x) \; . \tag{1}$$

2.1. Selfish individuals. Suppose all individuals are selfish in the sense that no individual cares about any other individual's health state. Suppose, moreover, that all individuals have the same von Neumann-Morgenstern preferences over health lotteries, with Bernoulli function $u : \mathbb{R}_+ \to \mathbb{R}$, where u is twice differentiable, with $u(0) = 0$, $u(1) = 1$, $u' > 0$, $u'' < 0$, and $\lim_{h \to 0} u'(h) = +\infty$. All individuals are thus risk-averse, with infinite marginal utility at health state zero.

In the absence of altruism, the individual health states in the three groups are $h_A = 1$, $h_B = 1 - p$, and $h_C = 0$, respectively. Every individual, being risk-averse, will thus choose to buy the insurance:

$$u(1 - p) > (1 - p)\, u(1) + p u(0) = 1 - p \; . \tag{2}$$

2.2. Altruistic individuals. Assume now that all individuals have the same altruistic preferences, defined by the sum of the utility from own health and a scalar

[5] In a large population with statistically *independent* individual health states, the population shares are close to their expected values, by the law of large numbers. Alternatively, individual health states may be thought of as statistically *dependent* random variables such that the population shares are deterministic.

$\alpha \in (0,1)$ times the sum of all others' utility from health:[6]

$$U_i = u(h_i) + \alpha \sum_{j \neq i} u(h_j) = (1-\alpha)\, u(h_i) + \alpha \sum_{j=1}^{n} u(h_j) \ . \tag{3}$$

Because of their altruistic concern for others, individuals may want to give voluntary transfers to those in worse health state than themselves. We therefore assume that, when all individuals' health states are realized in the beginning of period 1, the number of ill is known to all, and all individuals who are not ill decide independently whether or not to give a transfer to the ill, and, if so, how much. We thus model the interaction as a two-stage game with simultaneous moves in each stage. In stage one all individuals simultaneously choose whether or not to buy the insurance. In stage two, each individual learns his or her health state, and the number of uninsured ill. All individuals then simultaneously choose their personal voluntary (non-negative) transfer to the ill. We analyze the subgame perfect equilibria of this game.

Let $t_i \geq 0$ denote i's transfer. We imagine that all these individual transfers are handed over to a charity organization which helps all individuals who are ill.[7] Every such individual receives the same support[8]. Hence, with x insured individuals, all uninsured ill end up in the same health state

$$h_C = \frac{1}{(n-x)\, p} \sum_j t_j \ . \tag{4}$$

Consequently, for any individual i in group A or B, the optimal voluntary transfer, if positive, must satisfy the following first-order condition:

$$u'(h_i) = \alpha \frac{d}{dt_i} \sum_{j \neq i} u(h_j) = \alpha (n-x)\, p \frac{du(h_C)}{dt_i} = \alpha u'(h_C) \ , \tag{5}$$

[6] A more profound model of altruism is to let individuals care about others' altruistic concern for others, etc. in an infinite regress. It can be shown that an additively separable such model is mathematically equivalent with the present model, in a population of fixed size, see appendix at the end of the paper.

[7] In particular, this presumes that this organization without cost or error can verify the health state of any individual. In the more realistic case when the organization cannot do this, an additional moral hazard problem arises as healthy individuals may claim that they are ill in order to receive support.

[8] Due to the concavity of individual utility functions, this allocation maximizes the utility sum across all receivers.

where h_i is the health state of a giver i *after* giving the transfer, t_i.[9] The right-hand side is the same for all givers. Hence, all givers end up in the same health state

$$h_G = (u')^{-1}[\alpha u'(h_C)] \ . \tag{6}$$

Not surprisingly, each giver's health state is superior to that of any receiver: $h_G > h_C$ (since $\alpha < 1$).

Figure 1 depicts the decision tree for each individual. Here $\Delta > 0$ denotes the difference between giving the optimal transfer and not giving any transfer, in terms of the sum of all other individuals' consumption utilities.

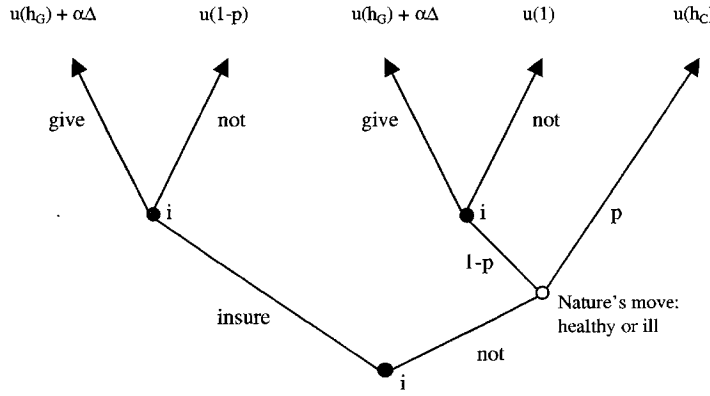

Figure 1: The decision tree for every individual i.

In equilibrium, either all buy insurance, no-one buys insurance, or some interior population share buys insurance. We refer to these equilibria as type I, II and III, respectively. In type-I equilibrium, no transfers are given and the resulting allocation is Pareto efficient. In type-II equilibrium, all healthy individuals give transfers to the ill. (If no transfers were given, then it would be better for each individual to buy insurance, see section 2.1.) In type-III equilibrium, finally, all individuals are indifferent between buying and not buying insurance. However, it is not possible that both the insured and those uninsured who are healthy give transfers. For by the above observation, the resulting health levels must then be the same for both groups of givers, and this would destroy the required indifference: it would then be

[9]No uninsured ill individual has any incentive to give a transfer to other uninsured ill (since $\alpha < 1$). If the giver i in equation (5) belongs to group A, then $h_i = 1 - t_i$, while if $i \in B$, then $h_i = 1 - p - t_i$.

suboptimal to not buy insurance, since the only effect of that would be to run the risk of becoming ill and receiving support (and hence ending up in a worse health state). It follows that in type-III equilibrium the insured do not give transfers, but the uninsured healthy do. However, this can occur only if the following equation happens to hold,

$$(1-p)\,u(h_G) + pu(h_G) + (1-p)\,\Delta = u(1-p) \ . \tag{7}$$

This equation fails generically for all subutility functions u, risks p, and levels of altruism α. Therefore, virtually every subgame perfect equilibrium is either of type I or II, and we accordingly do not analyze type-III equilibria.

Type-I equilibrium. In type-I equilibrium, all individuals buy insurance, no voluntary transfers are given, and the market allocation is Pareto efficient. We here examine under what circumstances such subgame perfect equilibria exist. Suppose therefore that all but one individual, i, buy insurance. The single deviator's *ex ante* expected consumption utility is $1-p+pu(T)$, where T is the sum of transfers from the $n-1$ insured individuals in case the deviator falls ill (we used the assumption that $u(1)=1$). The question is if T is large enough to make such a deviation worthwhile. Suppose all insured individuals give the same transfer, t_B, to the deviator in case the latter fall ill. Hence, $T=(n-1)\,t_B$, where t_B is individually optimal for every insured individual expecting all other insured individuals to transfer the same amount. In order to find t_B, note that, if i falls ill, then each individual $j \neq i$ faces the following decision problem:

$$[M1] \qquad \max_{t_j \in [0,1-p]} u(1-p-t_j) + \alpha u\left[t_j + (n-2)\,t_B\right].$$

The associated first-order condition is

$$u'(1-p-t_j) = \alpha u'\left[t_j + (n-2)\,t_B\right] \ . \tag{8}$$

Hence, a necessary (and in fact also sufficient) condition for t_B to be an equilibrium transfer in this subgame is $u'(1-p-t_B) = \alpha u'\left[t_B + (n-2)\,t_B\right]$, or, equivalently,

$$(n-1)\,t_B = (u')^{-1}\left[\frac{u'(1-p-t_B)}{\alpha}\right] \ . \tag{9}$$

The right-hand side of this equation is a continuously decreasing function of t_B, starting from a positive value and running down to zero as t_B approaches $1-p$. Hence,

there exists precisely one solution. We see from this equation that the equilibrium *per capita* transfer t_B is increasing in the degree of altruism α, and decreasing in the population size n, with limit value zero as $n \to \infty$. However, the total sum of transfers given to the deviator, $T = (n-1)\, t_B$, is increasing both in α and in n. To see the latter property, note that equation (9) can be re-written to yield

$$T = (u')^{-1}\left[\frac{u'(1-p-t_B)}{\alpha}\right] \to T_\infty = (u')^{-1}\left(\frac{u'(1-p)}{\alpha}\right) > 0\ , \tag{10}$$

as the population size n goes to infinity. In sum: while individual transfers tend to zero - "altruists free-ride on each others' altruism" - the total amount of transfers tends to a positive limit value.

