Contagion of Self-Fulfilling Financial Crises
Due to Diversification of Investment Portfolios

Itay Goldstein* and Ady Pauzner**

First version: March, 2000
This version: 5 June, 2001

ABSTRACT
We explore a model with two countries. Each might be subject to a self-fulfilling crisis, induced by agents withdrawing their investments in the fear that others will do so. While the fundamentals of the two countries are independent, the fact that they share the same group of investors may generate a contagion of crises. The realization of a crisis in one country reduces agents’ wealth and thus makes them more risk averse (we assume decreasing absolute risk aversion). This reduces their incentive to keep their investments in the second country since doing so exposes them to the strategic risk associated with the unknown behavior of other agents. Consequently, the probability of a crisis in the second country rises. This yields a positive correlation between the returns on investments in the two countries even though the two are completely independent in terms of fundamentals.

* We thank David Frankel, Elhanan Helpman, Stephen Morris, and Assaf Razin for helpful comments.
* Department of Economics, Princeton University. E-mail: igoldste@princeton.edu.
** The Eitan Berglas School of Economics, Tel Aviv University. E-mail: pauzner@post.tau.ac.il. Web: http://www.tau.ac.il/~pauzner
1. Introduction

In recent years, financial markets have become increasingly open to international capital flows.\(^1\) This process of globalization is usually praised for creating opportunities to diversify investment portfolios. At the same time, the financial world has witnessed a number of cases in which financial crises spread among different countries.\(^2\) This phenomenon of contagion has occurred despite the fact that the different countries did not seem to have any common economic fundamentals.

In this paper we present a model in which contagion of financial crises arises precisely because investment portfolios are diversified across countries. The fact that different countries share the same group of investors leads to the transmission of negative shocks from one part of the world to another. Thus, the realization of a financial crisis in one country can induce crises in other countries as well. This generates a positive correlation between the returns on the investments in different countries, and thus reduces the effectiveness of diversifying investments across countries.

We focus on self-fulfilling crises: crises that occur just because agents believe they are going to occur. This feature is important as financial crises are often viewed as the result of a coordination failure among economic agents.\(^3\) While recent literature has provided theoretical foundations for either the contagion of crises or for the possibility of self-fulfilling crises, models in which both co-exist have rarely been studied. The difficulty in trying to demonstrate contagion in a model of self-fulfilling beliefs comes from the fact that models of self-fulfilling beliefs are often characterized by multiple equilibrium outcomes, and thus do not account for the likelihood of each equilibrium. As a result, such models cannot capture a contagion effect in which a crisis in one country affects the likelihood of a crisis in another.

---

1 See, e.g., Bordo, Eichengreen, and Irwin (1999).
2 See, e.g., Krugman (2000).
To tackle this difficulty, we employ a technique introduced by Carlsson and van-Damme and recently applied in a number of papers that explore financial crises. This technique allows us to determine the likelihood of each outcome and relate it to observable variables. We find that the likelihood of a crisis decreases with agents’ wealth. Hence, the occurrence of a crisis in one country, which reduces this wealth, increases the likelihood of a self-fulfilling crisis in a second country.

Agents in our model hold investments in two countries. Investment can either be held to maturity, in which case the return increases with the fundamentals of the country and the number of agents who keep their investments there, or can be withdrawn prematurely for a fixed payoff. In most cases, if no one withdraws their investments early in a certain country, then each agent can obtain a higher return by keeping her investment in that country until it matures. But if all agents withdraw early, the long-term return becomes lower than the return for early withdrawal. As a result, agents might coordinate on withdrawing early in a country, even though they could get higher returns by coordinating on keeping their investments there. The beliefs of agents regarding the behavior of other agents in a country will determine whether there will be a financial crisis, i.e., a mass withdrawal of investments.

We examine a sequential framework in which the events in country 2 take place after the aggregate outcomes in country 1 (which depend on fundamentals and the behavior of agents there) are realized and known to all agents. Following Carlsson and van-Damme (1993), we assume that agents do not have common knowledge of the fundamentals of country 2, but rather get slightly noisy signals about them after they are realized. This can be due to agents having access to different sources of information or to slight differences in their interpretation of publicly available information. This structure of information enables us to uniquely determine the beliefs and behavior of agents in country 2 as a function of the fundamentals of country 2 and of the outcomes in country 1. We show

---


5 This can be due, e.g., to increasing returns to scale in aggregate investment or to liquidity constraints.

6 This kind of financial crisis is similar to the one described by Diamond and Dybvig (1983).
that agents will withdraw early in country 2 only if the fundamentals there are below a certain threshold. Importantly, this threshold level depends on the outcomes in country 1. In most circumstances, the coordination of agents on withdrawing their investments in country 1 early increases the threshold and thus increases the probability of observing a crisis in country 2. We refer to this effect as ‘contagion’.

The mechanism that generates contagion in our model originates in a wealth effect. In most cases, the occurrence of a crisis in country 1 reduces the wealth of agents. We assume that agents have decreasing absolute risk aversion. Thus, a crisis in country 1 makes them more risk averse when choosing their actions in country 2. Since keeping their investments in country 2 is a risky action, agents will have weaker incentives to do so following a crisis in country 1.

It is important to note that the risk involved in keeping one’s investment in country 2 does not result from the uncertainty about the level of the fundamentals in that country. This uncertainty is negligible since agents get rather precise signals about the level of these fundamentals. Rather, it is a strategic risk: a risk that originates in the unknown behavior of other agents in country 2. When an agent chooses to maintain her investment, her return depends on the actions of other agents. Thus, when she has less wealth, her incentive to withdraw early and obtain a return that does not depend on others’ behavior is increased.

While strategic risk seems to be a natural ingredient of an interaction that has strategic complementarities, such a risk is not captured in models that assume common knowledge of the fundamentals. In these models, each agent is certain about the equilibrium behavior of other agents and thus the strategic risk disappears. Here, however, an agent who observe a signal which is close to the threshold at which agents switch actions is uncertain about the behavior of other agents. Thus, the change in wealth has a direct effect on her behavior. This has a considerable effect on the threshold signal at which agents switch between keeping or withdrawing their investment.

Having demonstrated the existence of contagion in our model, we analyze the behavior of agents in country 1. We show that there exists an equilibrium in country 1 in which
agents withdraw early in country 1 only if the realization of the fundamentals in this country is below a certain level.\(^7\) In this equilibrium, an endogenous positive correlation exists between the returns on investments in the two countries. When fundamentals in country 1 are low, a crisis occurs there and the return on investment is low. Following this, a crisis is more likely to occur in country 2 as well, implying a higher likelihood of obtaining a low return there also. It is important to note that this positive correlation is obtained even though we have assumed that the fundamentals of the two countries are completely independent of one another. Thus, the positive correlation can only be the result of the contagion effect described above.

