Social Externalities, Overlap and the Poverty Trap

Young Chul Kim　　Glenn C. Loury

- Address for Young Chul Kim:
Korea Development Institute, Hoegiro 47, Dongdaemoon-gu, Seoul 130-868, South Korea
Email: yckim.kdi@gmail.com

- Address for Glenn C. Loury:
Department of Economics, Brown University, Providence, RI 02912, USA
Email: Glenn_Loury@brown.edu

Abstract

Previous studies find that some social groups are stuck in poverty traps because of network effects. However, these studies do not carefully analyze how these groups overcome low human capital investment activities. Unlike previous studies, the model in this paper includes network externalities in both the human capital investment stage and the subsequent career stages. This implies that not only the current network quality, but also the expectations about future network quality affect the current investment decision. Consequently, the coordinated expectation among the group members can play a crucial role in the determination of the final state. I define "overlap" for some initial skill ranges, whereby the economic performance of a group can be improved simply by increasing expectations of a brighter future. I also define "poverty trap" for some ranges, wherein a disadvantaged group is constrained by its history, and I explore the egalitarian policies to mobilize the group out of the trap.

Keywords: Group Inequality, Network Externality, Overlap, Poverty Trap.

JEL Code: I30, J15, Z13

1 Introduction

A human being is socially situated such that familial and communal resources explicitly influence his acquisition of human capital through various routes, including through training resources, nutritional and medical provision, after-school parenting, peer effects, role models, and even the psychological processes that shape one's outlook on life. Even after the skill acquisition period, one's social network influences his economic success through various routes, such as mentoring, job searches, business connections, and information channeling. Empirical studies concerning social externalities demonstrate the persistence of network effects, which include community effects (Kain 1968, Borjas 1995, Cutler and Glaeser 1997, Weinberg et al. 2004), job information effects (Rees and Shultz 1970, Blau and Robins 1990, Munshi 2003), peer effects (Hoxby 2002, Falk and Ichino 2005, Hanushek et al. 2009), and business network effects (Fafchamps and Minten 1999, Khwaja et al. 2009). An extensive body of sociological literature has been devoted to the topic, beginning with the seminal work by Granovetter (1974).[1] These studies imply that the persistence of a social group's low economic status is generated by group-level influences on individuals' skill investment activities and economic outcomes.

A number of theoretical works emphasize network effects and the subsequent development bias that generates between-group disparities. Some of these studies focus on network externalities in the human capital investment stage. For example, Becker and Tomes (1979) and Loury (1981) explain the persistence in relative economic status across generations via the effects of parental income on offspring's education. Akerlof (1997) argues that individual concern for status and conformity are the primary determinants of an individual's educational attainment, child bearing, and law-breaking behaviors. Lundberg and Startz (1998) argue that group disparities in earnings can persist indefinitely when the average level of human capital in a community affects the accumulation of human capital of the following generations. In a related work, Bowles, Loury and Sethi (2007) prove the instability of an equal society in a highly segregated economy under the interpersonal spillovers in human capital accumulation and the production complementarity between high and low skill labor. Other works focus on network externalities in the subsequent career stages. For example, Montgomery (1991, 1994) suggests that the widespread use of employee referrals combined with a tendency to refer others within individuals' social networks might generate persistent inequalities between groups of workers. More recently, Calvó-Armengol and Jackson (2004, 2007) argue that differences in collective employment histories and the consequent asymmetry of job information produce sustained inequality in wages and drop-out rates

[1] Ioannides and Loury (2004) and Durlauf and Fafchamps (2006) provide a rich survey of the literature on this topic.

across social groups.

Although theoretical works on the topic find that some groups are stuck in poverty traps because of network effects, they do not provide a rich analysis of how these groups can overcome the low human capital investment activities. In their models, multiple equilibria of skill investment rates are self-confirmed under the presence of strong network externalities. What can we say about the skill levels between those stationary points? How can a group with a lower skill investment rate change the skill investment behaviors of its group members? In the theoretical models, players are assumed to be myopic. However, what if they are farsighted and can cooperate with each other? This paper tries to answer these questions by suggesting that coordinated expectations among group members play a critical role in the determination of a group's skill investment activities.

As mentioned earlier, previous studies focus on network externalities in either the human capital investment stage or the subsequent career stages. Externalities in the former stage alter the cost to achieve skills and those in the latter stages alter the benefits from skill investments. Unlike previous studies, the theoretical model in this paper allows both types of externalities. Therefore, not only the current network quality, but also the expectations regarding future network quality affect current invest decisions. This implies that multiple equilibrium paths can exist for a certain range of initial skill levels. Within that range, a group's economic performance can be improved simply by increasing expectations for a brighter future. This provides rational support for the argument that some social groups are not constrained by their history but can raise themselves up "by their bootstraps." However, some groups may suffer from very low network quality if their skill levels are far below a certain range. This provides rational support for the argument that some groups may not escape their low skill investment activities without external interventions.

This paper illustrates the mechanism through concrete examples. The model considers both kinds of network externalities. These two effects operate via different channels. For network externalities in the education period, the change in a group's status tends to be caused by factors in the past; by altering skill investment cost, the current stock of network human capital directly affects the investment rate in a newborn cohort. In contrast, for network externalities in the working period, the change in a group's status tends to be affected by factors in the future; by altering the future benefits to skill acquisition, the expected success of one's network influences skill investment in an entering cohort. The latter effect implies a unique feature of collective action: the possibility of group members acting together to improve or deteriorate the quality of a group's social network. For instance, suppose that a group's network quality is relatively poor but that a newborn cohort happens to believe that the quality of the group's network will be better in the future. If this belief leads more newborn group members to acquire skills, then the next cohort will find that the overall network quality has improved due to the enhanced skill investment of the previous cohort. If they and the following cohorts continue to hold the optimistic view toward the future, they will maintain the enhanced skill investment rate and the quality of group's social network will improve over time – thereby justifying the optimistic beliefs of earlier cohorts. However, suppose that the newborn cohort holds a pessimistic view that the network quality will be even worse in the future. Fewer members invest in skill achievement because the expected benefits are fewer. If the following cohorts continue to hold the pessimistic view, the network quality will deteriorate over time. Thus, this pessimistic belief could also be self-fulfilling. This nature of possible collective actions stresses the importance of coordinated expectations across different time cohorts. Whether optimism or pessimism persists across cohorts determines the final economic state of a social group.

However, collective action through coordinated beliefs may not be feasible for all social groups with unequal network quality. The potential impact of altered beliefs is restricted by the strength of education period network externalities. The skill improvement that occurs due to coordinated optimism may not be feasible if the negative influence of the current network quality is too strong. This is the situation of "the past" that traps disadvantaged groups.

Therefore, the analysis of the dynamic structure of network externalities focuses on the identification of the network quality range in which both the optimistic and the pessimistic expectations are feasible for the group members. Krugman (1991) denoted the range with multiple equilibrium paths by *overlap* in his influential argument for the relative importance of history and expectations. In his argument, within an overlap, the final economic state is determined by expectations toward the future, while it is determined by history outside an overlap. Unlike his model with a fixed population, my model is developed based on the overlapping generation framework. The model emphasizes the importance of belief coordination over the long-term horizon: the expectations coordinated across the different time cohorts impact the dynamic path to be taken.

If there is no overlap, then history is always decisive. With greater overlap, expectations play a greater role in the determination of the final state. The model developed in this paper proves that the size of an overlap is determined by the relative strength of working-period network externalities over the education period network externality. This is consistent with the facts that the education period network externality operates as *a historical force* that restricts a group to be subject to the current network quality and that the working period network externalities operate as *a mobilization force* that leads a group to enhance (shrink) the skill investment activities by holding an optimistic (pessimistic) view about the future network quality.

Finally, the model in this paper provides some new perspectives on egalitarian policies such as affirmative action by considering economic agents' forward-looking behaviors. This point distinguishes the paper from other papers concerning the egalitarian policies whose main focus is on the equilibrium analysis (e.g., Coate and Loury 1993, Fryer and Loury 2007). If the initial network quality of a social group is far below the "overlap", the group may be trapped by the negative influence of network effects, and an active state role is required to enhance the group's skill level to enter the "overlap". However, if the network quality of a social group is already in the overlap range, the active state role would not have a significant impact on the group's skill level. Instead, an emphasis on coordinated optimism among the disadvantaged group members should be pursued, although this fact is often ignored in policy debates. Civic leaders, civic organizations, religious groups, and governments may all contribute to the encouragement of collective optimism. Therefore, an effective policy to mobilize a disadvantaged group out of the poverty trap first requires active governmental intervention and then requires societal belief coordination. A policy that fails in either respect cannot be successful in helping the group to advance as much as an advantaged group. This point is illustrated further later in this paper in the examination of a multiple group society.

The paper is organized into the following sections. Section 2 describes the basic structure of the model with social network externalities. Section 3 develops the dynamic model with the newborn cohort's forward-looking decision making and the dynamic evolution of group skill levels. Section 4 identifies multiple stationary states in the dynamic model. Section 5 identifies the equilibrium paths to those stationary states and the consequent overlap. In Section 6, we extend the model to examine a multiple group society with social interactions between groups and discuss the egalitarian policies to mobilize the disadvantaged groups out of the poverty trap. Section 7 provides study conclusions.

2 Social Externalities and Skill Investment Decision

Consider a social group with a large population of workers. A worker is subject to the "Poisson death process" with parameter α: in a unit period, each individual faces α chances to die. We assume that the total population of the group is constant at N, implying that the α fraction of the group's population is replaced by newborn group members in a unit period. Each worker is either skilled or unskilled. Let $s_t \in [0,1]$ denote the fraction of skilled workers in the group at time t, which is called *the group skill level at time t*. An agent's neighbors are n random draws from the group's population, and n is large enough that the quality of an agent's network is approximately equal to the group skill level s_t.[2]

At this stage, we examine the dynamics of social network externalities within a social group. In the second part of this paper, we extend this to allow social interactions between social groups. With the extension, one's network quality is no longer equal to the group skill level s_t, but is also affected by the skill levels of other social groups. The degree of segregation in the society plays a crucial role in the determination of the influence of one's own group skill level.[3]

A newborn's innate ability a is a random draw from a distribution $G(a)$. Each newborn agent born with an innate ability level a decides whether to be skilled or not during his early days of life. Each newborn individual at time t makes a skill investment decision by comparing the cost of skill acquisition with the expected benefits of investment. The cost to achieve a skill at time t depends on innate ability $a \in (-\infty, \infty)$ and the quality of social network at time t as suggested by Bowles, Loury and Sethi (2007): $C_t \equiv C(a, s_t)$. The $C(a, s_t)$ is a decreasing function in both arguments a and s_t. The cost includes both the mental and physical costs that are incurred for the skill achievement. The lower one's innate ability or the worse the quality of one's social network, the more mentally stressful the skill acquisition process is or the more materials he must spend on the achievement.

The expected benefits of investment to a newborn individual born at time t, $\Pi_t \in (0, \infty)$, is the net benefits of his skill investment to be realized over his whole lifetime from time t until he dies. Let us assume the base level salary for a skilled worker is w_1 and w_0 for an unskilled worker, in which $w_1 > w_0$. A worker's neighbors at time t are composed of the ns_t number of skilled workers and the $n(1 - s_t)$ number of unskilled workers. Let $\phi_{ij}(x)$ denote the extra benefits of having x number of j type workers in a i type worker's social network, in which $\phi_{ij}(0) = 0$ for any i, $j \in \{s, u\}$.

First, $\phi_{ss}(ns_t)$ denote the extra benefits of having ns_t skilled workers in a skilled worker's social network, which is an increasing function of ns_t. The benefits are both psychological and material. For instance, job

[2]In the extreme case that n is equal to the population size minus one $(N - 1)$, agents are all connected to one another. This extreme network is called *complete network* in the network literature (Jackson, 2008).

[3]Note that we do not allow the difference in the network quality between a skilled and an unskilled worker in the given model. However, the extension to this dimension would not generate meaningful changes in major results as long as the network quality of a skilled worker is significantly affected by the skill level of his own group.

information flows along the synapses of the social network (Granovetter 1974). The more skilled workers he has in his network, the more likely he is to find an appropriate job position for his specific skills (Holzer 1988). A skilled worker can be more efficient in contacting customers and handling specific work troubles when he has more skilled workers in his network. (Ozgen and Baron 2007). He may gain comfort and mentoring from the informal network, and the cost of maintaining jobs may decline with more skilled workers around him (Castilla 2005, Rockoff 2008).

In the same way, unskilled workers such as car mechanics, construction workers and shop attendants would obtain extra benefits from having unskilled workers in his network. $\phi_{uu}(n(1-s_t))$ denote the extra benefits of having $n(1-s_t)$ unskilled workers in an unskilled worker's social network, which is an increasing function of $n(1-s_t)$. For example, a car mechanic searching for a place to work for will find a car center that fits his specialty better when he has more car mechanics in his network. He can be more efficient in handling specific mechanical problems in his work when he can confer with more mechanics. A construction worker (or a shop attendant) will have more chances to find new job openings with more construction workers (or shop attendants) in his network.

Even a skilled worker may obtain extra benefits from having unskilled workers in his network, denoted by $\phi_{su}(n(1-s_t))$, but to a lesser degree than an unskilled worker would obtain: $\phi'_{su}(n(1-s_t)) < \phi'_{uu}(n(1-s_t))$. In a symmetric way for having skilled workers, we have $\phi'_{us}(ns_t) < \phi'_{ss}(ns_t)$. Thus, the net benefits of being a skilled worker realized at time τ is $w_s + \phi_{ss}(ns^e_\tau) + \phi_{su}(n(1-s^e_\tau)) - w_u - \phi_{uu}(n(1-s^e_\tau)) - \phi_{us}(ns^e_\tau)$, in which s^e_τ indicates the expected network quality at the future point of time τ. Replacing the baseline salary differential $w_s - w_u$ with $\bar{\delta}$, and $\phi_{ss}(ns^e_\tau) + \phi_{su}(n(1-s^e_\tau)) - \phi_{uu}(n(1-s^e_\tau)) - \phi_{us}(ns^e_\tau)$ with $f(s^e_\tau)$, the lifetime net benefits of skill investment to an agent born at time t are summarized as

$$\Pi^e_t = \int_t^\infty [\bar{\delta} + f(s^e_\tau)] e^{-(\rho+\alpha)(\tau-t)} \mathrm{d}\tau, \tag{1}$$

where ρ is a time-discounting factor and α is a Poisson death rate. Also note that $f'(s^e_\tau) > 0$, which implies the social increasing returns emphasized by Acemoglu(1996). We assume that $\bar{\delta}$ is big enough that $\bar{\delta} + f(0) > 0$ so that the net benefits of being skilled is always positive. Thus, the net benefits of skill investment Π^e_t are an increasing function of both the baseline salary differential $\bar{\delta}$ and the sequences of expected network quality $\{s^e_\tau\}^\infty_{\tau=t}$: $\Pi^e_t \equiv \Pi(\bar{\delta}, \{s^e_\tau\}^\infty_{\tau=t})$. The higher rate of return on skill investment with the more skilled workers in the social group seems natural with the existence of social network externalities as long as the group's skill level does not affect significantly the aggregate skill composition of the economy.[4]

An agent born at time t with an innate ability a commits a skill investment if and only if

$$C(a, s_t) \leq \Pi(\bar{\delta}, \{s^e_\tau\}^{\infty}_{\tau=t}). \tag{2}$$

The notable feature of this argument is that one's skill investment is affected by both the current network externalities in the skill acquisition period and the future network externalities over the labor market phase of his career. The interplay between the two kinds of network externalities has not yet been explored by other theoretical works including recent developments by Bowles et al. (2007) and Calvó-Armengol and Jackson (2004). The above formula generates a unique threshold ability level such that newborn individuals of the social group whose innate ability is at least the threshold invest in the skill acquisition. Let us define a function A that represents the unique threshold ability $\tilde{a}$: $\tilde{a}_t \equiv A(s_t, \Pi^e_t)$. Using the distribution of the innate ability level $G(a)$, the fraction of individuals born at time t who invest in the skill, denoted by x_t, is expressed by

$$x_t = 1 - G(A(s_t, \Pi^e_t)). \tag{3}$$

3 Dynamic System with Social Network Externalities

The skill investment of the newborns can be approximated by the following procedure. Consider a very short time interval between t and $t + \Delta t$. Suppose that, at the beginning of the interval, the randomly chosen $\alpha \Delta t$ fraction of the group's population, which is the $N \cdot \alpha \Delta t$ number of workers, die, and the same number of agents are newly born. The $N \cdot (1 - \alpha \Delta t)$ workers of the group survive until $t + \Delta t$. At the end of the interval, the newborn agents incur the cost of skill achievement. Then, the threshold level of ability $\tilde{a}$ for the skill investment is determined by the following equation: $C(\tilde{a}, s_t) = \Pi(\bar{\delta}, \{s^e_\tau\}^{\infty}_{\tau=t+\Delta t})$. The fraction of the individuals who are born at time t who invest in skill (x_t) is $1 - G(A(s_t, \Pi^e_{t+\Delta t}))$. The total number of skilled workers at time $t + \Delta t$ will be the sum of skilled workers in the surviving population and those in the newborn cohort: $N \cdot (1 - \alpha \Delta t) \cdot s_t + N \cdot \alpha \Delta t \cdot [1 - G(A(s_t, \Pi^e_{t+\Delta t}))]$. Thus, the group skill level at time $t + \Delta t$

[4] However, some scholars even suggest that the skill premium may depend positively on the the aggregate skill composition given the market frictions. For example, Acemoglu (1996) develops a mechanism for social increasing returns that the rate of return on human capital of a worker is increasing in the average human capital of the workforce when the labor market is characterized by costly search.

is approximated as

$$s_{t+\Delta t} \approx (1-\alpha\Delta t)\cdot s_t + \alpha\Delta t\cdot[1-G(A(s_t,\Pi^e_{t+\Delta t}))]. \tag{4}$$

Rearranging the equation gives us

$$\frac{\Delta s_t}{\Delta t} \equiv \frac{s_{t+\Delta t}-s_t}{\Delta t} \approx \alpha\left[1-G(A(s_t,\Pi^e_{t+\Delta t}))-s_t\right]. \tag{5}$$

Taking $\Delta t \to 0$, we achieve the evolution rule of a group's skill level s_t,

$$\dot{s}_t = \alpha\left[1-G(A(s_t,\Pi^e_t))-s_t\right]. \tag{6}$$

This can be expressed as $\dot{s}_t = \alpha\left[x_t - s_t\right]$ because of equation (3). If the fraction of newborn agents who invest in skill (x_t) is greater than the current skill level of the group (s_t), the network quality improves at time t. Otherwise, it declines.

Also, taking the derivative with respect to time t in equation (1), we have the evolution rule of the net benefits of skill investment Π^e_t,

$$\dot{\Pi}^e_t = (\rho+\alpha)\left[\Pi^e_t - \frac{\bar{\delta}+f(s_t)}{\rho+\alpha}\right]. \tag{7}$$

If a normalized level of the currently accrued benefits of being skilled $\left(\frac{\bar{\delta}+f(s_t)}{\rho+\alpha}\right)$ is greater than the lifetime benefits of being skilled that are expected to accrue from now until death (Π^e_t), the lifetime benefits of being skilled expected to accrue from the next time point $t+\Delta t$ to the death ($\Pi^e_{t+\Delta t}$) would be smaller than its current level: $\Pi^e_{t+\Delta t} < \Pi^e_t$. If they are equal, the lifetime net benefits of skill investment would not change within the short time interval Δt: $\Pi^e_{t+\Delta t} = \Pi^e_t$.

