The Dissertation Committee for Jennifer Lynn Steele

certifies that this is the approved version of the following dissertation:

Essays on Incomplete Information in Economic Development

Committee:

Maxwell Stinchcombe, Supervisor

Thomas Wiseman

Russell Cooper

Kim Ruhl

Stephen Magee

Essays on Incomplete Information in Economic Development

by

Jennifer Lynn Steele, MS, BComm

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2007

For my family and friends

Acknowledgments

From the idea conception to this 'final' draft, I am extremely grateful for the help and guidance of Max Stinchcombe and Tom Wiseman. I would also like to thank those who supported me throughout my studies, especially my parents without whom I would not have undertaken graduate studies.

JENNIFER LYNN STEELE

The University of Texas at Austin

May 2007

Essays on Incomplete Information in Economic Development

Publication No. ____________________

Jennifer Lynn Steele, Ph.D.
The University of Texas at Austin, 2007

Supervisor: Maxwell Stinchcombe

This thesis studies the effects of incomplete information on economic development. Relaxing the assumption that information is complete allows for corruption to occur, even in equilibrium, and for poverty traps to develop. The first paper looks at how the lack of enforcement mechanisms affects contracts, and how a more efficient contracting mechanism can be developed in aid settings. I find that as the level of corruption increases, the contract will encompass more stages. In the second paper, the agent's level of corruption is unknown, and the principal may screen agents by inducing corruption with positive probability. This would account for the corruption seen in development projects as an equilibrium effect. The third paper looks at the effect of uncertainty about foreign productivity on a firm's foreign direct investment

(FDI) decision. Dependent on the form of the information, this may result in either an underinvestment of FDI, or no FDI at all.

Contents

Chapter 1

The Optimal Sequencing of Carrots: Project Financing with Limited Enforcement

Abstract

I consider a dynamic principal-agent problem where the principal designs a multi-stage contract to induce the agent to undertake a mutually beneficial public project. The benefit in each stage depends on the cumulative contributions so far. At each stage, the agent must choose between applying the funds to the project and diverting them. As the agent's utility from diversion increases, the optimal contract requires contributions in more stages and the contribution in each stage decreases. When the project takes more than one

stage there will be overinvestment.

In many situations the principal's ability to punish deviation does not allow for efficient outcomes in one-stage contracts. As a result, multi-stage contracts, in which future allocations provide incentives for honesty, can increase efficiency and the principal's welfare.

If the principal is able to commit to a sequence of allocations, overinvestment may occur, where the allocation in the final stage discourages deviation in previous stages in addition to increasing the benefit from the project. The principal's ability to punish is constrained to two instruments; a reputation cost suffered by the agent if he deviates, and the termination of the contract upon deviation.

When projects are funded solely by one party it is often assumed that there is to be some repayment.[1] In this model the project can be thought of as a public good, where principal and agent both receive positive utility from applying funds to the project and the principal does not require repayment to undertake the project. Requiring repayment would lower the agent's utility, and therefore his incentives, as well as requiring additional enforcement mechanisms.

Papers concerned with public good projects generally involve multiple parties contributing funds (see for example Admati and Perry (1991), Marx and Matthews (2000) and Varian (1994)). The authors look at projects that rely on contributions

[1] For examples see Atkeson (1991) and Spear and Strivastava (1987)

by two parties with or without commitment, and analyze the strategic interaction between the parties. This paper is a principal-agent model applied to public projects, where the contribution of the principal is the resources for the project, while the agent decides on their use.

The most complete dynamic treatment of this public goods problem is in Marx and Matthews (2000) who find that the project is only completed when the number of periods is large, period length is small, or when players have similar evaluation of the project. By contrast, in my model, there are fewer restrictions on the benefit function and one side of the transaction, the donor, can commit. As a result the project is either not undertaken, or if undertaken, there is some overinvestment.

The benefit function is cumulative so the allocation in each stage adds to the benefit from the project. As a result each stage is a unique subgame, as opposed to a repeated game environment. Levin (2003) looks at a repeated game environment with relational incentive contracts, in which there is self-enforcement when external enforcement mechanisms are insufficient. He finds that the limitation can be characterized in additional bounds on compensation. Similarly, in this model, the lack of enforcement mechanisms limits the allocation in each stage. However, the ability of the principal to commit fully to the contract increases the efficiency of the contract.

Malcomson and Spinnewyn (1988) state that the advantage of the principal being able to commit is it “allows a contract to tie the principal to an expected loss [in the future].” They find that long-term contracts cannot provide Pareto improvements on short-term contracts. Here the principal’s commitment to a long term contract increases efficiency due to enforceability problems, and because funds are transferred before the agent chooses an action.

An important application of the model is found in development economics. Aid organizations that wish to fund large scale projects in developing countries must often contract with local officials to undertake them.

As debt forgiveness and renegotiation are becoming more common and more countries are becoming overly burdened with debt, many in the aid community are turning to grants as a more efficient option. Bulow and Rogoff (2005) argue that multilateral development banks should switch from administering loans to administering grants, reasoning that the transparency of grants could improve efficiency and decrease renegotiation. In June 2002 donors reached an agreement that 18-21% of IDA (a branch of the World Bank) aid should be in the form of grants. Currently donors are finalizing an agreement which states that 30% of disbursements over the next three years (approximately $11B) should be in the form of grants (Sanford (2004)). This comes amid attempts by donor countries to make aid more dependent on outcomes.

Breaking a contract into multiple stages allows the principal to make future allocations contingent on the agent’s history. Adam and Gunning (2002) discuss the benefits of ex-post conditionality in aid contracts, where the focus is on project outcomes rather than inputs.

The main question I address is how multi-stage contracts can provide incentives for honesty when the agent derives positive utility from diversion. Incentives can be a powerful tool for increasing project efficiency when there is no mechanism for contract enforcement. However, the principal's use of incentives may result in overinvestment in the project, as the marginal benefit of the last dollar includes both the benefit from the project, and the effect on the agent's incentives in previous stages.

When the agent's benefit from applying the funds to the project is cumulative, I find that the allocations are increasing over time, up to some finite stage. Watson (2002) finds that incomplete information can lead to 'starting small', in which levels of interaction between partners can grow over time as they overcome the informational asymmetries. In this paper 'starting small' occurs because the benefit from the project is a result of cumulative allocations, and is increasing over time, so the opportunity cost of diverting funds is increasing over time.

Section 3.1 gives an overview of the model. In Section 3.3, I characterize the optimal contract and the effects of increasing marginal utilities of diversion and reputation costs. Section 3.5 is the conclusion.

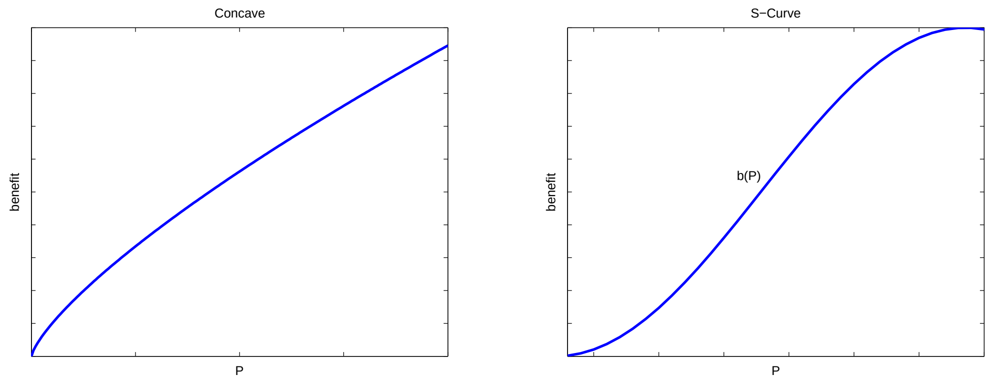


Figure 1.1: Benefit Functions

1.1 The Model

The principal is designing a contract to fund a project that the agent is undertaking. Both the agent and the principal receive positive benefits from the project. Their utilities are strictly monotonic with respect to the benefits produced by the project and are a function of the cumulative contributions to the project. The total funding level in stage t is $P_t = \sum_{s \leq t} p_s$ when all contributions in stages $s \leq t$ were applied to the project.

When applied to aid contracts, it is assumed that the principal is acting altruistically, funding the project for the benefit of the agent (or the country in which the agent resides). More generally when the contract is for the completion of a public project, it is assumed that the agent is benefitting from the public project in the same manner as other beneficiaries. At the least the agent is assumed to receive a small cumulative benefit, a 'warm glow' that increases with the total funding level.

The principal's benefit derived from the project in stage t is $b(P_t)$ and the agent's benefit from the project in stage t is $h(P_t)$. Both benefit functions are assumed to take the same general shape, examples of which are shown in Figure 1.1.

Assumption 1.1.1. *Benefit Function:* $g = \{h, b\}$

a. *g(P) is weakly increasing,* $\lim_{P \to \infty} g'(P) = 0$

b. *One of the following two shapes:*

Concave: $g''(P) \leq 0$ *for all* $P > 0$

s-shaped: $g''(P) \geq 0$ *for all* $P < c$ *and* $b''(P) \leq 0$ *for all* $P > c$

c. $\frac{b(P)}{1-\delta} > P$ *for some* $P > 0$

Given that the contract is for the financing of a defined project, the marginal benefit is weakly increasing for all funding levels. However, at some point additional resources have little effect and the marginal benefit approaches zero.

A concave benefit function realizes its potential immediately, and the marginal benefit is decreasing thereafter while an s-shaped benefit function represents a project where there are fixed costs to overcome before a substantial benefit is realized.

The principal's and agent's benefit functions have the same inflection point when s-shaped because it is a feature of the underlying project rather than the magnitude of their benefit from the project.

The principal suffers a cost of capital $\delta^t(p_t + s_t)$ in stage t therefore item $c.$ assumes that the project is worth undertaking when there are no constraints.

At the beginning of the game the principal outlines a contract, C, represented by a sequence of funding levels, p_t, and side payments, s_t.

$$C = \{p_t, s_t\}_{t=1}^{\infty}$$

She commits to giving the agent $p_t + s_t$ in each stage t provided the agent applies p_k to the project (is honest) in every stage $k < t$. Ex-ante commitment is optimal, and the principal is better off when commitment is credible, so it is implicitly assumed that the principal's cost of deviating is high enough to prevent the principal from ever deviating.

In stage t the agent receives $\{p_t, s_t\}$ and must decide between applying p_t to the project and diverting the funds. s_t is a side payment, and is not applied to the project. If he chooses to divert the project funds, p_t, the game ends. His benefit from applying the funds to the project is $h(P_t)+d(s_t)$ while his payoff from diversion is $d(p_t + s_t) - Z$. The reputation cost of diversion is Z and allows the enforceability of contracts and reputation costs to be built into the model. The more capacity the principal has to punish the agent, the higher Z will be.

Assumption 1.1.2. *The agent's utility from diversion $d(\cdot)$ is weakly concave, where $d'(\cdot) > 0$ and $d''(\cdot) \leq 0$ and $d(0) = 0$.*

In the next section I outline the optimal contract and examine the effects of changes in the parameters.

1.2 Optimal Contract

A contract inducing diversion in stage t with $p_t > 0$, $s_t > 0$ gives the principal the same payoff as a contract where $p_t = 0$ and $\tilde{s_t} = p_t + s_t$. However the agent strictly prefers the latter contract because they would not suffer reputation costs, and the principal would be better off offering $(0, \tilde{s_t} - \varepsilon)$ in stage t. Therefore the optimal contract will never involve the principal inducing diversion and can be written as follows:

$$\max_{\{p_t, s_t\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \delta^t [b(\sum_{n \leq t} p_n) - p_t - s_t] \tag{1.1}$$

$$\text{s.t.} \sum_{n=t}^{\infty} \delta^n [h(P_n) + d(s_n)] - \delta^t [d(p_t + s_t) - Z] \geq 0 \text{ for all } t \tag{1.2}$$

The principal is maximizing her benefit from the project less the cost of capital subject to the agent's incentive compatibility constraints.

Due to the cumulative nature of the benefit function, the principal wants to complete the project as quickly as possible, which is seen in two results. First, the

ICC will bind in every stage $t < T$.[2] Second there will be overinvestment in the project as the principal uses 'carrots' in the final stage to induce honesty from the agent in all previous stages.

As the agent's utility of diversion, $d(\cdot)$, increases, the number of stages with positive allocations in the optimal contract weakly increases, and the total benefit to the principal decreases. If the project is undertaken it will be completed but the incentive effects may cause overinvestment. Finally, the effect of reputation costs, Z, will be outlined. As Z increases, the total benefit to the principal increases, and more projects will be undertaken.

The best the principal can do is the allocation where she is not constrained by the agent's incentive compatibility constraint in any stage.

Definition 1.2.1. *The* ***unconstrained optimal funding level*** *, P°, is the funding level at which an efficient level of funds are invested in the project.*

Whether dealing with s-shaped or concave benefit functions, the unconstrained optimal funding level is where:

$$\frac{b'(P^{\circ})}{1-\delta} = 1$$

However, in the case of an s-shaped benefit function, the unconstrained op-

[2] Except in the case of some s-shaped curves, where the ICC will not bind in the first stage, see 1.5 for details

timal funding level is the allocation to the right of the inflection point, c, where the above equation holds.

Any funding level beyond the unconstrained optimal involves investment where the marginal benefit from the project is less than the marginal cost, and is therefore wasteful.

The *unconstrained optimal contract* is the contract where the allocation in the first stage is the unconstrained optimal funding level, $p_1 = P^{\circ}$, and all subsequent allocations are zero, i.e. the sequence of allocations is $\{P^{\circ}, 0, 0, ...\}$.

Definition 1.2.2. *An increase (decrease) in an agent's utility of diversion is a weak increase (decrease) in the slope of* $d(n)$ *such that* $d'_{\circ}(n) \leq d'_1(n)$ *(*$d'_{\circ}(n) \geq d'_1(n)$*) for all* n*.*

Proposition 1.2.3. *The* ***Optimal Contract*** *is a sequence of allocations,* $\{p_t, s_t\}_{t=1}^{\infty}$ *that has the following characteristics:*

a. *There is some finite stage,* T*, after which all allocations are zero*

b. *The agent is honest in every stage*

c. *As* d' *increases* T *increases*

    If d' *is sufficiently low* $T = 1$

    If d' *is sufficiently high* $s_T > 0$

d. *When the optimal contract takes more than one stage there will be overinvestment in the project*

When the agent is sufficiently honest the principal can complete the contract in one stage, and achieve their first-best solution. However, as the agent's marginal utility of diversion increases, they will optimally divert the funds if given the first-best contract. As a result, the principal will break the contract down into multiple stages.

Proof. *a. There is some finite stage, T, after which all allocations are zero.*

The proof is broken down into two steps. First, I show that if the ICC binds in stage t, the allocation in $t+1$ must be positive, and vice versa. Second, due to the marginal benefit from the project approaching zero in the limit, the ICC can not bind ad infinitum.

i. If the ICC binds in stage t, the allocation in $t+1$ is strictly positive.

The FOC for p_t is:

$$\sum_{s=t}^{\infty} \delta^s b'(P_s) - \delta^t + \sum_{n=1}^{t} \lambda_n \sum_{s=t}^{\infty} \delta^s h'(P_s) - \lambda_t \delta^t d'(p_t) + \sum_{n=t+1}^{\infty} \lambda_n \sum_{s=n}^{\infty} \delta^s h'(P_s) = 0 \tag{1.3}$$

and the FOC for p_{t+1} is:

$$\sum_{s=t+1}^{\infty} \delta^s b'(P_s) - \delta^{t+1} + \sum_{n=1}^{t} \lambda_n \sum_{s=t+1}^{\infty} \delta^s h'(P_s) - \lambda_{t+1}\delta^{t+1} d'(p_{t+1}) + \sum_{n=t+1}^{\infty} \lambda_n \sum_{s=n}^{\infty} \delta^s h'(P_s) = 0 \tag{1.4}$$

When the marginal benefit from the project is decreasing over time (as is the case with a concave benefit function), and $p_{t+1} = 0$ while the ICC binds in t, (1.4) would be strictly positive. Therefore if the ICC is binding in stage t, the allocation in $t+1$ must be positive.

Likewise, if the ICC is not binding in stage t ($\lambda_t = 0$), then the FOC for p_{t+1} is equal to zero when $p_{t+1} = 0$.

This shows the principal's preference for allocating funds earlier. She will only allocate funds in stage $t+1$ if the agent's ICC were binding in stage t.