Having studied in some detail the support T given to a single uninsured individual who falls ill, we conclude that the strategy profile which has every individual buying insurance constitutes a subgame perfect equilibrium if and only if the total utility to the deviator does not exceed his or her equilibrium utility,

$$(1-p)\left[1+\alpha(n-1)\, u(1-p)\right] + p\left[u(T) + \alpha(n-1)\, u(1-p-t_B)\right] \tag{11}$$

$$\leq \left[1+\alpha(n-1)\right] u(1-p)\ , \tag{12}$$

where t_B and T are given in equations (9) and (10). Note that the deviator is altruistic: he cares about the negative effect on others' consumption when giving a transfer to him. Equivalently, everyone buying insurance is a subgame perfect equilibrium if and only if

$$1-p-u(1-p)+pu(T) \leq \alpha p(n-1)\left[u(1-p)-u(1-p-t_B)\right]\ . \tag{13}$$

In the limit, as $n \to \infty$, this inequality becomes[10]

$$1-p-u(1-p)+pu(T_\infty) \leq \alpha p T_\infty u'(1-p)\ . \tag{14}$$

We illustrate the qualitative properties of these conditions in a special case.

Example. Suppose that $u(h) \equiv \sqrt{h}$. Equation (9) then becomes $(n-1)\, t = \alpha^2(1-p-t)$, so

$$t_B = \frac{(1-p)\,\alpha^2}{n-1+\alpha^2}\ , \tag{15}$$

[10]Note, however, that in the more profound model of altruism outlined in the appendix, $\alpha \to 0$ as $n \to \infty$, at a certain rate, calling for a different limit analysis.

and the condition (13) for subgame perfect equilibrium becomes

$$\alpha p \left[1 - n \left(\frac{\sqrt{1+\alpha^2/(n-1)} - 1}{\sqrt{1+\alpha^2/(n-1)}} \right) \right] \leq 1 - \sqrt{1-p} \,, \tag{16}$$

In the limit as the population sizes goes to infinity, this condition boils down to

$$\alpha \left(1 - \alpha^2/2\right) \leq \frac{1 - \sqrt{1-p}}{p} \,, \tag{17}$$

see Figure 2 below. Full insurance is a subgame perfect equilibrium outcome for all combinations of α and p outside the loop in the figure. For all α below 0.6, such full insurance is an equilibrium at all risk levels p. The degree of altruism must be quite high, and the risk of falling ill low, in order to cause market failure in a large economy. However, we have not yet studied the possibility of multiple equilibria for certain parameter values - another potential source of market failure.

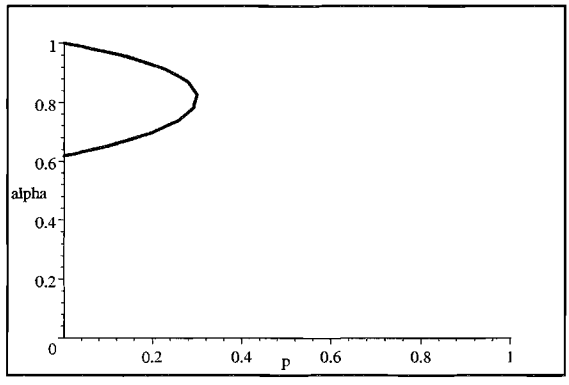

Figure 2: Combinations of α and p for which type-I equilibrium exists.

Type-II equilibrium. In type-II equilibria, no individual buys insurance, and healthy individuals give some transfers to the ill. The resulting resource allocation is Pareto inferior. The source of this inefficiency is the disincentive effect on buying insurance due to one's own altruism and the presence of many uninsured ill. Even if I would buy insurance for myself, there will be so many needy ill around that I will want to give a transfer to these, and this (voluntary) transfer would be so large that, *ex ante*, I would have been better off not insuring myself in the first place. We now proceed to examine the characteristics of such market-failure equilibria. In particular, we ask for what parameter values and utility functions they exist.

We start by computing the equilibrium transfer, t_A, given by each healthy individual in a type II equilibrium. It is clear that t_A is positive, since α is positive, and the marginal utility of health is infinite for every individual who has fallen ill and who has not received any support. The *ex ante* expected equilibrium utility to every individual i is, by (3),

$$U_i = [1 + \alpha (n-1)] [(1-p) u (h_G) + pu (h_C)]) , \tag{18}$$

where $h_G = 1 - t_A$ is the health state of the givers and $h_C = (1-p) t_A / p$ the health state of the recipients. Since the transfer t_A is positive, it necessarily satisfies the first-order condition

$$u' (1 - t_A) = \alpha u' \left[\left(\frac{1-p}{p} \right) t_A \right] , \tag{19}$$

which arises from each of the $(1-p) n$ healthy individuals' decision problem, namely to choose one's own personal transfer t to the charity organization, when expecting all other healthy individuals to each give t_A:

$$[M2] \qquad \max_{t \in [0, 1-p]} \left[u (1-t) + \alpha pnu \left(\frac{t + [(1-p) n - 1] t_A}{pn} \right) \right] .$$

Note that the left-hand side of equation (19) is continuous and increasing in t_A, from a positive value ($u'(1) > 0$), while the right-hand side is continuous and decreasing from plus infinity towards zero. Hence, the two curves intersect at exactly one point, and this is the equilibrium transfer t_A from each healthy individual to the charity organization. We note that this equilibrium transfer is independent of the population size. The diluting effect of more givers on the incentive to give is exactly matched by the enhancing effect of more receivers. Moreover, and not surprisingly, the equilibrium transfer is increasing in the degree of altruism. Note, finally, that the equilibrium transfer from each healthy individual is increasing in p, the risk of falling ill, as well as the number of individuals who are ill.

In order to investigate for what parameter values and utility functions the strategy profile of not buying insurance and giving a transfer if healthy constitutes a subgame perfect equilibrium, consider an individual i who deviates unilaterally from this profile by purchasing insurance and *not* giving any transfer to the ill.[11] Such a deviator's total utility is

[11] More generally, the deviating individual also has the choice of giving a transfer to the ill. If such a transfer would enhance the deviator's total utility, then the present non-profitability condition is only necessary, but not sufficient, for the strategy profile to be subgame perfect.

$$U_i = u(1-p) + \alpha(n-1)\left[(1-p)\,u\left(h_G^i\right) + pu\left(h_C^i\right)\right] ,$$

where h_C^i and h_G^i are the equilibrium health levels of welfare recipients and givers, respectively, in the subgame following upon i's defection. We see in equation (19) that transfers are independent of the total population size, and hence h_C^i and h_G^i are the same as in equilibrium. Hence, the deviation is unprofitable for individual i if and only if his utility from own health does not exceed his expected utility from own health in equilibrium, i.e., iff

$$u(1-p) \leq (1-p)\,u(h_G) + pu(h_C) .$$

This is equivalent to

$$u(1-p) \leq (1-p)\,u(1-t_A) + pu\left[\left(\frac{1-p}{p}\right)t_A\right] ,$$

where t_A is the unique solution of equation (19). This relation implicitly defines a critical level of altruism, as a function of p, below which the type II equilibrium does not exist. We denote this function $\widetilde{\alpha}(p)$. In general, type-II equilibria need not exist for all $\alpha > \widetilde{\alpha}(p)$, but we conjecture that they do exist at least for α somewhat above the critical value $\widetilde{\alpha}(p)$.

Example. Returning to the special case $u(h) = \sqrt{h}$, (19) yields

$$t_A = \frac{\alpha^2}{\alpha^2 + p/(1-p)} , \qquad h_C = \frac{(1-p)^2\alpha^2}{(1-p)p\alpha^2 + p^2} \qquad \text{and} \qquad \widetilde{\alpha}(p) = \frac{p}{1-p} .$$

In figure 3 we plot the critical altruism values for type-I equilibria and type-II equilibria, as functions of the risk of illness, p. For some combinations of health risk levels and levels of altruism, two equilibria exist simultaneously. For values of α and p that are situated in the region outside the loop and above the curve representing type-II equilibria, it is both an equilibrium that *all* purchase insurance and that *no one* purchases insurance. These values represent perhaps the most realistic values of both the altruism parameter and the risk parameter. Consequently, both complete market failure and full insurance are compatible with reasonable parameter values. We note that if the risk is large, it can no longer be an equilibrium that the individuals forego insurance: no matter how strong the altruistic concerns are in the population, buying

insurance is better than relying on altruistic transfers from others.

