On the Magnet Effect of Foreign Direct Investment

Pao-Li Chang, Chia-Hui Lu

September 2006

On the Magnet Effect of Foreign Direct Investment

Pao-Li Chang[*][†]   Chia-Hui Lu [‡]

Singapore Management University   Academia Sinica

September 29, 2006

Abstract

We extend Antràs and Helpman (2004) on firm heterogeneity and organizational choice to a dynamic setting with FDI uncertainty, in which the probability of investment failure decreases with the host country's infrastructure level and increases with the technological complexity facing each firm. Moreover, it decreases over time as the accumulated mass of firms succeeding in FDI increases. We show that a minimum level of infrastructure is required to trigger a first wave of industrial migration. We then formalize the often noted "magnet effect" of FDI–the first wave of industrial migration generates positive externality (information spillover) for subsequent investors, which stimulates a second wave of industrial migration. The process continues until the power of the "magnet" reaches its steady-state level. In contrast with the predictions in Antràs and Helpman (2004), we show that firms with intermediate productivity levels are the ones migrate first, while the most productive and the least productive firms tend to stay behind. This non-monotonic relationship between firms' productivity and their FDI propensities is consistent with the patterns of Taiwanese firms undertaking FDI in China.

---

[*]Tel: +65-6828-0830; fax: +65-6828-0833. *E-mail address:* plchang@smu.edu.sg (P.-L. Chang).
[†]This project was supported by the SMU research grant.
[‡]Tel: +886-2-2782-2791 ext 503; fax: +886-2-2785-3946 *E-mail address:* clu@econ.sinica.edu.tw (C.-H. Lu).

1 Introduction

We extend the seminal paper of Antràs and Helpman (2004) on firm productivity and organizational choice to the setting with investment uncertainty, whereby international vertical integration is associated with higher risk of failure than domestic vertical integration, with the risk being higher for firms producing more technologically sophisticated products. This modification alters the basic prediction of Antràs and Helpman (2004) that more productive firms in a given industry are more likely to undertake foreign direct investment (FDI). Instead, in this paper, we show that firms with intermediate productivity levels are the first ones to migrate, while the most productive and the least productive firms tend to stay behind. Thus, there is a non-monotonic relationship between firm productivity and FDI propensity.

Next, we show that the migration of Northern firms to the South is not automatic. We pin down the minimum level of infrastructure that the South must possess in order to attract the first wave of industry migration from the North. Migrating firms' identities as reflected by their productivity levels are also traced out. We then formalize the often noted *"magnet effect"* of FDI that international industry migration often occurs in clusters and in orders. As will be shown in the paper, when more firms have succeeded in FDI and produced in the South, the general investment uncertainty decreases, which lowers the FDI risk of failure. Thus, the fist wave of industry migration generates positive externality for subsequent investors and stimulates a second wave of industry migration, which then triggers a third wave of migration. The process of industry migration continues until the power of the "magnet" reaches its steady-state level and an equilibrium structure of international production emerges.

We conduct simulations to illustrate the dynamics of this magnet effect and to derive the structure of international production at the steady state. A series of simulations are carried out to study the effects of several key parameters on the dynamics of the migration process. This includes the level of infrastructure in the South, the economic distance between the host and home country of the FDI, the relative wage advantage of the South, the relative management cost disadvantage of international versus domestic vertical integration, sectoral characteristics such as headquarter-service intensity and demand conditions, and relative bargaining powers between the headquarter and the supplier of intermediate inputs.

2 Empirical Observations

The non-monotonic relationship between firm productivity and FDI propensity is observed in a sample of Taiwanese firms over a period of 15 years since 1991 when the Taiwanese government lifted the ban on westward FDI in mainland China. The details are shown in Figure 1. In the figure, the timing of the first-time FDI undertaken by a Taiwanese firm, measured as the number of quarters elapsed since 1991, is indicated along the vertical axis, and the productivity of a firm along the horizontal axis. Firms in six major industries are sampled, with each panel indicating the pattern for each industry.

The firms in our sample are firms listed on either the Taiwan Stock Exchange Corporation or the Over the Counter market. We compile the information on the timing of FDI for each firm based on the Foreign Direct Investment Database published by Taiwan Economic Journal Data Bank (TEJ) and the Yearly Report published by the Investment Commissions, Ministry of Economic Affairs, Taiwan. The productivity of a firm is measured by the ratio of Return on Assets (ROA). We take the average of the ratios for a firm between 1991 and 1993, as an indication of a firm's initial productivity level. The ROA data are taken from the TEJ Financial Statement Database.

As shown in the figure, across all industries, the firms with intermediate productivity levels tend to undertake FDI earlier than the most productive and the least productive firms. The phenomenon is most notable in the Electronics industry, where firms are spread over a wider range of productivity levels. This finding contradicts the conventional perception that higher-productivity firms have a higher tendency to undertake FDI than lower-productivity firms, as captured by the theoretical model of Antràs and Helpman (2004) in a static setting. The panel for the Electronics industry also suggests that the range of intermediate firms who undertake FDI expands over time, resembling a 'cone'. This is an interesting dynamics that a static model could not readily explain. These observations motivate us to develop a theoretical model that accounts for the non-monotonic relationship between firm productivity and FDI propensity, and a dynamic one that predicts the 'cone' dynamics of FDI.

3 The Model

The world has a unit measure of population with the following preference structure:

$$U = x_0 + \frac{1}{\alpha_j} \sum_{j=1}^{J} X_j^{\alpha_j}, \quad 0 < \alpha_j < 1$$

$$X_j = \left[\int x_j(i)^{\alpha_j} di \right]^{1/\alpha_j},$$

where x_0 is the numeraire, homogeneous, goods produced both in the North and in the South. The labor is the only factor of production and the production technology of x_0 exhibits constant returns to scale; thus the labor productivity of a country in producing the numeraire goods determines its wage rate. The North is assumed to have an absolute advantage in producing the numeraire goods, and therefore commands a higher wage: $w^N > w^S$. X_j is the aggregate consumption of differentiated products $x_j(i)$ in sector j. The elasticity of substitution between any two varieties within a sector ($\sigma_j = \frac{1}{1-\alpha_j}$) is allowed to differ across sectors.

The derived inverse demand function for each variety i in sector j is:

$$p_j(i) = x_j(i)^{\alpha_j - 1} \tag{1}$$

Let $h_j(i)$ denote the headquarter service input and $m_j(i)$ the intermediate input used in the production of variety i in sector j. It is assumed that only the North has the ability to produce the headquarter service, while the intermediate input can be produced either in the North or the South. The production technology of variety i in sector j is:

$$x_j^N(i) = \theta \left[\frac{h_j(i)}{\eta_j} \right]^{\eta_j} \left[\frac{m_j(i)}{1 - \eta_j} \right]^{1-\eta_j} \tag{2}$$

when the intermediate input is produced in the North. The parameter η_j denotes the intensity of the headquarter service component in the production of the final goods in sector j. The higher η_j is, the more important is the headquarter service component. The parameter θ indicates the productivity level of the headquarter producing the variety. We will also use this parameter to indicate the level of technology sophistication employed to produce the product. Thus, it is assumed that a more

4

productive headquarter produces a variety that embodies more sophisticated technology.

On the other hand, if the intermediate input is produced in the South, the headquarter faces a potential risk that the intermediate input produced abroad may fail to match the exact specification designed by the North. In this case, the output is zero. In brief, the production technology in this case is:

$$x_j^S(i) = \begin{cases} \theta \left[\frac{h_j(i)}{\eta_j}\right]^{\eta_j} \left[\frac{m_j(i)}{1-\eta_j}\right]^{1-\eta_j}, & \text{in case of success;} \\ 0, & \text{in case of failure.} \end{cases}$$

The headquarter chooses between acquiring the ownership of a manufacturing plant in the North (domestic vertical integration) or one in the South (international vertical integration). It is assumed that the organizational fixed cost to coordinate the headquarter and the supplier of intermediate components is higher in the case of international vertical integration than domestic vertical integration:

$$f^S > f^N. \tag{3}$$

In the following, we focus on one sector and suppress the sector index to simplify the notation. By (1) and (2), the revenue function for variety i given the amounts of the inputs is:

$$R^l(i) = \gamma^l \theta^\alpha \left[\frac{h(i)}{\eta}\right]^{\alpha\eta} \left[\frac{m(i)}{1-\eta}\right]^{\alpha(1-\eta)}, \quad l = \{N, S\} \tag{4}$$

where $\gamma^N = 1$, and $0 < \gamma^S < 1$ is the probability of successful matching of the headquarter service and the intermediate input produced in the South.

As the contract between the headquarter and the supplier of intermediate inputs is non-enforceable, the two parties negotiate the division of the sales revenue ex post. The division depends on the relative negotiating powers of the two parties and their outside options. Suppose the negotiation fails. Given the headquarter's ownership in the manufacturing plant, it can potentially seize the intermediate component at the cost of a fraction $(1-\delta)$ of the output. Also suppose that the headquarter is able to extract at the ex post negotiation, a fraction β of the surplus from the contract with the supplier. Then, the division of the sales revenue for the headquarter is:

$$\beta^N = \beta^S = (\delta)^\alpha + \beta[1 - (\delta)^\alpha]. \tag{5}$$

In the first stage, the headquarter and the supplier of the intermediate input determine respectively how much headquarter service and intermediate input to produce, taking into account the fraction of the sales revenue they will extract from the negotiation in the second stage. The production of either input is assumed to have a unit labor requirement of one. With profit maximization by both parties, the joint profit of the two parties in either of the two potential organizational forms is:

$$\pi^l(\theta, \eta) = \gamma^{l \frac{1}{1-\alpha}} \theta^{\frac{\alpha}{1-\alpha}} \psi^l(\eta) - w^N f^l, \quad l = \{N, S\}, \tag{6}$$

where

$$\psi^l(\eta) = \frac{1 - \alpha[\beta^l \eta + (1 - \beta^l)(1 - \eta)]}{\{(1/\alpha)(w^N/\beta^l)^\eta [w^l/(1-\beta^l)]^{1-\eta}\}^{\alpha/(1-\alpha)}}. \tag{7}$$

It can be shown that the headquarter will choose the organizational form that maximizes the above joint profit. Given that $w^S < w^N$, it follows that $\psi^S(\eta) > \psi^N(\eta)$. Thus, headquarters choosing to engage FDI in the South enjoy a larger variable profit margin, but at the same time, incur a higher fixed organizational cost. In addition, the risk of mismatch involved in FDI further reduces the variable profit gain of producing in the South when compared to the North. We assume that the probability of successful matching takes the following specific form:

$$\gamma^S(\theta) = \left(\frac{b}{\theta}\right)^{\alpha z}, \quad \theta \geq b > 0, \tag{8}$$

$$z = \frac{1}{K + \underline{X}^S}, \quad K > 1, \ \underline{X}^S \geq 0, \tag{9}$$

where the parameter K represents the level of infrastructure in the South and $\underline{X}^S$ the degree of FDI penetration in the South. The level of infrastructure encompasses various aspects of a nation's capacity in absorbing FDI. This includes physical infrastructure, social capital, human capital, and governance infrastructure. Equation (8) and (9) imply that the higher the level of infrastructure and the larger the presence of FDI in the South, the higher the probability that a FDI undertaking will succeed for a given θ. Observe that the FDI success probability is strictly decreasing in product sophistication, θ, with the probability equal to one for the lowest-tech product and approaching zero as the product sophistication increases toward infinity. Thus, $d\gamma^S/d\theta < 0$, $\gamma^S(b) = 1$, and $\lim_{\theta \to \infty} \gamma^S(\theta) = 0$.