This raises an interesting issue: In a case where each agent invests in only one country, the returns on investments in the two countries are completely independent of one another. In that case, agents can benefit from diversifying investments among the two countries. However, because of the contagion of crises described above, this diversification generates a positive correlation between the two investments and thus reduces the benefits from diversification. Thus, the globalization of capital that enables agents to invest in different countries bears a cost: It generates positive correlation between assets that were otherwise independent and thus reduces the means to diversify risk. This correlation can be viewed as a negative externality of agents’ diversifying their investments. Consequently, one may wonder whether full openness of financial markets is indeed the optimal regime.

A few recent papers have studied contagion. Masson (1998) discusses the possibility that self-fulfilling crises will be contagious, but does not present a mechanism through which a crisis in one country might induce a change in beliefs in another. Dasgupta (2000) uses Carlsson and van-Damme’s technique in order to provide such a mechanism. However, the mechanism in his paper is different from ours; it relies on the existence of capital links between different financial institutions. Other authors show that contagion can be the result of optimal portfolio allocations made by investors, but do not discuss self-fulfilling crises. Among these, Kyle and Xiong (2000) present a model that is more closely related to ours, since it also describes contagion as a consequence of a wealth 

\(^7\) We are not, however, able to prove that this is the unique equilibrium in country 1.
effect. Allen and Gale (2000) and Lagunoff and Schreft (1999) present models in which the connections between banks or projects induce a chain of crises. Again, they do not analyze contagion of self-fulfilling crises. Some authors analyze contagion as a transmission of information. In these models, a crisis in one market reveals some information about the fundamentals in the other, and thus might induce a crisis in the other market as well. Examples are King and Wadhwani (1990), Calvo (1999) and Chen (1999). Calvo and Mendoza (2000) suggest that the high cost of gathering information on individual countries might induce rational contagion.

The remainder of this paper is organized as follows: Section 2 presents the basic model. In section 3 we study the equilibrium behavior of agents in country 2 given the aggregate outcomes in country 1. We then show that contagion exists in our model. In section 4 we study the equilibrium behavior of agents in country 1 and demonstrate the positive correlation between the returns on the two investments in this equilibrium. Section 5 concludes. Proofs are relegated to the Appendix.

2. The Model

There is a continuum [0,1] of identical agents. Their utility from consumption, \( u(c) \), is twice continuously differentiable, increasing, and satisfies decreasing absolute risk aversion, that is, \(-u''(c)/u'(c)\) is decreasing. Each agent holds an investment of 1 in each of two countries (1 and 2).

An agent can choose when to withdraw each of her two investments. The (gross) return on investment in country \( i \) is 1 if withdrawn prematurely or \( R(\theta_i, n_i) \) if withdrawn at maturity. Long-term return \( R \) in country \( i \) is increasing in the fundamentals \( \theta_i \) of that country and decreasing in the proportion \( n_i \) of agents who withdraw their investments in this country early. The fact that the return is decreasing in \( n_i \) may represent increasing returns on aggregate investment at country \( i \) or liquidity constraints.

---

8 Kodres and Pritsker (1998) and Schinasi and Smith (1999) also relate contagion to portfolio allocations.
9 While we assume that agents initially split their investments equally between the two countries, this could be an endogenous property if each country was, ex-ante, as likely to become country 1.
An agent decides when to withdraw her investment in country $i$ after receiving information on the fundamentals in that country. The fundamentals $\theta_1$ and $\theta_2$ are independent and drawn from a uniform distribution on $[0,1]$. We assume that the fundamentals are not publicly reported after they are realized. Instead, each agent $j$ obtains a noisy signal $\theta_j^i$ on the fundamentals of country $i$, where $\theta_j^i = \theta_i + \epsilon_j^i$ and $\epsilon_j^i$ are error terms which are uniformly distributed over the interval $[-\epsilon,\epsilon]$ and independent across agents and countries. We will focus on the case in which signals are very precise, i.e. that $\epsilon$ is close to 0.

Clearly, an agent’s incentive to wait until her investment in country $i$ matures is higher when the country's fundamentals are good and when the number of agents who are going to withdraw early in this country is low. However, while the optimal behavior of an agent in country $i$ usually depends on her belief regarding the behavior of other agents in this country, we assume that there are small ranges of the fundamentals in which agents have dominant actions. More specifically, when the fundamentals of country $i$ are very good, an agent will prefer to keep her investment there until it matures no matter what she believes other agents will do. Similarly, when the fundamentals in country $i$ are very bad, the agent will withdraw her investment in that country prematurely even if she believes that all the other agents will maintain their investments there.

Formally, we assume that there exist $0 < \theta < \overline{\theta} < 1$ such that $R(\theta,0)=1$ and $R(\overline{\theta},1)=1$. As a result, when an agent observes a signal $\theta_j^i < \theta - \epsilon$, she knows that $R_i < 1$ no matter what other agents are going to do in country $i$. Thus, she will decide to withdraw her investment in country $i$ in this case. Similarly, if an agent observes $\theta_j^i > \overline{\theta} + \epsilon$, she will decide to keep her investment in country $i$ until it matures. Again, for most possible signals, i.e. when $\theta_j^i$ is between $\theta - \epsilon$ and $\overline{\theta} + \epsilon$, the optimal behavior of an agent in country 2 will depend on her belief regarding the behavior of other agents there.

We look at a sequential model in which activity takes place first in country 1 and then in country 2. In the first stage, the value of the fundamentals in country 1 is realized, agents receive signals regarding the fundamentals and decide whether to withdraw their investments there prematurely or not. In the second stage, the value of the fundamentals in
country 2 is realized, agents observe signals regarding this value and decide on their actions in this country as well. The exact realization of country 1 fundamentals, as well as the aggregate behavior in country 1, is known to agents before they choose their actions in country 2. The order of events is depicted in Figure 1:

![Order of Events Diagram](image)

We solve the model backwards. First, we analyze the equilibrium behavior of agents in country 2 for each possible outcome in country 1. This enables us to explore the effect of the behavior of agents in country 1 on their behavior in country 2. Then, we analyze the equilibrium behavior of agents in country 1, when they take into account the effect of the outcomes in country 1 on the equilibrium in country 2. This enables us to explore the correlation between the returns on the investments in the two countries in equilibrium. The solution of the model is described in the next two sections.

---

10 In equilibrium, it is sufficient that agents receive information regarding either the fundamental or the aggregate behavior, since one can be inferred from the other.
3. The Behavior of Agents in Country 2: Contagion of Crises

Equilibrium in Country 2

In her decision whether to run or not in country 2, an agent must take into account all relevant and available information. This includes her signal $\theta^j_2$ of country 2’s fundamentals and her wealth $w^j_1$ resulting from her investment in country 1, since these directly affect her incentive to run. Moreover, since her payoff depends on other agents’ behavior and since this behavior might depend on their own wealth, the agent must also consider the distribution of wealth in the population. (The agent is also concerned about the signals observed by the other agents; however, all she knows regarding them is her own signal $\theta^j_2$.)