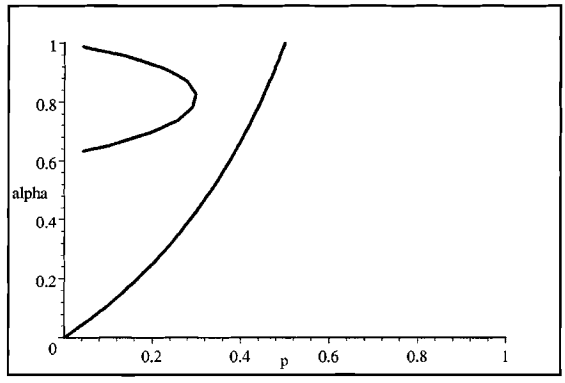

Figure 3: Combinations of α and p for which type-I and type-II equilibria exist.

3. Heterogeneity

In this section we consider how the insurance decisions, and equilibrium outcomes, are affected when individuals differ with respect to their risk of falling ill. As indicated in the introduction, we will analyze this case in two information settings: (C) complete information about the health risks, and (A) asymmetric information, where insurance companies lack information about individual health risks.

We imagine a population consisting of n individuals, $i = 1, 2, ..., n$, who are identical in all respects except for their individual risk p_i of falling ill. In case (C), all individuals and insurance companies knows each individual's risk p_i. In case (A), only individual i knows his or her own risk p_i, while all others, including the insurance companies, only know the average health risk in society,

$$\bar{p} = \frac{1}{n} \sum_{j=1}^{n} p_j \ .$$

In both settings, let $\Phi(p)$ be the population share of individuals i with health risks $p_i \leq p$. Hence, $\Phi : [0,1] \to [0,1]$ is a cumulative distribution function, with mean value $\bar{p}$.

3.1. Complete information. In the case of complete information, the insurance premium exactly matches the individual health risk. Hence, individuals with high health risks may prefer to forgo their expensive insurance and instead rely on others' altruism. Moreover, individuals with low health risks may both buy insurance for themselves and give transfers to the uninsured ill.

We first note that in the absence of altruism, each individual, being risk averse by assumption, will buy the actuarially fair insurance, just as in the case of a homogeneous population (see subsection 2.1). Secondly, in the presence of altruism - assuming all individuals to be equally altruistic - we note that, just as in the homogeneous case, all transfer recipients end up in the same health state h_C, and all transfer givers end up with the same health state h_G, where

$$u'(h_G) = \alpha u'(h_C) \ , \tag{20}$$

(see equations (6) and (5)). Unlike in the case of identical individuals, however, it is now possible that both uninsured healthy and insured individuals make positive transfers to the uninsured ill. In particular, individuals with low health risks p may prefer to both insure and transfer, whereas individuals with high health risks p may prefer not to insure themselves, but give transfers in case of a good health outcome for themselves. Individuals with intermediate health risks, on the other hand, may prefer to insure themselves but not give any transfer.

We now follow up on this equilibrium conjecture, and assume that individuals with health risks below some critical level p_G will both insure and make positive transfers, while individuals with health risks above some critical level $p_C > p_G$ will not buy insurance, and give transfers in case of their own good luck. Individuals with health risks between p_G and p_C will buy insurance but not make any transfers. Are there values of p_G and p_C that make this a subgame perfect equilibrium outcome?

We solve for equilibrium by backwards induction. The equilibrium health level h_G of givers not only has to satisfy the first-order condition (20), but it also has to make individuals with health risk p_G indifferent between giving and not giving a transfer. Thus

$$h_G = 1 - p_G \ . \tag{21}$$

Moreover, the transfer from each individual $i \in A \cup B$ is given by

$$t_i = \begin{cases} 1 - h_G & \text{if } i \in A \\ 1 - h_G - p_i & \text{if } i \in B \text{ and } p_i \leq p_G \\ 0 & \text{if } i \in B \text{ and } p_i > p_G \end{cases} \ . \tag{22}$$

Using equation (21), this becomes

$$t_i = \begin{cases} p_G & \text{if } i \in A \\ p_G - p_i & \text{if } i \in B \text{ and } p_i \leq p_G \\ 0 & \text{if } i \in B \text{ and } p_i > p_G \end{cases} \ . \tag{23}$$

Adding up the transfers from both groups of givers, we obtain the following equation for the total amount of transfers given, $T = \sum_j t_j$:

$$\begin{aligned} T/n &= p_G \int_{p_C}^{1} (1-p)\, d\Phi(p) + \int_{0}^{p_G} (p_G - p)\, d\Phi(p) \\ &= p_G \left[1 + \Phi(p_G) - \Phi(p_C) - \overline{p} + \Psi(p_C)\right] - \Psi(p_G) \ , \end{aligned}$$

where $\Psi : \mathbb{R}_+ \to [0,1]$ is defined by $\Psi(s) = \int_0^s p d\Phi(p)$. Hence, Ψ is continuous and increasing from 0 to $\bar{p}$, with $\Psi(p) \leq \Phi(p)$ for all $p \in [0,1]$. Given all individuals' transfers, h_C, the resulting health state of the uninsured ill, is thus given (just as in equation (4)) by

$$h_C = T\left[n \int_{p_C}^{1} p d\Phi(p)\right]^{-1} = \frac{p_G\left[1 - \Phi(p_C) + \Phi(p_G)\right] - \Psi(p_G)}{\bar{p} - \Psi(p_C)} - p_G \ , \tag{24}$$

The critical risk level p_C, finally, has to make the individual indifferent between, on the one hand, buying insurance and not giving any transfer, and, on the other hand, not buying insurance and giving a transfer if lucky:

$$u(1 - p_C) = (1 - p_C)\, u(h_G) + p_C u(h_C) \ . \tag{25}$$

We thus have four equations in equally many unknowns, p_G, p_C, h_G and h_C. Eliminating the health states, we obtain the following two equations in the two critical risk levels, p_G and p_C:

$$\begin{cases} \alpha u'\left[\left(p_G\left[1 - \Phi(p_C) + \Phi(p_G)\right] - \Psi(p_G)\right) / \left(\bar{p} - \Psi(p_C)\right) - p_G\right] = u'(1 - p_G) \\ p_C u\left(\left(u'\right)^{-1}\left[u'(1 - p_G)\right]/\alpha\right) = u(1 - p_C) - (1 - p_C)\, u(1 - p_G) \end{cases} \tag{26}$$

We return to our example, with the additional assumption that health risks are uniformly distributed in the population.

Example. In the special case when $u(h) = \sqrt{h}$ and $\Phi(p) \equiv p$, the two equations (26) boil down to

$$\begin{cases} \alpha^2 (1 - p_G)(1 - p_C^2) = p_G (1 + p_G - 2p_C + p_C^2) \\ \sqrt{1 - p_C} = \left[1 - p_C(1 - \alpha)\right] \sqrt{1 - p_G} \end{cases} .$$

Figures 4.1 and 4.2 show these two equations as curves in the (p_C, p_G)-plane, for different degrees of altruism.

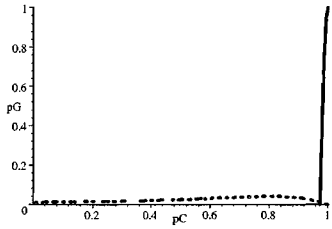

Figure 4.1: $\alpha = 0.1$.

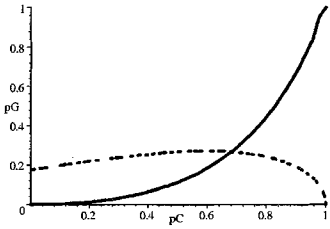

Figure 4.2: $\alpha = 0.5$.

From these graphs we see that if the degree of altruism is low ($\alpha = 0.1$), then only individuals with very high health risks will prefer to forego insurance and count on the help of others. Likewise, the only insured individuals who donate money are those with very low health risks. If the degree of altruism is large ($\alpha = 0.5$), more individuals find it better to become charity recipients in case they do not fall ill. However, even at this high level of altruistic concerns in the population, nearly three individuals in four prefer to insure rather than to rely on others' altruism. We can therefore conclude that in the complete information case, where the individuals have heterogeneous health risks, the altruistic concerns need to be quite substantial to have any significant adverse effect on the insurance market.