To simplify notation in the following exposition, define $\tilde{\theta} = \theta^{\frac{\alpha}{1-\alpha}}$, and $\tilde{b} = b^{\frac{\alpha}{1-\alpha}}$. It follows that

$$\pi^S(\tilde{\theta}, z) = \psi^S \tilde{b}^z \tilde{\theta}^{1-z} - w^N f^S, \tag{10}$$

$$\pi^N(\tilde{\theta}) = \psi^N \tilde{\theta} - w^N f^N. \tag{11}$$

As $0 < z < 1$, Equation (10) implies that the profit function of undertaking FDI is increasing and concave in $\tilde{\theta}$, which is in contrast with the linear profit function of producing in the North. Note that the above setup includes Antràs and Helpman (2004) as a special case when z approaches zero (and the probability of successful matching approaches one for all θ). In this case, there is no uncertainty involved in FDI and the profit function of FDI becomes a linear function in $\tilde{\theta}$: $\psi^S \tilde{\theta} - w^N f^S$.

We adopt the following assumption to avoid a taxonomy of cases.

Assumption 1 The parameters satisfy the following conditions: (i) $\tilde{b} < \frac{w^N f^N}{\psi^N}$; (ii) $\frac{f^N}{\psi^N} < \frac{f^S}{\psi^S}$.

Define $\tilde{\theta}_N \equiv \frac{w^N f^N}{\psi^N}$ and $\tilde{\theta}' \equiv \frac{w^N f^S}{\psi^S}$. Headquarters (hereafter firms) with the productivity level $\tilde{\theta}_N$ break even when producing in the North; similarly, firms with the productivity level $\tilde{\theta}'$ break even when producing in the South, under the scenario of no uncertainty. By Assumption 1(ii), $\tilde{\theta}_N < \tilde{\theta}'$, and there exists $\tilde{\theta}_S \equiv \frac{w^N (f^S - f^N)}{\psi^S - \psi^N} > \tilde{\theta}'$ such that firms with the productivity level $\tilde{\theta}_S$ are indifferent between producing in the North and producing in the South, under the scenario of no uncertainty. Together, Assumptions 1(i) and (ii) imply that the least productive firm $\tilde{b}$ is below the lower threshold to produce in the North, $\tilde{\theta}_N$, which is further below the lower threshold to produce in the South under no uncertainty, $\tilde{\theta}_S$. Thus, firms are partitioned according to their productivity levels into the least productive ones who do not produce, the less productive ones who produce in the North, and the most productive ones who produce in the South. There is no complete specialization by the South in the production of intermediate components even in the best scenario, for the South, of no FDI uncertainty. This is the benchmark scenario taken in Antràs and Helpman (2004) . We develop our analysis below conditional on this setup.

We start by characterizing the curve $\pi^S(\tilde{\theta}, z)$, which is illustrated in Figure 2. In the figure, $\tilde{\theta}$ is indicated on the horizontal axis and π^S (or π^N) on the vertical axis. Note that $\pi^S(\tilde{b}, z) = \psi^S \tilde{b} - w^N f^S$. Thus, regardless of z, the curve $\pi^S(\tilde{\theta}, z)$ always passes through the fixed point

$(\tilde{b}, \psi^S \tilde{b} - w^N f^S)$, denoted Point A in the figure. In the limiting case where there is no FDI uncertainty $(z \to 0)$, $\pi^S(\tilde{\theta}, z)$ approaches a linear function with a vertical intercept of $-w^N f^S$ and a slope of ψ^S. For $0 < z < 1$, $\pi^S(\tilde{\theta}, z)$ is increasing and concave in $\tilde{\theta}$. Note that $\pi^S(\tilde{\theta}, z)$ in (10) can be rewritten as $\pi^S(\tilde{\theta}, z) = \psi^S \tilde{\theta} (\frac{\tilde{b}}{\tilde{\theta}})^z - w^N f^S$. Thus, $\pi^S(\tilde{\theta}, z)$ is decreasing in z for $\tilde{\theta} > \tilde{b}$ and increasing in z for $\tilde{\theta} < \tilde{b}$. As z increases, the curve $\pi^S(\tilde{\theta}, z)$ shifts down for $\tilde{\theta} > \tilde{b}$ and shifts up for $\tilde{\theta} < \tilde{b}$. As $z \to 1$, the curve $\pi^S(\tilde{\theta}, z)$ approaches a step function with $\pi^S = -w^N f^S$ for $\tilde{\theta} = 0$ and $\pi^S = \psi^S \tilde{b} - w^N f^S$ for all $\tilde{\theta} > 0$. In contrast, $\pi^N(\tilde{\theta})$ is linear in $\tilde{\theta}$, with a vertical intercept of $-w^N f^N$ and a slope of ψ^N. By Assumption 1, it follows that $\pi^S(\tilde{b}, z) < \pi^N(\tilde{b})$.

Proposition 1 *Under Assumption 1, there exists a unique $z^* \in (0, 1)$ such that the curve π^S is tangent to π^N, and*

(i) for all $z \in (0, z^*)$, $\exists \ \tilde{\theta}_S < \tilde{\theta}_0 < \tilde{\theta}_1$, such that $\pi^S(\tilde{\theta}, z) > \pi^N(\tilde{\theta}) > 0$ for all $\tilde{\theta} \in (\tilde{\theta}_0, \tilde{\theta}_1)$, and $\pi^S(\tilde{\theta}, z) < \pi^N(\tilde{\theta})$ for all $\tilde{\theta} \in [\tilde{b}, \tilde{\theta}_0) \cup (\tilde{\theta}_1, \infty)$;

(ii) for all $z \in (z^*, 1)$, $\pi^S(\tilde{\theta}, z) < \pi^N(\tilde{\theta})$ for all $\tilde{\theta} \in [\tilde{b}, \infty)$.

Proof of Proposition 1. Define $\tilde{\theta}^+$ such that $\frac{d\pi^S}{d\tilde{\theta}} = \frac{d\pi^N}{d\tilde{\theta}}$; that is, $\tilde{\theta}^+$ is the productivity level where the two curves π^S and π^N have the same slope. It is straightforward to verify that

$$\tilde{\theta}^+ = \tilde{b}[(1-z)\frac{\psi^S}{\psi^N}]^{1/z}.$$

Thus, $\tilde{\theta}^+ \geq \tilde{b}$ if and only if $z \leq 1 - \frac{\psi^N}{\psi^S} \equiv \bar{z}$.

(*i*) For $0 < z < \bar{z}$, we have the difference between the two curves at $\tilde{\theta}^+$ as

$$\pi^S(\tilde{\theta}^+, z) - \pi^N(\tilde{\theta}^+) = \frac{z}{1-z}\tilde{\theta}^+\psi^N - w^N(f^S - f^N).$$

Define $\phi(z) \equiv \frac{z}{1-z}\tilde{\theta}^+$ and $g(z) \equiv \frac{1-z}{z}$. It is straightforward to show that

$$\frac{d\tilde{\theta}^+}{dz} = \frac{d \ln\tilde{\theta}^+}{dz} \tilde{\theta}^+$$
$$= \left\{ -\frac{1}{z^2}\ln\left[(1-z)\frac{\psi^S}{\psi^N}\right] - \frac{1}{z(1-z)} \right\}\tilde{\theta}^+ < 0,$$

where the last inequality follows from the fact that $(1-z)\frac{\psi^S}{\psi^N} > 1$ for $0 < z < \bar{z}$. As $\lim_{z \to 0} \tilde{\theta}^+ \to \infty$

8

and $\lim_{z \to 0} g(z) \to \infty$, by L'Hôpital rule,

$$\lim_{z \to 0} \phi(z) = \lim_{z \to 0} \frac{d\tilde{\theta}^+/dz}{dg(z)/dz}$$
$$= \lim_{z \to 0} \left\{ ln \left[(1-z) \frac{\psi^S}{\psi^N} \right] + \frac{z}{1-z} \right\} \tilde{\theta}^+ \to \infty.$$

Thus,

$$\lim_{z \to 0} \pi^S(\tilde{\theta}^+, z) - \pi^N(\tilde{\theta}^+) \to \infty.$$

Furthermore, it is straightforward to show that $\lim_{z \to \bar{z}} \phi(z) = \frac{\bar{z}}{1-\bar{z}} \tilde{b} = \frac{\psi^S - \psi^N}{\psi^N} \tilde{b}$. Therefore,

$$\lim_{z \to \bar{z}} \pi^S(\tilde{\theta}^+, z) - \pi^N(\tilde{\theta}^+) = \tilde{b}(\psi^S - \psi^N) - w^N(f^S - f^N) < 0,$$

where the last inequality follows by Assumption 1. Finally, observe that for $0 < z < \bar{z}$,

$$\frac{d\phi(z)}{dz} = [\frac{d \, ln\tilde{\theta}^+}{dz} - \frac{d \, ln \, g(z)}{dz}]\phi(z)$$
$$= \left\{ -\frac{1}{z^2} ln \left[(1-z) \frac{\psi^S}{\psi^N} \right] - \frac{1}{z(1-z)} + \frac{1}{z(1-z)} \right\} \phi(z)$$
$$= \left\{ -\frac{1}{z^2} ln \left[(1-z) \frac{\psi^S}{\psi^N} \right] \right\} \phi(z) < 0.$$

Therefore,

$$d \left(\pi^S(\tilde{\theta}^+, z) - \pi^N(\tilde{\theta}^+) \right) /dz < 0.$$

In summary, as $z \to 0$, $\pi^S(\tilde{\theta}, z)$ becomes linear, and $\tilde{\theta}^+ \to \infty$. The difference between $\pi^S(\tilde{\theta}, z)$ and $\pi^N(\tilde{\theta})$ at $\tilde{\theta} = \tilde{\theta}^+$ approaches infinity. As z increases, $\tilde{\theta}^+$ decreases and the difference between $\pi^S(\tilde{\theta}^+, z)$ and $\pi^N(\tilde{\theta}^+)$ also decreases monotonically. At $z = \bar{z}$, $\tilde{\theta}^+ = \tilde{b}$ and the difference between $\pi^S(\tilde{\theta}^+, z)$ and $\pi^N(\tilde{\theta}^+)$ becomes negative. Therefore, there must exist a unique $z^* \in (0, \bar{z})$ such that $\pi^S(\tilde{\theta}^+, z^*) - \pi^N(\tilde{\theta}^+) = 0$. In other words, the curve $\pi^S(\tilde{\theta}, z)$ is tangent to $\pi^N(\tilde{\theta})$ at $z = z^*$.