Suppose that agent $j$ believes that the proportion of other agents who will run in country 2, as a function of country 2’s fundamentals, is given by $n^j_2(\theta_2)$. The difference between the utility she expects in the case that she keeps her investment in country 2 until it matures and the utility she expects if she withdraws early is:

$$\Delta^j_2(\theta^j_2, n^j_2(\cdot), w^j_1) = \frac{1}{2\varepsilon} \int_{\theta_2 = \theta^j_2 - \varepsilon}^{\theta^j_2 + \varepsilon} [u(R(\theta_2, n^j_2(\theta_2)) + w^j_1) - u(1 + w^j_1)]d\theta_2.$$

(Note that when agent $j$ observes signal $\theta^j_2$, her posterior belief over the fundamentals in country 2 is uniformly distributed between $[\theta^j_2 - \varepsilon]$ and $[\theta^j_2 + \varepsilon]$. This is because her prior belief over $\theta_2$ is uniformly distributed and the signal error $\varepsilon^j_2$ is uniformly distributed over $[-\varepsilon, \varepsilon]$.)

For a given distribution of wealth among the population, an agent’s strategy is a function from her signal to an action – run ($r$) or not run ($nr$). The profile of strategies of all agents induces a function $n^j_2(\theta_2)$ which determines the number of agents who run given the true state of fundamentals (this number is deterministic since there is a continuum of agents).
In equilibrium, all agents know $n_2(\theta_1)$ (i.e., $n_1^j(\theta_2) = n_2^j(\theta_2)$ for all $j$). Thus, in equilibrium, it must be that each agent $j$ runs if and only if $\Delta_2(\theta_2^j, n_2^j(\cdot), w_i^j) < 0$.

The distribution of wealth consists of two mass points: the $n_1$ agents who ran in country 1 have wealth 1, whereas the 1-$n_1$ who did not have wealth $R(\theta_1, n_1)$.

As a result, an agent’s equilibrium strategy may depend on her group. Proposition 1 says that for any distribution of wealth (as determined by $n_1$ and $\theta_1$), there is a unique equilibrium in country 2. The equilibrium is characterized by two threshold signals: Each agent runs if she observes a signal below the threshold corresponding to her group, and does not if her signal is above it.

**PROPOSITION 1:** For any $\theta_1$ and $n_1 \in [0,1]$, there exists a unique equilibrium in country 2. In this equilibrium, each agent who ran in country 1 runs in country 2 if her signal $\theta_2^1$ is below $\theta_2^{1,}\nu$ and does not run if the signal is above, whereas an agent who did not run in country 1 runs in country 2 if her signal is below $\theta_2^{1,}\nu$ and does not run above.

It is important to note that although the behavior of agents is uniquely determined, crises (i.e. mass withdrawals of investment) in country 2 are self-fulfilling. In the range in which agents do not have dominant actions, i.e., between $\theta_1$ and $\overline{\theta}$, whenever agents run they do so only because they believe other agents are going to. The crucial point is that the special structure of information in the model uniquely determines agents’ beliefs and these, in turn, determine their behavior.

Models in which such a structure of information leads to a unique equilibrium are now common in the literature. However, in most of them (see, for example, Morris and Shin (1998) and Goldstein and Pauzner (2000)) agents are homogeneous. Here, in contrast, agents who act in country 2 belong to two different groups. This makes the analysis more complicated. Yet, because strategies are complementary not only within a group but also across groups (i.e., the incentive of an agent to withdraw early in country 2 increases if

---

It might be that $n_1$ equals 0 or 1 or that $R(\theta_1, n_1) = 1$. In these cases all agents have the same wealth.
more agents of either group withdraw early), the uniqueness of equilibrium holds also in our case.\textsuperscript{12}

While the two thresholds are distinct, they must be very close if agents’ signals are very precise. Lemma 1 says that the distance between the two is of order $\varepsilon$.

**LEMMA 1:** $|\theta_{2,r}^* - \theta_{2,nt}^*| \leq 2\varepsilon$

The intuition is as follows. If distance were larger than $2\varepsilon$, than the support of the posterior distribution over $\theta$ for an agent who observes the higher threshold signal would be above that of an agent who observes the lower threshold signal. (This is because the noise in signals is no more than $\varepsilon$.) Likewise, the support of her distribution over the number of agents who run would be below. Thus, independent of her wealth, she would have a higher incentive to maintain her investment, contradicting the fact that both should be indifferent.

Lemma 2 says that the two thresholds converge to some limit as $\varepsilon$ shrinks:

**LEMMA 2:** As $\varepsilon$ approaches 0, both $\theta_{2,r}^*$ and $\theta_{2,nt}^*$ converge to $\theta_2^*$, which is implicitly defined by the following two equations (with unknowns $\theta_2^*$ and $x$):

$$0 = \frac{1}{\theta_2 - (1 - \theta_2^*)} \left[ u(1 + R(\theta_2^*, ((1 - n_1) \cdot n(\theta_2, \theta_2^*) + n_1 \cdot n(\theta_2, \theta_2^* + x)))) - u(1 + 1) \right]$$

$$0 = \frac{1}{\theta_2 - (1 - \theta_2^*)} \left[ u(R(\theta_1, n_1) + R(\theta_2^*, ((1 - n_1) \cdot n(\theta_2, \theta_2^*) + n_1 \cdot n(\theta_2, \theta_2^* + x)))) - u(R(\theta_1, n_1) + 1) \right]$$

where $n(\theta_2, z) = \begin{cases} 1 & \text{if } z - 1 > \theta_2 \\ \frac{z + 1 - \theta_2}{2} & \text{if } z + 1 \geq \theta_2 \geq z - 1 \\ 0 & \text{if } \theta_2 > z + 1 \end{cases}$.

\textsuperscript{12} Another application in which a unique equilibrium is obtained with two types of agents can be found in Goldstein (2000). Frankel, Morris and Pauzner (2000) provide a general theoretical treatment; they show that a unique equilibrium must exist in any many player, many action game with strategic complementarities.
We know that, in the limit, all agents run below $\theta_2^*$ and do not run above. Because $\theta_2$ is uniformly distributed between 0 and 1, $\theta_2^*$ also represents the probability of a crisis in country 2. From now on, we will focus on the properties of this limit threshold.

**The effect of wealth from country 1 operations**

Threshold $\theta_2^*$ depends on the distribution of wealth, as determined by $n_1$ and $\theta_1$. We now turn to study the effect of these parameters on the behavior of agents in country 2. Theorem 1 says that if the population is wealthier (in distribution), the probability of a crisis in country 2 is decreased:

**THEOREM 1**: If the distribution of agents’ wealth corresponding to $n_1'$ and $\theta_1'$ first-order stochastically dominates that corresponding to $n_1$ and $\theta_1$, then $\theta_2^*(\theta_1,n_1) < \theta_2^*(\theta_1',n_1')$.