3.2. Asymmetric information. As seen in the previous section, it is possible that insurance markets partially fail in the presence of altruism, even when health risks are perfectly verifiable by the insurers. Under complete information, high risk individuals prefer to forego insurance and instead rely on the charity transfers. Suppose now instead that each individual's health risk is private information. A natural question is then how the presence of altruism influences insurance behavior within this information structure. In particular, we examine whether altruism worsens, or alleviates, the adverse selection effect that arises under asymmetric information of the traditional type (A) mentioned above. In the absence of altruism, adverse selection potentially occurs when individuals with low health risks prefer not to insure due to for them unfavorable contracts: insurers must offer a homogenous contract based on the average health risk in the population in order to make up for the losses on high risk individuals, and this may cause the insurance market to fail. In a population with altruistic individuals and private information about health risks, we may still expect the high risk individuals to insure unless the altruistic concerns are very strong.

We assume that the average health risk level in society, $\bar{p}$, is known by everyone. An individual with health risk p loses one unit of health with probability p.

Without altruism. As a benchmark, we first consider the case where the individuals are purely selfish, that is, when $\alpha = 0$. As is well known, insurance markets with asymmetric information may break down due to adverse selection. We maintain the assumption of an actuarially fair insurance market, in the following precise sense. All insurers offer full coverage, i.e., they fully cover the cost of failing ill, one unit of health at a premium which equates the insurer's expected revenue with expected costs. Hence, if all individuals with a health risk p above some critical level p^* buy insurance, then the premium w has to equate the sum of premiums paid with the sum of compensations to those who fall ill,[12]

$$w \int_{p^*}^{1} d\Phi(p) = \int_{p^*}^{1} p d\Phi(p) \ , \tag{27}$$

or, equivalently,

$$[1 - \Phi(p^*)] w = \bar{p} - \Psi(p^*) \ . \tag{28}$$

The critical risk level for buying the insurance, p^*, has to make an individual with precisely this health risk indifferent between buying and not buying the insurance, i.e.,

$$u(1 - w) = 1 - p^* \ , \tag{29}$$

(where we have used the normalizations $u(1) = 1$ and $u(0) = 0$). The two equations (28) and (29) are together necessary and sufficient for insurance market equilibrium, and they together imply the following equation:

$$\Psi(p^*) + [1 - \Phi(p^*)] \left[1 - u^{-1}(1 - p^*)\right] = \bar{p} \tag{30}$$

The corresponding insurance premium w^* is given by equation (29), as

$$w^* = 1 - u^{-1}(1 - p^*) \ . \tag{31}$$

Equations (30) and (31) together fully characterize insurance market equilibrium. Note that one such equilibrium is that only individuals with risk 1 buy insurance at

[12] We assume that the fraction of those with health risk p who fall ill is exactly p, for all health risks p.

price 1: $p^* = w^* = 1$ satisfies both (30) and (31). This "no trade" equilibrium means *complete market break-down*: no risk is eliminated from any individual. The intuition behind this equilibrium is that if insurers expect $p^* = 1$, i.e., that only individuals who fall ill with probability one buy insurance, then they set the premium accordingly, $w = 1$, and then, indeed, no other individuals want to buy insurance since by not buying their health state is never worse and can be better with positive probability.

For certain utility functions and risk distributions, other equilibria exists as well. It is never an equilibrium, however, that all individuals buy insurance. This is seen in equation (30), which does not have $p^* = 0$ as a solution. Hence, all other equilibria are *interior* in the sense that $0 < p^* < 1$.

Example. Returning to the special case where $u(h) \equiv \sqrt{h}$ and $\Phi(p) \equiv p$, equation (30) has exactly two solutions, $p^* = 1$ and $p^* = 1/2$, and (31) gives $w^* = 1$ and $w^* = 3/4$, respectively. In the first equilibrium, effectively no one buys insurance, while in the second half the population (those with health risk above one half) buys insurance at a premium of 3/4.

With altruism. We now return to the case where all individuals have altruistic preferences - the same for all. Like in the complete information case, each insured individual will end up in the same health state, $1 - w$, before making any transfer. Consequently, all insured individuals will in equilibrium make the same transfer to the uninsured ill, $t_B \geq 0$. Similarly, all uninsured healthy individuals will make the same transfer $t_A \geq 0$. Hence, if both the insured and uninsured would give positive transfers, they would end up in the same health state, as givers, just as in the homogeneous case in section two (see first-order condition (20)). This cannot be an equilibrium: it would be profitable for the uninsured to defect by buying insurance and thereby eliminating the risk of getting ill and thus become a transfer recipient. So either only the insured, or only the uninsured lucky, or no-one, give transfers. The first case is excluded, however, since the insured will have a lower health level, $1 - w$, than the uninsured lucky, whose health level is 1. Hence, if the former have an incentive to give transfers then so do the latter, once they have learned their own health state. Thus only two possibilities remain: either only the uninsured lucky give transfers, or no one gives transfers. The latter case is not possible, however, since it presupposes that all individuals buy insurance, which is not an equilibrium, due to the adverse selection effect, see preceding subsection.

In sum: the only possible subgame perfect equilibrium is that all individuals with health risk above some critical risk level p^* buy insurance, and those with health risk below p^* do not, and, once the health states have been known, the uninsured health give some transfer, t_A, to the uninsured ill. The resulting health level for the uninsured healthy is h_G, and that of the transfer recipients h_C, where we now have

$$[1 - \Phi(p^*)]\, w = \bar{p} - \Psi(p^*) \ , \tag{32}$$

$$u'(h_G) = \alpha u'(h_C) \ , \tag{33}$$

and, if $0 < p^* < 1$:[13]

$$u(1 - w) = (1 - p^*)\, u(h_G) + p^* u(h_C) \ . \tag{34}$$

With $n\Phi(p^*)$ uninsured individuals, the number of ill individuals will be $n\Psi(p^*)$. The resulting balance equation for transfers is $n\Psi(p^*)\, h_C = T$, where the total amount of transfers to the ill, T, is given by

$$T = n(1 - h_G) \int_0^{p^*} (1 - p)\, d\Phi(p) \ . \tag{35}$$

The transfer-balance condition can be written

$$\Psi(p^*)\, h_C = (1 - h_G)\, [\Phi(p^*) - \Psi(p^*)] \ . \tag{36}$$

We thus have five equations in equally many unknown, p^*, w, h_G, h_C, and T, characterizing all interior solutions (i.e. with $0 < p^* < 1$)

What about corner solutions, i.e., solution where either everyone buys insurance ($p^* = 0$) or no-one buys insurance ($p^* = 1$)? We know from the previous subsection that the first possibility is incompatible with equilibrium due to the asymmetry of information; the premium is too high for individuals with the lowest health risks. What about the opposite extreme case, which, after all, was an equilibrium outcome in the absence of altruism?

Suppose $p^* = 1$. Then the insurance premium is maximal, $w = 1$, and hence the necessary condition $u(1 - w) \leq u(h_C)$ is met, for any positive transfer to the uninsured ill. It remains to find a pair of health levels (h_G, h_C) that satisfy equations (33) and equation (36), which now boils down to

$$\bar{p} h_C = (1 - \bar{p})\, (1 - h_G) \ . \tag{37}$$

[13] For $p^* = 0$, the equality sign in (34) should be replaced by $\geq$, and for $p^* = 1$ by $\leq$.

These two equations together imply the following fixed-point equation in h_C:

$$h_C = (1/\bar{p} - 1)\left(1 - (u')^{-1}\left[\alpha u'(h_C)\right]\right) . \tag{38}$$

The right-hand side is continuously decreasing in h_C, from $1/\bar{p} - 1$ to minus infinity. Hence, there exists exactly one subgame perfect equilibrium with $p^* = 1$, i.e., in which the insurance market collapses. What about other equilibria. We return to the previous example.

Example. Consider again the special case where $u(h) \equiv \sqrt{h}$ and $\Phi(p) \equiv p$. The five equations characterizing all interior solutions then are:

$$\begin{cases} 2w = 1 + p^* \\ h_C = \alpha^2 h_G \\ w = 1 - \left[(1-p^*)\sqrt{h_G} + p^*\sqrt{h_C}\right]^2 \\ T = n(1-h_G)(1-p^*/2)p^* \\ p^* h_C = (1-h_G)(2-p^*) \end{cases} . \tag{39}$$

These equations imply the following fixed-point equation in p^*:

$$p^* = 1 + \alpha p^* - \sqrt{\frac{(1-p^*)(2-p^*+\alpha^2 p^*)}{4-2p^*}} . \tag{40}$$

Figure 5 below shows this fixed point equation for two cases with positive altruism values, as well as for the case without altruism.