The only exception to this is when the benefit is s-shaped, and the size of the allocation in the first stage does not affect the marginal benefit. In this case it may be optimal to allocate funds in stage one such that the ICC does not bind, a case outlined in 1.5.

ii) The ICC cannot bind ad infinitum

Above we showed that the ICC in stage $t > 1$ must bind if the allocation in $t+1$ is positive. Therefore, if there is no finite stage after which all allocations are zero, the ICC must bind ad infinitum.

The allocation p_t at which the ICC binds is strictly positive:

$$d(p_t + s_t) = \sum_{n=t}^{\infty} \delta^{n-t}[h(P_n) + d(s_n)] + Z \tag{1.5}$$

Therefore as the number of stages increases the total funding level (P_t) strictly increases. As the total funding level becomes large, the marginal benefit from the project (to both the principal and the agent) decreases, and the FOC for p_t becomes:

$$\sum_{s=t}^{\infty} \delta^{s-t} b'(P_s) + \sum_{n=1}^{t} \lambda_n \sum_{s=t}^{\infty} \delta^{s-t} h'(P_s) - \lambda_t d'(p_t) + \sum_{n=t+1}^{\infty} \lambda_n \sum_{s=n}^{\infty} \delta^{s-t} h'(P_s) = 1 \tag{1.6}$$

$b'(P_t)$ and $h'(P_t)$ are both approaching zero as P_t grows large. Therefore there must be some P_t above which equation (1.6) would not hold and the ICC will not bind.

With an s-shaped curve, the project will never optimally end to the left of the inflection point, c, because the marginal benefit is increasing up to c. As a result, the total funding level for the project will occur in the concave section of the benefit function, and the marginal benefit will be decreasing.

As a result the contract consists of the ICC binding in every stage $t \in [2, T-1]$

and some stage T with a positive allocation where the ICC does not bind and all future allocations are zero.

b. The agent is honest in every stage

Outlined in the introduction to this section.

c. As d' increases T increases

If d' is sufficiently low $T = 1$

If d' is sufficiently high $s_T > 0$

The proof will be broken into four steps. First, I define the utility of diversion below which the unconstrained optimal contract can be obtained. Second, as the agent's utility of diversion increases, the number of stages in the contract increases. Third, for sufficiently high utilities of diversion a positive side payment in stage T is optimal. Finally, as the agent's utility of diversion continues to increase, side payments in earlier stages become optimal under certain conditions.

i) If $d(P^\circ) < \frac{h(P^\circ)}{1-\delta} + Z$ the principal will offer the unconstrained optimal contract, and the agent will not divert funds.

If feasible, the unconstrained optimal contract is the best the principal can do. This allows her to fund the project up to the point where the marginal benefit from the project equals the marginal cost in the first stage, maximizing her benefit. However, if x, the utility of alternative uses of grant money, is sufficiently high, the unconstrained optimal contract may induce diversion in the first stage. If the ICC is met with equality at P° in the first stage the principal can obtain the unconstrained optimal contract:

$$d(P^{\circ}) = \frac{h(P^{\circ})}{1-\delta} + Z \tag{1.7}$$

For any decrease in the utility of diversion the ICC will not bind and the unconstrained optimal contract induces honesty. At the unconstrained optimal allocation, P°, the marginal benefit from the project equals the marginal cost. If it is incentive compatible to offer the unconstrained optimal allocation in the first stage, the principal cannot be made better off.

b) If $d(P^{\circ}) > \frac{h(P^{\circ})}{1-\delta} + Z$ the unconstrained optimal contract is not incentive compatible and the optimal contract, if non-trivial, will involve multiple stages with positive allocations.

The principal is strictly better off allocating funds in stage one subject to the constraint that the agent will not divert funds. The most they can offer in the first stage is the allocation at which the ICC binds and is not met for any $p > p_1$.

$$d(p_1) = \sum_{s=1}^{\infty} \delta^{s-1} h(P_s) + Z \tag{1.8}$$

$$d(p) - Z > \sum_{s=1}^{\infty} \delta^{s-1} h(P_s) \qquad \text{for all } p > p_1 \tag{1.9}$$

Corollary 1.2.4. *Let the set of utilities of diversion for which the above equations hold be defined as* $D^\circ = \{d^\circ(\cdot) :$ *equations* (1.8) *and* (1.9) *hold*$\}$. *The optimal contract will consist of more than one stage for any agent whose utility of diversion is an increase from some* $d^\circ(\cdot) \in D^\circ$.

From *a.* we know that if the ICC binds in stage one, the allocation in stage two will be positive.

If $d(\cdot)$, the utility of diverting funds, is sufficiently low, the contract will end in two stages:

$$\frac{b'(P_2)}{1-\delta} - 1 + \lambda_1 \left[\frac{h'(P_2)}{1-\delta}\right] = 0 \tag{1.10}$$

If P_2 cannot be achieved in two stages without inducing the agent to divert funds, the contract will take more than two stages to complete.

I now show that as the agent's utility of diversion increases the number of stages in the optimal contract is increasing, and the principal's benefit is decreasing.

In every stage $t < T$ the ICC binds and

$$d(p_t) = \sum_{s=t}^{\infty} \delta^{s-t} h(P_s) + Z \tag{1.11}$$

Thus the allocation in every stage $t < T$ is dependent on the allocation in stage T. Using the FOC for p_T:

$$\frac{\delta^T b'(P_T)}{1-\delta} + \sum_{t=1}^{T-1} \lambda_s \delta^s h'(P_s) = \delta^T \tag{1.12}$$

the FOCs for p_t for all stages $t < T$

$$\sum_{s=t}^{\infty} \delta^s b'(P_s) + \sum_{s=1}^{t} \lambda_s \sum_{n=t}^{\infty} \delta^n h'(P_n) - \lambda_t \delta^t d'(p_t) + \sum_{s=t+1}^{j-1} \lambda_s \sum_{n=s}^{\infty} \delta^n h'(P_n) = \delta^t \tag{1.13}$$

and the ICC's for all stages $t < T$, by defining p_T allocations p_t can be defined recursively.

As the agent's utility of diversion increases the allocation in every stage $t < T$ decreases, and the principal's benefit in every stage $t < T$ decreases. In addition, because P_t is decreasing in $d(\cdot)$, the contract will take weakly more stages to complete.

iii) If the agent has a sufficiently high utility of diversion, the principal will optimally give the agent a positive side payment in stage T.

A positive side payment in stage T is optimal when the gain from higher allocations in each stage $t < T$ outweighs the cost of capital in stage T. The allocation in stage T can be thought of as a 'bonus' for the agent, a reward for honesty in all stages $t \leq T$.

The FOCs for s_t and s_T are:

$$\sum_{n=1}^{t} \lambda_n d'(s_t) - \lambda_t d'(p_t + s_t) \leq 1 \qquad \text{for all } t < T \tag{1.14}$$

$$\sum_{n=1}^{t} \lambda_n d'(s_T) \leq 1 \tag{1.15}$$

For a sufficiently dishonest agent the principal will optimally allocate a positive side payment in the last stage. If the agent's utility of diversion is strictly concave, as the agent's utility of diversion increases the principal will allocate side payments in weakly earlier stages. The expected benefit from the contract is decreasing in $d(\cdot)$ regardless of whether the principal allocates positive side payments. However, when the principal does allocate positive side payments, an increase in $d(\cdot)$ causes a smaller decline in benefit than a contract with no side payments.

Corollary 1.2.5. *Let the set of utilities of diversion such that* (1.15) *holds with equality when* $s_T = 0$ *be defined by* $D^c = \{d^c(\cdot)$*: equation* (1.15) *holds with equality when* $s_T = 0\}$. *The principal will optimally give a positive side payment to any agent with utility of diversion* $d(\cdot)$ *such that* $d(\cdot)$ *is an increase from some* $d^c(\cdot) \in D^c$

When (1.15) holds with equality, any increase in $d(\cdot)$ causes the LHS to increase, and as a result s_T must be positive. If the agent's utility of diversion becomes more concave ($d''(\cdot)$ increases) the principal may choose to give side payments in earlier stages.

d. When the optimal contract takes more than one stage there will be overinvestment in the project

For every $d(\cdot)$ where the optimal contract is constrained by the agent's incentive compatibility constraint, the total investment in the project is greater than the unconstrained optimal funding level.

Provided $T > 1$,[3] in period T the incentive effects are positive:

$$\frac{b'(P_T)}{1-\delta} + \sum_{n=1}^{T-1} \lambda_n \delta^{n-T} h'(P_n) = 1 \tag{1.16}$$

The marginal benefit from the project in stage T is less than the marginal cost due to the positive incentive effects. Because the marginal benefit is decreasing in stage T, the total investment in the project must be greater than in the unconstrained optimal contract, where there are no incentive effects.

Any funding level, P_T in excess of the unconstrained optimal funding level represents an overinvestment in the project as the benefit from the project is less than the cost for all funding beyond P°. □

If the agent's utility of diversion is sufficiently low the unconstrained optimal contract is incentive compatible, and the principal receives the maximum benefit from the project.

As $d(\cdot)$ increases the number of stages required to complete the contract

[3]With an s-shaped benefit function this theorem holds when the ICC binds in $T-1$

increases as the allocation in each stage decreases. The benefit to the principal decreases as it takes longer to realize the benefits from the project.

Eventually $d(\cdot)$ increases to the point where the principal will allocate a positive side payment. When $d(\cdot)$ is large the allocations in each stage are sufficiently small that she will offer a 'bonus' that allows her to allocate more funds earlier.

Proposition 1.2.6. *With an s-shaped curve, $b(\cdot)$ where $b'(0) < 1$, there exists a set of utilities of diversion, $D^s = \{d^s(\cdot)\}$, where the principal's expected return is zero. For any utility of dishonesty that is an increase over some $d^s(\cdot) \in D^s$ the principal will not offer the agent a contract with positive funding levels.*

Proof. When there are no constraints the principal will undertake the project (from Assumption 1.1.1).

From Equation (1.5) when the ICC binds p_t is decreasing in $d(\cdot)$, and for a sufficiently large $d(\cdot)$ the payoffs are negative for the principal in the beginning until P_t approaches c, and the marginal benefit from the project is greater than one.[4] $b(P_t) < p_t$ for P_t sufficiently small.

In every stage $t \in [2, j-1]$, the ICC binds and p_t is decreasing in the agent's utility of diversion. Then we can always find a sufficiently high $d(\cdot)$ where $b(P_2) < p_2$, and the principal's benefit in the first two stages is less than her cost. As we continue to increase $d(\cdot)$, the number of stages with a negative return increases

[4] c is the inflection point for the s-shaped curve, and can be thought of as the point at which all fixed costs are overcome

as the allocation in each stage decreases. We can then find a utility of diversion, $d^s(\cdot)$ such that for a given $\delta < 1$ the total payoff from the project for the principal is zero for any stream of positive funding levels.

$$\sum_{t=1}^{\infty} \delta^t [b(P_t) - p_t - s_t] = 0 \tag{1.17}$$

For any $d(\cdot) > d^s(\cdot)$ the payoff the principal gets from the project is negative and the principal will offer a trivial contract of zero in each stage. □

With both a concave and an s-shaped benefit function the allocation in each stage where the ICC binds is increasing over time. Typically 'starting small' is explained through uncertainty in the benefit function or in relationships. Here an alternate hypothesis is presented. The cumulative nature of the allocations applied to a project allow for a stream of increasing allocations in the optimal contract.

Proposition 1.2.7. *As Z, the ability to punish or reputation cost, increases, the allocation in each stage $t < T$ increases. The benefit to the principal increases and the number of stages in the optimal contract weakly decreases.*

Proof. From the agent's ICC (equation (1.2)) as Z increases, the allocation where the ICC binds increases in each stage. As a result the principal gets a higher benefit in each stage. P_t, the total funds allocated to the project as of stage t, also increases, and as a result the project will have weakly fewer stages with positive allocations. □

The marginal effect of Z on p_t is d^{-1}, meaning that the more dishonest the agent is the smaller the effect of any punishment or reputation cost.

In a country with strong institutions for contract enforcement Z may be sufficiently high to achieve efficient contract completion. However, many international aid settings lack strong institutions, requiring much longer completion times for some desirable projects, and making others infeasible. Alternatively, if punishing the agent is costly (legal fees, time required, etc.), small projects may not warrant full punishment, and Z may be low.

1.3 Conclusion

When the principal and agent share similar benefit functions from a project, the question of how to finance the project involves how best to induce the agents to use the funds for the project. The difference in the relative values that the principal and agent place on the alternate uses of the funds can prevent the efficient outcome from being obtained.

Due to the cumulative nature of the allocations, and the incentive compatibility constraint binding in every stage up until the final stage, the allocations in all stages where the ICC binds are increasing over time. This 'starting small' effect is due to the cumulative nature of the benefit function. As the project grows, the cost of diversion (giving up the future benefit) increases, so the principal can increase the allocations as the project nears completion.

If the marginal utility of diversion is sufficiently high, the principal cannot achieve the unconstrained optimal funding level in the first stage. Instead, the principal would optimally extend the contract to multiple stages, resulting in over-investment in the project. If there were only one stage the principal could only fund the project up to the point where the agent would divert funds, resulting in under-investment. As the marginal utility of diversion of the agent increases the ability to split the project into multiple stages has more of an impact on the principal's payoff.

For sufficiently dishonest agents it will be optimal for the principal to give him a 'bonus' in the final stage that will not be applied to the project. If the agent's utility of diversion is sufficiently concave, the principal may also give side payments in earlier stages.

In the model monitoring is assumed to be costless. A model with endogenous stage lengths, which are determined by the cost of monitoring, could illustrate the tradeoff between shorter stages allowing the contract to be completed in a shorter period of time, and the costs of having more stages.

1.4 Supplement: Changes in the ICC

Agents continue receiving benefit from the project after diverting.

If once an agent diverts funds, the principal no longer offered him any allocations, but they were able to keep receiving the benefit from the project ad infinitum. The ICC would look as follows:

$$\sum_{n=t}^{\infty} \delta^{n-t}[h(P_n) + d(s_n)] \geq d(p_t + s_t) + \frac{h(P_{t-1})}{1-\delta} - Z$$

and our first order condition for p_t would change as the allocation in previous stages now increases the utility of diversion in the future.

$$\sum_{s=k}^{\infty} \delta^s b'(P_s) - \delta^k + \sum_{t=1}^{k} \lambda_t \sum_{s=k}^{\infty} \delta^s h'(P_s) - \lambda_k \delta^k d'(p_k + s_k) + \sum_{t=k+1}^{\infty} \lambda_t \left[\sum_{n=t}^{\infty} \delta^n h'(P_n) - \delta^n \frac{h'(P_{n-1})}{1-\delta} \right] \leq 0 \tag{1.18}$$

The boundary solution would change, as the ICC now binds at a lower level for all agents,

$$d(p_t) = \sum_{s=t}^{\infty} \delta^{s-t} h(P_s) - \frac{h(P_{t-1})}{1-\delta} + Z < \sum_{s=t}^{\infty} \delta^{s-t} h(P_s) + Z$$

As with an increase in the agent's utility of diversion, it will now require more stages to complete the contract, and the allocation in each stage will be lower. Because of both effects, the benefit from the contract for the principal is lower whenever the optimal contract involves more than one stage.

1.5 Supplement: s-shaped case

With an s-shaped benefit function the marginal benefit is weakly increasing until c is reached. As a result, the shape of the agent's utility function is important in determining the sequence of allocations. Allowing T to be the minimum number of stages the project can be completed in without violating the incentive compatibility constraints, the principal may reallocate funds from the first stage to stage T in order to reduce costs. In this case the contract would be 'backloaded' where the ICC would not bind in the first or the last stage.

