

Industrial targeting, experimentation and long-run specialization

Mikhail M. Klimenko*

*Graduate School of International Relations and Pacific Studies, University of California,
San Diego, La Jolla, CA 92093, USA*

Received 1 June 1999; accepted 1 September 2002

Abstract

This paper emphasizes the experimental nature of industrial targeting policies under uncertainty in a small open economy. A government promotes entry of new firms in selected industries and updates its beliefs about the country's comparative advantage in a Bayesian way. This selective targeting policy is analyzed in the framework of a special type of statistical decision problem known as the multi-armed bandit. The paper analyzes how the costs and benefits of learning about a country's comparative advantage depend on the characteristics of the targeted industries. The framework suggests that even an optimally designed industrial targeting policy may eventually steer the country away from specializing according to its true comparative advantage.

© 2003 Elsevier B.V. All rights reserved.

JEL classification: F13; D83; C44; O14

Keywords: Industrial targeting; Bayesian learning; Multi-armed bandits

1. Introduction

The practice of selective promotion of industries deemed critical for accelerating economic growth was popular among the governments of many emerging economy nations throughout the second half of the 20th century. Such industrial targeting policies can be justified by the existence of some market failure that makes government nonintervention suboptimal. For example, earlier literature on strategic trade policy showed that targeting policies can be welfare-enhancing when industries are characterized

* Tel.: +1-858-534-5014; fax: +1-858-534-3939.

E-mail address: mklimenko@ucsd.edu (M.M. Klimenko).

by imperfect competition or information asymmetries.¹ State intervention in selected industries can also be justified by the existence of multiple equilibria leading to coordination problems in economic development.² Another reason for industrial targeting policy may be a less than socially desirable rate of entry of new firms in the targeted industries caused by non-appropriability of external benefits from innovation and investment.³

Important foci of the industrial targeting literature has been on the timing and type of policy interventions as well as on the specifics of the market failures justifying such interventions on efficiency grounds. The problem that has received less attention in the literature is the government's ability to allocate its limited targeting capacity among the industries that deserve intervention, given that the effectiveness of the targeted interventions is uncertain. In this paper, I consider the characteristics of the process by which the government selects the most worthy among all industries in which policy interventions are assumed to be justified on the grounds of welfare efficiency. Specifically, this paper emphasizes the role of policy-makers' learning and experimentation in promoting the entry in new industries, and considers the long-term outcome of such experimental policy activism.

That experimentation is an important characteristic of industrial targeting and export promotion policies has been made evident in the application of these policies worldwide. In Japan, Ministry of International Trade and Industry (MITI) adopted the policy of capacity-licensing with staggered and contingent entry to ensure sequential entry of firms in new industries in the late 1950s. Similar policies were adopted by the Industrial Development Bureau of the Ministry of Economic Affairs in Taiwan and the Ministry of Commerce and Industry in Korea between the 1950s and 1970s. Formally, the rationale for sequential capacity-licensing and technology-licensing was to ensure minimum efficient scale of operation for the early entrants and to force them to innovate by creating a credible

¹ Brander and Spencer (1985) and Spencer and Brander (1983) analyze models that demonstrate how government intervention in imperfectly competitive industries allows a country to extend its market share and thereby "capture" more rents. Grossman and Horn (1988) and Bagwell and Staiger (1989) employ models of *adverse selection* to explore the possibility that infant-industry protection or export subsidies may be welfare-improving in the markets for experience goods, the quality of which can be learned only after the purchase. Dinopoulos et al. (1995) uses techniques from the literature on contracting under asymmetric information to examine the optimal industrial targeting when domestic firms have private knowledge of their learning economies.

² For example, in vertically integrated industries, miscoordination in the entry of upstream and downstream firms may lead to multiple equilibria. Industrial targeting can be used to steer the economy away from the bad equilibria by stimulating entry of new firms producing essential intermediate inputs and complementary producer services. Potential coordination problems inherent in economic development are analyzed, for example, in Azariadis and Drazen (1990), Matsuyama (1995b), and Rodrik (1996).

³ Firms' entry may be slower than optimal because early entrants do not capture the external benefits that they create. These external benefits occur either as a result of increasing the total output of the industry (in adaptive-learning models) or because every new entry reduces the uncertainty about the publicly observed productivity or cost parameter common to all firms in the industry (in Bayesian learning models). A government subsidy can help to overcome firms' insufficient incentives for entry. Bardhan (1971) analyzes the optimal subsidy to a learning industry in the context of infant-industry protection in a model with dynamic externalities in production.

threat of new entry after a prespecified period of time. However, another important characteristic of the policy, based on sequential and contingent licensing, was flexible adaptation of the main directions and intensity of the policy based on feedback obtained from the continual monitoring of the performance of the targeted industries.⁴

The agencies responsible for industrial policy often experimented by concentrating their targeting efforts on new industries in which they had little experience, at the expense of promoting entry in industries in which their countries were believed to have comparative advantage. Such experimental policy activism was not insured against mistakes, and many instances of “picking the losers” instead of winners can be cited from the experience of even the most successful industrial policies worldwide. However, the long-run success of any policy is determined by the policy-maker’s ability to learn from past mistakes.⁵

To focus on the special concerns raised by the experimental nature of the government’s targeting policy, this paper does not consider the specific market failure that could prompt government intervention. I simply assume that an investment subsidy intended to facilitate the firms’ entry is justified by the presence of failures in multiple markets for new goods in a small developing country.⁶ Similar to Romer (1994), foreign multinational firms are the main source of technology for the production of new goods in this country.⁷ The multinationals’ technologies have to be matched with local labor and the quality of the match is a random variable with unknown probability distribution. The government’s limited targeting capacity and the uncertainties associated with the promotion of new economic activities mean that the welfare-maximizing targeting policy has to include elements of experimentation and learning.

The problem faced by a government trying to allocate its limited capacity “to pick the winners” among the industries based on foreign technologies with uncertain characteristics can be analyzed within a statistical decision framework known as the multi-armed bandit (MAB).⁸ Using the MAB framework, the paper examines how the economic

⁴ Kuo (1983), Krueger (1985), Pack and Westphal (1986), Wade (1990), Rodrik (1995, 1996), Wah and Jomo (1995), Lall (1996), and Kim and Ma (1997) provide multiple examples of experimentation in industrial targeting policies by governments in East Asia.

⁵ As Stiglitz points out, “good decisionmaking by the government necessarily involves making mistakes: a policy that supported only sure winners would have taken no risks. The relatively few mistakes speak well for the government’s ability to pick winners” (see Stiglitz, 1996).

⁶ While an investment subsidy is not the only instrument of targeting policy available to the government, it is widely used in modeling state intervention in the contexts involving introduction of new goods. For example, Aizenman (1998) shows that the investment subsidy to new entrants is optimal when introduction of new inputs creates externalities not captured by the entrants and the entry conditions are uncertain. Hoff (1997) considers a two-period model of infant industry protection in which a subsidy is used as a means of overcoming informational barriers to entry.

⁷ However, in contrast with the Romer’s (1994) model, which considers multinationals investing in the production of intermediate goods used in local economy, in the present framework foreign firms invest in the country to take advantage of its low-cost labor and produce *final* goods that are exported abroad.

⁸ I use the MAB framework because it is the most convenient and simple way to illustrate the main ideas of this paper. The classical MAB framework considers stochastic reward processes, or ‘arms’, with different unknown success probabilities. The arms are ‘pulled’ one at a time and the objective is to achieve a high overall success rate by selecting at each stage the next arm to be pulled on the basis of current information about the different probabilities for success.

characteristics of the new industries affect their selection for targeted promotion. Specifically, I analyze how the value of experimentation with promoting a given technology depends on such parameters as the set-up costs, the variance of the unknown productivity characteristic and the size of the ‘experimentation field’ determined by the number of potential entrants (i.e., entrepreneurs) that are willing to test the new foreign technology in the country’s economic environment.

After describing an industrial targeting policy as an optimal experimentation strategy of an incompletely informed government, I turn to an examination of the set of industries in which a country will specialize in the long-run as a result of such policy. I show that for any set of targeted industries, it is possible to determine whether a country will specialize in this set in the limit with positive or zero probability. The model considered in this paper highlights the factors that determine which technologies will not survive the targeting policy in the long-run (i.e., will survive only with zero probability).

The industrial targeting policy not only helps the government to mitigate market failures, but also transforms the decentralized market learning process into the process of experimentation by a single policy-maker, who selectively influences the firms’ entry and exit decisions in the targeted industries. In Section 5 of this paper, I show that an optimally designed industrial targeting policy may lead in the long-run to the country’s specialization in industries in which it does not have comparative advantage. Depending on the policy-maker’s prior beliefs, it may be that the country with probability one eventually abandons the industries in which it has true comparative advantage. When this happens the country will specialize eventually in the ‘wrong’ or inferior technologies, because these technologies will perform well enough not to induce the government to look for better ‘targets’. This long-run divergence of the government’s optimal experimentation strategy from the full-information optimal policy can be interpreted as a failure of the industrial targeting policy. In this regard, I advance arguments in support of the conjecture that, despite the concomitant market failures, the outcome of market learning process through private ‘unassisted’ experimentation by firms may, under certain conditions, be closer to the full-information social optimum than the centralized experimentation by a government engaged in industrial targeting. Although my analysis of the long-run results of government intervention under uncertainty is based on the familiar possibility of incomplete learning by an experimenting decision-maker,⁹ this is the first paper that shows that such a possibility is relevant in the discussion of the role of industrial policy in promoting economic development.

On the technical side, the main contribution of this paper is that it analyzes the long-run stochastic properties of the experimentation strategy in the setting where the decision-maker can experiment with more than one alternative reward process at a time (i.e., can target more than one industry at a time), and each alternative can be of a continuum of types. A similar issue was studied by [Banks and Sundaram \(1992\)](#). However, they restricted their analysis to the special case of the multi-armed bandit framework in which the gambler has the capacity to pull only a single arm at a time and the arms can be of a finite number of types. The technical results of this paper make the present model of

⁹ In economics, this possibility is known as the Rothschild effect (see [Rothschild, 1974](#)).

industrial targeting and potentially many other economic applications of the MAB framework significantly more realistic by allowing them to assume that the policy-maker can engage in parallel experimentation with many risky projects and by expanding the range of probability distributions representing the experimenter's uncertainty.

Although the issue of optimal experimentation has never been explored in the context of industrial policy, it has been studied in other economic contexts. There is extensive theoretical literature on optimal Bayesian learning in uncertain economic environments when an individual agent's action not only provides immediate reward, but also delivers information leading to the reduction of uncertainty. This information may be valuable for selecting future actions. For example, a monopolist faced with an uncertain demand curve may decide to produce to maximize expected profits based on his current beliefs, or to vary his output from period to period deliberately to accumulate information about the location of the demand curve. Rothschild (1974), McLennan (1984), and Aghion et al. (1991) study problems of this type.¹⁰

Another strand of literature closely related to the present paper is based on rational expectations models. Anderson and Sonnenschein (1982), Blume and Easley (1984), Jordan (1985), Easley and Kiefer (1988), and Banks and Sundaram (1992) explore properties of the learning process and conditions under which economies generate enough information to assure that optimally acting agents eventually learn the true distributions of the unknown parameters in the economy. This issue motivated the current study of the long-run results of government experimentation with industrial targeting in a small open economy.

The remainder of the paper is organized as follows. Section 2 introduces the basic model and describes the government's objective function and the environment in which the government conducts the industrial targeting policy. Section 3 formulates the government's problem of selective industry promotion using the multi-armed bandit framework. In Section 4, the paper turns to the properties of the government's optimal industrial targeting strategy, which is based on promoting the industries with the highest dynamic allocation indices. Section 5 analyzes the country's long-run pattern of export specialization, which results from the government's optimal targeting policy. Some open questions are discussed in Section 6, which concludes the paper. Proofs of results that are omitted in the main body of the paper can be found in Appendix A.

2. The basic model

In this section, I consider a *highly stylized* model of industrial targeting policy in a small open economy, which takes the world prices as given. Specifically, assuming a two-country world, we can identify the small country as South and its trade partner as country North. South's terms of trade are determined by the supply and demand conditions in the rest of the world, i.e., in North.

¹⁰ The analytical framework of my paper is related to models of optimal sequential investment decisions under uncertainty about future payoffs. For example, Weitzman (1979), Roberts and Weitzman (1981), Bernanke (1983), and Dixit and Pindyck (1994) examine projects with sequential outlays using models that stress the trade-off between the costs of waiting for the arrival of new information and taking uninformed actions immediately.