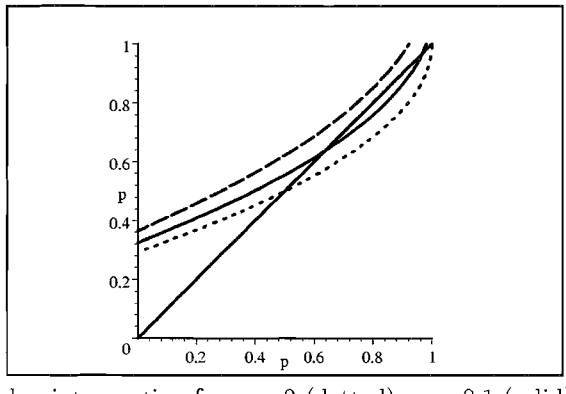

Figure 5: The fixed-point equation for $\alpha = 0$ (dotted), $\alpha = 0.1$ (solid), $\alpha = 0.2$ (dashed).

Any solution to this equation determines uniquely all the other four variables via (39). Note that for all α-values in an interval from zero to some value $\bar{\alpha} \approx 0.16$, there exists two interior solutions, see Figure 6 below. As $\alpha \to 0$, both interior equilibria approach the two equilibria in the absence of altruism ($p^* = 1/2$ and $p^* = 1$). For $\alpha = \bar{\alpha}$, the two interior equilibria collapse to one, and for $\alpha > \bar{\alpha}$ there exists no interior equilibrium. For such high values of α, the insurance market collapses completely.

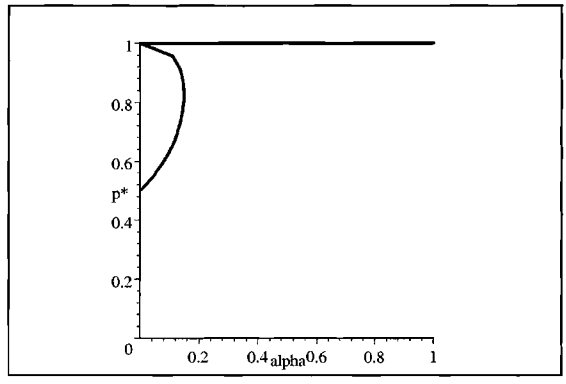

Figure 6: The equilibrium correspondence.

Also note that, as α increases from 0 to $\bar{\alpha}$, the two interior solutions move closer to each other. In other words: altruism in this range splits the interior equilibrium into two, both with less insurance than in the "good" equilibrium without altruism ($p^* = 1/2$ when $\alpha = 0$). In this somewhat complex sense, altruism does not mitigate the adverse selection effect, but reinforces it. The presence of altruism reduces the "punishment" when falling ill uninsured, and hence dilutes the incentive to buy insurance. We conjecture that the "middle" equilibrium is unstable in plausible out-of-equilibrium dynamics in the insurance markets.

4. Appendix

A more profound model of altruism is to let individuals care about others' altruistic concern for others, etc. in an infinite regress. More specifically, suppose

$$U_i = u(h_i) + \beta \sum_{j \neq i} U_j$$

for some $\beta \in (0,1)$. Addition of all n individuals' total utilities gives

$$\sum_{i=1}^{n} U_i = \sum_{i=1}^{n} u(h_i) + (n-1)\beta \sum_{i=1}^{n} U_i$$

and hence

$$\sum_{i=1}^{n} U_i = \frac{1}{1+(n-1)\beta} \sum_{i=1}^{n} u(h_i) \ .$$

Consequently,

$$U_i = u(h_i) - \beta U_i + \frac{\beta}{1+(n-1)\beta} \sum_{j=1}^{n} u(h_j)$$

or, equivalently,

$$(1+\beta) U_i = \left(1 + \frac{\beta}{1+(n-1)\beta}\right) u(h_i) + \frac{\beta}{1+(n-1)\beta} \sum_{j \neq i}^{n} u(hj)$$

In sum:

$$U_i = \frac{1}{1+\beta} \left[\frac{1+n\beta}{1+(n-1)\beta} u(h_i) + \frac{\beta}{1+(n-1)\beta} \sum_{j \neq i}^{n} u(hj) \right]$$

For fixed n, this is mathematically equivalent to the model in the main text, if we set

$$\alpha = \frac{\beta}{1+n\beta} \ .$$

Note, however, that now we have $\alpha \to 0$ as $n \to \infty$, for any fixed β, a relevant concern for the limit considerations in the paper.

References

[1] Andreoni, J. [1988] "Privately Provided Public Goods in a Large Economy: The Limits of Altruism" , *Journal of Public Economics* 35 (1), 57-73.

[2] Andreoni, J. [1989] "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence", *Journal of Political Economy* 97 (6), 1447-1458.

[3] Becker, G. S. [1974] "A Theory of Social Interactions", *Journal of Political Economy* 82 (6), 1063-1093.

[4] Becker, G. S. [1976] "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology" , *Journal of Economic Literature* 14, 817-826.

[5] Bruce, N. and M. Waldman [1990] "The Rotten-Kid Theorem Meets the Samaritan's Dilemma", *Quarterly Journal of Economics* 105, 155-165.

[6] Buchanan, J. M. "The Samaritan's Dilemma", in Edmund S. Phelps, ed., *Altruism, morality, and economic theory.* New York: Russell Sage Foundation, 1975, 71-85.

[7] Coate, S. [1995] "Altruism, the Samaritan's Dilemma, and Government Transfer Policy", *American Economic Review* 85 (1), 46-57.

[8] Lindbeck, A. and J. W. Weibull [1988] "Altruism and Time Consistency: The Economics of Fait Accompli", *Journal of Political Economy* 96 (6), 1165-1182.

[9] Rose-Ackerman, S. [1996] "Altruism, Nonprofits, and Economic Theory" , *Journal of Economic Literature* 34 (2), 701-728.

Essay V: Non-reciprocal altruism in dictator games

Non-reciprocal altruism in dictator games

MAGNUS JOHANNESSON AND BJÖRN PERSSON*
STOCKHOLM SCHOOL OF ECONOMICS

ABSTRACT. We carry out a double blind dictator game experiment where the anonymous recipients are randomly drawn from the Swedish general population, and any donations are mailed to the recipients. About a third of the subjects donate some money.

1. INTRODUCTION

In dictator games, a dictator divides a sum of money between himself and a recipient that must accept the allocation. The standard game theoretic model of payoff-maximization predicts that the dictator will keep all the money himself and leave nothing to the recipient. This prediction has been refuted in numerous experiments (e.g. Kahnemann, Knetsch and Thaler, [1986]; Forsythe et al., [1994]).

A possible explanation for the deviations from payoff-maximization in dictator games is that subjects have a taste for altruism or fairness. This explanation has, however, been disputed by, among others, Hoffman et al. [1994] and Hoffman, McCabe and Smith [1996]. They argued that a main motivation for positive donations in dictator games may be expectations of reciprocity, i.e. by donating some money the subject expects to gain in future interactions with the experimenter, the recipient, or others who know the dictator's decision. They hypothesized that by increasing the social distance between the dictator and the experimenter donations would decrease. This was tested using a double blind experimental design where the decision of each dictator would be anonymous both towards other subjects in the experiment and towards the experimenter. The double blind procedure significantly decreased the proportion of subjects that donated anything to 36 percent and the average donation to 9 percent of the amount allocated. Subsequent experiments with the same double blind design have yielded similar results with average donations ranging between 8 percent and 16 percent of the amount allocated (Eckel and Grossman, [1996], [1998]).

Although the double blind experimental design decreased donations, about a third of the subjects still donated some money to the recipients. The reciprocity between the dictator and the recipient may, however, not have been completely removed with the double blind procedure. The student subjects acting as dictators can infer that the anonymous person is someone like themselves participating in the experiment.

*We thank an anonymous referee for comments.

Therefore, they may not believe that anonymity can be completely guaranteed. This was also suggested as a possible explanation for the remaining level of donations in the experiment by Hoffman, McCabe and Smith [1996].

In this experiment we increase the social distance between the dictators and the recipients by randomly drawing recipients from the adult general population in Sweden. If money is donated, it is mailed to the recipients, who are unaware that they are participating in the experiment. This design guarantees anonymity between dictators and recipients and removes any possible remaining reciprocity in the double blind design used by Hoffman et al. [1994]. If donations in dictator games are motivated solely by reciprocity, donations should therefore drop to zero with this experimental treatment. We also replicate the standard double blind procedure used in recent dictator game experiments (Hoffman et al., [1994]; Hoffman, McCabe and Smith, [1996]; Eckel and Grossman, [1996],[1998]). The null hypothesis we test is that the distribution of donations does not differ between the two experimental treatments.

