ECONOMIC GROWTH CENTER

YALE UNIVERSITY

P.O. Box 208629
New Haven, CT 06520-8269
http://www.econ.yale.edu/~egcenter/

CENTER DISCUSSION PAPER NO. 924

Coordinating Development: Can Income-based Incentive Schemes Eliminate Pareto Inferior Equilibria?

Philip Bond
University of Pennsylvania

and

Rohini Pande
Yale University

October 2005

Notes: Center Discussion Papers are preliminary materials circulated to stimulate discussions and critical comments.

We are grateful to Dilip Mookherjee, Debraj Ray and two anonymous referees for detailed comments. We also thank Dino Gerardi, Stephen Morris and Andrew Newman for helpful comments. Bond thanks the Institute for Advanced Study for hospitality and financial support (in conjunction with Deutsche Bank) over the academic year 2002-03. Any errors remain ours.

Coordinating Development: Can Income-based Incentive Schemes Eliminate Pareto Inferior Equilibria?

Philip Bond and Rohini Pande

Abstract

Individuals' inability to coordinate investment may significantly constrain economic development. In this paper we study a simple investment game characterized by multiple equilibria and ask whether an income-based incentive scheme can uniquely implement the high investment outcome. A general property of this game is the presence of a crossover investment point at which an individual's incomes from investment and non-investment are equal. We show that arbitrarily small errors in the government's knowledge of this crossover point can prevent unique implementation of the high investment outcome. We conclude that informational requirements are likely to severely limit a government's ability to use income-based incentive schemes as a coordination device.

Keywords: Coordination, Public Policy, Income Taxation, Implementation

JEL codes: O21, H23

1 Introduction

Whenever investment decisions are decentralized but individual returns to investment remain dependent on the choices made by others, economic activity may be limited by a coordination failure among the agents. This observation has often been taken to be particularly relevant for low-income countries. A large theoretical literature, starting with Rosenstein and Rodan (1943) and Hirschman (1957), has explored how coordination failures can cause poor countries and people to stay poor (see, e.g., Murphy, Shleifer and Vishny 1989, Matsuyama 1996 and Becker, Murphy and Tamura 1990).

If one takes seriously the view that economic growth is inhibited by coordination failures, then a natural question is whether suitably designed public policies can prevent such failures. Import substitution industrialization policies, for instance, are widely viewed as just such an attempt, although there is widespread scepticism about their success (Krueger 1997 and Bhagwati 1993). In general, however, discussions of whether or not a government can coordinate economic development have remained largely informal and have rarely progressed beyond the observation that a government's inability to directly observe individual investment decisions may prevent such coordination (Rodrik 1996, and again Krueger 1997). Put differently, the government's inability to observe individual actions makes coordination a multi-agent moral hazard problem.

However, as is now well known, the appropriate design of contracts can mitigate agency concerns in moral hazard problems. In the standard example, an employee is paid according to observable output. A suitably designed wage scheme will induce the agent to take the employer's preferred, though unobservable, action. In a public policy context (e.g., Mirrlees' 1971 seminal work on income taxation) the analogue is a tax and transfer policy which depends on observed outcomes (an individual's market income), but not on unobserved actions. In keeping with this line of research, this paper examines whether a tax and transfer policy can be used as a coordination device when the government does not observe individual actions.

In the canonical coordination game an agent's income depends both on her action and the actions of all other agents in the economy. As a result the game has multiple Pareto-ranked equilibria.[1] In this spirit, we consider an economy in which a continuum of agents must each decide whether or not to invest.[2] The returns to both investment and non-investment depend on the proportion of investors, which we denote as η. The social optimum is for all agents to invest. Individual incomes from non-investment exceed those from investment when $\eta = 0$, with the opposite true when $\eta = 1$. Hence, although it is an equilibrium for all agents to invest, it is also an equilibrium for no agent to invest. Moreover, there is an intermediate level of investment, $\tilde{\eta}$ say, at which investment and non-investment generate the same returns. We cast the problem of coordination of economic activity as a principal-agent problem in which a single principal — the government — seeks to provide multiple agents with the incentive to invest.

If the government can directly observe individual investments then coordinating the full-investment equilibrium is straightforward. Indeed, Rodrik (1996) has argued that information gathering by institutions such as the Japanese MITI was key to the success of the industrial policies pursued by East Asian governments. However, previous discussions of government coordination have also noted that, in practice, a government bureaucrat may find it difficult to distinguish between genuine investment and consumption disguised as investment. That is, a government's lack of information constrains its ability to foster coordination and hence economic development.

Our main results both reinforce and develop the view that information is critical to successful coordination — and that in practice, the informational hurdle is hard to clear. We start from the point where previous informal discussions end: is there anything a government can do if it *cannot* observe individual investment decisions? We allow a government to make transfers to individuals based on their incomes, and likewise to engage in income taxation.[3]

[1] Cooper (1999) reviews the coordination game literature, while Hoff (2000) and Ray (2000) discuss their relevance for modeling market outcomes in low income countries.

[2] Our description of the investment game is similar to those in "big push" models of development – industrialization is individually profitable only if a sufficiently large number of agents industrialize (see, for instance, Murphy, Shleifer and Vishny (1989)).

[3] Since we do not impose budget balancing, these two options allow for the same outcomes.

Given this, we ask whether there exists tax/transfer schemes such that all agents investing is the only stable equilibrium.

Recall that when $\tilde{\eta}$ individuals invest, investment and non-investment yield the same income. Investors and non-investors appear identical to the government at the investment level $\tilde{\eta}$; as such, there is necessarily an equilibrium at that point. Our main results relate to the difficulties this engenders for successful coordination. Specifically, it is reasonable to regard a government coordination policy as successful only if it leaves $\eta = 1$ as the only stable equilibrium. In other words, $\tilde{\eta}$ must be an *unstable* equilibrium when the policy is in place. To ensure this the government's policy must leave investors with higher post-tax/transfer income both when just more and just less than $\tilde{\eta}$ of the population invest.

Consider first the class of government policies in which each individual's transfer depends only on that individual's pre-transfer income. Coordination is only feasible if the government knows the *exact* value of $\tilde{\eta}$, along with the corresponding income at that point. We show (Proposition 2) that if instead it makes even an arbitrarily small mistake in identifying these values, then coordination fails. The reason is that the policy creates a stable equilibrium in the vicinity of $\tilde{\eta}$. To the extent small misperceptions are inevitable, we interpret this result as indicating that coordination using this class of policies is impossible.

Second, we consider a broader class of government policies in which each individual's transfer can also depend on the full distribution of incomes in the "economy." In this case coordination is often feasible. Specifically, if the externality associated with investment is sufficiently strong then a government can successfully use information on the proportion of investors at different income levels to ensure that post-transfer returns from investment always exceed those from non-investment. However, we believe that such policies are much harder for a government to successfully implement. In particular, the relevant "economy" is the set of individuals actually affected by the externalities in question. If the government fails to correctly identify this, then it likewise misidentifies the income distribution, and its coordination policy will fail. This problem is made even more relevant by the fact that investment games are typically played out at the sectoral level, and an economy consists of multiple sectors. Again, the informational requirements that a coordination policy must meet appear substantial.

Related literature

Our analytic approach is related to the literature on implementation in moral-hazard settings — see Ma (1988), Arya, Glover and Young (1995), Arya, Glover and Hughes (1997), and Chung (1999).[4] The first three of these papers address a problem noted by Mookherjee (1984): when a principal seeks to provide incentives to many agents, the optimal incentive scheme may possess multiple equilibria. Ma shows that, in general, the use of "integer" message games can eliminate the multiplicity of equilibria. The papers of Arya, Glover and their co-authors rule out the use of such mechanisms on grounds of realism, and show that in specific instances of the many-agent incentive problem unwanted equilibria can still be eliminated. Closest to us is Chung (1999) who examines whether a specific class of policies, affirmative action policies, can implement the unique efficient (human capital) investment equilibrium.

Paper outline

The remainder of the paper is structured as follows. Section 2 describes the economic environment. Section 3 identifies necessary features of a tax scheme which implements the high investment outcome. Section 4 considers implementation via policies based on individual income, and Section 5 policies which also condition on the income distribution. Section 6 discusses the robustness of our results under alternative assumptions and Section 7 concludes.

[4] The standard implementation problem concerns an adverse-selection setting (see Maskin 1999). That is, agents first learn their type, which a planner then seeks to induce them to report truthfully.

2 Environment

2.1 Basics

The economy consists of a continuum of risk-neutral agents. Every agent faces a discrete investment choice, $a \in \{0, 1\}$, where $a = 1$ is interpreted as investment. Investment choice a yields a monetary return of $y \equiv f_a(\eta)$, where η is the fraction of agents who invest.

We assume $f_a(\cdot)$ is a continuous, twice-differentiable function. An agent's decision to invest generates positive spill-overs for other investors, i.e., $f_1'(\cdot) > 0$. In contrast, $f_0'(\cdot) \gtrless 0$. However, $f_0'(.) = 0$ at only finitely many values of η, i.e., investment affects the returns to non-investment almost everywhere. The externalities investment generates for investors exceed those generated for non-investors. That is, $|f_1'(.)| > |f_0'(.)|$. F_η denotes the income distribution when a fraction η invest. The income distribution collapses to a one-point distribution whenever incomes from investment and non-investment coincide. Otherwise, it is a two-point distribution — a proportion $1 - \eta$ of the population has income $f_0(\eta)$ and a proportion η income $f_1(\eta)$.

A principal, who we refer to as the government, seeks to design a mechanism to affect agents' investment decisions. We restrict the government to transfer schemes $T(y)$ that determine the post-transfer income of an agent with pre-tax income y. We consider, in turn, the cases where T *only* depends on an agent's income, y, and where it can also depend on the full income distribution F. In both cases we assume T is a continuous function of y and F, in the following sense. Let $\mathcal{F}$ denote the set of distribution functions on the real line that place mass on at most two points. For any pair of sequences $\{y^n\} \subset \Re$ and $\{F^n\} \subset \mathcal{F}$ such that $y^n \to y$ and $F^n(\epsilon) \to F(x)$ for all x except the mass points of F, $T(y^n, F^n) \to T(y, F)$.[5] For the most part we restrict attention to mechanisms that do not make use of messages; in Section 6 we discuss more general mechanisms of this type.