North has a set of N competitive industries $\mathcal{N}=\{1, \dots, N\}$ characterized by increasing marginal cost technologies. The technologies are employed by risk-neutral firms or entrepreneurs. The total costs of a firm, which produces output Q_i using technology i in country North, is defined by $C_i^*(Q_i)=K_i+w^*\gamma_i^*Q_i^2$.¹¹ (An asterisk will always indicate parameters specific to North.) γ_i^* is a cost/productivity parameter representing the quality of the match of technology i with labor in North. The fixed cost K_i consists of capital while the variable cost $w^*\gamma_i^*Q_i^2$ consists of the wages w^* paid to hired labor. In competitive industries, profits are equal to zero and, therefore, prices equal average costs. Because firms maximize profits, the industry price is also equal to marginal cost.¹² Thus, by setting the average cost AC_i^* equal to the marginal cost MC_i^* , I can determine the equilibrium price $\theta_i^*=2(K_i/w^*\gamma_i^*)^{1/2}$ and output $Q_i^*=K_i^{1/2}(w^*\gamma_i^*)^{-1/2}$ in industry i in North.

Although South has lower wage than North, $w^*>w$, there is uncertainty regarding the quality of the match between the skills of local labor in South and technologies $i \in \mathcal{N}$ transplanted from North. This uncertainty is captured by the random cost/productivity parameter $\tilde{\gamma}_i$, which is drawn from the same distribution every time a new firm with technology i is established in South. The cost parameter, $\tilde{\gamma}_i$, can have either low realization (high productivity) γ_i^{low} with some objective probability p_i or high realization (low productivity) γ_i^{high} with probability $(1-p_i)$. The cost parameter enters the total cost equation for a firm with technology i in South the same way as for a firm with the same technology in North: $C_i(Q_i)=K_i+w\tilde{\gamma}_iQ_i^2$. The probability p_i measures the quality of the match between technology i and local labor in South. The South's government as well as new firms entering with technology i in country South have only imperfect information about p_i . Their beliefs about p_i are represented by a probability distribution of the random variable \tilde{p}_i defined on the interval $[0,1]$.

Suppose that there are \bar{n}_i entrepreneurs wishing to establish new firms using technology i in South. The domestic market for good i in country South is very small and the transportation costs between the countries are negligible. Therefore, it only makes sense to produce in South in order to sell in North. If a firm decides to enter in industry i in South, it

¹¹ The assumption of quadratic total cost function is not essential. The main results of the paper can be extended to other forms of increasing marginal costs.

¹² Perfect competition is assumed in order to simplify the exposition and to focus on experimentation and learning in industrial targeting policies. Therefore I abstract from the specifics of market failures that cause policy intervention. The actual market failures and policy remedies against them may vary across targeted industries. In a study focused on specific historical incidents of government intervention in market mechanism, my analytical framework can be extended to incorporate specific market failures. One possible extension is to merge my model of experimental targeting policies with Romer's (1994) framework in which imperfectly competitive producers of differentiated intermediate goods have to make an up-front capacity investment before they can face the downward sloping demand curve from the final good industry. The entry of intermediate good suppliers increases the productivity of the final good industry and continues until the cost of up-front capacity investment exceeds the entrants' ex post monopoly revenue. In this setting, the government can target different vertically integrated industries by promoting the entry of alternative groups of suppliers producing intermediate goods specific to the targeted final good. The index-based approach to identifying the optimal experimentation strategy discussed in Section 3 of this paper can still be applied in this setting. This extension can be used to analyze the Japanese and Korean governments' experimentation in "picking the winners" among different industries with scale economies. I thank a referee for pointing this out.

will have to sell its product at the price prevailing in that industry in the larger country North. I assume that the number \bar{n}_i is small relative to the total number of existing firms in industry i in North. Therefore, the exports of \bar{n}_i new firms from South would not affect the domestic competitive price in North. Given the price in North, the output and the expected profit of a risk-neutral firm i producing in South and exporting to North can be determined from the expected profit maximization equation:

$$\max_{Q_i} E(\tilde{\Pi}_i) = \max_{Q_i} \left[2(K_i w^* \gamma_i^*)^{\frac{1}{2}} Q_i - K_i - w \bar{\gamma}_i Q_i^2 \right] \tag{1}$$

where $\bar{\gamma}_i = E(\tilde{\gamma}_i) = E(\tilde{p}_i \gamma_i^{\text{low}} + (1 - \tilde{p}_i) \gamma_i^{\text{high}}) = \bar{p}_i \gamma_i^{\text{low}} + (1 - \bar{p}_i) \gamma_i^{\text{high}}$.

New firms entering in country South make production plans based on their expectations of the quality of the match between their technologies and the South’s labor force. The planned output of a firm i in South is:

$$Q_i = (K_i w^* \gamma_i^*)^{\frac{1}{2}} (w \bar{\gamma}_i)^{-1} \tag{2}$$

The planned output defined in Eq. (2) is produced over the lifetime of the firm with technology i in South. However, the random match-quality parameter $\tilde{\gamma}_i$ is realized by the end of the first period after the firm makes its investment in South. Since the operation of the firm after the first period does not bring any new information about the match between technology i and labor in South, we can simplify the exposition without any loss of generality by assuming that the firm’s entire planned output is produced during one period. At the end of the period, the information about the firm’s total cost and, therefore, the realization of random variable $\tilde{\gamma}_i$ become public, and the beliefs about \bar{p}_i are updated according to Bayes rule. The firm’s profit and the expected profit are given by

$$\tilde{\Pi}_i = K_i [\tilde{a}_i - 1] \text{ and } E(\tilde{\Pi}_i) = K_i [\bar{a}_i(\bar{p}_i) - 1] \tag{3}$$

where the random variable \tilde{a}_i and its expected value $\bar{a}_i(\bar{p}_i)$ are defined respectively by

$$\tilde{a}_i = \frac{w^* \gamma_i^*}{w \tilde{\gamma}_i} \left(2 - \frac{\tilde{\gamma}_i}{\bar{\gamma}_i} \right) \text{ and } \bar{a}_i(\bar{p}_i) = E[\tilde{a}_i | p_i = \bar{p}_i] = \frac{w^* \gamma_i^*}{w \bar{\gamma}_i}$$

Assuming that entrepreneurs with access to Northern technologies have already established firms in country North, they would have incentives to establish firms in South only if these new firms can generate nonnegative profits net of taxes or license fees imposed by the South’s government. If a new firm in South is expected to be unprofitable, to attract a new firm the South’s government must offer it a subsidy equal to the expected loss.¹³

¹³ To avoid the time inconsistency problem of the “obsolescing bargain” (see Brander and Spencer, 1987) in which the relative bargaining powers of the government and the new firm change systematically over the life of an agreement, I assume that the Southern government can credibly precommit itself to the announced level of investment tax or subsidy. Therefore, the Southern government in my model does not exploit the change in the relative bargaining position arising after a firm has incurred the sunk costs.

Therefore, the welfare of country South from the entry of a firm with technology i is defined by the tax revenue collected from (or subsidy paid to) that firm:

$$\tilde{W}_i = K_i[\tilde{a}_i - 1] \tag{4}$$

Then, given the expectation of the probability of the good match \bar{p}_i , the welfare-maximizing risk-neutral government of country South has the immediate expected welfare reward function:¹⁴

$$\bar{W}_i(\bar{p}_i) = K_i[\bar{a}_i(\bar{p}_i) - 1] \tag{5}$$

The expected value, $\bar{a}_i(\bar{p}_i) = (w^* \gamma_i^*) / (w \bar{\gamma}_i)$, can be interpreted as a measure of comparative advantage of South in industry i . Given the prior beliefs, South has comparative advantage over North in all industries characterized by $\bar{a}_i(\bar{p}_i) \geq 1$. The Southern government expects to receive positive welfare reward from targeting these industries. In the remainder of the paper, I refer to $\bar{a}_i(\bar{p}_i)$ as the coefficient of comparative advantage.

Finally, the following assumption is made with respect to the nature of the process of selecting the industries for targeted promotion. The Southern government has the capacity to license at most M new firms ($M \leq N$) from M different industries at a time. This can be explained by a two-tier structure of the targeting process. Each tier is endowed with a limited amount of a critical resource required for the evaluation and selection of new projects/entrants. At every moment of time, each of the N industry-specific agencies of the lower tier has enough capacity to evaluate at most one new project. After that the single inter-industry agency of the higher tier has the capacity to evaluate and select at most M ($M \leq N$) from N new projects selected at the lower tier. Given the described decision-making process, the targeted set at time t includes M new firms from M different industries. Using the two-tier decision structure, the South’s government tries to allocate its targeting capacity to those industries that would allow it to maximize the total expected welfare from the entry of new firms.¹⁵

Although the assumption about the structure of the targeting process might appear to be very restrictive, it reflects the characteristics of the government decision-making that was practiced in many of the newly industrializing countries when they were implementing industrial targeting policies. For example, in high-performing Asian economies in the 1950s and the 1960s, the role of the higher-tier agency was played by different inter-

¹⁴ Tax or licensing revenue is not the only way to represent the welfare gains from the entry of targeted firms. For example, in some newly industrializing economies, industrial policies were conducted in the environment with labor-market rigidities that fixed wages above the market clearing level. For these economies, reduction of unemployment could be an important factor in selection of industries for targeted promotion. If we introduce unemployment in our model, the welfare that South would receive from the entry of a new firm i will be the sum of the collected tax (or subsidy) and wages paid by firm i to the newly employed local workers: $\tilde{W}_i = E(\tilde{\Pi}_i) + w \bar{\gamma}_i (K_i w^* \gamma_i^*) (w \bar{\gamma}_i)^{-2}$. The expected welfare reward will then be given by $\bar{W}_i(\bar{p}_i) = K_i [2w^* \gamma_i^* / w \bar{\gamma}_i - 1]$, which is very similar to Eq. (5). Therefore, the introduction of unemployment in our model would not change the qualitative results of this paper concerning experimentation and learning in targeting policy.

¹⁵ Although the targeting capacity is assumed to be fixed over time, a welfare-maximizing government should be able to expand it by increasing the size of the bureaucracy targeting new entrants. However, it will be much harder analytically (although not impossible) to demonstrate the validity of the main results of this paper for a government with varying targeting capacity.

industry ministries such as the MITI in Japan, the Ministry of Commerce and Industry in Korea, and the Industrial Development Bureau in Taiwan. These agencies designed development programs for their economies by strategically allocating subsidies and credits among the projects recommended by the lower-tier industry-specific ministries and deliberation councils composed of government officials and representatives from the private sector.¹⁶

3. The government's optimal targeting policy

The Southern government's task of maximizing the discounted flow of welfare rewards while trying to target industries that have the best match with local labor can be formulated as a 'multi-armed bandit' (MAB) problem. In statistical decision theory, the MAB framework describes a Markov decision problem in which a gambler must find an optimal strategy against a gambling machine with N independent arms and has the option to pull M arms each time. M is referred to as the capacity of the gambler. The reward, which the gambler receives for pulling an arm, is a random variable that is a function of a stochastically evolving state.¹⁷ Since a detailed discussion of the MAB framework can be found in Mandelbaum (1986), which uses the same formulation of the problem as I do (i.e., one of controlling a process with a multidimensional time parameter), I provide here only a brief outline of the main elements of this framework and place it into the context of the government's industrial targeting policy.

The state variable in the context of industrial targeting is defined by the parameters of the posterior distribution of the probability of a good match between a firm with technology i and skills of Southern workers. I assume that the prior beliefs about the probability of the good match \tilde{p}_i are represented by the beta probability density function with the parameters (r_i, n_i) . Therefore, after any sequence of n_i' trials (i.e., n_i' entries of new firms with technology i in South) with r_i' successes (i.e., r_i' realizations of good match between local workers and technology i), the posterior beliefs will be the beta distribution with parameters $(r_i + r_i', n_i + n_i')$.¹⁸ Using this property of the beta p.d.f., one

¹⁶ See Aoki et al. (1997) on the role of deliberation councils in government-market coordination in East Asian High Performing economies, and Kim and Ma (1997) on the role of MITI and the Petrochemical Cooperative Discussion Group in the development of the petrochemical industry in Japan in the 1950s.

¹⁷ Bellman (1956) derived the dynamic programming solution to the discrete-time version of the two-armed bandit problem in which the decision-maker can pull only one arm at a time. However, Bellman's solution was extremely complicated. Later, Gittins and Jones (1974) found more simple and elegant solution to the problem. They proved that if the states of the arms evolve in a Markovian fashion, it is possible to associate each arm with a function of its state, now called the Gittins index, such that the gambler can maximize the total discounted reward by sequentially engaging the arms with the maximal current Gittins index. These ideas are extended in Varaiya et al. (1985) and Mandelbaum (1986) where the assumption about the Markovian structure of reward is relaxed by assuming arbitrary reward processes. Banks and Sundaram (1992) considered the case of a denumerable-armed (i.e., countably infinite-armed) bandit.