2. Experimental design

2.1. Double blind standard procedure. In this experiment thirty subjects are recruited to room A (the dictator room) and twenty-nine subjects are recruited to room B (the recipient room). The subjects are paid a SEK 50 show up fee, and asked to sit at assigned seats positioned to keep subjects as separate as possible (SEK=Swedish crowns; Exchange rate June 1999 $1=SEK 8.50). Before the experiment starts the subjects are reminded that there should be no talking during the experiment. The subjects are given the experimental instructions that are read aloud by the experimenter. A monitor is chosen in the dictator room and he/she conducts the experiment and verifies that the procedures are followed as described in the instructions.

Twenty-nine envelopes are randomly distributed to the dictators by the monitor. Twenty-seven envelopes contain five SEK 20 bills and five slips of paper of the same size and the remaining two envelopes contain ten slips of paper. The envelopes with ten slips of paper are included as an additional guarantee of anonymity, since even if no dictator donates any money to the recipients the decision of a single dictator cannot be inferred by the dictator. In private behind a screen, dictators remove five units from the envelope (bills or slips of paper), seal the envelope and then drop it in a box and leave the room. The box with the twenty-nine sealed envelopes is taken by the monitor to room B and the envelopes are randomly distributed to the recipients. The envelopes are opened and the contents are recorded by the monitor. The recipients are given the contents of each envelope and leave the room. The experiment is then over. The procedure is the same as the design used by Hoffman et al. [1994], Hoffman, McCabe and Smith [1996] and Eckel and Grossman [1996],

[1998], with some modifications.[1] We carried out one session of this experiment.[2] Subjects were recruited among undergraduate business and economics students at the Stockholm School of Economics.

2.2. Double-blind with randomly drawn recipients from the general population. In this experiment thirty subjects are recruited to a room. The subjects are paid a SEK 50 show up fee, and asked to sit at assigned seats positioned to keep subjects as separate as possible. Before the experiment starts the subjects are reminded that there should be no talking during the experiment. The subjects are given the experimental instructions that are read aloud by the experimenter. A monitor is chosen and he/she conducts the experiment and verifies that the procedures are followed as described in the instructions.

Twenty-nine envelopes are randomly distributed to the dictators by the monitor. Twenty-seven envelopes contain five SEK 20 bills and five slips of paper of the same size and the remaining two envelopes contain ten slips of paper. In private behind a screen, dictators remove five units from the envelope (bills or slips of paper), seal the envelope and then drop it in a box and leave the room. The monitor is given a box with twenty-nine stamped and addressed envelopes. Each envelope in the box has been provided with the address to a randomly selected person in the Swedish population between the ages 18-74 years, who does not know that he/she has been randomly selected for the experiment.[3] The monitor chooses one addressed envelope and one envelope from the box with the dictators' envelopes. The dictator envelope is opened by the monitor and its contents recorded. If the envelope contains any money, the monitor puts the money in the addressed envelope and seals the envelope. If the envelope does not contain any money, the monitor discards the addressed envelope. The monitor continues until all the envelopes have been opened. The monitor then takes the sealed envelopes with money in and, together with the experimenter, goes

[1] We use a larger group size (thirty subjects rather than fifteen in the dictator room). Another difference is that the amount allocated to the dictators in the previous experiments was $10, whereas it was SEK 100 ($12.5) in this experiment. The use of SEK rather than US dollars also means that the possible divisions of the money will differ. In the previous experiments ten one dollar bills were used, whereas we used five SEK 20 bills (SEK 20 is the smallest denomination of bills in Sweden).

[2] Since six subjects did not arrive on time in the dictator group the number of subjects was twenty-four rather than thirty in the dictator room. The number of subjects was accordingly reduced to twenty-three in the recipient room.

[3] The random sample of recipients was provided by a survey research firm (SEMA-Gruppen). They provided a random sample of all persons registered in Sweden in the ages 18-74 years (drawn from SPAR, the person and address register of the state). The only individuals excluded from the random draw were individuals with "protected addresses" (due to that they are exposed to some threat) which is very uncommon (about 0.01% of the population), and persons that have asked to be excluded from all direct mail advertising (about 2% of the population).

to a mailbox and mails the envelopes. The experiment is then over. The exact instructions for this experiment are reproduced in the Appendix. We carried out two sessions of this experiment. Dictators were recruited among undergraduate business and economics students at the Stockholm School of Economics.

3. Results

The results are summarized in Table 1, showing the percent of decisions for each amount donated.

Table 1: Percent of decisions for each amount donated

Amount donated	Double blind standard	Double blind randomly drawn
SEK 0	66.66	68.52
SEK 20	14.29	22.22
SEK 40	14.29	5.56
SEK 60	0.00	3.70
SEK 80	0.00	0.00
SEK 100	4.76	0.00
Average donation (STD)	SEK 13.33 (24.77)	SEK 8.89 (15.38)
Number of observations	21	54

In the double blind standard procedure 33.34 percent of the dictators donate some money and the mean donation is SEK 13.33 (13.33 percent of the amount allocated to the dictators). In the other experimental group with randomly drawn recipients from the general population 31.48 percent of the dictators donate some money and the mean donation is SEK 8.89 (8.89 percent of the amount allocated to the dictators).

Table 2 contains results from six statistical tests.

Table 2: Statistical test results

Test description	$P-$value
t-test of equal donation	0.45
t-test of equal donation, conditional on giving	0.33
Chi-square test of equal medians	0.88
Chi-square contingency table test that donations are independent of the experimental group	0.26
Chi-square contingency table test that, conditional on giving, donations are independent of the experimental group	0.16
Mann-Whitney test of donations distributions	0.70
Mann-Whitney test of donations distributions, conditional on giving	0.24
Kolmogorov-Smirnov test of donation distributions	0.99
Kolmogorov-Smirnov test of donation distributions, conditional on giving	0.84
Logit analysis of probability donating	0.88

We carry out the tests both for the entire distributions of donations and the distributions conditional on giving (donating something). We fail to reject the null hypothesis of no difference between the experimental groups at the 10 percent level for all the statistical tests.

4. Concluding Remarks

In our dictator experiment according to the standard double blind procedure about a third of the dictators made positive donations, and the mean donation was 13 percent of the total amount allocated. This result is similar to the results of previous studies in the US with the same design, where the mean donation has ranged between 8

percent and 16 percent of the amount allocated (Hoffman et al, [1994]; Hoffman, McCabe and Smith, [1996]; Eckel and Grossman [1996], [1998]).

We also introduced a variation in the double blind design to remove any remaining reciprocity between the dictators and recipients. The anonymous recipients were randomly drawn from the Swedish general population and any donations were mailed to the recipients. Also with this experimental treatment about a third of the dictators made positive donations. The mean donation was 9 percent of the total amount allocated, which is slightly lower than with the standard double blind procedure. This difference was, however, not statistically significant and we could not reject the null hypothesis of no difference between the experimental groups.

We conclude that after removing any possible reciprocity between dictators and recipients, about one third of the dictators still deviated from the standard assumption of pure selfishness and donated some money. We interpret this as evidence of other-regarding behavior not motivated by reciprocity.

5. Appendix

Instructions for "double-blind with randomly drawn recipients from the general population"

You have been asked to participate in an economics experiment. We have paid you SEK 50 in cash for your participation. You may also earn an additional amount of money during the experiment.

In this experiment each of you will be paired with a different person who has been randomly selected from the Swedish population between the ages 18-74 years. You will not be told who these people are, either during or after the experiment. You will notice that there are other people in the room with you who are also participating in the experiment. You will not be paired with any of these people.

One of you will be chosen to be the monitor for the experiment. The monitor will be paid SEK 100 in addition to the SEK 50 already paid. The monitor will be in charge of the envelopes as explained below. In addition the monitor will verify that the instructions have been followed as they appear here.

The experiment is conducted as follows: Twenty-nine unmarked envelopes have been placed in a box. Twenty-seven of these envelopes contain five SEK 20 bills and five blank slips of paper of the same size. The remaining two envelopes contain ten blank slips of paper. The monitor will call one person at a time and hand each person an envelope from the box. The person who was called will take the envelope and go behind the screen at the end of the room. The envelope will then be opened privately behind the screen.

Each person in room A must decide how many bills and how many slips of paper to leave in the envelope. The number of bills plus the number of slips of paper must

add up to five. The person then pockets the remaining dollar bills and slips of paper. Example: (1) Leave SEK 20 and four slips of paper in the envelope and pocket SEK 80 and one slip of paper; (2) Leave SEK 60 and two slips of paper in the envelope and pocket SEK 40 and three slips of paper. These are examples only; the actual decision is up to each person. Also note that no one else, including the experimenter, will know the personal decision made by each person.

Once you have made your decision you will seal the envelope behind the screen and then place it in the box at the front marked "return envelopes". You may then leave the room.