The FOC for p_1 is as follows:

$$\sum_{t=1}^{\infty} \delta^t b'(P_t) + \sum_{t=1}^{j-1} \lambda_t \sum_{s=t}^{\infty} \delta^s h'(P_s) - \lambda_1 \delta d'(p_1) \leq \delta \tag{1.19}$$

When the marginal benefit in the second stage is larger than the marginal benefit in the first stage and when $b''(p_t) = 0$ for all equilibrium values of p_t, an increase in p_1 does not affect the marginal benefit of the project. Therefore, it may be optimal to offer an allocation in stage one such that the ICC does not bind, but the project can still be completed in T stages.[5]

In stage two the marginal benefit will increase, at the minimum because

[5]This would occur in a benefit function where the benefit is zero until c, at which point the project comes into operation and the benefit jumps, then constant up to the optimal funding level

they are one stage closer to completion of the project and the ICC for p_t in stages $t \in [2, T-1]$ must bind. If the ICC in stages one and two did not bind, but the allocations were positive, the FOCs would be:

$$\sum_{t=1}^{\infty} \delta^{t-1} b'(P_t) + \sum_{t=3}^{j-1} \lambda_t \sum_{s=t}^{\infty} \delta^{s-1} h'(P_s) = 1 \tag{1.20}$$

$$\sum_{t=2}^{\infty} \delta^{t-2} b'(P_t) + \sum_{t=3}^{j-1} \lambda_t \sum_{s=t}^{\infty} \delta^{s-2} h'(P_s) = 1 \tag{1.21}$$

The discounted marginal benefit is higher in stage two than stage one and the ICC must bind in stage two. Because every stage brings the project closer to c, increasing the marginal benefit, the ICC will bind in every stage $t > 1$ such that $P_t < c$.

For any stage t such that $P_t > c$ the marginal benefit is decreasing, and the remainder of the project has concave benefits, therefore, as shown above, the ICC will bind in every stage $t < T$.

In general, when the benefit function is s-shaped, the ICC will bind in every stage $t \in [2, T-1]$ (and it may bind in the first stage). With a concave benefit function, the ICC binds in every stage $t < T$.

When the marginal benefit plus future incentive effects are equal to the marginal cost in stage one, and an increase in p_1 has no effect on marginal benefit,

$$\sum_{t=1}^{\infty} \delta^{t-1} b'(P_t) + \sum_{t=2}^{j-1} \lambda_t \sum_{s=t}^{\infty} \delta^{s-1} h'(P_s) = 1 \quad \text{and} \quad \frac{\partial \sum_{t=1}^{\infty} \delta^t b'(P_t)}{\partial p_1} = 0 \tag{1.22}$$

the ICC will not bind in the first stage.

Chapter 2

Project Financing with Asymmetric Information: Inducing Diversion as a Means of Screening

Abstract

When undertaking a large project over multiple stages the benefit from the project is cumulative, with the application of funds in any given stage adding to the total benefit from the project. The principal must contract with an agent to undertake a project, but does not know

> the agent's type. Due to a lack of enforcement mechanisms, the agent chooses between applying the funds to the project and diverting them in each stage. When there is sufficient uncertainty about the agent's type, the principal may induce diversion with positive probability. I show that the principal will offer only one contract, with types separating by their choice of stage in which to divert.

Aid-granting organizations typically face two main issues which determine the efficiency of contracts when funding projects in developing countries. First, a lack of enforcement of property rights and contracts makes punishment difficult. If punishment is inadequate a multi-stage contract which provides incentives for the local authorities increases the organization's expected benefit.

Second, authorities have varying utilities of diversion, both between countries and within countries. This results in inefficiencies in the provision of public projects. When the agent's utility is not known, the aid-granting organization may find that the effectiveness of aid is decreased, as funds are diverted to other purposes with positive probability.

Hirschman (1967) studies eleven large projects undertaken by the World Bank, and analyzes the causes of the variable levels of success attained. He identifies different kinds of uncertainty that have an effect on the success of the project, in particular how uncertainty about the domestic administration of a project can derail it. He also discusses the use of sequential problem-solving to deal with this and other forms of uncertainty. The present paper presents a more thorough analysis of the ability of sequential contracts to minimize the effects of this kind of uncertainty.

In the model resources are transferred from the principal to the agent with the express purpose of funding a project which benefits both parties. This specification can apply to any sort of public project where enforcement mechanisms are weak and the agent has incentive to divert funds.

Asymmetric information games typically involve separating or pooling equilibria, with the principal offering a menu of contracts in order to separate types.[1] Because all types receive the same benefit from the project I show that the optimal contract is the same for all types, with separating occurring through their choice of which stage to divert funds in.

Ma (1991) looks at dynamic contracting in principal agent models where screening may occur in the second period, and where the contract may not be undertaken. He focusses on renegotiation proofness and allows the principal to screen using a set of incentive schemes.

Malcomson and Spinnewyn (1988) find that long-term contracts cannot provide Pareto improvements on short-term contracts. However, because of limited

[1]See Laffont and Tirole (1986) who look at optimal contracting with cost uncertainty, Sappington (1983) focusses on contracts with limited liability, but where types are realized after the contract is agreed upon while Harris and Townsend (1981) study optimal contracts with asymmetric information ex-ante.

liability and the fact that funds are transferred before the agent chooses an action, in this environment the principal is able to increase their expected return by committing to a long term contract. In Freixas, Guesnerie and Tirole (1985) they look at a central planning problem and find that commitment is optimal for the principal, for similar reasons as outlined in this paper.

The environment is similar to that in Steele (2007), but in this case there is asymmetric information with respect to the agent's type. With asymmetric information the principal can use multiple stages to learn about the agent's type. If the uncertainty is sufficiently high, the principal will screen in each stage, meaning that there is a positive probability the agent will divert funds in each stage.

In the optimal contract I find that the allocations are increasing over time, up to some finite stage. Watson (2002) finds that incomplete information can lead to 'starting small', in which levels of interaction between partners can grow over time as they overcome the informational asymmetries. In this paper 'starting small' occurs for two reasons: first, the benefit from the project is a result of cumulative allocations, and is increasing over time, so the opportunity cost of diverting funds is increasing over time. Second, with incomplete information the principal may be screening over time and more honest agents can be given a higher allocation without inducing diversion. This second effect is similar to Watson's analysis.

Section 3.1 gives an overview of the model. Section 2.2 examines the effects of incomplete information, when the agent's type is not known to the principal, studying how screening increases the principal's payoff. Finally in Section 2.3 I look at how changes in the distribution of types affect the optimal contract.

2.1 Model Overview

In order to construct the project, funds must be transferred from the principal to the agent, who must then choose between using them for the project and diverting them. The transactions may occur over a number of stages, with the benefit from the project growing over time as more funds are applied.

The principal's benefit derived from the project in stage t is then $b(P_t)$ and the agent's benefit from the project in stage t is $h(P_t)$. P_t represents the sum of allocations given in stages $s \leq t$ that were applied to the project $[P_t = \sum_{s \leq t} p_s]$.

Assumption 2.1.1. *Benefit Function:* $g = \{h, b\}$

i. *g(P) is weakly increasing,* $\lim_{P\to\infty} g'(P) = 0$

ii. *One of the following two shapes:*

Concave: $g''(P) \leq 0$ *for all* $P > 0$

s-shaped: $g''(P) \geq 0$ *for all* $P < c$ *and* $b''(P) \leq 0$ *for all* $P > c$

iii. $\frac{b(P)}{1-\delta} > P$

Given that the contract is for the financing of a defined project, I assume that the principal's marginal benefit from the project is weakly positive for all funding levels. However, at some point additional resources have little effect on the benefit from the project, and the marginal benefit approaches zero.

A concave benefit function realizes its potential immediately, and the marginal benefit is decreasing thereafter, while an s-shaped benefit function represents a project where there are fixed costs to overcome before a substantial benefit is realized.

The principal's and agent's benefit functions have the same inflection point when s-shaped because it is a feature of the underlying project rather than the magnitude of their benefit from the project.

In addition, the project must be worth undertaking if there were no constraints, otherwise the principal's optimal contract would be trivial for any utility of diversion.

The principal's cost of capital in stage t is normalized to $\delta^t p_t$, where δ is the discount rate.

If the agent diverts funds his payoff is the utility from diversion $xp_t - Z$, with Z representing the reputation cost of diversion and x his marginal utility from diversion. The benefit from diversion only involves the allocation in stage t (it is not a function of previous or future allocations). The agent's type is determined by his marginal utility of diversion, x.

The reputation cost parameter, Z, allows the enforceability of contracts and reputation costs to be built into the model. The more capacity the principal has to punish the agent, the higher Z will be.

In the next section I outline the optimal contract for a given distribution of types. In section 2.3 I show how parametric changes in the distribution affect the optimal contract. Finally, the conclusion is section 3.5.

2.2 Benchmark Model

I show that optimally the principal will commit to one sequence of funding levels, and in each stage the agent will choose between diverting the funds and applying them to the project. If he diverts the funds all future allocations are zero.

Any set of actions the principal would like to take in each stage can be decided ex-ante and committed to in a contract specifying a sequence of allocations $\{p_t\}_{t=1}^{\infty}$. Requiring the principal to maximize their utility over this sequence ex-ante increases their expected utility.

The maximum punishment the principal can enforce for deviation is withdrawing any future allocations. Therefore, if the agent diverts funds the principal will optimally allocate zero in all future stages.

If the agent diverts funds in stage t the principal gets a continuation value $\frac{b(P_{t-1})}{1-\delta}$, from the portion of the project completed. The agent forfeits the benefit from the project in stage t had they applied the funds, $h(P_t)$, plus the most he could have gotten in future stages δV_{t+1} dependent on his optimal actions in the future.

He also loses any continuation value from the portion already built, perhaps as a result of them having to flee the country with the funds, or that they no longer get credit for the project from the electorate.

Ex-ante commitment is optimal, and the principal is better off when commitment is credible, so it is implicitly assumed that the principal's cost of deviating is sufficiently high to prevent the principal from ever deviating.

In development settings projects are often left unfinished, or never started. With incomplete information this may be expected to occur as a byproduct of screening. The principal may not know the agent's marginal utility of diversion. Instead she may have an expectation based on the visible characteristics of the agent. There is a continuum of possible types assigned by nature.

Assumption 2.2.1. *The set of all possible types is* $X = [\underline{x}, \overline{x}]$ *where* $\underline{x} > 0$

With incomplete information, the principal may choose to increase the allocation in a given stage to 'screen' the agent, that is, to induce diversion from those types with higher marginal utilities of diversion. By inducing diversion amongst those most tempted to divert, the principal has effectively screened out some of the possible 'bad apples', and is able to give the agent a higher allocation in the future provided they are honest today. The cost is the increased risk of diversion, and subsequent non-completion of the project.

Proposition 2.2.2. *For any sequence of allocations* $\{p_t\}_{t=1}^{\infty}$*, the stage in which the* *agent chooses to divert funds is weakly decreasing in his utility of diversion, x.*

Proof. If type x' never diverts funds, his payoff from diversion must be less than his payoff from always being honest for every stage t:

$$x'p_t - Z \leq \sum_{s=t}^{\infty} \delta^{s-t} h(P_s) \qquad \text{for all } t \tag{2.1}$$

If $x < x'$ the equation will never bind, so any type $x < x'$ will never divert. As x increases, at some point (2.1) will be violated and the agent will prefer diversion in some stage t.

If agent x' diverts funds in stage k, then it must be optimal for him to divert funds in k:

$$x'p_k - Z \geq \sum_{s=k}^{\infty} \delta^{s-k} h(P_s) \tag{2.2}$$

As well, he must prefer diversion in stage k to diversion in any other stage $t \neq k$:

$$x'p_k - Z \geq \sum_{s=k}^{t-1} \delta^{s-k} h(P_s) + \delta^{t-k}(x'p_t - Z) \qquad \text{for all } t > k \tag{2.3}$$

$$x'p_t - Z \leq \sum_{s=t}^{k-1} \delta^{s-t} h(P_s) + \delta^{k-t}(x'p_k - Z) \qquad \text{for all } t < k \tag{2.4}$$

For $x > x'$ equations 2.2 and 2.3 will always hold, so a more dishonest type will divert funds, and will never prefer diversion in a later stage, to diversion in k. However if x is sufficiently greater than x' equation 2.4 will bind for some $t < k$, which implies that sufficiently dishonest agents would prefer diversion in some earlier stage to diversion in k.

As x increases the agent will divert funds in a weakly earlier stage, and therefore an earlier time in the game. Due to the lumpiness of discrete time there is some set of marginal utilities, X_k, where $x' \in X_k$, and any other $x \in X_k$ will have the same optimal strategy (divert funds in the same stage) as agent x'. □

Corollary 2.2.3. *For any sequence of allocations $\{p_t\}_{t=1}^{j}$ there is a sequence $\overline{x} \geq x_1 \geq x_2 \geq ...$ where all types x in $[x_{t-1}, x_t)$ will choose to divert in stage t.*

Since more dishonest agents will always divert funds earlier given any stream of allocations, any screening done by the principal will involve inducing any agent with marginal utility $x > x_t$ to divert funds in stage t or earlier. x_t is defined as the marginal utility of diversion at which an agent is indifferent between diverting funds in stage t and applying them to the project.

Typically screening models involve offering a menu of contracts, and different types partially identify themselves by their choices. Here we need offer only one contract because the agents will sort themselves by the stage at which they choose to divert.

Proposition 2.2.4. *Any optimal menu of contracts $\mathcal{C}$ that separates agents by type can be represented by a single contract, $C = \{p_t\}_{t=1}^{\infty}$, where types may differ in terms of the stage in which they choose to divert funds*

Proof. Let $\mathcal{C}$ be a menu of contracts optimally chosen by the principal, and $C(x) = \{p_t(x)\}_{t=1}^{\infty}$ be the contract in $\mathcal{C}$ chosen by type x. $\overline{x}$ is the type with the highest utility of diversion.

The proof will be broken into two cases. First, I show that when it is not optimal for the principal to induce diversion from type $\overline{x}$ the optimal menu of contracts is to offer a single contract that is incentive compatible for type $\overline{x}$. Second, I will show that when it is optimal for the principal to induce diversion, in every contract where honesty is induced in all stages $n < t$ the allocation in stage t must be the same.

i. If type $\overline{x}$ never diverts funds, but applies them to the project in every stage, the following incentive constraint must hold for every contract offered and every stage n:

$$\sum_{t=1}^{\infty} \delta^t h(P_t(\overline{x})) \geq \sum_{t=1}^{n-1} \delta^t h(P_t(x)) + \delta^n(\overline{x} p_n(x) - Z) \quad \text{for all } x \in X, n \leq T \tag{2.5}$$

$$\geq \sum_{t=1}^{\infty} \delta^t h(P_t(x)) \quad \text{for all } x \in X \tag{2.6}$$

If the incentive compatibility constraint for contract $C(\overline{x})$ does not bind in every stage n, the maximum incentive compatible allocation in any other contract $C(x)$ decreases (the LHS of equation (2.5) decreases). By offering the maximum incentive compatible allocation in every stage $t < j$ such that the ICC binds (as in the complete information contract outlined in Steele (2007)) the principal maximizes their benefit and maximizes the incentive compatible allocation for any contract $C(x) \neq C(\overline{x})$.

If $C(\overline{x})$ is the complete information contract for type $\overline{x}$, any contract $C(x)$ chosen by type x that satisfies the above constraints, but has a different allocation in some stage, must have a lower benefit than contract $C(\overline{x})$. If contract $C(\overline{x})$ is incentive compatible for type $\overline{x}$ it is also incentive compatible for any type $x < \overline{x}$, and therefore in order to maximize her benefit, the principal would offer $C(\overline{x})$ to all types.

ii. If by offering contract $C(\overline{x})$ the principal induces diversion from type $\overline{x}$ in stage k, the following incentive compatibility constraints must hold for every stage $n > 1$ in contract $C(x)$ for all $x \in X$:

$$\sum_{t=1}^{k-1} \delta^t b(P_t(\overline{x})) + \delta^k(\overline{x}p_k(\overline{x}) - Z) \geq \sum_{t=1}^{n-1} \delta^t b(P_t(x)) + \delta^n(\overline{x}p_n(x) - Z) \tag{2.7}$$

$$\sum_{t=1}^{k-1} \delta^t b(P_t(\overline{x})) + \delta^k(\overline{x}p_k(\overline{x}) - Z) \geq \sum_{t=1}^{\infty} \delta^t b(P_t(x)) \tag{2.8}$$

Starting in stage one, the allocation in any contract is bound by:

$$\sum_{t=1}^{k-1} \delta^{t-1} b(P_t(\overline{x})) + \delta^{k-1}(\overline{x} p_k(\overline{x}) - Z) \geq \overline{x} p_1(x) - Z$$

$$\Rightarrow p_1(x) \leq \delta^{k-1} p_k(\overline{x}) + \frac{\sum_{t=1}^{k-1} \delta^{k-1} b(P_t(\overline{x})) + (1 - \delta^{k-1})Z}{\overline{x}} \tag{2.9}$$

If this equation is incentive compatible for type $\overline{x}$ it is also incentive compatible for any type $x < \overline{x}$. Therefore if type $\overline{x}$ prefers not to divert funds in stage one, no other type will divert funds in stage one.