We impose no other restrictions on T. In particular, we do not require budget balancing — there is no reason the investment externality should encompass the whole economy.

[5]That is, convergence in distribution (see, e.g., Billingsley 1995). We discuss the restriction to continuous tax/transfer functions T in Section 6.

Finally, when we refer to the economy absent government intervention, we simply mean $T(y, F) \equiv y$, i.e., no net taxes or transfers.

2.2 Equilibrium

Given a fraction η of investors, let $\Delta_T(\eta)$ denote the gain to an agent from investing instead of not investing :

$$\Delta_T(\eta) \equiv T(f_1(\eta), F_\eta) - T(f_0(\eta), F_\eta)$$

We assume that, absent government intervention, full investment ($\eta = 1$) maximizes social welfare, i.e.,

$$1 \in \arg\max_\eta \eta f_1(\eta) + (1 - \eta) f_0(\eta)$$

We define our concept of equilibrium and stability for this economic environment, and use these definitions to characterize a coordination game.

Definition 1 ***(Nash equilibrium)***
Given a government policy T, a fraction η of agents investing is a Nash equilibrium if and only if $\eta = 0$ and $\Delta_T(0) \leq 0$, OR *$\eta = 1$ and $\Delta_T(1) \geq 0$* OR *$\eta \in (0, 1)$ and $\Delta_T(\eta) = 0$.*

Throughout we refer to an equilibrium by the fraction that invest in that equilibrium, η.

Definition 2 ***(Stability)***
An equilibrium η is unstable if there exists δ such that either $\Delta_T(\eta - \varepsilon) < 0$ for all $\varepsilon \in (0, \delta)$, or $\Delta_T(\eta + \varepsilon) > 0$ for all $\varepsilon \in (0, \delta)$. An equilibrium that is not unstable is stable.

The notion of equilibrium stability consists of asking whether individuals' optimal investment strategies are immune to small perturbations of the proportion of investors around the presumed equilibrium, i.e., to investor noise. Below (see Proposition 2) we show that arbitrarily small government misperceptions of the economic fundamentals create new stable equilibria. The sizes of the basins of attraction of the newly created stable equilibria (i.e., the extent of investor noise) depend on the extent of government misperceptions. It

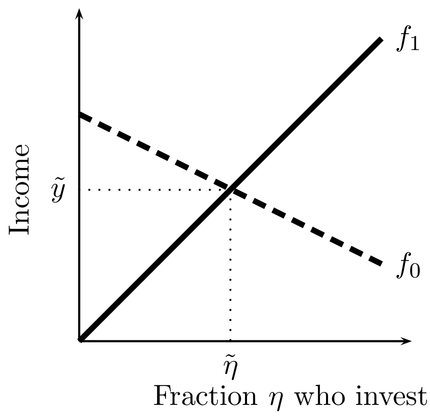

Figure 1: A linear example of a coordination game

follows that our definition of stability is premised on the assumption that investor noise is less than government noise. If we interpret the notion of equilibrium stability as reflecting the outcome of a learning process among investors (in the face of strategic uncertainty) then an economic interpretation of this assumption is that, relative to the government, investors are faster at 'learning' about their economic environment.

Since we have assumed that $f_1' > 0$ and $|f_1'| > |f_0'|$, the function Δ_T is strictly increasing absent government intervention (i.e., $T(y) \equiv y$). Consequently, absent intervention there are three possible equilibrium sets in our economy: $\eta = 0$ is the only stable equilibrium; $\eta = 1$ is only stable equilibrium; and $\eta = 0$ and $\eta = 1$ are both stable equilibria, with an unstable equilibrium at a unique investment level point $\tilde{\eta} \in (0, 1)$. For the remainder of the paper we assume that the last of these three cases holds, and refer to our economy as a *coordination game*. Figure 1 illustrates a coordination game. (In Section 7, we briefly discuss the problem of ensuring full investment when the only stable equilibrium absent intervention is $\eta = 0$. This case is, of course, a manifestation of the "free-rider" problem.[6])

[6]A simple example of the free-rider problem in our economy is $f_0(\eta) = 5/8 + \eta/2$ and $f_1(\eta) = \eta$. Investment has positive externalities for both investors and non-investors, the socially efficient investment level is $\eta = 1$, yet the only equilibrium is at $\eta = 0$.

Since investment and non-investment incomes coincide at the investment level $\tilde{\eta} \in (0, 1)$, necessarily $\Delta_T(\tilde{\eta}) = 0$ for all government policies T. That is, investment by a fraction $\tilde{\eta}$ of agents is an equilibrium. Given this, the most that can be hoped of a government policy is that it leave $\eta = 1$ as a stable equilibrium, but $\tilde{\eta}$ as an unstable equilibrium:

Definition 3 ***(Implementation)***
A policy T implements η if η is the only stable equilibrium.

Implementation when individual actions are publicly observable is straightforward: tax the socially inefficient action, or subsidize the socially efficient action.[7] More formally, consider the policy $T(a = 1, y, F) = y$ and $T(a = 0, y, F) = f_1(0) - \varepsilon$, some $\varepsilon > f_1(0) - y$. This policy only taxes non-investors, such that they prefer investing. Under this choice of T, full-investment is the only equilibrium. The remainder of the paper considers implementation when agents' incomes, but not actions, are observable.

3 Implementation

In selecting the parameters of the tax and transfer scheme the government's aim is to eliminate the no-investment equilibrium ($\eta = 0$), while preserving full investment as a stable equilibrium ($\eta = 1$). We start by noting:

Proposition 1 ***(Δ must be positive)***
Suppose that either (1) $\Delta_T \equiv 0$ over some interval $[\eta - \varepsilon, \eta] \subset [0, 1]$, or (2) $\Delta_T(\eta) < 0$ for some $\eta \in [0, 1]$. Then there is a stable equilibrium to the left of η.

Proof of Proposition 1: The returns to investment and non-investment f_0 and f_1 are continuous in η. Continuity of T implies that Δ_T is continuous. If condition (1) holds, clearly any point $\eta' \in [\eta - \varepsilon, \eta]$ is a stable equilibrium. If condition (2) holds then either there exists $\eta' < \eta$ such that $\Delta_T(\eta') > 0$, and so there must be a stable equilibrium between η' and η; or else $\Delta_T(0) < 0$, in which case $\eta = 0$ is a stable equilibrium. **QED**

[7] This was the case considered by Pigou (1932).

Proposition 1 immediately implies two restrictions on implementation via income-based incentive schemes.

The first restriction relates to the fact that an individual's pre-transfer income from non-investment exceeds that from investment when the fraction of investors is less than $\tilde{\eta}$. From Proposition 1, we know that any tax and transfer scheme that implements full investment must leave investors with a higher after-transfer income than non-investors. Consequently, the tax and transfer scheme must be non-monotonic in pre-transfer income: over some range, income must be taxed at more than 100%. An immediate corollary is that implementation is infeasible if economic agents can costlessly burn money.

More formally, assume an agent with true income y can costlessly reduce her income to any $\theta \in [0, y]$. This transforms a tax/transfer scheme as follows. Given T, an agent chooses income $\theta \in \arg\max_{\tilde{\theta} \in [0,y]} T\left(\tilde{\theta}\right)$. (Since T is continuous and $[0, y]$ is compact, $T(\theta)$ has a well-defined maximum over $[0, y]$.) So the effective transfer scheme is

$$T^{eff}(y) = \max_{\tilde{\theta} \in [0,y]} T\left(\tilde{\theta}\right).$$

Trivially T^{eff} is weakly increasing in y. It is easily shown to be continuous. Consequently:

Corollary 1 *If agents can costlessly burn money, implementation of full investment is impossible.*

The second restriction relates to the existence of an investment level $\tilde{\eta}$ at which investor and non-investor incomes coincide. Proposition 1 tells us that implementation requires that the post-transfer returns from investment must everywhere (weakly) exceed those from non-investment. However, if investment entails even a small non-pecuniary cost ϵ then under any tax/transfer policy investors are worse off than non-investors when the investment level is $\tilde{\eta}$. As such, a stable equilibrium will exist to the left of $\tilde{\eta}$:

Corollary 2 *If investment entails a strictly positive non-pecuniary cost then implementation of full investment is impossible.*

Corollaries 1 and 2 identify two serious impediments to the possibility of implementing full investment. In the next section, we show that even if a non-monotonic tax/transfer scheme is viable (i.e., burning money is not possible), and there are no non-pecuniary costs of investment implementation of full investment remains extremely challenging. Consistent with informal discussions of coordination policies, our analysis highlights the key role played by government information (or its lack thereof).

4 Policies based on individual income

We start with the case where the transfer policy, T, can only be conditioned on an agent's income. We know from Proposition 1 that implementing full investment requires that *after taxes and transfers* investment yields a higher income level for all η. Non-investment yields a higher income when less than $\tilde{\eta}$ invest but not when more than $\tilde{\eta}$ invest. Hence, T must ensure that higher incomes are penalized when fewer than $\tilde{\eta}$ invest, but rewarded when more than $\tilde{\eta}$ invest.