After all twenty-nine envelopes have been returned the monitor will be given another box. This box is marked "addresses" and contains twenty-nine stamped envelopes. Each envelope has been provided with the address to a randomly selected person in the Swedish population between the ages 18-74 years. This person does not know that he/she has been randomly selected for the experiment. In each envelope there is also a letter that describes this experiment. The monitor will then choose one envelope from the box with "return envelopes" and one envelope from the box with "addresses". The monitor will then open the envelope from the box with "return envelopes" and record its contents. If the envelope contains any money the monitor will put the money in the envelope from the box with "addresses" and seal the envelope. If the envelope does not contain any money the addressed envelope will be discarded, i.e. it will not be mailed to the randomly selected person in the population. The monitor will continue until all the envelopes have been opened. The monitor will then take the sealed envelopes with money in and together with the experimenter go to the closest mailbox and mail the envelopes. The experiment is then over.

References

[1] Eckel, C.C., Grossman, P.J. [1996] "Altruism in anonymous dictator games", *Games and Economic Behavior* 16, 181-191.

[2] Eckel, C.C., Grossman, P.J. [1998] "Are women less selfish than men?: Evidence from dictator experiments", *Economic Journal* 108, 726-735.

[3] Forsythe, R., Horowitz, J.L., Savin, N.E., Sefton, M. [1994] "Fairness in simple bargaining experiments", *Games and Economic Behavior* 6, 347-369.

[4] Hoffman, E., McCabe, K., Shachat, K., Smith, V. [1994] "Preferences, property rights, and anonymity in bargaining games", *Games and Economic Behavior* 7, 346-380.

[5] Hoffman, E., McCabe, K., Smith, V. [1996] "Social distance and other-regarding behavior in dictator games", *American Economic Review* 86, 653-660.

[6] Kahneman, D., Knetsch, J., Thaler, R. [1986] "Fairness and the assumptions of economics" In: Hogarth, R.M., Reder, M.W. (Eds.), Rational Choice. University of Chicago Press. Chicago, pp. 101-116.

EFI
The Economic Research Institute

Reports since 1995
A complete publication list can be found at www.hhs.se/efi

Published in the language indicated by the title

2001

Eriksson, R., Price Responses to Changes in Costs and Demand.
Hill, M., Essays on Environmental Policy Analysis: Computable General Equilibrium Approaches Applied to Sweden.
Lange, F. och Wahlund, R., Category Management – När konsumenten är manager.
Liljenberg, A., Customer-geared competition – A socio-Austrian explanation of Tertius Gaudens -
Skoglund, J., Essays on Random Effects model and GARCH.

2000

Berg-Suurwee, U., Styrning före och efter stadsdelsnämndsreform inom kultur och fritid – Resultat från intervjuer och enkät.
Bergkvist, L., Advertising Effectiveness Measurement: Intermediate Constructs and Measures.
Brodin, B., Lundkvist, L., Sjöstrand, S-E., Östman, L., Koncernchefen och ägarna.
Bornefalk, A., Essays on Social Conflict and Reform.
Charpentier, C., Samuelson, L.A., Effekter av en sjukvårdsreform.
Edman, J., Information Use and Decision Making in Groups.
Emling, E., Svenskt familjeföretagande.
Ericson, M., Strategi, kalkyl, känsla.
Gunnarsson, J., Wahlund, R., Flink, H., Finansiella strategier i förändring: segment och beteenden bland svenska hushåll.
Hellman, N., Investor Behaviour – An Empirical Study of How Large Swedish Institutional Investors Make Equity Investment Decisions.
Hyll, M., Essays on the Term Structure of Interest Rates.
Håkansson, P., Beyond Private Label – The Strategic View on Distributor Own Brands.
I huvudet på kunden. Söderlund, M., (red).
Karlsson Stider, A., Familjen och firman.
Ljunggren, U., Styrning av grundskolan i Stockholms stad före och efter stadsdelsnämndsreformen – Resultat från intervjuer och enkät.
Ludvigsen, J., The International Networking between European Logistical Operators.
Nittmar, H., Produktutveckling i samarbete – Strukturförändring vid införande av nya Informationssystem.
Robertsson, G., International Portfolio Choice and Trading Behavior.
Schwarz, B., Weinberg, S., Serviceproduktion och kostnader – att söka orsaker till kommunala skillnader.
Stenström, E., Konstiga företag.
Styrning av team och processer – Teoretiska perspektiv och fallstudier.
Bengtsson, L., Lind, J., Samuelson, L.A., (red).

Sweet, S., Industrial Change Towards Environmental Sustainability - The Case of

Replacing Chloroflouorocarbons.
Tamm Hallström, K., Kampen för auktoritet – Standardiseringsorganisationer i arbete.

1999

Adler, N., Managing Complex Product Development.
Allgulin, M., Supervision and Monetary Incentives.
Andersson, P., Experto Credite: Three Papers on Experienced Decision Makers.
Ekman, G., Från text till batong – Om poliser, busar och svennar.
Eliasson, A-C., Smooth Transitions in Macroeconomic Relationships.
Flink, H., Gunnarsson, J., Wahlund, R., Svenska hushållens sparande och skuldsättning– ett konsumentbeteende-perspektiv.
Gunnarsson, J., Portfolio-Based Segmentation and Consumer Behavior: Empirical Evidence and Methodological Issues.
Hamrefors, S., Spontaneous Environmental Scanning.
Helgesson, C-F., Making a Natural Monopoly: The Configuration of a Techno-Economic Order in Swedish Telecommunications.
Japanese Production Management in Sunrise or Sunset. Karlsson, C., (red).
Jönsson, B., Jönsson, L., Kobelt, G., Modelling Disease Progression And the Effect of Treatment in Secondary Progressive MS. Research Report.
Lindé, J., Essays on the Effects of Fiscal and Monetary Policy.
Ljunggren, U., Indikatorer i grundskolan i Stockholms stad före stadsdels-nämndsreformen – en kartläggning.
Ljunggren, U., En utvärdering av metoder för att mäta produktivitet och effektivitet i skolan – Med tillämpning i Stockholms stads grundskolor.
Lundbergh, S., Modelling Economic High-Frequency Time Series.
Mägi, A., Store Loyalty? An Empirical Study of Grocery Shopping.
Mölleryd, B.G., Entrepreneurship in Technological Systems – the Development of Mobile Telephony in Sweden.
Nilsson, K., Ledtider för ledningsinformation.
Osynlig Företagsledning. Sjöstrand, S-E., Sandberg, J., Tyrstrup, M., (red).
Rognes, J., Telecommuting – Organisational Impact of Home Based – Telecommuting.
Sandström, M., Evaluating the Benefits and Effectiveness of Public Policy.
Skalin, J., Modelling Macroeconomic Time Series with Smooth Transition Autoregressions.
Spagnolo, G., Essays on Managerial Incentives and Product-Market Competition.
Strauss, T., Governance and Structural Adjustment Programs: Effects on Investment, Growth and Income Distribution.
Svedberg Nilsson, K., Effektiva företag? En studie av hur privatiserade organisationer konstrueras.
Söderström, U., Monetary Policy under Uncertainty.
Werr, A., The Language of Change The Roles of Methods in the Work of Management Consultants.
Wijkström, F., Svenskt organisationsliv – Framväxten av en ideell sektor.

1998

Andersson, M., On Testing and Forecasting in Fractionally Integrated Time Series Models.
Berg-Suurwee, U., Styrning av kultur- och fritidsförvaltning innan stadsdelsnämnds-reformen.