The intuition behind this result is as follows: In country 2, each agent has to choose between two actions. The first action is safe: The agent withdraws her investment in country 2 early and receives a certain return of 1. The second action is risky: The agent keeps her investment in country 2 until it matures and receives an uncertain return then. Because agents' risk aversion decreases with wealth, those agents with increased wealth from their country 1 operations will be more willing to bear risks. As a result, these agents will coordinate on keeping their investments in country 2 over a wider range of realizations of the fundamentals in this country. Consequently, and because of the strategic complementarities, those agents whose wealth has not changed will also have a stronger incentive to maintain their investment. As a result, the threshold below which all agents run in country 2 will fall.

It is important to note that the risk involved in not withdrawing early in country 2 is not a result of the uncertainty about the level of the fundamentals in that country. This is because agents have very precise information about the level of these fundamentals, which makes this uncertainty negligible. Rather, agents face a *strategic risk*: When they choose to maintain their investments, their return depends on the unknown behavior of
other agents. In other words, agents in our model are averse to being in situations where their payoff depends on the behavior of others. This aversion, however, decreases with their level of wealth.

The implications of Theorem 1 go beyond our model. Consider, for example, a hypothetical case in which all agents have an identical level of wealth which is given exogenously. According to the theorem, the likelihood of a run in country 2 decreases with that level of wealth. This case can be interpreted as a situation where country 1 is a developed country in which investments are already established: the returns do not depend on the number of investors but do depend on the fundamentals of the economy. (Thus, the return in country 1 would simply be $R(\theta_1)$.) Country 2 could be thought of as an emerging market, in which investments are still reversible and returns do depend on the aggregate level of investment (i.e., the return is $R(\theta_2, n_2)$). In such a scenario, bad news regarding the developed country’s fundamentals might generate a crisis in the emerging economy.

Contagion of crises

Returning to our model (in which the two countries can be thought of as emerging markets), Theorem 1 implies that there is a contagion effect: the behavior of agents in country 1 affects their behavior in country 2. For a given realization of $\theta_1$ above the lower dominance region, when there is a run in country 1, the distribution of wealth is lower than in the case of no run. Thus, $\theta_2' (\theta_1, 1)$ is above $\theta_2' (\theta_1, 0)$. This implies that a run in country 1 increases the likelihood of a run in country 2. This is stated in the next corollary and shown in Figure 2:

COROLLARY: Assume that $R(\theta_1, 0) > 1$. There is a range of country 2 fundamentals in which: if there was a run in country 1 ($n_1=1$) then there will be a run in country 2, and if there was no run in country 1 ($n_1=0$) then there will be no run in country 2.

---

13 In fact, only agents who get signals that are very close to the threshold have uncertainty about the behavior of other agents. Thus, a change in the level of wealth will have a direct effect only on the behavior of these agents. However, since the optimal behavior of agents who observe other signals depends on the behavior of these agents, the change in wealth will have an indirect effect on the behavior of other agents as well. Thus, a change in wealth will change the behavior of a large group of agents.
To complete the analysis, we study the effect of changes in $n_1$ between 0 and 1 on the equilibrium behavior in country 2. Recall that $R(\theta_1,n_1)$ is decreasing in $n_1$, exceeds 1 when $n_1$ is small and falls below 1 when $n_1$ is close to 1. In the range where $R(\theta_1,n_1)$ is greater than 1, an increase in $n_1$ has a negative effect on the distribution of wealth among agents. The reason is that the number of agents who run and receive 1 becomes larger and the number of agents who wait and receive $R(\theta_1,n_1)$ becomes smaller. Moreover, the wealth of agents in the second group is reduced since $R(\theta_1,n_1)$ decreases in $n_1$. Thus, by theorem 1, when more agents run in country 1 there is a higher likelihood of a run in country 2.

In the range where $R(\theta_1,n_1)$ is below 1, however, the effect of $n_1$ on $\theta_2^*$ becomes ambiguous. When an additional agent decides to run, her wealth is increased from $R(\theta_1,n_1)$ to 1. On the other hand, the wealth of those agents who do not run is decreased. Nonetheless, we do know that if $n_1$ is increased to 1, agents’ wealth is increased since in that case all agents receive a return of 1. Figure 3 summarizes these results:

Figure 3: The effect of $n_1$ on $\theta_2^*$ for a given realization of $\theta_1$ at which $R(\theta_1,0)>1$
4. Behavior of Agents in Country 1 and the Endogenous Correlation Between the Returns on Investment

Equilibrium in country 1

When agent $j$ chooses her action in country 1, the only information she has is her signal $\theta_j^i$ regarding the fundamentals of that country. Her decision also depends on her belief $n_j^i(\theta_j)$ regarding the number of agents who are going to withdraw early in country 1 as a function of the fundamentals in that country. Another relevant factor is the wealth, $w_j^i$, that she expects to obtain from her investment in country 2. This wealth will depend on $\theta_2$ and is therefore considered to be random at this stage. However, according to the results of the previous section, it can be uniquely determined (for any level of $\theta_2$) by $\theta_1$ and $n_1$. It is approximately (for very small noise in signals) given by:

$$w_j^i(\theta_j, n_j^i(\theta_j); \theta_2) = \begin{cases} 1 & \theta_2 \leq \Delta_2^i(\theta_j, n_j^i(\theta_j)) \\ R(\theta_2, 0) & \theta_2 > \Delta_2^i(\theta_j, n_j^i(\theta_j)) \end{cases}.$$ 

Agent $j$ will withdraw early in country 1 if and only if $\Delta_1(\theta_j^i, n_j^i(\theta_j)) < 0$, where $\Delta_1$ denotes the difference between the utility that the agent expects to achieve in the case that she keeps her investment in country 1 until it matures and the utility she expects to achieve in the case that she withdraws early. It is given by:

$$\Delta_1(\theta_j^i, n_j^i(\theta_j)) = \int_{\theta_j = \theta_j^i - \varepsilon}^{\theta_j^i + \varepsilon} \int_{\theta_2 = 0}^{1} \left[ u(R(\theta_j^i, n_j^i(\theta_j)) + w_j^i(\theta_j^i, n_j^i(\theta_j), \theta_2)) ight] d\theta_2 d\theta_j.$$

The analysis of equilibrium behavior in country 1 is more involved than that of country 2 (Section 3). The reason is that apart from the effect of $\theta_1$ and $n_1$ on $R(\theta_1, n_1)$, which directly affects the desirability of early withdrawal, there is also an indirect effect: $\theta_1$ and

---

14 When $\varepsilon$ is positive, the expression is more complicated since there are two distinct thresholds describing the behavior of agents who did or did not run in country 1. Moreover, there are nontrivial ranges of $\theta_2$ where the number of agents who run in country 2 is strictly between 0 and 1. While for the purpose of exposition we describe here the limit as $\varepsilon \to 0$, the proofs that appear in the appendix use the ‘correct’ specification.
$n_1$ determine $w_2$, which in turn affects $\Delta_i$. As a result, we are unable to show that the equilibrium in country 1 is unique. We can, however, show the existence of a threshold equilibrium, i.e., an equilibrium in which all agents withdraw early when they observe a signal below some common threshold $\theta_i^*$ and wait if they observe a signal above.