It follows that the tax/transfer policy T must be at a minimum at the income level at which investment and non-investment incomes coincide, i.e., at $\tilde{y} = f_0(\tilde{\eta}) = f_1(\tilde{\eta})$. When T is differentiable, a somewhat loose argument is as follows. Since $\Delta_T(\tilde{\eta}) = 0$, it follows from Proposition 1 that Δ_T must be at a local minimum at $\tilde{\eta}$ — else, there exists a point η such that $\Delta_T(\eta) < 0$, which is inconsistent with implementation of full-investment. Taking derivatives of Δ_T gives

$$\begin{aligned}
\Delta_T'(\eta) &= f_1'(\eta)\, T'(f_1(\eta)) - f_0'(\eta)\, T'(f_0(\eta)) \\
\Delta_T''(\tilde{\eta}) &= \left(f_1'(\tilde{\eta})^2 - f_0'(\tilde{\eta})^2\right) T''(\tilde{y})
\end{aligned}$$

Then $\Delta_T'(\tilde{\eta}) = 0$ implies $T'(\tilde{y})$, while $\Delta_T''(\tilde{\eta}) \geq 0$ and $|f_1'| > |f_0'|$ imply $T''(\tilde{y}) \geq 0$. More formally:

Lemma 1 ***(Intersection points)***

Let $\tilde{\eta} \in (0,1)$ be an investment level at which investment and non-investment incomes

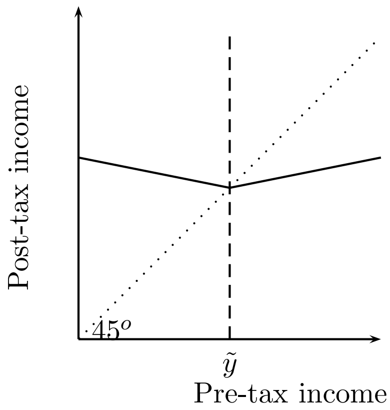

Figure 2: The tax/transfer scheme $T = \tilde{y} + \gamma|y - \tilde{y}|$

coincide. Any scheme T that implements full-investment ($\eta = 1$) must have a local minimum at $\tilde{y} = f_1(\tilde{\eta})$.

Proof of Lemma 1: See appendix.

The following example illustrates Lemma 1. Assume income from investment is $f_1(\eta) = \eta + x_1$, while that from non-investment is $f_0(\eta) = x_0 - \beta\eta$. Also, assume $\beta \in (0, 1)$, $x_0 > x_1$ and $1 + x_1 \geq x_0 - \beta$, so that absent government intervention no-investment and full-investment are both stable equilibria.

When a fraction $\tilde{\eta} = \frac{x_0 - x_1}{1+\beta}$ invest both investment and non-investment incomes equal $\tilde{y} = x_1 + \frac{x_0 - x_1}{1+\beta}$. Lemma 1 suggests it is natural to consider a tax/transfer policy with a minimum at $\tilde{y}$. One class of policies with this property, illustrated in Figure 2, is

$$T = \tilde{y} + \gamma|y - \tilde{y}|$$

This specification of T is continuous in y and implements $\eta = 1$ as follows. First, if $\eta < \tilde{\eta}$ invest then the pre-transfer incomes of investors are lower than those of non-investors; but their post-transfer incomes are higher since the externalities experienced by investors are stronger, and so $|f_1(\eta) - \tilde{y}| > |f_0(\eta) - \tilde{y}|$. Second, if $\eta > \tilde{\eta}$ invest then both the pre-transfer and post-transfer incomes of investors exceed those of non-investors. So $\Delta_T(\eta) > 0$ for all η other than $\tilde{\eta}$, and $\eta = 1$ is the only stable equilibrium. Figure 3 provides a graphical illustration.

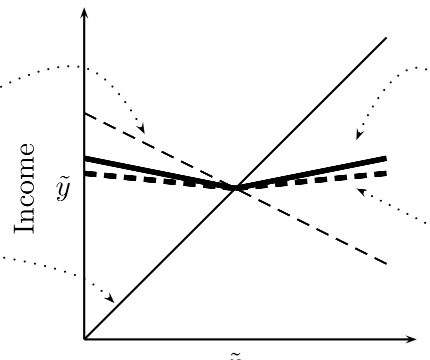

Figure 3: Implementing full-investment: the solid lines represent investment incomes, the dashed lines non-investment incomes, and the heavy lines after-tax/transfer incomes. For all η after-tax/transfer incomes are higher for investors than non-investors.

However, implementation is extremely sensitive to small imperfections in government information. To see this, consider the specific policy with $\gamma = 1$ (so that $T(y) = 2\tilde{y} - y$ for $y \leq \tilde{y}$, and $T(y) = y$ otherwise) and assume the government wrongly estimates the returns to investing as $f_1(\eta) + (\hat{x}_1 - x_1)$, where $\hat{x}_1 \neq x_1$. We depict this possibility in Figure 4.

Suppose it overestimates investment returns, i.e., $\hat{x}_1 > x_1$. In this case, incomes in the neighborhood of $\tilde{y}$ are unaffected, and consequently $\Delta_T(\tilde{\eta}) = 0$, with $\Delta_T(\eta) < 0$ immediately to the left of $\tilde{\eta}$. There is now a stable equilibrium to the left of $\tilde{\eta}$.

Suppose instead the government underestimates investment returns, i.e., $\hat{x}_1 < x_1$. Then incomes around $\tilde{y}$ are reversed, making $\tilde{\eta}$ a stable equilibrium — whereas absent the tax/transfer policy it was unstable.

In this example arbitrarily small errors in the government's knowledge of the economy cause implementation to fail. We show that this sensitivity of implementation to the government's information is a general property of the environment.

In more general terms than before, a government may misperceive the returns to investment as g, in place of the true returns, f_1.[8] Let $\mathcal{G}$ be the metric space consisting of

[8]Proposition 2 (see below) also holds if we restrict attention to the specific case analyzed in the example, i.e., mistakes of the form $f_1(\eta) + k$, for some *constant* k.

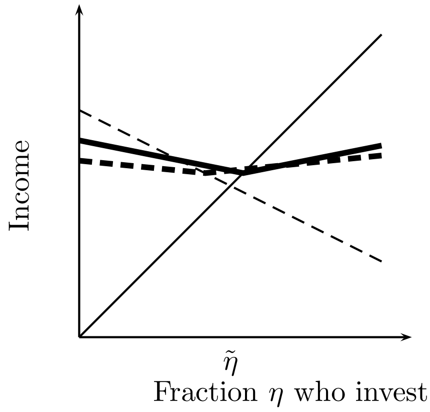

Figure 4: Errors create interior stable equilibrium: The solid lines represent investment incomes, the dashed lines non-investment incomes, and the heavy lines after-tax/transfer incomes. We see that a small upwards bias in identifying the income level $\tilde{y}$ at which incomes coincide creates a stable equilibrium at $\tilde{\eta}$.

all continuous, increasing and twice-differentiable functions from $[0, 1]$ to the positive real line, with metric $\|g - h\| = \int |g - h| \, \forall g, h \in \mathcal{G}$. It is reasonable to regard the government's misperception as small whenever $\|f_1 - g\|$ is small, i.e., when

$$g \in \{h \in \mathcal{G} | \, \|h - f_1\| < \varepsilon\} \equiv B_\varepsilon (f_1)$$

for some small ε. Given these preliminaries, we define fragile implementation:

Definition 4 ***(Fragility)***
*Suppose that T implements η. Implementation is **fragile** if for any $\varepsilon > 0$, implementation of η is non-generic within $B_\varepsilon (f_1)$, i.e., if the set*

$$\{g \in B_\varepsilon (f_1) \, | T \text{ implements } \eta \text{ in economy } (f_0, f_1 = g)\}$$

*is nowhere dense in $B_\varepsilon (f_1)$. Implementation is **non-fragile** if it is not fragile.*

It is worth pausing to compare the definitions of stability and fragility. In the above example the government's misperception of f_1 creates interior stable equilibria with basins of attraction that extend only a distance $\frac{2\beta(\hat{x}_1 - x_1)}{(1-\beta)(1+\beta)}$ to the right of the newly created stable equilibrium (see Appendix C). Consequently, when the government's misperception is small, the size of trembles required for agents to escape the interior equilibrium is likewise small.

By definition, equilibrium stability incorporates some notion of strategic uncertainty among investors – an equilibrium is stable if investors' optimal strategy is invariant to small fluctuations in the proportion of investors. In contrast, implementation is fragile if generic small misperceptions of investment returns f_1 by the government cause implementation to fail. It follows that implicit in our definition of fragility is the assumption that a government's informational misperceptions are of greater "magnitude" then exogenous "trembles" in agents' investment decisions.

It is possible to provide an economic interpretation of this assumption. If we interpret stability of equilibrium in terms of the dynamic learning process among agents (in the face of strategic uncertainty), then (loosely speaking) our definition of fragility assumes that the government learns more slowly than investors.

The observation that implementation was very sensitive to small informational errors in the above example generalizes to:

Proposition 2 ***(Implementation is fragile)***
Implementation of full investment by a tax/transfer policy T is fragile.

Proof of Proposition 2: See appendix.

How bad is it when implementation fails? Suppose that a tax/transfer policy T fails to achieve a local minimum at a point $\tilde{y} = f_1(\tilde{\eta}) = f_0(\tilde{\eta})$, where $\tilde{\eta}$ is the investment level at which investment and non-investment incomes coincide. Then Lemma 1 implies that a stable equilibrium exists in the interval $[0, \tilde{\eta} + \varepsilon]$, where $\varepsilon > 0$ can be made arbitrarily small. A stable equilibrium $\eta > 0$ implies a positive level of investment, whereas the equilibrium $\eta = 0$ implies no investment. In the example considered above investment generates a negative externality on non-investors ($f_0' < 0$). Hence, any equilibrium $\eta \in (0, \tilde{\eta} + \varepsilon]$ is socially worse than the no-investment equilibrium $\eta = 0$.

5 Policies based on the income distribution

Implementation of full investment requires the use of a tax/transfer scheme that is non-monotonic in income. That is, post-transfer income must be higher for agents with lower income when $\eta < \tilde{\eta}$ with the opposite true when $\eta > \tilde{\eta}$. If it has a minimum somewhere else — even if very close by — then implementation fails. Failure of implementation is generic when the tax/transfer scheme only conditions on individual income and the government lacks perfect information about economic fundamentals, specifically the value of $\tilde{\eta}$.

We now examine whether non-fragile implementation is feasible if the government can also condition the tax/transfer scheme on the income distribution F_η.[9,10] We show that non-fragile implementation is possible if $\tilde{\eta} \neq 1/2$ and the externality on investment returns is sufficiently strong.