Berg-Suurwee, U., Nyckeltal avseende kultur- och fritidsförvaltning innan stadsdelsnämndsreformen.
Bergström, F., Essays on the Political Economy of Industrial Policy.
Bild, M., Valuation of Takeovers.
Charpentier, C., Samuelson, L.A., Effekter av en sjukvårdsreform – en analys av Stockholmsmodellen.
Eriksson-Skoog, G., The Soft Budget Constraint: The Emergence, Persistence and Logic of an Institution. The Case of Tanzania 1967-1992.
Gredenhoff, M., Bootstrap Inference in Time Series Econometrics.
Ioannidis, D., I nationens tjänst? Strategisk handling i politisk miljö – en nationell teleoperatörs interorganisatoriska, strategiska utveckling.
Johansson, S., Savings Investment, and Economic Reforms in Developing Countries.
Levin, J., Essays in Company Valuation.
Ljunggren, U., Styrning av grundskolan i Stockholms stad innan stadsdelsnämndsreformen.
Mattsson, S., Från stat till marknad – effekter på nätverksrelationer vid en bolagiseringsreform.
Nyberg, A., Innovation in Distribution Channels – An Evolutionary Approach.
Olsson, P., Studies in Company Valuation.
Reneby, J., Pricing Corporate Securities.
Roszbach, K., Essays on Banking Credit and Interest Rates.
Runsten, M., The Association Between Accounting Information and Stock Prices. Model development and empirical tests based on Swedish Data.
Segendorff, B., Essays on Bargaining and Delegation.
Sjöberg, L., Bagozzi, R., Ingvar, D.H., Will and Economic Behavior.
Sjögren, A., Perspectives on Human Capital: Economic Growth, Occupational Choice and Intergenerational Mobility.
Studier i kostnadsintäktsanalys. Jennergren, P., (red)
Söderholm, J., Målstyrning av decentraliserade organisationer. Styrning mot finansiella och icke-finansiella mål.
Thorburn, K., Cash Auction Bankruptcy and Corporate Restructuring
Wijkström, F., Different Faces of Civil Society.
Zethraeus, N., Essays on Economic Evaluation in Health Care. Evaluation of Hormone Replacement Therapy and Uncertainty in Economic Evaluations.

1997

Alexius, A., Essays on Exchange Rates, Prices and Interest Rates.
Andersson, B., Essays on the Swedish Electricity Market.
Berggren, N., Essays in Constitutional Economics.
Ericsson, J., Credit Risk in Corporate Securities and Derivatives. Valuation and Optimal Capital Structure Choice.
Charpentier, C., Budgeteringens roller, aktörer och effekter. En studie av budgetprocesserna i en offentlig organisation.
De Geer, H., Silfverberg, G., Citizens´Trust and Authorities´ Choices, A Report from The Fourth International Conference on Ethics in the Public Service, Stockholm June 15-18, 1994.
Friberg, R., Prices, Profits and Exchange Rates.
Från optionsprissättning till konkurslagstiftning. Bergström, C., Björk, T., (red)
Hagerud, G.E., A New Non-Linear GARCH Model.
Haksar, A., Environmental Effects of Economywide Policies: Case Studies of Costa Rica and Sri Lanka.

He, C., Statistical Properties of Garch Processes.
Holmgren, M., Datorbaserat kontrollrum inom processindustrin; erfarenheter i ett tidsperspektiv.
Jennergren, P., Tutorial on the McKinsey Model for Valuation of Companies.
Lagerlöf, J., Essays on Political Economy, Information, and Welfare.
Lange, F., Wahlund, R., Planerade och oplanerade köp - Konsumenternas planering och köp av dagligvaror.
Löthgren, M., Essays on Efficiency and Productivity; Contributions on Bootstrap, DEA and Stochastic Frontier Models.
Nilsson, H.E., Using Conceptual Data Modelling in Management Accounting: A Case Study.
Sjöberg, L., Ramsberg, J., En analys av en samhällsekonomisk bedömning av ändrade säkerhetsföreskrifter rörande heta arbeten.
Säfvenblad, P., Price Formation in Multi-Asset Securities Markets.
Sällström, S., On the Dynamics of Price Quality.
Södergren, B., På väg mot en horisontell organisation? Erfarenheter från näringslivet av decentralisering och därefter.
Tambour, M., Essays on Performance Measurement in Health Care.
Thorén, B., Berg-Surwee, U., Områdesarbete i Östra Hökarängen - ett försök att studera effekter av decentralisering.
Zhang Gang., Chinese Rural Enterprises Between Plan and Market.
Åhlström, P., Sequences in the Profess of Adopting Lean Production.
Åkesson, G., Företagsledning i strategiskt vakuum. Om aktörer och förändrings-processer.
Åsbrink, S., Nonlinearities and Regime Shifts in Financial Time Series.

1996

Advancing your Business. People and Information Systems in Concert.
Lundeberg, M., Sundgren, B (red).
Att föra verksamheten framåt. Människor och informationssystem i samverkan.
red. Lundeberg, M., Sundgren, B.
Andersson, P., Concurrence, Transition and Evolution - Perspectives of Industrial Marketing Change Processes.
Andersson, P., The Emergence and Change of Pharmacia Biotech 1959-1995. The Power of the Slow Flow and the Drama of Great Events.
Asplund, M., Essays in Industrial Economics.
Delmar, F., Entrepreneurial Behavior & Business Performance.
Edlund, L., The Marriage Market: How Do You Compare?
Gunnarsson, J., Three Studies of Financial Behavior. Research Report.
Hedborg, A., Studies of Framing, Judgment and Choice.
Holgersson, C., Höök, P., Ledarutveckling för kvinnor - Uppföljning av en satsning på Volvo - Research Report.
Holmberg, C., Stores and Consumers - Two Perspectives on Food Purchasing.
Håkansson, P., Wahlund, R., Varumärken. Från teori till praktik.
Karlsson, A., The Family Business as an Heirloom. Research Report.
Linghag, S., Man är handelsstudent. Research Report.
Molin, J., Essays on Corporate Finance and Governance.
Mägi, A., The French Food Retailing Industry - A Descriptive Study.
Mölleryd, B., Så byggdes en världsindustri - Entreprenörskapets betydelse för svensk mobiltelefoni. Research Report.
Nielsen, S., Omkostningskalkulation for avancerede produktions-omgivelser - en

sammenligning af stokastiske og deterministiske omkost-ningskalkulationsmodeller.

Normark, P., Danielsson, L., Larsson, A., Lundblad, P., Kooperativa nyckeltal. Research Report.
Sandin, R., Heterogeneity in Oligopoly: Theories and Tests.
Sandén-Håkansson, U., Från kampanjmål till mediemix - en studie av samarbete mellan annonsörer, reklambyråer och mediebyråer. Research Report.
Stein, J., Söderlund, M., Framgång i arbetet, strategier för att utföra arbetsuppgifterna, arbetsuppgifternas karaktär och utbildningskvalitet. En empirisk studie av civil-ekonomer. Research Report.
Strömberg, P., Thorburn, K., An Empirical Investigation of Swedish Corporations in Liquidation Bankruptcy. Research Report.
Söderlund, M., Och ge oss den nöjda kunden. En studie av kundtillfredsställelse och dess orsaker och effekter. Research Report.
Thodenius, B., Användningen av ledningsinformationssystem i Sverige: Lägesbild 1995. Research Report.
Ulfsdotter, U., Internationalisering för expansion eller hemmamarknadsförsvar? De nordiska marknaderna för fruktyoghurt 1982-1994.
Westelius, A., A Study of Patterns of Communication in Management Accounting and Control Projects.
Wijkström, F., Den svenska ideella sektorn och pengarna. Research Report.
Örtendahl, M., Health and Time - A Problem of Discounting.

1995

Becker, T., Essays on Stochastic Fiscal Policy, Public Debt and Private Consumption.
Blomberg, J., Ordning och kaos i projektsamarbete - en socialfenomenologisk upplösning av en organisationsteoretisk paradox.
Brodin, B., Lundkvist, L., Sjöstrand, S-E., Östman, L., Styrelsearbete i koncerner
Brännström, T., Bias Approximation and Reduction in Vector Autoregressive Models. Research Report.
Ekonomisk politik i omvandling. Jonung, L (red).
Gunnarsson, J., Wahlund, R., Hushållens finansiella strategier. En explorativ studie.
Höök, P., Chefsutveckling ur könsperspektiv - Mentorskap och nätverk på Vattenfall - Research Report.
Levin, J., Olsson, P., Looking Beyond the Horizon and Other Issues in Company Valuation. Research Report.
Mägi, A., Customer Statisfaction in a Store Performance Framework. Research Report.
Nittmar, H., Produktutveckling i samarbete.
Persson, P-G., Modeling the Impact of Sales Promotion on Store Profits.
Roman, L., Institutions in Transition. A Study of Vietnamese Banking.
Sandberg, J., Statistisk metod - dess vetenskapliga hemvist, grundläggande principer och möjligheter inom samhällsvetenskapen. Research Report.
Sandberg, J., How Do We Justify Knowledge Produced by Interpretative Approaches? Research Report.
Schuster, W., Redovisning av konvertibla skuldebrev och konvertibla vinstandelsbevis - klassificering och värdering.
Söderberg, K., Farmartjänsten - Ny kooperation inom lantbruket. Research Report.
Söderqvist, T., Benefit Estimation in the Case of Nonmarket Goods. Four Essays on Reductions of Health Risks Due to Residential Radon Radiation.
Tamm Hallström, K., Kampen för auktoritet – standardiseringsorganisationer i arbete.
Thorén, B., Användning av information vid ekonomisk styrning - månadsrapporter och andra informationskällor.