¹⁸ The mean and the variance of any beta probability distribution with parameters (r, n) are given by r/n and $(r/n)(1 - r/n)/(n + 1)$, respectively. A uniform $U[0,1]$ prior distribution of \tilde{p}_i is equivalent to the beta distribution with the parameters (1,2) and the mean 1/2. After a successful trial the mean of the posterior beliefs becomes 2/3, etc.

can simplify the notation by identifying the state $x_i(t)$ of the stochastic reward process for technology i at time t with the number of high-productivity firms in the total number of new entrants with technology i by time t . Therefore, after r'_i successful matches out of n'_i new entries of firms with technology i in South, the state variable is defined by¹⁹

$$x_i(t) = (r_i + r'_i, n_i + n'_i)$$

The control variable for the optimal targeting problem $\Delta_i(t)$ is equal to one if the government licenses a new firm with a Northern technology i at time t , and $\Delta_i(t)$ is zero if no new firm with technology i is licensed in country South at time t .

The control sequence of N -dimensional vectors $\{\Delta(t)\}_{t=0}^{\infty}$ represents the choice of industries that are targeted as time progresses by licensing new firms with respective technologies. In choosing between a firm with a technology in which South is believed to have a comparative advantage (i.e., characterized by a high \bar{p}_i) and a firm from an industry in which it has little experience (and so a poor estimate of the true probability of the good match), the government experiences a conflict between exploitation and exploration. From the point of view of the immediate expected reward, it would be better to license an investment by the first firm. However, if the economy's long-term performance is important (i.e., the discount factor is not too small), it may be preferable to test the second industry, which may turn out to have a higher productivity than the first. Therefore, from the long-term point of view, the acquisition of additional information may be worth more than the immediate yield and experimentation may be justified.

The optimal industrial targeting policy maximizes the expected discounted stream of investment tax (or license fee) revenues received by the Southern government, i.e., the limit of the sequence of partial sums $E \sum_{i=1}^N \sum_{t=0}^{\infty} \beta^t \tilde{W}_i(x_i(t)) \Delta_i(t)$.²⁰ According to a well-known result for the MAB problem, the optimal policy for the government targeting only one industry at a time ($M=1$) is to target at every period of time the industry with the greatest current Gittins index $v_i(x_i)$. The industry-specific Gittins index, $v_i(x_i)$, also known as the dynamic allocation index (DAI), represents the amount that makes the policy maker indifferent between the options of continuing the targeting of industry i and abandoning that industry forever in exchange for the guaranteed one-time reward $v_i(x_i)$.²¹

¹⁹ Another common set-up for the MAB problem is the one in which the decision-maker perceives the reward as a random variable having a continuous distribution (e.g., a normal random variable) with unknown mean and known variance. See, for example, Lai (1988). Benkherouf et al. (1992) use the MAB framework to model oil exploration by a company which can drill sequentially in several areas with unknown amount of oil. They assume that the company's priors are represented by either Euler or Heine families of probability distributions.

²⁰ To ensure the existence of the sequence's limit, I assume that expected total discounted reward from targeting any of the industries is finite: $E \sum_{i=1}^{\infty} \beta^t | \tilde{W}_i(x_i(t)) | < \infty$ for $i \in \mathcal{N}$.

²¹ The formal definition of the Gittins index is given in the Proof of Proposition 1 in Appendix A. The optimal policy for the Southern government is to license at each stage a firm in the industry that has the highest index $v_i(x_i)$ greater than zero. If there are no $v_i(x_i)$ greater than zero for $i \in \mathcal{N}$, then there are no industries in the set \mathcal{N} that have positive experimentation value for the South's government given its priors about the comparative advantage of country South in these industries.

Formally, the optimal strategy for $M=1$ is defined by:

$$\Delta_i(t) = \begin{cases} 1 & \text{if } v_i(x_i(t)) = \max_{1 \leq j \leq N} \{v_j(x_j(t))\} \\ 0 & \text{if otherwise} \end{cases} \quad (6)$$

It is known that for the multi-armed bandit problem, the policy of selecting the arms with the M largest Gittins indices may be suboptimal.²² Therefore, the optimal strategy for a version of the targeting problem, in which the policy-maker can at each time select up to $M>1$ industries, cannot be described in terms of the Gittins indices.

However, the Gittins index policy for the MAB problem with $M>1$ provides a computationally simple heuristic, which in the long-run is almost optimal both in terms of the value it generates (i.e., asymptotically optimal) and in terms of the actions it prescribes (i.e., turnpike optimal).²³ Therefore, by following the Gittins index heuristic over a long horizon, the policy-maker will choose optimal actions most of the time.

In the context of industrial targeting policy, asymptotic optimality and turnpike optimality of the Gittins index strategy suggest that the government, which has the targeting capacity $M>1$, can satisfactorily solve a problem for which the optimal solution is intractable. Therefore, I assume that the government conducting the industrial targeting policy uses a simple, albeit only almost optimal, index-based strategy rather than the computationally complex optimal strategy.²⁴

4. Properties of the experimental industrial targeting policy

The value of experimentation in industrial targeting is reflected in the option-like properties of the Gittins dynamic allocation indices determining the order in which the government “picks the winners” among industries. The index increases with the variance of reward, very much like the value of a call option increases with the variance of the underlying stock. Therefore, through their effect on the variance of the welfare reward both an increase in the fixed cost K_i and an increase in uncertainty about the quality of the match between the foreign technology and local labor lead to an increase in the dynamic

²² The version of the MAB problem with $M>1$ belongs to the class of \mathcal{NP} -complete problems which is characterized by the extreme computational complexity of finding an optimal strategy. The problems of this class are usually solved by enumerative methods such as the branch-and-bound method. However, a simple heuristic whose performance is arbitrarily close to optimal strategy often exists even in those cases in which the problem of finding an optimal solution has been proved \mathcal{NP} -hard.

²³ Whittle (1988) and Weber and Weiss (1990) were the first to propose the index policy heuristic for the multi-armed bandit problems in which a gambler can pull more than one arm at a time. The asymptotic optimality of the Gittins index policy for $M>1$ was demonstrated by Ishikida and Varaiya (1994). The turnpike optimality requires that in the long-run the fraction of decision times during which the prescription of the index policy contravenes the prescriptions of the optimal policy is zero. The approach developed by Weiss (1994) for optimal scheduling of stochastic jobs on parallel servers can be used to prove the turnpike optimality of the Gittins index policy for $M>1$.

²⁴ Schmalensee (1975) advocated consideration of asymptotically optimal learning strategies in those economic contexts where such strategies are more behaviorally plausible than the optimizing behavior, which can be ruled out based on its complexity.

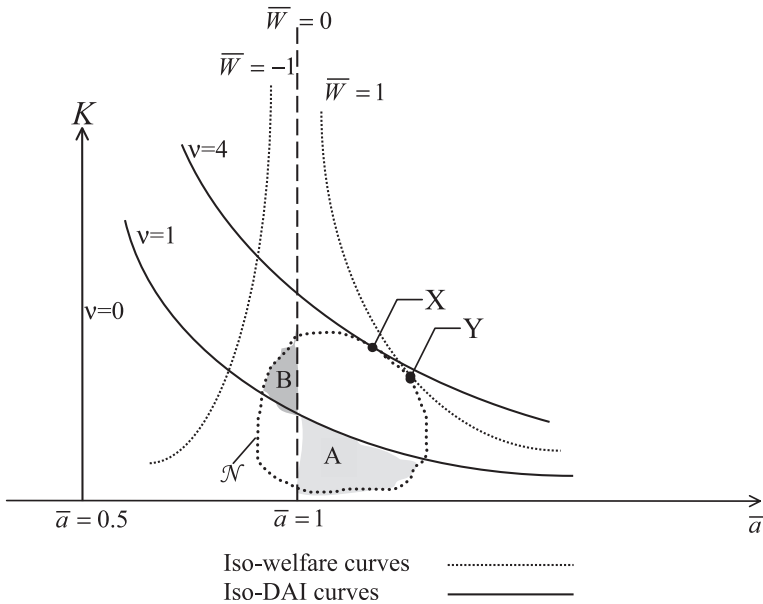


Fig. 1. Iso-welfare and iso-DAI curves.

allocation index. The following proposition summarizes this and other properties of the dynamic allocation index relevant for the industrial targeting problem.

Proposition 1. *The dynamic allocation index $v_i(\cdot)$ for industry i increases in the fixed cost K_i , the variance of the match quality parameter \hat{p}_i , and the number of entrepreneurs willing to establish firms using technology i in country South, \bar{n}_i .*

Proof. See Appendix A. □

Fig. 1 provides the graphic illustration of the index-based industrial targeting policy. The points on the plane represent industries characterized by the fixed cost (measured on the vertical axis) and South’s expected comparative advantage (measured on the horizontal axis). A countable set of industries available for targeting is represented by area \mathcal{N} . The figure also shows iso-welfare and iso-DAI curves (i.e., the loci of industries characterized by the same levels of immediate expected welfare and the dynamic allocation index).

It is assumed that the Southern government’s initial beliefs about the quality of a match between labor and technology in the set \mathcal{N} is represented by the probability density function beta (1,2).²⁵ The number of entrepreneurs willing to establish new firms in South is sufficiently large and assumed to be the same for all industries. Thus, industries differ only in terms of the fixed cost and the expected comparative advantage coefficient.

²⁵ I assume that the variance of the comparative advantage coefficient is same in all industries. Since this assumption is ensured by setting $|a_i^{\text{high}} - a_i^{\text{low}}| = 1$ for all i and since for simplicity I also assume that the coefficient of comparative advantage is non-negative (i.e., $a_i^{\text{high}}, a_i^{\text{low}} \geq 0$), it follows that $\bar{a}_i(1/2) \geq 1/2$. An industry i with $\bar{a}_i(1/2) = 1/2$ must have $a_i^{\text{high}} = 1$. Therefore for any posterior beliefs (r,n) , South will not have comparative advantage in that industry. Hence, an industry that has $\bar{a}_i(1/2) = 1/2$ will be characterized by the dynamic allocation index equal to zero for any posterior beliefs, i.e., $v_i(r,n) = 0$.

As shown in Fig. 1, the Southern government would prefer industry Y to industry X according to the myopic criterion of maximizing immediate expected welfare reward but not according to the index policy criterion. One can see that, depending on the shape of the “cloud” of industries \mathcal{N} , there are industries which yield negative expected welfare and in which South does not have any “myopic” comparative advantage (i.e., for these industries $\bar{a}_i < 1$). However, some of these industries might get priority under the government’s optimal targeting policy over the industries, which yield positive immediate welfare and in which South has myopic comparative advantage (i.e., those industries for which $\bar{a}_i > 1$). For example, the government will prefer to target the industries in the area B over the industries in the area A even though in the short-term such policy is expected to bring welfare loss. One can also infer from Fig. 1 that unlike immediate welfare reward the dynamic allocation index always increases as we move northeast from the origin. This is because larger sunk cost K_i implies larger variance of reward, which always makes an industry more attractive for targeting according to the index criterion but not according to the myopic criterion.

As the number of potential entrants \bar{n}_i in the industry i in country South increases (for example, with the size of the industry in North), the scope for experimentation by the Southern government also increases, and this leads to a higher dynamic allocation index. As \bar{n}_i reduces to one, the index converges to the immediate reward. Therefore, we can hypothesize that if the size of the industry in country North is an indicator of the number of entrepreneurs seeking to establish firms in that industry in country South, then industries, which are large in North, are more likely to have higher priority in the industrial policy of the South’s government even when prior information indicates that South does not have a comparative advantage in that industry.²⁶

This effect can be illustrated by examples from the history of automotive industry in Latin American countries. For example, the governments of Mexico and Brazil in the 1950s and 1960s focused their efforts on promoting the car assembly by instituting various subsidies and tax preferences for this downstream segment of the automotive industry. In many cases, these governments relaxed or even waived for multinational automotive companies the existing local content rules that otherwise would have favored production of automotive components rather than cars.²⁷

²⁶ It is possible that in industries with high fixed costs, the total number of firms will be low. If the main determinant of the number of potential entrants in South, \bar{n}_i , is the total number of firms in industry i in North, then an increase in K_i might indirectly reduce the industry’s dynamic allocation index by leading to a lower \bar{n}_i . By contrast, according to Proposition 1, the direct effect of an increase in K_i is to increase the dynamic allocation index. To establish which of the two effects will be stronger it is necessary to have an explicit model connecting K_i and \bar{n}_i . Such a model is beyond the scope of this paper.