For $z \in (0, z^*)$, $\pi^S(\tilde{\theta}^+, z) - \pi^N(\tilde{\theta}^+) > 0$. Thus, the curve $\pi^S(\tilde{\theta}, z)$ must have intersected the line $\pi^N(\tilde{\theta})$ twice. Label the corresponding productivity levels $\tilde{\theta}_0$ and $\tilde{\theta}_1$ with $\tilde{\theta}_0 < \tilde{\theta}_1$. Then it follows from the concavity of $\pi^S(\tilde{\theta}, z)$ that $\pi^S(\tilde{\theta}, z) - \pi^N(\tilde{\theta}) > 0$ for all $\tilde{\theta} \in (\tilde{\theta}_0, \tilde{\theta}_1)$. Next, note that for $z > 0$, the curve $\pi^S(\tilde{\theta}, z)$ falls below the linear line $\psi^S \tilde{\theta} - w^N f^S$, which intersects the line $\pi^N(\tilde{\theta})$ at $\tilde{\theta}^S$. Thus, it must be the case that $\tilde{\theta}_0 > \tilde{\theta}^S$ for $z > 0$. Moreover, because $\tilde{\theta}_0 > \tilde{\theta}_N$, it follows that

9

$\pi^N(\tilde{\theta}) > 0$ for $\tilde{\theta} \in (\tilde{\theta}_0, \tilde{\theta}_1)$. Therefore, $\pi^S(\tilde{\theta}, z) > \pi^N(\tilde{\theta}) > 0$ for all $\tilde{\theta} \in (\tilde{\theta}_0, \tilde{\theta}_1)$, where $\tilde{\theta}_S < \tilde{\theta}_0 < \tilde{\theta}_1$. It follows from the concavity of $\pi^S(\tilde{\theta}, z)$ that $\pi^S(\tilde{\theta}, z) < \pi^N(\tilde{\theta})$ for all $\tilde{\theta} \in [\tilde{b}, \tilde{\theta}_0) \cup (\tilde{\theta}_1, \infty)$.

(ii) For $z^* < z < \bar{z}$, the curve $\pi^S(\tilde{\theta}, z)$ falls completely below the line $\pi^N(\tilde{\theta})$ for all $\tilde{\theta} \in [\tilde{b}, \infty)$. Thus, $\pi^S(\tilde{\theta}, z) < \pi^N(\tilde{\theta})$ for all $\tilde{\theta} \in [\tilde{b}, \infty)$.

For $\bar{z} \leq z < 1$, $\tilde{\theta}^+ \leq \tilde{b}$, and because $\frac{d\pi^S}{d\tilde{\theta}}$ is decreasing in $\tilde{\theta}$, it follows that $\frac{d\pi^S}{d\tilde{\theta}} < \frac{d\pi^N}{d\tilde{\theta}}$, for all $\tilde{\theta} > \tilde{b}$. In addition, at $\tilde{\theta} = \tilde{b}$, $\pi^S(\tilde{b}, z) < \pi^N(\tilde{b})$ by Assumption 1. It follows that $\pi^S(\tilde{\theta}, z) < \pi^N(\tilde{\theta})$ for all $\tilde{\theta} \in [\tilde{b}, \infty)$. The desired result in (ii) therefore follows. ∎

Corollary 2 *For $z \in (0, z^*)$, firms are partitioned according to their productivity levels as follows: (i) for firms with $\tilde{\theta} \in [\tilde{b}, \tilde{\theta}_N]$, they exit the market; (ii) for firms with $\tilde{\theta} \in [\tilde{\theta}_N, \tilde{\theta}_0] \cup [\tilde{\theta}_1, \infty)$, they integrate the production in the North; (iii) for firms with $\tilde{\theta} \in [\tilde{\theta}_0, \tilde{\theta}_1]$, they undertake FDI.*

On the other hand, for $z \in (z^*, 1)$, FDI is not viable: (i) for firms with $\tilde{\theta} \in [\tilde{b}, \tilde{\theta}_N]$, they exit the market; (ii) for firms with $\tilde{\theta} \in [\tilde{\theta}_N, \infty)$, they integrate the production in the North.

Proposition 3 *For $z \in (0, z^*)$, the range of firms undertaking FDI, as denoted by $(\tilde{\theta}_1 - \tilde{\theta}_0)$, is monotonically decreasing in z:*
(i) as $z \to 0$, $\tilde{\theta}_1 - \tilde{\theta}_0 \to \infty$;
(ii) as $z \to z^*$, $\tilde{\theta}_1 - \tilde{\theta}_0 \to 0$

Proof of Proposition 3. Recall that $\pi^S(\tilde{\theta}, z)$ is decreasing in z for a given $\tilde{\theta} > \tilde{b}$, so the increasing concave curve $\pi^S(\tilde{\theta}, z)$ shifts down as z increases for $\tilde{\theta} > \tilde{b}$. It follows that with a larger $z < z^*$, the curve $\pi^S(\tilde{\theta}, z)$ will intersect the line $\pi^N(\tilde{\theta})$ at a higher $\tilde{\theta}_0$ and at a lower $\tilde{\theta}_1$. Thus, the range of firms undertaking FDI, $\tilde{\theta}_1 - \tilde{\theta}_0$, decreases, as z increases toward z^*. At $z = z^*$, $\pi^S(\tilde{\theta}, z)$ becomes tangent to $\pi^N(\tilde{\theta})$. It follows that $\tilde{\theta}_0 = \tilde{\theta}_1$. As $z \to 0$, $\pi^S(\tilde{\theta}, z) \to \psi^S \tilde{\theta} - w^N f^S$. It follows that $\tilde{\theta}_0 \to \tilde{\theta}_S$ and $\tilde{\theta}_1 \to \infty$. The desired result therefore follows. ∎

4   The Magnet Effect of FDI

In this section, we extend the static model introduced above to a dynamic setting with multiple time periods. In so doing, we formalize the stylized fact that earlier FDI often reveals information regarding the host country's investment environment, which helps lower the uncertainty faced by

later investors. The information spillover is assumed to be external to firms and to affect the whole economy. Specifically, we assume that

$$\gamma_t^S(\theta) = \left(\frac{b}{\theta}\right)^{\alpha z_t}, \ \theta \geq b > 0, \tag{12}$$

$$z_t = \frac{1}{K + X_{t-1}^S}, \quad K > 1, \ t = 1, 2, \ldots \tag{13}$$

where X_{t-1}^S denotes the degree of FDI penetration in the South in period $t-1$. Equations (12) and (13) imply that a higher degree of FDI penetration in period $t-1$ raises the success probability of FDI in period t for any given θ. Thus, earlier FDI creates a positive externality for subsequent FDI. In particular, the degree of FDI penetration in period t is defined as the effective mass of firms producing in the South in period t:

$$X_t^S = \frac{G(\theta_{1,t}) - G(\theta_{0,t})}{d}, \ d > 0, \tag{14}$$

where $G(\cdot)$ is the cumulative distribution function of firm productivity levels and is chosen to be a Pareto distribution with shape k, i.e., $G(\theta) = 1 - \left(\frac{b}{\theta}\right)^k$. In (14), the magnitude $G(\theta_{1,t}) - G(\theta_{0,t})$ represents the absolute mass of firms producing in the South in period t, which is scaled by the economic distance, d, between the host and home country of FDI. The economic distance, d, summarizes the barriers to information exchange arising from physical distance, and cultural, language, and institutional differences. Thus, for a given absolute mass of firms transplanted in the South, the effective mass and the extent of information spillover it creates is larger, the closer the two countries are in terms of their economic distance.

We now formulate the magnet effect of FDI. Begin with an initial period ($t = 0$) when no FDI is present in the South ($X_0^S = 0$). It is straightforward to identify the minimum level of infrastructure that the South must possess to trigger the first wave of FDI. By Proposition 1, it follows that:

Corollary 4 *The minimum level of infrastructure that the South must possess to trigger the first wave of FDI is $K^* = 1/z^*$.*

The importance of the recipient country's infrastructure in influencing FDI flows has been documented by various empirical contributions. See, for example, Wei (2000) for a study of corruption

and its depressing effect on inward FDI. In another study, Globerman and Shapiro (2002) estimated the minimum threshold of infrastructure that a recipient country must achieve to attract positive FDI flows from the U.S., where they measured the level of infrastructure by estimating a linear combination of proxies for governance infrastructure, human capital, and physical capital.

Let $\theta_{0,t}$ and $\theta_{1,t}$ indicate the lower and upper bound of productivity levels, of which firms undertake FDI in period t. If the South possesses the minimal level of infrastructure $(K > K^*)$, FDI takes place at $t = 1$, and by Corollary 2, there exists $\theta_{0,1} < \theta_{1,1}$ such that firms with $\theta \in (\theta_{0,1}, \theta_{1,1})$ undertake FDI. By (12)–(14), it follows that $\gamma_2^S > \gamma_1^S$, given that $X_1^S > X_0^S = 0$. Thus, the first wave of FDI helps raise the success probability of FDI at $t = 2$. This is illustrated in Figure 3, where at $t = 2$, the curve π^S tilts upward. Firms with productivity levels in the range of $(\theta_{0,1} - \epsilon, \theta_{0,1})$ or $(\theta_{1,1}, \theta_{1,1} + \epsilon)$, who find it not profitable to undertake FDI at $t = 1$, now prefer moving the production process to the South, since the risk associated with FDI is lower than before. As a result, the first wave of migration induces a second wave of migration, $(\theta_{0,1}, \theta_{1,1}) \subset (\theta_{0,2}, \theta_{1,2})$, and the effective mass of FDI firms in the South is enlarged, $X_2^S > X_1^S > X_0^S = 0$. The larger mass of FDI firms further reduce the FDI uncertainty and trigger a third wave of FDI. The process of migration will continue until the power of the "magnet" reaches its steady-state level. The time paths of $\{\gamma_t^S\}_{t=1}^{\infty}$, $\{X_t^S\}_{t=1}^{\infty}$, and $\{(\theta_{0,t}, \theta_{1,t})\}_{t=1}^{\infty}$ can be derived accordingly by iteration.

5   Simulation

In this section, we conduct simulations to illustrate the dynamics of industry migration introduced above. By varying the parameter values of the model, we will also obtain insights into the qualitative effects of important parameters on the speed and range of FDI over time. We choose the following parameter values for our benchmark case: $w^N = 4$, $w^S = 1$, $f^N = 1$, $f^S = 2$, $\alpha = 0.5$, $\delta = 0.64$, $\beta = 0.5$, $\eta = 0.6$, $b = 100$, $K = 5$, $k = 1$, and $d = 1$. Substituting them into (7), one can derive ψ^N and ψ^S and verify that Assumption 1 holds. Following the steps discussed in the previous section, the time paths of $(\theta_{0,t}, \theta_{1,t})$ can be derived iteratively given the parameter values of the model. The result for the benchmark case is shown in the middle panel of Figure 4 and repeated in Figures 5–13 for comparisons with alternative parameter values.

We perform a series of experiments in which we perturb the value of one parameter at a time

with respect to the benchmark case and derive the corresponding time paths of $(\theta_{0,t}, \theta_{1,t})$. For each parameter, two alternative values to the benchmark value are tried to illustrate the effects of an increase and a decrease in the corresponding parameter value. We discuss the findings below.

The first experiment examines the effect of the level of infrastructure in the South on the speed and range of FDI. The benchmark is perturbed with respect to the parameter K: $K = \{5.25, 5, 4.75\}$. The results are shown in Figure 4, and they indicate that a better infrastructure in the South attracts a wider range of intermediate firms in the first wave of industry migration, which in turn creates a larger externality and leads to a bigger second wave of industry migration. The result: at steady state, a wider range of intermediate firms undertake FDI in the South that has a higher level of infrastructure. It is straightforward to see that if the level of infrastructure in the South falls significantly below $K = 4.75$, the first wave of FDI will not kick off at all.

The second experiment studies the effects of the economic distance between the host and the home country of FDI, where we vary the parameter d: $d = \{10, 1, 0.5\}$. The results are illustrated in Figure 5. Conditional on the same level of infrastructure in the South, K, the same range of firms undertake FDI in the first wave of industry migration. However, given the same absolute mass of firms producing in the South in the first period, a shorter distance between the host and the home country facilitates faster information spillover and leads to a larger subsequent wave of industry migration. At steady state, a wider range of intermediate firms produce in the South that is closer to the North.

Figure 6 illustrates the effects of the relative wage advantage of the South, w^N, in determining the extent of FDI. The range of variation for this parameter is: $w^N = \{4.2, 4, 3.8\}$. As shown by the figure, the higher the wage in the North relative to the South, the range of the first wave of industry migration shifts up, and so does the steady-state range of intermediate firms located in the South. The intuition is easier seen in the linear case where FDI incurs no risk. Recall that the lower bound of firms who remain in the North is $\tilde{\theta}_N \equiv \frac{w^N f^N}{\psi^N}$ and the lower bound of firms who undertake FDI is $\tilde{\theta}_S \equiv \frac{w^N (f^S - f^N)}{\psi^S - \psi^N}$. It follows that $\tilde{\theta}_S = \tilde{\theta}_N \frac{f^S/f^N - 1}{\psi^S/\psi^N - 1}$. By Equations (5) and (7), it can be verified that as the relative wage in the North rises, $\tilde{\theta}_N$ and $\tilde{\theta}_S$ both increase, but $\tilde{\theta}_S/\tilde{\theta}_N$ decreases. This corresponds to the higher lower bound θ_0 in Figure 6 with a higher w^N for the nonlinear case. On the other hand, the upper bound θ_1 is also higher with a higher w^N. This is because in the linear case, as w^N increases, although both the profit lines of the South and

of the North shifts down, the profit line of the South becomes relatively steeper compared to the profit line of the North ($\psi^S/\psi^N = w^{N(1-\eta)\frac{\alpha}{1-\alpha}}$ increases as w^N increases). This implies that in the nonlinear case, the concave profit curve of the South π^S will intersect the profit line of the North at a higher upper bound.

The next two experiments look at how the relative management cost disadvantage of international vertical integration affects the extent of FDI. This is reflected by the fixed organizational cost f^S when the intermediate input is produced in the South and f^N when it is produced in the North. The experiments are $f^S = \{2.05,\ 2,\ 1.95\}$ and $f^N = \{1.08,\ 1,\ 0.92\}$, respectively. The results in Figure 7 show that the required fixed organizational cost to have the intermediate input produced in the South is negatively correlated with the extent of FDI. As f^S increases, the curve π^S shifts down uniformly, thus intersecting the schedule π^N at a larger lower bound and a smaller upper bound. Thus, a smaller range of intermediate firms undertake FDI in the first wave of migration. The effect then propagates to the subsequent waves of migration. At steady state, a smaller range of intermediate firms produce in the South that entails a higher fixed organizational cost. Figure 8 shows that the fixed organizational cost of having the intermediate input produced in the North has exactly the opposite effect. As f^N increases, a larger range of intermediate firms relocate their production of intermediate inputs to the South in the first wave of migration. The trend continues until the steady state.

Figure 9 demonstrates the effects of sectoral production technology, as reflected by the intensity of headquarter service, on the extent of FDI. The experiments are $\eta = \{0.615,\ 0.6,\ 0.585\}$. The results suggest that as the intensity of headquarter service increases, fewer intermediate firms will undertake FDI. This is because the reduction in the variable cost by producing in the South becomes less important when the intermediate input is a smaller proportion of the final product ($\psi^S/\psi^N = w^{N(1-\eta)\frac{\alpha}{1-\alpha}}$ decreases as η increases).

Next, recall that the contract between the headquarter and the component supplier is non-enforceable ex post. Given this institutional constraint, Figures 10–11 indicate that if the headquarter is more powerful in terms of either its ability to appropriate the intermediate input when the ex post negotiation fails (δ), or its ability to extract from the ex post surplus (β), the extent of FDI in the South will be smaller. This is because when the headquarter has the upper hand in the dealing with the component supplier, the distortion of underinvestment by the component supplier

14

will be larger, and the significance of variable cost saving by producing in the South relatively smaller. As a result, the extent of FDI will be smaller at the steady state. [check]

Figure 12 illustrates the effects of sectoral demand condition α on the extent of FDI. With a larger α, the demand elasticity for the product $\sigma = \frac{1}{1-\alpha}$ is larger. Thus, the saving in variable cost by producing in the South becomes more important. This is also reflected by the fact that $\psi^S/\psi^N = w^{N(1-\eta)\frac{\alpha}{1-\alpha}}$ increases as α increases. In this case, a higher α will lead to a wider range of intermediate firms undertaking FDI in the first wave of migration, as shown in Figure 12. As a result, the extent of FDI is also larger at the steady state. [add discussions of Figure 13]

6   Conclusion

In this paper, we develop a dynamic theoretical model of FDI with firm heterogeneity in productivity. We show that with the presence of FDI uncertainty, firms with intermediate productivity levels will undertake FDI ahead of firms in the lower tier and upper tier of productivity levels. This is contrary to the results of Antràs and Helpman (2004) where FDI propensity rises with firm productivity. Given a first wave of industry migration, we then demonstrate the magnet effect of FDI, where the presence of some FDI in the South creates positive externality for the later comers by reducing the general uncertainty. This triggers a second wave of industry migration; a wider range of intermediate firms now find it profitable to produce in the South. The enlarged mass of firms located in the South further reduce FDI uncertainty and attract a third wave of migration. The process continues until the power of the magnet effect reaches its steady state and an equilibrium structure of international specialization emerges.

We conduct simulations to study the effects of several key parameters on the equilibrium structure of international production. We find that the extent of FDI undertaken by firms from the North (measured by the range of productivity levels of the intermediate firms) is positively correlated with the level of infrastructure in the South, the fixed organizational cost of producing in the North, and the demand elasticity for a product. On the other hand, it is negatively correlated with the economic distance between the North and the South, the fixed organizational cost of producing in the South, the intensity of headquarter service, and the relative bargaining power of the headquarter versus the component supplier. As the relative wage of the North increases, however,

both the lower bound and the upper bound of migrated firms shift up.

References

Antràs, P., Helpman, E., 2004. Global sourcing. Journal of Political Economy 112, 552–580.

Globerman, S., Shapiro, D., 2002. Governance infrastructure and U.S. foreign direct investment. Manuscript.

Wei, S.-J., 2000. How taxing is corruption on international investors. Review of Economics and Statistics 82, 1–11.

Figure 1: Firm Productivity and FDI Timing: FDI in China by Taiwanese Firms
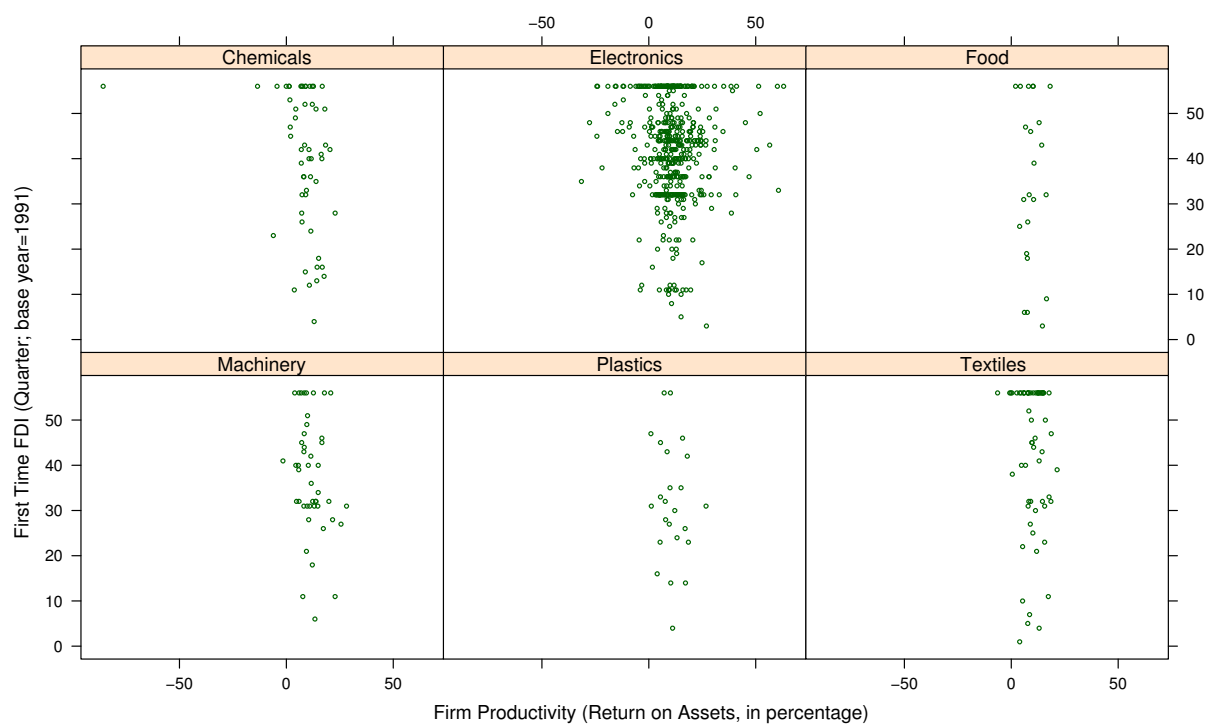
Note: This figure shows the relationship between the timing that a Taiwanese firm undertakes its first FDI in mainland China and its productivity. The timing is measured as the number of quarters elapsed since the Taiwanese government lifted the ban in 1991, and the productivity of a firm is measured by the ratio of Return on Assets. We take the average of the ratios for a firm between 1991 and 1993. The data source is described in the text.

Figure 2: $\pi^S(\tilde{\theta}, z)$ as z varies

$\pi^S(\tilde{\theta}, z)|\ z \to 0$

$\pi^N(\tilde{\theta})$

π^S, π^N

$\pi^S(\tilde{\theta}, z)|\ z < z^*$

$\pi^S(\tilde{\theta}, z)|\ z = z^*$

$\tilde{\theta}_N$

$\pi^S(\tilde{\theta}, z)|\ z > z^*$

0

$\tilde{b}$

$\tilde{\theta}_0$

$\tilde{\theta}_1$

$\tilde{\theta}$

$\tilde{\theta}'$

$\tilde{\theta}_S$

$\psi^S \tilde{b} - w^N f^S$

$\psi^S \tilde{b} - w^N f^N$

$\pi^S(\tilde{\theta}, z)|\ z \to 1$
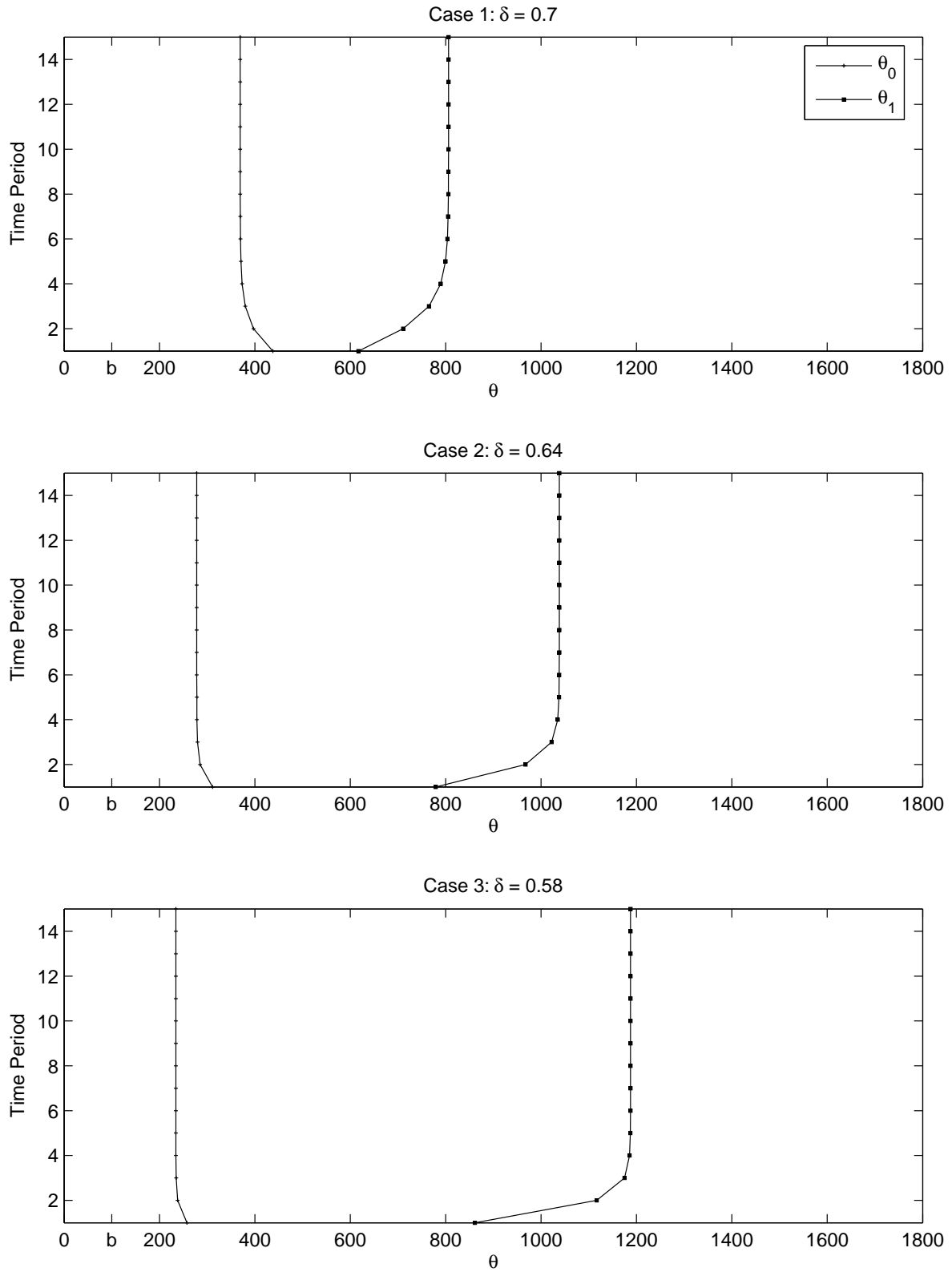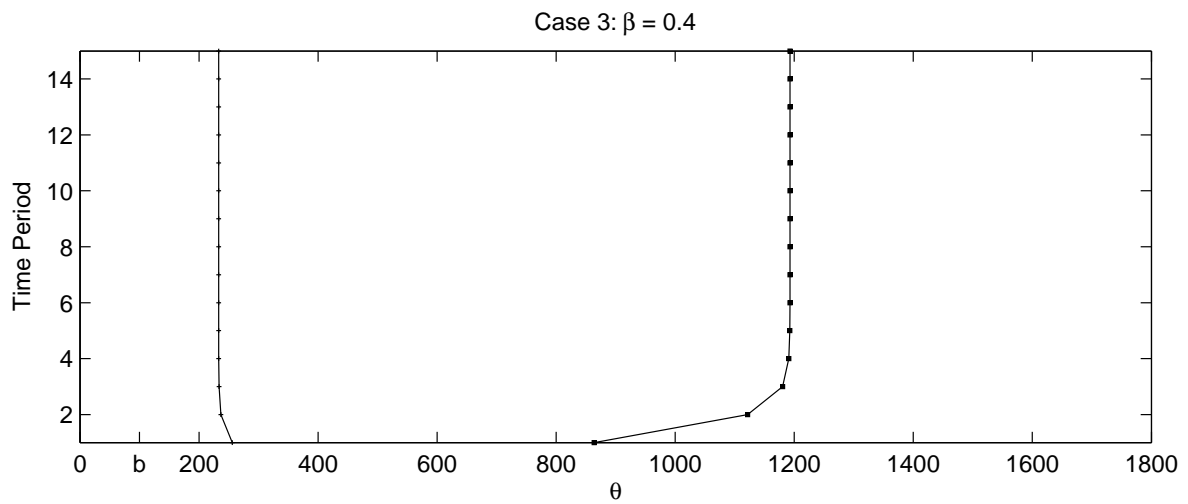
A

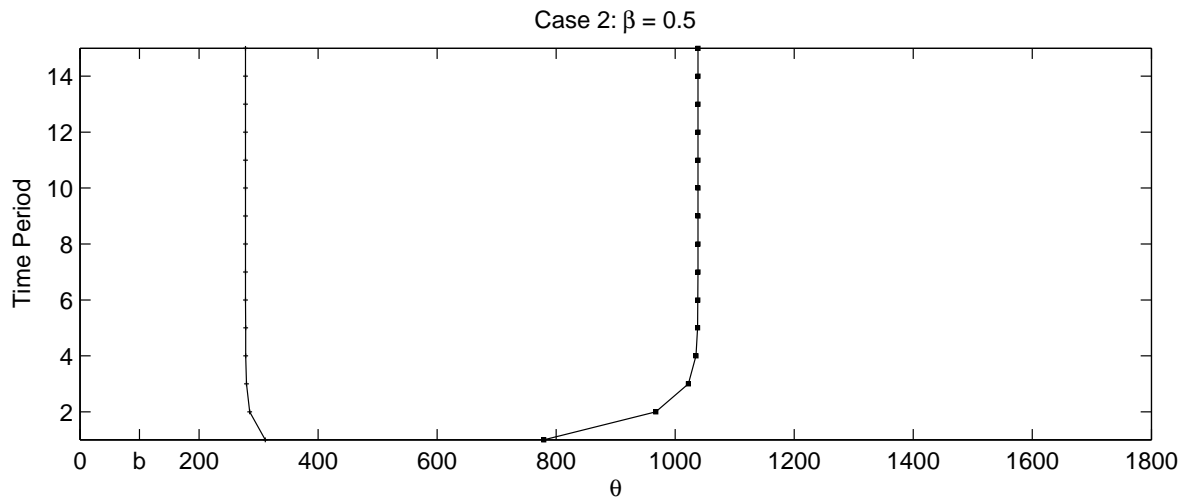$-w^N f^S$
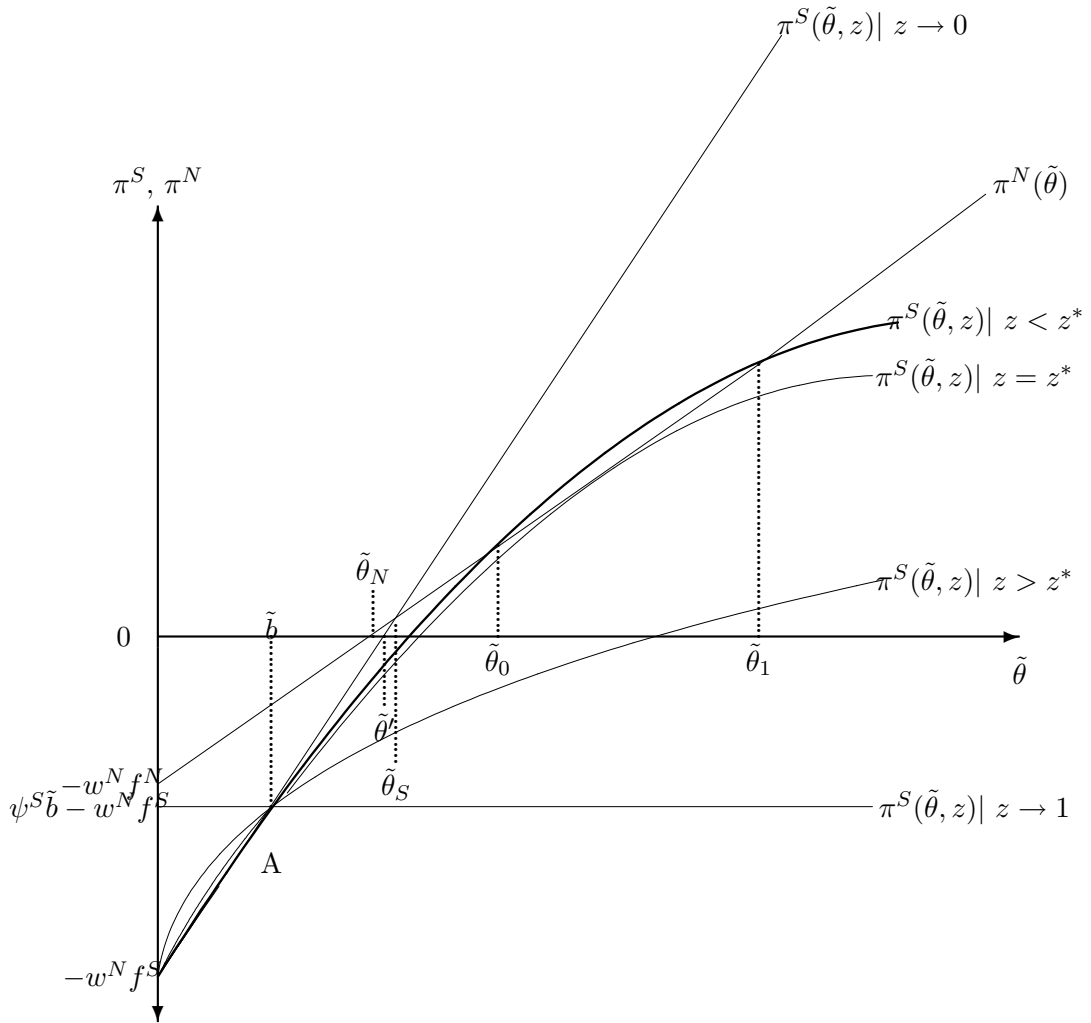
Figure 3: Magnet Effect of FDI
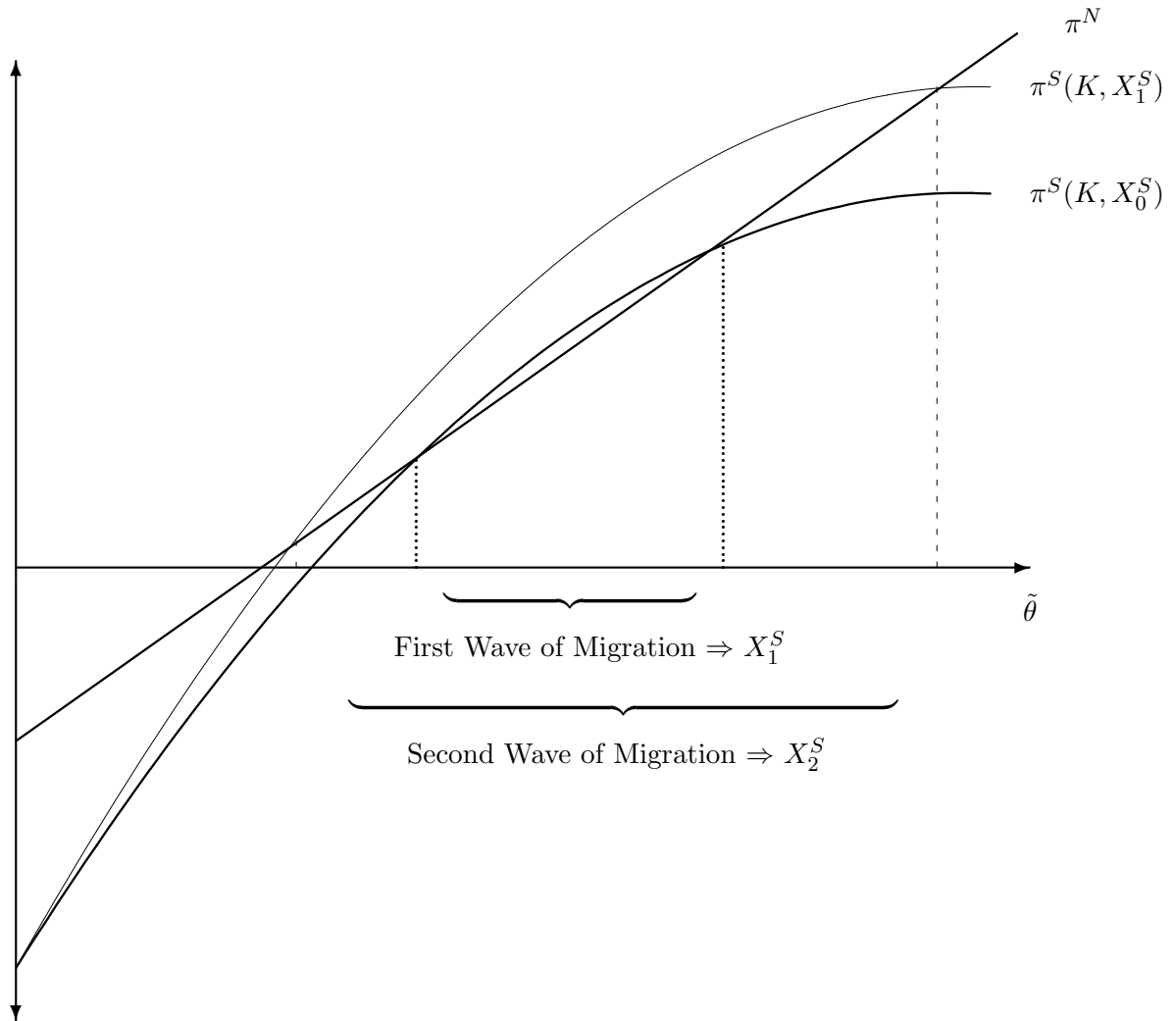
Figure 4: Effects of K on the Speed and Range of FDI
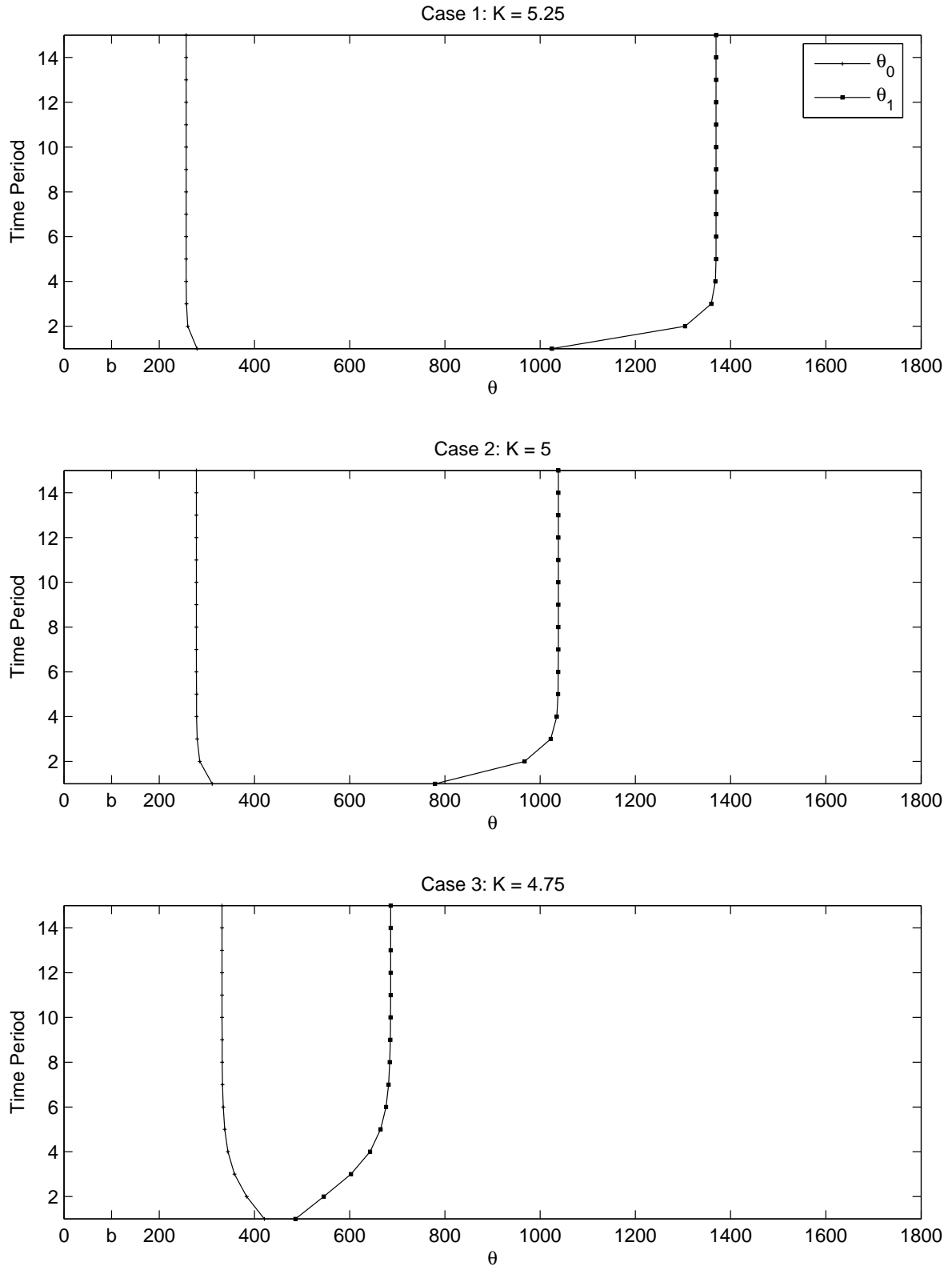
Figure 5: Effects of d on the Speed and Range of FDI
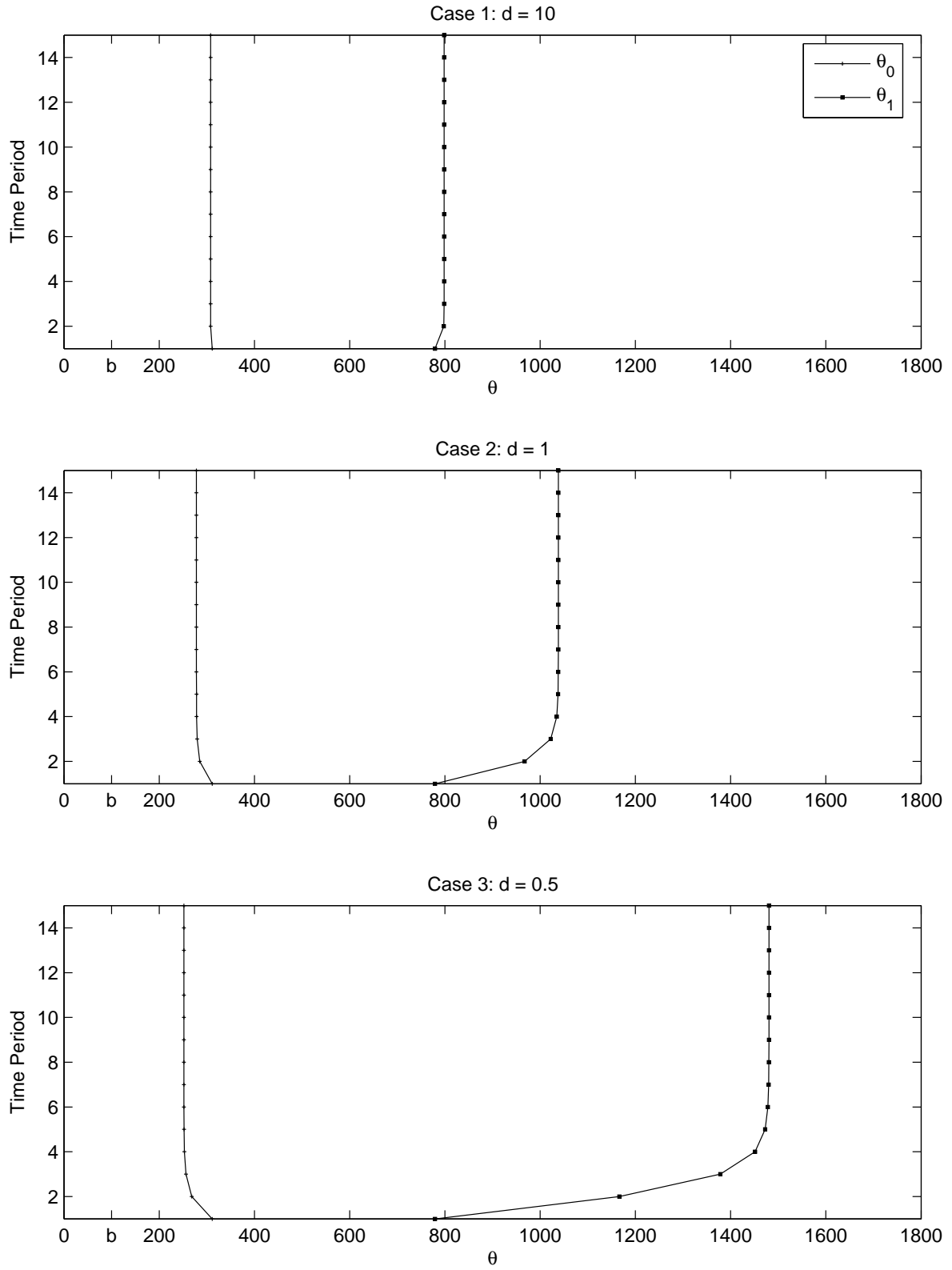
Figure 6: Effects of w^N on the Speed and Range of FDI
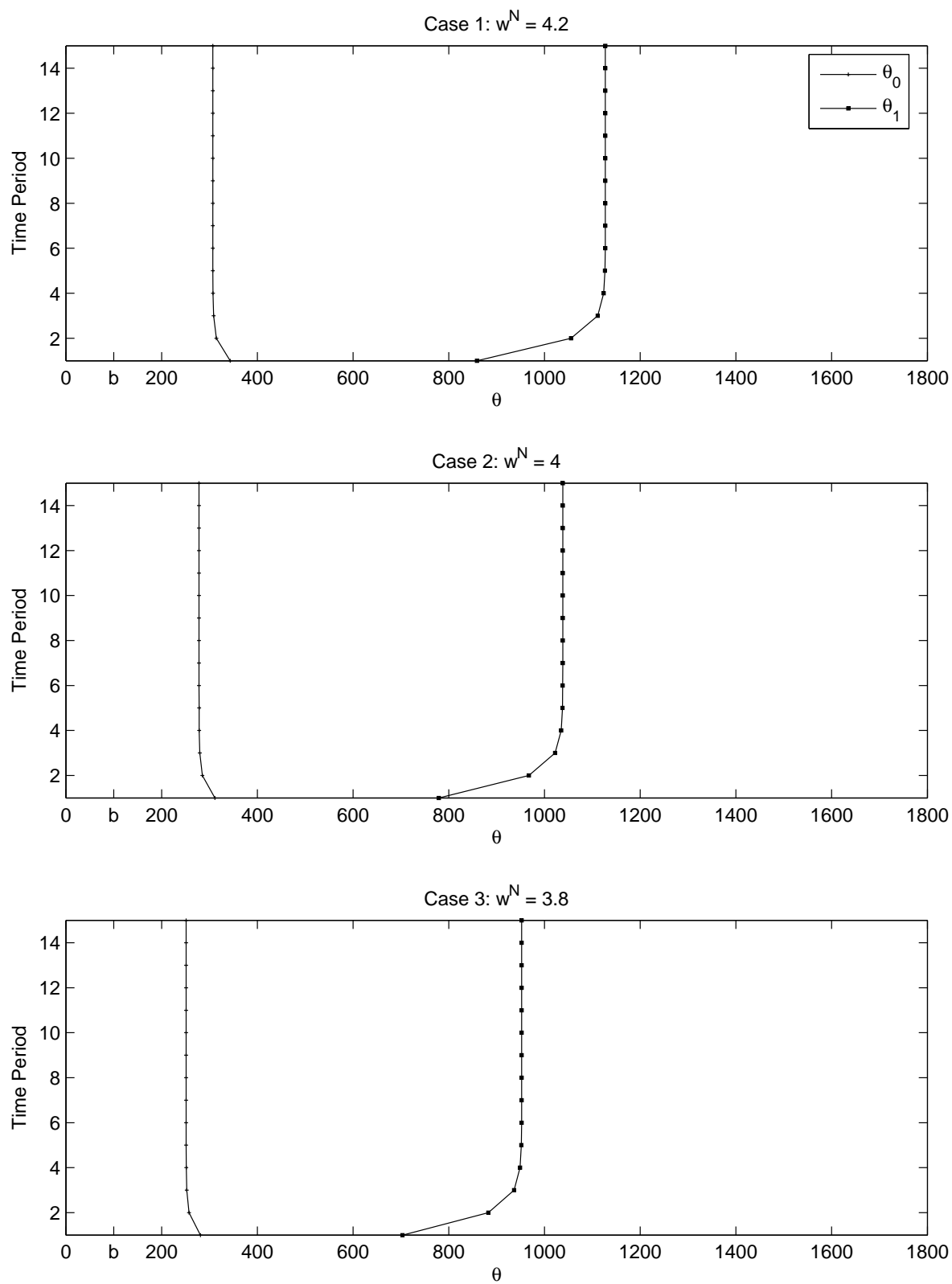
Figure 7: Effects of f^S on the Speed and Range of FDI
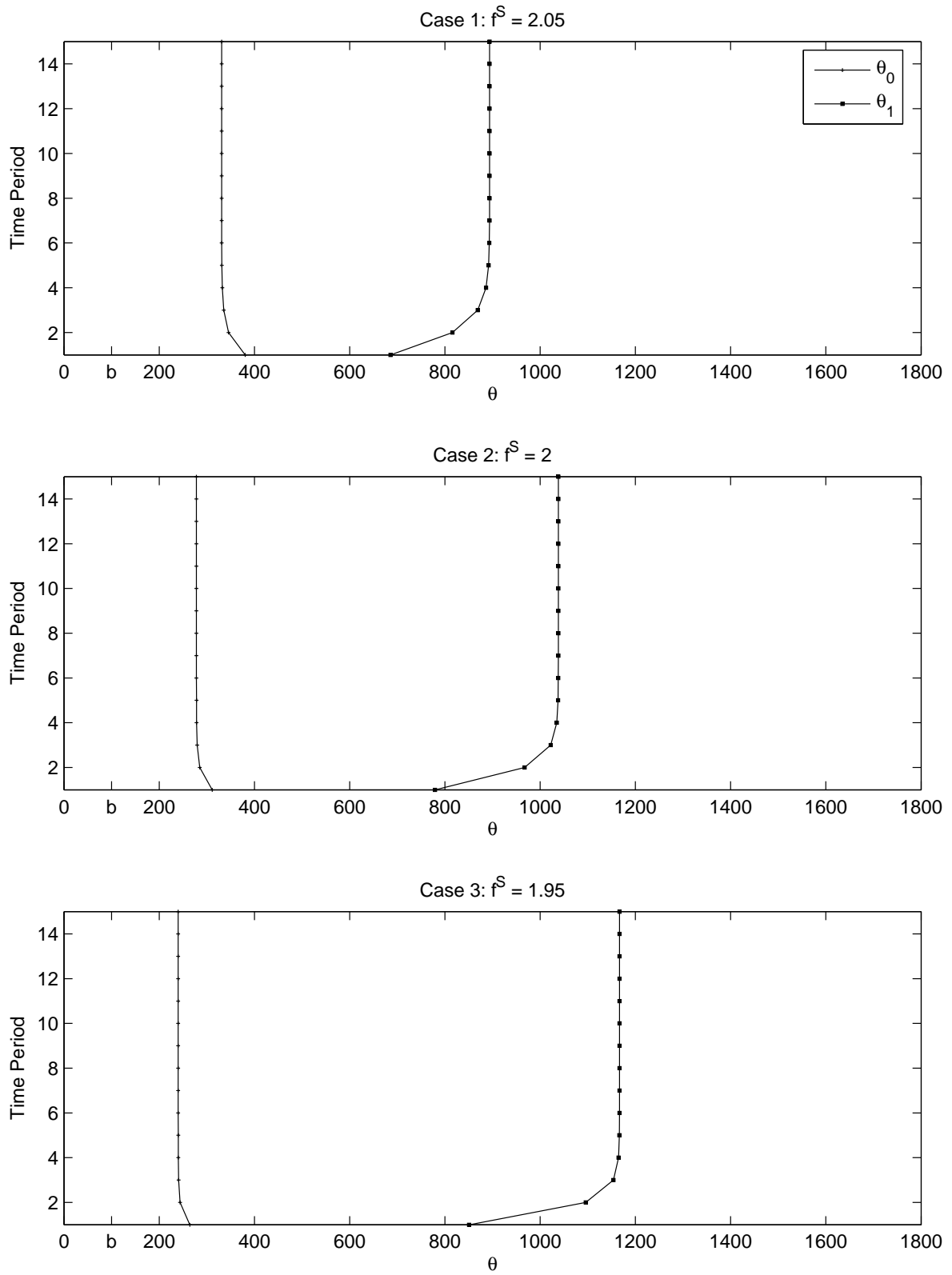
Figure 8: Effects of f^N on the Speed and Range of FDI
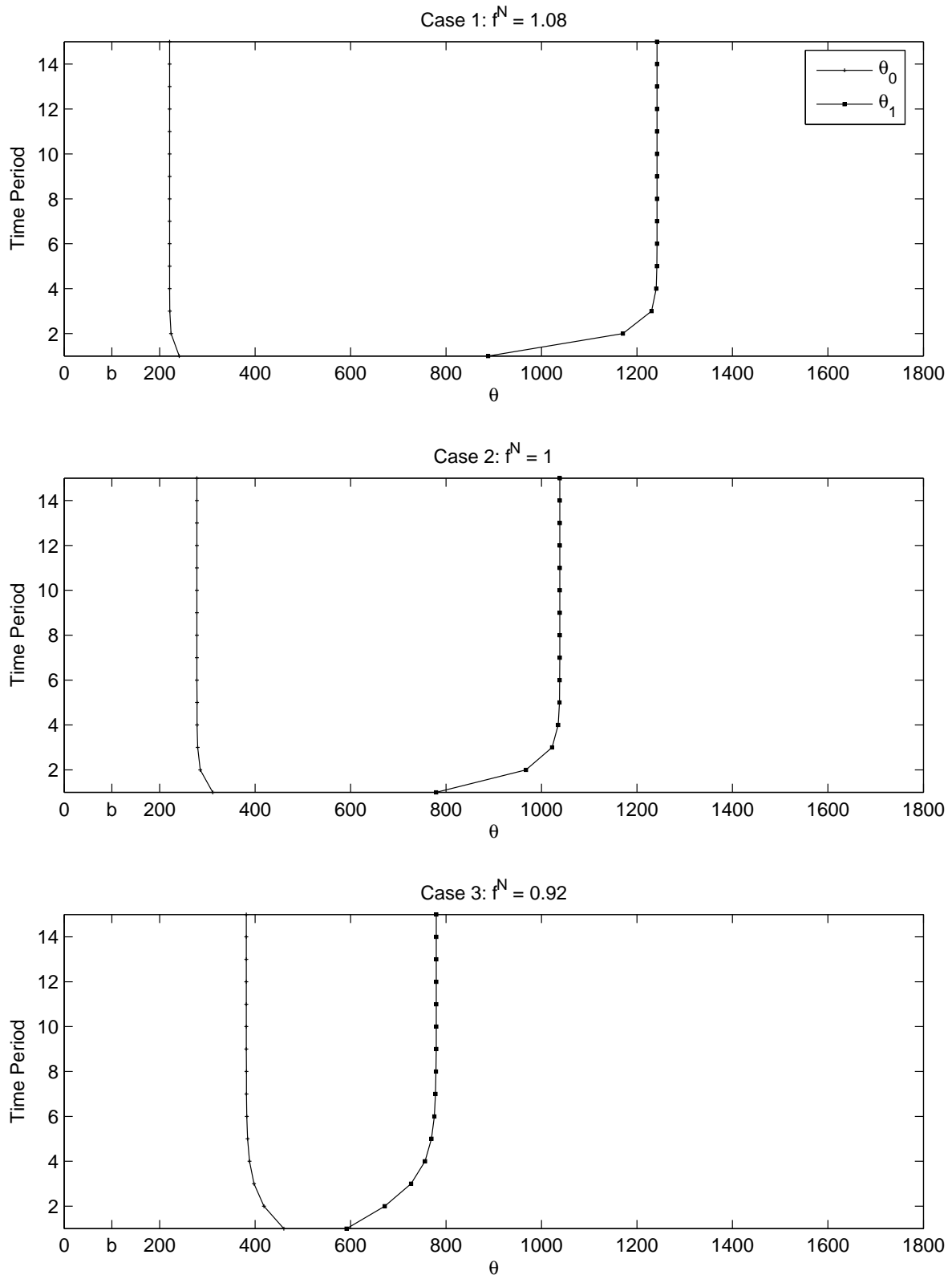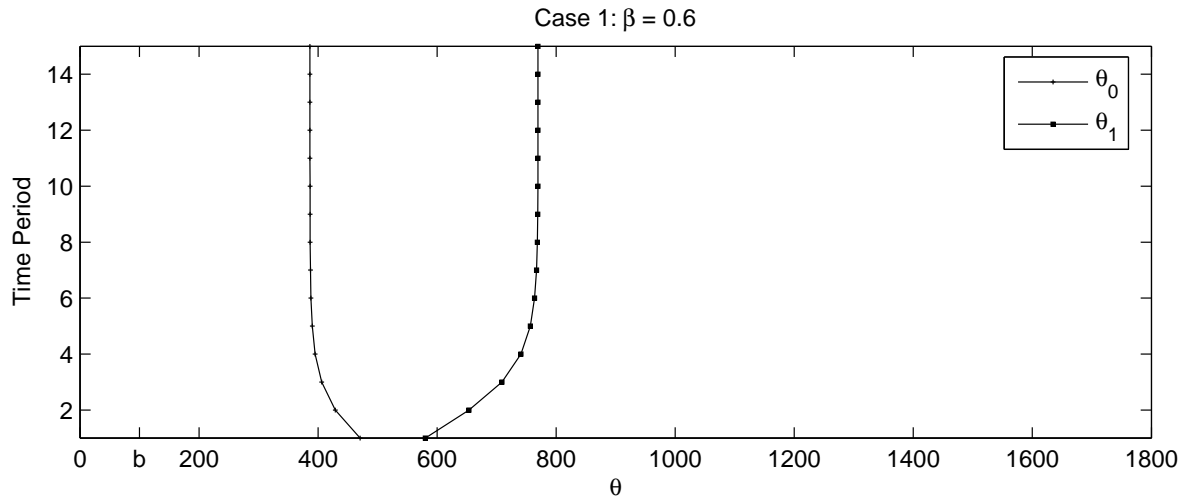
Figure 9: Effects of η on the Speed and Range of FDI
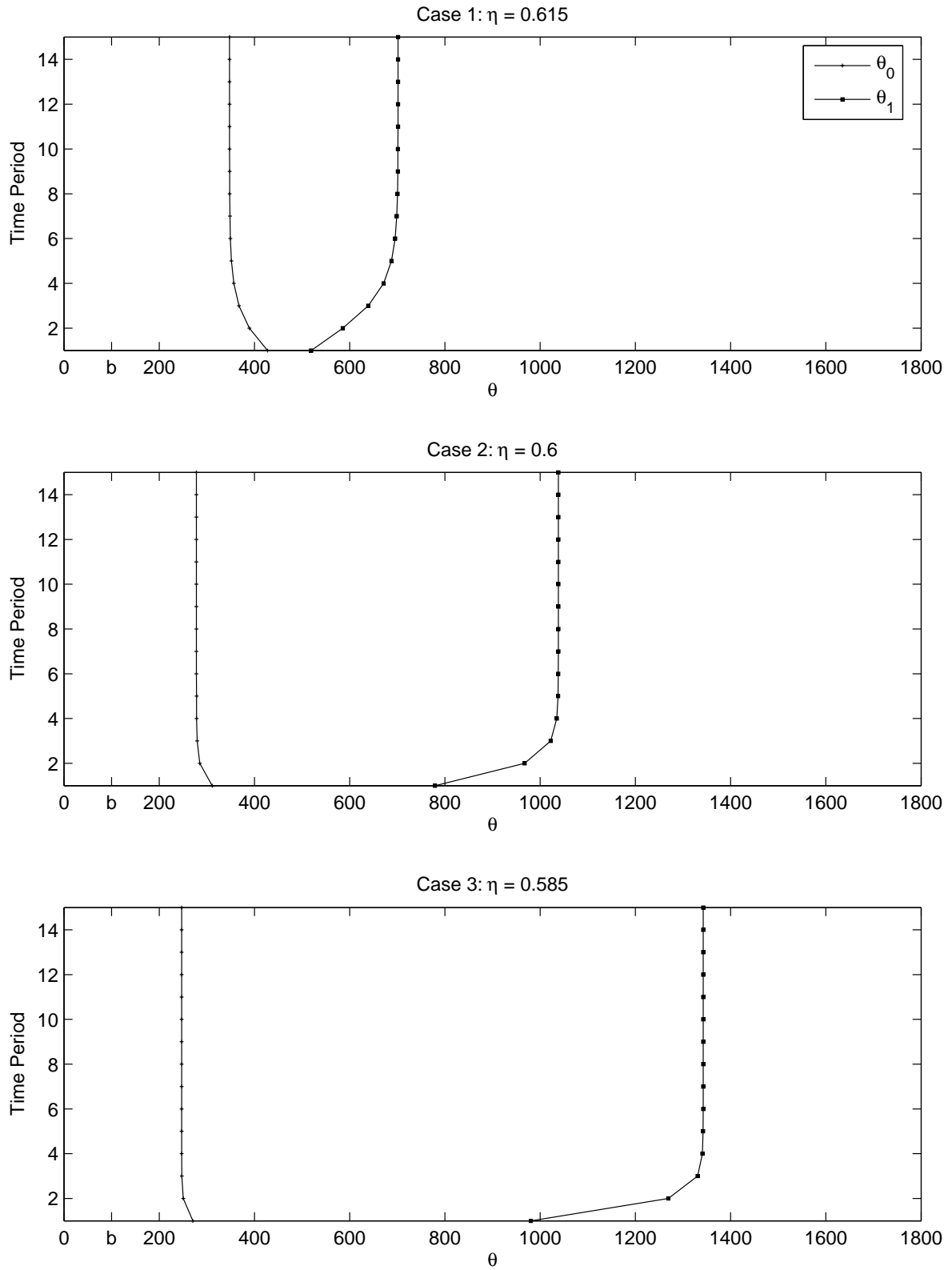
Figure 10: Effects of δ on the Speed and Range of FDI

Figure 11: Effects of β on the Speed and Range of FDI
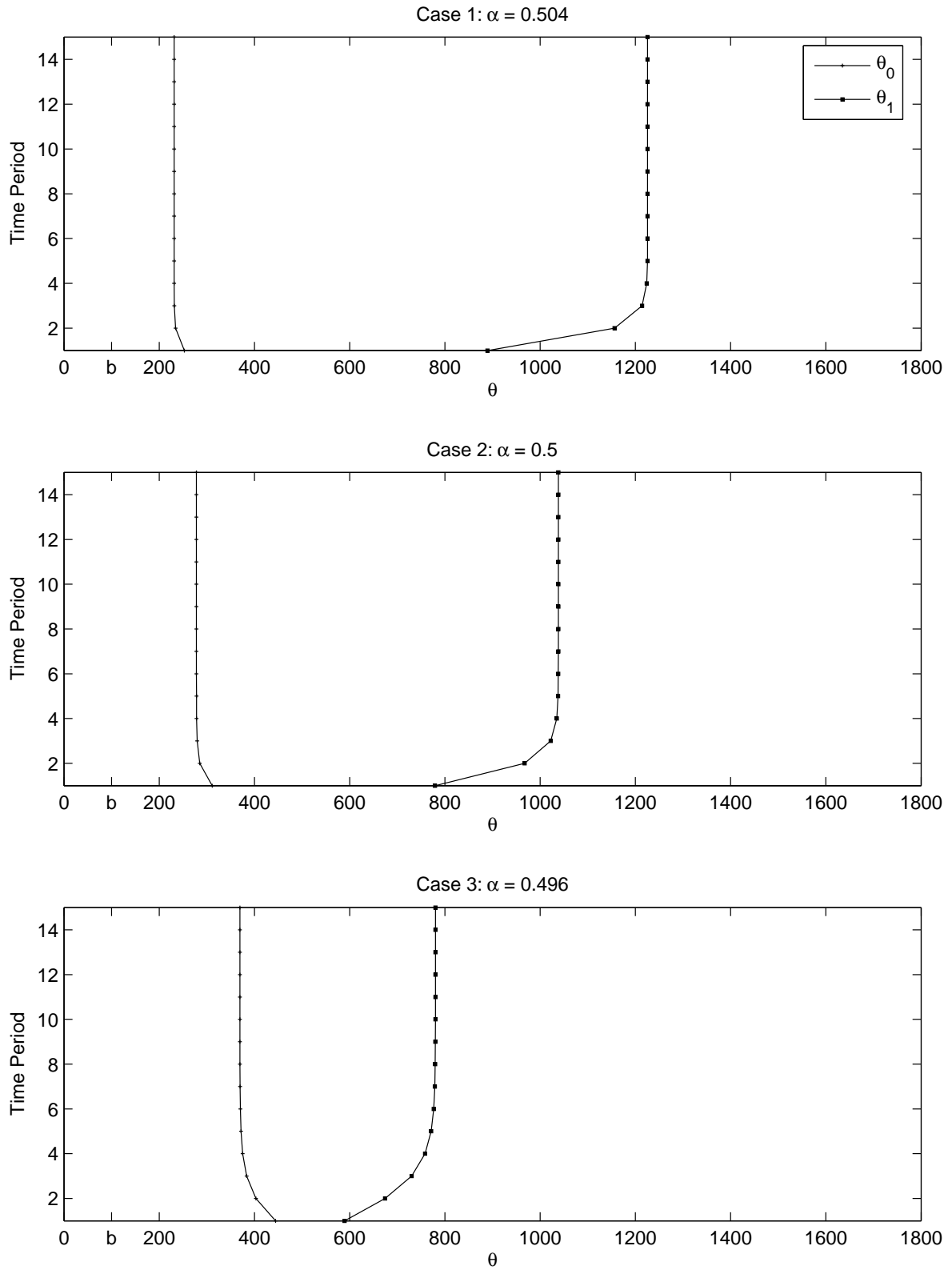
Figure 12: Effects of α on the Speed and Range of FDI

Figure 13: Effects of k on the Speed and Range of FDI

Case 1: k = 2

Case 2: k = 1

Case 3: k = 0.5