Figure 5 provides the intuition. For investment levels other than $\tilde{\eta}$, the income distribution has two mass points. We denote the proportion of the population at the higher of these mass points by ξ. Figure 5 shows the relation of ξ to η when $\tilde{\eta} < 1/2$. When $\eta < \tilde{\eta}$ non-investors have higher incomes, and so $\xi = 1 - \eta$; conversely, when $\eta > \tilde{\eta}$ investors have higher incomes and $\xi = \eta$.

For non-fragile implementation, the important point is that income distributions with $\xi \leq 1/2$ only arise when the investment level is between $\tilde{\eta}$ and $\hat{\eta} = \frac{1}{2} > \tilde{\eta}$. Moreover, for a large enough investment externality, the income of investors when $\eta > 1/2$ exceeds the income level of both non-investors and investors for any $\eta < \tilde{\eta}$.

[9]Section 7 discusses whether it is realistic to assume that T can depend on the income distribution.

[10]The implementation problem when transfers can depend on the full income distribution is related to Battaglini (Forthcoming). The main difference between our paper and his is that we show unique implementation of the full-investment equilibrium in our environment. Battaglini's more general results only establish uniqueness of implementation in coalition-proof equilibria. In our environment no-investment is not a coalition-proof equilibrium whenever $f_0(\eta) < f_1(1)$ (this is implied by $f_0' < 0$ and the assumption that full-investment is socially efficient). Thus the coordination problem at the heart of our paper (and, indeed, in much of the literature on coordination failures) simply disappears if we only consider coalition-proof equilibria.

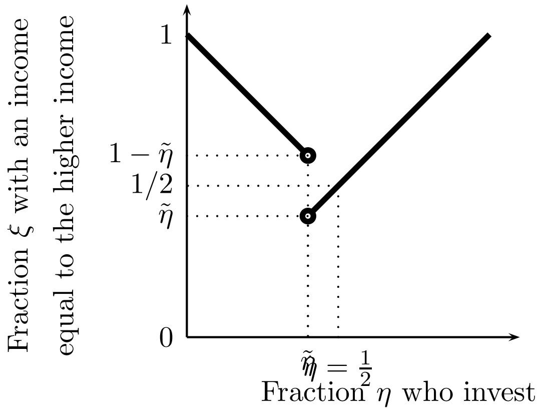

Figure 5: The fraction ξ with an income equal to the higher income as a function of the investment level (shown for $\tilde{\eta} < 1/2$).

Non-fragile implementation can then be achieved as follows. If $\xi \le 1/2$, it must be that $\eta > \tilde{\eta}$, and so individuals with higher incomes should receive higher post-transfer incomes. Moreover, if $\xi > 1/2$ but the higher income exceeds $\max_{\eta\in[0,\tilde{\eta}]} f_0(\eta)$, then again it must be that $\eta > \tilde{\eta}$, and so individuals with higher incomes should receive higher post-transfer incomes. Finally, if $\xi > 1/2$ but the higher income level is less than $\max_{\eta\in[0,\tilde{\eta}]} f_0(\eta)$ then it must be that $\eta < \tilde{\eta}$, and so individuals with higher incomes should receive *lower* post-transfer incomes.

A similar scheme is possible when $\tilde{\eta} > 1/2$ (see the proof of Proposition 3). Moreover, it is worth stressing that the requirement that the externality is sufficiently strong is sufficient but not necessary. That is, all that is required is that when $\xi > 1/2$, the government can robustly distinguish the distribution when $\eta = 1 - \xi$ invest from that when $\eta = \xi$ invest. When externalities are strong this is easily established, but it will be true much more generally. However, for the purposes of presenting a reasonably concise constructive proof, we focus on the case of strong externalities. Formally, our result is:

Proposition 3 ***(Non-fragile implementation)***

Provided the externality on investment is sufficiently large (i.e., f_1' large enough), and $\tilde{\eta} \neq 1/2$, there exists a tax/transfer policy T such that for all $\hat{f}_1$ sufficiently close to f_1 full-

investment is the only stable equilibrium.

Proof of Proposition 3: The text above gives the main idea. The main difficulty encountered in the formal proof is defining T such that it is continuous on the space $\mathcal{F}$ of distributions. The proof is relegated to the appendix.

6 Some extensions

We conclude our analysis with a discussion of the robustness of our results to four features of our model.

6.1 Continuity of Tax Scheme

Our analysis assumes that the tax/transfer policy T is continuous — both in an agent's income y and in the income distribution F when this is also an argument. Does this assumption matter? The main issue with allowing T to depend on incomes y in a discontinuous fashion is that it allows the government to eliminate equilibria in very unconvincing ways. For instance, consider the following simple example. The return to investment is $f_1(\eta) = \eta$ while that to non-investment is $f_0(\eta) = \frac{1+2\eta}{4}$. Consider the tax/transfer policy

$$T(y) = \begin{cases} -y & \text{if } y \leq 3/8 \\ y & \text{if } y > 3/8 \end{cases}$$

Since $f_0(\eta) \leq 3/8$ whenever $\eta \leq 1/4$ while $f_1(\eta) \leq 3/8$ whenever $\eta \leq 3/8$, this implies

$$\Delta_T(\eta) = \begin{cases} -\eta + \frac{1+2\eta}{4} & \text{if } \eta \in \left[0, \frac{1}{4}\right] \\ -\eta - \frac{1+2\eta}{4} & \text{if } \eta \in \left(\frac{1}{4}, \frac{3}{8}\right] \\ \eta - \frac{1+2\eta}{4} & \text{if } \eta \in \left(\frac{3}{8}, 1\right] \end{cases}$$

That is, $\Delta_T(\eta)$ is strictly positive for $\eta \leq 1/4$, jumps to being negative and decreasing over $(1/4, 3/8]$, and then is increasing over $(3/8, 1]$. The only stable equilibrium is $\eta = 1$.

However, there is a strong sense in which $\eta = 1/4$ should be considered a stable equilibrium. If just less than $\eta = 1/4$ invest, then more agents will invest; while if just more than $\eta = 1/4$ invest, then fewer agents invest. Requiring that the tax/transfer policy T be continuous in y avoids such problems.

With regard to tax/transfer policies based on the income distribution, it suffices to note that, *a fortiori*, Proposition 3 would still hold if T were allowed to be discontinuous.

6.2 Message Games

Thus far we have restricted attention to tax/transfer policies that rely solely on pre-tax incomes observed by the government. We now examine whether the government can do better by using a scheme where it also collects messages from agents.

We restrict attention to finite message spaces; we denote the message space by M.[11] We also continue to distinguish between schemes which only make use of individual incomes, and those that use the full distribution.

A scheme that uses only individual incomes and individual messages is equivalent to one using only individual incomes. To see this, consider a scheme $T(y, m)$ in which the agent's after-transfer income is determined by his own pre-transfer income y and message $m \in M$. Define the scheme $\hat{T}(y) \equiv \max_m T(y, m)$: that is, for any income y the government knows which message(s) provide an agent with the highest payoff, and so can design an alternate scheme that directly embeds this information. Clearly $\hat{T}$ has the same equilibria as T.

Second, suppose a scheme can condition both on the distribution of incomes and the distribution of messages. Similarly to above, the equilibria under such a mechanism are the same as those under a suitably defined scheme in which messages are not used. We provide the details of this claim in the appendix; broadly speaking, the government can define a message-free scheme $\hat{T}$ that matches T when equilibrium messages are reported.

6.3 Nash Implementation

We have focused throughout on Nash implementation: full-investment is implemented if and only if it is a Nash equilibrium under the tax/transfer policy in effect. However, alternate

[11] This avoids the possibility of "tail-chasing" schemes.

notions of implementation are possible. For example, one might instead ask whether full-investment is *undominated Nash implementable*: that is, is there a policy T under which full investment is the only equilibrium when agents play undominated strategies?[12]

In our setting, undominated Nash implementation is straightforward. Recall that, by assumption, $f_1(1) > \max_\eta f_0(\eta)$. Choose $\hat{y} \in (\max_\eta f_0(\eta), f_1(1))$, and consider the policy

$$T(y) = \begin{cases} 0 & \text{if } y < \hat{y} \\ f_1(1)\frac{y-\hat{y}}{f_1(1)-\hat{y}} & y \geq \hat{y} \end{cases}.$$

Under this policy, investment is strictly preferred when everyone else invests. Moreover, at any investment level η, an agent's income from not investing is $f_0(\eta) < \hat{y}$, and so $T(\hat{y}) = 0$. This is at least weakly less than the agent's income from investing. Thus not investing is a weakly dominated strategy, while investing is an undominated strategy. Consequently, if we require that agents use undominated strategies a tax/transfer scheme which conditions on own income implements full investment. Moreover, this scheme is robust to small informational mistakes.[13]

More generally, and in terms of Δ_T, a policy implements full investment in undominated Nash if $\Delta_T(\eta) \geq 0$ for all η, with $\Delta_T(\eta) > 0$ for some η. The proof of Proposition 2 is easily adapted to deal with this alternate implementation concept, and implies that for non-fragile implementation in undominated Nash the policy T must be flat over some neighborhood around $\tilde{y}$.

The above example makes clear that this is straightforward to accomplish in our existing model. However, at least under some perturbations of the model the requirement of complete flatness of the tax/transfer scheme is surprisingly hard to satisfy. In particular, consider the following change to the timing of our model to capture the idea that with a small probability the government will abandon its announced tax/transfer policy:

(1) The government announces its policy T.

[12] A weakly dominated strategy is one such that an alternative strategy is weakly preferred no matter what other agents do, and is strictly preferred for at least some choice of strategies by other agents. An undominated strategy is one that is not weakly dominated.

[13] We thank the editor for bringing this example to our attention.

(2) Agents observe whether the state of the world is $\omega \in \{L, H\}$.

(3) Agents decide whether or not to invest.

(4) With probability ε_ω a "regime change" occurs: the government abandons its tax policy T, and instead adopts T_0. For concreteness, we take T_0 to be the no-intervention policy, $T_0(y) \equiv y$.