²⁷ What makes automotive industry in these countries a particularly good example for discussing industrial targeting is the fact that the governments in this region were often confronted with a very stark choice between promoting car assembly and promoting production of automotive components. The reason for this is that multinational automotive companies very rarely make greenfield investments in developing countries by building production plants on completely new sites. Instead, they prefer to make brownfield investments by redeveloping the existing production sites or not to invest at all. While allocation of greenfield sites was not difficult for the governments because they are abundantly available, the availability of brownfield sites was limited even in large developing countries. For every site that was suitable for automotive manufacturing, the governments typically had several competing offers from automotive multinationals to build new assembly or component-manufacturing facilities on this site (see Havas, 1997; Shaiken and Herzenberg, 1987). By choosing between the assembly and component-manufacturing projects these governments were making clear-cut industrial targeting decisions.

Initially, there were many foreign automakers willing to take the advantage of these policies and cheap labor in these countries. However, since most components for the assembly had to be delivered from abroad, the labor cost advantages of these countries were not sufficient to compensate for their poor infrastructure, which inhibited the just-in-time delivery of components and resulted in high overhead cost of assembly. Multinationals began to demand greater incentives in exchanges for their commitment to continue the assembly operations in these countries. Consequently, the governments shifted the emphasis in their promotion policies to manufacturing of automotive components, which was less dependent on the infrastructure but required investments into more capital- and skilled-labor-intensive technologies.²⁸ Prompted by this policy shift, some automotive multinationals ventured to build sophisticated stamping, body shop and engine production facilities, which they initially considered unsuited to these Latin American countries.²⁹ As it turned out later, upstream production of supply parts, in particular car engines and transmissions, was more efficient than downstream assembly in these countries (see *Shaiken, 1995*).

According to Proposition 1, among the reasons for the described sequencing of the automotive industry segments in the industrial targeting policies conducted by these governments could be the number and characteristics of potential projects that foreign car makers were willing to undertake in these countries. Since there were more potential entrants,³⁰ the host countries' governments had more room to experiment in the assembly segment in order to find out how well it was suited to their economies. Although automotive component manufacturing had an advantage of being less dependent on the local infrastructure, it had fewer potential entrants that could be targeted with special entry incentives in the host country. Therefore, given the uncertainty of the eventual outcome in both segments of the automotive, the learning considerations led first to the promotion of foreign investments in the local assembly even though prior beliefs were suggesting that components manufacturing could give greater immediate payoff.³¹

²⁸ For example, both the Brazilian and the Mexican governments imposed stricter local content rules on assemblers and redirected subsidies and tax incentives toward projects aimed at developing component manufacturing. A commonly cited alternative explanation for this policy shift is that large volume of component imports for the assembly plants worsened the countries' trade balances and came into conflict with the import substitution goals of these governments. However, this explanation does not answer the question why the governments did not concentrate their targeting efforts on components manufacturing from the start of their targeting policies in the automotive industry.

²⁹ The most notable examples in Mexico are Ford's automotive engine plant in Chihuahua, built in 1983, which exports 90% of its output to the U.S., and Nissan's engine and gearbox plant in Aguascalientes, serving the company's U.S. car assemblers. Mexico's engine exports soared from 320,000 units in 1982 to about 1,256,100 in 1994, and the value of its component exports rose from \$212 million to over \$2 billion (see *Shaiken, 1995*).

³⁰ Automotive multinationals were more willing to set up kit assembly plants in these countries rather than build component plants. One of the reasons for this had to do with the high degree of modularity and flexibility of the automotive assembly, which made it easily fungible. The scale of operations could be reduced or expanded without significant additional investments by changing the degree of subassembly in the knocked down car kits that were imported from abroad. This made investments into assembly plants somewhat more reversible and less risky than investments into less fungible component production (see *Lung, 2000*).

³¹ The motive to use industrial policy for generating new information on the suitability of automotive technology to local conditions was also emphasized by *Hoff (1997)* in the context of the Brazilian government's policy toward automotive industry in the 1950s and 1960s. Using a model that is different from ours, *Hoff (1997)* analyzed how the government intervention can create more information on relative efficiency of different sectors and how this information can improve resource allocation.

Some readers might object to this interpretation of events by suggesting that the governments were too myopic to take learning considerations into account and that they targeted the assembly projects simply because they underestimated the importance of good infrastructure for automobile assembly.³² That might be the case for the Latin American governments that started their policies in the 1950s, i.e., in the beginning of internationalization of automotive production. However, it is unlikely that some 20 years later there were developing countries, in which policy-makers were still unaware of the insufficiency of cheap labor alone without good infrastructure to ensure the success of local automotive assembly. Nevertheless, many developing country governments throughout the world followed the pattern of targeting the assembly segment first despite the criticism that they were committing the same mistake in their policies that their counterparts in Latin American developing nations had made in the 1950s and 1960s.³³

If considered in the context of optimal experimentation, there is nothing irrational in the behavior of developing country governments which in the 1970s and the 1980s began to pursue the automotive industry policies that had already been abandoned by other governments in the economic environments with very similar labor market and infrastructure characteristics. In fact, this historical incidence of “policy dispersion” bears much resemblance to the phenomenon of price dispersion that was first explained by Rothschild (1974) using the concept of incomplete learning. He showed that if there are several players who do not perfectly observe each other’s histories of experimentation, there is a positive probability that they will settle on different arms (or stochastic reward processes) all thinking that it is the other players who are playing the wrong arm after having had bad draws on the right arm.³⁴ Similarly, given the inherent uncertainty of the

³² An alternative interpretation of the described pattern of automotive industry development in these countries can be based on the adaptive-learning rather than the Bayesian-learning framework. This interpretation would emphasize the active role of the automotive multinationals in adapting their production networks to the local conditions of developing economies. According to this interpretation, the car makers started with building assembly plants because they needed to accumulate experience about the local economic environment using relatively simple small-scale technologies. The acquired experience would later be used in component manufacturing, which required large-scale investments. The local governments just passively reacted to the multinationals’ strategies by providing production sites for all their projects. However, there are problems with this interpretation. First, as discussed above, the governments had to make active targeting decisions every time they were selecting between assembly and component-manufacturing proposals for a given brownfield site. Second, because of the differences in technologies, very little uncertainty regarding the match between local labor and component manufacturing technology can be resolved by operating assembly plants (see Shaiken and Herzenberg, 1987). A more nuanced and complex model would interpret the process of international fragmentation of the automotive industry as bilateral search process that combines elements of Bayesian learning on the sides of both host country governments and multinationals. In such a model, both companies and governments would be testing the match between the host countries’ economic environment and the foreign companies’ technologies. However, the focus of this paper is on the effect of the government’s experimental policy activism on the country’s ability to specialize according to comparative advantage. This effect can be best analyzed using a simpler model of one-sided Bayesian learning considered here.

³³ See Meth-Cohn (1995), Calbreath (1995), and Havas (1997).

³⁴ It is worth quoting Rothschild (1974) in this regard: “One could well ask whether they (stores) would be content charging the prices that they think are best while observing that other stores—presumably rational—are charging different prices. I do not think this is a particularly compelling point. Unless store A has access to store B’s books, the mere fact that store B is charging a price different from A’s and not going bankrupt is not conclusive evidence that A is doing the wrong thing. Who is to say A’s experience is not a better guide to the true state of affairs than B’s?”

match between the foreign automotive technology and the local economic environment, the governments that targeted the automotive assembly segment in the 1970s and 1980s might have interpreted the earlier experience of the other governments with the assembly plants as bad luck and decided to promote this segment of the industry because of its high experimentation value.

5. The long-run pattern of specialization

I now turn to the examination of the set of industries in which the country will specialize in the long-run if its industrial targeting policy is governed by the optimal (or almost optimal) strategy as defined in Section 3. To this end, I analyze the stochastic convergence properties of the index-based industrial targeting process.

In this section, I assume that for each Northern technology $i \in \mathcal{N}$ there is unlimited number of entrepreneurs wishing to establish firms in country South. As in the previous sections, all new firms in South produce to sell in the Northern market. Moreover, since I assume that, despite entry of new firms, South remains much smaller than North, the terms of trade between the two countries continue to be determined by the supply and demand conditions in North. Therefore, the prices, the expected profits or losses of new firms, and the expected welfare of the Southern government are the same as in Section 2.³⁵

Suppose at each time t there is a priority ordering of the available industries $J(t)=(1, 2, \dots, N_t)$ which ranks them by the size of their indices

$$v_{1_t}(x_{1_t}(t)) \geq v_{2_t}(x_{2_t}(t)) \geq \dots \geq v_{N_t}(x_{N_t}(t))$$

Then the *set of targeted sectors* (i.e., the set of industries targeted at time t by licensing new firms in these industries) is formally defined by³⁶

$$T_t = \{j_t : j_t \leq M\}$$

The initial set of targeted sectors T_0 consists of the industries with the M highest indices given the prior beliefs $x_i(0)=(r_i, n_i)$, $i = 1, 2, \dots, N$. Let

$$m^* = \sup_{j \in \mathcal{N} \setminus T_0} v_j(x_j(0)) \tag{7}$$

Define $\mathbf{x}(\tau)=(x_1(\tau), \dots, x_N(\tau))$ to be the vector of posterior beliefs about the quality of the match between the available technologies and Southern labor at time τ . The set of the

³⁵ It is also assumed that the set \mathcal{N} of available Northern technologies is such that it makes sense for the Southern government to experiment with any of them; that is, I rule out technologies characterized by such a random productivity parameter that licensing a firm with such a technology can yield negative (positive) expected welfare even under the most optimistic (pessimistic) beliefs about the quality of match with local labor. More formally, I rule out the technologies for which $E(\tilde{W}_i | \bar{p}_i = 1) < 0$ ($E(\tilde{W}_i | \bar{p}_i = 0) > 0$), where \bar{p}_i is the expected probability of a good match.

³⁶ The results of this section can be easily restated for the special case when $M=1$ and the set T_t is a singleton. Therefore, our characterization of the long-run stochastic properties of the almost optimal policy is also applicable for the optimal targeting policy under the constraint $M=1$.

targeted sectors T_0 will remain unchanged by time t if the vector of beliefs updated after every new entry satisfies the inequalities:

$$v_i(x_i(\tau)) \geq m^* \text{ for } i \in T_0 \text{ and } \tau = 1, 2, \dots, t$$

To analyze the set of the targeted sectors in the long-run, or as time t goes to infinity, I will characterize the probability distribution of the number of periods during which the above inequalities remain valid.³⁷

The support of the random variable $\tilde{\gamma}_i$ is denoted by $\Gamma_i = \{\gamma_i^{\text{low}}, \gamma_i^{\text{high}}\}$. The two outcomes in the support correspond to the realizations of the productivity parameter γ_i under the good and bad matches between local labor and technology i . The vectors of productivity realizations in the targeted industries are the elements of the M -fold Cartesian product of Γ_i denoted by $\Gamma = \times_{i=1}^M \Gamma_i$.

The sequences of vectors of productivity realizations after t consecutive entries of new firms in each of the industries in the set T_0 are the elements of the t -fold Cartesian product of Γ denoted by $\Gamma^t = \times_{\tau=1}^t \Gamma$. Given the vector of the true probabilities of high productivity realizations $\mathbf{p} = (p_1, \dots, p_M)$, the probability of observing a specific sequence of vectors $\gamma^t \in \Gamma^t$ is defined by

$$P(\gamma^t | \mathbf{p}) = \prod_{i \in T_0} \text{Prob}(\tilde{\gamma}_{ik} = \gamma_{ik}, 1 \leq k \leq t | p_i) = \prod_{i \in T_0} p_i^{s_i} (1 - p_i)^{t - s_i} \tag{8}$$

where γ_{ik} is the realization of the productivity parameter of the k th entrant in industry i and $s_i = \sum_{k=1}^t \mathbf{1}_{\{\tilde{\gamma}_{ik} = \gamma_{ik}^{\text{low}}\}}$ is the number of high productivity (low cost) realizations after t entries of firms in industry i .

Denote by G^t the set of all t -element sequences of realizations of productivity vectors after licensing t firms in each of the industries in the targeted set T_0 such that after these realizations the elements of the targeted set are the same as in the beginning of the policy implementation. Clearly, $G^t \subset \Gamma^t$. Then the probability that the elements of the targeted set are unchanged by time t given that the vector of the true probabilities of high productivity (good match) is \mathbf{p} is defined by

$$\Psi_{\mathbf{p}}^{T_0}(t) = \int_{G^t} P(\gamma^t | \mathbf{p}) d\gamma^t \tag{9}$$

Define $A_{\mathbf{p}}^{T_0} = \lim_{t \rightarrow \infty} \Psi_{\mathbf{p}}^{T_0}(t)$ to be the probability that the country will specialize in the long-run in the same set of industries in which it specializes at time $t=0$ given that the vector of the true probabilities of the good match between the technologies and local labor is \mathbf{p} . The limit exists because $\Psi_{\mathbf{p}}^{T_0}(t)$ is non-increasing in t and bounded.