Given that no type chooses to divert funds in stage one, the principal will maximize their benefit by setting $p_1(x)$ subject to the constraint that Equation (2.9) binds in all contracts. If this were not the case, and there were some contract $C(x)$ where the inequality is strict, the principal could increase their benefit in contract $C(x)$ by increasing $p_1(x)$ by ε and decreasing $p_2(x)$ by ε, strictly increasing their benefit.

Given that the allocations are the same for all contracts in stage one, all contracts face the same incentive compatibility constraint for stage two:

$$\Rightarrow p_2(x) \leq \delta^{k-2} p_k(\overline{x}) + \frac{\sum_{t=2}^{k-1} \delta^{k-2} b(P_t(\overline{x})) + (1 - \delta^{k-2})Z}{\overline{x}} \tag{2.10}$$

Again the maximum benefit the principal can obtain is from allocating the maximum incentive compatible $p_2(x)$ in every contract. This continues for every stage $t < k$, resulting in identical allocations in all stages $t < k$ in all contracts.

Because (2.7) must hold in any contract, and assuming $C(\overline{x})$ is the optimal contract for type $\overline{x}$ the maximum benefit the principal can obtain in any contract is where the maximum incentive compatible allocation for type $\overline{x}$ is offered in every stage $n < k$.

In stage k, Equation (2.7) must also hold, so $p_k(x) \leq p_k(\overline{x})$ in every contract. If all future allocations were zero in every contract, then $p_k(x) = p_k(\overline{x})$ and any separation of types is trivial. If $p_k(x) < p_k(\overline{x})$ then all types would prefer contract $C(\overline{x})$. If the future allocation is positive in some stage $m > k$ in contract $C(x)$, and $p_k(x) < p_k(\overline{x})$ then the principal could be made better off increasing the allocation in k by ε and decreasing the allocation in m by ε. Therefore the allocations in every stage $t \leq k$ are the same in all contracts.

For every contract $C(x)$ the allocation in any stage $n > k$ must also satisfy Equation (2.7). Let x_k denote the highest utility of diversion who chooses to be honest in stage k. Then the following equations must hold for every $n > k$.

$$x_k p_k(x_k) - Z \geq \sum_{t=k-1}^{n-1} \delta^{t-k} b(P_t(x)) + \delta^{n-k}(x_k p_n(x) - Z) \tag{2.11}$$

If this equation binds for some $x_k < \overline{x}$ it is non-binding for any type $x > x_k$ and we no longer need to worry about Equation (2.7). Any type with $x < x_k$ strictly prefers honesty in stage k while any type $x > x_k$ strictly prefers diverting funds. Again, the above equation holds up to some stage n in which type x_k chooses to divert funds, at which point a new type with the highest utility of diversion is present in $n + 1$.

Therefore in stage one, all allocations are identical. For all types remaining in stage two, the allocations in their respective contracts are also identical. This continues for every stage n, where all contracts that induce honesty in every stage $t < n$ must have identical allocations in stage n.

In a contract where the agent diverts funds in stage k, the allocations in stages $t > k$ are irrelevant, provided they meet the agent's ICC constraint (when an agent diverts funds, they forfeit all future allocations). From Equation (2.11) all ICC's for $t > k$ are met for all agents who optimally chose to divert funds in stage k, and therefore any stream of allocations for stages $t > k$ in any other contract will be incentive compatible for agents who divert in k. As a result, if the contract for type x_k who diverts in stage k is offered to type x_n, who diverts in stage $n < k$, the outcome will not change.

The optimal menu of contracts offered can then be obtained by offering all types one contract, where types may differ in choosing the stage in which they divert funds. $\square$

In Steele (2007) we find that due to the cumulative nature of the benefit function, the principal will allocate payments earlier rather than later when possible, to the extent that it does not induce the agent to divert funds. With asymmetric information she will want to allocate payments subject to the constraint that, at a minimum, the most dishonest agent remaining in the game is indifferent between diverting and not diverting (except for the last stage with a positive allocation).

In light of this result, the optimization problem for the principal can be seen as a sequence of allocations that may induce diversion with positive probability (screen) in some or all stages. The optimization problem can be written as follows:

$$Max_{\{p_t\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \delta^t [F(x_t)b(P_t) + [F(x_{t-1}) - F(x_t)] \frac{b(P_{t-1})}{1-\delta} - F(x_{t-1})p_t]$$

$$s.t. \qquad x_t p_t - Z - [h(P_t) + \delta[x_t p_{t+1} - Z]] \geq 0 \qquad \forall t \leq T-1 \tag{2.12a}$$

$$x_{T-1}p_{T-1} - Z - \left[h(P_{T-1}) + \frac{\delta}{1-\delta} h(P_T)\right] \geq 0 \tag{2.12b}$$

$$x_T = x_{T-1} \tag{2.12c}$$

x_t is the type indifferent between diverting funds and honesty in stage t. $F(x_t)$, the probability of the agent's type being less than x_t, is the probability of the agent applying the funds to the project in stage t.

Which leads to the following FOC:

$$
\begin{aligned}
p_k: \qquad & \sum_{t=k}^{\infty} \delta^t \left[\frac{F(x_t) - \delta F(x_{t+1})}{1-\delta} \right] b'(P_t) - \delta^k F(x_{k-1}) \\
& + \sum_{t=1}^{T-1} f(x_t) \frac{\partial x_t}{\partial p_k} \left[\frac{b(P_t) - b(P_{t-1})}{1-\delta} - \delta p_{t+1} \right] A_t \leq 0
\end{aligned}
\tag{2.13}
$$

The first term is the marginal benefit from the project provided the agent is honest in stage t less the discounted probability of him being honest in $t+1$, the second term is the marginal cost in stage k. The last term is the benefit to the principal of the agent being honest in stage t multiplied by the marginal effect of an increase in p_k on the probability of the agent diverting funds. An increase in p_k increases the probability of the agent diverting funds in stage k, while it weakly decreases the probability of the agent diverting funds in earlier or later stages.

If she was not subject to any incentive compatibility constraints, the principal would offer an allocation in the first stage such that the marginal cost equals the marginal benefit:

$$
\frac{b'(P^\circ)}{1-\delta} = 1 \tag{2.14}
$$

The sequence of funding levels where P° is offered in the first stage, and the allocation is zero thereafter is the *unconstrained optimal contract.*

This contract is only incentive compatible for types with $x \leq x^\circ$ where x° is defined by the following:

$$x^{\circ}P^{\circ} - Z = \frac{h(P^{\circ})}{1-\delta} \tag{2.15}$$

For any type with $x > x^{\circ}$ the unconstrained optimal contract would induce diversion in the first stage.

Proposition 2.2.5. *The optimal contract involves a sequence of allocations* $\{p_t\}_{t=1}^{T}$ *that induce honesty with positive probability such that in each stage* $t < T$ *the most dishonest type that does not divert funds in any stage* $k < t$, x_{t-1}, *is at least indifferent between diverting funds and applying them to the project in stage* t.

Proof. If the principal offers the agent the unconstrained optimal contract, and $\overline{x} > x^{o}$ the agent would divert funds with positive probability in the first stage. The FOC for stage one would be as follows:

$$F(x_1)\frac{b'(P^{\circ})}{1-\delta} - 1 + f(x_1)\frac{\partial x_1}{\partial p_1}A_1 \leq 0 \tag{2.16}$$

By definition, $x_1 = x^{\circ}$, so $F(x_1)$ is less than one, and the screening effects are negative (increasing p_1 decreases the probability that the agent will be honest). The FOC for the first stage would be strictly less than zero. Therefore p_1 must decrease, and if $p_2 = 0$ the FOC for stage two would be

$$F(x_1)\frac{\delta b'(P_1)}{1-\delta} - \delta + f(x_1)\frac{\partial x_1}{\partial p_2}A_1 \leq 0 \tag{2.17}$$

In this case the screening effects are positive, so the FOC is violated, and p_2 must be greater than zero.

In general, if there is screening in stage t ($x_t < x_{t-1}$), the allocation in stage $t+1$ must be positive. If the allocation in stage $t+1$ were zero, the following two equations could not both hold:

$$\begin{aligned} p_t : \sum_{s=t}^{\infty} F(x_s)\delta^s b'(P_s) + [F(x_{t-1}) - F(x_t)]\delta^t \frac{b'(P_{t-1})}{1-\delta} - F(x_{t-1})\delta^t \\ + f(x_{t-1})\frac{\partial x_{t-1}}{\partial p_t}A_{t-1} + f(x_t)\frac{\partial x_t}{\partial p_t}A_t = 0 \end{aligned} \tag{2.18}$$

$$p_{t+1} : \sum_{s=t+1}^{\infty} F(x_s)\delta^s b'(P_s) - \delta^{t+1}F(x_t) + f(x_t)\frac{\partial x_t}{\partial p_{t+1}}A_t < 0 \tag{2.19}$$

If (2.18) holds, then (2.19) would be greater than zero, and p_{t+1} must increase. Therefore if there is screening in stage t, the allocation in stage $t+1$ must be positive.

As in the complete information model, the allocations must be such that the most dishonest agent remaining in the game in each stage t is at least indifferent between honesty and diversion in every stage $t \in [2, T-1]$. This is because the principal strictly prefers allocating funds earlier when possible due to the cumulative nature of the benefit function. Generally the ICC will bind in the first stage as well, with the exception of certain s-shaped benefit functions. $\square$

If agents are sufficiently dishonest, it may be optimal to offer a 'bonus' in stage $T+1$. By inducing the agent to divert funds in the final stage, the incentive to be honest in earlier stages increases, and the principal is able to complete the contract earlier. The conditions under which this is optimal are outlined in 2.5.

From the first order condition we see that the optimal contract is dependent on the distribution of x. In the next section we look at two parametric examples outlining changes in the optimal contract resulting from shifts in the distribution of types.

2.3 Two Parametric Examples

The following two propositions are assuming that x is distributed along a uniform distribution. This results in a constant density $f = F'(\cdot)$ for a given distribution $F(\cdot)$. The first example involves a mean-preserving contraction of the distribution. This is the opposite of a mean-preserving spread, as defined in Rothschild and Stiglitz (1970). The second example involves a shift in the distribution, while maintaining the same density.

Proposition 2.3.1. *For any mean-preserving contraction of the distribution, the optimal contract will involve less screening, up until some density f^c, where the optimal contract will mirror the complete information optimal contract for the most dishonest type, $\overline{x}$.*

Proof. Due to the uniform distribution of x, $f = \frac{1}{\overline{x}-\underline{x}}$ and a mean-preserving contraction of the distribution results in a density, $f' > f$ where $\overline{x}' = \overline{x}-\varepsilon$ and $\underline{x}' = \underline{x}+\varepsilon$ for any positive value $\varepsilon < \frac{\overline{x}-\underline{x}}{2}$.

If the allocations were the same in each stage after the contraction, the type which is indifferent between diversion and honesty would be the same as well ($x_t' = x_t$). Dividing (2.13) by the density, the following equation is derived:

$$\sum_{t=k}^{T-1} \delta^{t-1}(x_t - \underline{x})b'(P_t) + \delta^{T-1}(x_{T-1} - \underline{x}) \left[\frac{b'(P_T)}{1-\delta}\right] - \delta^{k-1}(x_{k-1} - \underline{x}) + \sum_{t=1}^{T-1} \frac{\partial x_t}{\partial p_k} A_t \leq 0 \tag{2.20}$$

Because the screening effects are negative in early stages, the first order condition is strictly negative if $\underline{x}$ increases. As a result, if the density increases, p_k must decrease in early stages, and the marginal utility of diversion at which an agent is indifferent between diversion and not increases, $x_t' > x_t$.

As t increases, the screening effects increase, and the effect of a change in the density on the first order condition decreases. At some point the screening effects become positive, and the first order condition becomes positive as $\underline{x}$ increases. In these stages, the principal will increase the allocation, and if there is screening, the marginal utility at which an agent is indifferent between diversion and honesty decreases $x_t' < x_t$.

As the density increases, the cost of screening increases. The marginal effect of an increase in p_t on the probability of the agent diverting is greater. As a result the principal will keep more dishonest types in the game longer.

The overall result of an increase in the density is that in early stages the marginal utility of diversion at which the agent is indifferent will increase. In the later stages the marginal utility of diversion at which the agent is indifferent may decrease, as the principal screens less in early stages.

As the density continues to increase, at some point the costs of screening are sufficiently large that the principal will not screen any agents in any stage. From the complete information case, we know that the best the principal can do, if no screening is occurring, is to offer the complete information optimal contract for $\overline{x}$, $\{p_t^c\}_{t=1}^T$.

Looking at the first order condition for the first stage, the marginal cost is 1 because all agents are in the game:

$$\sum_{t=1}^{T-1} \delta^{t-1} F(x_t) b'(P_t) + \delta^{T-1} F(x_{T-1}) \left[\frac{b'(P_T)}{1-\delta} \right] - 1 + f \sum_{t=1}^{T-1} \frac{\partial x_t}{\partial p_1} A_t \leq 0 \qquad (2.21)$$

As f increases, p_t decreases until $x_1 = \overline{x}$, at which point no screening occurs in the first stage. Once $x_1 = \overline{x}$, then any further increase in f has no effect on $F(x_1)$ because $F(x) = 1$ for all $x \geq \overline{x}$. The principal will always allocate funding earlier rather than later so the incentive compatibility constraint for $\overline{x}$ must be met with equality. Therefore as f continues to increase screening in future stages must decrease to compensate. As a result, when f is sufficiently high, no screening will take place. $\square$

Decreases in uncertainty increase the principal's cost of screening. As the cost increases, the principal screens less in early stages, until at some sufficiently high cost of screening, they stop screening altogether.

An increase in uncertainty would act in the opposite manner. The marginal utility of diversion at which the agent is indifferent would decrease when the screening effects are negative, and the probability of the agent surviving in each stage would decrease.

When the distribution is sufficiently compact no screening will take place, and the complete information optimal contract will apply.

Proposition 2.3.2. *If the expected marginal utility of diversion of the agent in-*

creases, while the density remains the same, the probability of the agent diverting funds in any given stage increases.

Proof. Increases in the expected marginal utility of diversion of the agent, without any change in the density would work in a similar manner to changes in density. In this case, the upper and lower limits $(\overline{x}, \underline{x})$ of marginal utilities increase, while the density remains the same.

Once again, the increase in the lower limit decreases the first order condition in stages where the screening effects are negative, and the marginal utility of diversion at which agents are indifferent increases.

The effect on the probability of the agent remaining honest in each stage is different. In order to keep the probabilities the same, the principal would have to decrease the allocation in each stage, in order to increase x_t. This would increase the marginal benefit from the project, while keeping the marginal cost the same. The screening effects would decrease, as the benefit from agents diverting tomorrow rather than today decreases.

$$\sum_{t=k}^{T-1} \delta^{t-k} \frac{b'(P'_t)}{1-\delta}[F_o(x_t) - \delta F_o(x_t+1)] + \delta^T F_o(x_{T-1}) \frac{b'(P'_T)}{1-\delta} - \delta^{k-1} F_o(x_{k-1}) + f \sum_{t=1}^{T-1} \frac{\partial x_t}{\partial p'_k} A_t \leq 0 \qquad (2.22)$$

$F_o(x_t)$ is the ex-ante probability of the agent remaining honest in stage t. Above, assuming the probabilities remain the same, the increase in marginal benefits outweigh any changes in screening effects, and as a result the first order condition would be positive. The allocation in each stage with negative screening effects would increase to satisfy the equation, and the probability of the agent diverting in each stage would increase. $\square$

Therefore as the agent's expected marginal utility of diversion increases, the principal will screen with a higher probability, because the benefit of keeping them in the game is lower.

Overall, a decrease in the density of the distribution decreases the cost of screening, resulting in more screening occurring, while an increase in the expected marginal utility of diversion of the agent increases screening because the marginal benefit of keeping the agents honest decreases.

2.4 Conclusion

When the principal and agent both derive positive benefit from a project, the question of how to finance the project involves how best to induce the agents to use the funds for the project. The difference in value that types place on alternate uses of the funds can prevent the efficient outcome from being obtained. When there is incomplete information and there is a positive probability of the agent being sufficiently dishonest the question becomes when should the principal cut their losses and screen the more dishonest agents (the 'bad apples'). The cost of inducing diversion with positive probability can be outweighed by the benefit from being able to offer higher allocations to relatively more honest types, and completing the contract earlier.