**THEOREM 2:** There exists a threshold equilibrium in country 1.

**Correlation**

We now focus on a given threshold equilibrium which is characterized by the threshold signal $\theta_i^*$. With a small amount of noise in signals, the behavior of agents in country 1 is approximately as follows: All agents run in country 1 when the fundamentals there are below $\theta_i^*$, whereas none of them does so when the fundamentals in country 1 are above $\theta_i^*$. In the first case, all agents possess wealth $w_i = 1$, while in the second each has wealth $R(\theta_i,0) > 1$. By the results of section 3, in the first case agents will run in country 2 when fundamentals are below threshold $\theta_2^*(w_i = 1)$, which is higher than the threshold $\theta_2^*(w_i = R(\theta_i,0))$ corresponding to the second case. This is illustrated in Figure 4 (note that since $\theta_2$ is uniformly distributed between 0 and 1, $\theta_2^*$ is simply the probability of a run in country 2):

![Figure 4: The probability of a run in country 2 as a function of the fundamentals in country 1.](image-url)
As shown in the figure, when the fundamentals in country 1 are below $\theta_1^*$, the probability of a crisis in country 2 is fixed at $\theta_1^*(w_i = 1)$. When the fundamentals in country 1 reach the level of $\theta_1^*$ there is a sharp decline in the probability of a run in country 2. This occurs because at this level of fundamentals, agents switch from running in country 1 to waiting and therefore enjoy a sharp increase in the wealth they attain from their operations in that country. Finally, when the fundamentals in country 1 keep increase beyond $\theta_1^*$, the probability of a run in country 2 gradually decreases. This occurs because at this level of fundamentals, agents do not run in country 1 and their wealth $w_i = R(\theta_1,0)$ increases gradually with the level of fundamentals there.

Thus, an endogenous spillover effect exists: the level of fundamentals in country 1 affects the probability of crisis in country 2. This generates a positive correlation between the returns on investments in the two countries. Importantly, this correlation is achieved in spite of the fact that the fundamentals in the two countries are completely independent of each other. This positive correlation is the result of the wealth effect which induces agents to coordinate to a larger extent on the unfavorable equilibrium in country 2 when they obtain lower returns on their investments in country 1.

Figure 5 demonstrates the positive correlation between the returns on investments in the two countries (in the threshold equilibrium previously described). The return on investment in country 1 is $w_i(\theta_1)$. It equals 1 below $\theta_1^*$ and $R(\theta_1,0)$ above. The expected return in country 2 is $Ew_2(\theta_1) = \theta_2^*(n_i(\theta_1)) + \int_{\theta_2 = \theta_2^*(n_i(\theta_1))}^{R(\theta_2,0)} d\theta_2$. ($n_i(\theta_1)$ equals 1 below $\theta_1^*$ and 0 above). A positive correlation exists because both $w_i(\theta_1)$ and $Ew_2(\theta_1)$ are increasing in $\theta_1$. 
5. Conclusions

We studied contagion of self-fulfilling financial crises. The mechanism that generates contagion in our model is based on a wealth effect. Following a crisis in one country, agents’ wealth is reduced. They are, then, less willing to bear the strategic risk that originates in the unknown behavior of other agents in the other country. As a result, they have a higher tendency to run in this country. This means that the occurrence of a crisis in one country increases the probability of a crisis occurring in the other country as well. This contagion effect generates positive correlation between the returns of the two investments even though the investments are independent in terms of their fundamentals.

This result raises an interesting issue. In a world where each agent invests only in one country, the returns on investments in two different countries are independent of one another (recall that fundamentals are uncorrelated). Thus, each agent may benefit by investing in both countries since the diversification reduces her risk. However, as shown in the above analysis, when each agent invests in both countries, the returns on the two investments become positively correlated. This weakens the potential benefits from diversification. In other words, there is a negative externality: when one agent diversifies her portfolio, the benefits to other agents from diversification are reduced.
This leaves an open question: can a regime that only allows agents to *partially* diversify their investment portfolio Pareto-dominate the full globalization regime in which investors can freely allocate their investments between countries?
Appendix

Proof of Proposition 1

The proof for the cases in which all agents ran in country 1 \((n=1)\) or none did \((n=0)\) is standard – see e.g., Morris and Shin (1999). The proof for the case in which both groups are nonempty is given below. It is based on the technique of Frankel, Morris and Pauzner (2000).

Let \(n_{1,r}^j(\theta_2)\) \((n_{1,\text{nr}}^j(\theta_2))\) denote agent’s \(j\)’s belief regarding the number of agents who ran (did not run) in country 1 and are going to run in country 2. Let \(\Delta_{2,r}(\theta_2, n_{2,\text{nr}}^j(\theta_2), n_{2,r}^j(\theta_2))\) \((\Delta_{2,\text{nr}}(\theta_2, n_{2,\text{nr}}^j(\theta_2), n_{2,r}^j(\theta_2)))\) denote the difference in expected utility between waiting in country 2 and running there for an agent who ran (did not run) in country 1. These functions are given by:

\[
\Delta_{2,r}(\theta_2, n_{2,\text{nr}}^j(\theta_2), n_{2,r}^j(\theta_2)) = \frac{1}{2\epsilon} \int_{\theta_2 - \epsilon}^{\theta_2 + \epsilon} \left[ u(1 + R(\theta_2, n_{2,\text{nr}}^j(\theta_2) + n_{2,r}^j(\theta_2))) - u(1 + 1) \right] d\theta,
\]

and

\[
\Delta_{2,\text{nr}}(\theta_2, n_{2,\text{nr}}^j(\theta_2), n_{2,r}^j(\theta_2)) = \frac{1}{2\epsilon} \int_{\theta_2 - \epsilon}^{\theta_2 + \epsilon} \left[ u(R(\theta_2, n_{2,\text{nr}}^j(\theta_2) + n_{2,r}^j(\theta_2))) - u(R(\theta_2, n_{2,\text{nr}}^j(\theta_2) + 1 + 1) \right] d\theta.
\]