³⁷ To eliminate the trivial cases, I assume that the set \mathcal{N} contains only “competitive” technologies. This means that for each technology i in \mathcal{N} , the random match-quality parameter includes in its support an outcome which allows i to “overtake” by some time τ other technologies and be included in the set of the targeted sectors T_τ under some vector of posterior probabilities of the good match $\mathbf{x}(\tau) = (x_1(\tau), \dots, x_N(\tau))$. Under this condition those technologies that were not included in the set of the targeted sectors under one vector of beliefs have a chance of eventually replacing the technologies that were initially there if the latter perform in an unsatisfactory way and the beliefs about their match quality deteriorate. Formally, this condition is determined by: $\cap_{i \in \mathcal{N}} [E(\tilde{W}_i | \bar{p}_i = 0), E(\tilde{W}_i | \bar{p}_i = 1)] \neq \emptyset$.

Now I can characterize the conditions under which the country will eventually specialize in the set of industries, which its government targeted at the start of its learning process based on the available information and the index-based strategy. The main result of this section is formulated in the following proposition.

Proposition 2. *For each industry $i \in T_0$ there exists a probability p_i^* such that (1) if $p_i < p_i^*$ for any $i \in T_0$, then $\Lambda_p^{T_0} = 0$, and (2) if $p_i \geq p_i^*$ for all $i \in T_0$, then $\Lambda_p^{T_0} > 0$.*

Proof. See Appendix A. □

Proposition 2 suggests that for each industry i in the initial set T_0 we can find a certain probability p_i^* such that if for at least some $i \in T_0$ the true probability of the good match p_i is less than p_i^* , then the country will eventually abandon some or all of the technologies in the set T_0 in favor of some other technologies, which are not in the set T_0 .³⁸ The converse is correct if the true probability of the good match between local labor and the technology exceeds a certain level p_i^* for all $i \in T_0$. Therefore, Proposition 2 makes it possible to determine how good the technologies have to be in reality so that once they are selected for promotion, they have a chance of continuing to be promoted in the long-run based on the performance of new firms using these technologies.

Since $\bar{a}_i(p_i)$ is an increasing function of p_i , the results of Proposition 2 can also be formulated in terms of the true coefficient of comparative advantage $\bar{a}_i(p_i)$. In these terms, the main implication of Proposition 2 is the possibility that a country will not specialize in the long-run in an industry in which it has the greatest true comparative advantage $\bar{a}_i(p_i)$. While some technologies may be better than the others in terms of the country's true comparative advantage, they may be not good enough (i.e., $\bar{a}_i(p_i) < \bar{a}_i(p_i^*)$) to remain in the targeted set forever even if they were targeted at some time. With probability one the government will target these technologies only in a finite number of periods and will never learn the true competitiveness of its country with these technologies, because its posterior beliefs will be updated only a finite number of times. When this happens, the long-run specialization in these truly best technologies occurs with probability zero, and the country eventually specializes in the “wrong” set of technologies. These are not the best among available technologies but their comparative advantage coefficients exceed the critical bounds $\bar{a}_i(p_i^*)$ and, therefore, they can perform well enough not to induce the government to look for better technologies.

The result presented in this section offers a new angle in the debate about the benefits of government intervention in the market mechanism.³⁹ Specifically, this result indicates that

³⁸ Note that p_i^* will depend on the prior distribution of all other alternatives in the targeted set T_0 . Moreover, if all industries have identical fixed cost and productivity parameters (i.e., $K_i = K$ and $\gamma_i = \gamma$, $\forall i \in \mathcal{N}$) and differ only in terms of the parameters of the prior p.d.f. beta (r_i, n_i), then it is easy to see that all targeted industries have a common cut-off probability $p^* = r^*/n^*$ (i.e., $p_i^* = p^*$) where r^* and n^* can be obtained using an inverse of the DAI function of the industry with the $M+1$ highest index defined by the equation (G1) in Appendix A: $(r^*, n^*) = v_{(M+1)_0}^{-1}(m^*)$. I thank the referee for pointing out this fact to me.

³⁹ One of the arguments against such intervention emphasizes the conflict between the efficiency conditions for the introduction of new goods and the conditions for the efficient allocation of existing goods. In this regard, Romer (1994) and Matsuyama (1995a, 1997) argued that although government policies aimed at overcoming monopolistic rents eliminate distortions in the markets for existing goods, these policies also reduce firms' incentives for innovation and entry into new markets. They showed that losses from non-entry to new industries can, under many circumstances, be even larger than the loss from allowing the distortions in the existing markets to persist.

because of the government's failure to learn the country's true comparative advantage, the targeted promotion of firms' entry may eventually steer the country away from industries in which it can be truly competitive. When contrasted with the known results on multi-agent market learning, the result of this section gives us reasons to believe that because of the government's learning failure, the small open country's long-run pattern of specialization under incomplete information may be closer to the full information optimum pattern if it is achieved through decentralized "unassisted" entry and exit decisions of firms rather than through the industrial targeting policy.

Incomplete learning implies that realization paths of the optimal policy for the underlying MAB problem are distributed "near" the full information optimal policy (i.e., the optimal policy chosen by the policy maker who has prior knowledge of the true distribution parameters of rewards in the MAB problem). Therefore, in a setting with many independently experimenting agents, there is a positive probability that at any period of time they will hold differing posterior beliefs even if all of them began with the same prior beliefs and acted optimally at all times. This dispersion of beliefs improves the chances that mistaken views will not dominate in the long-run. Using this argument and the results about the stochastic survival of a set of targeted industries formulated in Proposition 2, it can be conjectured that if specialization according to full information comparative advantage is taken as a criterion, then parallel experimentation with alternative technologies pursued independently by a large number of entrepreneurs will yield a better outcome than promotional experimentation by a single policy maker.

A more detailed examination of this conjecture requires explicit modeling of the market failures that characterize unassisted experimentation by firms, which given the difficulty of this task is better left for a separate paper. Therefore, here we deliberately discuss the conjectured effect and its supporting heuristics in an informal way. However, the arguments in support of the conjecture can be made fully rigorous by building on the material contained in the papers by Anantharam et al. (1987), Ishikida and Varaiya (1994), and Smith (1991).⁴⁰

One difficulty in modeling experimentation in the multi-agent setting is that the agents can learn by observing each others' experiments (observational learning) as well as from their own experience (experiential learning).⁴¹ The informational spillovers from private experimentation create an externality that transforms the bandit problem into a game of strategic experimentation. In this game, agents might underinvest in gaining information through their own experience and rely too heavily on observational learning from the experimentation of others. The latter effect creates a possibility of informational herding that can arise in the environment in which agents must choose actions among a number of alternatives, having observed their own private signals about the relative value of actions

⁴⁰ Formal proof of the conjecture would involve the examination of conditions under which a collection of optimal policies for M independent agents confronting the common-value MAB problem and able to target a single reward process at a time yields a higher expected welfare than the optimal policy for the same MAB problem with a single agent who has the capacity to target M reward processes at a time.

⁴¹ Unlike the model analyzed in this paper, a model of decentralized learning would have to consider the firms that have the ability to experiment by changing their adopted technologies at least once during their lifetime.

and the actions chosen by other agents (see Banerjee, 1992; Bikhchandani et al., 1992). Therefore, similar to incomplete learning effect of optimal experimentation in a single-agent model, misguided herds may arise from optimal actions of individual agents in the multi-agent learning context.

However, the herding externality does not arise under a set of conditions that prevent agents from relying too much on learning from each other (see Banerjee, 1992). Appropriate “compartmentalization” of the learning environment balances agents’ experiential and observational learning incentives and ensures that each agent’s individual risk of the learning failure is minimized without leading to misguided herding. For example, Smith (1991) considers agents that enter sequentially and learn by observing a “snapshot” of the actions taken by a fraction of agents from previous generations before solving their own bandit problem. In Smith’s framework, strategic behavior is limited because agents are not allowed to observe each other once they have entered. Smith shows that the incomplete learning result described by Rothschild (1974) is not robust to this form of market learning. The most profitable action is eventually chosen by the market with probability one, as opposed to the case of an isolated agent, who might take less than the most profitable actions forever even under the optimal learning policy.⁴² The condition that the agent can learn from fragments of the past experience of a subset of other agents rather than from all accumulated experience of all agents is not that restrictive and is very likely to obtain in the context of learning by entrepreneurs investing in the foreign country.⁴³ Therefore, learning in which industries the country has true comparative advantage and eventual specialization according to full information are more likely to occur under decentralized experimentation by many firms, than under the targeting policy experimentation by the single policy-maker.

6. Conclusion

A standard justification for industrial targeting policy involves either externalities or coordination failures that inhibit the introduction of new goods, or lead to a less than socially optimal rate of entry by new firms. However, the sheer fact that the industries in which the government intervenes are new for the country (or even nonexistent before the intervention) implies that there is always a great deal of uncertainty about the outcomes of such intervention. The government’s inability to intervene in all the industries, in which its

⁴² The definition of complete learning in the market will be different from the definition of single-agent complete learning. While defining complete market learning formally is outside the scope of this paper, intuitively such a definition will have to incorporate the almost sure convergence to the full-information optimal actions by individual agents and the convergence of the proportion of such agents in the entire agent population to one.

⁴³ Banerjee and Fudenberg (1995) and Smith and Sørensen (1996) considered models that correspond to word-of-mouth type of social learning. In their models individuals observe not the entire action history of all agents, but rather unordered action samples from the pool of history. They showed that when the sampling mechanism favors the more recent past full learning can usually be achieved. However, learning is incomplete if the sampling mechanism over-samples the more remote past.

intervention may be desirable, implies that under uncertainty it will necessarily experiment in allocating its scarce targeting capacity. The set of targeted industries will change in the course of such policy experimentation as the uncertainties about the industries are resolved.

In this paper, the optimal industrial targeting policy of the government, which wants to discover in which industries its country has the greatest comparative advantage and, at the same time, maximize the flow of welfare during the process of search for such industries, is represented as the Gittins index-based solution to the multi-armed bandit problem. Using this solution, the paper analyzes how the government's choice of industries under the optimal targeting policy is affected by such parameters as the set-up cost, the variance of the unknown productivity parameter, and the size of the 'experimentation field' determined by the number of entrepreneurs willing to test the new technology in the economic environment of the country.

Having clarified the main features of the industrial targeting policy pursued by a welfare maximizing government under uncertainty, this paper analyzes the long-term outcome of such a policy. It is demonstrated that even under the optimal targeting policy the country might eventually specialize in the industries in which it does not have comparative advantage. The possibility of failure to specialize according to comparative advantage suggests that industrial targeting policy as a form of government intervention in economic development does not guarantee improved efficiency in the long-run, even if it is justified in the short run by the need to overcome market failures and facilitate entry of new firms. The paper advances arguments suggesting that although decentralized, unassisted entry and exit decisions by firms may be costlier in the short run in terms of possible negative externalities or coordination failures, in the long-run such decentralized experimentation has a better chance of leading to a more efficient specialization of the country as a whole.

One issue that the paper leaves open is characterization of targeting policy in the environment in which the set of industries that can be targeted changes as a result of innovation and the advent of new technologies. For example, assuming a stochastic arrival of new technologies, the government's targeting problem can be formulated in the context of the so-called "arms-acquiring bandit", which also has an index-based optimal solution. Further, to make the multi-armed bandit model of industrial targeting policy more realistic, one could introduce a cost of adjusting the set of targeted industries in the course of policy implementation. The main difficulty in analyzing such an extension is in presenting the optimal policy in an analytically tractable way because, under the condition of costly experimentation, the index-based formulation of the optimal solution for the multi-armed bandit problem does not exist.

Acknowledgements

I would like to thank Paul Krugman, John Roberts, Philipp Afèche, Michael Herron, T. L. Lai, three anonymous referees, and participants at seminars at the Stanford University Graduate School of Business, and at the Graduate School of International Relations and Pacific Studies at the University of California at San Diego for helpful discussions.

Appendix A

Proof of Proposition 1. (I) The Gittins index for the industrial targeting problem is defined by:⁴⁴

$$v_i(x_i) = \max_{\tau > 0} \frac{E \left\{ \sum_{t=0}^{\tau-1} \beta^t \tilde{W}_i(x_i(t)) \mid x_i(0) = x_i \right\}}{E \left\{ \sum_{t=0}^{\tau-1} \beta^t \mid x_i(0) = x_i \right\}} \tag{GI}$$

where τ is a stopping time with respect to the filtration generated by the past history of targeting the industry i . Assuming that at each decision time the government has the option to suspend targeting any industry with guaranteed zero reward, the index is nonnegative. Therefore, rewriting the index in the following form immediately shows that it is increasing in K_i :

$$v_i(x_i) = \max_{\tau > 0} \frac{K_i E \left\{ \sum_{t=0}^{\tau-1} \beta^t [\tilde{a}_i(x_i(t)) - 1] \mid x_i(0) = x_i \right\}}{E \left\{ \sum_{t=0}^{\tau-1} \beta^t \mid x_i(0) = x_i \right\}} \tag{Q.E.D.}$$

(II) Omitting the industry subscript, the variance of an industry’s welfare reward in the state beta (r,n) is given by

$$\text{Var}(\tilde{W}(r, n)) = K^2 \text{Var}(\tilde{p}) = K^2 \frac{(1 - r/n)r/n}{n + 1}$$

Since the variance of a random variable \tilde{p} with p.d.f. beta (r,n) is defined by $\text{Var}(\tilde{p}) = [(1 - r/n)r/n]/(n + 1)$, a mean-preserving increase in the variance is equivalent to a decrease in n for a fixed ratio r/n .

Now, let us consider a problem in which the Southern government can promote only a single foreign technology. Let the prior beliefs about the quality of the match between the foreign technology and labor in country South be determined either by the p.d.f. beta (r,n) or by the p.d.f. beta (r',n') , such that $r/n = r'/n'$, and $n > n'$. Hence, the beliefs about the quality of the match in the industry characterized by the imported foreign technology can evolve according to one of the two Markov chains $x(t)$ and $x'(t)$, $t = 0, 1, 2, \dots$, defined respectively on the state spaces $\{(a,b), a = r, r + 1, r + 2, \dots; b = n, n + 1, n + 2\}$ and $\{(a',b'), a' = r', r' + 1, r' + 2, \dots; b' = n', n' + 1, n' + 2, \dots\}$.

Define for $t = 1, 2, \dots$, the rewards

$$\tilde{W}_t := \tilde{W}(x(t - 1)) \text{ and } \tilde{W}'_t := \tilde{W}(x'(t - 1))$$

Denote by $\mathcal{F}(t)$ and $\mathcal{F}'(t)$ the sigma algebras representing information accumulated by observing the productivity of new firms in the industry by time t . Let $\tilde{W} := \{\tilde{W}_t, \mathcal{F}(t)\}_{t=1}^\infty$

⁴⁴ See Mandelbaum (1986).

and $\tilde{W}' := \{\tilde{W}'_t, \mathcal{F}'(t)\}_{t=1}^\infty$ be the stochastic reward processes associated with targeting the industry based on the foreign technology when the beliefs evolve respectively from the initial states $x(0)=(r,n)$ and $x'(0)=(r',n')$.

Assume that the number of new firms, which can enter the industry, is bounded by n^* . Define

$$v^{n^*}(x) = \max_{n^* \geq \tau > 0} \frac{E \left\{ \sum_{t=0}^{\tau-1} \beta^t \tilde{W}(x(t)) \mid x(0) = x \right\}}{E \left\{ \sum_{t=0}^{\tau-1} \beta^t \mid x(0) = x \right\}}$$

and let

$$\sigma(n^*) = \min[n^*, \inf\{t : v^{n^*-t}(r + s_t, n + t) < v^{n^*}(r, n)\}]$$

where $s_t = \sum_{j=1}^t \mathbf{1}_{\{\tilde{\gamma}_j = \gamma^{low}\}}$. The following result is due to Gittins (1989, p. 47):

$$v^{n^*}(x) = \sup_{n^* \geq \tau > 0} \frac{E \left\{ \sum_{t=0}^{\tau-1} \beta^t \tilde{W}(x(t)) \mid x(0) = x \right\}}{E \left\{ \sum_{t=0}^{\tau-1} \beta^t \mid x(0) = x \right\}} \tag{A.1}$$

For $n^*=1$ we have

$$v^1(r, n) = \bar{W}(r, n) = \bar{W}(r', n') = v^1(r', n')$$

If $n^*=2$, one obtains

$$\sigma(n^*) = \begin{cases} 1, & \text{if } \tilde{\gamma}_1 = \gamma^{high} \\ 2, & \text{if } \tilde{\gamma}_1 = \gamma^{low} \end{cases}$$

Hence,

$$\begin{aligned} v^2(r, n) &= \frac{(1 - r/n)\bar{W}(r, n) + [\bar{W}(r, n) + \beta\bar{W}(r + 1, n + 1)]r/n}{(1 - r/n) + [1 + \beta]r/n} \\ &= \frac{1 + \frac{r}{n}\beta \frac{\bar{W}(r + 1, n + 1)}{\bar{W}(r, n)}}{1 + \frac{r}{n}\beta} \bar{W}(r, n) \end{aligned} \tag{A.2}$$

Since $[(r' + 1)/(n' + 1)] > [(r + 1)/(n + 1)]$, we have $\bar{W}(r' + 1, n' + 1) > \bar{W}(r + 1, n + 1)$ and $v^2(r', n') > v^2(r, n)$.

Assume now that for $t=2, \dots, n^*-1$, if $r/n = r'/n'$, and $n > n'$, then $v^t(r', n') > v^t(r, n)$. We need to show that the following is also true:

$$v^{n^*}(r', n') > v^{n^*}(r, n).$$

Given the assumption of the inductive hypothesis, for $t=2, \dots, n^*-1$, from $[(r+s_t)/(n+t)] \leq (r/n)$ it follows that $v^{n^*-t}(r+s_t, n+t) \leq v^{n^*-t}(r, n)$. Also, it follows from the definition of $v^{n^*}(r, n)$ that $v^{n^*-t}(r, n) \leq v^{n^*}(r, n)$, $t=2, \dots, n^*-1$. Therefore, when $[(r+s_t)/(n+t)] \leq (r/n)$ we have $v^{n^*-t}(r+s_t, n+t) \leq v^{n^*}(r, n)$.

Now notice that from $t < \sigma(n^*)$ it follows that $[(r+s_t)/(n+t)] > (r/n)$, which implies that $(s_t/t) > (r/n)$.

Define for the Markov chains $\{x(t)\}$ and $\{x'(t)\}$, $t=0, 1, 2, \dots$, the random productivity parameters $\tilde{\gamma}_t := \tilde{\gamma}(x(t-1))$ and $\tilde{\gamma}'_t := \tilde{\gamma}(x'(t-1))$.

Let $\tilde{F} := \{\tilde{\gamma}_t, \mathcal{F}\}_{t=1}^\infty$ and $\tilde{F}' := \{\tilde{\gamma}'_t, \mathcal{F}'\}_{t=1}^\infty$ be the stochastic sequences of productivity parameters observed after each new entry in the industry given that the beliefs about these parameters evolve from the initial states $x(0)=(r, n)$ and $x'(0)=(r', n')$.

Recall that $s_t = \sum_{j=1}^t \mathbf{1}_{\{\tilde{\gamma}_j = \gamma^{\text{low}}\}}$ and define $s'_t = \sum_{j=1}^t \mathbf{1}_{\{\tilde{\gamma}'_j = \gamma^{\text{low}}\}}$. Also recall that

$$\bar{\gamma}_{t+1} = E(\tilde{\gamma}_{t+1}) = \gamma^{\text{low}} P(\tilde{\gamma}_{t+1} = \gamma^{\text{low}} \mid s_t = s) + \gamma^{\text{high}} (1 - P(\tilde{\gamma}_{t+1} = \gamma^{\text{low}} \mid s_t = s)) \tag{A.3}$$

and

$$\bar{\gamma}'_{t+1} = E(\tilde{\gamma}'_{t+1}) = \gamma^{\text{low}} P(\tilde{\gamma}'_{t+1} = \gamma^{\text{low}} \mid s'_t = s) + \gamma^{\text{high}} (1 - P(\tilde{\gamma}'_{t+1} = \gamma^{\text{low}} \mid s'_t = s)). \tag{A.3'}$$

Notice that it follows from the inequality $(s_t/t) > (r/n)$ that

$$P(\tilde{\gamma}_{t+1} = \gamma^{\text{low}} \mid s_t = s) = \frac{r+s}{n+t} < \frac{r'+s}{n'+t} = P(\tilde{\gamma}'_{t+1} = \gamma^{\text{low}} \mid s'_t = s)$$

and therefore $\bar{\gamma}_{t+1} > \bar{\gamma}'_{t+1}$.

Hence, for $s_t = s'_t = s$,

$$\begin{aligned} E(\tilde{W}_{t+1} \mid x(t) = (r+s, n+t)) &= K \left(2 \frac{w^N \gamma^N}{w^S \bar{\gamma}_t} - 1 \right) < K \left(2 \frac{w^N \gamma^N}{w^S \bar{\gamma}'_t} - 1 \right) \\ &= E(\tilde{W}'_{t+1} \mid x'(t) = (r'+s, n'+t)) \end{aligned} \tag{A.4}$$

From Eq. (A.4) and using the fact that the product of random variables $\mathbf{1}_{\{\sigma(n^*) \geq t\}}$ \tilde{W}'_t is a $\sigma\{\mathcal{F}(t) \cup \mathcal{F}'(t)\}$ -measurable random variable we can find:

$$E \sum_{t=1}^{\sigma(n^*)} \beta^{t-1} \tilde{W}_t = \sum_{t=1}^{n^*} \beta^{t-1} E \left[\tilde{W}_t \mathbf{1}_{\{\sigma(n^*) \geq t\}} \right] < \sum_{t=1}^{n^*} \beta^{t-1} E \left[\tilde{W}'_t \mathbf{1}_{\{\sigma(n^*) \geq t\}} \right] = E \sum_{t=1}^{\sigma(n^*)} \beta^{t-1} \tilde{W}'_t \tag{A.5}$$

Then using Eq. (A.5) together with Eq. (A.1) we obtain:

$$v^{n^*}(r, n) = \frac{E \sum_{t=1}^{\sigma(n^*)} \beta^{t-1} \tilde{W}_t}{E \sum_{t=1}^{\sigma(n^*)} \beta^{t-1}} < \frac{E \sum_{t=1}^{\sigma(n^*)} \beta^{t-1} \tilde{W}'_t}{E \sum_{t=1}^{\sigma(n^*)} \beta^{t-1}} \leq \sup_{0 < \tau \leq n^*} \frac{E \sum_{t=1}^{\tau} \beta^{t-1} \tilde{W}'_t}{E \sum_{t=1}^{\tau} \beta^{t-1}} = v^{n^*}(r', n'). \tag{A.6}$$

Q.E.D.

(III) Assume that there are \bar{n}_i entrepreneurs with technology i who would like to establish firms in country South. Let the prior beliefs about the quality of the match between technology i and Southern labor be determined by the p.d.f. beta (r, n) . Note that the beliefs about the quality of the match evolve according to a $(2^{\bar{n}_i} + 2)$ -state Markov chain $x_i(t)$, $t = 0, 1, 2, \dots$, with the states: $\{(a, b), a = r, r + 1, r + 2, \dots, r + \bar{n}_i; b = n, n + 1, n + 2, \dots, n + \bar{n}_i; \emptyset\}$ where \emptyset denotes the absorbing state and r/n is the initial state.

Let $P = \{P_{ij}\}$ denote the $(2^{\bar{n}_i} + 2) \times (2^{\bar{n}_i} + 2)$ transition matrix for the chain $\{x_i(t)\}$. Then the \bar{n}_i th power of the transition matrix, $P^{\bar{n}_i}$, has the following elements:

$$P_{kj}^{\bar{n}_i} = 0, j = 1, \dots, 2^{\bar{n}_i} + 1 \text{ and } P_{k, 2^{\bar{n}_i} + 2}^{\bar{n}_i} = 1, k = 1, \dots, 2^{\bar{n}_i} + 2.$$

For any Markov chain with such transition matrix, it takes no more than \bar{n}_i transitions to evolve from any initial state into the absorbing state. Therefore after the establishment of \bar{n}_i new firms in South, beliefs evolve into the absorbing state \emptyset , in which the government receives the deterministic zero reward: $W_i(\emptyset) = 0$.

For the Markov chain $\{x_i(t)\}$ we define

$$v_{\bar{n}_i}^L(x_i) = \max_{L \geq \tau > 0} \frac{E \left\{ \sum_{t=0}^{\tau-1} \beta^t \tilde{W}_i(x_i(t)) \mid x_i(0) = x_i \right\}}{E \left\{ \sum_{t=0}^{\tau-1} \beta^t \mid x_i(0) = x_i \right\}}$$

where $L \geq 1$ is a nonrandom integer and the maximum is taken over stopping times τ for which $P(\tau \leq L) = 1$. Then $v_{\bar{n}_i}^{L+1}(x_i) \geq v_{\bar{n}_i}^L(x_i)$. However, since the reward is zero after \bar{n}_i transitions, we have $v_{\bar{n}_i}^{\bar{n}_i+k}(x_i) = v_{\bar{n}_i}^{\bar{n}_i}(x_i)$, $k = 1, 2, \dots$

Suppose σ is such that the Gittins index for the Markov chain $\{x_i(t)\}$ is equivalent to

$$v_{\bar{n}_i}(x_i) = \frac{E \left\{ \sum_{t=0}^{\sigma-1} \beta^t \tilde{W}_i(x_i(t)) \mid x_i(0) = x_i \right\}}{E \left\{ \sum_{t=0}^{\sigma-1} \beta^t \mid x_i(0) = x_i \right\}}$$

where the subscript \bar{n}_i indicates that we are considering the Gittins index for a $(2^{\bar{n}_i} + 2)$ -state Markov chain.

Then we have

$$v_{\bar{n}_i}(x_i) = \frac{\sum_{t=0}^{\infty} \beta^t E[\tilde{W}_i(x_i(t)) \mathbf{1}_{\{\sigma > t\}} \mid x_i(0) = x_i]}{\sum_{t=0}^{\infty} \beta^t P\{\sigma > t \mid x_i(0) = x_i\}} \tag{A.7}$$

and

$$v_{\bar{n}_i}^L(x_i) \geq \frac{\sum_{t=0}^L \beta^t E[\tilde{W}_i(x_i(t)) \mathbf{1}_{\{\sigma > t\}} \mid x_i(0) = x_i]}{\sum_{t=0}^L \beta^t P\{\sigma > t \mid x_i(0) = x_i\}}. \tag{A.8}$$

Therefore

$$v_{\bar{n}_i}(x_i) - v_{\bar{n}_i}^L(x_i) \leq \frac{\sum_{t=0}^{\infty} \beta^t E[\tilde{W}_i(x_i(t)) \mathbf{1}_{\{\sigma > t\}} \mid x_i(0) = x_i]}{\sum_{t=0}^{\infty} \beta^t P\{\sigma > t \mid x_i(0) = x_i\}} - \frac{\sum_{t=0}^T \beta^t E[\tilde{W}_i(x_i(t)) \mathbf{1}_{\{\sigma > t\}} \mid x_i(0) = x_i]}{\sum_{t=0}^T \beta^t P\{\sigma > t \mid x_i(0) = x_i\}} \leq \frac{\beta^L E(\tilde{W} \mid x_i = (r + \bar{n}_i, n + \bar{n}_i))}{1 - \beta} \tag{A.9}$$

Hence, $\lim_{L \rightarrow \infty} v_{\bar{n}_i}^L(x_i) = v_{\bar{n}_i}(x_i)$ and since $v_{\bar{n}_i+k}^{\bar{n}_i}(x_i) = v_{\bar{n}_i}(x_i)$, $k = 1, 2, \dots$, it follows that $v_{\bar{n}_i}(x_i) = v_{\bar{n}_i}^{\bar{n}_i}(x_i)$.

If we now assume that there are $\bar{n}_i - 1$ entrepreneurs with technology i wishing to establish firms in country South, the evolution of beliefs as these firms are established can be described by a Markov chain $\{x_i'(t)\}$, with $2^{\bar{n}_i-1} + 2$ states and a $(2^{\bar{n}_i-1} + 2) \times (2^{\bar{n}_i-1} + 2)$ transition matrix P' which is a truncation of the matrix P . When the Markov chains $\{x_i(t)\}$ and $\{x_i'(t)\}$ have the same initial state, we have

$$v_{\bar{n}_i-1}^{\bar{n}_i-k}(x_i(0)) = v_{\bar{n}_i}^{\bar{n}_i-k}(x_i(0)), \quad k = 1, 2, \dots, \bar{n}_i - 2$$

Therefore

$$v_{\bar{n}_i-1}(x_i(0)) = v_{\bar{n}_i-1}^{\bar{n}_i-k}(x_i(0)) = v_{\bar{n}_i}^{\bar{n}_i-k}(x_i(0)) \leq v_{\bar{n}_i}^{\bar{n}_i}(x_i(0)) = v_{\bar{n}_i}(x_i(0)),$$

$$k = 1, 2, \dots, \bar{n}_i - 2$$

Hence, $v_{\bar{n}_i-1}(x_i(0)) \leq v_{\bar{n}_i}(x_i(0))$, which establishes the last claim of Proposition 1. □

Proof of Proposition 2. (I) Consider an industry $i \in T_0$, where T_0 is a set of industries targeted at the beginning of the industrial policy. Recall that \bar{p}_i denotes the probability of a high realization of productivity, γ_i^{high} . Define a sequence of random variables $X_{i,n} = E(\bar{p}_i \mid \mathcal{F}_{i,n})$, $n = 1, 2, \dots$, where $\mathcal{F}_{i,n} = \sigma(\tilde{\gamma}_{i,1}, \dots, \tilde{\gamma}_{i,n})$ is a sigma algebra generated by a sequence of random variables $\tilde{\gamma}_{i,1}, \dots, \tilde{\gamma}_{i,n}$. Then $\{\mathcal{F}_{i,n}\}$ is a natural filtration and $\mathcal{F}_{i,\infty} = \sigma(\bigcup_{n=1}^{\infty} \mathcal{F}_{i,n})$. Set $X_{i,0}$ equal to the expected probability of high productivity conditional on the prior beliefs, i.e., $X_{i,0} = \bar{p}_i^0 = (r_i^0/n_i^0)$.

Since the probability parameter \bar{p}_i is bounded by definition, we can directly use Lévy’s “Upward” theorem to claim that $X_{i,n}$, $n = 1, 2, \dots$ is a uniformly integrable (UI)

martingale relative to $\{\mathcal{F}_{i,n}\}$ and $X_{i,n} \rightarrow X_{i,\infty} := E(\tilde{p}_i | \mathcal{F}_{i,\infty})$ almost surely and in \mathcal{L}^1 (see Williams, 1991, p. 134; Billingsley, 1986, Theorem 35.5, p. 492).

Recall that $m^* = \sup_{j \in \mathcal{N} \setminus T_0} v_j(x_j(0))$. Let $X_i^* = E(\tilde{p}_i | x_i = (r_i, n_i))$ where the parameters r_i and n_i are such that $v_i(X_i^*) = m^*$. Since v_i is increasing in \tilde{p}_i^0 (see Proposition 1), for $i \in T_0$ we must have $\tilde{p}_i^0 > X_i^*$.

Now define the random variables τ_i and X_{τ_i} :

$$\tau_i = \min\{n: X_{i,n} < \tilde{p}_i^0\} \tag{B.1}$$

$$X_{\tau_i} = \begin{cases} X_{i,\tau_i(\omega)}(\omega) & \text{if } \tau_i(\omega) < \infty \\ X_{i,\infty}(\omega) & \text{if otherwise} \end{cases} \tag{B.2}$$

τ_i is a stopping time, since $\mathbf{1}_{\{\tau_i \leq n\}}$ is $\mathcal{F}_{i,n}$ -measurable random variable, i.e., $\mathbf{1}_{\{\tau_i \leq n\}} \in m\mathcal{F}_{i,n}$.

We now prove that:

$$E(X_{\tau_i}) = E(X_{i,1}) = E[E(\tilde{p}_i | \sigma(\tilde{\gamma}_{i,1}))] = E^0(\tilde{p}_i) = \tilde{p}_i^0 \tag{B.3}$$

From the definition of $X_{i,n}$ and from the fact that \tilde{p}_i is uniformly bounded it follows that $X_{i,n}$ is a UI martingale relative to the natural filtration $\{\mathcal{F}_{i,n}\}$. Therefore, applying Lévy’s ‘Upward’ theorem again, we have $X_{i,n} \rightarrow X_{i,\infty} := E(\tilde{p}_i | \mathcal{F}_{i,\infty})$ in \mathcal{L}^1 . Since $E|X_{i,n} - X_{i,\infty}| \geq |E(X_{i,n} - X_{i,\infty})| = |E(X_{i,n}) - E(X_{i,\infty})|$, it follows that $|E(X_{i,n}) - E(X_{i,\infty})| \rightarrow 0$. Moreover, we have $E(X_{i,1}) = E(X_{i,\infty})$ because $X_{i,n}$ is a martingale.

Now consider a stopped process:

$$X_{\tau_i \wedge n} = \begin{cases} X_{i,n} & \text{if } n \leq \tau_i \\ X_{\tau_i} & \text{if } n > \tau_i \end{cases} \tag{B.4}$$

Then as an immediate consequence of Theorem 10.9 of Williams (1991), $X_{\tau_i \wedge n}$ is a martingale and $E(X_{\tau_i \wedge n}) = E(X_{i,1})$. Moreover, $X_{\tau_i \wedge n}$ is UI, because $X_{i,n}$ is a uniformly bounded martingale and $X_{i,n} \rightarrow X_{i,\infty}$ in \mathcal{L}^1 .

Finally, applying Lévy’s ‘Upward’ theorem to the UI martingale $X_{\tau_i \wedge n}$ we conclude that $X_{\tau_i \wedge n} \rightarrow X_{\tau_i \wedge \infty}$ in \mathcal{L}^1 . But by definition of $X_{\tau_i \wedge n}$,

$$X_{\tau_i \wedge \infty}(\omega) = \begin{cases} X_{i,\infty}(\omega) & \text{if } \tau_i = \infty \\ X_{\tau_i(\omega)}(\omega) & \text{if } \tau_i < \infty \end{cases} \tag{B.5}$$

Hence, $X_{\tau_i \wedge \infty}(\omega) = X_{\tau_i(\omega)}(\omega)$ almost everywhere and we obtain:

$$E(X_{i,1}) = E(X_{\tau_i \wedge n}) = E(X_{\tau_i \wedge \infty}) = E(X_{\tau_i}). \tag{B.6}$$

(II) Recall that $\Gamma^t = \times_{\tau=1}^t \Gamma$ is a t -fold Cartesian product of the support of the random productivity vector corresponding to entry of new firms in the industries in the set T_0 . Now, define $\Gamma^\infty = \times_{\tau=1}^\infty \Gamma$ and $\Gamma^{-t} = \times_{\tau=t+1}^\infty \Gamma$. Define sigma algebras: (i) $\mathcal{F}_n = \sigma(\cup_{i \in S_0} \mathcal{F}_{i,n})$, (ii) $\mathcal{F}_\infty = \sigma(\cup_{n=1}^\infty \mathcal{F}_n)$ and (iii) $\mathcal{F}(\Omega) = \sigma(\times_{i=1}^M [0,1] \times \Gamma^\infty)$ on the set $\Omega = \times_{i=1}^M [0,1] \times \Gamma^\infty$.

Recall that for the vector of the true probabilities of high productivity realizations $\mathbf{p}=(p_1, \dots, p_M)$, the probability of observing a specific sequence of vectors $\gamma^t=(\gamma_1, \dots, \gamma_t) \in \Gamma^t$ is defined by

$$P(\gamma^t | \mathbf{p}) = \prod_{i \in T_0} \text{Prob}(\tilde{\gamma}_{ik} = \gamma_{ik}, k \leq t | p_i) = \prod_{i \in T_0} p_i^{s_i} (1 - p_i)^{t-s_i} \tag{B.7}$$

where γ_{ik} is the realization of the productivity parameter of the k th new firm in industry i and $s_i = \sum_{k=1}^t \mathbf{1}_{\{\tilde{\gamma}_{ik} = \gamma_{ik}^{\text{low}}\}}$ is the number of high productivity (low cost) realizations after entry of t new firms in industry i . Thus, for any $A \in \mathcal{F}_\infty$ such that $A = C \times \Gamma^{-t}$, we can calculate the probability $P(A | \mathbf{p})$ by integrating Eq. (B.7) over the set of sequences $(\gamma_1, \dots, \gamma_t) \in C$.

Given the prior probability density function $f^0(\mathbf{p})$, we can define a probability measure on the set Ω of sample paths $\{\mathbf{p}, \gamma_1, \dots, \gamma_t\}$ as follows:

$$P^0 = \int_0^1 \dots \int_0^1 \left(\prod_{i \in T_0} p_i^{s_i} (1 - p_i)^{t-s_i} \right) f^0(\mathbf{p}) dp_{1_0} \dots dp_{M_0} \tag{B.8}$$

where the subscripts $1_0, \dots, M_0$ denote the industries in the set T_0 ranked by the size of their dynamic allocation indices at time zero. Thus for any $D \in \mathcal{F}(\Omega)$ such that $D = \Theta \times A$ for some $\Theta \in \times_{i=1}^M [0,1]$ and $A \in \mathcal{F}_\infty$, we can write $P^0(D) = \int \cdot \Theta \cdot \int P(A | \mathbf{p}) f^0(\mathbf{p}) dp_{1_0} \dots dp_{M_0}$.

Remark: The a.s. convergence in the above statement of Lévy’s “Upward” theorem was with respect to the measure P^0 .

(III) We showed in (I) that $E(X_{\tau_i}) = \bar{p}_i^0$ for any $i \in T_0$. Therefore, $P^0(\tau_i = \infty, \forall i \in T_0) > 0$, because if otherwise:

$$P^0(\tau_i = \infty, \forall i \in T_0) = 0 \Rightarrow P^0(\tau_i < \infty \text{ for some } i \in T_0) = 1 \Rightarrow E(X_{\tau_i}) < \bar{p}_i^0 \text{ for some } i \in T_0,$$

i.e., τ_i is a.s. finite and we obtain a contradiction.

By definition of P^0 in (II), $P^0(\tau_i = \infty, \forall i \in T_0) > 0$ implies that there exists some set of vectors of true probabilities $\Theta \in \times_{i=1}^M [0,1]$ and an \mathcal{F}_∞ -measurable set of realizations of productivity $A \subset \Gamma^\infty$, such that $E^0(\mathbf{1}_{\mathbf{p} \in \Theta}) > 0$ and $P(A | \mathbf{p}) > 0, \forall \mathbf{p} \in \Theta$. But this just says that if the vector of true probabilities is \mathbf{p} , then the set of industries will survive infinitely long with positive probability under a rejection rule stricter than the one which the government is actually applying in the selection of industries for targeting, because $\{\omega: X_{i,n}(\omega) < \bar{p}_i\} \subset \{\omega: v_i(X_{i,n}(\omega)) < v_i(X_i^*) = m^*\}$. Therefore, if for a given set of true parameter vectors Θ the set of industries survives in the long-run with positive probability for the strict rule, then it will survive for the same or even larger set of true parameter vectors Θ' under the government’s actual policy rule based on the Gittins indices, i.e. $\Lambda_p^{T_0} > 0, \forall \mathbf{p} \in \Theta'$ such that $\Theta \subset \Theta'$.

Let $p_i^* = \inf\{p_i: p \in \Theta'\}$. Note that since the Bernoulli distribution satisfies the monotone likelihood ratio property, we can conclude using a result proved by [Banks and Sundaram](#)

(1992) that $P(A|p)$ is increasing in p . Therefore, the probability of the long-run survival of a technology in the targeted set is increasing in the true probability of the good match between this technology and the local labor. But then for all true probabilities $p \geq p^*$, the technology will survive forever with positive probability.

Conversely, if the inequality $p \geq p^*$ is not satisfied, then it is certain that the initial set of targeted technologies will not survive in the long-run and at least some of them will eventually be abandoned. \square

References

- Aghion, P., Bolton, P., Harris, C., Jullien, B., 1991. Optimal learning by experimentation. *Review of Economic Studies* 58, 621–654.
- Aizenman, J., 1998. New activities, the welfare cost of uncertainty and investment policies. *International Journal of Finance and Economics* 3, 97–110.
- Anantharam, V., Varaiya, P., Walrand, J., 1987. Asymptotically efficient allocation rules for multi-armed bandit problem with multiple plays. Part I: I.I.D. Rewards, Part II: Markovian Rewards. *IEEE Transactions on Automatic Control* AC-32 (11), 968–982.
- Anderson, R., Sonnenschein, H., 1982. On the existence of rational expectations equilibrium. *Journal of Economic Theory* 26, 261–278.
- Aoki, M., Murdock, K., Okuno-Fujiwara, M., 1997. Beyond the East-Asian miracle: introducing the market-enhancing view. In: Aoki, M., et al. (Eds.), *The Role of Government in East Asian Economic Development: Comparative Institutional Analysis*. Oxford Univ. Press, New York, pp. 1–40.
- Azariadis, C., Drazen, A., 1990. Threshold externalities and economic development. *Quarterly Journal of Economics* 105, 501–526.
- Bagwell, K., Staiger, R., 1989. The role of export subsidies when product quality is unknown. *Journal of International Economics* 27, 69–89.
- Banerjee, A.V., 1992. A simple model of herd behavior. *Quarterly Journal of Economics* 107, 797–817.
- Banerjee, A., Fudenberg, D., 1995. Word of mouth learning. Mimeo, MIT.
- Banks, J., Sundaram, R., 1992. Denumerable-armed bandits. *Econometrica* 60, 1071–1096.
- Bardhan, P.K., 1971. On optimum subsidy to a learning industry: an aspect of the theory of infant-industry protection. *International Economic Review* 12, 54–70.
- Bellman, R.E., 1956. A problem in the sequential design of experiments. *Sankhya* 16, 221–229.
- Benkherouf, L., Glazebrook, K., Owen, R., 1992. Gittins indices and oil exploration. *Journal of the Royal Statistical Society. Series B* 54, 229–241.
- Bermanke, B., 1983. Irreversibility, uncertainty, and cyclical investment. *Quarterly Journal of Economics* 98, 85–106.
- Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as information cascades. *Journal of Political Economy* 100, 992–1026.
- Billingsley, P., 1986. *Probability and Measure*. Wiley, New York.
- Blume, L., Easley, D., 1984. Rational expectations equilibrium: an alternative approach. *Journal of Economic Theory* 34, 116–129.
- Brander, J., Spencer, B., 1985. Export subsidies and international market share and rivalry. *Journal of International Economics* 18, 83–100.
- Brander, J., Spencer, B., 1987. Foreign direct investment with unemployment and endogenous taxes and tariffs. *Journal of International Economics* 22, 257–279.
- Calbreath, D., 1995. Volkswagen and Skoda: together forever? *Business Central Europe* 3 (3), 7–10.
- Dinopoulos, E., Lewis, T., Sappington, D., 1995. Optimal industrial targeting with unknown learning-by-doing. *Journal of International Economics* 38, 275–295.
- Dixit, A., Pindyck, R., 1994. *Investment Under Uncertainty*. Princeton Univ. Press, Princeton.
- Easley, D., Kiefer, N.M., 1988. Controlling a stochastic process with unknown parameters. *Econometrica* 56, 1045–1064.

- Gittins, J.C., 1989. *Multi-Armed Bandit Allocation Indices*. Wiley, Chichester.
- Gittins, J.C., Jones, D., 1974. A dynamic allocation index for the sequential allocation of experiments. In: Gani, J., et al. (Eds.), *Progress in Statistics*. North-Holland, Amsterdam, pp. 241–266.
- Grossman, G., Horn, H., 1988. Infant-industry protection reconsidered: the case of informational barriers to entry. *Quarterly Journal of Economics* 103, 767–787.
- Havas, A., 1997. Foreign direct investment and intra-industry trade: the case of the automotive industry in Central Europe. In: Dyker, D. (Ed.), *The Technology of Transition: Science and Technology Policies for Transition Economies*. Central European Univ. Press, Budapest, pp. 211–240.
- Hoff, K., 1997. Bayesian learning in an infant industry model. *Journal of International Economics* 43, 409–436.
- Ishikida, T., Varaiya, P., 1994. Multi-armed bandit problem revisited. *Journal of Optimization Theory and Applications* 83, 113–154.
- Jordan, J.S., 1985. Learning rational expectations: the finite state case. *Journal of Economic Theory* 36, 257–276.
- Kim, H.-K., Ma, J., 1997. The role of government in acquiring technological capability: the case of the petrochemical industry in East Asia. In: Aoki, M., et al. (Eds.), *The Role of Government in East Asian Economic Development: Comparative Institutional Analysis*. Oxford Univ. Press, New York, pp. 101–133.
- Krueger, A., 1985. The experience and lessons of Asia's super exporters. In: Corbo, V., et al. (Eds.), *Export-Oriented Development Strategies: The Success of Five Newly Industrializing Countries*. Westview Press, Boulder, CO, pp. 187–212.
- Kuo, S., 1983. *The Taiwan Economy in Transition*. Westview Press, Boulder, CO.
- Lai, T.L., 1988. Asymptotic solutions of bandit problems. In: Fleming, W., Lions, P.-L. (Eds.), *Stochastic Differential Systems, Stochastic Control Theory and Applications*. Springer, New York, pp. 275–292.
- Lall, S., 1996. *Learning from Asian Tigers*. Macmillan, London.
- Lung, Y., 2000. Is the rise of emerging countries as automobile producers an irreversible phenomenon? In: Humphrey, J., et al. (Eds.), *Global Strategies and Local Realities: The Auto Industry in Emerging Markets*. MacMillan Press, London, pp. 16–41.
- Mandelbaum, A., 1986. Discrete multi-armed bandits and multi-parameter processes. *Probability Theory and Related Fields* 71, 129–147.
- Matsuyama, K., 1995a. New goods, market formations, and pitfalls of system design. *Journal of the Japanese and International Economies* 9, 376–402.
- Matsuyama, K., 1995b. Complementarities and cumulative processes in models of monopolistic competition. *Journal of Economic Literature* 33, 701–729.
- Matsuyama, K., 1997. Economic development as coordination problems. In: Aoki, M., et al. (Eds.), *The Role of Government in East Asian Economic Development: Comparative Institutional Analysis*. Oxford Univ. Press, New York, pp. 134–162.
- McLennan, A., 1984. Price dispersion and incomplete learning in the long run. *Journal of Economic Dynamics and Control* 7, 331–347.
- Meth-Cohn, D., 1995. Breeding lame ducks. *Business Central Europe* 3 (6), 15–16.
- Pack, H., Westphal, L., 1986. Industrial strategy and technological change: theory vs. reality. *Journal of Development Economics* 22, 87–128.
- Roberts, K., Weitzman, M., 1981. Funding criteria for research, development, and exploration projects. *Econometrica* 49, 1261–1288.
- Rodrik, D., 1995. Getting interventions right: how South Korea and Taiwan grew rich. *Economic Policy* 20, 53–107.
- Rodrik, D., 1996. Coordination failures and government policy: a model with applications to East Asia and Eastern Europe. *Journal of International Economics* 40, 1–22.
- Romer, P., 1994. New goods, old theory, and the welfare costs of trade restrictions. *Journal of Development Economics* 43, 5–38.
- Rothschild, M., 1974. A two-armed bandit theory of market-pricing. *Journal of Economic Theory* 9, 185–202.
- Schmalensee, R., 1975. Alternative models of bandit selection. *Journal of Economic Theory* 10, 333–342.
- Shaiken, H., 1995. *Technology and Work Organization in Latin American Motor Vehicle Industries*. Economic Commission for Latin America and the Caribbean, Santiago, Chile.

- Shaiken, H., Herzenberg, S., 1987. Automation and Global Production: Automobile Engine Production in Mexico, the United States, and Canada. Center for U.S.–Mexican Studies, University of California, San Diego, La Jolla, CA.
- Smith, L., 1991. Error persistence, and experiential versus observational learning. Working Paper. The Foerder Institute for Economic Research, Tel Aviv University.
- Smith, L., Sørensen, P., 1996. Rational social learning with random sampling. Mimeo, Department of Economics, M.I.T., Cambridge, MA.
- Spencer, B., Brander, J., 1983. International R&D rivalry and industrial strategy. *Review of Economic Studies* 50, 707–722.
- Stiglitz, J., 1996. Some lessons from the East Asian miracle. *World Bank Research Observer* 11 (2), 151–177.
- Varaiya, P., Walrand, J., Buyukkoc, C., 1985. Extensions of the multi-armed bandit problem: the discounted case. *IEEE Transactions on Automatic Control* AC-30, 426–439.
- Wade, R., 1990. *Governing the Market*. Princeton Univ. Press, Princeton.
- Wah, T.K., Jomo, K.S., 1995. Economic theory and industrial policy in East Asia. In: Fitzgerald, R. (Ed.), *The State and Economic Development: Lessons from the Far East*. Frank Cass, Portland, OR, pp. 17–54.
- Weber, R.R., Weiss, G., 1990. On an index policy for restless bandits. *Journal of Applied Probability* 27, 637–648.
- Weiss, G., 1994. On almost optimal priority rules for preemptive scheduling of stochastic jobs on parallel machines. *Advances in Applied Probability* 27, 821–839.
- Weitzman, M., 1979. Optimal search for the best alternative. *Econometrica* 47, 641–654.
- Whittle, P., 1988. Restless bandits: activity allocation in a changing world. In: Gani, J. (Ed.), *A Celebration of Applied Probability*. *Journal of Applied Probability*, vol. 25A, pp. 287–298.
- Williams, D., 1991. *Probability with Martingales*. Cambridge Univ. Press, Cambridge.