Due to the cumulative nature of the allocations, and the incentive compatibility constraint binding in every stage up until the final stage, the allocations in all stages where the ICC binds are increasing over time. This 'starting small' effect can be broken into two main components. The first, similar to that outlined in Watson (2002), is the ability of the agent to give larger allocations once the more dishonest types had been screened. This effect occurs due to the incomplete information. The second effect is due to the cumulative nature of the benefit function. As the project grows, the cost of diversion (giving up the future benefit) increases, so the principal can increase the allocations as the project nears completion.

With incomplete information multi-stage contracts allow the principal to screen more dishonest agents. They can induce honesty from more dishonest types when the marginal benefit is high in the earlier stages, then induce diversion with positive probability, and finish the contract with only relatively more honest types. As the expected marginal utility of diversion increases the agent will make it to the last stage with decreasing likelihood. As the density of the distribution of types increases screening becomes more expensive, and the principal decreases the amount of screening they undertake, until the contract is the same as in the complete information case, with no screening.

Overall, multi-stage contracts increase efficiency due to the incentives the principal may offer, and the opportunities for screening at different stages of the project.

Possible extensions of the model include a closer look at how screening occurs for different benefit functions. Making different assumptions on the second and third derivatives can greatly alter the screening contract.

In addition, a more general framework with the effects of changes in the distribution on screening would be useful.

2.5 Supplement: Inducing Diversion with Incomplete Information

If it were optimal for the principal to induce diversion in stage $T+1$ when there is incomplete information, the maximization problem would look as follows:

$$\max_{\{p_t\}_{t=1}^{T+1}} \sum_{t=1}^{T} \delta^t [F(x_t)b(P_t)+[F(x_{t-1})-F(x_t)]\frac{b(P_{t-1})}{1-\delta}-p_t]+\delta^{T+1}F(x_{T-1})\left[\frac{b(P_T)}{1-\delta}-p_{T+1}\right] \tag{2.23}$$

s.t.

$$x_t p_t - Z - [h(P_t) + \delta[x_t p_{t+1} - Z] \geq 0 \tag{2.24a}$$

$$x_{T-1}p_{T-1} - Z - [h(P_{T-1}) + \delta\frac{h(P_T)}{1-\delta} + \delta^2[x_{T-1}p_{T+1}] \geq 0 \tag{2.24b}$$

Which leads to the following first order conditions:

$$\begin{aligned} p_k: &\sum_{t=k}^{T-1} \delta^t F(x_t)b'(P_t) + \sum_{t=k+1}^{T-1} \delta^t [F(x_{t-1}) - F(x_t)]\frac{b'(P_{t-1})}{1-\delta} \\ &+ \delta^T F(x_{T-1})\frac{b'(P_T)}{1-\delta} - \delta^k F(x_{k-1}) + f(x_k)\sum_{t=1}^{T-1}\frac{\partial x_t}{\partial p_k}A_t \leq 0 \end{aligned} \tag{2.25a}$$

$$p_{T+1}: -\delta^{T+1}F(x_{T-1}) + f(x_{T-1})\sum_{t=1}^{T-1}\frac{\partial x_t}{\partial p_{T+1}}A_t \leq 0 \tag{2.25b}$$

If k is the last stage in which screening occurs ($x_t = x_k$ for all $t > k$) then inducing diversion is optimal if and only if

$$\frac{f(x_k)}{F(x_k)}x_k \geq \frac{1}{\delta^{-k}[\sum_{t=k}^{T-1} \delta^t h(P_t) + \frac{\delta^T h(P_T) - \delta^k h(P_k)}{1-\delta} - \sum_{t=k+1}^{T+1} \delta^t p_t]} \tag{2.26}$$

Chapter 3

The Effect of Foreign Productivity on FDI Decisions

Abstract

With varying levels of productivity in the home country, firms' decisions of whether or not to invest in a foreign country will be based on their expected foreign productivity, and may differ between firms. Incomplete information about foreign productivity can explain both clustering, and the underprovision of foreign direct investment. If there is no learning, then foreign direct investment will be underprovided. With learning, clustering within industries can occur, and ex-post, investment may not follow comparative advantage.

Factor price equalization suggests that large differences in wages across countries are reflective of differences in productivities between workers. However, due to limited factor mobility and imperfect information, wage differentials may not accurately reflect the differences in productivities, but instead reflect other factors, such as level of industry operating in the country, opportunity costs for workers and trade barriers.

Estimates of foreign productivity indicate that levels of foreign direct investment in less developed countries should be much higher than they are. FDI occurs most often between north countries rather than from north to south countries. Lucas (1990) explains the asymmetry through differences in human capital, externalities and capital market imperfections.

Other explanations for the discrepancy between North-North and North-South FDI include exchange rate volatility and financing options (see Cushman (1985) and Goldberg and Kolstad (1994)). Cushman argues that more exchange rate volatility decreases the expected return from FDI. In Goldman et al. they argue that when demand shocks are correlated with exchange rate shocks the profitability of FDI increases.

Most trade literature treats FDI as a method of avoiding transport costs when serving a new market. Using the heterogeneous firms model outlined in Melitz (2003), Helpman, Melitz and Yeaple (2004) explore the choice between export and FDI with heterogeneous firms. Every firm faces the decision of whether or not to serve a foreign market, and whether to serve it through FDI or exporting. Their framework suggests that heterogeneity in productivity determines firms' FDI decisions, the least productive produce domestically and serve the domestic market only, more productive firms export to the foreign market and the most productive firms set up subsidiaries in the foreign market. This approach fits with the majority of north-north investment, but the asymmetries between north and south countries lead to different equilibria. In this paper I attempt to explain the disparities between the data and the theory by allowing for differences in productivity between the two countries, asymmetric markets and incomplete information.

FDI in developing economies may be undertaken to take advantage of lower costs of production. At times the firm may not even serve the market in the country where they produce, choosing to export all production. With asymmetric markets I find equilibria where foreign demand may not be sufficient to warrant producing in the foreign country just to serve the foreign market, where firms may produce in the foreign country just to serve the home market.

Hausmann and Rodrik (2003) suggest that industrial investment is lower than expected due to incomplete information. Entrepreneurs do not know what they are good at producing. They explain why Latin America has not experienced the growth that Asia has, although Latin America has made more attempts to follow the growth-creating policies suggested by development economists. Using a two sector model with both traditional and modern sectors, they allow uncertainty to enter in the production functions of the modern sector goods. They conclude that uncertainty can result in too little investment and entrepreneurship, but once firms have invested there is too much product variety in the market, as less productive firms stay in once the sunk costs are committed.

Empirically they look at three 'building blocks' for their argument, i) a large element of uncertainty as to what a country will be good at producing, ii) significant difficulties importing technology off-the-shelf, with many changes required for local adaptation, iii) imitation follows quickly once the first two blocks are overcome.

Entry to the industry is only allowed through FDI as opposed to domestic investment. Razin, Sadka and Tong (2005) motivate this assumption by arguing that domestic entrepreneurs face higher setup costs than more knowledgable foreign firms. In this paper it is not motivated explicitly, but foreign firms have already entered the industry in their home country, and therefore may have some knowledge that will lower their setup costs of FDI. It may also be due to within-firm public goods that increase a firm's productivity as in Horstmann and Markusen (1989).

Empirically incomplete information could show up in many different forms. Benassy-Quéré, Coupet and Mayer (2005) evaluate the role of institutions in FDI. They find that information about firms and quality of goods in the destination country increased the amount of FDI.

The clustering effect that arises in the paper is due to knowledge spillovers, 'learning', that represent the transfer of information about productivity once one firm has entered the market. In the literature clusters have been studied as agglomeration effects. Matsuyama (1996) argues that specialization occurs as a result of trade, and the world is split into rich and poor because countries specialize in different economic activities. This paper explains how this specialization may occur as a result of imperfect information.

3.1 Model Overview

Differences between a firm's productivity in the home country and the foreign country is divided into two main sources. The first is derived from the individual firms' technologies, and their ability to adapt that technology to a foreign workplace. Firms are heterogeneous with respect to labor productivity in the home market, which is based on their firm-specific technology. Some technologies may be more productive than others in a foreign workplace, and may not differ in the same way as in the home country (i.e. more productive technologies in the home market are not necessarily more productive in the foreign market).

The second source is industry-specific differences in productivity. This could be influenced by differences in education, cultural practices, or institutions. Some countries may be more productive in textile industries because of cultural emphasis on handicrafts, or because of the group culture leading to more productive assembly lines, while other countries may be more adept at work involving machines because of the emphasis on machinery in the educational system. These are not to be thought of as innate abilities or characteristics of certain countries. Rather they should be thought of as learned skills from social interaction in different cultural environments or the education system, which can vary dramatically from country to country in terms of both level and focus of education.

If there is uncertainty about either source the level of FDI can be inefficient. In section 3.4.1 the inefficiencies from firm-level uncertainty are outlined. If uncertainty is firm-level FDI will be underprovided. Firms whose true productivity would warrant investment may not enter the industry, while firms whose true productivity is low, would enter, discover their productivity and leave, creating waste. In aggregate the number of firms operating in the foreign country will be below the efficient level.

Industry-specific uncertainty can allow learning to occur after one firm enters, an extension characterized in Section 3.4.2. If all firms' productivities are known once one firm has entered clustering can occur. It can also lead to no entry when entering is optimal, because the first-mover cannot internalize the externalities. It may be optimal to increase the first-mover advantage by extending their lead time, in order to encourage investment.

In 3.4, I assume that firms receive a productivity shock in the foreign country, either high or low. A firm will always invest in the foreign country if they receive a high shock. Because the FDI I focus on is rich-poor I assume that there is adequate surplus labor in the poor country that a firm undertaking FDI does not affect the wage rate. In Razin et al. (2005) the effect of FDI on wages counteracts the positive productivity shock.

The spillovers from FDI make it a very appealing means to promote economic growth. Traditionally countries have been urged to follow free market policies in order to increase growth. Countries which have resisted this open economy push in the short run, like South Korea and Taiwan, have seen strong growth, while countries making the requisite open economy changes, like Brazil or Argentina, have not.

Incomplete information also suggests a reason why when FDI does occur, it often clusters in a few select industries, often quite disparate between countries. The typical clustering examples are Bangladesh with hat and pant producers, Columbia cut flowers, and India software design. If firms learn about the productivity of their industry from other firms' investment decisions, incomplete information can be overcome with sufficient investment once investment has occurred within the industry.

In section 3.2 the basic model is introduced and the autarkic equilibrium is outlined when there is no production in the foreign country. In section 3.3 the autarkic and export-only equilibria are presented to remind the reader of the equilibria found in Melitz's work (2003) and introduce slight differences in parameterizations and assumptions. The model with FDI expands on Helpman et al's (2004) work to include the option for FDI that does not serve the foreign market, and for different productivities in the home country versus the foreign country. In section 3.4 the concept is expanded to evaluate firms' FDI decisions when they do not know their foreign productivity. Increases in uncertainty lead to the underprovision of FDI under certain conditions. They may also lead to 'poverty traps' where for sufficient levels of risk no investment will be undertaken.

3.2 Setup of the Model

This section reviews the basic heterogeneous firm model set up in Melitz (2003), with slightly different assumptions and parameterizations of the variables.

Using a two country, general equilibrium model, the effects of differences in factor prices, incomplete information, and heterogeneity in productivity levels on the decisions of the firm is outlined. The focus is on trade between a developed country (the home country), and a less developed country (the foreign country). Consumers in the home country have higher wages and spend a larger amount of money on modern goods than consumers in the foreign country.

There are $M+1$ industries in the world. In each industry m a modern good is produced, of which an infinite number of varieties are available for production. In the remaining industry a homogeneous good is produced. For ease of analysis it is assumed that each firm employs labor as the sole factor of production. In each industry m the set of varieties being sold in country c's market is denoted as V^c.

Each firm i in industry m and country c thus produces variety $v_i \in V^c$. Each firm produces output by the following production function:[1]

$$Y_i^c = L_i^c \tilde{a}_i^c \tag{3.1}$$

where Y_i^c is the output of firm i in country c, L_i^c is the labor hired by firm i in country c, $\tilde{a}_i^c$ is the number of units of output produced using one unit of labor. Letting w^c indicate the wage in country c.

$$a_i^c = \frac{\tilde{a}_i^c}{w^c} \tag{3.2}$$

Now a_i^c represents the units of output produced per dollar and is henceforth referred to as firm i's productivity in country c.

Consumer Preferences

As is typical in the literature, consumers in both countries have CES preferences, represented by the log utility function:

$$u = (1 - \sum_{m=1}^{M} \beta_m^c) \log(z) + \sum_{m=1}^{M} \frac{\beta_m^c}{\alpha_m} \log(\int_{v \in V_m^c} x_m(v)^{\alpha_m} dv) \tag{3.3}$$

Where $x_m(i)$ is the consumption of variety i (from firm i) in sector m, V_m^c

[1] All analysis is done from the perspective of industry $m \in M$ and the industry subscript is dropped

is the set of all firms in sector m selling their good in country c and z is the units of homogeneous goods consumed. The elasticity of substitution is $\varepsilon = \frac{1}{1-\alpha}$ and is assumed to be greater than one. The proportion of income spent on goods in sector m in country c is fixed at β_m^c and $\Sigma_m \beta_m^c < 1$. The remaining income, $1 - \Sigma_m \beta_m^c$, is spent on the homogeneous good.

From these preferences we find that the demand for good i in sector m is

$$x_m^c(i) = \left[\frac{\beta_m^c E^c}{\sum_{v \in V^c} p_c(v)^{1-\varepsilon}}\right] p^c(i)^{-\varepsilon} \tag{3.4}$$

E^c is the income of country c and $\tau^{c\neg c}$ represents the iceberg transport costs of shipping a good from country c to country $\neg c$. It is assumed that transport costs are symmetric, so $\tau^{c\neg c} = \tau^{\neg cc} = \tau > 1$. $P^c = \Sigma_{j \neq i} p^c(j)$. Firm i offers their good in country c at p_i^c which when producing their good in country c is $p_i^{cc} = (a_i^c \alpha)^{-1}$ and when producing in country $\neg c$ for sale in c is $p_i^{\neg cc} = \tau(a_i^{\neg c}\alpha)^{-1}$. This gives us the following gross profit for a firm producing in country c and serving country c:[2]

$$\pi^{cc} = (1-\alpha)(a_i^c \alpha)^{\varepsilon - 1} A^c \tag{3.5}$$

And for a firm producing in country $\neg c$ and serving country c

$$\pi^{\neg cc} = (1-\alpha)(a_i^{\neg c}\alpha)^{\varepsilon-1} A^c \tau^{1-\varepsilon} \tag{3.6}$$

[2] $A^c = \frac{\beta^c E^c}{\Sigma_{v \in V^c} p^c(v)^{1-\varepsilon}} \quad c \in \{h, f\}$

Additionally each firm faces fixed costs from three sources. First, they must pay f_e to enter the industry, after which they draw their productivity a^c from the distribution $G(a)$. Then, if they choose to produce, they bear fixed costs f_p for each plant location. Finally, for each market in which they wish to sell their goods, they must pay distribution costs f_x.

Autarkic Equilibrium

Melitz (2003) solves for the autarkic equilibrium, when no firms are exporting to or operating in the foreign country. Firms are heterogeneous with respect to their productivity in the home country. Firms that are currently producing goods know their productivity in the home country, and firms that are not currently producing know the distribution of productivities they could draw from, were they to enter the market. Each industry consists of an equilibrium determined number of heterogeneous firms, differentiated by their productivity levels. Firm i producing at home has productivity a_i^h

Using the variables as defined earlier, his cutoff point for realized productivity level a^{h*} above which firms will produce, and below which firms will choose not to produce becomes:

$$(a^{h*})^{\varepsilon-1} = \frac{f_x + f_p}{(1-\alpha)A^h\alpha^{\varepsilon-1}} \tag{3.7}$$

Given the cutoff point, firms will choose to enter and draw a productivity if:

$$(1-\alpha)\alpha^{\varepsilon-1}A^h\int_{a^{h*}}^{\infty}a^{\varepsilon-1}f(a)da \geq (1-F(a^{h*}))[f_x+f_p]+f_e \tag{3.8}$$

Because we are summing over the prices of all firms in the industry (in A^c), an increase in the number of firms decreases the expected profits of a potential entrant, so firms will continue to enter until the expected profits reach zero. The cutoff point, a^{h*}, is increasing in the number of firms in the market thus given the fixed entry costs f_e the number of firms serving the home market can be determined.

Export Equilibrium

If we open up the foreign country, and allow the firms producing in the autarkic equilibrium in the home country to start exporting to the foreign country, firms now have three options: b_0 do not produce, $b_{h,h}$ produce in the home country and serve the home market only, $b_{h,b}$ produce in the home country and serve both the home and foreign markets. It is assumed that the foreign country is sufficiently small[3] so it will never be optimal to produce in the home country and serve only the foreign market or produce in the foreign country, only serving the foreign market.

There are a number of reasons why at some point in time a foreign country may 'open up'. It may be that previously it was simply too expensive to export to

[3] $A^f < \frac{f_x}{f_x+f_p}A^h\tau^{\varepsilon-1}$ which implies that $\beta^f E^f$ is sufficiently small, so it could mean that they have a lower GDP or that they spend a smaller amount of their income in that specific industry

that market (τ was too large). This encompasses both trade restrictions (tariffs) or transport costs. A decrease in transport costs or tariffs could lead to the foreign country suddenly being attractive to home firms. Another reason could be the size of the foreign market. If the foreign economy is growing, at some point home firms will choose to export to the foreign market to take advantage of the market growth.

Melitz (2003) also solves for the export equilibrium. Using the variables as defined earlier, his productivity cutoff points become:

$$(a_p^h)^{\varepsilon-1} = \frac{f_p + f_x}{(1-\alpha)A^h\alpha^{\varepsilon-1}} \tag{3.9}$$

$$(a_x^h)^{\varepsilon-1} = \frac{f_x}{(1-\alpha)A^f\tau^{1-\varepsilon}\alpha^{\varepsilon-1}} \tag{3.10}$$

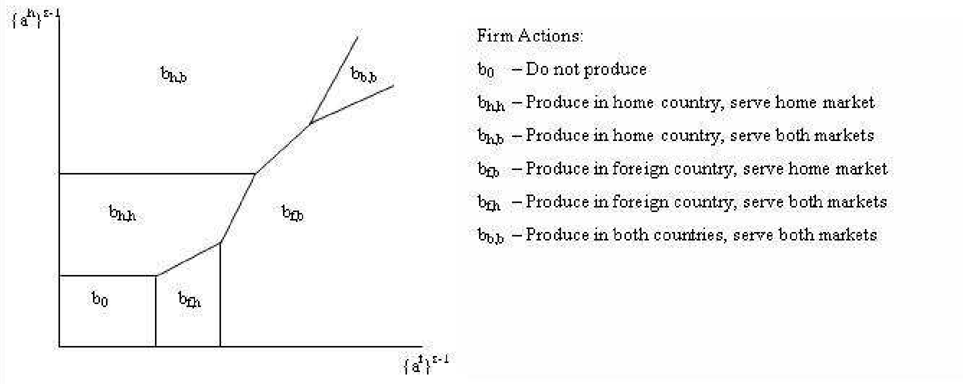
Once firm i has drawn a productivity level, they will not produce if $a_i^h < a_p^h$, will produce and only serve the home market if $a_p^h < a_i^h < a_x^h$ and will produce and serve both the home and foreign markets if $a_i^h > a_x^h$. The expected profit from entering the home market is

$$\int_{a_p^h}^{\infty} a^{\varepsilon-1} f(a) da \alpha^{\varepsilon-1}(1-\alpha)A^h - (1-F(a_p^h))(f_x + f_p) \tag{3.11}$$

$$+ \int_{a_x^h}^{\infty} a^{\varepsilon-1} f(a) da \alpha^{\varepsilon-1}(1-\alpha)A^f\tau^{1-\varepsilon} - (1-F(a_x^h))(f_x) - f_e \tag{3.12}$$

Compared to the autarkic equilibrium, the expected profit from entering the market is higher because the firms retain the payoff from serving the home market,

Figure 1: Firm's optimal actions for any set $\{a^h, a^f\}$



plus get profits from serving the foreign market. As a result, more firms will now enter the market. Thus if firms are allowed to export their products to a foreign market, the number of firms, n_e, producing in the home market in equilibrium will be higher. As well, the cut off a_p^h below which the firm will not produce, is higher in the export equilibrium. Ex-ante a firm is less likely to produce upon entering the home country, but if it does produce, its expected profits are higher.

3.3 Benchmark Model

In the benchmark model production in both the home and foreign countries is allowed. Ex-ante the home country is operating in an export equilibrium in industry m, and there are n_e firms producing and selling their good in the home market. There are no firms in industry m producing in the foreign country although those producing at home may be exporting their goods to the foreign market.

Given these actions and the firm's profit function (equations (3.5) and (3.6)), we can determine the optimal actions for any set of productivities $\{a^h, a^f\}$. Throughout the market I assume that the foreign market is sufficiently small[4], it may be profitable to produce in the foreign market and not serve the foreign market. For example, Banana Republic makes many of its items in the Philippines, but does not sell its goods in the Philippines.

Looking at figure 1, the optimal actions for each possible set of productivities are outlined. The set of productivities for which each of the above actions are optimal can be found in 3.6. If the firm is relatively unproductive ($\{a^h, a^f\}$ are sufficiently close to zero) then the firm will choose not to produce in either country. As a^h increases the firm chooses to produce in the home country, serving the home market, and as a^h continues to increase, they will produce in the home market and export to the foreign market. If a^f is increasing, and the foreign market is sufficiently small, the firm may go from not producing to producing in the foreign market and only serving the home market, to producing in the foreign market and serving both markets. If both home productivity and foreign productivity are sufficiently large the firm may choose to produce in both the home and foreign markets, serving their respective markets.

[4] $A^f < A^h \tau^{1-\varepsilon} \frac{f_x}{f_p + f_x}$

3.3.1 Foreign Direct Information - Complete Information

The complete information case is when, once a firm has drawn their home productivity, they know their foreign productivity. Throughout this section it is assumed that the foreign market is smaller than the home market, in order to allow for outsourcing in equilibrium.

Once a firm has entered the home market, their expected returns from their various actions are as follows:

- b_0 = do not produce

$$\text{Profit} = 0$$

- $b_{h,h}$ = produce in the home country and serve the home market only

$$\pi^{hh} - f_x - f_p = (1-\alpha)\alpha^{\varepsilon-1}A^h(a^h)^{\varepsilon-1} - f_x - f_p \tag{3.13}$$

- $b_{h,b}$ = produce in the home country and serve both the home and foreign markets

$$\pi^{hh} + \pi^{hf} - 2f_x - f_p = (1-\alpha)\alpha^{\varepsilon-1}(A^h + \tau^{1-\varepsilon}A^f)(a^h)^{\varepsilon-1} - 2f_x - f_p \tag{3.14}$$

- $b_{f,h}$ = produce in the foreign country and serve the home market

$$\pi^{fh} - f_x - f_p = (1-\alpha)\alpha^{\varepsilon-1}\tau^{1-\varepsilon}A^h(a^f)^{\varepsilon-1} - f_x - f_p \quad (3.15)$$

- $b_{f,b}$ = produce in the foreign country and serve both the home and foreign markets

$$\pi^{fh} + \pi^{ff} - 2f_x - f_p = (1-\alpha)\alpha^{\varepsilon-1}(\tau^{1-\varepsilon}A^h + A^f)(a^f)^{\varepsilon-1} - f_x - f_p \quad (3.16)$$

- $b_{b,b}$ = produce in the home country and the foreign country to serve their respective markets

$$\pi^{hh}+\pi^{ff}-2(f_p+f_x) = (1-\alpha)\alpha^{\varepsilon-1}[A^h(a^h)^{\varepsilon-1}+A^f(a^f)^{\varepsilon-1}]-2(f_p+f_x) \quad (3.17)$$

Helpman, Melitz and Yeaple (2004) find the FDI equilibrium if firms have the same productivity in both countries. With differences in productivity between the home and foreign countries, under certain conditions the firm might choose to produce in the foreign country to serve both the home and foreign markets ('outsourcing').

If the firm has the same productivity in the home and foreign country, we find the same equilibrium outcome as Helpman et al. Above some minimum productivity,

a_p^h, firms will choose to produce in the home country and serve the home country only.

$$(a_p^h)^{\varepsilon-1} = \frac{f_p + f_x}{A^h(1-\alpha)\alpha^{\varepsilon-1}} \tag{3.18}$$

As the productivity increases, at a_x^h, provided transport costs are sufficiently low, the firm will choose to export to the foreign country.

$$(a_x^h)^{\varepsilon-1} = \frac{f_x}{A^f \tau^{1-\varepsilon}(1-\alpha)\alpha^{\varepsilon-1}} \tag{3.19}$$

Finally, above some productivity level, a_b^h the firm will produce in both the home and foreign countries.[5]

$$(a_b^h)^{\varepsilon-1} = \frac{f_p}{A^f(1-\tau^{1-\varepsilon})(1-\alpha)\alpha^{\varepsilon-1}} \tag{3.20}$$

If the firm's foreign productivity is a linear transformation of their home productivity, $a^f = ca^h$, we can see the effects of differences in wage, or in human capital levels. If the number of workers required to make one unit of output is the same in both the home and foreign countries, then the only differentiating factor would be the wage and $c = \frac{w^h}{w^f}$ for all values of a^h. If there were factor price equalization, where wages counteracted any difference in output per unit of labor, c would equal one, and we would have the Helpman et al. equilibrium.

[5]assuming $\frac{f_x}{f_p} < \frac{\tau^{1-\varepsilon}}{(1-\tau^{1-\varepsilon})}$

The concept of comparative advantage suggests that relative productivities across industries may differ from country to country. We could then denote the industry specific productivity difference as c_m. If the home country had a comparative advantage in industry m and the foreign country had a comparative advantage in industry n then $c_n > c_m$.

Looking at a firm in industry m (but dropping the subscript) with both the large foreign market and the small foreign market, as c grows the firm becomes increasingly likely to engage in FDI. However, because we assume that fixed costs are the same in both countries the firm's actions depend largely on the relationship between the conversion and shipping costs.

If $c < \tau^{-1}$ no firms will produce in the foreign country. This is because the increased costs of production are greater than the savings from no longer shipping to the foreign market, even for the most productive firms.

For any values of $c \in [\tau^{-1}, \tau]$ the firm will decide between producing at home, and producing in the foreign country based on their productivity level at home. As their home productivity level increases they are more likely to produce in the foreign country. For every productivity a^h there is some c above which they undertake FDI, and below which they don't.

Finally, if $c > \tau$ all firms will produce in the foreign country. In this case the lower cost of production in the foreign country outweighs the cost of shipping the products from the foreign country to the home market even for the least productive firms.

More generally, c can be a function of a^h where $a^f = c(a^h) * a^h$ and is the firm's 'conversion parameter'. If the conversion parameter is sufficiently high for some firm the firm will undertake FDI and consumer welfare will weakly increase in both countries.

Using the export equilibrium as the base level of production, if FDI is undertaken, the conversion parameter, $c(a^h)$ must be sufficiently large that it is optimal for the firm to produce in the foreign country:

- If the firm would not produce in the home country, $a^h < a_p^h$:

$$[c(a^h)]^{\varepsilon-1} > \left[\frac{a_p^h}{a^h}\right]^{\varepsilon-1} \tag{3.21}$$

- If the firm would produce in the home country, but not export, $a^h \in [a_p^h, a_x^h]$

$$[c(a^h)]^{\varepsilon-1} > \min\{\tau^{\varepsilon-1}, \frac{A^h}{A^f + \tau^{1-\varepsilon}A^h}[1 + \frac{A^f}{A^h}\left[\frac{a_x^h}{a^h\tau}\right]^{\varepsilon-1}]\} \tag{3.22}$$

- If the firm would produce in the home country, and export to the foreign country, $a^h > a_x^h$

$$[c(a^h)]^{\varepsilon-1} > \min\{\frac{A^h + \tau^{1-\varepsilon}A^f}{A^f + \tau^{1-\varepsilon}A^h}, \tau^{\varepsilon-1}[1 - \frac{A^f}{A^h}\left[\frac{a_b^h}{a^h\tau}\right]^{\varepsilon-1}]\} \tag{3.23}$$

If any of the above equations hold, then at least one firm will enter the foreign country. In that case, the consumer welfare overall will improve. If no firms produce in the foreign country for the home market, consumer welfare in the home country may stay the same, but consumer welfare in the foreign country increases, as the price of some goods decrease (it must be cheaper for the firm to produce in the foreign market, than export from the home country to the foreign market). If firms undertake FDI for the sole purpose of shipping the goods to the home country, consumer welfare in the foreign country stays the same (although employment may increase), but consumer welfare in the home country increases as prices decrease.

In terms of comparative advantage and gains from trade, if the productivity gain is great enough, all firms in an industry will produce in the foreign country. Likewise, if there is no productivity gain all firms will produce in the home country, and only the most productive firms will produce in the foreign country to serve the foreign market (in order to save the shipping costs). Opening up the foreign economy to imports from the home country will result in more firms producing in the home country. Further opening up the foreign economy to FDI may result in additional firms producing, but where the firms choose to produce depends on the relative productivity of the foreign country to the home country.

3.4 Incomplete Information

In many cases a firm may not be able to infer their foreign productivity from their home productivity. Instead they may have some beliefs about the distribution of the productivity, where the cdf of their beliefs is $G(a^f|a^h)$ and $G'(a^f|a^h) = g(a^f|a^h)$. We assume that this cdf is common knowledge. The uncertainty is how well the firm's technology will adapt to the foreign country relative to another firm in the same industry. Some firms' technologies may adapt better, and yield a higher productivity than other firms. Here no learning would occur because the firm knows how well their industry will adapt in general, but does not know how well their firm-specific technology will adapt.

Assumption 3.4.1. *If $a_i^h > a_j^h$ then $G(a^f|a_i^h) < G(a^f|a_j^h)$.*

If firm i's productivity at home is higher than firm j's productivity at home, then firm i's expected productivity in the foreign country must first order stochastically dominate firm j's. So if a firm is less productive in the home country, ex-ante they expect to be less productive in the foreign country as well.

Now firms choose whether or not to invest in the foreign country based on their expected productivity in the foreign country and their beliefs about the distribution of that productivity. They must also choose whether or not to serve the foreign market when producing in the foreign country, and whether or not to close their home country production facilities when producing abroad. If a firm is already operating in the home country, their expected profits for each action are now a function of the expected distribution of productivity levels in the foreign country. Their expected profits from actions b_0, $b_{h,h}$, and $b_{h,b}$ remain the same, as they do not include producing in the foreign country. The expected profits from actions $b_{f,h}$, $b_{f,b}$ and $b_{b,b}$ are as follows:

- $b_{f,h}$ = produce in the foreign country and serve the home market

$$\pi^{fh} - f_x - f_p = (1-\alpha)(A^h \tau^{1-\varepsilon})\alpha^{\varepsilon-1} \int_{a^f} a^{\varepsilon-1} g(a|a^h)da - f_x - f_p \quad (3.24)$$

- $b_{f,b}$ = produce in the foreign country and serve both the home and foreign markets

$$\pi^{fh} + \pi^{ff} - 2f_x - f_p = (1-\alpha)\alpha^{\varepsilon-1}(A^h \tau^{1-\varepsilon} + A^f) \int_{a^f} a^{\varepsilon-1} g(a|a^h)da - 2f_x - f_p \quad (3.25)$$

- $b_{b,b}$ = produce in the home country and the foreign country to serve their respective markets

$$\pi^{hh} + \pi^{ff} - 2(f_x + f_p) = (1-\alpha)\alpha^{\varepsilon-1}[A^h (a^h)^{\varepsilon-1} + A^f \int_{a^f} a^{\varepsilon-1} g(a|a^h)da] - 2(f_x + f_p) \quad (3.26)$$

Choosing between b_0, $b_{h,h}$ and $b_{h,b}$ depend solely on the firm's home productivity (a^h) and exogenous variables, specifically the size of the home and foreign markets, and the fixed costs. The payoffs from actions $b_{f,h}$, $b_{f,b}$ and $b_{b,b}$ depend on the distribution of foreign productivities, $G(a^f|a^h)$.

Assuming that the fixed cost of drawing a productivity in the foreign country is the same as in the home country (f_e), a firm currently producing in the home market has two options: draw (and pay f_e) or don't draw and limit their action set to $\{b_0, b_{h,h}, b_{h,b}\}$. If the firm draws they can then choose from actions $\{b_0, b_{h,h}, b_{h,b}, b_{f,h}, b_{f,b}, b_{b,b}\}$. When deciding whether or not to draw the firm looks at the expected payoff from drawing (the probability of $b_{f,h}$, $b_{f,b}$ or $b_{b,b}$ being optimal multiplied by the payoff less the optimal action had the firm chosen not to draw the foreign productivity), and if it's higher than f_e the firm will draw. As the number of firms who draw their foreign productivity grows, the number of firms producing in the foreign country will also grow, and expected profits will decrease until the market is in equilibrium.

However, it's not necessarily true that firms with higher home productivity are more likely to draw their foreign productivity. If they are more productive at home the gains from drawing the foreign productivity may be lower, dependent upon the distribution $G(a^f|a^h)$ If the first moment of the distribution increases less than proportionally with home productivity ($E(a^f|a^h)$ is a concave function of a^h) then the more productive a firm is, the lower their expected profit from drawing a foreign

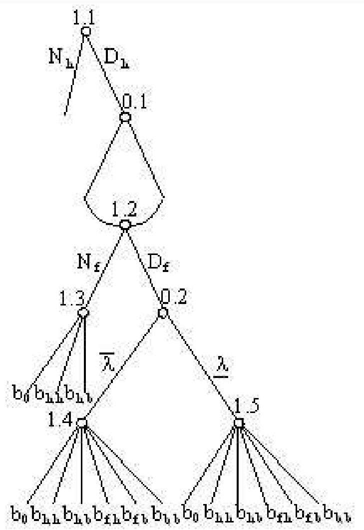


Figure 2: Game Tree for incomplete Information without learning
1.1 Firm decides whether or not to enter home country and draw home productivity (D_h), or not enter (N_h)
0.1 Nature determines firm's home productivity given that firm has entered
1.2 Firm decides to enter foreign country (D_f) or not (N_f)
1.3 Firm decides between not producing (b_0), producing in home country for home market (b_{hh}) or in home country for both markets (b_{hb})
0.2 Nature determines firm's foreign productivity shock, λ
1.4 Firm decides action given it's received a high productivity shock
1.5 Firm decides action given it's received a low productivity shock

productivity (because their opportunity cost is higher). As a result, the equilibrium could consist of only the firms that are less productive at home producing in the foreign country.

Using backwards induction, if a firm is not currently producing in the home country, the expected profit from drawing a home productivity level (given the set of all actions) must be weakly greater than that in the export equilibrium.

As a result, with incomplete information we would expect to see more firms active in equilibrium than in the export equilibrium. Given certain distributions, G, we could see fewer firms than in the complete information model, as the uncertainty may decrease the expected profits, leading fewer firms to draw their foreign productivity.

Some distributions G yield an equilibrium where only the more unproductive home firms will draw their foreign productivity, yielding a weakly less productive foreign market than a complete information equilibrium. In the next section the effects of incomplete information with a specific distribution of foreign productivities is outlined.

3.4.1 Firm-Level Uncertainty

In this section it is assumed that once firm i, draws their home productivity, a_i^h, they have the following expectation of their foreign productivity,

$$a^f = \begin{cases} \underline{\lambda} a^h & \text{with probability } \delta \\ \overline{\lambda} a^h & \text{with probability } 1-\delta \end{cases}$$

The distribution of foreign productivities can also be thought of as shocks to a firm's home productivity. They expect to receive a high shock with probability δ and a low shock with probability $1-\delta$.

To begin, it is assumed that there are no firms in the industry producing in the foreign country. For a firm that is not currently producing in the home country, the game tree is illustrated in figure 2. If the firm is currently producing in the home country, it is playing the subgame that begins at node 1.2.

Given the above setup, the game can be solved using backwards induction. At this point $\overline{\lambda}$ is restricted to be greater than τ, so that any firm that draws $\overline{\lambda}$ will never choose to produce in the home country. $\underline{\lambda}$ is restricted to be less than τ^{-1} so any firm that draws $\underline{\lambda}$ will never choose to produce in the foreign country. It is also assumed at this point that the foreign country is sufficiently small, as defined earlier.

Starting from node 1.5

At this point the firm has drawn the low productivity shock, $\underline{\lambda}$, resulting in a low relative productivity in the foreign country. Because it is assumed that $\underline{\lambda} < \tau^{-1}$ it will never be optimal for the firm to produce in the foreign country because it is less costly to produce the good in the home country and ship it to the foreign country than to produce in the foreign country.

Therefore if a firm draws a low productivity shock it will act as in the export equilibrium. It will not produce given a sufficiently low home productivity, if $a^h < a_p^h$. If they have a slightly higher home productivity, specifically if $a^h \in [a_p^h, a_x^h]$, they will produce in the home country just for the home market. Finally, given a sufficiently high home productivity, $a^h > a_x^h$, the firm will produce in the home country for both the home and foreign markets.

From node 1.4

At this node the firm has drawn the high productivity shock, $\overline{\lambda}$, resulting in a high relative productivity in the foreign country. Since $\overline{\lambda} > \tau$ it is less costly for the firm to produce in the foreign country, and ship to the home market, than to produce in the home country. As a result, their optimal actions can be limited to $b_{f,h}$, $b_{f,b}$ and

$b_{b,b}$. Again, we can find cutoffs in their home productivity that determine which action they will choose.

More specifically, if

$$a^h < \frac{\tau a_p^h}{\overline{\lambda}} \tag{3.27}$$

then the firm will choose to not produce in the foreign country (action b_0), even though they've received a high productivity shock.

If the firm's home productivity is higher, they may choose to produce in the foreign country, and export the good to the home market, not serving the foreign market (action $b_{f,h}$). The range of home productivities for which this would be optimal are as follows:

$$a^h \in \left[\frac{\tau a_p^h}{\overline{\lambda}}, \frac{a_x^h}{\tau \overline{\lambda}}\right] \tag{3.28}$$

Finally, given a sufficiently high home productivity,

$$a^h > \frac{a_x^h}{\tau \overline{\lambda}} \tag{3.29}$$

The firm will choose to serve the foreign and home markets while producing in the foreign country (action $b_{f,b}$)

From node 1.3

If the firm chooses not to draw its foreign productivity, it cannot produce in the foreign country. Therefore it can only choose whether or not to produce at home, and which markets to serve. The home productivity cutoffs, and optimal actions are then the same as at node 1.5.

From node 1.2

The probability of a high productivity shock is δ, leading to the following expected returns for specific ranges of a^h. The optimal actions at nodes 1.4 and 1.5 are listed as $\{b(1.4), b(1.5)\}$

i. $a^h < \frac{\tau}{\overline{\lambda}} a_p^h$

$$E[\Pi|D_f, a^h] = -f_e \tag{3.30}$$

$\{b_0, b_0\}$The firm's home productivity is sufficiently low such that the firm will not produce in the foreign market, even if they receive a high productivity shock.

ii. $a^h \in \left[\frac{\tau}{\overline{\lambda}} a_p^h, a_p^h\right]$

$$E[\Pi|D_f, a^h] = (1-\delta)[(\overline{\lambda} a^h)^{\varepsilon-1} \alpha^{\varepsilon-1} (1-\alpha) \tau^{1-\varepsilon} A^h - f_p - f_x] - f_e \tag{3.31}$$

$\{b_0, b_{f,h}\}$ Here the firm's home productivity is too low to warrant production in the home market, but if they received a high productivity shock they would produce in the foreign country.

iii. $a^h \in \left[a_p^h, \frac{1}{\tau}a_x^h\right]$

$$E[\Pi|D_f, a^h] = \delta(a^h)^{\varepsilon-1}\alpha^{\varepsilon-1}(1-\alpha)A^h+$$
$$(1-\delta)(\overline{\lambda}a^h)^{\varepsilon-1}\alpha^{\varepsilon-1}(1-\alpha)\tau^{1-\varepsilon}A^h - f_p - f_x - f_e \quad (3.32)$$

$\{b_{h,h}, b_{f,h}\}$ In this case the firm would produce in the home country if they received a low productivity shock, and produce in the foreign country if they received a high productivity shock, but in both cases they would just serve the home market.

iv. $a^h \in [\frac{1}{\tau}a_x^h, a_x^h]$

$$E[\Pi|D_f, a^h] = \delta(a^h)^{\varepsilon-1}\alpha^{\varepsilon-1}(1-\alpha)A^h$$
$$+ (1-\delta)[(\overline{\lambda}a^h)^{\varepsilon-1}\alpha^{\varepsilon-1}(1-\alpha)(A^f + \tau^{1-\varepsilon}A^h) - f_x] - f_p - f_x - f_e \quad (3.33)$$

$\{b_{h,h}, b_{f,b}\}$The firm's home productivity is now high enough that if they were to receive a high productivity shock they would produce in the foreign country and serve both the home and foreign markets, but if they receive a low productivity shock, they would produce in the home country, only serving the home market

v. $a^h > a^h_x$

$$E[\Pi|D_f, a^h] = \delta[(a^h)^{\varepsilon-1}\alpha^{\varepsilon-1}(1-\alpha)(A^h + \tau^{1-\varepsilon}A^f)]$$
$$+ (1-\delta)[(a^h\overline{\lambda})^{\varepsilon-1}\alpha^{\varepsilon-1}(1-\alpha)(\tau^{1-\varepsilon}A^h + A^f)] - f_p - 2f_x - f_e \quad (3.34)$$

$\{b_{h,b}, b_{f,b}\}$ Finally, given a sufficiently high home productivity, the firm would produce in the home country if it received a low productivity shock, and in the foreign country if it received a high productivity shock, in both cases serving both markets.

At this point the firm's optimal actions if they don't enter the foreign country, and their expected return if they do enter have been determined, and the next step is determining the conditions under which entering the foreign country is optimal. In (i.) it is never optimal for the firm to enter because they will never choose to produce. For ($ii.-v.$), given the cutoffs listed above, the following conditions are the incentive compatibility constraints for a firm, within it's respective range of home productivities, to produce:

$$ii. a^h > a_p^h \left[\frac{f_e}{(1-\delta)(f_p+f_x)} + 1 \right]^{\frac{1}{\varepsilon-1}} \frac{\tau}{\overline{\lambda}} \tag{3.35a}$$

$$iii. (a^h)^{\varepsilon-1} > (a_p^h)^{\varepsilon-1} \left[\frac{f_e}{(f_x+f_p)(1-\delta)} \right] \left[\frac{\tau^{\varepsilon-1}}{\overline{\lambda}^{\varepsilon-1} - \tau^{\varepsilon-1}} \right] \tag{3.35b}$$

$$iv. (a^h)^{\varepsilon-1} > \frac{f_e + (1-\delta) f_x}{(1-\delta)\alpha^{\varepsilon-1}(1-\alpha)[\overline{\lambda}^{\varepsilon-1} A^f + \left[\left(\frac{\overline{\lambda}}{\tau} \right)^{\varepsilon-1} - 1 \right] A^h]} \tag{3.35c}$$

$$v. (a^h)^{\varepsilon-1} > \frac{f_e}{(1-\delta)\alpha^{\varepsilon-1}(1-\alpha)[\left[\left(\frac{\overline{\lambda}}{\tau} \right)^{\varepsilon-1} - 1 \right] A^h + \left[\overline{\lambda}^{\varepsilon-1} - \tau^{1-\varepsilon} \right] A^f]} \tag{3.35d}$$

Proposition 3.4.2. *If the distribution of shocks is constant across productivities, the expected return from entering the foreign country is monotonically increasing in productivity, a^h, and there is a unique productivity level, a_d^h above which entering is always optimal.*

Proof. As a^h increases, the benefit from entering and drawing their foreign productivity is increasing in each of the four equations. Starting with equation (3.35a), as a^h increases the benefit from drawing increases. When $a^h = a_p^h$ the firm is indifferent between *ii.* and *iii.*, which implies that the benefit from drawing must be the same. In equation (3.35b) the benefit is again increasing with a^h, and therefore must be higher than in case *ii.*. This continues for *iii.*, *iv.*, and *v.*, with the benefit from drawing monotonically increasing in a^h.

If the benefit from drawing is positive for the lowest home productivity, $\underline{a}^h$, it is positive for all $a^h > \underline{a}^h$, and all firms will optimally draw their foreign productivity.

If the benefit is negative for the lowest home productivity, $\underline{a}^h$, because the benefit of entering the foreign country is monotonically increasing, if the benefit is positive for some (a^h) there must be a unique productivity level a_d^h for which the benefit is positive for all $a^h > a_d^h$ and the benefit is negative for all $a^h < a_d^h$.

Therefore given the distribution of firms, there is a single crossing point a_d^h above which all firms will optimally enter the foreign country and draw a productivity, and below which no firm will enter the foreign country. $\square$

Each of the equations is decreasing in the probability of drawing a low productivity, δ. Therefore as δ increases, the productivity level at which entry is optimal strictly increases. As a result fewer firms will enter the country if there is more 'risk', when risk is defined as the probability of a low draw.

Proposition 3.4.3. *Incomplete information is costly in that the number of firms operating in the foreign country in the incomplete information equilibrium is less than in the complete information equilibrium.*

If there were complete information the LHS of the inequalities would be strictly greater (the $(1-\delta)$ would disappear) for firms that receive a high shock, and the cutoff above which a firm would enter would be strictly lower. Firms receiving a low shock will not change their actions, but there is less waste, because they will not have entered to draw a foreign productivity. The difference between the number of firms producing in the foreign country with incomplete information, and that with complete information is the distortionary effect of incomplete information on FDI.

This information cost can be eliminated by subsidizing firms' cost of entry, f_e up until the point where the cutoff with complete information is the same as that with incomplete information. There would be no adverse selection problem because the higher the firm's productivity, the more incentive there is to enter.

3.4.2 Industry Level Uncertainty

If the uncertainty is about how well the industry-specific technology will adapt rather than the firm-specific technology, then learning can occur. In this case the firm expects the relative productivities among firms within the industry to stay the same, but does not know the productivity of the industry in general.

$$a^f = \begin{cases} \overline{\mu} a^h & \text{with probability } 1-\gamma \\ \underline{\mu} a^h & \text{with probability } \gamma \end{cases}$$

Once a firm has drawn its foreign productivity the other firms can see whether they decide to produce in the foreign country or not. If the firm decides to produce in the foreign country the other firms can deduce that the true $\hat{\mu} = \overline{\mu}$. If the firm decides not to produce in the foreign country, the other firms know the true $\hat{\mu} = \underline{\mu}$.

However, it takes the other firms time to observe and copy the first mover(s). In this time the first firm(s) are able to extract an additional rent as a result of moving first if the true $\hat{\mu} = \overline{\mu}$. Once again it is assumed that $\overline{\mu} > \tau$, $\underline{\mu} < \tau^{-1}$ and the foreign country is sufficiently small.

Now the firms can gain from waiting, so the choice is no longer enter or don't enter, it's now enter today, wait and enter tomorrow, or don't enter.

In this case the Nash equilibrium concept is used to define the equilibria. The set of equilibria are constrained to be pure strategy equilibria.[6]

The timing is as follows. In period $t = 0$ each firm i decides whether or not to enter the foreign country. Ex ante they have the beliefs about their industry's productivity as given above.

If they enter and draw their foreign productivity, they can then choose whether or not to produce in the foreign country. In period $t = 1$ every firm sees the action undertaken by every other firm and can deduce their own foreign productivity if at least one firm chooses to enter in period $t = 0$.

If no firm chooses to enter in $t = 0$, the game at $t = 1$ would be the same as at $t = 0$. If all firms choose not to enter in period $t = 0$ their decision is the same in $t = 1, 2, 3, ...$, so the equilibria focussed on are those where at least one firm enters in $t = 0$ and where no firm enters in any period $t \geq 0$.

[6]The focus is on the conditions that lead to a firm entering the foreign country. Any conditions that support a mixed strategy equilibrium will also support a pure strategy equilibrium, so although I recognize the existence of mixed equilibria, they won't be outlined here

Once one firm has entered there is complete information, and all firms will act according to the equilibrium set out in section 3.1. Periods $t = 0$ and $t > 0$ are weighted by discount factor ρ, the profit to the firm is a weighted average of profits in $t = 0$ $(1 - \rho)$ and profits in $t > 0$ (ρ).