Because \(R\) is decreasing in \(n\), both \(\Delta_{2,r}\) and \(\Delta_{2,\text{nr}}\) are weakly decreasing in \(n_{2,r}^j\) and in \(n_{2,\text{nr}}^j\) (i.e., if \(n_{2,r}^j(\theta_2) \geq n_{2,\text{nr}}^j(\theta_2)\) for all \(\theta_2\) then \(\Delta_{2,r}(\theta_2, n_{2,\text{nr}}^j(\theta_2), n_{2,r}^j(\theta_2)) \leq \Delta_{2,r}(\theta_2, n_{2,\text{nr}}^j(\theta_2), n_{2,r}^j(\theta_2))\), and similarly for \(n_{2,\text{nr}}^j\)). Thus, the game between the agents satisfies strategic complementarities: an agent’s incentive to run is higher if more agents run at each \(\theta_2\). It is also easy to see that if \(n_{2,r}^j(\theta_2)\) \((n_{2,\text{nr}}^j(\theta_2))\) are weakly decreasing, then, because \(R\) is increasing in \(\theta_2\), \(\Delta_{2,r}\) and \(\Delta_{2,\text{nr}}\) are increasing in \(\theta_2^j\). Moreover, by the assumption of dominance regions, we know that for any beliefs \(n_{2,r}^j\) \((n_{2,\text{nr}}^j, \Delta_{2,r}\) and \(\Delta_{2,\text{nr}}\) are negative for all \(\theta_2^j \leq \theta - \epsilon\) and positive for \(\theta_2^j \geq \overline{\theta} + \epsilon\). Functions \(\Delta_{2,r}\) and
\(\Delta_{2, \text{nr}}\) are also continuous in \(\theta_2^j\) since a small change in \(\theta_2^j\) slightly shifts the interval over which the integrals are computed (and because \(R\) is bounded).

We show that the equilibrium is unique by iterative dominance. We start with the belief that makes agents least willing to run: that all other agents never run (i.e., \(n_{2, \text{nr}}^j = n_{2, r}^j = 0\)). Since \(\Delta_{2, r}\) and \(\Delta_{2, \text{nr}}\) are decreasing in \(\theta_2^j\), there are thresholds \(\theta_{2, r}^1\) and \(\theta_{2, \text{nr}}^1\) such that agents run if they observe a signal below it and do not run if they observe a signal above it (\(\theta_{2, r}^1\) corresponds to agents who ran in country 1 and \(\theta_{2, \text{nr}}^1\) to those who did not).

Because of the strategic complementarities, we know that if agents run below these thresholds under the belief that others never run, then they must do so under any belief.

We now consider the belief that makes agents least willing to run among those that are consistent with the fact that they must run below \(\theta_{2, r}^1\) and \(\theta_{2, \text{nr}}^1\). This is the belief that they run below these thresholds and do not run above them. We obtain new thresholds, \(\theta_{2, r}^2\) and \(\theta_{2, \text{nr}}^2\) below which agents must run. These thresholds are higher than \(\theta_{2, r}^1\) and \(\theta_{2, \text{nr}}^1\), respectively, since they are computed using higher functions \(n_{2, r}^j\) and \(n_{2, \text{nr}}^j\). We iterate this process ad infinitum and denote the limits by \(\theta_{2, r}^\infty\) and \(\theta_{2, \text{nr}}^\infty\). We know that agents run below these thresholds. Moreover, since the iteration stopped there, we know that under the belief that agents run below these thresholds and do not run above them, agents would not run above them.

We now start an iterative process from above. However, this time we work with a translation of the pair \((\theta_{2, r}^\infty, \theta_{2, \text{nr}}^\infty)\). We start with the belief that makes agents most willing to run: that all other agents always run (i.e., \(n_{2, r}^j \equiv n\) and \(n_{2, \text{nr}}^j \equiv 1 - n\)). Let \(x^1\) be the smallest number such that, under this belief, agents do not run above \(\theta_{2, r}^\infty + x^1\) and \(\theta_{2, \text{nr}}^\infty + x^1\) (note that \(x^1\) must be positive since we are using a belief that generates a higher incentive to run relative to the belief that defines \(\theta_{2, r}^\infty\) and \(\theta_{2, \text{nr}}^\infty\)). Knowing that agents do not run above these thresholds, we can obtain a number \(0 < x^2 < x^1\) such that agents do not run above
\( \theta_{2,r}^\infty + x^2 \) and \( \theta_{2, nr}^\infty + x^2 \). We iterate this process ad infinitum and denote the limit of the sequence by \( x^\infty \). We know that agents run above \( \theta_{2,r}^\infty + x^\infty \) and \( \theta_{2, nr}^\infty + x^\infty \). Moreover, because the iteration stopped there, we know that under the belief that agents run below these thresholds and do not run above them, it cannot be the case that there is a positive interval below each one of the thresholds in which agents do not run. This means that under this belief, either \( \Delta_{2,r} = 0 \) at \( \theta_{2,r}^\infty + x^\infty \) or \( \Delta_{2, nr} = 0 \) at \( \theta_{2, nr}^\infty + x^\infty \).

Suppose first that \( \Delta_{2,r} = 0 \) at \( \theta_{2,r}^\infty + x^\infty \). By definition of \( \Delta_{2,r} \), we have

\[
\int_{\theta_{2,r}^\infty + x^\infty - \varepsilon}^{\theta_{2,r}^\infty + x^\infty + \varepsilon} \left[ u(1 + R(\theta_{2}, n_{2, nr}(\theta_{2}, \theta_{2, nr}^\infty + x^\infty) + n_{2, r}^j(\theta_{2}, \theta_{2, r}^\infty + x^\infty))) - u(1 + 1) \right] d\theta = 0
\]

where \( n_{2, j}(\theta_{2}, \theta') \) denotes the belief over the number of agents who run resulting from the belief that agents who ran in country 1 run in country 2 below \( \theta' \) and do not run above, and where \( n_{2, nr}(\theta_{2}, \theta') \) is defined similarly. Changing variables to \( \tilde{\theta}_{2} = \theta_{2} + x^\infty \), we obtain:

\[
\int_{\tilde{\theta}_{2}^\infty - x^\infty - \varepsilon}^{\tilde{\theta}_{2}^\infty - x^\infty + \varepsilon} \left[ u(1 + R(\tilde{\theta}_{2} - x^\infty, n_{2, nr}(\tilde{\theta}_{2}, \theta_{2, nr}^\infty) + n_{2, r}^j(\tilde{\theta}_{2}, \theta_{2, r}^\infty))) - u(1 + 1) \right] d\theta = 0 .
\]

But from the first iteration we know that an agent at \( \theta_{2,r}^\infty \) is indifferent between running or not on the belief that others run below \( \theta_{2,r}^\infty \) and \( \theta_{2, nr}^\infty \), and do not run above them:

\[
\int_{\theta_{2} - \theta_{2,r}^\infty - \varepsilon}^{\theta_{2} - \theta_{2,r}^\infty + \varepsilon} \left[ u(1 + R(\theta_{2}, n_{2, nr}(\theta_{2}, \theta_{2, nr}^\infty) + n_{2, r}^j(\theta_{2}, \theta_{2, r}^\infty))) - u(1 + 1) \right] d\theta = 0 .
\]

Since \( R \) is decreasing in \( \theta \), the two equations can be satisfied only if \( x^\infty = 0 \). In a similar way, we can show that \( x^\infty \) must also equal 0 if \( \Delta_{2, nr} = 0 \) at \( \theta_{2, nr}^\infty + x^\infty \). This means that the limits of the iterations from above and from below coincide. Hence, there is a unique equilibrium in which agents who ran in country 1 run if they observe a signal below \( \theta_{2,r}^\infty \).
and do not run above, and agents who did run in country 1 run if they observe a signal below \( \theta_{2,nr}^* \) and do not run above. QED.