(5) The government observes incomes and executes its policy.

From the agents' perspective, the tax policy when they make their investment decisions is $T_\omega = (1 - \varepsilon_\omega) T + \varepsilon_\omega T_0$. The key point is that T_ω cannot be flat around $\tilde{y}$ for both states $\omega = L, H$. As such, non-fragile implementation in undominated Nash equilibrium is not possible in both states $\omega = L, H$. (In contrast, this modeling perturbation does not eliminate the possibility of fragile implementation.[14])

6.4 Equilibrium stability

A related issue is our definition of stability. For concreteness, consider the following example: $f_1(\eta) = \eta$, $f_0(\eta) = 3/4 - \eta/2$, and a tax policy defined by

$$T(y) = \begin{cases} \frac{1}{2} + \gamma\left(\frac{3}{8} - y\right) & \text{if } y < \frac{3}{8} \\ \frac{1}{2} & \text{if } y \in \left[\frac{3}{8}, \frac{5}{8}\right] \\ \frac{1}{2} + \gamma\left(y - \frac{5}{8}\right) & \text{if } y > \frac{5}{8} \end{cases}.$$

It is readily verified that $\Delta_T(\eta)$ is positive up to 3/8, zero from 3/8 to 5/8, and positive from 5/8 to 1.

Given our current definition of stability, this tax scheme fails to implement full investment: investment levels in the range 3/8 to 5/8 are stable equilibria. However, these equilibria are clearly not as stable as an equilibrium η for which $\Delta_T(\eta) = 0$ and Δ_T is

[14] In particular, in the example we saw that $T(y) = \tilde{y} + \gamma |y - \tilde{y}|$ Nash implements full investment. Under the possibility of regime change,

$$T_\omega = \begin{cases} (1 - \varepsilon_\omega)(1 + \gamma)\tilde{y} + (\varepsilon_\omega - (1 - \varepsilon_\omega)\gamma) y & \text{if } y \le \tilde{y} \\ (1 - \varepsilon_\omega)(1 - \gamma)\tilde{y} + ((1 - \varepsilon_\omega)\gamma + \varepsilon_\omega) y & \text{if } y > \tilde{y} \end{cases}$$

Provided $(1 - \varepsilon_\omega)\gamma > \varepsilon_\omega$ for $\omega = L, H$, full investment is the only stable Nash equilibrium for $\omega = L, H$.

strictly decreasing at η. In particular, if stability relates to small mistakes by agents in playing the game, over time agents will eventually "escape" to the right from any equilibrium in the interval 3/8 to 5/8. In essence, these equilibria may be seen as "short-run" stable, but "long-run" unstable.

If we alter our definition of stability so as to classify equilibria of this type as unstable, then full investment could be (at least) sometimes implemented in a non-fragile way. One such instance is the above example. However, as with alternate implementation concepts, Nash implementation here depends critically on the government being able to impose a tax scheme that is completely flat over the income range around $\tilde{y}$. Again, although easily accomplished in our existing model, this is not possible in the perturbed model sketched above.

7 Discussion

This paper examines the use of income-based incentive schemes as a coordination device. Generic implementation is infeasible if the scheme only conditions on individual income, and is sometimes feasible if it conditions on individual income and the income distribution.

The main problem with implementation arises from the existence of a crossover point $\tilde{\eta}$ at which incomes from investment and non-investment coincide, not from the existence of externalities *per se*. This is clearly seen when we consider implementation in a *free-rider game*, that is a game with externalities but a unique (inefficient) equilibrium.[15] In a free-rider game individual returns from investment vary with the number of investors but are always below those from not-investing. Implementation is straightforward: choose a tax/transfer that rewards those with lower pre-tax incomes. Observe that $f_1' > f_0'$ implies

[15]The economics literature on how to achieve efficiency in the presence of externalities has, almost exclusively, focused on free-rider games. Two much studied examples are the "tragedy of the commons" games where common access to resources lead agents to choose consumption levels in excess of the social optimum (see, e.g., Gordon 1954), and "team production games " where, given a joint production process, team members under-provide effort (Holmström 1982; also see Legros and Matthews 1993; and Battaglini, Forthcoming).

that $f_1(\eta) - f_0(\eta)$ is strictly increasing. Since full-investment is not an equilibrium $f_1(\eta) < f_0(\eta)$ for all $\eta < 1$. A policy of the form $T(y) = Y - \lambda y$, where Y and $\lambda > 0$ are constants rewards those with lower pre-transfer incomes. Thus $T(f_1(\eta)) > T(f_0(\eta))$ for all $\eta < 1$. That is, full-investment ($\eta = 1$) is the only stable equilibrium. Moreover, implementation is insensitive to small errors in the government's knowledge of the returns to investment, f_1.

Overall, we are inclined to interpret our results as both reinforcing and developing the widely held view that limited information constrains a government's ability to conduct a successful development policy based on coordination of investment activity. Coordination using policies that only condition on individual income fails whenever the government has even arbitrarily small misperceptions regarding the fundamentals of the economy. In contrast, the high-investment equilibrium can be implemented with less exact knowledge if the tax/transfer policy is allowed to depend on the full distribution of incomes. However, the successful execution of such a policy, in turn, presents its own informational challenges. In particular, it requires that the policy-maker know which individuals are actually affected by the multiple-equilibria generating externality. In practice, this is likely to be a non-trivial matter. To see this, consider an economy populated by many different groups. Individuals in each group (separately) decide whether to invest in a group-specific technology. The positive externality generated by investment may only apply across individuals adopting the *same* technology. If the principal does not know the group identity of agents, and if there are many groups, then the income distribution for each group is unobservable (Lanjouw and Lanjouw 2001 discuss the widespread coexistence of multiple heterogeneous sectors in developing countries).

A quite different concern about mechanisms which depend on the whole income distribution is that they may require more commitment by the government than those which do not. For instance, suppose only the government observes all agents' incomes, while each agent only observes his/her income. Then *ex post* the government will be tempted to implement the transfer policy $\min_{F \in \mathcal{F}} T(y, F)$: that is, give each agent the smallest transfer consistent with his/her individual income. Only tax/transfer policies that are independent of the income distribution are immune from this type of problem.[16]

We would like to conclude with a methodological counter to our policy pessimism. A rich literature has developed and used the techniques of implementation theory to analyze the efficient provision of public goods. However, few researchers have applied these techniques to the field of development economics.[17] This paper, we hope, both demonstrates the importance of these techniques for development economics and, more specifically, suggests their usefulness for identifying policy mechanisms which are both feasible and robust to informational constraints.

[16]The auctions literature analyzes the possibility that mechanisms in which an agent's allocation depends on other agents' messages may be prone to cheating by the principal. Vickrey (1961) notes that a "second-price method may not be automatically self-policing to quite the same extent as the top-price method." Porter and Shoham (2003) show that a second-price auction collapses to a first-price auction when the seller can costlessly "invent" bids.

[17]Exceptions include Rai (2002), who studies the design of credit programs, and Besley and Coate (1992), who analyze targeting of public work programs.

References

Anil Arya, Jonathan Glover, and John S. Hughes. Implementing coordinated team play. *Journal of Economic Theory*, 74(1):218–232, May 1997.

Anil Arya, Jonathan Glover, and Richard Young. Virtual implementation in separable Bayesian environments using simple mechanisms. *Games and Economic Behavior*, 9: 127–138, 1995.

Marco Battaglini. Moral hazard in teams. *Journal of Economic Theory*, Forthcoming.

Gary S Becker, Kevin Murphy, and Robert Tamura. Human capital, fertility, and economic growth. *Journal of Political Economy*, 98(5 Part 2):S12–37, 1990.

Timothy Besley and Steve Coate. Workfare versus welfare: Incentive arguments for work requirements in poverty alleviation programs. *American Economic Review*, 82(1):249–261, 1992.

J. Bhagwati. *India's Economy: The Shacked Giant*. Oxford University Press, 1993.

Patrick Billingsley. *Probability and Measure*. Wiley-Interscience, 1995.

Kim-Sau Chung. Affirmative action as an implementation problem. *Working Paper*, 1999.

Russell Cooper. *Coordination Games: Complimentarities and Macroeconomics*. Cambridge University Press, 1999.

H. Scott Gordon. The economic theory of a common-property resource: The fishery. *Journal of Political Economy*, 62(2):124–142, April 1954.

A.O. Hirschman. *The Strategy of Economic Development*. New Haven, CT: Yale University Prfess, 1957.

Karla Hoff. Beyond Rosenstein-Rodan: The modern theory of coordination problems in development. *Proceedings of the Annual World Bank Conference on Development Economics.*, 2000.

Bengt Holmström. Moral hazard in teams. *Bell Journal of Economics*, 13(2):324–340, Autumn 1982.

A. O. Krueger. Trade policy and economic development: how we learn. *American Economic Review*, 87(1):1–22, 1997.

J. O. Lanjouw and Peter Lanjouw. The rural non-farm sector: Issues and evidence from developing countries. *Agricultural Economics*, 26:1–23, 2001.

Patrick Legros and Steven A. Matthews. Efficient and nearly-efficient partnerships. *Review of Economic Studies*, 60(3):599–611, July 1993.

Eric Maskin. Nash equilibrium and welfare optimality. *Review of Economic Studies*, 66: 23–38, 1999.

Kiminori Matsuyama. *The Role of Government in East Asian Development: Comparative Institutional Analysis*, chapter Economic Development as Coordination Problems. Oxford University Press, 1996.

James A Mirrlees. An exploration in the theory of optimal income taxation. *Review of Economic Studies*, 38:175–208, 1971.

Dilip Mookherjee. Optimal incentive schemes with many agents. *Review of Economic Studies*, 51(3):433–446, July 1984.

Kevin Murphy, Andrei Shleifer, and Robert Vishny. Industrialization and the big push. *Journal of Political Economy*, 97, 1989.

Arthur C Pigou, editor. *The economics of welfare*. McMillan, 1932.

Ryan Porter and Yoav Shoham. On cheating in sealed-bid auctions. *Working Paper*, 2003.

Ashok S Rai. Targeting the poor using community information. *Journal of Development Economics*, 69:71–84, October 2002.

Debraj Ray. What's new in development economics? *The American Economist*, 441:3–16, 2000.

Dani Rodrik. Coordination failures and government policy: A model with applications to East Asia and Eastern Europe. *Journal of International Economics*, 40(1-2):1–22, February 1996.

Paul Rosenstein-Rodan. Problems of industrialization of eastern and south eastern europe. *Economic Journal*, 53:202–211, June SeptembeR 1943.

Ching to Ma. Unique implementation of incentive contracts with many agents. *Review of Economic Studies*, 55:555–571, 1988.

William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 16(1):8–37, March 1961.

A Appendix: Proofs

The proof of Lemma 1 requires the following result:

Lemma 2

Let $T : \Re \to \Re$ be a continuous function, with $\hat{y}$, $\tilde{y} \in \Re$ such that $T(\hat{y}) < T(\tilde{y})$. Then there exists $\breve{y} \neq \tilde{y}$ such that $|\breve{y} - \tilde{y}| \leq |\hat{y} - \tilde{y}|$ and $T(y) > T(\breve{y})$ for all y such that $|y - \tilde{y}| < |\breve{y} - \tilde{y}|$. That is, we can find a point $\breve{y}$ such that at all points closer to $\tilde{y}$ the function T is strictly greater than at $\breve{y}$ (and $\breve{y}$ itself is closer to $\tilde{y}$ than is $\hat{y}$).

Proof of Lemma 2: For any $\varepsilon > 0$ and $y \in \Re$ let $B_\varepsilon(y)$ denote the open interval $(y - \varepsilon, y + \varepsilon)$, while $\bar{B}_\varepsilon(y)$ denotes the closed interval $[y - \varepsilon, y + \varepsilon]$. The proof is by contradiction. Suppose that contrary to the claimed result for any $\breve{y} \in \bar{B}_{|\hat{y}-\tilde{y}|}(\tilde{y}) \setminus \{\tilde{y}\}$ there exists a $y \in B_{|\breve{y}-\tilde{y}|}(\tilde{y})$ such that $T(y) \leq T(\breve{y})$.

Since the function T is continuous it must obtain its infimum value over the closed interval $\bar{B}_{|\hat{y}-\tilde{y}|}(\tilde{y})$. That is, the set $A \equiv \arg\min_{y \in \bar{B}_{|\hat{y}-\tilde{y}|}(\tilde{y})} T(y)$ is non-empty. Define $y_- = \sup\{y \in A : y \leq \tilde{y}\}$ and $y_+ = \inf\{y \in A : y \geq \tilde{y}\}$, and then define $\breve{y}$ as whichever of y_+ and y_- is closer to $\tilde{y}$.[18] That is, $\breve{y}$ is essentially the element of A that is closest to $\tilde{y}$. By continuity $T(\breve{y}) = \min_{y \in \bar{B}_{|\hat{y}-\tilde{y}|}(\tilde{y})} T(y) \leq T(\hat{y}) < T(\tilde{y})$. By supposition, there exists $y \in B_{|\breve{y}-\tilde{y}|}(\tilde{y})$ such that $T(y) \leq T(\breve{y})$. But then y must lie in A, and is strictly closer to $\tilde{y}$ than $\breve{y}$ — a contradiction. **QED**

Proof of Lemma 1: The proof is by contradiction. Suppose to the contrary that T does not have a local minimum at $\tilde{y}$.

Recall $|f_1'| > |f_0'|$, and that both f_0 and f_1 are twice-differentiable. A straightforward application of the mean-value theorem implies that there exists an $\varepsilon > 0$ such that

$$|f_1(\eta) - \tilde{y}| > |f_0(\eta) - \tilde{y}| \text{ for all } \eta \in (\tilde{\eta} - \varepsilon, \tilde{\eta} + \varepsilon) \tag{1}$$

[18] In the case of ties, set $\breve{y} = y_-$. Also, if $\{y \in A : y \geq \tilde{y}\}$ is empty then simply define $\breve{y} = y_-$, while if $\{y \in A : y \leq \tilde{y}\}$ is empty define $\breve{y} = y_+$.

That is, for all investment levels close enough to $\tilde{\eta}$ the income from investing is further from $\tilde{y}$ than is the income from not investing.

By assumption $\tilde{y}$ is not a local minimum of T. So since f_1 is strictly increasing and continuous, there must exist some $\hat{\eta} \in (\tilde{\eta} - \varepsilon, \tilde{\eta} + \varepsilon)$ such that $T(f_1(\hat{\eta})) < T(\tilde{y})$. Since T is strictly below $T(\tilde{y})$ at $\hat{y} \equiv f_1(\hat{\eta})$, it follows that we can find a point $\check{y}$ such that T is strictly lower at $\check{y}$ than at all points nearer $\tilde{y}$ (Lemma 2 in the appendix gives a formal proof). Moreover, the point $\check{y}$ is itself at least as close to $\tilde{y}$ as is $\hat{y}$. Let $\check{\eta} \in (\tilde{\eta} - \varepsilon, \tilde{\eta} + \varepsilon)$ be such that $f_1(\check{\eta}) = \check{y}$. Then by (1),

$$\Delta_T(\check{\eta}) = T(f_1(\check{\eta})) - T(f_0(\check{\eta})) < 0$$

But then by Proposition 1, T cannot implement full-investment, a contradiction. **QED**

Proof of Proposition 2: Take any $\varepsilon > 0$, and let $\mathcal{G}^* \subset B_\varepsilon(f_1)$ be those specifications of investment returns for which T implements full-investment. We show that the set $\mathcal{G}^*$ is nowhere dense in $B_\varepsilon(f_1)$, i.e., that its closure $\overline{\mathcal{G}^*}$ has an empty interior.

Suppose to the contrary that there exists an open set $\mathcal{H} \subset \overline{\mathcal{G}^*}$. All elements of $\mathcal{H}$ must lie either in $\mathcal{G}^*$ itself, or else arbitrarily close to a member of $\mathcal{G}^*$. Select an element $f_1^* \in \mathcal{H} \cap \mathcal{G}^*$, and take $\upsilon > 0$ such that $B_\upsilon(f_1^*) \subset \mathcal{H}$. Let η^* be the investment level at which $f_1^*(\eta^*) = f_0(\eta^*)$.

Without loss we can assume that $f_0'(\eta^*) \neq 0$. Take η_L and η_H so that $f_0(\eta_L) < f_0(\eta^*) < f_0(\eta_H)$. The heart of the proof consists of establishing:

Claim: For all η_L and η_H sufficiently close to η^*, then $T(f_0(\eta_L)) = T(f_0(\eta_H))$.

Proof of claim: Suppose to the contrary that $\gamma \equiv |T(f_0(\eta_L)) - T(f_0(\eta_H))| > 0$. Choose functions $\overline{f_1^H}$ and $\overline{f_1^L}$ from $\mathcal{H}$ such that $\overline{f_1^H}(\eta_H) = f_0(\eta_L)$ and $\overline{f_1^L}(\eta_L) = f_0(\eta_H)$. Whenever η_L and η_H are close enough to η^*, such a choice is possible. Since T is continuous, there exist δ_H and δ_L such that $\left|T\left(\overline{f_1^H}(\eta_H)\right) - T(y)\right| < \gamma$ for $\left|\overline{f_1^H}(\eta_H) - y\right| < \delta_H$, and $\left|T\left(\overline{f_1^L}(\eta_L)\right) - T(y)\right| < \gamma$ for $\left|T\left(\overline{f_1^L}(\eta_L)\right) - y\right| < \delta_L$.

While $\overline{f_1^H}$ and $\overline{f_1^L}$ may not themselves lie within the set $\mathcal{G}^*$ over which T implements full-investment, by supposition there exist functions $f_1^H, f_1^L \in \mathcal{G}^*$ such that $\left|\overline{f_1^H}(\eta_H) - f_1^H(\eta_H)\right| <$

δ_H and $\left|\overline{f_1^L}(\eta_L) - f_1^L(\eta_L)\right| < \delta_L$. For these functions, Proposition 1 implies that

$$T\left(f_1^H(\eta_H)\right) \geq T(f_0(\eta_H)) \quad (2)$$

$$T\left(f_1^L(\eta_L)\right) \geq T(f_0(\eta_L)). \quad (3)$$

By continuity of T, there exist constants $\theta_H, \theta_L \in (-\gamma, \gamma)$ such that

$$T\left(f_1^H(\eta_H)\right) = T\left(\overline{f_1^H}(\eta_H)\right) + \theta_H \quad (4)$$

$$T\left(f_1^L(\eta_L)\right) = T\left(\overline{f_1^L}(\eta_L)\right) + \theta_L. \quad (5)$$

To complete the proof of our claim, suppose first that $T(f_0(\eta_H)) = T(f_0(\eta_L)) + \gamma$. Then (2) and (4) together imply that

$$T\left(\overline{f_1^H}(\eta_H)\right) + \theta_H = T(f_0(\eta_L)) + \theta_H \geq T(f_0(\eta_H)) = T(f_0(\eta_L)) + \gamma,$$

a contradiction since $|\theta_H| < \gamma$. Likewise, if instead $T(f_0(\eta_L)) = T(f_0(\eta_H)) + \gamma$ then (3) and (5) together imply that

$$T\left(\overline{f_1^L}(\eta_L)\right) + \theta_L = T(f_0(\eta_H)) + \theta_L \geq T(f_0(\eta_L)) = T(f_0(\eta_H)) + \gamma,$$

which is again a contradiction.

Back to main proof: The above claim implies that T must be constant over some open neighborhood around $f_1^*(\eta^*)$. To see this, suppose to the contrary that $T(y) \neq T(f_1^*(\eta^*))$ for some y close to $f_1^*(\eta^*)$. Then $T(\check{y}) = T(y) \neq T(f_1^*(\eta^*))$ for all $\check{y}$ close to $f_1^*(\eta^*)$ on the opposite side of $f_1^*(\eta^*)$ from y, contradicting the continuity of $T(\cdot)$.

It then follows that $\Delta^T \equiv 0$ over some interval around η^*. But this contradicts Proposition 1, completing the proof. **QED**

Proof of Proposition 3: We focus on the case $\tilde{\eta} < 1/2$. The case $\tilde{\eta} > 1/2$ is similar, and is sketched in less detail below. As noted in the main text, the main difficulty in proving this result arises from the need to define the tax/transfer function T such that it is continuous on the space $\mathcal{F}$ of distributions. We proceed as follows:

1. Partition $\mathcal{F}$ into $\mathcal{F}_u$, the set of distributions with mass at one point only, and $\mathcal{F}_B$, the set of distributions with mass at two points. For all $F \in \mathcal{F}_u$, define $T(\cdot, F) \equiv y$.

2. Any two-point distribution $F \in \mathcal{F}_B$ is uniquely represented by a triple (y_-, y_+, ξ), where y_- denotes the lower mass point, $y_+ \neq y_-$ the higher mass point, and ξ the mass at the higher mass point. Observe that if a sequence $\left\{\left(y_-^n, y_+^n, \xi^n\right)\right\}$ is such that $\left(y_-^n, y_+^n, \xi^n\right) \to (y_-, y_+, \xi)$ for some triple (where $y_-^n \neq y_+^n$ and $y_- \neq y_+$), then the corresponding sequence of two-point distributions $\{F^n\}$ converges to the distribution F that corresponds to (y_-, y_+, ξ).

3. We use this representation to define $T(\cdot, F)$ on $\mathcal{F}_B$. Take $\bar{y}$ greater than the maximum income achieved by non-investors when $\eta < \tilde{\eta}$, i.e., $\max_{z \in [0,\tilde{\eta}]} f_0(z)$. Then if $\xi \geq 1/2$ and $y_+ \leq \bar{y}$, define a function $t(y, (y_-, y_+, \xi))$ by

$$t(y, (y_-, y_+, \xi)) = -y(y_+ - y_-)^2 \exp((\xi - 1/2)(y - \bar{y}))$$

while otherwise define $t(y, (y_-, y_+, \xi))$ by

$$t(y, (y_-, y_+, \xi)) = y(y_+ - y_-)^2 \exp((\xi - 1/2)(y - \bar{y})).$$

Given t, we then define the policy T by

$$T(y, (y_-, y_+, \xi)) = t(y, (y_-, y_+, \xi)) + \xi(y_+ - t(y_+, (y_-, y_+, \xi))) + (1 - \xi)(y_- - t(y_-, (y_-, y_+, \xi)))$$

By construction, T is budget-balancing: for all distributions (y_-, y_+, ξ),

$$\xi T(y_+, (y_-, y_+, \xi)) + (1 - \xi) T(y_-, (y_-, y_+, \xi)) = \xi y_+ + (1 - \xi) y_-.$$

4. Clearly the specification of T above is continuous as a function of y and (y_-, y_+, ξ). So for any pair of sequences $\{y^n\} \subset \Re$ and $\{F^n\} \subset \mathcal{F}_B$ such that $y^n \to y$ and $F^n \to F \in \mathcal{F}_B$, then $T(y^n, F^n) \to T(y, F)$. This same property holds if $F^n \to F \in \mathcal{F}_U$. For this, note that the corresponding sequence of triples $\left\{\left(y_-^n, y_+^n, \xi^n\right)\right\}$ must converge to a point of the form (y, y, ξ). So $T(y^n, F^n) \to y = T(y, F)$.

5. Thus far, we have defined T in such a way that is continuous in both its arguments. We now show it implements full-investment. To this end, we evaluate $\Delta_T(\eta)$ for all η.

(a) For $\eta < \tilde{\eta}$, the distribution of incomes is $(f_1(\eta), f_0(\eta), 1-\eta)$. Note that $1-\eta > 1/2$, and $f_0(\eta) \leq \bar{y}$. So T is such that higher pre-tax incomes are associated with lower post-tax incomes, and so $\Delta_T(\eta) > 0$.

(b) At $\eta = \tilde{\eta}$ the distribution has mass at one point, $T(y, F_\eta) \equiv y$ and so $\Delta_T(\eta) = 0$.

(c) For $\eta \in (\tilde{\eta}, 1/2)$, the distribution of incomes is $(f_0(\eta), f_1(\eta), \eta)$. Since $\eta < 1/2$, T is such that higher pre-tax incomes are associated with higher post-tax incomes, and so $\Delta_T(\eta) > 0$.

(d) For $\eta \geq 1/2$, the income at the higher mass point is again $f_1(\eta)$. Provided the externality in investment is sufficiently pronounced, this exceeds $\bar{y}$. So as in the case of $\eta \in (\tilde{\eta}, 1/2)$ it follows that $\Delta_T(\eta) > 0$.

Given this, the only stable equilibrium is at $\eta = 1$.

6. It is easily verified that the previous step remains valid even if f_1 is perturbed slightly.

This completes the proof of the case $\tilde{\eta} < 1/2$. For the case $\tilde{\eta} > 1/2$, a similar mechanism works. As before, for all $F \in \mathcal{F}_u$, define $T(\cdot, F) \equiv y$. For $F \in \mathcal{F}_B$, define T as follows. Take $\underline{y}$ lower than the minimum income achieved by non-investors when $\eta > \tilde{\eta}$, i.e., $\min_{z \in [\tilde{\eta}, 1]} f_0(z)$. Provided the externality is sufficiently strong, then if $\eta < 1/2$ then $f_1(\eta) < \underline{y}$ for suitably chosen $\underline{y}$. Then if $\xi \leq 1/2$, or if $\xi > 1/2$ but $y_- < \underline{y}$ define $t(y, (y_-, y_+, \xi))$ so that incomes are reversed. Specifically,

$$t(y, (y_-, y_+, \xi)) = -y(y_+ - y_-)^2 \exp((\xi - 1/2)(y - \bar{y}))$$

Otherwise, define $t(y, (y_-, y_+, \xi))$ so that incomes are not reversed:

$$t(y, (y_-, y_+, \xi)) = y(y_+ - y_-)^2 \exp((\xi - 1/2)(y - \bar{y}))$$

The actual transfer function T is then defined in terms of t in the same way as before.

Full investment is implemented as follows. For $\eta < 1/2$ then there are $1-\eta > 1/2$ at the higher mass point, and the lower income is the investment income. Since by assumption this is less than $\underline{y}$, incomes are reversed and investment yields more than non-investment.

For $\eta \in [1/2, \tilde{\eta})$, there are $1-\eta \leq 1/2$ at the higher mass point; incomes are again reversed and investment yields more than non-investment. Finally, if $\eta > \tilde{\eta}$ then there are $\eta > 1/2$ at the higher mass point, and the lower income is the non-investment income. Since this is certainly above $\underline{y}$, incomes are not reversed and investment yields more than non-investment.
QED

B Appendix: Message Games

Let $M = \{m_1, \ldots, m_n\}$ be a finite set of possible messages. Let $\Lambda(M)$ be the set of probability distributions over M, with typical member $(\lambda^1, \ldots, \lambda^n)$. For any $\lambda \in \Lambda(M)$, let $\text{supp}\lambda$ denote the support of λ, i.e., $\text{supp}\lambda \equiv \{m_i : \lambda^i > 0\}$.

Write F^i for the distribution of incomes for agents reporting message m_i, and let $G = (F^1, \ldots, F^n, \lambda^1, \ldots, \lambda^n) \in \mathcal{F}^n \times \Lambda(M)$ represent the joint distribution of incomes and messages.

For use below, given a fraction of agents who invest, η, define $G_\eta(\zeta_0, \zeta_1)$ to be the distribution corresponding to the η investors playing mixed-strategy ζ_1 and the $1-\eta$ non-investors playing mixed-strategy ζ_0.

Formally, a tax/transfer policy that uses messages is a function $T(y, m, G)$: an agent's after-transfer income depends on his pre-tax income y, his message m, and the joint distribution over messages and incomes G. As in our main analysis, we assume that T is continuous in its arguments.

Taking η as given, and with some abuse of language, we will say that message strategies $(\lambda_0, \lambda_1) \in \Lambda(M) \times \Lambda(M)$ constitute an *equilibrium given* η if for $a \in \{0, 1\}$ and $m \in \text{supp}\lambda_a$, then $m \in \arg\max_{m'} T(f_a(\eta), m', G(\lambda_0, \lambda_1))$. That is, given the message-income distribution $G(\lambda_0, \lambda_1)$, every message reported with strictly positive probability by investors (respectively, non-investors) is a best response given the investor's (respectively, non-investor's) income level. Standard arguments imply that for any η an equilibrium given η exists.[19]

An equilibrium is an investment level η and a pair of reporting strategies (λ_0, λ_1) such that (λ_0, λ_1) is an equilibrium given η, and such that

$$\begin{aligned}
\max_{m' \in M} T\left(f_1(\eta), m', G(\lambda_0, \lambda_1)\right) &\geq \max_{m' \in M} T\left(f_0(\eta), m', G(\lambda_0, \lambda_1)\right) \text{ if } \eta \in (0,1] \\
\max_{m' \in M} T\left(f_1(\eta), m', G(\lambda_0, \lambda_1)\right) &\leq \max_{m' \in M} T\left(f_0(\eta), m', G(\lambda_0, \lambda_1)\right) \text{ if } \eta \in [0,1).
\end{aligned}$$

Proposition 4 *Suppose that $\eta_1, \ldots, \eta_k$ are the equilibrium investment levels given T. Then there is a transfer scheme $\hat{T}$ that makes no use of messages such that $\eta_1, \ldots, \eta_k$ are the equilibrium investment levels given $\hat{T}$ also.*

Proof

Define $\hat{T}$ as follows: for each η, define $\hat{T}(f_a(\eta), F_\eta) = T(f_a(\eta), m, G_\eta(\lambda_0, \lambda_1))$ where (λ_0, λ_1) is an equilibrium in the message game given η, and $m \in \text{supp}\lambda_a$. For the equilibria levels of investment $\eta_1, \ldots, \eta_k$, use message strategies $(\lambda_0, \lambda_1)_1, \ldots, (\lambda_0, \lambda_1)_k$ such that for each $i = 1, \ldots, k$, investment level η_i and message strategies $(\lambda_0, \lambda_1)_i$ do indeed together constitute an equilibrium of investment-message game.

By construction, for $i = 1, \ldots, k$, investment level η_i is indeed an equilibrium given $\hat{T}$.

Conversely, consider any other investment level $\eta \notin \{\eta_1, \ldots, \eta_k\}$. Suppose for now that $\eta \in (0,1)$. Since this is not an equilibrium investment level given T, for any message strategies (λ_0, λ_1) that are an equilibrium given η, then

$$\max_{m' \in M} T\left(f_1(\eta), m', G(\lambda_0, \lambda_1)\right) \neq \max_{m' \in M} T\left(f_0(\eta), m', G(\lambda_0, \lambda_1)\right).$$

[19] To see this, define a correspondence $Z : \Lambda(M) \times \Lambda(M) \rightarrow\rightarrow \Lambda(M) \times \Lambda(M)$ by

$$\begin{aligned}
Z(\lambda_0, \lambda_1) = &\{(\zeta_0, \zeta_1) \in \Lambda(M) \times \Lambda(M) \text{ such that for } a = 0,1, \\
&\text{if } m \in \text{supp}\zeta_a \text{ then } m \in \arg\max_{m'} T\left(f_a(\eta), m', G(\lambda_0, \lambda_1)\right).\}
\end{aligned}$$

That is, for strategies λ_0 and λ_1, the correspondance Z gives the strategies that have support over messages that are best responses given $G(\lambda_0, \lambda_1)$. Trivially $\Lambda(M) \times \Lambda(M)$ is compact, convex and non-empty. $Z(\lambda_0, \lambda_1)$ is non-empty and convex for all $(\lambda_0, \lambda_1) \in \Lambda(M) \times \Lambda(M)$. Finally, given the continuity of T, the correspondance Z is closed. Kakutani's fixed point theorem thus applies, and implies the existence of some (λ_0, λ_1) such that $(\lambda_0, \lambda_1) \in Z(\lambda_0, \lambda_1)$. The strategies (λ_0, λ_1) constitute an equilibrium.

But from this is follows that

$$\hat{T}(f_1(\eta), F_\eta) \neq \hat{T}(f_0(\eta), F_\eta),$$

so that η is not an equilibrium given $\hat{T}$ either. The cases $\eta = 0$ and $\eta = 1$ follow similarly.
QED

C Appendix: Basins of attraction in the linear example

In the main text we presented a simple linear example: $f_1(\eta) = \eta + x_1$ and $f_0(\eta) = x_0 - \beta\eta$ with $x_0 > x_1$ and $1 + x_1 \geq x_0 - \beta$. We showed that if the government is able to precisely estimate the functions f_1 and f_0, then it can compute the income level

$$\tilde{y} = x_1 + \frac{x_0 - x_1}{1+\beta} = \frac{x_0 + \beta x_1}{1+\beta}$$

corresponding to the investment level $\tilde{\eta} = \frac{x_0 - x_1}{1+\beta}$ at which investment and non-investment incomes coincide. It can then impose the tax/transfer policy

$$T(y) = \begin{cases} (1+\gamma)\tilde{y} - \gamma y & \text{if } y \leq \tilde{y} \\ (1-\gamma)\tilde{y} + \gamma y & \text{if } y \geq \tilde{y} \end{cases}$$

and successfully implement full investment. In contrast, if the government misestimates the parameter x_1 by δ it misestimates $\tilde{y}$ by $\frac{\beta\delta}{1+\beta}$ and imposes the tax/transfer policy

$$T(y) = \begin{cases} (1+\gamma)\tilde{z} - \gamma y & \text{if } y \leq \tilde{z} \\ (1-\gamma)\tilde{z} + \gamma y & \text{if } y \geq \tilde{z} \end{cases} \tag{6}$$

where $\tilde{z} = \tilde{y} + \frac{\beta\delta}{1+\beta} = \frac{x_0 + \beta(x_1+\delta)}{1+\beta}$.

As we discussed in the main text, whenever $|\delta| > 0$ this tax/transfer policy results in a stable equilibrium close to $\tilde{\eta}$. We also noted that the basin of attraction for this stable equilibrium is small. Here, we explicitly calculate the basin of attraction.

Under the tax/transfer policy given in (6),

$$T(f_1(\eta)) = \begin{cases} (1+\gamma)\tilde{z} - \gamma(\eta + x_1) & \text{if } \eta \le \tilde{z} - x_1 \\ (1-\gamma)\tilde{z} + \gamma(\eta + x_1) & \text{if } \eta \ge \tilde{z} - x_1 \end{cases}$$

$$T(f_0(\eta)) = \begin{cases} (1-\gamma)\tilde{z} + \gamma(x_0 - \beta\eta) & \text{if } \eta \le \frac{x_0 - \tilde{z}}{\beta} \\ (1+\gamma)\tilde{z} - \gamma(x_0 - \beta\eta) & \text{if } \eta \ge \frac{x_0 - \tilde{z}}{\beta} \end{cases}.$$

Observe that $\frac{x_0-\tilde{z}}{\beta} < \tilde{\eta}$ is equivalent to $(1+\beta)(x_0 - \tilde{z}) < \beta(x_0 - x_1)$, or equivalently to $(1+\beta)\tilde{z} > x_0 + \beta x_1$. Likewise, $\tilde{\eta} < \tilde{z} - x_1$ is equivalent to $x_0 - x_1 < (1+\beta)(\tilde{z} - x_1)$, or again equivalently to $x_0 + \beta x_1 < (1+\beta)\tilde{z}$.

It follows that if $\delta > 0$,

$$\frac{x_0 - \tilde{z}}{\beta} < \tilde{\eta} < \tilde{z} - x_1,$$

while if $\delta < 0$,

$$\tilde{z} - x_1 < \tilde{\eta} < \frac{x_0 - \tilde{z}}{\beta}.$$

As such, when $\delta > 0$ there are three subcases to consider:

If $\eta \le \frac{x_0-\tilde{z}}{\beta}$, then

$$\begin{aligned} \Delta_T(\eta) &= \gamma(2\tilde{z} - (\eta + x_1) - (x_0 - \beta\eta)) \\ &= \gamma(2\tilde{z} - x_1 - x_0 - (1-\beta)\eta) \\ &\ge \gamma\left(2\tilde{z} - x_1 - x_0 - \frac{(1-\beta)(x_0 - \tilde{z})}{\beta}\right) \\ &= \frac{\gamma}{\beta}((2\beta + (1-\beta))\tilde{z} - \beta x_1 - (\beta + (1-\beta))x_0) \\ &= \frac{\gamma}{\beta}((1+\beta)\tilde{z} - \beta x_1 - x_0) \\ &= \frac{\gamma}{\beta}\beta\delta > 0. \end{aligned}$$

If $\eta \in \left(\frac{x_0-\tilde{z}}{\beta}, \tilde{z} - x_1\right)$, then

$$\begin{aligned} \Delta_T(\eta) &= \gamma(-(\eta + x_1) + x_0 - \beta\eta) \\ &= \gamma(x_0 - x_1 - (1+\beta)\eta). \end{aligned}$$

This is positive if and only if

$$\eta \leq \frac{x_0 - x_1}{1+\beta} = \tilde{\eta}.$$

Finally, if $\eta \geq \tilde{z} - x_1$ then

$$\begin{aligned}\Delta_T(\eta) &= \gamma\left(-2\tilde{z} + (\eta + x_1) + (x_0 - \beta\eta)\right) \\ &= \gamma\left(x_1 + x_0 - 2\tilde{z} + (1-\beta)\eta\right)\end{aligned}$$

At $\eta = \tilde{z} - x_1$ this is equal to

$$\gamma\left(\beta x_1 + x_0 - (1+\beta)\tilde{z}\right)$$

which is negative. It is strictly increasing over the range, and is zero at

$$\eta = \frac{2\tilde{z} - (x_1 + x_0)}{1-\beta}.$$

It follows that Δ_T is negative if and only if

$$\eta \in \left(\frac{x_0 - x_1}{1+\beta}, \frac{2\tilde{z} - (x_1 + x_0)}{1-\beta}\right).$$

Consequently, there is a stable equilibrium at $\eta = \frac{x_0 - x_1}{1+\beta} = \tilde{\eta}$. The basin of attraction for this equilibrium is $[0, \frac{2\tilde{z}-(x_1+x_0)}{1-\beta})$. The righthand boundary of this interval lies a distance

$$\begin{aligned}&\frac{2\tilde{z} - (x_1 + x_0)}{1-\beta} - \frac{x_0 - x_1}{1+\beta} \\ =\ &\frac{2\left(x_0 + \beta\left(x_1 + \delta\right)\right) - (1+\beta)(x_1 + x_0) - (1-\beta)(x_0 - x_1)}{(1-\beta)(1+\beta)} = \frac{2\beta\delta}{(1-\beta)(1+\beta)}\end{aligned}$$

away from the stable equilibrium $\tilde{\eta}$. That is, the distance from the newly created stable equilibrium to the righthand boundary of its basin of attraction is proportional to δ, the size of the government's misestimation of x_1.

The analysis of the case in which $\delta < 0$ is similar. In this case, Δ_T is negative if and only if

$$\eta \in \left(\frac{2\tilde{z} - (x_1 + x_0)}{1-\beta}, \frac{x_0 - x_1}{1+\beta}\right),$$

and there is a stable equilibrium at $\frac{2\tilde{z}-(x_1+x_0)}{1-\beta}$. The basin of attraction is $[0, \frac{x_0-x_1}{1+\beta})$; again, the right hand boundary lies a distance $\frac{2\beta\delta}{(1-\beta)(1+\beta)}$ to the right of this equilibrium.