There are two incentive compatibility constraints that help define the possible equilibria:

1) If no other firms are entering, firm i will enter if and only if

$$
\begin{aligned}
(1-\gamma)\beta^h E^h & \left[\frac{(1-\rho)\overline{\mu}^{\varepsilon-1}\tau^{1-\varepsilon}}{P_x^{fh}} + \frac{\rho\overline{\mu}^{\varepsilon-1}\tau^{1-\varepsilon}}{P_e^{fh}}\right] \\
& + (1-\gamma)\beta^f E^f \left[\frac{(1-\rho)\overline{\mu}^{\varepsilon-1}}{P_x^{ff}} + \frac{\rho\overline{\mu}^{\varepsilon-1}}{P_e^{ff}}\right] - \frac{(a_i^h)^{1-\varepsilon} f_e}{\alpha^{\varepsilon-1}(1-\alpha)} \\
& > (1-\gamma)\beta^h E^h \frac{1}{P_x^{hh}} + (1-\gamma)\beta^f E^f \frac{\tau^{1-\varepsilon}}{P_x^{hf}} \qquad (3.36)
\end{aligned}
$$

The LHS is the gain from being a first mover, it's the rent the firm can get from entering and being able to produce for a period of time before other firms can follow, given that $\mu = \overline{\mu}$, plus the future gain, since firm i is the only one entering, less the fixed costs of entering. The RHS is what the firm would have gotten had they not entered, which is just the export equilibrium profits (since it's assumed that if no firms enter in $t = 0$ no firms will enter in $t = 1$). $P_j^{dc} = \sum_{v \in V^c} p^c(v)^{1-\varepsilon}$ is the sum of the prices of varieties distributed in country c when the firm is producing in country d, for scenario j.[7]

$$[a_i^h]^{\varepsilon-1}\left[\beta^h E^h\left[\frac{\overline{\mu}^{\varepsilon-1}}{\tau^{\varepsilon-1}P_d^{fh}}-\frac{a}{f}\right]\right] \tag{3.37}$$

$$(a_i^h)^{\varepsilon-1} > \frac{f_e}{\alpha^{\varepsilon-1}(1-\alpha)(1-\gamma)}$$
$$\left[\beta^h E^h\left(\left(\frac{\overline{\mu}}{\tau}\right)^{\varepsilon-1}\left(\frac{1-\rho}{P_x^{fh}}+\frac{\rho}{P_e^{fh}}\right)-\frac{1}{P_x^{hh}}\right)+\beta^f E^f\left(\overline{\mu}\left(\frac{1-\rho}{P_x^{ff}}+\frac{\rho}{P_e^{ff}}\right)-\frac{\tau^{1-\varepsilon}}{P_x^{hf}}\right)\right]^{-1} \tag{3.38}$$

ρ is a function of how much of a lead the first mover gets on the rest of the firms. As the period of time increases, before which the following firms can enter, the expected profit increases. In (3.36) this is seen by $(1-\rho)$ increasing, which increases the LHS of the equation. Therefore, although a longer lead time for the first mover is detrimental to the consumers, because there are fewer varieties available for longer, it encourages firms to enter the foreign market.

It may be the case that even if another firm is entering in the first period, it is still optimal for firm i to enter. This is seen in the following equation:

[7] $j \in \{x, e, d\}$ where x is the export equilibrium, e is the FDI with complete information equilibrium, and d is the strategic scenario where more than one firm enters the country in t=0

$$(1-\gamma)\beta^h E^h \frac{\overline{\mu}^{\varepsilon-1}\tau^{1-\varepsilon}}{P_d^{fh}} + (1-\gamma)\beta^f E^f \frac{\overline{\mu}^{\varepsilon-1}}{P_d^{ff}} - \frac{(a_i^h)^{1-\varepsilon} f_e}{\alpha^{\varepsilon-1}(1-\alpha)}$$
$$> (1-\gamma)\beta^h E^h \frac{1}{P_d^{hh}} + (1-\gamma)\beta^f E^f \frac{\tau^{1-\varepsilon}}{P_d^{hf}} \quad (3.39)$$

If (3.39) holds for firm i it means that firm i would optimally enter in $t = 0$ even if other firm(s) choose to enter in $t = 0$. Here the gains from being a first mover must outweigh the benefit of waiting. The LHS of (3.39) is the rent from being a first mover, as opposed to waiting (RHS). Again as a_i^h increases this equation becomes less binding.

From here we can determine the possible equilibria. If equation (3.36) is positive only for the most productive firm, the equilibria will consist of only that firm entering in period $t = 0$. If it's not positive for any firm, then no firms will choose to enter the foreign country.

Letting j denote the most productive firm, if (3.39) is negative for any $i \neq j$ given that only j is entering in $t = 0$, only one firm will choose to enter in equilibrium. However, the firm that enters is not necessarily firm j, it could be a less productive firm, because if firm i enters, it may not be optimal for firm j to enter. The longer the lead time for the first mover, the more likely a firm is to enter, and the higher the number of firms that will invest in $t = 0$. However, a longer lead time also delays the equilibrium with the highest number of varieties available, and may decrease consumer welfare.

Proposition 3.4.4. *As γ, the probability of drawing a low productivity, increases, the expected return from entering decreases. There exists some $\gamma*$ such that for any $\gamma < \gamma*$ entry is optimal and for any $\gamma > \gamma*$ no firms will enter.*

As entering becomes more 'risky' (the probability of drawing a low productivity increases), the expected return from entering decreases. We can therefore find the appropriate level of risk above which no firms will enter the market.

Proof. The relevant equation for entry is Equation (3.36). If this equation is met entry will occur in equilibrium. As γ increases the LHS strictly decreases relative to the RHS. Assuming entry is optimal when $\gamma = 0$, and that the cost of entry is strictly positive, we can find the level of risk such that the benefits of entry equal the costs. □

3.5 Conclusion

Firms conduct FDI for many reasons. It may be to take advantage of an emerging market when shipping costs are high, or to take advantage of lower costs and outsource.

While firms that are heterogeneous in terms of productivity levels have been widely studied, differences in firm-level productivity between countries has been overlooked. In this paper three models of heterogeneity both between firms and between countries are outlined.

In the first I look at the complete information model, and find that given market size and preferences the relation of foreign to home productivities across productivity levels largely determines the FDI equilibrium. If the foreign productivity is linearly related to home productivity, only the more productive firms conduct FDI.

In the second model I look at the effect of firm-level uncertainty in productivity. There are no externalities from entry, and an inefficient level of FDI is undertaken. This explains why FDI may be underprovided in many countries where it seems there may be a cost advantage. Policy implications suggest subsidy of entry costs may allow the industry to reach its efficient level.

Finally I look at the effect of industry-level uncertainty, where once one firm enters, all firms know their foreign productivity. Once entry has occurred the industry reaches its efficient level, but because firms can not internalize the externalities of early entry, entry may not occur. In this case one would expect clusters to emerge as industries which have experienced entry will grow quickly, while other industries may not emerge at all if entry is too expensive. In this case it can be argued that government should subsidize infant-industries, or first movers.

If one expects that less developed countries are more 'risky' (the probability of drawing a high productivity are lower), underinvestment in FDI and in the industries that do enter, more clusters can be expected relative to developed countries.

Future work includes data analysis on the export level, which will help substantiate the claims made by the model.

3.6 Supplement: Optimal Actions with Complete Information

- b_0 is optimal if

$$(a^h)^{\varepsilon-1} \in \left[0, \frac{f_p + f_x}{A^h(1-\alpha)\alpha^{\varepsilon-1}}\right] \quad (a^f)^{\varepsilon-1} \in \left[0, \frac{f_p + f_x}{A^h\tau^{1-\varepsilon}(1-\alpha)\alpha^{\varepsilon-1}}\right] \tag{3.40}$$

- $b_{h,h}$ is optimal if

$$\begin{aligned} &(a^h)^{\varepsilon-1} \in \left[\frac{f_p + f_x}{A^h(1-\alpha)\alpha^{\varepsilon-1}}, \frac{f_x}{A^f\tau^{1-\varepsilon}(1-\alpha)\alpha^{\varepsilon-1}}\right] \\ &(a^f)^{\varepsilon-1} \in \left[0, \min\{(a^h\tau)^{\varepsilon-1}, (a^h)^{\varepsilon-1}\left[\frac{A^h}{A^f + \tau^{1-\varepsilon}A^h}\right] + \frac{f_x}{(A^f + \tau^{1-\varepsilon}A^h)(1-\alpha)\alpha^{\varepsilon-1}}\}\right] \end{aligned} \tag{3.41}$$

- $b_{h,b}$ is optimal if

$$(a^h)^{\varepsilon-1} \in \left[\frac{f_x}{A^f\tau^{1-\varepsilon}(1-\alpha)\alpha^{\varepsilon-1}}, \infty\right] \tag{3.42}$$

$$(a^f)^{\varepsilon-1} \in \left[0, min\{(a^h)^{\varepsilon-1}\frac{A^h+\tau^{1-\varepsilon}A^f}{A^f+\tau^{1-\varepsilon}A^h}, \left[\frac{a^h}{\tau}\right]^{\varepsilon-1} + \frac{f_p}{A^f(1-\alpha)\alpha^{\varepsilon-1}}\}\right] \tag{3.43}$$

- $b_{f,h}$ is optimal if

$$(a^h)^{\varepsilon-1} \in \left[0, (a^f)^{\varepsilon-1}\tau^{1-\varepsilon}\right] \quad (a^f)^{\varepsilon-1} \in \left[\frac{f_p+f_x}{A^h\tau^{1-\varepsilon}(1-\alpha)\alpha^{\varepsilon-1}}, \frac{f_x}{A^f(1-\alpha)\alpha^{\varepsilon-1}}\right] \tag{3.44}$$

- $b_{f,b}$ is optimal if

$$(a^h)^{\varepsilon-1} \in [0, min\{(a^f)^{\varepsilon=1}\left[\frac{A^f + \tau^{1-\varepsilon}A^h}{A^h}\right] - \frac{f_x}{A^h(1-\alpha)\alpha^{\varepsilon-1}}, \tag{3.45}$$

$$(a^f)^{\varepsilon-1}\frac{A^f + \tau^{1-\varepsilon}A^h}{A^h + \tau^{1-\varepsilon}A^f}, (a^f)^{\varepsilon-1}\tau^{1-\varepsilon} + \frac{f_p}{A^h(1-\alpha)\alpha^{\varepsilon-1}}\}] \tag{3.46}$$

$$(a^f)^{\varepsilon-1} \in [max\{(a^h)^{\varepsilon-1}\left[\frac{A^h}{A^f + \tau^{1-\varepsilon}A^h}\right] + \frac{f_x}{(A^f + \tau^{1-\varepsilon}A^h)(1-\alpha)\alpha^{\varepsilon-1}}, \tag{3.47}$$

$$\frac{f_x}{A^f(1-\alpha)\alpha^{\varepsilon-1}}, (a^h\tau)^{\varepsilon-1}\frac{A^h + \tau^{1-\varepsilon}A^f}{A^f + \tau^{1-\varepsilon}A^h} - \frac{f_p\tau^{\varepsilon-1}}{A^h(1-\alpha)\alpha^{\varepsilon-1}}\}, \infty] \tag{3.48}$$

- $b_{b,b}$ is optimal if

$$(a^h)^{\varepsilon-1} \in [(a^f)^{\varepsilon-1}\tau^{1-\varepsilon} + \frac{f_p}{A^h(1-\alpha)\alpha^{\varepsilon-1}}, \infty] \tag{3.49}$$

$$(a^f)^{\varepsilon-1} \in [(a^h)^{\varepsilon-1}\tau^{1-\varepsilon} + \frac{f_p}{A^f(1-\alpha)\alpha^{\varepsilon-1}}, \infty] \tag{3.50}$$

Bibliography

Adam, C. S. and J. W. Gunning (2002, December). Redesigning the Aid Contract: Donors' Use of Performance Indicators in Uganda. *World Development 30*(12), 2045–2056.

Admati, A. R. and M. Perry (1991, April). Joint Projects without Commitment. *Review of Economic Studies 58*(2), 259–76.

Atkeson, A. (1991, July). International Lending with Moral Hazard and Risk of Repudiation. *Econometrica 59*(4), 1069–89.

Benassy-Quere, A., M. Coupet, and T. Mayer (2005, April). Institutional determinants of foreign direct investment. Working Papers 2005-05, CEPII research center.

Bulow, J. and K. Rogoff (2005, May). Grants versus Loans for Development Banks. *American Economic Review 95*(2), 393–397.

Cushman, D. O. (1985, May). Real exchange rate risk, expectations, and the level of direct investment. *The Review of Economics and Statistics 67*(2), 297–308.

Freixas, X., R. Guesnerie, and J. Tirole (1985, April). Planning under incomplete information and the ratchet effect. *Review of Economic Studies 52*(2), 173–91.

Goldberg, L. S. and C. D. Kolstad (1994, August). Foreign direct investment, exchange rate variability and demand uncertainty. NBER Working Papers 4815, National Bureau of Economic Research, Inc.

Hausmann, R. and D. Rodrik (2003, December). Economic development as self-discovery. *Journal of Development Economics 72*(2), 603–633.

Helpman, E., M. J. Melitz, and S. R. Yeaple (2004, March). Export versus FDI with Heterogeneous Firms. *American Economic Review 94*(1), 300–316.

Hirschman, A. O. (1967). *Development Projects Observed*. Washington, DC: The Brookings Institute.

Horstmann, I. J. and J. R. Markusen (1989, February). Firm-specific assets and the gains from direct foreign investment. *Economica 56*(221), 41–48.

Laffont, J.-J. and J. Tirole (1986, June). Using cost observation to regulate firms. *Journal of Political Economy 94*(3), 614–41.

Levin, J. (2003, June). Relational Incentive Contracts. *American Economic Review 93*(3), 835–857.

Lucas, Robert E, J. (1990, May). Why Doesn't Capital Flow from Rich to Poor Countries? *American Economic Review 80*(2), 92–96.

Ma, C.-t. A. (1991, February). Adverse selection in dynamic moral hazard. *The Quarterly Journal of Economics 106*(1), 255–75.

Malcomson, J. M. and F. Spinnewyn (1988, July). The Multiperiod Principal-Agent Problem. *Review of Economic Studies 55*(3), 391–407.

Marx, L. M. and S. A. Matthews (2000, April). Dynamic Voluntary Contribution to a Public Project. *Review of Economic Studies 67*(2), 327–58.

Matsuyama, K. (1996, August). Why are there rich and poor countries?: Symmetry-breaking in the world economy. NBER Working Papers 5697, National Bureau of Economic Research, Inc.

Melitz, M. J. (2003, November). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica 71*(6), 1695–1725.

Milton, H. and R. M. Townsend (1981, January). Resource allocation under asymmetric information. *Econometrica 49*(1), 33–64.

Razin, A., E. Sadka, and H. Tong (2005, September). Bilateral FDI Flows: Threshold Barriers and Productivity Shocks. NBER Working Papers 11639, National Bureau of Economic Research, Inc.

Rothschild, M. and J. E. Stiglitz (1970, September). Increasing risk: I. a definition. *Journal of Economic Theory 2*(3), 225–243.

Sanford, J. E. (2004, September). IDA Grants and HIPC Debt Cancellation: Their Effectiveness and Impact on IDA Resources. *World Development 32*(9), 1579–1607.

Sappington, D. (1983, February). Limited liability contracts between principal and agent. *Journal of Economic Theory 29*(1), 1–21.

Spear, S. E. and S. Srivastava (1987, October). On Repeated Moral Hazard with Discounting. *Review of Economic Studies 54*(4), 599–617.

Steele, J. L. (2007). The optimal sequencing of carrots: The timing and size of aid grants for large projects. *Mimeo*.

Varian, H. R. (1994, February). Sequential contributions to public goods. *Journal of Public Economics 53*(2), 165–186.

Watson, J. (2002, January). Starting Small and Commitment. *Games and Economic Behavior 38*(1), 176–199.

Vita

Jennifer Lynn Steele

Permanent Address: 9231 Arrowsmith Drive
Richmond, BC
Canada V7A 4Z5

This dissertation was typeset with LaTeX 2_ε[8] by the author.

[8] LaTeX 2_ε is an extension of LaTeX. LaTeX is a collection of macros for TeX. TeX is a trademark of the American Mathematical Society. The macros used in formatting this dissertation were written by Dinesh Das, Department of Computer Sciences, The University of Texas at Austin, and extended by Bert Kay, James A. Bednar, and Ayman El-Khashab.