**Proof of Lemma 1**

Suppose that \( \theta_{2,r}^* - \theta_{2,nr}^* > 2\epsilon \). Then, an agent who ran in country 1 and observes \( \theta_{2,r}^* \) in country 2 believes that all the agents who ran in country 1 will not run in country 2. This agent is also indifferent between her two options in country 2. Thus, \( \Delta_{2,r} \left( \theta_{2,r}^*, 0, n_{2,r} \left( \theta_{2}, \theta_{2,r}^* \right) \right) = 0 \). Because the long-term return \( R \) on the investment in country 2 is increasing in \( \theta_{2} \) and decreasing in \( n_{2} \), a necessary condition for this equation to hold is that \( R \left( \theta_{2,r}^* - \epsilon, n_{1} \right) \) be lower than 1. Now consider an agent who did not run in country 1 and observes \( \theta_{2,nr}^* \). She believes that all the agents who ran in country 1 will run in country 2. This agent is also indifferent between her two options in country 2. Thus, \( \Delta_{2,nr} \left( \theta_{2,nr}^*, n_{2,nr} \left( \theta_{2}, \theta_{2,nr}^* \right) \right) = 0 \). A necessary condition for this equation to hold is that \( R \left( \theta_{2,nr}^* + \epsilon, n_{1} \right) \) be higher than 1. However, since \( \theta_{2,r}^* - \theta_{2,nr}^* > 2\epsilon \), this requirement contradicts the former – that \( R \left( \theta_{2,r}^* - \epsilon, n_{1} \right) \) be lower than 1. Similarly, one can show that \( \theta_{2,nr}^* - \theta_{2,r}^* \) cannot be higher than \( 2\epsilon \). QED.

**Proof of Lemma 2**

Consider the set of equations:

\[
0 = E_{\theta_{2}, U \left( \theta_{2}, \theta_{2}^*, n_{2} \left( \theta_{2} \right) \right)} \left[ u(1 + R(\theta_{2}^*, (1-n_{1}) \cdot n^c(\theta_{2}, \theta_{2}^*) + n_{1} \cdot n^c(\theta_{2}, \theta_{2}^* + x \cdot \epsilon))) - u(1 + 1) \right]
\]

\[
0 = E_{\theta_{2}, U \left( \theta_{2} + x \epsilon, \theta_{2}^*, n_{2} \left( \theta_{2} + x \epsilon \right) \right)} \left[ u(R(\theta_{2} + x \epsilon, n_{1}) + R(\theta_{2}^*, (1-n_{1}) \cdot n^c(\theta_{2}, \theta_{2}^*) + n_{1} \cdot n^c(\theta_{2}, \theta_{2}^* + x \cdot \epsilon))) - u(R(\theta_{2} + x \epsilon, n_{1}) + 1) \right]
\]

where \( n^c(\theta_{2}, z) = \begin{cases} 1 & \text{if } z - \epsilon > \theta_{2} \\ \frac{z + \epsilon - \theta_{2}}{\epsilon} & \text{if } z + \epsilon \geq \theta_{2} \geq z - \epsilon \\ 0 & \text{if } \theta_{2} > z + \epsilon \end{cases} \)
It is easy to see that the solution \((\theta^*_2, x)\) is the same as that of the equations in the Lemma. Moreover, because \(R\) is continuous in the first argument and strictly increasing in the second, for small \(\varepsilon\) the solution doesn’t change much if we replace the fixed \(\theta^*_2\) in the first argument of \(R\) by the variable \(\theta_2\). (Note that the effect of \(\theta^*_2\) on \(R\) through the second argument becomes arbitrarily large as \(\varepsilon\) shrinks.) Now, substituting \(\theta^*_2 + x\varepsilon\) by \(\theta^*_{2,r}\), we obtain the exact equations that define \(\theta^*_{2,r}\) and \(\theta^*_{2,nr}\). QED.

**Proof of Theorem 1**

Denote the two groups of agents corresponding to \(n_1\) and \(\theta_1\) by “rich” and “poor”, where the rich are those with higher wealth from their country 1 operations. It is easy to see that the threshold \(\theta^*_{2,rich}\) is below the threshold \(\theta^*_{2,poor}\). If it were above, then a rich agent observing \(\theta^*_{2,poor}\) would strictly prefer to wait. This is because she has the same belief over the distribution of the number of agents who run as a poor agent who would have observed that signal. Since the poor agent is indifferent at that signal and because of decreasing absolute risk aversion, the richer agent must strictly prefer the risky prospect.

Now consider the distribution of wealth corresponding to \(n'_1\) and \(\theta'_1\). Since it dominates the distribution corresponding to \(n_1\) and \(\theta_1\), it must be that the group of rich agents has grown or that the wealth of some group has increased (or both). To show that \(\theta^*_{2,rich}\) and \(\theta^*_{2,poor}\) must have decreased, we will eliminate all other possibilities.

Assume first that \(\theta^*_{2,rich}\) has increased. Since the wealth of rich agents has not decreased, and since they are now indifferent at a higher signal, it must be that at the new threshold their belief over the number of agents who run in country 2 (in the new equilibrium) is above that corresponding to the old threshold (and old equilibrium). However, since the size of the rich group has not decreased, it must be that they believe that a higher proportion of the poor group are now running. This necessitates that \(\theta^*_{2,poor}\) has increased by even more than the increase in \(\theta^*_{2,rich}\). (Note that the distribution of the proportion of rich
agents who run is unchanged.) But now consider a poor agent at the new threshold. She observes a higher signal than before, her wealth has not decreased and her belief over the number of agents who run in country 2 has become lower – both because the size of the rich group has not decreased and because \( \theta^*_{2,\text{rich}} \) has increased by less than \( \theta^*_{2,\text{poor}} \) (note that her distribution over the proportion of poor agents who run is unchanged). Thus, she must now strictly prefer to wait, in contradiction to the fact that she must be indifferent at her threshold.

A symmetric argument shows that \( \theta^*_{2,\text{poor}} \) must not have increased either. Now assume that both thresholds have not changed. If the wealth of agents from one group had increased, they would have a higher incentive to wait. The incentive would move in the same direction if more agents belonged to the rich group which has a lower threshold. Thus, at least one group is no longer indifferent.

To conclude the proof we claim that if one threshold had decreased, then the second would also have decreased. To see why, note that in such a case an agent from the other group who observes her old threshold would have a lower distribution over the number of agents who run. A change in the size of the groups will only contribute to the decrease in the distribution. Since the wealth of her group has not decreased, the only way she can remain indifferent is if her signal is lower. QED.