The group skill level s_t is constantly adjusted by the level of skill investments among the newborn cohort, which means that it is a flow variable, which cannot make a sudden jump at a point of time. However, the lifetime benefits of skill investment Π^e_t depend on the expectations about future network quality. By altering the expectation of $\{s^e_\tau\}^\infty_{\tau=t}$, the lifetime benefits can make a sudden jump at any point of time. Thus, it is a jumping variable. The dynamic system with network externalities that includes a flow variable s_t and a jumping variable Π_t^e is therefore summarized by the following equations:

$$\begin{aligned} \dot{s}_t &= \alpha(x_t - s_t), \text{ in which } x_t = 1 - G(A(s_t, \Pi_t^e)), && (8) \\ \dot{\Pi}_t^e &= (\rho + \alpha)\left[\Pi_t^e - \frac{\bar{\delta} + f(s_t)}{\rho + \alpha}\right], && (9) \end{aligned}$$

and the two isoclines of the time dependent variables are represented by

$$\begin{aligned} \dot{s}_t = 0 \text{ Locus} &: \quad s_t = 1 - G(A(s_t, \Pi_t^e)), && (10) \\ \dot{\Pi}_t^e = 0 \text{ Locus} &: \quad \Pi_t^e = \frac{\bar{\delta} + f(s_t)}{\rho + \alpha}. && (11) \end{aligned}$$

For further analysis of this dynamic system without damaging its essential structure, we introduce the following linear functional forms of the cost function $C(a, s_t)$ and the benefits function $f(s_t)$: $C(a, s_t) = c_0 - \psi a - p s_t$ and $f(s_t) = f_0 + q s_t$, where p represents the influence of the education period network externality, and q represents the influence of the working period network externality. ψ represents the cost sensitivity to the innate ability level a. The threshold ability level ($\tilde{a}$) for the skill achievement is obtained from the equation $C(\tilde{a}, s_t) = \Pi_t^e$:

$$A(s_t, \Pi_t^e) = \frac{c_0 - p s_t - \Pi_t^e}{\psi}. \tag{12}$$

In order to avoid the massive complications, the model in this paper uses the above linear functional forms. Readers will be able to find that the major results of this paper are derived with the general forms of the cost function $C(a, s_t)$ and the benefits function $f(s_t)$ without the artificial linearization.

4 Multiple Steady States

In this section, we check the possible multiple steady states in the given dynamic system. We start with the simplest case wherein the innate ability is equal across the population: $a \equiv \bar{a}$. Then, we examine the case with the general form of the ability distribution.

4.1 Steady States with Unique Ability Level

In the unique system with the equal innate ability $\bar{a}$, the investment rate among the newborns is either $x_t = 1$ or $x_t = 0$. All of the newborn agents at time t invest in skill achievements when the expected benefits are no less than the cost: $x_t = 1$ with $\Pi_t^e \geq C(\bar{a}, s_t)(= c_0 - \psi\bar{a} - ps_t)$. This implies that the equation $\dot{s}_t = 0$ holds when $\Pi_t^e \geq c_0 - \psi\bar{a} - p$ and $s_t = 1$ as $x_t = s_t = 1$. Also, for the given $s_t = 1$, the equation $\dot{\Pi}_t^e = 0$ holds when $\Pi_t^e = \frac{\bar{\delta}+f_0+q}{\rho+\alpha}$, according to equation (11). Therefore, there exists a steady state with $s_t = 1$ when $\frac{\bar{\delta}+f_0+q}{\rho+\alpha} \geq c_0 - \psi\bar{a} - p$. Let us normalize $\frac{\bar{\delta}}{\rho+\alpha}$ as $\bar{\delta}'$, which represents the lifetime level of the wage differential $\int_t^\infty \bar{\delta} \cdot e^{-(\rho+\alpha)(\tau-t)} d\tau$, and normalize $\frac{\rho}{\rho+\alpha}$ and $\frac{f_0}{\rho+\alpha}$ as ρ' and f_0'. The high investment rate with $x_t = 1$ and the high quality of the social network with $s_t = 1$ is a steady state when the composite network externalities $(p + q')$ are large enough that $p + q' \geq c_0 - \psi\bar{a} - \bar{\delta}' - f_0'$.

None of the newborns invest in the skill when the expected benefits is smaller than the cost, $x_t = 0$ with $\Pi_t^e < C(\bar{a}, s_t)(= c_0 - \psi\bar{a} - ps_t)$. This implies that the equation $\dot{s}_t = 0$ holds when $\Pi_t^e < c_0 - \psi\bar{a}$ and $s_t = 0$ as $x_t = s_t = 0$. For the given $s_t = 0$, the equation $\dot{\Pi}_t^e = 0$ holds when $\Pi_t^e = \frac{\bar{\delta}+f_0}{\rho+\alpha}$ (equation (11)). Therefore, another steady state that includes the low investment rate $x_t = 0$ and the low quality social network with $s_t = 0$ exists when the following holds: $c_0 - \psi\bar{a} - \bar{\delta}' - f_0' > 0$. The two steady states are displayed in Figure 1 with the two isocline loci.

Proposition 1. *Given a unique innate ability level $\bar{a}$ for all the newborns and assuming $c_0 - \psi\bar{a} - \bar{\delta}' - f_0' > 0$, multiple steady states exist, $(s_l, \Pi_l) = (0, \bar{\delta}' + f_0')$ and $(s_h, \Pi_h) = (1, \bar{\delta}' + f_0' + q')$, if and only if the composite network externalities $(p + q')$ are big enough that $p + q' \geq c_0 - \psi\bar{a} - \bar{\delta}' - f_0'$.*

This proposition proves that sufficient network externalities must be present for multiple steady states to exist. In addition, under the absence of network externalities, a unique steady state always exists. If the base salary differential $\bar{\delta}'$ is big enough, then the high network quality $(s_t = 1)$ is self-confirmed at the unique steady state. Otherwise, the low network quality $(s_t = 0)$ is self-confirmed.

Corollary 1. *Given a unique innate ability level $\bar{a}$ for all the newborns, there exists a unique steady state under the absence of network externalities $(p = q = 0)$. The unique steady state is $(1, \bar{\delta}' + f_0')$ if $\bar{\delta}' \geq c_0 - \psi\bar{a} - f_0'$ and $(0, \bar{\delta}' + f_0')$ otherwise.*

4.2 Steady States with Ability Distribution $G(a)$

In the simplest case with a unique ability label $\bar{a}$, we have shown that the multiplicity of steady states is generated by the influence of network externalities, and, without the network externalities, the multiplicity is not achieved. In this section, we confirm this conclusion with the more general form of the innate ability function $G(a)$. Suppose $G(a)$ is S-shaped. Then, there exists $\hat{a}$ such that $G(a)'' > 0$ for any $a \in (-\infty, \hat{a})$ and $G(a)'' < 0$ for any $a \in (\hat{a}, \infty)$, which implies that its PDF $g(a)$ has one peak at $\hat{a}$ (e.g., a bell-shaped $g(a)$).

The $\dot{s}_t = 0$ locus in equation (10) is represented by (s_t, Π_t^e)s that satisfy the following two equations, which are associated with the threshold ability level ($\tilde{a}$) for the skill achievement.

$$\begin{cases} s_t = 1 - G(\tilde{a}) \\ \tilde{a} = A(s_t, \Pi_t^e) \end{cases} \tag{13}$$

The first is denoted by the solid curve in Panel A of Figure 2 in the $(s_t, \tilde{a})$ domain, and the second is denoted by the dotted lines for each level of Π_t^e (iso-Π lines) in the same panel. The slope of the $\dot{s}_t = 0$ locus is obtained through the implicit function theorem: defining a function F as $F = 1 - G(A(s_t, \Pi_t^e)) - s_t$, we have $\frac{\mathrm{d}\Pi_t^e}{\mathrm{d}s_t}\big|_{\dot{s}_t=0} \left(\equiv -\frac{F_s}{F_\Pi}\right) = -\frac{G'(\tilde{a})\cdot A_s + 1}{G'(\tilde{a})\cdot A_\Pi}$. Because $A_s = -p\psi^{-1}$ and $A_\Pi = -\psi^{-1}$ (equation (12)), we have the following lemma.

Lemma 1. *The slope at an arbitrary point (s', Π') on the $\dot{s}_t = 0$ locus is $\frac{\psi}{g(\tilde{a}')} - p$, in which $\tilde{a}' = A(s', \Pi')$ and $\tilde{a}' = G^{-1}(1 - s')$.*

For the specific ability level $\hat{a}$ under which $g(a)$ is maximized, the corresponding $\hat{s}$ on the $\dot{s}_t = 0$ locus is $\hat{s} \equiv 1 - G(\hat{a})$, and the corresponding $\hat{\Pi}$ is the Π_t^e that satisfies $\hat{a} = A(\hat{s}, \Pi_t^e)$: $\hat{\Pi} \equiv c_0 - \psi\hat{a} - p\hat{s}$.

Suppose the $\dot{\Pi}_t^e = 0$ locus passes through the specific point $(\hat{s}, \hat{\Pi})$, as displayed in Panel B of Figure 2. Note that the slope of the $\dot{s}_t = 0$ locus, $\frac{\psi}{g(\tilde{a})} - p$, is minimized at the point because $g(\tilde{a})$ is maximized with $\tilde{a} = \hat{a}$. In this case, it is obvious that multiple steady states exist if and only if the slope of the $\dot{\Pi}_t^e = 0$ locus is greater than that of the $\dot{s}_t = 0$ locus at the point $(\hat{s}, \hat{\Pi})$: $q' > \frac{\psi}{g(\hat{a})} - p$. Thus, we conclude that the multiplicity of the steady states is achieved when the composite influence of the network externalities measured by $p + q'$ is big enough that $p + q' > \frac{\psi}{g(\hat{a})}$.

Proposition 2. *Suppose that the $\dot{\Pi}_t^e = 0$ locus passes through $(\hat{s}, \hat{\Pi})$ on the $\dot{s}_t = 0$ locus, in which $A(\hat{s}, \hat{\Pi}) =$*

$\hat{a}$. *Three steady states exist if and only if the composite network externalities* $(p + q')$ *are big enough that* $p + q' > \frac{\psi}{g(\hat{a})}$.

Let us denote the three steady states by $E_l(s_l, \Pi_l)$, $E_m(s_m, \Pi_m)$ and $E_h(s_h, \Pi_h)$, in which $s_l < s_m (= \hat{s}) < s_h$, as displayed in Figure 2. With the low quality social network s_l now and in the future, the expected benefits of skill investment are low (Π_l). With the low level of the network quality s_l together with the low level of benefits to investment Π_l, the ability threshold for skill achievement among the newborns is high and a relatively small fraction of the newborns invest in skills. In this manner, the low quality social network is self-confirmed, which is represented by the steady state $E_l(s_l, \Pi_l)$. With the high quality social network s_h now and in the future, the expected benefits of skill investment are high (Π_h). With the high levels of network quality (s_h) and the net benefits to investment (Π_h), the ability threshold for skill achievement is lower and a relatively large fraction of the newborns invest in skills. Thus, the high quality social network is also self-confirmed, which is represented by the steady state $E_h(s_h, \Pi_h)$.

The existence of multiple steady states is possible only when the influence of network externalities is sufficiently strong. Consider the economy under the absence of network externalities $(p = q = 0)$. The $\dot{\Pi}^e_t = 0$ locus is flat because of $q = 0$: $\Pi^e_t = \bar{\delta}' + f'_0$. The $\dot{s}_t = 0$ locus is $\Pi^e_t = -\psi G^{-1}(1 - s_t) + c_0$ because of $p = 0$, according to equations (10) and (12), which is a monotonically increasing function. Therefore, a unique steady state always exists at $(1 - G((c_0 - \bar{\delta}' - f'_0)\psi^{-1}), \bar{\delta}' + f'_0)$ without the network externalities.

Corollary 2. *A unique steady state* $(1 - G((c_0 - \bar{\delta}' - f'_0)\psi^{-1}), \bar{\delta}' + f'_0)$ *exists under the absence of the network externalities* $(p = q = 0)$.

Proposition 2 can be generalized further as follows.

Theorem 1. *If and only if the composite network externalities* $(p + q')$ *are big enough that* $p + q' > \frac{\psi}{g(\hat{a})}$, *does a range of the base salary differential* $[\bar{\delta}_2, \bar{\delta}_1]$ *exist such that the multiple steady states exist with any* $\bar{\delta}$ *within the range:* $\bar{\delta}_j = (\rho + \alpha)(c_0 - \psi G^{-1}(1 - s_j) - (p + q')s_j - f'_0), \forall j \in \{1, 2\}$, *in which both* s_1 *and* s_2 *satisfy* $g(G^{-1}(1 - s_j)) = \psi(p + q')^{-1}$ *and* $s_1 < \hat{s} < s_2$. *Otherwise, a unique steady state exists regardless of the base salary differential level* $\bar{\delta}$.

Proof. For the proof, see the appendix. ■

The theorem implies there are three steady states with $\bar{\delta}$ between $\bar{\delta}_2$ and $\bar{\delta}_1$ given $p + q' > \frac{\psi}{g(\hat{a})}$ and two with either $\bar{\delta} = \bar{\delta}_2$ or $\bar{\delta} = \bar{\delta}_1$ because the $\dot{\Pi}^e_t = 0$ locus is tangent to the $\dot{s}_t = 0$ locus. The above theorem confirms the importance of network externalities in generating multiple steady states. If the influence of network externalities ($p + q'$) is weak, the different steady states never include multiple social groups. When the influence is sufficiently strong, however, we can have multiple social groups at the steady states with the different network qualities. However, multiplicity is only possible when the base salary differential $\bar{\delta}$ is in a certain range $[\bar{\delta}_2, \bar{\delta}_1]$. If it is either large enough that $\bar{\delta} > \bar{\delta}_1$ or small enough that $\bar{\delta} < \bar{\delta}_2$, a unique steady state exists $(\bar{s}, \bar{\Pi})$ in spite of the strong presence of network externalities. With $\bar{\delta} > \bar{\delta}_1$, the benefits of skill investment are so great that a large fraction of the newborns invest in skills, and the high network quality $\bar{s}(> \hat{s})$ is solely self-confirmed. In the same manner, when $\bar{\delta} < \bar{\delta}_2$, a small fraction of the newborns invest in skills and the low network quality $\bar{s}(< \hat{s})$ is solely self-confirmed.

Corollary 3. *Even when the composite network externalities* $(p + q')$ *are greater than* $\frac{\psi}{g(\hat{a})}$*, there exists a unique steady state* $(\bar{s}, \bar{\Pi})$ *with the base salary differential* $\bar{\delta}$ *out of the range* $[\bar{\delta}_2, \bar{\delta}_1]$*. The steady state satisfies* $\bar{s} > \hat{s}$ *and* $\bar{\Pi} > \hat{\Pi}$ *with* $\bar{\delta} > \bar{\delta}_1$*, and* $\bar{s} < \hat{s}$ *and* $\bar{\Pi} < \hat{\Pi}$ *with* $\bar{\delta} < \bar{\delta}_2$*.*

5 Dynamic Equilibrium Paths

In this section, we identify the converging dynamic paths to the steady states and provide the economic interpretation of the paths.

5.1 Dynamic Paths with Unique Ability Level

In the simplest case of the identical ability level ($\bar{a}$) across the population, we have identified two possible steady states, assuming the conditions in Proposition 1 are satisfied. Denoting them by E_l and E_h, they are $E_l(0, \bar{\delta}' + f_0')$ and $E_h(1, \bar{\delta}' + f_0' + q')$. To examine the converging dynamic paths to the steady states, we need a phase diagram with direction arrows, which are displayed in Figure 3, in which the four dynamic regimes are classified by the two straight lines, $\Pi_t^e = c_0 - \psi\bar{a} - ps_t$ and $\Pi_t^e = \bar{\delta}' + f_0' + q's_t$. Let us denote the intersection of the two straight lines by E_m: $E_m(\frac{c_0-\psi\bar{a}-\bar{\delta}'-f_0'}{p+q'}, \frac{p(\bar{\delta}'+f_0')+q'(c_0-\psi\bar{a})}{p+q'})$. The two converging paths to E_h and E_l spiral out of the intersection E_m. The dynamic path converging to E_h above the two straight lines is determined by the following dynamic system, $\dot{s}_t = -\alpha s_t - \alpha$ and $\dot{\Pi}_t^e = (\rho+\alpha)\Pi_t^e - qs_t - \bar{\delta} - f_0$ because of $x_t = 1$ in the regime (equations (8) and (9)). The optimistic path is summarized by $\Pi_t^e = \frac{q}{\rho+2\alpha}s_t + \bar{\delta}' + f_0' + \frac{\alpha q'}{\rho+2\alpha}$. Also, the dynamic path converging to E_l below the two straight lines is determined by the dynamic system,

$\dot{s}_t = -\alpha s_t$ and $\dot{\Pi}_t^e = (\rho+\alpha)\Pi_t^e - qs_t - \bar{\delta} - f_0$ because $x_t = 0$. This pessimistic path is summarized by $\Pi_t^e = \frac{q}{\rho+2\alpha}s_t + \bar{\delta}' + f_0'$. Therefore, we can calculate both the lower bound of s_t for the converging path to E_h and the upper bound of s_t for the converging path to E_l, denoted by e_o and e_p for each:

$$\begin{cases} e_o = \max\left\{\left[p + \frac{q}{\rho+2\alpha}\right]^{-1}\left(c_0 - \psi\bar{a} - \bar{\delta}' - f_0' - \frac{\alpha q'}{\rho+2\alpha}\right), 0\right\} \\ e_p = \min\left\{\left[p + \frac{q}{\rho+2\alpha}\right]^{-1}\left(c_0 - \psi\bar{a} - \bar{\delta}' - f_0'\right), 1\right\}. \end{cases} \tag{14}$$

With an initial social network quality s_0 between the two bounds, $s_0 \in [e_o, e_p]$, two coordinated equilibrium paths are available to the social group: the optimistic path to E_h and the pessimistic path to E_l. If the coordinated expectation about the future is optimistic across the generations, the expected benefits of the skill achievement at time zero (Π_0^{op}) are $\frac{q}{\rho+2\alpha}s_0 + \bar{\delta}' + f_0' + \frac{\alpha q'}{\rho+2\alpha}$, and the expected benefits of the skill achievement among the following newborn cohorts are greater than the level: $\Pi_t^{op} > \Pi_0^{op}, \forall t \in (0,\infty)$. The newborn cohort and all following cohorts invest in skills: $x_t = 1, \forall t \in [0,\infty)$, which means that the skill level of the group s_t improves over time until it reaches one. However, if the coordinated expectation is pessimistic across the generations, the expected benefits at the initial point (Π_0^{pe}) is $\frac{q}{\rho+2\alpha}s_0 + \bar{\delta}' + f_0'$, which is smaller than Π_0^{op} by as much as $\frac{\alpha q'}{\rho+2\alpha}$, and the expected benefits of investment among the following newborn cohorts are smaller than the level for the current newborns: $\Pi_t^{pe} < \Pi_0^{pe}, \forall t \in (0,\infty)$. The newborn cohort and all following cohorts do not invest in skills: $x_t = 0, \forall t \in [0,\infty)$, which means that the skill level of the group deteriorates over time until it reaches zero.

Let us denote the range $[e_o, e_p]$ by *overlap*, in which multiple coordinated equilibrium paths are available, as suggested by Krugman (2001). Outside the overlap, a unique equilibrium path exists that is either optimistic or pessimistic. If the initial network quality is good enough that $s_0 > e_p$, the only reasonable expectation about the newborns' investments is $x_t = 1, \forall t \in [0,\infty)$. If it is poor enough that $s_0 < e_o$, the only reasonable expectation is $x_t = 0, \forall t \in [0,\infty)$.

The size of the overlap, denoted by $L(\equiv e_p - e_o)$, is essential to understand the characteristics of the economy: $L = \frac{\alpha}{(\rho+2\alpha)(\alpha+\rho)\frac{p}{q}+(\alpha+\rho)}$, as far as $e_o, e_p \in (0,1)$. The bigger the L, the more likely it is that the coordinated expectation about the future is critical in the determination of the final skill level of the group. Because L is a decreasing function with respect to p/q, we have the following result.

Proposition 3. *Given a unique innate ability level $\bar{a}$ for all newborns, the size of overlap (L) is positively* *related to the relative strength of the working period network externality in comparison with the education period network externality (q/p). As the influence of the working period network externalities (q) increases and the influence of the education period network externalities (p) decreases, the more likely it is that the coordinated expectation will determine the final state of a social group's skill level instead of the history.*

Moreover, the overlap does not exist under the absence of working period network externalities: $L = 0$ when $q = 0$. This means that the skill investment activities of the newborns are subject to the "past" if network externalities are active only over the education period. That is, the initial quality of the social network determines the future, and the belief coordination across the generations does not have an effect. However, the size of the overlap is maximized under the absence of the education period network externality: $\arg\max_p L = 0, \forall q \in (0, \infty)$. The belief coordination across generations and the consequent collective action are most crucial when network externalities are not active over the education period.

Suppose a group's current skill level is in the overlap range. Then, the economic performance of the group can be improved simply by increasing expectations of a brighter future. This rationalizes the arguments of those who suggest that social groups are not constrained by their history but can raise themselves up "by their bootstraps." However, suppose that a group is poor enough that its current skill level (s_0) is below the lower bound of the optimistic path (e_o). The group cannot escape its miserable condition through belief coordination or collective actions among the group members. Under this situation, a disadvantaged group is trapped by its own history.

Definition 1 (Poverty Trap)**.** *A social group is in the poverty trap if its network quality (s_0) is below the lower bound of the optimistic path: $s_0 < e_o$.*

Consider two social groups, A and B, at the different steady states. Group A's skill level is one, and group B's is zero. Assuming that the overlap range is between zero and one, group B is in the poverty trap and group A is out of it. The disparity between the two groups cannot be overcome without governmental intervention. Suppose that the government helps group B improve its skill level and enter the overlap range. At this stage, the crucial point is the belief coordination among group B members. If optimism prevails, the newborn cohorts invest in skills and the group's skill level improves consistently up to the level of group A. If pessimism prevails, the newborn cohorts do not invest in skills and the skill level can even deteriorate over time until reaching group B's original level of zero, making the earlier governmental intervention useless.

In addition, at the lower steady state, group B can escape the poverty trap if the economic condition changes favorably. The optimistic path to E_h becomes available to the group when $e_o = 0$ in the formula (14). This implies that the group escapes the poverty trap when the base salary differential ($\bar{\delta}$) increases up to $(\rho + \alpha)\left(c_0 - \psi\bar{a} - f_0' - \frac{\alpha q'}{\rho+2\alpha}\right)$. With the greater benefits of being skilled, it becomes easier to induce collective actions among the young cohorts in the disadvantaged group.

5.2 Dynamic Paths with Ability Distribution $G(a)$

In this section, we confirm the findings presented in the earlier section using the more general form of the ability distribution, S-shaped $G(a)$. Suppose that the condition for multiplicity is satisfied in Theorem 1. Suppose that the following three distinct steady states exist: $E_l(s_l, \Pi_l)$, $E_m(s_m, \Pi_m)$ and $E_h(s_h, \Pi_h)$, in which $s_l < s_m < s_h$. This is achieved with $\bar{\delta} \in (\bar{\delta}_2, \bar{\delta}_1)$ in Theorem 1. Using equations (8) and (9), $\dot{s}_t$ is positive (negative) above (below) the $\dot{s}_t = 0$ locus, $\dot{\Pi}_t^e$ is positive (negative) above (below) the $\dot{\Pi}_t^e = 0$ locus. The phase diagram with direction arrows is displayed in Figure 4. The characteristics of the steady states are summarized by the following lemma.

Lemma 2. *Among the three steady states, $E_l(s_l, \Pi_l)$, $E_m(s_m, \Pi_m)$ and $E_h(s_h, \Pi_h)$, in which $s_l < s_m < s_h$, E_l and E_h are saddle points and E_m is a source.*

Proof. For the proof, see the appendix. ■

We can identify the equilibrium path (saddle path) to each saddle point, E_l and E_h, as described in Figure 4. The equilibrium paths spiral out of a source E_m. The lower bound of the optimistic path to E_h (e_o) is smaller than the upper bound of the pessimistic path (e_p): $e_o < e_p$. Within the overlap range $[e_o, e_p]$, multiple coordinated equilibrium paths exist. If the coordinated expectation is optimistic, the upper path is taken and the skill level approaches s_h. If it is pessimistic, the lower path is taken and the skill level approaches s_l. Outside the overlap, a unique reasonable equilibrium path exists, which is either an optimistic path to s_h or a pessimistic path to s_l. Thus, within the overlap, the coordinated expectation determines the final state, and the history determines the final state outside the overlap.

The existence of the overlap range is related to the existence of working period network externalities. First, consider the case where working period network externalities are absent ($q = 0$). Because the benefits of skill acquisition are fixed as $\bar{\delta}' + f_0'$ in this case, the expectation about the future does not play any role. This is displayed in the phase diagram of Panel A in Figure 5 as the flat $\dot{\Pi}_t^e = 0$ locus ($\Pi_t^e \equiv \bar{\delta}' + f_0'$) and the non-existence of overlap. If the initial network quality is poor, below s_m, the group skill level converges to the low skill equilibrium s_l. If it is good, above s_m, the skill level converges to the high skill equilibrium s_h. Therefore, the final economic state depends entirely on the initial network quality of the group. Also, the existence of the overlap range is guaranteed by the existence of working period network externalities ($q > 0$). When the future benefits of skill acquisition are affected by the future network quality, one's skill investment should be influenced by other group members' skill investments now and in the future. Collective action to mange expectation can therefore play an important role in the determination of the overall skill investment rate among the newborn cohorts.

Proposition 4. *Suppose that the condition in Theorem 1 is satisfied with $\bar{\delta} \in (\bar{\delta}_2, \bar{\delta}_1)$ such that there exist three distinct steady states (E_l, E_m and E_h). The overlap range does not exist under the absence of the working period network externalities ($q = 0$), and it always exists with the existence of the working period network externalities ($q > 0$).*

Proof. For the proof, see the appendix. ■

The size of the overlap is determined by the relative influence of the working period network externalities over the education period network externalities. This is because, with the greater q given a fixed level of p, the expected benefits of being skilled are more affected by future network quality. The coordinated expectation across the generations can play a greater role in the determination of the final state. Thus, with the greater q, the size of the overlap, in which the coordinated expectation rather than history determines the final state, tends to be larger. This point is illustrated by the following theorem.

Theorem 2. *Suppose that the condition in Theorem 1 is satisfied with $\bar{\delta} \in (\bar{\delta}_2, \bar{\delta}_1)$ such that there exist three distinct steady states (E_l, E_m and E_h). With the increased influence of the working period network externalities (greater q) while holding a steady state E_m at (s_m, Π_m), the optimistic path to E_h with greater q is placed above the original optimistic path to E_h for any $s_t \geq s_m$, and the pessimistic path to E_l with greater q is placed below the original pessimistic path to E_l for any $s_t \leq s_m$ (Refer to Panel B of Figure 5).*

Proof. For the proof, see the appendix. ■

The above theorem implies that, with the greater q given any p, the overlap is more likely to be bigger as the distance between two equilibrium paths gets wider. The result is consistent with the earlier result with a unique ability level $\bar{a}$ in Proposition 3. Also, the theorem suggests that, given any initial s_0 within the overlap, the difference between the expected benefits of investments with an optimistic view (Π_0^{op}) and that with a pessimistic view (Π_0^{pe}) tends to be greater with the increased influence of working period network externalities (greater q) because the distance between the optimistic path and the pessimistic path gets wider: the greater $\Pi_0^{op} - \Pi_0^{pe}$ with the greater q given p. This means that, with the greater q relative to p, we are more likely to observe a greater difference in the newborns' skill investment activities between the case with the coordinated optimism and that with the coordinated pessimism because $x_0^{op} - x_0^{pe} = G(A(s_0, \Pi_0^{pe})) - G(A(s_0, \Pi_0^{op}))$ in equation (3). This further illustrates that the coordinated expectation can generate a greater difference in newborns' skill investment activities with the greater influence of working period network externalities relative to the education period network externalities.

6 Extension to Social Interactions between Groups

In the preceding sections, we examined the dynamics of network externalities within a social group. We extend this to allow social interactions between social groups. As mentioned earlier, one key difference is that one's network quality is no longer equal to the skill level of the group that he belongs to, but is also affected by the skill levels of other social groups. With this extension, we can examine dynamic structures with different levels of societal integration and analyze the effectiveness of policy interventions.

The analysis in this section focuses on the two-group economy because the most interesting features associated with social interactions are contained in the two-group economy. The two social groups are denoted by group one and group two. Population shares among the two group members are denoted by β^1 and β^2, respectively, with $\beta^1 + \beta^2 = 1$. Let s_t^i denote the fraction of skilled workers in group $i \in \{1, 2\}$ at time t, which is called *group i skill level at time t*. The fraction of skilled workers among the two groups' populations is then $\bar{s}_t \equiv \beta^1 s_t^1 + \beta^2 s_t^2$. Let σ_t^i denote the fraction of skilled workers in the social network of an individual belonging to group i at time t, which is called *group i network quality at time t*. This depends on the levels of human capital in each of the two groups and the extent of segregation η: $\sigma_t^i \equiv \eta s_t^i + (1 - \eta)\bar{s}_t$, as used in Chaudhuri and Sethi (2008) and Bowles, Loury and Sethi (2007). With full integration ($\eta = 0$), σ_t^i is equal to $\bar{s}_t$ for any group i, thus indicating no difference in the network quality across social groups. With full segregation

($\eta = 1$), σ_t^i equals s_t^i. The σ_t^i is a convex combination of s_t^i and s_t^j with weights k^i and $1 - k^i$,

$$\sigma_t^i \equiv k^i s_t^i + (1 - k^i) s_t^j, \text{ where } k^i = \eta + (1 - \eta)\beta^i. \tag{15}$$

The k^i represents the degree of influence of the skill level of one's own and $1 - k^i$ represents that of the other group's skill level.

The cost to achieve the skill for a group i agent who is born with an innate ability a is $C(a, \sigma_t^i)$, instead of $C(a, s_t^i)$. The lifetime net benefits of skill investment to a group i agent born at time t is $\Pi_t^i = \int_t^\infty [\bar{\delta} + f(\sigma_\tau^i)] e^{-(\rho+\alpha)(\tau - t)} \mathrm{d}\tau$, in which $f(\sigma_\tau^i) = \phi_{ss}(n\sigma_\tau^i) + \phi_{su}(n(1-\sigma_\tau^i)) - \phi_{uu}(n(1-\sigma_\tau^i)) - \phi_{us}(n\sigma_\tau^i)$, because a group i worker's neighbors at time τ, which are n random draws from the two groups' populations, are composed of $n\sigma_\tau^i$ number of skilled workers and $n(1 - \sigma_\tau^i)$ number of unskilled workers. An agent born in time t belonging to group i with an innate ability a commits a skill investment if and only if

$$C(a, \sigma_t^i) \leq \Pi(\bar{\delta}, \{\sigma_\tau^i\}_{\tau=t}^\infty). \tag{16}$$

This generates the unique threshold ability: $\tilde{a}_t^i \equiv A(\sigma_t^i, \Pi_t^i)$. The fraction of individuals born in time t of group i who invest in the skill, denoted by x_t^i, is expressed by $x_t^i = 1 - G(A(\sigma_t^i, \Pi_t^i))$. Because the evolution rule of group skill level can be expressed as $\dot{s}_t^i = \alpha(x_t^i - s_t^i)$ (equation (8)), we have the evolution rule for the group i skill level,

$$\dot{s}_t^i = \alpha[1 - G(A(\sigma_t^i, \Pi_t^i)) - s_t^i]. \tag{17}$$

Also, taking the derivative with respect to time t, we have the evolution rule of the net benefits of skill investment Π_t^i,

$$\dot{\Pi}_t^i = (\rho + \alpha)\left[\Pi_t^i - \frac{\bar{\delta} + f(\sigma_t^i)}{\rho + \alpha}\right]. \tag{18}$$

Therefore, the dynamic system with two flow variables s_t^1 and s_t^2 and two jumping variables Π_t^1 and Π_t^2 is summarized by the following four-variable differential equations:

$$\begin{aligned}\dot{s_t}^1 &= \alpha\left[1 - G(A(\sigma_t^1, \Pi_t^1)) - s_t^1\right] \\ \dot{s_t}^2 &= \alpha\left[1 - G(A(\sigma_t^2, \Pi_t^2)) - s_t^2\right] \\ \dot{\Pi}_t^1 &= (\rho+\alpha)\left[\Pi_t^1 - \frac{\bar{\delta} + f(\sigma_t^1)}{\rho+\alpha}\right] \\ \dot{\Pi}_t^2 &= (\rho+\alpha)\left[\Pi_t^2 - \frac{\bar{\delta} + f(\sigma_t^2)}{\rho+\alpha}\right],\end{aligned}$$

in which $A(\sigma_t^i, \Pi_t^i) = \frac{c_0 - p\sigma_t^i - \Pi_t^i}{\psi}$ (equation (12)) and $f(\sigma_t^i) = f_0 + q\sigma_t^i$, and σ_t^1 and σ_t^2 satisfy the following given the societal segregation level η:

$$\begin{cases}\sigma_t^1 = & k^1 s_t^1 + (1-k^1)s_t^2, \text{ with } k^1 = \eta + (1-\eta)\beta^1 \\ \sigma_t^2 = & k^2 s_t^2 + (1-k^2)s_t^1, \text{ with } k^2 = \eta + (1-\eta)\beta^2.\end{cases} \tag{19}$$

6.1 Steady States

Suppose that the condition for multiplicity is satisfied in Theorem 1 with $\bar{\delta} \in (\bar{\delta}_2, \bar{\delta}_1)$ such that the following three distinct steady states exist in a dynamic system for an isolated social group: $E_l(s_l, \Pi_l)$, $E_m(s_m, \Pi_m)$ and $E_h(s_h, \Pi_h)$, in which $s_l < s_m < s_h$. Let $(s^{1**}, s^{2**}, \sigma^{1**}, \sigma^{2**}, \Pi^{1**}, \Pi^{2**})$ denote a steady state satisfying $\dot{s}_t^1 = \dot{s}_t^2 = \dot{\Pi}_t^1 = \dot{\Pi}_t^2 = 0$, where two sets, (s^{1**}, s^{2**}) and $(\sigma^{1**}, \sigma^{2**})$, are bijective with parameters η, β^1 and β^2. First, let us identify "partial" steady states $(s^{i*}, \sigma^{i*}, \Pi^{i*})|_{s^j}$ which are $(s_t^i, \sigma_t^i, \Pi_t^i)$s that satisfy both $\dot{s}_t^i = \dot{\Pi}_t^i = 0$ and $\sigma_t = k^i s_t^i + (1-k^i)s_t^j$, given s_t^j. The following three conditions characterize the set of partial steady states $(s^{i*}, \sigma^{i*}, \Pi^{i*})|_{s^j}$:

$$\dot{s}_t^i = 0 \text{ condition} \quad : \quad s^{i*} = 1 - G(A(\sigma^{i*}, \Pi^{i*})) \tag{20}$$

$$\dot{\Pi}_t^i = 0 \text{ condition} \quad : \quad \Pi^{i*} = \frac{\bar{\delta} + f(\sigma^{i*})}{\rho+\alpha} \tag{21}$$

$$\text{clearing condition} \quad : \quad \sigma^{i*} = k^i s^{i*} + (1-k^i)s^j. \tag{22}$$

To combine the first and second conditions, let us introduce a function $\tilde{A}(\sigma^i) \equiv A(\sigma^i, \frac{\bar{\delta}+f(\sigma^i)}{\rho+\alpha})$. $\tilde{A}(\sigma^i)$ is a decreasing linear function of σ^i because $\tilde{A}(\sigma^i) = (c_0 - \bar{\delta}' - f_0' - (p+q')\sigma^i)\cdot\psi^{-1}$. Then, the partial steady states given s^j are determined by $(s^{i*}, \sigma^{i*})s$ that satisfy both $s^{i*} = 1 - G(\tilde{A}(\sigma^{i*}))$ and $\sigma^{i*} = k^i s^{i*} +$

$(1-k^i)s^j$. Appendix Figure 1 identifies the partial steady states using the two equations. Note that the $s^{i*} = 1 - G(\tilde{A}(\sigma^{i*}))$ locus must pass through three symmetric points, (s_l, s_l), (s_m, s_m) and (s_h, s_h) in the (σ^i, s^i) domain, because s_l, s_m and s_h satisfy equations (10) and (11).

The second step is to collect all partial steady states $(s^{2*}, \sigma^{2*}, \Pi^{2*})|_{s^1}$ and $(s^{1*}, \sigma^{1*}, \Pi^{1*})|_{s^2}$ in order to identify (global) steady states $(s^{1**}, s^{2**}, \sigma^{1**}, \sigma^{2**}, \Pi^{1**}, \Pi^{2**})$. Panel A of Appendix Figure 2 indicates the former and Panel B indicates the latter. In the top figure of Panel A, the slashed lines with different levels of s^1 help identify (s^{2*}, σ^{2*}) for each level of s^1. Note that the slope of the slashed line is $\frac{1}{k^2}$. Consequently, all partial steady states are denoted by $s^{2*}(s^1)$ locus in the bottom figure. In the same manner, in the top figure of Panel B, the slashed lines with different levels of s^2 help identify (s^{1*}, σ^{1*}) for each level of s^2. The slope of the slashed line is k^1. All partial steady states are denoted by $s^{1*}(s^2)$ locus in the bottom figure of the panel. As we overlap the two partial steady state loci, $s^{2*}(s^1)$ and $s^{1*}(s^2)$, we identify the (global) steady states in Panel C.[5]

Using equations (20), (21) and (22), each partial steady state locus $s^{i*}(s^j)$ is characterized by

$$s^{i*}(s^j) \text{ locus} : s^{i*} = 1 - G(\tilde{A}(k^i s^{i*} + (1-k^i)s^j)), \quad \forall s^j \in [0,1]. \tag{23}$$

This implies that a unique s^j exists that corresponds to a partial steady state s^{i*}. For further analysis, let us define a useful function $D^j(s^{i*})$ as the unique s^j given $s^{i*} \in [0,1]$ that satisfies the above formula (23). That is, using $\tilde{A}(\sigma^i) = (c_0 - \bar{\delta}' - f_0' - (p+q')\sigma^i) \cdot \psi^{-1}$,

$$D^j(s^{i*}) = \frac{-\psi G^{-1}(1-s^{i*}) + c_0 - \bar{\delta}' - f_0'}{(p+q')(1-k^i)} - \frac{k^i s^{i*}}{1-k^i}, \quad \forall s^{i*} \in [0,1]. \tag{24}$$

Then, the (global) steady states (s^{1**}, s^{2**}) are a set of (s^{1*}, s^{2*})s that satisfy the following two equations: $D^2(s^{1*}) = s^{2*}$ and $D^1(s^{2*}) = s^{1*}$.

Lemma 3. *The function $D^j(s^{i*})$ is concave with $s^{i*} < 1 - G(\hat{a})$, and convex with $s^{i*} > 1 - G(\hat{a})$.*

Proof. For the proof, see the appendix. ■

Therefore, the partial steady state loci are composed of a concave part and a convex part. The shapes of the loci for each level of segregation η are summarized by the following lemma.

[5] Note that $k^1 < k^2$ due to $\beta^1 < \beta^2$ in the given example in Appendix Figure 2, which generates the curvature difference between the two loci.

Lemma 4. *As η declines, the partial steady state locus $s^{i*}(s^j)$ tends to be "flatter", which means the distance between the locus and the diagonal $|D^j(s^{i*}) - s^{i^*}|$ shrinks as η declines. Given a segregation level η among the two social groups, the partial steady state locus $s^{i*}(s^j)$ is "steeper" for the larger population group (the bigger β^i) in the sense that the distance between the locus and the diagonal $|D^j(s^{i*}) - s^{i^*}|$ is greater for the larger population group than for the smaller population group.*

Proof. Note that $|D^j(s^{i*}) - s^{i*}| = \left|\frac{-\psi G^{-1}(1-s^{i*})+c_0-\bar{\delta}'-f_0'}{(p+q')} - s^{i*}\right| \cdot \frac{1}{1-k^i}$. The first derivative with respect to η gives $\frac{\partial|D^j(s^{i*})-s^{i*}|}{\partial\eta} = \left|\frac{-\psi G^{-1}(1-s^{i*})+c_0-\bar{\delta}'-f_0'}{(p+q')} - s^{i*}\right| \cdot \frac{1}{(1-\beta^i)(1-\eta)^2} > 0$. Also, given a segregation level η, the first derivative with respect to β^i gives $\frac{\partial|D^j(s^{i*})-s^{i*}|}{\partial\beta^i} = \left|\frac{-\psi G^{-1}(1-s^{i*})+c_0-\bar{\delta}'-f_0'}{(p+q')} - s^{i*}\right| \cdot \frac{1}{(1-\beta^i)^2(1-\eta)} > 0$. ■

The $D^j(s^{i*})$ curve gets closer to the diagonal as η declines, which implies the $s^{i*}(s^j)$ locus tends to be "flatter" as η declines. Figure 6 displays how the slopes of the loci change with the different levels of η. The lemma implies that the bigger the size of the group, the more distant the curve $D^j(s^{i*})$ is from the diagonal. Suppose group 1 is the minority and group 2 is the majority ($\beta^1 < \beta^2$). Figure 6 shows that the locus for group 2, the $s^{2*}(s^1)$ locus (the $D^1(s^{2*})$ curve), is "steeper" than that for group 1 for a given segregation level η and less sensitive to the change in the societal segregation level η because the population size of the group is bigger than that of group 1.

Let us denote four regions in the (s^1, s^2) plane by Regions 1, 2, 3 and 4 in clockwise order, which are divided by one vertical line ($s^1 = s_m$) and one horizontal line ($s^2 = s_m$). Additionally, the left and top region is denoted by Region 1, as displayed in Panel C of Appendix Figure 2. The total number of steady states should be nine with full segregation, as illustrated in Panel A of Figure 6. It should be three with full integration because the steady states will be equal to those in the case with one social group, as illustrated in Panel F of the figure. In general, using the above lemmas, we have the following results in terms of the number of steady states.

Proposition 5. *The total number of steady states decreases from nine to three as η declines from one to zero (from full segregation to full integration): the number of steady states decreases from three to zero in Regions 1 and 3, and there is always a unique steady state in Regions 2 and 4. Regardless of η, three symmetric steady states always exist (s_l, s_l), (s_m, s_m) and (s_h, s_h). All other steady states are asymmetric.*

Proof. For the proof, see the appendix. ■

6.2 Stable Manifolds and Convergence Ranges

For notational simplicity, we adopt a notation rule for steady states in the following sections.

Notation 1. *When $\eta = 1$, each steady state (s_i, s_j) is denoted by Q_{ij} for $i, j \in \{l, m, h\}$ (Refer to Panel A of Figure 6). As η declines, each steady state is denoted by its original notation at $\eta = 1$.*

The locations of the symmetric steady states do not change with varying η: Q_{ll}, Q_{mm} and Q_{hh}. Other asymmetric steady states (Q_{ij}s with $i \neq j$) move continuously as η changes, as displayed in Figure 6. With the above notation rule, when we have less than nine steady states, we can denote each by following its original notation in the economy with $\eta = 1$.

Let us define an *economically stable state* as a steady state (s^1, s^2) for which there exists a converging dynamic path for any (s_t^1, s_t^2) in the neighborhood of the state. For example, the following four states are economically stable with $\eta = 1$: Q_{ll}, Q_{hh}, Q_{lh} and Q_{hl}. In geometry, a collection of points on all converging paths to a limit set Q is called *a stable manifold to the limit set Q*.[6] The stable manifold to Q_{ij}, denoted by SM_{ij}, is a collection of $(s_0^1, s_0^2, \Pi_0^1, \Pi_0^2)$s that converge to an economically stable state Q_{ij}: $SM_{ij} \equiv \{(s_0^1, s_0^2, \Pi_0^1, \Pi_0^2) \in \mathbb{R}^4 | (s_t^1, s_t^2, \Pi_t^1, \Pi_t^2)|_{(s_0^1, s_0^2, \Pi_0^1, \Pi_0^2)} \to Q_{ij}\}$. The convergence range to Q_{ij}, denoted by M_{ij}, is defined as a projection of the stable manifold to Q_{ij} to the (s^1, s^2) plane: $M_{ij} \equiv \{(s_0^1, s_0^2) \in [0,1]^2 | (s_t^1, s_t^2, \Pi_t^1, \Pi_t^2)|_{(s_0^1, s_0^2, \Pi_0^1, \Pi_0^2)} \to Q_{ij}\}$.

Before we move to the analysis of the given dynamic system, let us check the simplest case, in which the social network externalities over the working period are negligible and only those over the education period are effective ($p > 0$ and $q = 0$). When $q = 0$, history is important such that an initial skill level determines the final state as we previously checked for the case with one social group. There are a maximum of four stable states in the given dynamic system: Q_{ll}, Q_{lh}, Q_{hl} and Q_{hh}. The basins of attractions (convergence ranges) for those states are separated by separatrices that connect Q_{mm} to saddle points (Q_{mh}, Q_{ml}, Q_{hm}, Q_{lm}). The convergence ranges to those states given $q = 0$ are illustrated in Figure 6. Note that the convergence ranges do not overlap with $q = 0$. As displayed in Panels E and F, when the two group society is relatively more integrated (smaller η), the final state tends to be an equal society regardless of the initial skill composition (s_0^1, s_0^2) because only two convergence ranges exist, M_{ll} and M_{hh}. When the two group society is relatively more segregated (larger η) and the initial skill disparity ($|s_0^1 - s_0^2|$) is great, the final state tends to be an

[6] A limit set in geometry is the state a dynamic system reaches after an infinite amount of time has passed by either going forward or backward in time.

unequal society, either M_{lh} and M_{hl}, as displayed in Panels A and B.

Now we move to the analysis concerning the existence of both kinds of network externalities ($p > 0$ and $q > 0$). The first step is to check the stability of the steady states.

Lemma 5. *Regardless of η and the combination (β^1, β^2), steady states Q_{ll}, Q_{hh}, Q_{lh} and Q_{hl} are economically stable states if they exist, and all other steady states are economically unstable states.*

Proof. For the proof, see the appendix. ■

The economically stable states are illustrated in Figure 7 with dark circles. The number of economically stable states decreases from four to two as η declines. The convergence ranges and their overlapped areas are depicted for each level η in the same figure. Two convergence ranges M_{lh} and M_{hl} disappear at some level of η, and two stable states Q_{lh} and Q_{hl} disappear. Two convergence ranges M_{ll} and M_{hh} tend to expand as η declines, and the other two convergence ranges, M_{lh} and M_{hl}, tend to shrink. All convergence ranges tend to be larger with the greater influence of working period network externalities (greater q) because, as we have shown, the size of overlap tends to be bigger with the greater q in the case with one social group (Theorem 2). The following theorem summarizes these characteristics of the convergence ranges.

Theorem 3. *As η declines, the convergence ranges M_{ll} and M_{hh} tend to expand, and the convergence ranges M_{lh} and M_{hl} tend to shrink. All manifold ranges tend to be larger with the stronger influence of working period network externalities (greater q).*

Proof. For the proof, see the appendix. ■

The interpretation of the overlapped areas is as follows. Consider a case in which all four economically stable states exist with a relatively large η, where the following holds $s_l < e_o < s_m < e_p < s_h$. (eg. Panels A and B in Figure 7). The four possible convergence ranges, M_{ll}, M_{hh}, M_{lh} and M_{hl}, overlap in the middle of the (s^1, s^2) domain. When both groups' skill levels are equally mediocre, the final state will be Q_{ll} (Q_{hh}) if both group members' coordinated expectations are optimistic (pessimistic), and the final state will be either Q_{lh} or Q_{hl} if the coordinated expectation among members of one social group is optimistic and that among members of the other group is pessimistic. Two of the four convergence ranges also overlap near the middle. Consider an overlap of M_{lh} and M_{hh} located in the upper middle area, in which group 2's skill level is relatively good and that of group 1 is mediocre. The final state is either Q_{lh} or Q_{hh} depending on the coordinated expectation among the members of group 1. However, the expectation among the members of group 2 does not affect the final state, because it is determined as optimistic. Finally, in each corner area, there exists a unique convergence range without overlaps. Consider a corner range in the upper left area, in which the skill level of group 2 is good enough but that of group 1 is very bad. Group 1 is in the poverty trap such that the current network quality of group 1 is so poor that there is no way to recover the skill investment rate of the group through belief coordination among group members.

Now consider the case of the existence of two economically stable states with a relatively small η (e.g., Panels E and F in Figure 7). The two social groups are sufficiently integrated that the final state is either a high skilled equal state (Q_{hh}) or a low skilled equal state (Q_{ll}). In the area with two convergence ranges overlapped, collective optimism in both groups leads to Q_{hh} and collective pessimism to Q_{ll}.

6.3 Policy Intervention for the Egalitarian Society

During the Jim Crow period of US history and until the civil rights movement in the 1960s, African Americans were segregated from whites and discriminated against in an overt manner in the US labor market. Although overt discrimination decreased in recent decades, we still observe persistent skill disparities between the blacks and the whites. In particular, Black youths have significantly lower academic achievement than white youths. The dynamic model in this paper suggests that Blacks with more low-skilled people in their social network tend to be trapped by the adverse effects of poor network quality in the segregated American society. Further, the model implies that the optimistic coordinated expectation is important for behavioral changes among Black youth. The persistent group disparity that occurs in many other segregated societies including South Africa, Australia and many countries in Latin America may also be attributable to the poverty trap through the social externalities channel and the failure of belief coordination.

Consider a society with two social groups in which group 1 is in a poverty trap and group 2 has good network quality. According to the following proposition, Q_{hh} is a Pareto-dominant steady state. The government has an incentive to pursue egalitarian policies to mobilize the society toward equality of skills (Q_{hh}).

Proposition 6. *For any segregation level $\eta < 1$, (s_h, s_h) is a strictly Pareto-dominant steady state and (s_l, s_l) is a strictly Pareto-inferior steady state.*

Proof. Because $|D^j(s^{i*}) - s^{i*}|$ is monotonically decreasing as η declines (Lemma 4), both groups' skill levels are less than s_h at any steady state with $\eta < 1$, except the fixed steady state (s_h, s_h). Also, both groups' skill levels are greater than s_l at any steady state with $\eta < 1$, except the fixed steady state (s_l, s_l). ■

6.3.1 Integration Policy

First, note that the asymmetric stable state Q_{lh} disappears with the facilitation of integration. The threshold level of segregation $\hat{\eta}$ for its disappearance depends on the population size of the disadvantaged group, as discussed in Bowles, Loury and Sethi (2007) with a simpler setup.

Lemma 6. *There exists $\hat{\beta}$ such that $\hat{\eta}(\beta^1)$ is strictly decreasing in $(0, \hat{\beta})$ and strictly increasing in $(\hat{\beta}, 1)$.*

Proof. For the proof, see the appendix. ■

As integration proceeds, either 1) Q_{lh} and Q_{mh} merge or 2) Q_{lh} and Q_{lm} merge before Q_{lh} disappears. Therefore, before the steady state Q_{lh} disappears, the state must get into either the convergence range M_{hh} or M_{ll} or both ranges because the convergence range M_{hh} (M_{ll}) always covers the unstable state Q_{mh} (Q_{lm}). Using this fact together with the Lemma above, we have the full picture of the integration effect, which is summarized in Figure 8. In the diagram, $\eta^*(\beta^1)$ indicates the threshold level of η for Q_{lh}'s entering M_{hh}. It is obvious that $\eta^*(\beta^1) > \hat{\eta}(\beta^1)$ with $\beta^1 \in (0, \hat{\beta})$, because Q_{lh} enters the convergence range M_{hh} before it disappears. $\eta^{**}(\beta^1)$ indicates the threshold level of η for Q_{lh}'s entering M_{ll}. It is clear that $\eta^{**}(\beta^1) > \hat{\eta}(\beta^1)$ with $\beta \in (\hat{\beta}, 1)$, because Q_{lh} enter the convergence range M_{ll} before it disappears. Let us denote the β^* with which $\eta^*(\beta^*) = \hat{\eta}(\beta^*)$ and the β^{**} with which $\eta^{**}(\beta^{**}) = \hat{\eta}(\beta^{**})$, where $\beta^* > \beta^{**}$, as noted in the figure.

The multiple convergence ranges available in each area are displayed in the figure. The multiplicity of convergence ranges means that the current state is in the overlapped area of convergence ranges, so the final economic state depends on the coordinated expectation among group members. For example, if the current state is in the area with two convergence ranges (M_{lh} and M_{hh}), the final state will be Q_{hh} if the coordinated expectation is optimism among group 1 members. Otherwise, it will be the unequal state Q_{lh}.

The results summarized in the figure suggest that, when the population size of the disadvantaged group β^1 is small enough ($\beta^1 < \beta^*$), integration ($\eta^*(\beta^1)$) can help the economic state enter the convergence range to the high skilled equal state Q_{hh}. (Refer to the shaded area in the figure.) Once the state enters the convergence range M_{hh}, the coordinated expectation among the disadvantaged group members plays a key role in moving into the Pareto-dominant equal society (s_h, s_h).[7] Therefore, the integration policy combined with collective

[7] However, note that the expectation among the advantaged group is also important in the determination of the final state if the size of the disadvantaged group is sufficiently large that $\beta^{**} < \beta^1 < \beta^*$. To achieve the high skilled equal state Q_{hh} with the facilitated integration, the expectation coordination of the advantaged group should not be pessimism because the economic state will move to Q_{ll} with the group's pessimism when the segregation level (η) falls below $\eta^{**}(\beta^1)$, as shown in Figure 8.

optimism in the disadvantaged group can help mobilize the group in the poverty trap. Without optimism, a much greater level of integration ($\hat{\eta}(\beta^1)$ instead of $\eta^*(\beta^1)$) is required to start the mobilization, which may not be achievable due to either political resistance from the advantaged group or homophily effects (Currarini et al. 2009).

6.3.2 Affirmative Action Policies

We consider two types of affirmative action policies, an asymmetric training subsidy and a quota system. First, we consider a training subsidy policy under which the government distributes educational resources in an asymmetric manner: more resources are distributed to disadvantaged group members and less are distributed to advantaged group members. This generates the decreased c_0 for group 1 and the increased c_0 for group 2, and consequently the downward shift of the $D^2(s^{1*})$ curve and the upward shift of the $D^1(s^{2*})$ curve, where $D^j(s^{i*})$ is an increasing function of c_0 in equation (24). The curve shifts are illustrated in Panel A of Figure 9. If the government provides enough subsidies for group 1, the economic state where group 1 is in a poverty trap can enter the convergence range to Q_{hh} (M_{hh}). If the expectation is coordinated properly for optimism among the disadvantaged group members, the economic state can gradually move toward the high skilled equal state (s_h, s_h). However, if the subsidy is insufficient or the disadvantaged group fails in the expectation coordination, the asymmetric subsidy policy will be useless to make significant behavioral changes among young members of the disadvantaged group.

Secondly, consider a quota policy under which government places some unskilled group 1 members in skilled positions that would otherwise go to skilled members of group 2. Suppose the current economic state is $Q_{lh}(s_l', s_h')$. The rate of skilled job positions in the economy is then fixed as $\bar{s} = \beta^1 s_1' + \beta^2 s_h'$. The government intervenes to mitigate the skilled job disparity between the two groups, $|s^2 - s^1|$, which means that a higher fraction of group 1 takes the skilled job positions, and a lower fraction of group 2 takes the skilled job positions under the constraint of $\bar{s} = \beta^1 s_l' + \beta^2 s_h'$. If this external intervention can lead the economic state (s^1, s^2) to enter the convergence range M_{hh} as displayed in Panel B of Figure 9, society can move toward the high-skilled equal state if the expectations of the disadvantaged group members are well coordinated for optimism. However, if the policy intervention is not sufficiently strong, the economic state cannot enter the overlapped area of two convergence ranges (M_{lh} and M_{hh}). The economic state will return to the original unequal state Q_{lh}. Further, even if the quota is imposed properly, the failure of the belief coordination among the disadvantaged group members will make the policy useless.

Therefore, if a social group is trapped by poor network quality, an effective egalitarian policy of either social integration or affirmative action requires both active governmental intervention and societal belief coordination to mobilize the group out of the poverty trap and to help it advance as far as the advantaged group. A policy that fails in either respect would not be successful in eliminating a persistent skill disparity between social groups. However, if the network quality of a disadvantaged group is already in the overlap range, the active governmental intervention would not have a significant impact on the advancement of the group. Instead, an emphasis on coordinated optimism should be pursued, although this point is often ignored in policy debates. Collective action to manage expectation can generate a significant impact on behavioral changes among young members of the disadvantaged group.

7 Further Discussion

In the above policy discussion, we assume that the returns to investment in skills do not depend on the skill composition in the economy. However, if we allow the skill complementarities between high and low skill labor in production, an increase in the skill share of one group can come at the expense of the welfare of the other group. The equal steady state Q_{hh} may no longer be Pareto-dominant: there is a potential for group conflict. A historically advantaged group would resist the introduction of egalitarian policies and would have an interest in maintaining pessimism about the future in the disadvantaged group. Further analysis that considers skill complementarity is left to future research.

Adsera and Ray (1998) argue that *overlap* is generated only when agents have an incentive to choose the option that offers less appealing benefits at the moment of decision. In my model, the incentive originates in the overlapping generation structure. Since agents are given only one chance to choose their occupational type at the early stage of their lives, they choose a type with less appealing benefits at the moment of the skill investment decision while expecting greater benefits to accrue in his lifetime. The dynamic structure with an overlapping generational framework provided here can be useful in the examination of other topics in which external economies exist and populations are replaced by new cohorts. Hauk and Saez-Marti(2002)'s work on the cultural transmission of corruption and Bisin and Verdier(2001)'s work on intergenerational transmission of values adopt the similar structure though they do not explicitly analyze the existence of multiple equilibrium paths and the size of overlap. I expect future theoretical works that include external economies and overlapping generations to adopt the given dynamic structure.

The coordinated expectations over the long-term horizon discussed in this paper is also useful to explain the classic example of the "Big Push" doctrine of Rosenstein-Rodan (1943). He begins by noting that the development of any particular industry may only be privately profitable if an entire set of interlocking industries emerged simultaneously. Thus, the willingness of firms to invest depends on their expectation that other firms will invest in the future. The role of governments in development policy is to create convergent expectations around the high investment. Murphy, Shleifer, and Vishny (1989) and Matsuyama (1991) formalized this example with theoretical models demonstrating that multiple stationary states exist because of increasing returns. The self-fulfilling expectations often allow an escape from the state of pre-industrialization. The model presented in this paper suggests a theoretical framework to formalize the "Big Push" doctrine in the context of human capital development. In a modern economy, economic development is determined by the level of human capital investment, as emphasized by Lucas (1988) and Barro (1991). Suppose that the return to skill investments is significantly affected by the overall skill investment rate in the economy, the microfoundation of which is well discussed in Acemoglu (1996) given the matching imperfections in the labor market. The willingness of newborn cohorts to invest in skills depends on others' skill investments now and in the future. Consider an open economy with a low skill investment rate. If the coordination failure of agents' expectations is the cause of poverty, the state's role in initiating the development process should be to promote optimism within the underdeveloped economy, as noted by Rosenstein-Rodan (1943).

8 Conclusion

This paper examines the effect of coordinated expectation on the improvement of a group' skill level under the presence of network effects. This point distinguishes the model presented here from other theoretical models that focus on the multiplicity of stationary states. The consideration of economic agents' forward-looking behaviors contributed to the development of this dynamic model.

Most theoretical works concerning network effects have examined policies that can eliminate a bad equilibrium by adjusting given parameters. This is appropriate for policy issues for which the number of equilibria can be changed flexibly, but it is not realistic or too mechanical for other issues for which the multiplicity with both good and bad equilibria is natural. The limitation is often generated by their omission of the dynamic perspectives. With the consideration of players' forward-looking behaviors, we can examine policies concretely without maintaining the change of the set of equilibria. The main focus in policy analysis needs to be on the dynamic equilibrium paths, and not limited to specific equilibria.

9 Appendix: Proofs

9.1 Proof of Theorem 1

Lemma 1 indicates the slope of the $\dot{s}_t = 0$ locus can be represented by $\frac{\mathrm{d}\Pi_t^e}{\mathrm{d}s_t}\Big|_{\dot{s}_t=0} = \frac{\psi}{g(G^{-1}(1-s))} - p$, in which $\frac{\mathrm{d}\Pi_t^e}{\mathrm{d}s_t}\Big|_{\dot{s}_t=0} = \infty$ as either $s \to 0$ or $s \to 1$. This implies that there must be at least one steady state. The slope is minimized at $\hat{s}(= 1 - G(\hat{a}))$ with its minimum $\frac{\psi}{g(\hat{a})} - p$, and the slope is decreasing in $(0, \hat{s})$, and increasing in $(\hat{s}, 1)$. Therefore, if the slope of the $\dot{\Pi}_t^e = 0$ locus (q') is greater than the minimum $\frac{\psi}{g(\hat{a})} - p$, we can find two network quality levels, s_1 and s_2, with which the slopes of the two loci are equalized: $\frac{\psi}{g(G^{-1}(1-s))} - p = q'$ for $s \in \{s_1, s_2\}$, in which $s_1 < \hat{s} < s_2$. When the $\dot{\Pi}_t^e = 0$ locus is tangent to the $\dot{s}_t = 0$ locus at either s_1 or s_2, the corresponding Π_1 and Π_2 on the $\dot{s}_t = 0$ locus are $\Pi_j = -\psi G^{-1}(1 - s_j) - ps_j + c_0, \forall j \in \{1, 2\}$, because of equations (10) and (12). Therefore, noting that the $\dot{\Pi}_t^e = 0$ locus shifts up with the greater $\bar{\delta}$, we can find the corresponding $\bar{\delta}_1$ and $\bar{\delta}_2$ at s_1 and s_2: $\bar{\delta}_j(= \Pi_j(\rho+\alpha) - qs_j - f_0) = (\rho+\alpha)(c_0 - \psi G^{-1}(1-s_j) - (p+q')s_j - f_0'), \forall j \in \{1, 2\}$. Given the slope q' of the $\dot{\Pi}_t^e = 0$ locus greater than $\frac{\psi}{g(\hat{a})} - p$, there exist multiple steady states with $\bar{\delta} \in [\bar{\delta}_2, \bar{\delta}_1]$ and the number of steady states is three with $\bar{\delta} \in (\bar{\delta}_2, \bar{\delta}_1)$.

If the slope of the $\dot{\Pi}_t^e = 0$ locus (q') is smaller than the minimum $\frac{\psi}{g(\hat{a})} - p$, we have only one intercept between the two loci for any level of $\bar{\delta}$, because the slope of the $\dot{s}_t = 0$ locus is greater than that of the $\dot{\Pi}_t^e = 0$ locus (q') for any $s_t \in (0, 1)$. In the same reason, if the slope of the $\dot{\Pi}_t^e = 0$ locus (q') equals the minimum $\frac{\psi}{g(\hat{a})} - p$, the two loci cross each other only one time, because the slope of the $\dot{s}_t = 0$ locus is greater than that of the $\dot{\Pi}_t^e = 0$ locus for any $s_t \in (0, 1)$, except for $s_t = \hat{s}$. ■

9.2 Proof of Lemma 2

Given the dynamic system in equations (8) and (9), its linearization around a steady state $(\bar{s}, \bar{\Pi})$ is

$$\begin{aligned} \dot{s}_t &= \alpha[-G'A_s' - 1](s_t - \bar{s}) + \alpha[-G'A_\Pi'](\Pi_t^e - \bar{\Pi}) \\ \dot{\Pi}_t^e &= -f'(s_t - \bar{s}) + (\rho + \alpha)(\Pi_t^e - \bar{\Pi}). \end{aligned}$$

The Jacobian matrix J_E evaluated at a steady state is

$$J_E \equiv \begin{bmatrix} -\alpha G'A_s' - \alpha & -\alpha G'A_\Pi' \\ -f' & \rho + \alpha \end{bmatrix}_{(\bar{s},\bar{\Pi})} = \begin{bmatrix} \alpha \cdot \frac{g(\tilde{a}')k}{\psi} - \alpha & \alpha \cdot \frac{g(\tilde{a}')}{\psi} \\ -q & \rho + \alpha \end{bmatrix}_{(\bar{s},\bar{\Pi})},$$

, in which $\tilde{a}' = A(\bar{s}, \bar{\Pi})$. Consequently, its transpose is $trJ_E = \alpha \cdot \frac{g(\tilde{a}')p}{\psi} + \rho$ and the determinant is $|J_E| = \frac{\alpha(\alpha+\rho)g(\tilde{a}')}{\psi}\left[p + q' - \frac{\psi}{g(\tilde{a}')}\right]$. Since trJ_E is positive, every steady state is unstable. The slope of the $\dot{s}_t = 0$ locus is $\frac{\psi}{g(\tilde{a}')} - p$ (Lemma 1) and the slope of the $\dot{\Pi}_t^e = 0$ locus is q'. At the steady states E_l and E_h, the slope of the $\dot{s}_t = 0$ locus is greater than the slope of the $\dot{\Pi}_t^e = 0$ locus: $p + q' < \frac{\psi}{g(\tilde{a}')}$. At the steady state E_m, the slope of the $\dot{s}_t = 0$ locus is smaller than the slope of the $\dot{\Pi}_t^e = 0$ locus: $p + q' > \frac{\psi}{g(\tilde{a}')}$. Therefore, $|J_E|$ is negative at E_l and E_h, and positive at E_m, which implies E_l and E_h are saddle points and E_m is a source. ■

9.3 Proof of Proposition 4

First, consider the case with $q = 0$. Since the benefits of being skilled is $\bar{\delta}$ at each point of time, the benefits for skill achievements is fixed as $\bar{\delta}' + f_0'$ because $\Pi_0^e = \int_0^\infty (\bar{\delta} + f_0) \cdot e^{-(\rho+\alpha)\tau} d\tau$, for any s_0. Therefore, there cannot be multiple choices of the level of Π_0^e for any given s_0. Second, consider the case with $q > 0$. Consider an initial network quality $s_0 = s_m + \epsilon$ and its converging sequences $\{s_\tau\}_0^\infty$ to s_h. Then, we have $\Pi_0^{op} = \int_0^\infty [\bar{\delta} + f_0 + q s_\tau] e^{-(\rho+\alpha)\tau} d\tau > \frac{\bar{\delta}+f_0+q(s_m+\epsilon)}{\rho+\alpha}$ because $\Pi_0^{op} = \frac{\bar{\delta}+f_0+q(s_m+\epsilon)}{\rho+\alpha} + \int_0^\infty q(s_\tau - (s_m+\epsilon)) e^{-(\rho+\alpha)\tau} d\tau$ and $s_\tau > s_m + \epsilon, \forall \tau \in (0, \infty)$. Π_0^{op} is above the $\dot{\Pi}_t^e = 0$ locus for $s_0 = s_m + \epsilon$ for any small ϵ. Therefore, the converging path to E_h is placed above the $\dot{\Pi}_t^e = 0$ locus around s_m, implying $e_o < s_m$. Also, consider an initial network quality $s_0 = s_m - \epsilon$ and its converging sequences $\{s_\tau\}_0^\infty$ to s_l. Then, in the same way, we can show that Π_0^{pe} is below the $\dot{\Pi}_t^e = 0$ locus for $s_0 = s_m - \epsilon$ for any small ϵ. Therefore, the converging path to E_l is placed below the $\dot{\Pi}_t^e = 0$ locus around s_m, implying $e_p > s_m$. These imply the existence of an overlap in the neighborhood of s_m. ■

9.4 Proof of Theorem 2

Holding a steady state E_m at (s_m, Π_m) requires the $\dot{\Pi}_t^e = 0$ locus to pass through (s_m, Π_m) for different levels of q: the combination $(\bar{\delta}, q)$ should satisfy $\Pi_m = \frac{\bar{\delta}+f_0+q s_m}{\rho+\alpha}$. The q increase by Δq must come with $\bar{\delta}$ decrease by $s_m \Delta q$: $\Delta\bar{\delta} = -s_m \Delta q$. Note that the $\dot{\Pi}_t^e = 0$ locus is rotated in a counter-clockwise direction as displayed in Panel B of Figure 5. We compare the converging dynamic paths in a dynamic system with $q_0 > 0$ with those in a dynamic system with $q_0 + \Delta q$. Let us denote two stable steady states in the dynamic system with q_0 by $E_l(s_l, \Pi_l)$ and $E_h(s_h, \Pi_h)$, and those in the system with $q_0 + \Delta q$ by $E_l'(s_l', \Pi_l')$ and $E_h'(s_h', \Pi_h')$. The $\dot{\Pi}_t^e$ given $q_0 + \Delta q$, denoted by $\dot{\Pi}_t^e(q_0 + \Delta q)$, is smaller (greater) than the $\dot{\Pi}_t^e(q_0)$ for any $s_t > s_m$ ($s_t < s_m$), because $\dot{\Pi}_t^e(q_0 + \Delta q)$ is, using equation (9):

$$\begin{aligned}
\dot{\Pi}_t^e(q_0 + \Delta q) &= (\rho+\alpha)\left[\Pi_t^e - \frac{(\bar{\delta} + \Delta\bar{\delta}) + f_0 + (q_0 + \Delta q)s_t}{\rho+\alpha}\right] \\
&= (\rho+\alpha)\left[\Pi_t^e - \frac{(\bar{\delta} - s_m \Delta q) + f_0 + (q_0 + \Delta q)s_t}{\rho+\alpha}\right] \\
&= \dot{\Pi}_t^e(q_0) - \Delta q(s_t - s_m) \qquad (25)
\end{aligned}$$

There always exist $s_c \in (s_m, s_h)$ such that the optimistic path to E_h in the dynamic system with q_0 intercepts the $\dot{\Pi}_t^e = 0$ locus in the dynamic system with $q_0 + \Delta q$ at $s_t = s_c$. Let us denote Π_t^{op} on the optimistic path in the dynamic system with q_0 ($q_0 + \Delta q$) by C (C') at $s_t = s_c$, A (A') at $s_t = 1$, and B (B') at $s_t = s_m$. First, it is obvious that, over the range $[s_c, s_h)$, the optimistic path to E_h' with $q_0 + \Delta q$ is above the $\dot{\Pi}_t^e = 0$ locus with $q_0 + \Delta q$, and the optimistic path to E_h with q_0 is below the $\dot{\Pi}_t^e = 0$ locus with $q_0 + \Delta q$. Secondly, we show that the optimistic path to E_h' is above the optimistic path to E_h over the range $[s_h, 1]$. Consider a start point at $s_t = 1$: $(1, x)$. In order that the path from the point in the dynamic system with $q_0 + \Delta q$ passes through E_h, x should be greater than A because the state moves down faster with $\dot{\Pi}_t^e(q_0 + \Delta q) < \dot{\Pi}_t^e(q_0)$. A' should be greater than x because the path from the point $(1, A')$ approaches E_h', in which $s_h' > s_h$. Therefore, we have $A' > x > A$. The same logic can be applied for any s_t over the range

$[s_h, 1]$. Thus, the optimistic path to E'_h is above the optimistic path to E_h at any $s_t \in [s_h, 1]$. Lastly, we show that the optimistic path to E'_h is above the optimistic path to E_h over the range $[s_m, s_c)$. Consider a start point at $s_t = s_m$: (s_m, y). In order that the path from the point in the dynamic system with $q_0 + \Delta q$ passes through C, y should be greater than B because the state moves up more slowly with $\dot{\Pi}^e_t(q_0 + \Delta q) < \dot{\Pi}^e_t(q_0)$. B' should be greater than y because the path from the point (s_m, B') approaches (s_c, C') and $C' > C$. Therefore, we have $B' > y > B$. The same logic can be applied for any s_t in the range $[s_m, s_c)$. In sum, we conclude that the optimistic path to E'_h with $q_0 + \Delta q$ is above the path to E_h with q_0 for any $s_t \geq s_m$. In a symmetric way, we can show that the pessimistic path to E'_l with $q_0 + \Delta q$ is below the path to E_l with q_0 for any $s_t \leq s_m$. ■

9.5 Proof of Lemma 3

The first and second derivatives of the function $D^j(s^{i*})$ are

$$\frac{\mathrm{d}D^j(s^{i*})}{\mathrm{d}s^{i*}} = \frac{\psi}{(p+q')(1-k^i)\cdot G'(G^{-1}(1-s^{i*}))} - \frac{k^i}{1-k^i}, \tag{26}$$

$$\frac{\mathrm{d}^2D^j(s^{i*})}{\mathrm{d}s^{i*2}} = \frac{\psi \cdot G''(G^{-1}(1-s^{i*}))}{(p+q')(1-k^i)\cdot [G'(G^{-1}(1-s^{i*}))]^3}. \tag{27}$$

Because $G''(\tilde{a})$ is negative for $\tilde{a} > \hat{a}$, $G''(G^{-1}(1-s^{i*}))$ is negative for $s^{i*} < 1 - G(\hat{a})$, implying the second derivative is negative. Also, $G''(\tilde{a})$ is positive for $\tilde{a} < \hat{a}$, $G''(G^{-1}(1-s^{i*}))$ is positive for $s^{i*} > 1 - G(\hat{a})$, implying the second derivative is positive. ■

9.6 Proof of Proposition 5

It is obvious that the total number of steady states is nine with $\eta = 1$. Regardless of η, the three symmetric states, (s_l, s_l), (s_m, s_m) and (s_h, s_h), are steady states because $\{s_l, s_m, s_h\}$ are skill levels of steady states for one social group case. First of all, I claim that there does not exist a symmetric steady state other than those three. Suppose that there exists a symmetric steady state $(\hat{s}, \hat{s})$ that is not one of the three. Since $\sigma^1 = \sigma^2 = s^1 = s^2 = \hat{s}$, this implies $\dot{s_t}^1 = \alpha\left[1 - G(A(\hat{s}, \Pi^1_t)) - \hat{s}\right] = 0$, $\dot{s_t}^2 = \alpha\left[1 - G(A(\hat{s}, \Pi^2_t)) - \hat{s}\right] = 0$, $\dot{\Pi}^1_t = (\rho + \alpha)\left[\Pi^1_t - \frac{\bar{\delta}+f(\hat{s})}{\rho+\alpha}\right] = 0$ and $\dot{\Pi}^2_t = (\rho + \alpha)\left[\Pi^2_t - \frac{\bar{\delta}+f(\hat{s})}{\rho+\alpha}\right] = 0$. This contradicts that there are only three skill levels (s_l, s_m, s_h) that satisfy equations (8) and (9). Secondly, let us prove that the total number of steady states is three with $\eta = 0$. This is true when there are no asymmetric steady states with $\eta = 0$. Suppose an asymmetric steady state $(\hat{s}^1, \hat{s}^2)$ exists, where $\hat{s}^1 \neq \hat{s}^2$. Since two groups are fully integrated, $\sigma^1 = \sigma^2 = \bar{s}$. Because it is a (global) steady state, it should be a partial steady state as well. By equations (20) and (21), s^{i*} is uniquely determined by σ^{i*}, which implies that $\hat{s}^1 = \hat{s}^2$ when $\sigma^1 = \sigma^2$. This contradicts that $(\hat{s}^1, \hat{s}^2)$ is an asymmetric steady state. Finally, the total number of steady states monotonically decreases from nine to three as η declines, because the distance between the partial steady state locus and the diagonal, $|D^j(s^{i*}) - s^{i*}|$, is monotonically decreasing as η declines (Lemma 4) and there is a unique inflection point $(s^{i*} = 1 - G(\hat{a}))$ in the partial steady state loci ($D^2(s^{1*})$ and $D^1(s^{2*})$) (Lemma 3). The above claims imply that the number of steady states decreases from three to zero as η declines, in Regions 1 and 3, and there is always a unique steady state in Regions 2 and 4. ■

9.7 Proof of Lemma 5

First, check the local stability at one steady state Q_{hh}. We have the following Jacobian matrix at the steady state (s_h, s_h, Π_h, Π_h):

$$\mathbf{J_{Q_{hh}}} = \begin{bmatrix} \alpha[-G'A'_\sigma(\eta+(1-\eta)\beta^1)-1] & \alpha[-G'A'_\sigma(1-\eta)\beta^2] & \alpha[-G'A'_\Pi] & 0 \\ \alpha[-G'A'_\sigma(1-\eta)\beta^1] & \alpha[-G'A'_\sigma(\eta+(1-\eta)\beta^2)-1] & 0 & \alpha[-G'A'_\Pi] \\ -f'_\sigma(\eta+(1-\eta)\beta^1) & -f'_\sigma(1-\eta)\beta^2 & \rho+\alpha & 0 \\ -f'_\sigma(1-\eta)\beta^1 & -f'_\sigma(\eta+(1-\eta)\beta^2) & 0 & \rho+\alpha \end{bmatrix}_{Q_{hh}}.$$

Let us denote $\mathbf{J_{Q_{hh}}} - \lambda\mathbf{I}$ using 2×2 matrices J_{ij}s: $\mathbf{J_{Q_{hh}}} - \lambda\mathbf{I} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}$. We need to calculate the determinant of $\mathbf{J_{Q_{hh}}} - \lambda\mathbf{I}$ in order to find eigenvalues. Note that $|\mathbf{J_{Q_{hh}}} - \lambda\mathbf{I}| \equiv |J_{22}|\cdot|J_{11} - J_{12}J_{22}^{-1}J_{21}|$. Let us denote the second term by J': $J' \equiv J_{11} - J_{12}J_{22}^{-1}J_{21}$. Using the explicit forms of J_{ij}s, J' is

$$\begin{aligned} J' &= J_{11} - \begin{bmatrix} \alpha[-G'A'_\Pi] & 0 \\ 0 & \alpha[-G'A'_\Pi] \end{bmatrix} \cdot \begin{bmatrix} (\rho+\alpha-\lambda)^{-1} & 0 \\ 0 & (\rho+\alpha-\lambda)^{-1} \end{bmatrix} \\ &\quad \cdot \begin{bmatrix} -f'_\sigma(\eta+(1-\eta)\beta^1) & -f'_\sigma(1-\eta)\beta^2 \\ -f'_\sigma(1-\eta)\beta^1 & -f'_\sigma(\eta+(1-\eta)\beta^2) \end{bmatrix}. \end{aligned} \tag{28}$$

After a bit messy calculation, we find its determinant

$$\begin{aligned} |J'| &= \left| J_{11} - \alpha\xi \begin{bmatrix} \eta+(1-\eta)\beta^1 & (1-\eta)\beta^2 \\ (1-\eta)\beta^1 & \eta+(1-\eta)\beta^2 \end{bmatrix} \right|, \text{ where } \xi = \frac{G'A'_\Pi f'_\sigma}{\rho+\alpha-\lambda}, \\ &= [\lambda - \alpha(-G'A'_\sigma\eta - 1) + \alpha\xi\eta]\cdot[\lambda - \alpha(-G'A'_\sigma - 1) + \alpha\xi]. \end{aligned} \tag{29}$$

Therefore, the determinant of $\mathbf{J_{Q_{hh}}} - \lambda\mathbf{I}$ is

$$\begin{aligned} |\mathbf{J}_{Q_{hh}} - \lambda\mathbf{I}| &= |J_{22}|\cdot[\lambda - \alpha(-G'A'_\sigma\eta - 1) + \alpha\xi\eta]\cdot[\lambda - \alpha(-G'A'_\sigma - 1) + \alpha\xi] \\ &= [\lambda^2 - \lambda(-\alpha G'A'_\sigma + \rho) - \alpha(\alpha+\rho)(G'A'_\sigma + 1) - \alpha G'A'_\Pi f'_\sigma]_{Q_{hh}} \\ &\quad \cdot[\lambda^2 - \lambda(-\alpha G'A'_\sigma\eta + \rho) - \alpha(\alpha+\rho)(G'A'_\sigma\eta + 1) - \alpha G'A'_\Pi f'_\sigma\eta]_{Q_{hh}}. \end{aligned} \tag{30}$$

Taking $|\mathbf{J}_{Q_{hh}} - \lambda\mathbf{I}| = 0$, we obtain four eigenvalues at the steady state. First, note that $[-\alpha(\alpha+\rho)(G'A'_\sigma + 1) - \alpha G'A'_\Pi f'_\sigma]_{Q_{hh}} = \frac{\alpha(\alpha+\rho)g(\tilde{a}')}{\psi}\left[p + q' - \frac{\psi}{g(\tilde{a}')}\right]\Big|_{\tilde{a}'=A(s_h,\Pi_h)} < 0$, as shown in the proof of Lemma 2. Thus, the first term of the determinant has one positive and one negative eigenvalue. That is, the local stability condition at E_h in the case with one social group implies one negative and one positive eigenvalue at Q_{hh} in this two-group economy. Also, in the second term, we have $[-\alpha(\alpha+\rho)(G'A'_\sigma\eta + 1) - \alpha G'A'_\Pi f'_\sigma\eta]_{Q_{hh}} < 0$ because $-\alpha(\alpha+\rho)(G'A'_\sigma\eta+1) - \alpha G'A'_\Pi f'_\sigma\eta = \eta(-\alpha(\alpha+\rho)(G'A'_\sigma+1) - \alpha G'A'_\Pi f'_\sigma) - \alpha(\alpha+\rho)(1-\eta)$. Therefore, there are two positive and two negative eigenvalues.

There exists a unique equilibrium path if the number of jumping variables equals the number of eigenvalues with a positive real part (Buiter, 1984). Since we have two jumping variables, Π_t^1 and Π_t^2, and two positive eigenvalues, we know the existence of the unique equilibrium path in the neighborhood of (s_h, s_h). Therefore, Q_{hh} is an economically stable state. The four steady states Q_{ll}, Q_{hh}, Q_{lh} and Q_{hl} are identical in terms of their local dynamic structures. Without loss of generality, we can infer that two positive and two negative eigenvalues are at those states, and they are economically stable states, if they exist.

All other steady states, Q_{lm}, Q_{mh}, Q_{ml}, Q_{hm} and Q_{mm}, are economically unstable steady states. For example, check the local stability of Q_{mm}. Using equation (30), we have the determinant $\mathbf{J}_{Q_{mm}} - \lambda \mathbf{I}$:

$$\begin{aligned} |\mathbf{J}_{Q_{mm}} - \lambda \mathbf{I}| &= [\lambda^2 - \lambda(-\alpha G' A'_\sigma + \rho) - \alpha(\alpha+\rho)(G' A'_\sigma + 1) - \alpha G' A'_\Pi f'_\sigma]_{Q_{mm}} \\ &\quad \cdot [\lambda^2 - \lambda(-\alpha G' A'_\sigma \eta + \rho) - \alpha(\alpha+\rho)(G' A'_\sigma \eta + 1) - \alpha G' A'_\Pi f'_\sigma \eta]_{Q_{mm}}. \end{aligned} \tag{31}$$

We know that $[-\alpha(\alpha+\rho)(G'A'_\sigma+1)-\alpha G'A'_\Pi f'_\sigma]_{Q_{mm}} = \frac{\alpha(\alpha+\rho)g(\tilde{a}')}{\psi}\left[p+q'-\frac{\psi}{g(\tilde{a}')}\right]\Big|_{\tilde{a}'=A(s_m,\Pi_m)} > 0$, as shown in the proof of Lemma 2. Also we know that $-\alpha G' A'_\sigma + \rho > 0$ because of $A'_\sigma = \frac{-p}{\psi} < 0$. Thus, the first term of the determinant implies two eigenvalues with positive real parts. The second term implies at least one eigenvalue with positive real part because $-\alpha G' A'_\sigma \eta + \rho > 0$. Therefore, at least three eigenvalues have positive real parts. Since we have only two jumping variables, we cannot always find a unique equilibrium path in the neighborhood of (s_m, s_m) (Buiter, 1984). Thus, Q_{mm} is an economically unstable state. Now check other states. Since all other four are identical in terms of their dynamic structures, we need to check only one of them. Consider Q_{mh}. When $\eta = 1$, there must be three eigenvalues with positive real parts and one negative eigenvalue, because group 1 is at an economically unstable state E_m and group 2 is at an economically stable state E_h in the case with one social group. Thus, Q_{mh} with $\eta = 1$ is an economically unstable state since the number of positive eigenvalues, which is three, exceeds the number of jumping variables: except for initial points that s_0^1 exactly equals s_m, there does not exist a converging path to the state in the neighborhood of (s_m, s_h). We cannot explicitly calculate the signs of eigenvalues with $\eta < 1$. However, the qualitative approach helps us to conclude that it cannot be an economically stable state, because we can easily find at least one point (s^1, s^2) in the neighborhood of Q_{mh}, in which a converging equilibrium path to Q_{mh} does not exist. ■

9.8 Proof of Theorem 3

First consider the case with the absence of working period network externalities $q = 0$, in which convergence ranges are not overlapped at all, as illustrated in Figure 6. With $q = 0$, the dynamic structure can ignore two jumping variables Π_t^1 and Π_t^2, because they are fixed as a constant $(\bar{\delta}')$. Each convergence range is a basin of attraction for an attractor (an economically stable state). The basins are separated by separatrices that are connecting saddle points (economically unstable states). As partial steady states loci, $D^1(s^{2*})$ and $D^2(s^{1*})$ curves, get closer to the diagonal with the declined η (Lemma 4), the basins for economically stable asymmetric states $(Q_{lh},\ Q_{hl})$ should be shrinking, while those for economically stable symmetric states $(Q_{ll},\ Q_{hh})$ should be expanding, as manifested in Figure 6. This analysis for the special case with $q = 0$ is directly applied to the general case with $q > 0$, because the only difference is the bigger convergence ranges with the greater working period network externalities (greater q). The second argument in the theorem is true because the overlap of equilibrium paths in one social group case is analogous to the overlaps of the convergence ranges in the given two group case. When $q = 0$, there exists no overlap of convergence ranges (Figure 6). With greater q, overlaps tend to get larger, as Theorem 2 implies. ■

9.9 Proof of Lemma 6

As integration proceeds, either Q_{lh} and Q_{mh} are merged together or Q_{lh} and Q_{lm} are merged together before Q_{lh} disappears. First, envision a threshold segregation level for a sufficiently small $\beta^{1\prime}$: $\hat{\eta}(\beta^{1\prime})$. With the threshold level, the $D^2(s^{1*})$ curve will be tangent to the $D^1(s^{2*})$ curve and Q_{lh} will be merged with Q_{mh}, as Panel C of Figure 7 illustrates roughly. Now, let us increase $\beta^{1\prime}$ to $\beta^{1\prime} + \epsilon$ holding $\eta = \hat{\eta}(\beta^{1\prime})$. With this increase, $D^2(s^{1*})$ moves away from a diagonal because of the increased β^1 and $D^1(s^{2*})$ curve moves closer to the the diagonal because of the decreased β^2, according to Lemma 4. Thus, two steady states, Q_{lh} and Q_{mh}, get more distant from each other. In order to merge them and to make $D^2(s^{1*})$ curve tangent to the $D^1(s^{2*})$ curve again, the lower segregation level is required. Therefore, $\hat{\eta}(\beta^{1\prime}) > \hat{\eta}(\beta^{1\prime} + \epsilon)$, which implies $\hat{\eta}(\beta^1)$ is a strictly deceasing function with respect to β^1 in the lower range of β^1. In the same way, we can prove that $\hat{\eta}(\beta^{1\prime\prime}) > \hat{\eta}(\beta^{1\prime\prime} - \epsilon)$ for a sufficiently large $\beta^{1\prime\prime}$, which implies that $\hat{\eta}(\beta^1)$ is a strictly increasing function with respect to β^1 in the higher range of β^1.

Finally, imagine a group 1 population size of $\hat{\beta}$, with which all three steady states, Q_{lh}, Q_{mh} and Q_{lm}, are merged together at some level of segregation: $\hat{\eta}(\hat{\beta})$. With an increase of β^1 to $\hat{\beta} + \epsilon$, the $D^2(s^{1*})$ curve moves away from a diagonal and the $D^1(s^{2*})$ curve moves close to the the diagonal, which means only one steady state Q_{mh} survives and the other two disappear. This implies the threshold level of segregation should be higher with $\hat{\beta} + \epsilon$: $\hat{\eta}(\hat{\beta}) < \hat{\eta}(\hat{\beta} + \epsilon)$. With a decrease of β^1 to $\hat{\beta} - \epsilon$, the $D^1(s^{2*})$ curve moves away from a diagonal and the $D^2(s^{1*})$ curve moves closer to the the diagonal, which means only one steady state Q_{lm} survives and the other two disappear. This implies the threshold level of segregation should be higher with $\hat{\beta} - \epsilon$: $\hat{\eta}(\hat{\beta}) < \hat{\eta}(\hat{\beta} - \epsilon)$. Therefore, $\hat{\eta}(\hat{\beta})$ is a local minima. ■

Reference

Acemoglu, Daron: A Microfoundation for Social Increasing Returns in Human Capital Accumulation, *The Quarterly Journal of Economics* 111: 779-804 (1996)

Adsera, A and D. Ray: History and Coordination Failure, *Journal of Economic Growth* 3: 267-276 (1998)

Akerlof, George: Social Distance and Social Division, *Econometrica* 65: 1005-1027 (1997)

Barro, Robert: The Economic Growth in a Cross Section of Countries, *The Quarterly Journal of Economics* 106: 363-94 (1991)

Becker, G. and N. Tomes: An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility, *Journal of Political Economy* 87: 1153-89 (1979)

Bisin, A. and T. Verdier: The Economics of Cultural Transmission and the Dynamics of Preferences, *Journal of Economic Theory* 97: 298-319 (2001)

Blau, D. M. and P. K. Robins: Job Search Outcomes for the Employed and Unemployed, *Journal of Political Economy* 98: 637655 (1990)

Borjas, George: Ethnicity, Neighborhoods, and Human Capital Externality, *American Economic Review*, 85: 365-390 (1995)

Bowles, S., G. Loury and R. Sethi: Group Inequality, unpublished manuscript, Barnard College, Columbia University (2007)

Buiter, Willem: Saddlepoint Problems in Continuous Time Rational Expectations Models: A General Method and Some Macroeconomic Examples, *Econometrica*, 52: 665-680 (1984)

Castilla, Emilio: Social Networks and Employee Performance in a Call Center, *American Journal of Sociology* 110: 1243-83 (2005)

Calvó-Armengol, A. and M. Jackson: The Effects of Social Networks on Employment and Inequality, *American Economic Review*, 94: 426-454 (2004)

——— : Networks in Labor Markets: Wage and Employment Dynamics and Inequality, *Journal of Economic Theory* 132: 27-46 (2007)

Chaudhuri, S. and R. Sethi: Statistical Discrimination with Peer Effects: Can Integration Eliminate Negative Stereotypes? *Review of Economic Studies* 78: 579 - 596 (2008)

Coate, S. and G. Loury: Will Affirmative-Action Policies Eliminate Negative Stereotype? *American Eco-*

nomic Review 83: 1220-1240 (1993)

Currarini, Sergio, Mattew Jackson and Paolo Pin: An Economic Model of Friendship: Homophily, Minorities and Segregation, *Econometrica* 77: 100345 (2009)

Cutler, David and Edward Glaeser: Are Ghettos Good or Bad? *Quarterly Journal of Economics* 112: 827-872 (1997)

Durlauf, Steven, and Marcel Fafchamps: Social Capital. In P. Aghion and S. Durlauf(eds.) *Handbook of Economic Growth* . Amsterdam: North Holland (2006)

Fafchamps, M. and B. Minten: Relationship and Traders in Madagascar, *Journal of Development Studies* 35: 1-35 (1999)

Falk, Armin and Andrea Ichin: Clean Evidence of Peer Effects, *Journal of Labor Economics* 25 (2005)

Fryer, Roland and Glenn C. Loury: Valuing Identity: The Simple Economics of Affirmative Action Policies, unpublished manuscript (2007)

Granovetter, Mark S.: *Getting A Job: A Study of Contracts and Careers*, Chicago: University of Chicago Press; 2nd Edition 1995 (1974)

Hauk, Esther and Maria Saez-Mart: On the Cultural Transmission of Corruption, *Journal of Economic Theory* 107(2): 311-335 (2002)

Hanushek, Eric, John Kain, and Steve Rivkin: New Evidence about Brown vs Board of Education: The Complex Effects of School Racial Composition on Achievement, *Journal of Labor Economics* 27: 349-383 (2009)

Holzer, Harry J.: Search Method Use by Unemployed Youth, *Journal of Labor Economics* 1-20 (1988)

Hoxby, Caroline: The Power of Peers: How Does Make-up of a Classroom Change Achievements, *Education Next* 2.2 (2002)

Ioannides, Yannis and Linda Loury: Job Information Networks, Neighborhood Effects, and Inequality, *Journal of Economic Literature* 42: 1056-93 (2004)

Jackson, Matthew: *Social and Economic Networks*, Princeton and Oxford: Princeton University Press (2008)

Kain, John F.: Housing Segregation, Negro Employment and Metropolitan Decentralization, *Quarterly Journal of Economics* 82: 175-197 (1968)

Khwaja, Asim, Atif Mian and Abid Qamar: The Value of Joining Giant Networks, WCFIA Working

Paper, Harvard University (2009)

Krugman, Paul: History Versus Expectations, *Quarterly Journal of Economics* 106: 651-667 (1991)

Lucas, Robert: On the Mechanics of Economic Development, *Journal of Monetary Economics* 22: 3-12 (1988)

Lundberg, Shelly and Richard Startz: On the Persistence of Racial Inequality, *Journal of Labor Economics* 16: 292-322 (1998)

Montgomery, James: Social Networks and Labor Market Outcomes: Toward an Economic Analysis, *American Economic Review* 815:1408-18 (1991)

——— : Weak Ties, Employment and Inequality: An Equilibrium Analysis, *American Journal of Sociology* 99: 1212-36 (1994)

Munshi, Kaivan: Networks in the Modern Economy: Mexican Migrants in the US Labor Market, *Quarterly Journal of Economics* 118: 549599 (2003)

Murphy, K., A. Shleifer and R. Vishny: Industrialization and the Big Push. *Journal of Political Economy* 97: 1003-26 (1989)

Ozgen, Eren and Robert A. Baron: Social sources of information in opportunity recognition: Effects of mentors, industry networks, and professional forums, *Journal of Business Venturing* 22: 174-192 (2007)

Rees, Albert and George P. Shultz: *Workers in an Urban Labor Market,* Chicago: University of Chicago Press (1970)

Rockoff, Jonah E.: Does Mentoring Reduce Turnover and Improve Skills of New Employees? Evidence from Teachers in New York City. NBER Working Papers No. 13868 (2008)

Rosenstein-Rodan, P.: Problems of Industrialization in Eastern and South-Eastern Europe, *Economic Journal* 53: 202-11 (1943)

Weinberg, Bruce, Patricia Reagan and Jeffrey Yankow: Do Neighborhoods Affect Hours Worked? *Journal of Labor Economics* 22: 891-921 (2004)

Figure 1. Steady States with Unique Ability Level

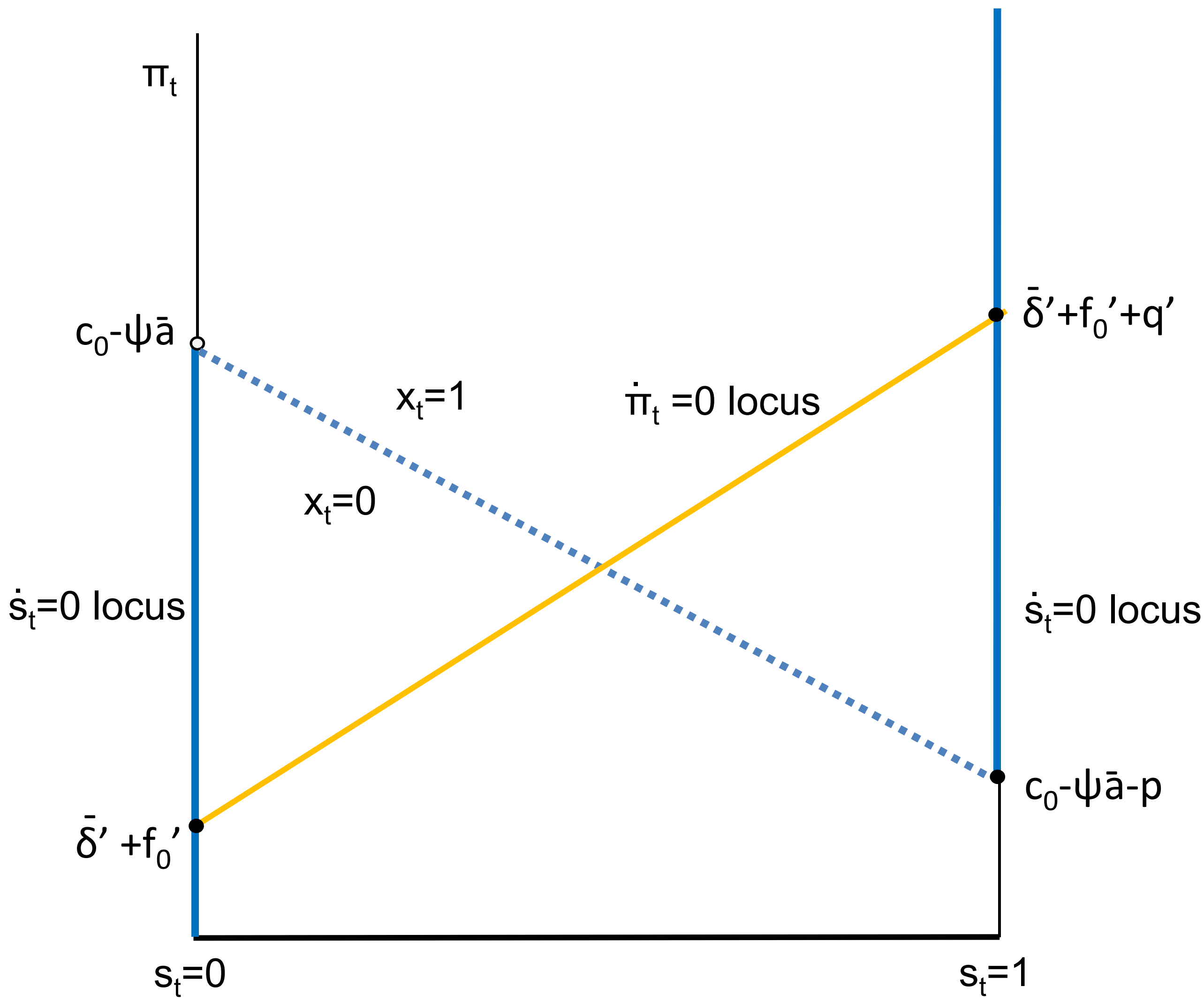
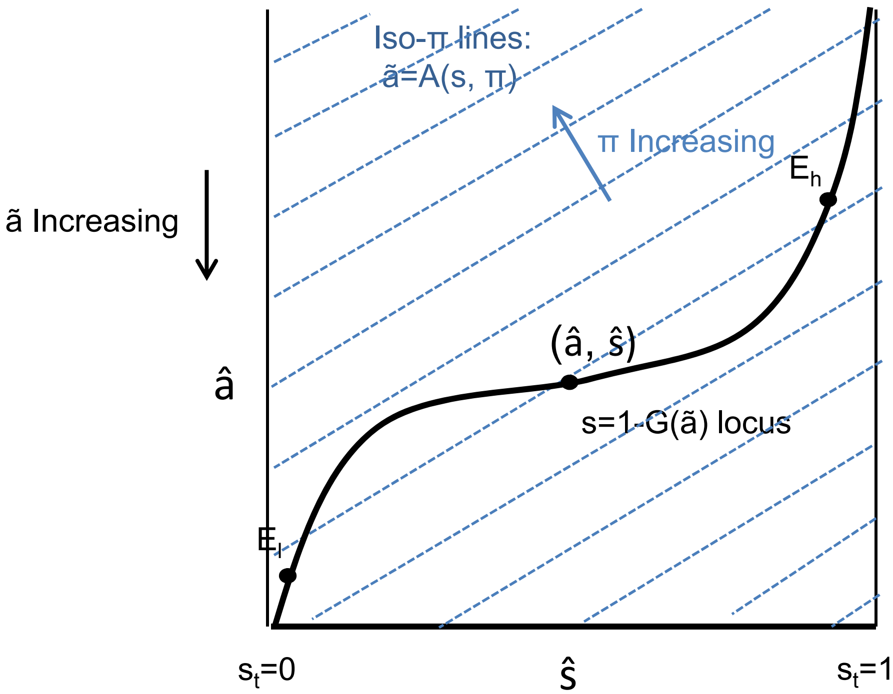


Figure 2. Steady States with Ability Distribution G(a)

Panel A Finding the $\dot{s}_t$=0 Locus



Panel B Steady States with the $\dot{s}_t$=0 and $\dot{\pi}_t$=0 Loci

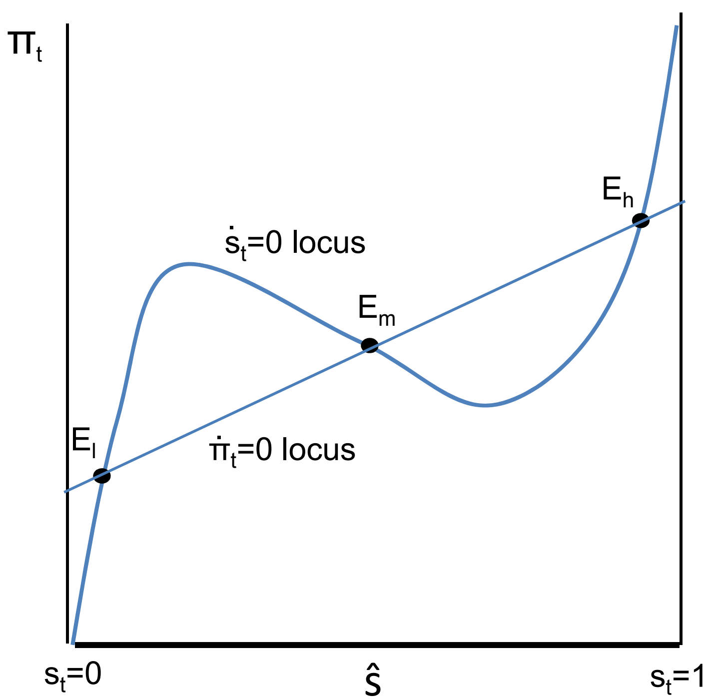


Figure 3. Equilibrium Paths with Unique Ability Level

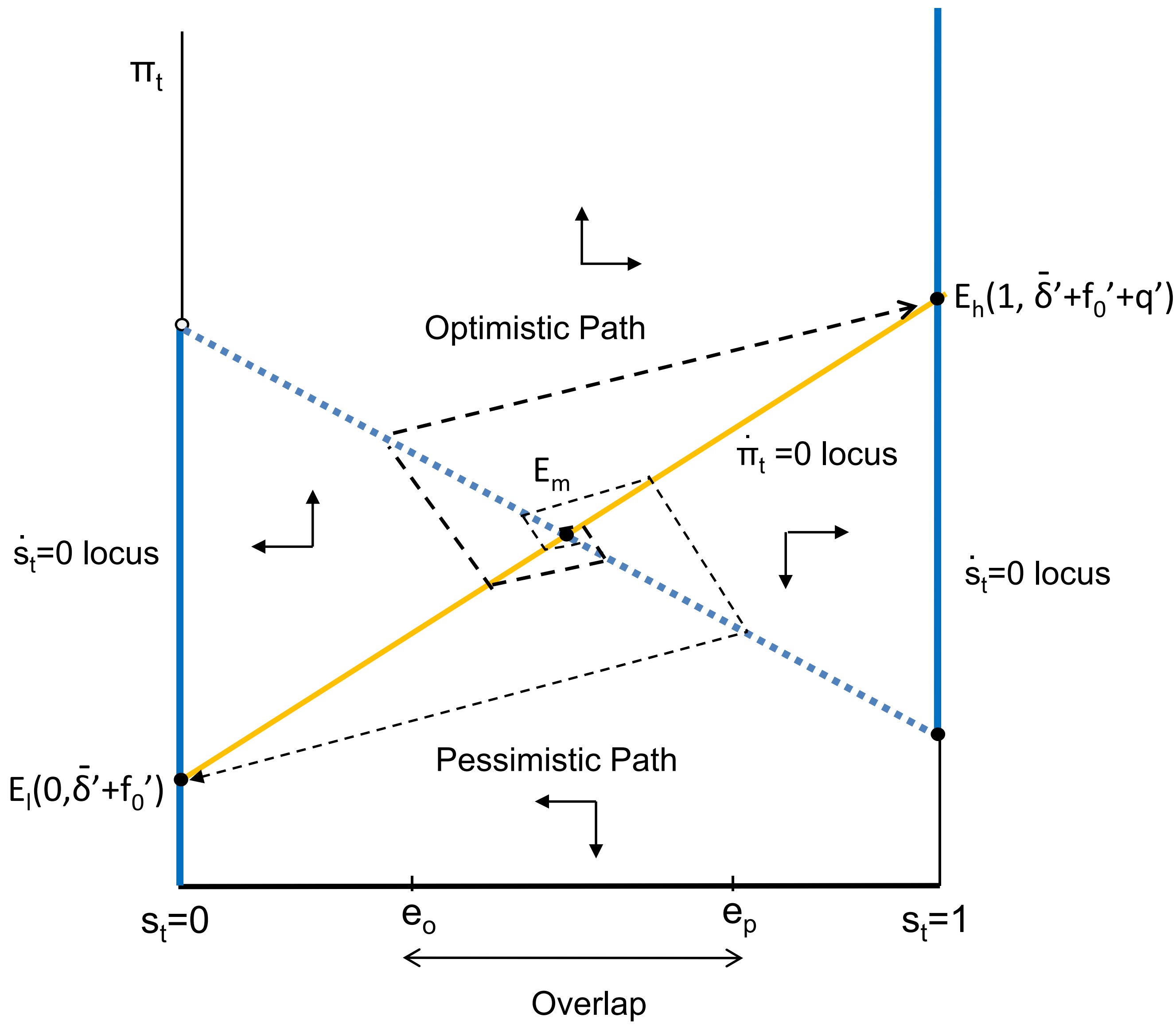


Figure 4. Equilibrium Paths with Ability Distribution G(a)

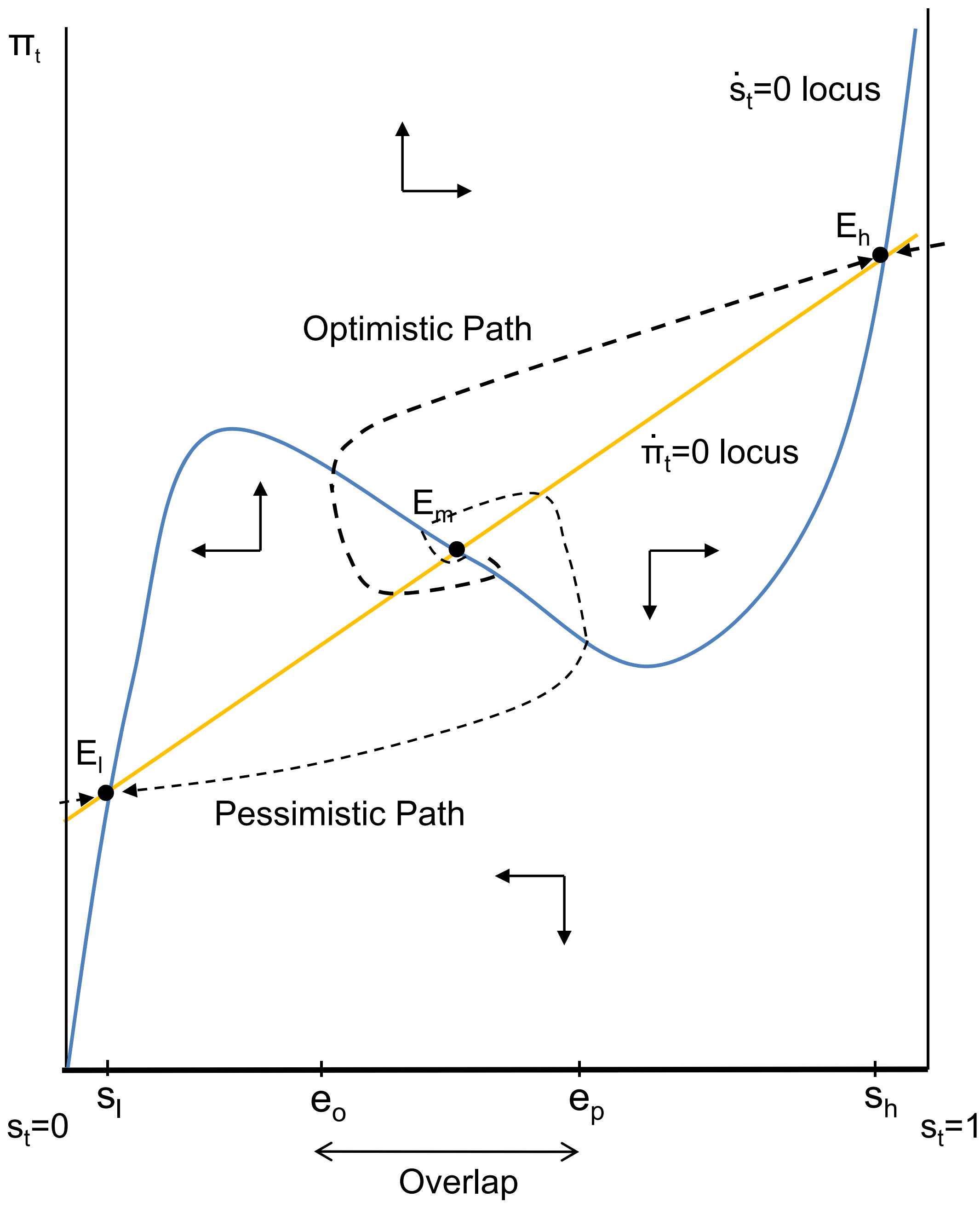


Figure 5. Different Levels of Working-period Network Externalities

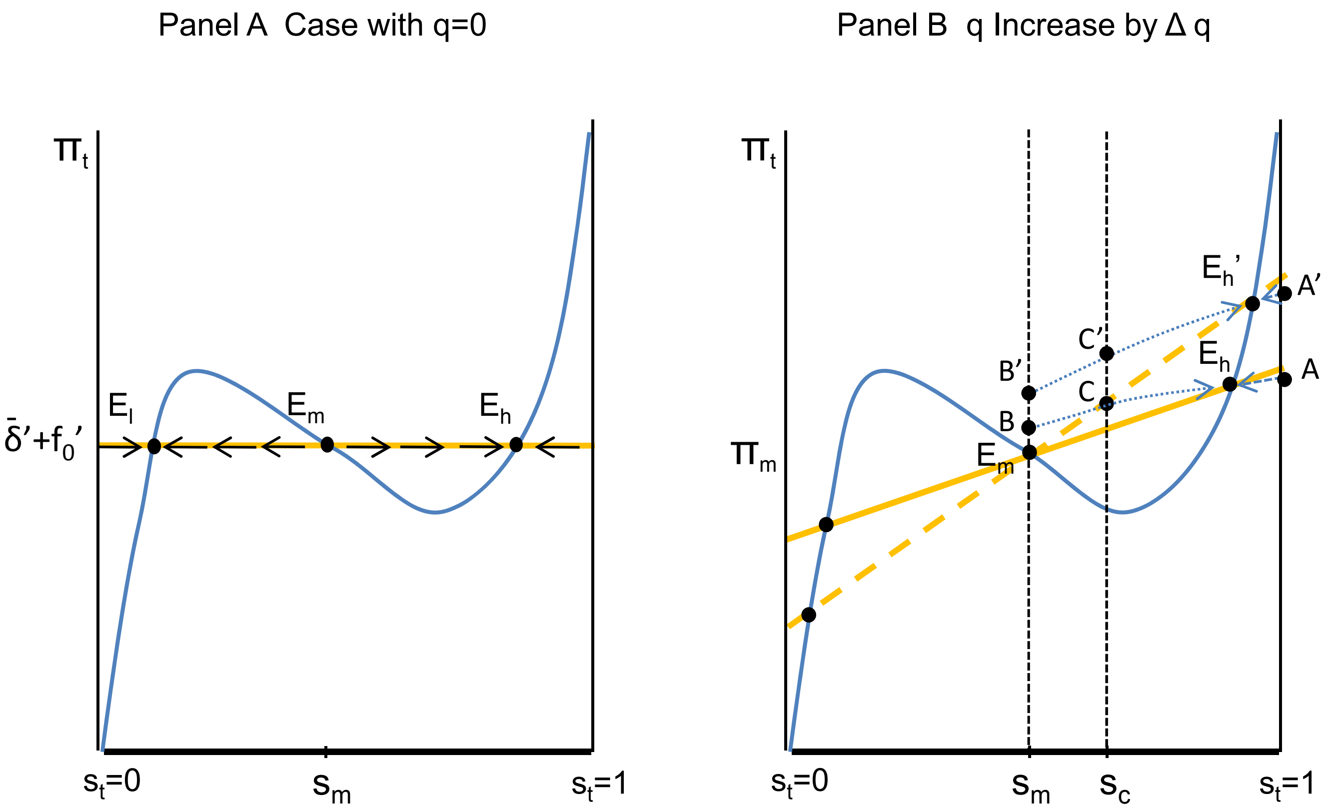


Figure 6. Steady States for Each Segregation Level η (given $\beta^1 < \beta^2$)

- η declines in Panel A, B, C, D, E, F order, with η=1 in Panel A and η=0 in Panel F.

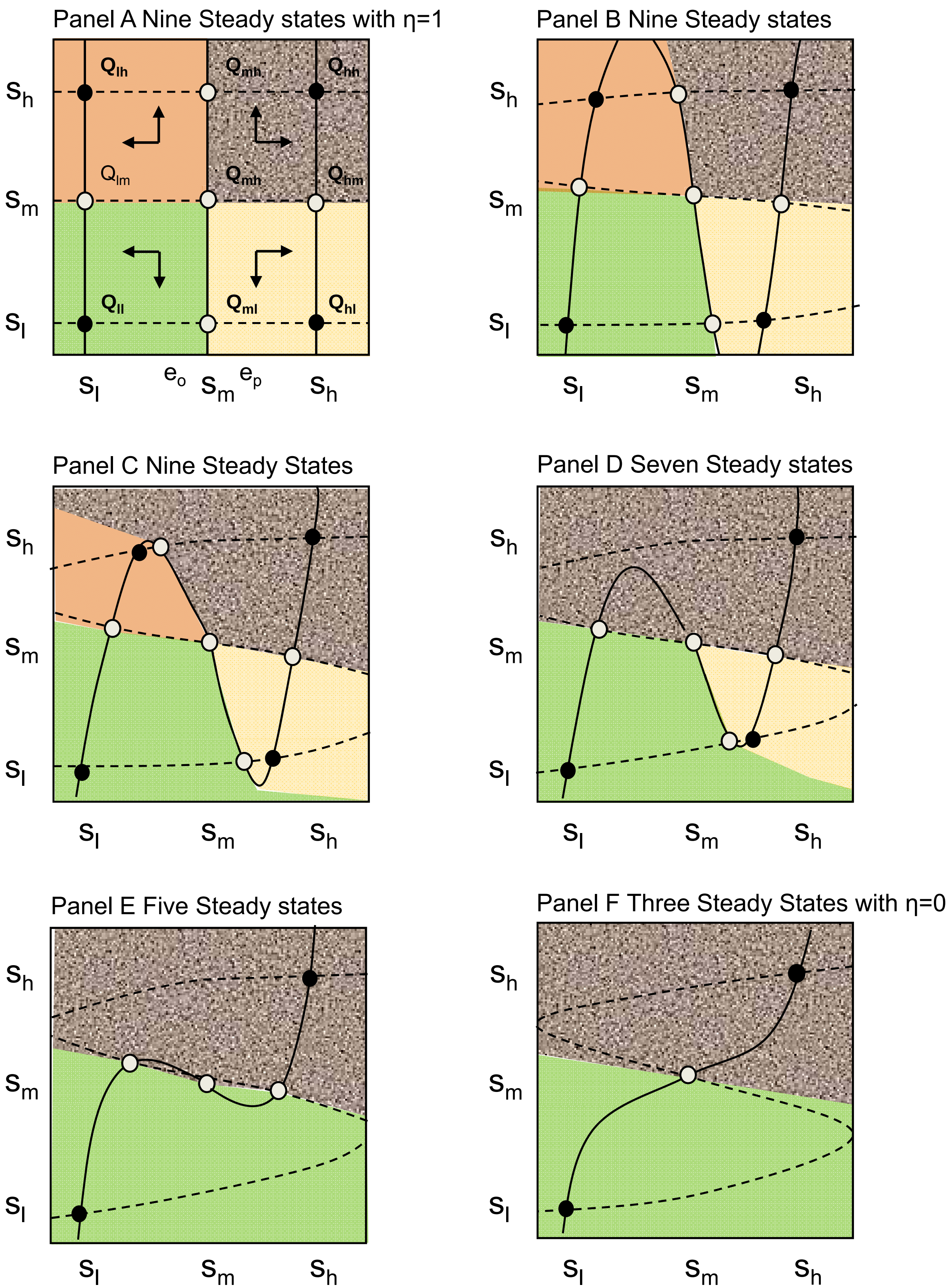


Figure 7. Convergence Ranges for Each Segregation Level η (given $\beta^1 < \beta^2$)

- η declines in Panel A, B, C, D, E, F order, with η=1 in Panel A and η=0 in Panel F.

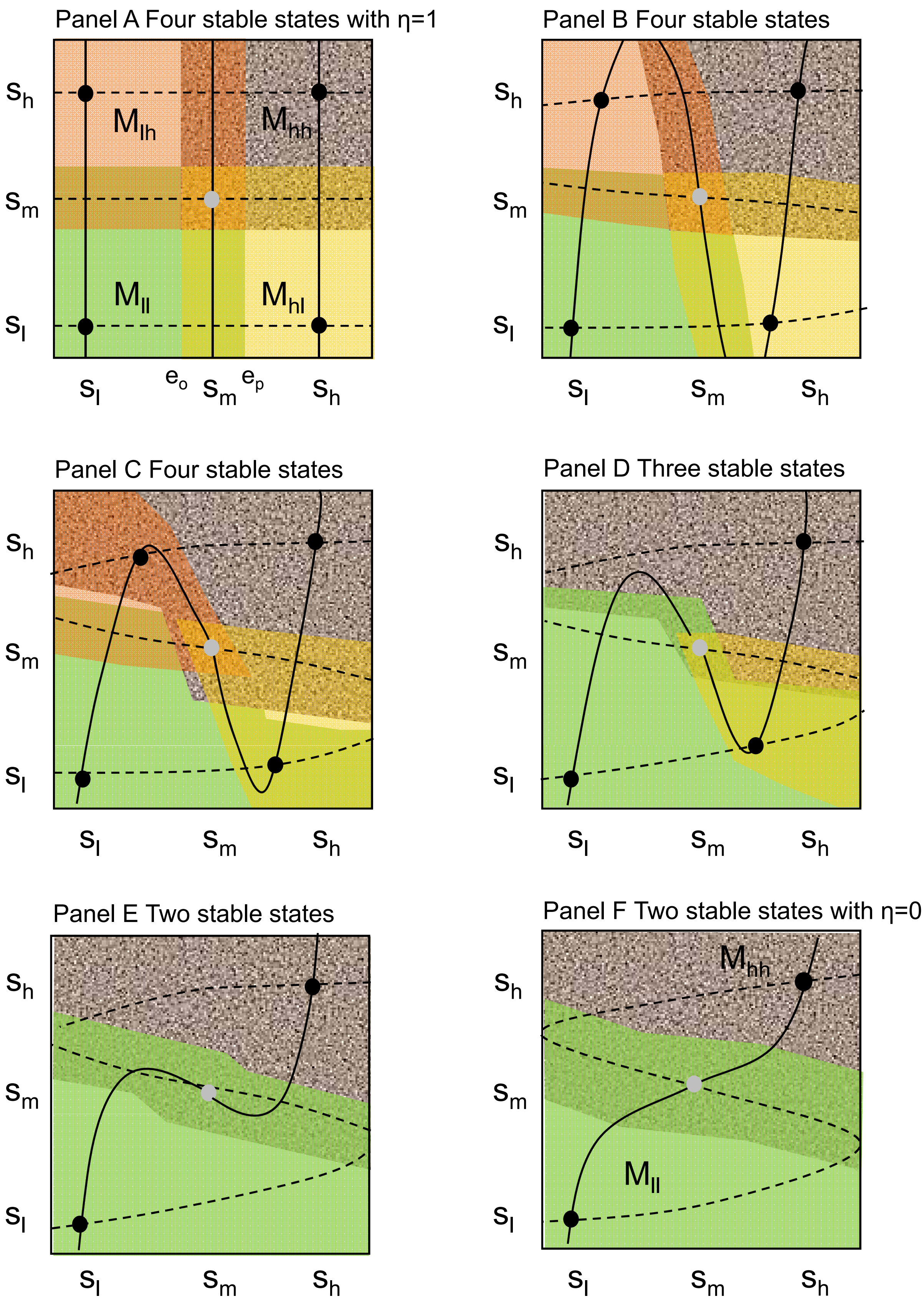


Figure 8. Integration Policy (with η Declining from One)

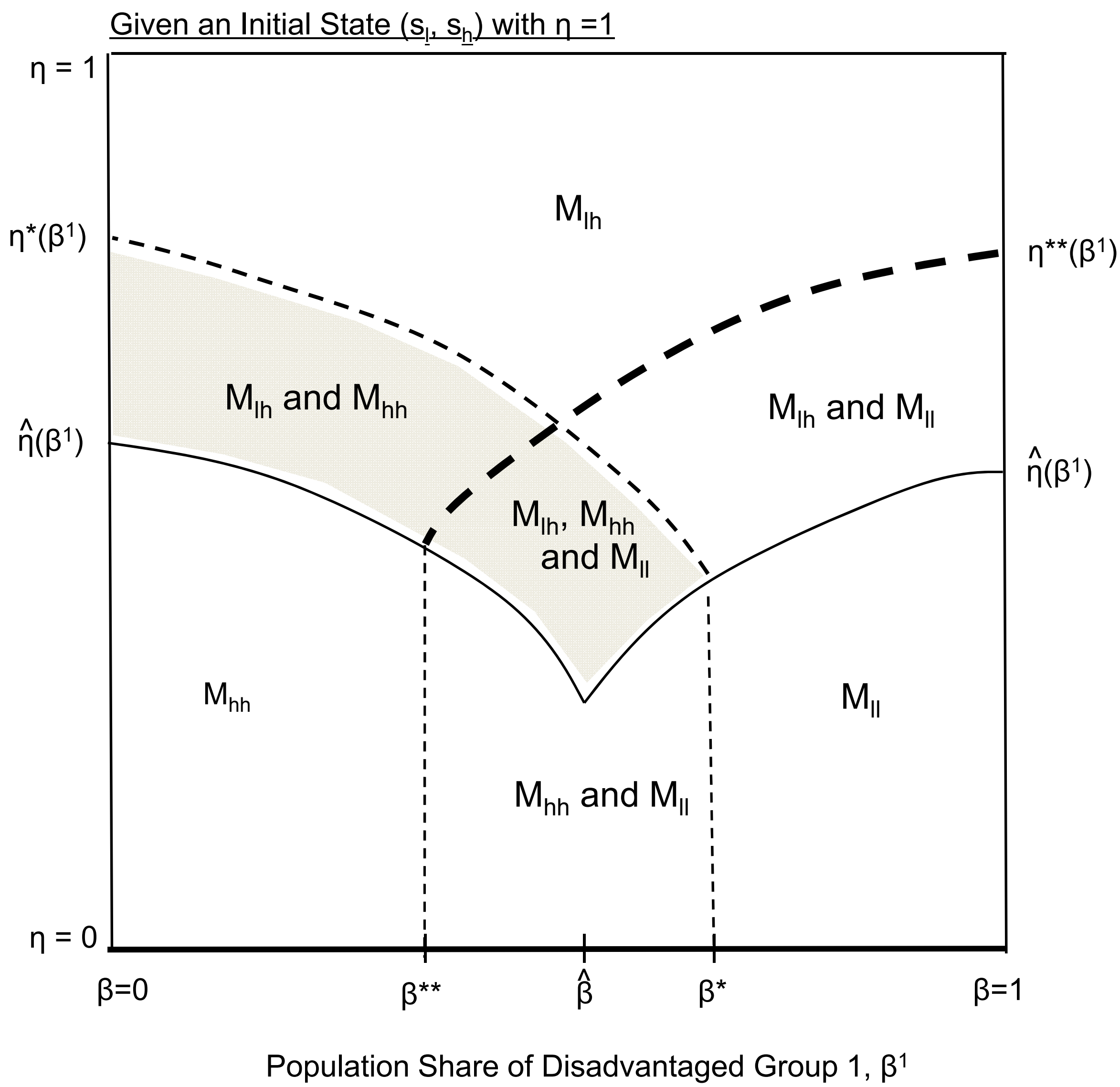
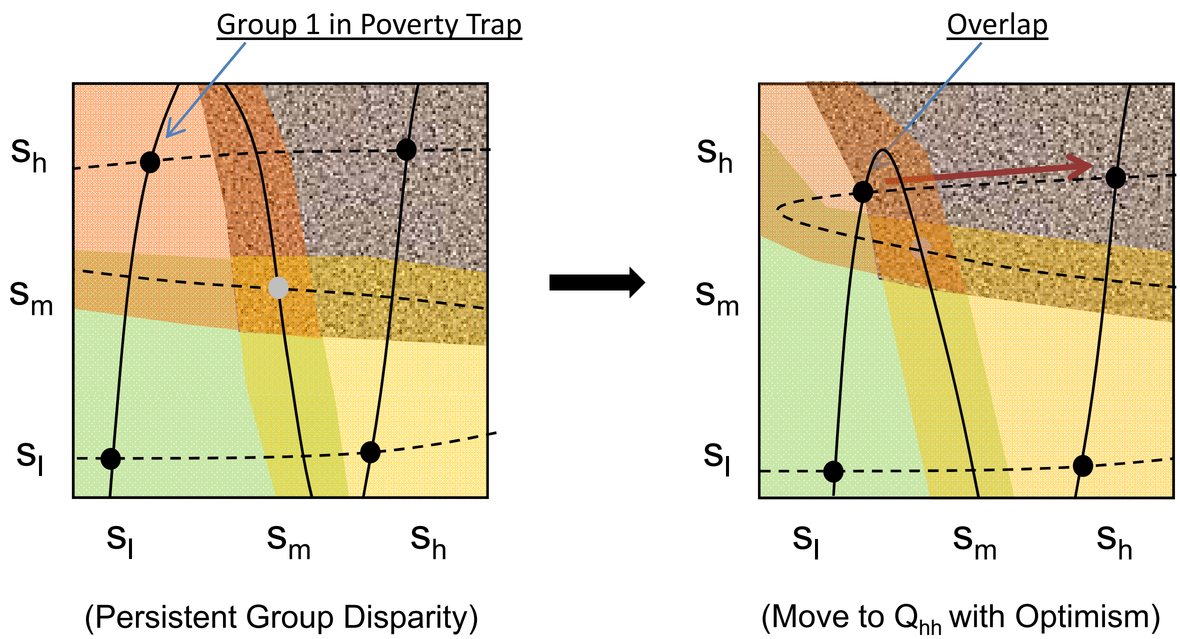
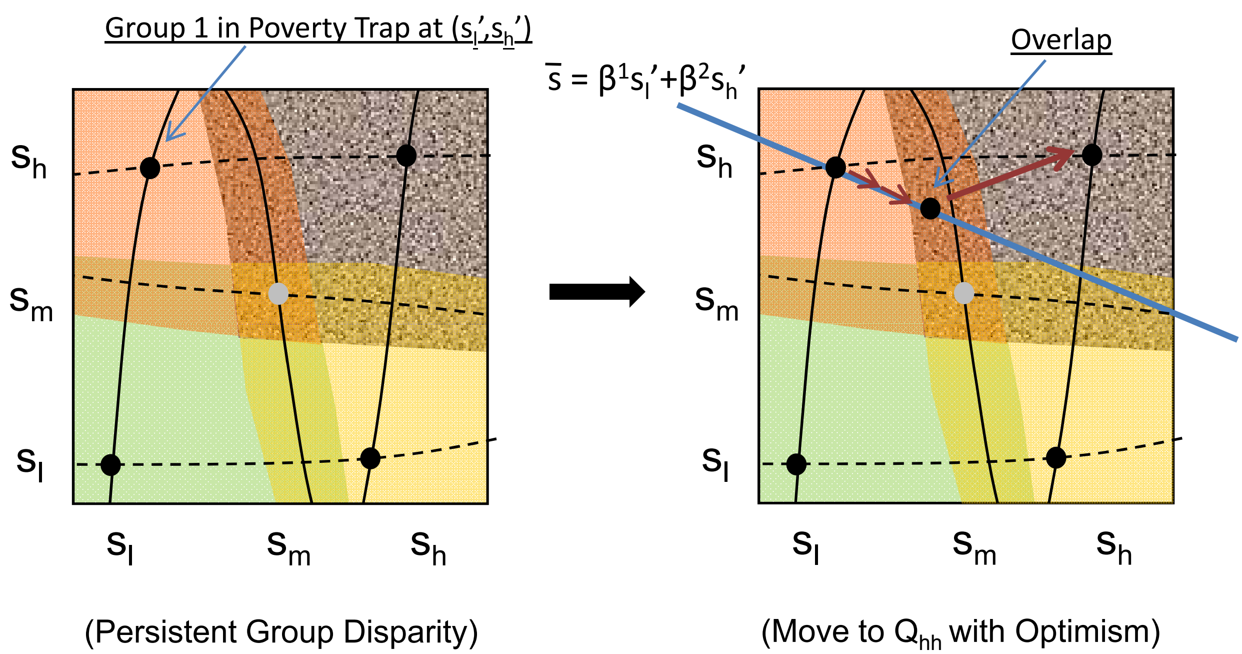


Figure 9. Affirmative Action Policies

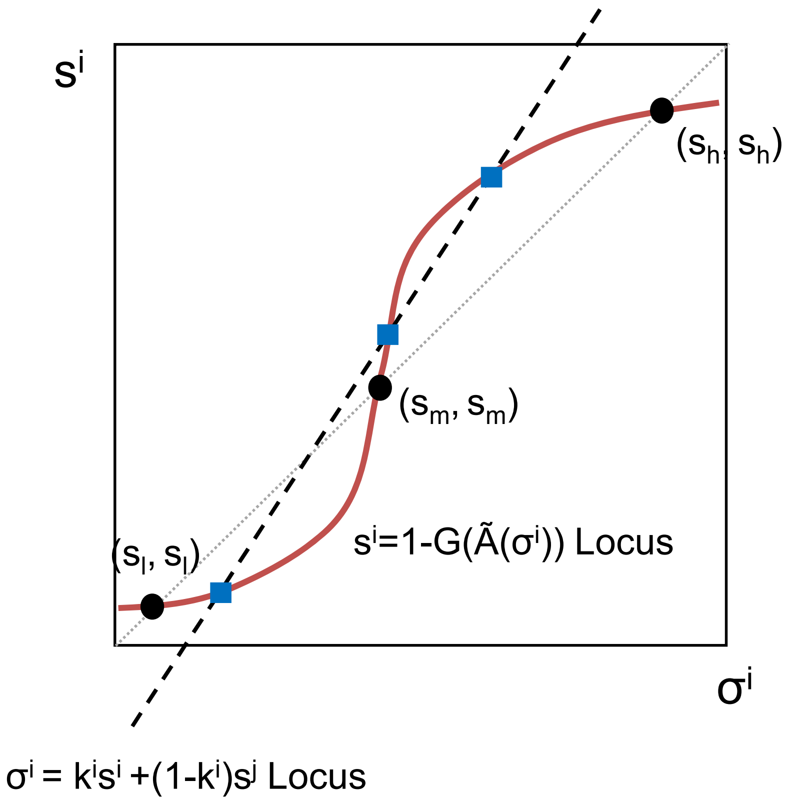
Panel A Asymmetric Training Subsidy

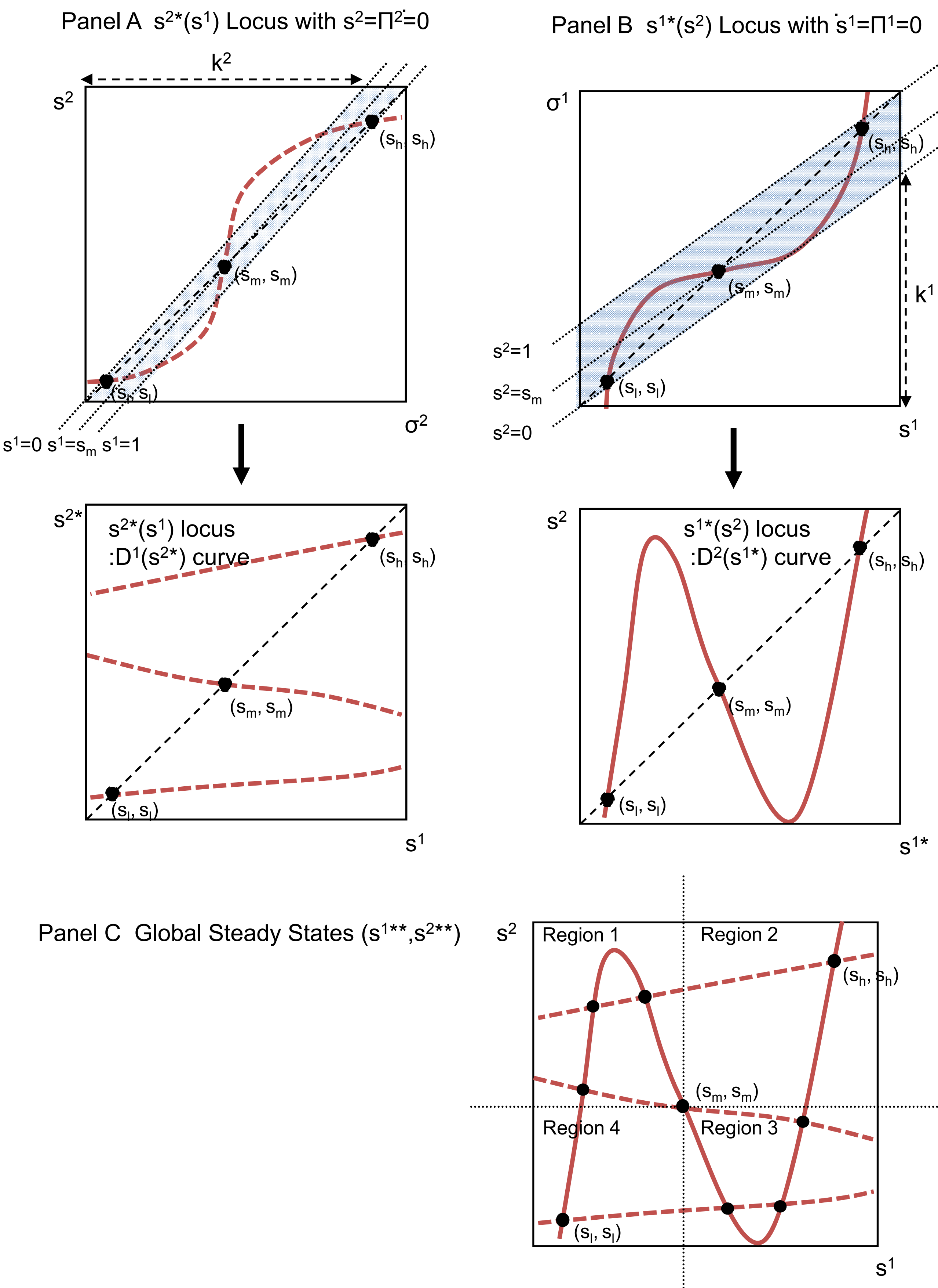


Panel B Quota System



Appendix Figure 1. Partial Steady States (s^{i*},σ^{i*}) Given s^j





Appendix Figure 2. Global Steady States (given $\beta^1 < \beta^2$)