**Proof of Theorem 2**

Let \( n_i^\varepsilon(\theta, \tilde{\theta}) \) denote the proportion of agents who run in country 1 as a function of \( \theta \), given that each agent runs in that country if she observes a signal below \( \tilde{\theta} \) and does not run above it (the index \( \varepsilon \) appears so as to make the dependence explicit). Let \( \Delta_i^\varepsilon(\theta, \tilde{\theta}) \) denote the difference between the utility that an agent expects to attain in the case that she keeps her investment in country 1 until it matures and the utility she expects to attain if she withdraws early, when she observes the signal \( \theta \) and has the belief \( n_i^\varepsilon(\theta, \tilde{\theta}) \). A
threshold equilibrium then exists in country 1 if there is some $\tilde{\theta}_i$, such that $
abla_i'\left(\tilde{\theta}_i, \tilde{\theta}_i\right) = 0$ and $\nabla_i'\left(\theta_i', \tilde{\theta}_i\right) < (>) 0$ for any $\theta_i' < (>) \tilde{\theta}_i$.

Consider the expression for $\nabla_i'\left(\tilde{\theta}_i, \tilde{\theta}_i\right)$:

$$
\nabla_i'(\tilde{\theta}_i, \tilde{\theta}_i) = \frac{1}{2\varepsilon} \int_{\theta_i = \tilde{\theta}_i - \varepsilon}^{\tilde{\theta}_i + \varepsilon} \int_{\theta_i = \tilde{\theta}_i - \varepsilon}^{\tilde{\theta}_i + \varepsilon} \left[ u(R(\theta_i, n_i^c(\theta_i, \tilde{\theta}_i)) + w_{2,ar}^c(\theta_i, n_i^c(\theta_i, \tilde{\theta}_i); \theta_i, \theta_i')) - u(1 + w_{2,ar}^c(\theta_i, n_i^c(\theta_i, \tilde{\theta}_i); \theta_i, \theta_i')) \right] d\theta_i' d\theta_i d\theta_i
$$

where $w_{2,ar}^c$ and $w_{2,r}^c$ denote the returns in country 2 in the case where the realization in country 1 is $(\theta_1, n_i^c(\theta_1, \tilde{\theta}_1))$, the realization in country 2 is $\theta_2$ and the agent observed the signal $\theta_1'$. By Lemma 2, $w_{2,ar}^c$ and $w_{2,r}^c$ are the same for all $\theta_2$, except for an interval with measure no more than $2\varepsilon$. Thus, for small enough $\varepsilon$, this expression must be positive when $\tilde{\theta}_i$ is (deep enough) in the upper dominance region of the fundamentals and negative when $\tilde{\theta}_i$ is in the lower dominance region of the fundamentals. Finally, $\nabla_i'(\tilde{\theta}_i, \tilde{\theta}_i)$ is continuous in $\tilde{\theta}_i$ since a small change in $\tilde{\theta}_i$ only slightly shifts the interval over which $\theta_i$ ranges, and since the integrand is continuous in $\tilde{\theta}_i$ and bounded (note that by Lemma 3-continuity, a small change in $n_i^c(\theta_i, \tilde{\theta}_i)$ leads to a small change in the threshold signals of country 2). This shows that there exists some $\tilde{\theta}_i$ at which $\nabla_i'(\tilde{\theta}_i, \tilde{\theta}_i) = 0$.

Assume now that $\theta_i'$ satisfies $\nabla_i'(\theta_i', \theta_i') = 0$ and assume that $\theta_i' < \theta_i^*$. We will show that $\nabla_i'(\theta_i', \theta_i') < 0$. (The proof that $\nabla_i'(\theta_i', \theta_i') > 0$ for $\theta_i' > \theta_i^*$ is analogous.) Denote $c = [\theta_i^- - \varepsilon, \theta_i^+ + \varepsilon] \cap [\theta_i' - \varepsilon, \theta_i' + \varepsilon]$ and $d' = [\theta_i' - \varepsilon, \theta_i' + \varepsilon] \setminus c$. Then

$$
\nabla_i'(\theta_i', \theta_i') = \frac{1}{2\varepsilon} \int_{\theta_i = c}^{- \varepsilon} \int_{\theta_i = c}^{\varepsilon} \left[ u(R(\theta_i, n_i^c(\theta_i, \theta_i')) + w_{2,ar}^c(\theta_i, n_i^c(\theta_i, \theta_i'); \theta_i, \theta_i')) - u(1 + w_{2,ar}^c(\theta_i, n_i^c(\theta_i, \theta_i'); \theta_i, \theta_i')) \right] d\theta_i' d\theta_i d\theta_i
$$

and

$$
\nabla_i'(\theta_i', \theta_i') = \frac{1}{2\varepsilon} \int_{\theta_i = d'}^{- \varepsilon} \int_{\theta_i = d'}^{\varepsilon} \left[ u(R(\theta_i, 1) + w_{2,ar}^c(\theta_i, 1; \theta_i, \theta_i')) - u(1 + w_{2,ar}^c(\theta_i, 1; \theta_i, \theta_i')) \right] d\theta_i' d\theta_i d\theta_i
$$
Since $\Delta_i(\theta_i^*, \theta_i^*) = 0$, we know that $\theta_i^* - \varepsilon$ must be below the upper dominance region. This implies that for all $\theta_i \in d^*$, $R(\theta_i, 1) < 1$, and thus, for small enough $\varepsilon$, the second component must be negative. To see why the first component is negative, consider the value of $R$ at the highest point in $c$: $R(\theta_i + \varepsilon, n_i^* \theta_i + \varepsilon, \theta_i^*)$. If it is less or equal than 1 then $R$ must be less than 1 at any point in $c$. Since the derivatives of $R$ are bounded away from 0, then for small enough $\varepsilon$ the effect of the difference between $w_{z, nr}^\varepsilon$ and $w_{z, r}^\varepsilon$ is negligible and the integrand is negative for all $\theta_i \in c$, implying that the first component is negative. Similarly, if $R(\theta_i + \varepsilon, n_i^* \theta_i + \varepsilon, \theta_i^*) > 1$, then for small enough $\varepsilon$ we must have

$$I(\theta_i) = \int_{\theta_i = 0}^{\theta_i + \varepsilon} \int_{\theta_i' = \theta_i}^{\theta_i' + \varepsilon} \left[ u(R(\theta_i, n_i^* (\theta_i, \theta_i^*))) + w_{z, nr}^\varepsilon (\theta_i, n_i^* (\theta_i, \theta_i^*); \theta_2, \theta_2^*) \right] d\theta_i' d\theta_2 > 0$$

for all $\theta_i \in d^* = [\theta_i^* - \varepsilon, \theta_i^* + \varepsilon] \setminus c$. This implies that $\int_d I(\theta_i) > 0$. Now, since $\Delta_i(\theta_i^*, \theta_i^*) = \int_c I(\theta_i) + \int_{\theta_i} I(\theta_i) = 0$, we must have that our first component, $\int_c I(\theta_i)$, is negative. QED.
References:


