4

Contemporary Models of Development and Underdevelopment

Individuals need not make the right tradeoffs. And whereas in the past we thought the implication was that the economy would be slightly distorted, we now understand that the interaction of these slightly distorted behaviors may produce very large distortions. The consequence is that there may be multiple equilibria and that each may be inefficient.

—*Karla Hoff and Joseph E. Stiglitz,* Frontiers in Development Economics, *2002*

Governments can certainly deter entrepreneurship when they try to do too much; but they can also deter entrepreneurship when they do too little.

—*Dani Rodrik,* One Economics, Many Recipes, *2007*

After more than a half century of experience with attempting to encourage modern development, we have learned that development is both possible and extremely difficult to achieve. Thus, an improved understanding of impediments and catalysts of development is of the utmost importance. Since the late 1980s, significant strides have been made in the analysis of economic development and underdevelopment. In some cases, ideas of the classic theories reviewed in Chapter 3 have been formalized, and in the process, their logical structure and their significance for policy have been clarified and refined. At the same time, the analysis has also led to entirely new insights into what makes development so hard to achieve (as witnessed in sub-Saharan Africa) but also possible to achieve (as witnessed in East Asia). Indeed, this is what makes the study of economic development so very important: It does not happen automatically; it requires systematic effort. But development is far from a hopeless cause; we know it can be done. Theory helps us think systematically about how to organize our efforts to help achieve development—a goal second to none in its importance to humanity.

In this chapter, we review a sample of some of the most influential of the new models of economic development. In some ways, these models show that development is harder to achieve, in that it faces more barriers than had previously been recognized. But greater understanding itself facilitates improvements in development strategy, and the new models have already influenced development policy and modes of international assistance. The

chapter concludes with a framework for appraising the locally **binding constraints** on the ability of a developing nation to further close the gap with the developed world.

Binding constraint The one limiting factor that if relaxed would be the item that accelerates growth (or that allows a larger amount of some other targeted outcome).

The new research has broadened considerably the scope for modeling a market economy in a developing-country context. One of its major themes is incorporating problems of coordination among **economic agents**, such as among groups of firms, workers, or firms and workers together. Other key themes, often but not always in conjunction with the coordination problem, include the formal exploration of situations in which increasing returns to scale, a finer division of labor, the availability of new economic ideas or knowledge, learning by doing, information externalities, and monopolistic competition or other forms of industrial organization other than perfect competition predominate. The new perspective frequently incorporates work in the "new institutional economics," such as that of Nobel laureate Douglass C. North, and introduced in Chapter 2. All of these approaches depart to some degree from conventional neoclassical economics, at least in its assumptions of perfect information, the relative insignificance of externalities, and the uniqueness and optimality of equilibria.[1]

Economic agent An economic actor—usually a firm, worker, consumer, or government official—that chooses actions so as to maximize an objective; often referred to as "agents."

4.1 Underdevelopment as a Coordination Failure

Many newer theories of economic development that became influential in the 1990s and the early years of the twenty-first century have emphasized **complementarities** between several conditions necessary for successful development. These theories often highlight the problem that several things must work well enough, at the same time, to get sustainable development under way. They also stress that in many important situations, investments must be undertaken by many agents in order for the results to be profitable for any individual agent. Generally, when complementarities are present, an action taken by one firm, worker, or organization increases the incentives for other agents to take similar actions.

Complementarity An action taken by one firm, worker, or organization that increases the incentives for other agents to take similar actions. Complementarities often involve investments whose return depends on other investments being made by other agents.

Models of development that stress complementarities are related to some of the models used in the endogenous growth approach (described in Appendix 3.3), in ways we will point out later in the chapter, but the **coordination failure** approach has evolved relatively independently and offers some significant and distinct insights.[2] Put simply, a coordination failure is a state of affairs in which agents' inability to coordinate their behavior (choices) leads to an outcome (equilibrium) that leaves all agents worse off than in an alternative situation that is also an equilibrium. This may occur even when all agents are fully informed about the preferred alternative equilibrium: They simply cannot get there because of difficulties of coordination, sometimes because people hold different expectations and sometimes because everyone is better off waiting for someone else to make the first move. This section spells out the meaning and implications of these perspectives in detail, through both simple models and examples.

Coordination failure A situation in which the inability of agents to coordinate their behavior (choices) leads to an outcome (equilibrium) that leaves all agents worse off than in an alternative situation that is also an equilibrium.

When complementarities are present, an action taken by one firm, worker, organization, or government increases the incentives for other agents to take

similar actions. In particular, these complementarities often involve investments whose return depends on other investments being made by other agents. In development economics, such network effects are common, and we consider some important examples later in this chapter, including the model of the **big push**, in which production decisions by modern-sector firms are mutually reinforcing, and the **O-ring model**, in which the value of upgrading skills or quality depends on similar upgrading by other agents. Curiously, such effects are also common in analyses of frontier technologies in developed countries, particularly information technologies, in which the value of using an operating system, word-processing program, spreadsheet program, instant messaging, and other software or product standard depends on how many other users also adopt it. In both cases, the circular causation of positive feedback is common.[3] This framework may also be used in analyses of the **middle-income trap**, in which countries develop to a degree but chronically fail to reach high-income status, often due to lack of innovation capacity.

Big push A concerted, economy-wide, and typically public policy–led effort to initiate or accelerate economic development across a broad spectrum of new industries and skills.

O-ring model An economic model in which production functions exhibit strong complementarities among inputs and which has broader implications for impediments to achieving economic development.

Middle-income trap A condition in which an economy begins development to reach middle-income status but is chronically unable to progress to high-income status. Often related to low capacity for original innovation or for absorption of advanced technology, and may be compounded by high inequality.

An important example of a complementarity is the presence of firms using specialized skills and the availability of workers who have acquired those skills. Firms will not enter a market or locate in an area if workers do not possess the skills the firms need, but workers will not acquire the skills if there are no firms to employ them. This coordination problem can leave an economy stuck in a bad equilibrium—that is, at a low average income or growth rate or with a class of citizens trapped in extreme poverty. Even though all agents would be better off if workers acquired skills and firms invested, it might not be possible to get to this better equilibrium without the aid of government. As we will see, such coordination problems are also common in initial industrialization, as well as in upgrading skills and technologies, and may extend to issues as broad as changing behavior to modern "ways of doing things." Such problems are further compounded by other market failures, particularly those affecting capital markets.[4]

Another example typical of rural developing areas concerns the commercialization of agriculture. As Adam Smith already understood, specialization is one of the sources of high productivity. Indeed, specialization and a detailed division of labor are hallmarks of an advanced economy. But we can specialize only if we can trade for the other goods and services we need. Producers must somehow get their products to markets while convincing distant buyers of their quality. As Shahe Emran and Forhad Shilpi stress, in the development of agricultural markets, middlemen play a key role by effectively vouching for the quality of the products they sell; they can do this because they get to know the farmers from whom they buy as well as the products. It is difficult to be an expert in the quality of many products, so in order for a specialized agricultural market to emerge, there needs to be a sufficient number of concentrated producers with whom a middleman can work effectively. But without available middlemen to whom the farmers can sell, they will have little incentive to specialize in the first place and will prefer to continue producing their staple crop or a range of goods primarily for personal consumption or sale within the village. The result can be an **underdevelopment trap** in which a region remains stuck in subsistence agriculture.[5]

Underdevelopment trap A poverty trap at the regional or national level in which underdevelopment tends to perpetuate itself over time.

In many cases, the presence of complementarities creates a classic "chicken and egg" problem: Which comes first, the skills or the demand for skills? Often the answer is that the complementary investments must come at the same time,

through coordination. This is especially true when, as is generally the case, there is a lag between making an investment and realizing the return on that investment.[6] In this case, even if, for some reason, all parties expect a change to a better equilibrium, they will still be inclined to wait until other parties have made their investments. Thus, there can be an important role for government policy in coordinating joint investments, such as between the workers who want skills that employers can use and the employers who want equipment that workers can use. Neither may be in a position (or find it in their self-interest) to take the first step; each may be better off waiting for the other parties to invest first.

As another example, a new or modernizing firm using new technologies may provide benefits to other firms that the adopting firm cannot capture; so each firm has an incentive to underinvest in the new technology unless a sufficient number of others invest. Some of these benefits may include raising demand for key industrial products such as steel, helping pay for the fixed costs of an essential infrastructure such as railroads or container ports, or learning from others' experiences. We will take a closer look at this problem later in the chapter.

The new work expands the scope for potentially valuable government policy interventions, but it does not take their success for granted. Rather, government itself is increasingly analyzed in contemporary development models as one of the components of the development process that may contribute to the problem as well as to the solution; government policy is understood as partly determined by (endogenous to) the underdeveloped economy (see Chapter 11). For example, a dictator such as Mobutu Sese Seko, the former ruler of the Democratic Republic of Congo when it was known as Zaire, may prefer to keep his country in an underdevelopment trap, knowing full well that as the economy develops, he will lose power. But rather than concluding that government generally exacerbates underdevelopment rather than facilitates development (as in extreme versions of the neoclassical counterrevolution school), many development specialists look actively for cases in which government policy can still help, even when government is imperfect, by pushing the economy toward a self-sustaining, better equilibrium. Such **deep interventions** move an economy to a preferred equilibrium or even to a higher permanent rate of growth in which there is no incentive to go back to the behavior associated with the bad equilibrium. In these cases, government has no need to continue the interventions, because the better equilibrium will be maintained automatically. Government can then concentrate its efforts on other crucial problems in which it has an essential role (e.g., in addressing problems of public health). This onetime-fix character of some multiple-equilibrium problems makes them worthy of special focus because they can make government policy that much more powerful in addressing problems of economic development. But it also makes the policy choices more momentous, because a bad policy today could mire an economy in a bad equilibrium for years to come.

Deep intervention A government policy that can move the economy to a preferred equilibrium or even to a higher permanent rate of growth, which can then be self-sustaining so that the policy need no longer be enforced because the better equilibrium will then prevail without further intervention.

In much of economics, such complementarities are not present. For example, in competitive markets, when there is excess demand, there is counterpressure for prices to rise, restoring equilibrium. Whenever **congestion** may be present, these counterpressures are very strong: The more people there are fishing in one lake, the more fishers try to move to another lake that is less crowded; the more people there are using one road, the more commuters

Congestion The opposite of a complementarity; an action taken by one agent that decreases the incentives for other agents to take similar actions.

try to find an alternative route. But in the process of economic development, joint externalities are common: Underdevelopment begets underdevelopment, while processes of sustainable development, once under way, tend to stimulate further development.

Where-to-meet dilemma A situation in which all parties would be better off cooperating than competing but lack information about how to do so. If cooperation can be achieved, there is no subsequent incentive to defect or cheat.

Coordination problems are illustrated by the **where-to-meet dilemma**: Several friends know that they will all be in Buenos Aires on a certain day but have neglected to settle on a specific location within the city. Now they are out of communication and can arrive at a common meeting point only by chance or by very clever guessing. They want to meet and consider themselves better off if they can do so; there is no incentive to "cheat." Thus, the where-to-meet problem is quite different from that of **prisoners' dilemma**, another problem often encountered in theories of economic development.[7] But the fact that all gain from coordination does not make the where-to-meet dilemma easy to solve. There are many famous places in Buenos Aires: the Plaza de Mayo, the Cathedral, the colorful Caminito neighborhood, the Café Tortoni, the Cementerio de la Recoleta, even the casino. Only with luck would the friends end up making the same guesses and meeting in the same place. Arriving at, say, the center of Caminito and not finding the others there, one of our travelers might decide to try the Plaza de Mayo instead. But en route she might miss another of the other travelers, who at that moment might be on his way to check out the Cementerio. So the friends never meet. Something analogous happens when farmers in a region do not know what to specialize in. There may be several perfectly good products from which to choose, but the critical problem is for all the farmers to choose *one* so that middlemen may profitably bring the region's produce to market.

Prisoners' dilemma A situation in which all parties would be better off cooperating than competing, but once cooperation has been achieved, each party would gain the most by cheating, provided that others stick to cooperative agreements—thus causing any agreement to unravel.

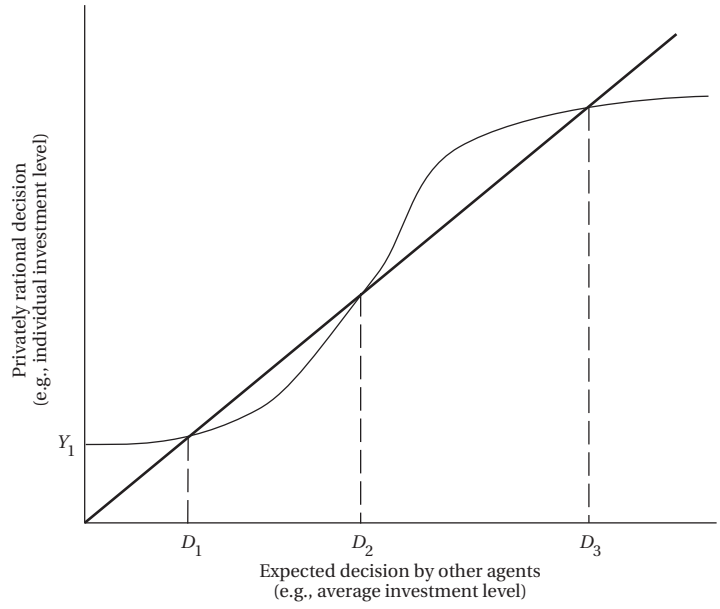
The story may lose a bit of its power in the age of texting, cell phones, and e-mail. For example, as long as the friends have each other's contact information, they can come to an agreement about where to meet. Sometimes what seems at first a complex problem of coordination is really a simpler one of communication. But anyone who has tried to establish a meeting time by phone or e-mail with a large number of participants with no formal leader knows that this can be a slow and cumbersome process. Without a clear leader and with a large enough number of participants, no meeting place may be agreed to on short notice before it is too late. And in real economic problems, the people who need to "meet"—perhaps to coordinate investments—do not even know the identity of the other key agents.[8] However, our example does point up possibilities for improved prospects for development with the advent of modern computing and telecommunications technology. Of course, peasant farmers may not have access to cell phones or e-mail (but see the case study for Chapter 11 on the Grameen Bank).

4.2 Multiple Equilibria: A Diagrammatic Approach

Multiple equilibria A condition in which more than one equilibrium exists. These equilibria sometimes may be ranked, in the sense that one is preferred over another, but the unaided market will not move the economy to the preferred outcome.

The standard diagram to illustrate **multiple equilibria** with possible coordination failure is shown in Figure 4.1. This diagram, in one version or another, has become almost as ubiquitous in discussions of multiple equilibria as the famous supply-and-demand ("Marshallian scissors") diagram in discussions of single equilibrium analysis.[9]

FIGURE 4.1 Multiple Equilibria



The basic idea reflected in the S-shaped function of Figure 4.1 is that the benefits an agent receives from taking an action depend positively on how many other agents are expected to take the action or on the extent of those actions. For example, the price a farmer can hope to receive for his produce depends on the number of middlemen who are active in the region, which in turn depends on the number of other farmers who specialize in the same product.

How do we find the equilibria in this type of problem? In the Marshallian supply-and-demand scissors diagram, equilibrium is found where the supply and demand curves cross. In the multiple-equilibria diagram, equilibrium is found where the "privately rational decision function" (the S-shaped curve in Figure 4.1) crosses the 45-degree line. This is because in these cases, agents observe what they expected to observe. Suppose that firms expected no other firms to make investments, but some firms did anyway (implying a positive vertical intercept in the diagram). But then, seeing that some firms did make investments, it would not be reasonable to continue to expect no investment! Firms would have to revise their expectations upward, matching their expectations to the level of investment they actually would see. But if firms now expected this higher level of investment, firms would want to invest even more. This process of adjustment of expectations would continue until the level of actual investment would just equal the level of expected investment: At that level, there would be no reason for firms to adjust their expectations any further. So the general idea of an equilibrium in such cases is one in which all participants are doing what is best for them, given what they expect others to do, which in turn matches what others are actually doing. This happens when the function crosses the 45-degree line. At these points, the values on the x-axis and y-axis are equal, implying in our example that the level of investment expected is equal to the level that all agents find best (e.g., the profit-maximizing level).

In the diagram, the function cuts the 45-degree line three times. Any of these points could be an equilibrium: That is what we mean by the possibility of multiple equilibria. Of the three, D_1 and D_3 are "stable" equilibria. They are stable because if expectations were slightly changed to a little above or below these levels, firms would adjust their behavior—increase or decrease their investment levels—in a way to bring us back to the original equilibrium. Note that in each of these two stable equilibria, the S-shaped function cuts the 45-degree line from above—a hallmark of a stable equilibrium.

At the middle equilibrium at D_2, the function cuts the 45-degree line from below, and so it is unstable. This is because in our example, if a little less investment were expected, the equilibrium would be D_1, and if a little more, were expected, the equilibrium would move to D_3. D_2 could therefore be an equilibrium only by chance. Thus, in practice, we think of an unstable equilibrium such as D_2 as a way of dividing ranges of expectations over which a higher or lower stable equilibrium will hold sway.

Typically, the S-shaped "privately rational decision function" first increases at an increasing rate and then at a decreasing rate, as in the diagram. This shape reflects what is thought to be the typical nature of complementarities. In general, some agents may take the complementary action (such as investing) even if others in the economy do not, particularly when interactions are expected to be with foreigners, such as through exporting to other countries. If only a few agents take the action, each agent may be isolated from the others, so spillovers may be minimal. Thus, the curve does not rise quickly at first as more agents take the action. But after enough agents invest, there may be a snowball effect, in which many agents begin to provide spillover benefits to neighboring agents, and the curve increases at a much faster rate. Finally, after most potential investors have been positively affected and the most important gains have been realized, the rate of increase starts to slow down.

In many cases, the shape of the function in Figure 4.1 could be different, however. For example, a very "wobbly" curve could cut the 45-degree line several times. In the case of telephone service, getting on e-mail or instant messaging, or buying a fax machine, where the value of taking the action steadily increases with the number of others in the network, the function may only increase at an increasing rate (like a quadratic or exponential function). Depending on the slope of the function and whether it cuts the 45-degree line, there can be a single equilibrium or multiple equilibria, including cases in which either no one ever adopts a new technology or virtually everyone does. In general, the value (utility) of the various equilibria (two in this case) is not the same. For example, it is very possible that everyone is better off in the equilibrium in which more people use the network. In this case, we say the equilibria are Pareto-ranked, with the higher rank to the equilibrium giving higher utility to everyone; in other words, moving to this equilibrium represents a **Pareto improvement** over the equilibrium with fewer users.

Pareto improvement A situation in which one or more persons may be made better off without making anyone worse off.

The classic example of this problem in economic development concerns coordinating investment decisions when the value (rate of return) of one investment depends on the presence or extent of other investments. All are better off with more investors or higher rates of investment, but the market may not get us there without the influence of certain types of government policy (but note that we may also not arrive at the preferred solutions if we have the wrong kinds of government policy). The difficulties of investment

coordination give rise to various government-led strategies for industrialization that we consider both in this chapter and later in the text (see especially Chapter 12).

The investment coordination perspective helps clarify the nature and extent of problems posed when technology spillovers are present, such as seen in the Romer model described in Appendix 3.3.[10] Given what was learned in examining endogenous growth theory about the possible relation between investment and growth, you can see that an economy can get stuck in a low growth rate largely because the economy is expected to have a low investment rate. Strategies for coordinating a change from a less productive to a more productive set of mutually reinforcing expectations can vary widely, as the example in Box 4.1 and the findings in Box 4.2 illustrate. However, changing expectations may not be sufficient if it is more profitable for

BOX 4.1 Synchronizing Expectations: Resetting "Latin American Time"

Kaushik Basu and Jorgen Weibull argue that while the importance of culture is undeniable, the innateness of culture is not. They present a model that shows that punctuality may be "simply an equilibrium response of individuals to what they expect others to do" and that the same society can benefit from a "punctual equilibrium" or get caught in a lateness equilibrium.

Estimates suggested that Ecuador lost between 4% and 10% of its GDP due to chronic lateness. As one commentator put it, "Tardiness feeds on itself, creating a vicious cycle of *mañana, mañana.*" Lately, Ecuador has tried to make up for lost time. Inspired by some in the younger generation who are fed up with "Latin American time," government and business have joined in a private-sector-funded drive to get people to show up at their scheduled appointment times. The country has launched a national *campaña contra la impuntualidad* (campaign against lateness), coordinated by Participación Ciudadana (Citizen Participation). The result is a test of the idea that a society can consciously switch from a bad to a good equilibrium through a change in expectations.

The campaign is a timely one. A newspaper is publishing a list each day of officials who are late for public events. A popular poster for the campaign against lateness describes the disease and says, "Treatment: Inject yourself each morning with a dose of responsibility, respect and discipline. Recommendation: Plan, organize activities and repair your watches." Hundreds of public and private institutions have signed up to a promise to be punctual. A popular notice for meeting rooms in the style of hotel "Do Not Disturb" signs has been making the rounds. On one side it says, "Come in: You're on time." When the meeting begins at its scheduled time, it is turned around to the other side, which reads, "Do not enter: The meeting began on time."

In Peru, a similar campaign is under way. If the campaign against lateness proves successful, it will be about more than time. If a social movement to change expectations about punctuality can be made to work, something similar might be tried around the world for fixing even more pernicious problems, such as public corruption.

Sources: Kaushik Basu and Jorgen Weibull, "Punctuality: A cultural trait as equilibrium," in *Economics for an Imperfect World: Essays in Honor of Joseph Stiglitz*, ed. Richard Arnott et al. (Cambridge, Mass.: MIT Press, 2003); Scott Wilson, "In Ecuador, a timeout for tardiness drive promotes punctuality," *Washington Post Foreign Service*, November 4, 2003, p. A22; "The price of lateness," *Economist*, November 22, 2003, p. 67; "Punctuality pays," *New Yorker*, April 5, 2004, p. 31. For an interesting critique, see Andrew M. Horowitz, "The punctuality prisoners' dilemma: Can current punctuality initiatives in low-income countries succeed?" Paper presented at the Northeast Universities Development Consortium Conference, Harvard University, October 2007.

BOX 4.2 FINDINGS Village Coordination and Monitoring for Better Health Outcomes

Chapter 4 explains the important role of improved information, shared expectations, and coordination across agents in making development progress. Coordination across households potentially can improve outcomes, for example, by changing norms toward lower fertility and ending harmful practices, and enforcing noncorrupt and efficient public-service provisions. A recent study by Martina Björkman and Jakob Svensson shows how these mechanisms may work by drawing on evidence from a randomized control trial. The researchers found that initially, villagers had little information about the scope of health problems in their village compared with outside standards, nor about what to reasonably expect from government-funded health workers. The program provided villagers with the knowledge and resources to enable them to monitor health workers individually and through their community organization. This is important to do as a community because both information gathering and monitoring have features of public goods. The results suggest that such a program can improve the behavior of health workers and lead to measurably better health outcomes—all for apparently very modest cost outlays.

The study questions were whether the intervention *caused* an increased quantity and quality of health care provision; and whether this resulted in improved health outcomes. The researchers were checking for impacts along the hypothesized "accountability chain" that treatment communities became more involved in monitoring health workers and that the intervention changed the behavior of health workers. The initial intervention had three components: first, a meeting of villagers; second, a meeting with health care workers; and finally, a meeting including both groups. This was followed by a plan of action and monitoring organized by villagers.

Initially, a "report card" comparing performance of the local health facility with others was prepared. Then facilitators in conjunction with local community leaders and community-based organizations organized a village meeting to hear and discuss the results and develop an action plan. (This is similar to the process of many community-based development activities in Africa and elsewhere.) Participation in the two-afternoon event was carefully planned to include—and hear from—diverse representatives to avoid elite capture. The facilitators "encouraged community members to develop a shared view on how to improve service delivery and monitor the provider," which were "summarized in an action plan." In these meetings, researchers observed some common concerns that "included high rates of absenteeism, long waiting-time, weak attention of health staff, and differential treatment."

The health facility meeting was a one-afternoon, all-staff event where facilitators contrasted the facility's information on service provision with findings from a household survey. Finally, an "interface meeting" was held with community representatives chosen at the community meetings and health workers, where rights, responsibilities, and suggestions for improvements were discussed, resulting in a "shared action plan…on what needs to be done, how, when and by whom." Then, "after the initial meetings, the communities were themselves in charge of establishing ways of monitoring the provider."

The program was associated with (and apparently caused) positive health outcomes, including relatively higher weights of infants, fewer deaths of children less than five years old, and greater utilization of health facilities. Evidence showed that as a result of the program, treatment practices also improved the "quality and quantity of health care provision," suggesting that increases "are due to behavioral changes." In particular, equipment (such as a thermometer) was used more often; waiting time was reduced; clinic cleanliness improved; better information was provided to patients; appropriate supplements and vaccines for children were provided more often; and absenteeism by health workers declined. The program was estimated to improve health outcomes to a degree similar

to findings from high-impact medical trials. However, such trials assume the health system is working fine and only benefits from improved procedures and medications; in contrast, this approach focused on getting health workers to do what they were supposed to do in the first place.

Some checks confirmed the program more likely had its impact through community participation rather than other mechanisms, but it is still possible that other mechanisms such as health workers responding to learning about patient rights rather than community pressure played some role; so we may not yet be certain how the program worked. This type of question is important to investigate because understanding mechanisms helps with designing other programs effectively.

Overall, the researchers surmised that "lack of relevant information and failure to agree on, or coordinate expectations of, what is reasonable to demand from the provider were holding back individual and group action to pressure and monitor the provider."

The authors caution that: "Before scaling up, it is also important to subject the project to a cost-benefit analysis....A back-of-the-envelope calculation suggests that....The estimated cost of averting the death of a child under five is around $300." If this estimate holds up to more systematic analysis, it would be an unusually cost-effective program. The authors concluded by noting that "future research should address long-term effects, identify which mechanisms or combination of mechanisms that are important, and study the extent to which the results generalize to other social sectors."

There remain some other questions. As hinted, it is uncertain whether these improvements can be sustained over time—at least without periodic outside facilitation—for example, if the initial interest for participants is in being part of a foreign-sponsored program and this motive fades over time, or if long-term threats to collective organization including free riding and capture rear their heads. So it would be valuable to return to these villages to look at conditions after a few years. It is not clear yet how well or how cost-effectively this approach would work elsewhere—the "external validity" question again. Even if the program does indeed work through the mechanism of empowerment, as seems quite likely, the real powers that be may not have allowed such outcomes if material interests of rulers were threatened by the program. Moreover, as the researchers note, an approach that combined more monitoring from the top of the health ministry in combination with the bottom-up monitoring of communities, as done in this program, could have even larger positive impacts. Finally, people and their communities have limited time; so inducing a shift of time to the health system monitoring activity in this program could cause a decrease in the amount of other valuable community activities.

But in sum, this is an exemplary design and evaluation of a community-based development program that provides substantive evidence of what can work to improve health (and empowerment) of villagers in a low-income rural area.

Sources: Martina Björkman and Jakob Svensson, "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda," *Quarterly Journal of Economics*, 124 (2), pp 735–769, May 2009; and supplementary appendix.

a firm to wait for others to invest rather than to be a "pioneer" investor. In that case, government policy is generally needed in addition to a change of expectations. This explains why attention to the potential presence of multiple equilibria is so important. Market forces can generally bring us to one of the equilibria, but they are not sufficient to ensure that the best equilibrium will be achieved, and they offer no mechanism to become unstuck from a bad equilibrium and move toward a better one.

A similar multiple-equilibria situation will be encountered in our analysis of the Malthus population trap in Chapter 6. In this population trap, fertility decisions need in effect to be coordinated across families—all are

better off if the average fertility rate declines, but any one family may be worse off by being the only one to have fewer children. We also see coordination failures in processes of urbanization and other key elements of economic development.

In general, when jointly profitable investments may not be made without coordination, multiple equilibria may exist in which the same individuals with access to the same resources and technologies can find themselves in either a good or a bad situation. In the view of many development economists, it is very plausible that many of the least developed countries, including many in sub-Saharan Africa, are essentially caught in such circumstances. Of course, other problems are also present. For example, political pressures from potential losers in the modernization process can also prevent shifts to better equilibria. In addition, modern technology may not yet be available in the country. The technology transfer problem is another important concern in economic development. In fact, another problem illustrated by the graph in Figure 4.1 could be that the amount of effort each firm in a developing region expends to increase the rate of technology transfer depends on the effort undertaken by other firms; bringing in modern technology from abroad often has spillover effects for other firms. But the possibility of multiple equilibria shows that making better technology available is generally a necessary but not a sufficient condition for achieving development goals.

4.3 Starting Economic Development: The Big Push

Whether an economy has been growing sustainably for some time or has been stagnant seems to make a very big difference for subsequent development. If growth can be sustained for a substantial time, say, a generation or more, it is much more unusual for economic development to later get off track for long (though, of course, there will be setbacks over the business cycle as the economy is affected by temporary shocks). Certainly, we have had too many disappointing experiences to assume, with Rostow, that once economic development is under way, it can in effect never be stopped. As noted in the case study in Chapter 3, a century ago, Argentina was regarded as a future powerhouse of the world economy, yet it later experienced relative stagnation for more than half a century. A look at the record, however, allows us to agree with Rostow at least in that it is very difficult to get modern economic growth under way in the first place and much easier to maintain it once a track record has been established.

Why should it be so difficult to start modern growth? Many models of development that were influential in earlier years, such as the Lewis model examined in Chapter 3, assume perfectly competitive conditions in the industrial sector. Under perfect competition, it is not clear why starting development would be so difficult, provided at least that the needed human capital is developed, the technology transfer problem is adequately addressed, and government provides other essential services. But development seems hard to initiate even when better technologies are available—they often go unused. Apparently, people do not have the incentives to put the new technology to

work. Beyond this, perfect competition does not hold under conditions of increasing returns to scale. And yet, looking at the Industrial Revolution, it is clear that taking advantage of returns to scale has been key. Many development economists have concluded that several market failures work to make economic development difficult to initiate, notably **pecuniary externalities**, which are spillover effects on costs or revenues.

Pecuniary externality A positive or negative spillover effect on an agent's costs or revenues.

Perhaps the most famous coordination failures model in the development literature is that of the "big push," pioneered by Paul Rosenstein-Rodan, who first raised some of the basic coordination issues.[11] He pointed out several problems associated with initiating industrialization in a subsistence economy, of the type introduced in Chapter 1. The problem is easiest to perceive if we start with the simplifying assumption that the economy is not able to export. In this case, the question becomes one of who will buy the goods produced by the first firm to industrialize. Starting from a subsistence economy, no workers have the money to buy the new goods. The first factory can sell some of its goods to its own workers, but no one spends all of one's income on a single good. Each time an entrepreneur opens a factory, the workers spend some of their wages on other products. So the profitability of one factory depends on whether another one opens, which in turn depends on its own potential profitability, and that in turn depends on the profitability of still other factories. Such circular causation should now be a familiar pattern of a coordination failure problem. Moreover, the first factory has to train its workers, who are accustomed to a subsistence way of life. The cost of training puts a limit on how high a wage the factory can pay and still remain profitable. But once the first firm trains its workers, other entrepreneurs, not having to recoup training costs, can offer a slightly higher wage to attract the trained workers to their own new factories. However, the first entrepreneur, anticipating this likelihood, does not pay for training in the first place. No one is trained, and industrialization never gets under way.

The big push is a model of how the presence of market failures can lead to a need for a concerted economy-wide and probably public-policy-led effort to get the long process of economic development under way or to accelerate it. Put differently, coordination failure problems work against successful industrialization, a counterweight to the push for development. A big push may not always be needed, but it is helpful to find ways to characterize cases in which it will be.

Rosenstein-Rodan's arguments became a major part of the way development economists thought about development problems in the 1950s and 1960s, and they have continued to be taught in development courses. But while some of the basic intuition has thus been around for decades, the approach received a huge boost following the 1989 publication of a technical paper by Kevin Murphy, Andrei Shleifer, and Robert Vishny, which for the first time demonstrated the formal logic of this approach more clearly.[12] Its recent appeal is also due in part to its perceived value in explaining the success of the East Asian miracle economies, notably that of South Korea. One value of using a formal model is to get a clearer sense of when the need for coordination is more likely to present a serious problem. The approach of these authors was in turn simplified and popularized by Paul Krugman in his 1995 monograph, *Development, Geography, and Economic Theory*, and became the classic model of the new development theories of coordination failure of the 1990s.[13]

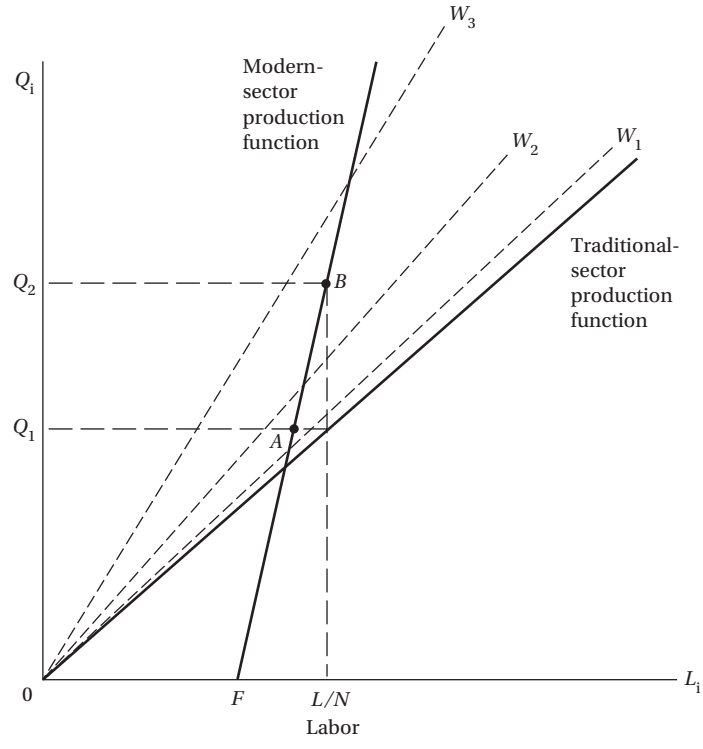
The Big Push: A Graphical Model

Assumptions In any model (indeed, in any careful thinking), we need to make some assumptions, sometimes seemingly large assumptions, to make any progress in our understanding. The analysis of the big push is no exception to this rule. The assumptions we use for the big push analysis here can be relaxed somewhat, though at the expense of requiring more mathematical technique, but it should be noted that we cannot relax our assumptions as much as we are accustomed to doing in simpler microeconomic problems, such as those that assume perfect competition. Here we cannot meaningfully assume perfect competition in the modern sector, where increasing returns to scale and hence natural monopoly, or at least monopolistic competition, prevail. To paraphrase Paul Krugman, if we think development has something significant to do with increasing returns to scale, then we will have to sacrifice some generality to address it. We will make six types of assumptions.

1. *Factors.* We assume that there is only one factor of production—labor. It has a fixed total supply, L.
2. *Factor payments.* The labor market has two sectors. We assume that workers in the traditional sector receive a wage of 1 (or normalized to 1, treating the wage as the numeraire; that is, if the wage is 19 pesos per day, we simply call this amount of money "1" to facilitate analysis using the geometry in Figure 4.2). Workers in the modern sector receive a wage $W > 1$ (that is, some wage that is greater than 1).

   As a stylized fact, this wage differential is found in every developing country, even if it needs some explanation (see Chapter 7). The underlying reason for this differential *may be* a compensation for disutility of modern factory types of work. If so, in equilibrium, workers would receive no net utility benefits from switching sectors during industrialization; but if economic profits are generated, this will represent a Pareto improvement (in this case because investors are better off and no one is worse off), and average income would rise (there can also be income redistribution so that everyone may be made better off, not just no one worse off). Moreover, if there is surplus labor in the economy or if modern wages are higher than opportunity costs of labor for some other reason,[14] the social benefits of industrialization are all the greater.[15] Finally, note that we are examining one example of a model in which a driving force for an underdevelopment trap is the relatively high wages that have to be paid in the modern sector. We do this because it is an approach that is easy to characterize graphically and that has received a lot of attention. As will be described later, however, high modern wages is only one circumstance in which a coordination problem may exist. In fact, we will see that there may be coordination failure problems even if modern-sector wages are no higher than those in the traditional sector.
3. *Technology.* We assume that there are N types of products, where N is a large number.[16] For each product in the traditional sector, one worker produces one unit of output (this is a less stringent assumption than it

FIGURE 4.2 The Big Push



appears because again we have a certain freedom in choosing our unit of measurement; if a worker produces three pairs of shoes per day, we call this quantity one unit). This is a very simple example of constant-returns-to-scale production. In the modern sector, there are increasing returns to scale. We want to introduce increasing returns in a very simple way. Assume that no product can be produced unless a minimum of, say, F workers are employed. This is a fixed cost. Because we are keeping things simple to facilitate analysis of the core issues, we have not put capital explicitly in the model; thus the only way to introduce a fixed cost is to require a minimum number of workers. After that, there is a linear production function in which workers are more productive than those in the traditional sector. Thus labor requirements for producing any product in the modern sector take the form $L = F + cQ$, where $c < 1$ is the marginal labor required for an extra unit of output. The trade-off is that modern workers are more productive, but only if a significant cost is paid up front. As this fixed cost is amortized over more units of output, average cost declines, which is the effect of increasing returns to scale. We assume symmetry: The same production function holds for producing any product in the modern sector.

4. *Domestic demand.* We assume that each good receives a constant and equal share of consumption out of national income. The model has only one

period and no assets; thus there is no saving in the conventional sense. As a result, if national income is Y, then consumers spend an equal amount, Y/N, on each good.[17]

5. *International supply and demand.* We assume that the economy is closed. This makes the model easy to develop. The most important conclusions will remain when trade is allowed, provided that there are advantages to having a domestic market. These advantages likely include initial economies of scale and learning to achieve sufficient quality, favorable product characteristics, and better customer support before having to produce for distant and unknown consumers. These are very realistic considerations: Evidence suggests that export-led economies such as South Korea have benefited enormously from the presence of a substantial domestic market to which early sales are directed.[18] Moreover, export-led economies have benefited from an active industrial policy aimed at overcoming coordination failures (see Chapter 12). The points will also hold if there are necessary inputs that are not tradable, such as certain types of services. Alternative models focusing on infrastructure investments can also imply the need for a big push even with a fully open world economy.[19]

6. *Market structure.* We assume perfect competition in the traditional (cottage industry) sector, with free entry and no economic profits. Therefore, the price of each good will be 1, the marginal cost of labor (which is the only input). We assume that at most, one modern-sector firm can enter each market. This limitation is a consequence of increasing returns to scale. Given the assumptions about preferences, the monopolist faces unit-elastic demand, so if this monopolist *could* raise its price above 1, it would be profitable to do so.[20] However, if price is raised above 1, competition from the traditional-sector producers will cause the modern-sector firm to lose all of its business. Therefore, the monopolist will also charge a price of 1 if it decides to enter the market.[21] Because the monopolist charges the same price, it will monopolize this particular market if it enters but will also produce the same quantity that was produced by the traditional producers. Because this firm is the only one using modern techniques and, in producing all other products, workers receive a wage of 1, national income will be essentially the same, so more units of output cannot be sold.[22] We also assume that at the point the monopolist would choose to produce, it is able to produce at least as much output as the traditional producers for that same level of labor; otherwise, it would make no sense to switch out of the traditional techniques.

Conditions for Multiple Equilibria With these six assumptions, we can characterize cases that will require a big push. To begin, suppose that we have a traditional economy with no modern production in any market. A potential producer with modern technology (i.e., a technology like the one described previously, with fixed costs and increasing returns) considers whether it is profitable to enter the market. Given the size of the fixed cost, the answer depends on two considerations: (1) how much more efficient the modern

sector is than the traditional sector and (2) how much higher wages are in the modern sector than in the traditional sector.

In Figure 4.2, production functions are represented for the two types of firms for any industry.[23] The traditional producers use a linear technique with slope 1, with each worker producing one unit of output. The modern firm requires F workers before it can produce anything, but after that, it has a linear technique with slope $1/c > 1$. Price is 1, so revenues PQ can be read off the Q axis. For the traditional firm, the wage bill line lies coincident with the production line (both start at the origin and have a slope of 1). For the modern firm, the wage bill line has slope $W > 1$. At point A, we see the output that the modern firm will produce if it enters, provided there are traditional firms operating in the rest of the economy. Whether the modern firm enters depends, of course, on whether it is profitable to do so.

Using Figure 4.2, first consider a wage bill line like W_1 passing below point A. With this relatively low modern wage, revenues exceed costs, and the modern firm will pay the fixed cost F and enter the market. In general, this outcome is more likely if the firm has lower fixed costs or lower marginal labor requirements as well as if it pays a lower wage. By assumption, production functions are the same for each good, so if a modern firm finds it profitable to produce one good, the same incentives will be present for producing all goods, and the whole economy will industrialize through market forces alone; demand is now high enough that we end up at point B for each product. This shows that a coordination failure need not always happen: It depends on the technology and prices (including wages) prevailing in the economy.

If a wage bill line like W_2 holds, passing between points A and B, the firm would not enter if it were the only modern firm to do so in the economy because it would incur losses. But if modern firms enter in each of the markets, then wages are increased to the modern wage in all markets, and income expands. We may assume that price remains 1 after industrialization. Note that the traditional technique still exists and would be profitable with a price higher than 1. So to prevent traditional firms from entering, modern firms cannot raise prices above 1.[24] The modern firm can now sell all of its expanded output (at point B), produced by using all of its available labor allocation (L/N), because it has sufficient demand from workers and entrepreneurs in the other industrializing product sectors. As can be seen in Figure 4.2, with prevailing wage W_2, point B is profitable after industrialization because it lies above the W_2 line. Workers are also at least as well off as when they worked in the traditional sector because they can afford to purchase an additional quantity of goods in proportion to their increased wage,[25] and they have changed sectors (from traditional to modern) voluntarily. All of the output is purchased because all of national income is spent on output; national income is equal to wages plus profits, the value of which is output of each product times the number of products N.[26]

Thus, with a prevailing wage like W_2, there are two equilibria: one in which producers with modern techniques enter in all markets, and profits, wages, and output are higher than before; and one in which no modern producer enters, and wages and output remain lower. The equilibrium with higher output is unambiguously better, but in general, the market will not get there by itself.

A final possibility is found in a wage bill line like W_3, passing above point *B*. In this case, even if a modern producer entered in all product sectors, all of these firms would still lose money, so again the traditional technique would continue to be used. In general, whenever the wage bill line passes below point *A*, the market will lead the economy to modernize, and whenever it passes above *A*, it will not. The steeper (i.e., more efficient) the modern-sector production technique or the lower the fixed costs, the more likely it is that the wage bill will pass below the corresponding point *A*. If the line passes above *B*, it makes no sense to industrialize. But if the wage line passes between points *A* and *B*, it is efficient to industrialize, but the market will not achieve this on its own. Be sure to note that these are three different wages that might exist, depending on conditions in a particular economy at one point in time, not three wages that occur successively.

Again, the problematic cases occur when the wage bill line passes between *A* and *B*, thus creating two equilibria: one in which there is industrialization and the society is better off (point *B*) and one without industrialization (point *A*). However, the market will not get us from *A* to *B* because of a coordination failure.[27] In this case, there is a role for policy in starting economic development. There is no easy test to determine where a traditional economy, such as Mozambique, is located on this continuum. But at least we can begin to understand why development often has not gotten under way, even when technology is available.

Note that in general, it is not necessary for all product sectors to industrialize to get a sufficient push for some to do so. It is only necessary that a sufficient number industrialize in order to generate enough national income (through the higher industrial wage and positive profits from the industrialized product sectors) to make industrialization minimally profitable. Also note that each firm's failure to take into account the impact of its investments on demand for other firms' goods represents a very small distortion by itself. But when added up across all of the product sectors, the resulting distortion—namely, the failure to industrialize at all—is very large indeed.

We could also have cases of semi-industrialization, in which benefits or costs accrue in different amounts to different product sectors or in which there are different types of spillovers from firm to firm. For example, this is plausible when the level of required fixed costs declines the more product sectors industrialize, because there are more local examples from which to learn.[28] With this alternative type of externality, no wage premium is necessary for multiple equilibria to be present. In this case, if there are clusters of two or more firms that have large effects on each other's fixed costs, *F*, but not on firms outside of the cluster, the result can be an equilibrium in which only the industries in this cluster change to modern techniques. Thus, in this circumstance, we could have three or more equilibria; we could also have enclave economies, in which a modern sector exists side by side with traditional cottage industries in other product sectors.[29]

Technological externality A positive or negative spillover effect on a firm's production function through some means other than market exchange.

Notice that this model has not assumed the existence of any type of **technological externality**, in which the presence of one advanced firm can, through "learning by watching" other firms' production methods or some similar effect, generate spillovers to other firms that can raise their productivity as well as lower their costs. This is another type of market failure that can

also lead to inefficiently low investment; we considered one such possibility when we examined the Romer endogenous growth model in Appendix 3.3.

Other Cases in Which a Big Push May Be Necessary

The need for a big push can result from four conditions beyond those described previously.

1. *Intertemporal effects.* Even if the industrial wage rate is 1 (i.e., the same as the traditional-sector wage), multiple equilibria can occur if investment must be undertaken in the current period to get a more efficient production process in the next period.[30] Investment in the first period depresses aggregate demand in the first period but increases it in the second (or later) period. But investment will be undertaken only if it is profitable, that is, if demand is expected to be high enough in the second period, and this may require that many product sectors invest simultaneously. Once again, however, the market does not ensure that industrialization will occur, even when it is (Pareto-)preferred, because of pecuniary externalities. Again the source of the multiple equilibria is that one firm's profits do not capture its external contribution to overall demand for modern-sector products because it also raises wage income in the future periods when other entering modern firms will be seeking to sell their own products. When there is a case for a big push, industrialization makes the society better off (is Pareto-preferred) because first-period income is decreased only by the fixed cost, but second-period income is sufficiently increased by both the wage and profits in other product sectors to more than offset this.[31] Note once again that a part of the profits can, in principle, also be subject to income redistribution so that everyone may be made better off rather than just some people made better off and no one made worse off.

2. *Urbanization effects.* If some of the traditional cottage industry is rural and the increasing-returns-to-scale manufacturing is urban, urban dwellers' demand may be more concentrated in manufactured goods (e.g., foods must be processed to prevent spoilage due to the time needed for transportation and distribution). If this is the case, one needs a big push to urbanization to achieve industrialization.[32]

3. *Infrastructure effects.* By using infrastructure, such as a railroad or a port, an investing modern firm helps defray the large fixed costs of that infrastructure. The existence of the infrastructure helps investing firms lower their own costs. But investing firms thereby contribute indirectly to lowering the costs of other firms (by lowering the average cost of infrastructure use). Infrastructure, such as roads, railroads, and ports, is not tradable; by definition, it is located in a particular region. And openness to foreign investment cannot always solve the problem because investors do not know whether firms will develop to make use of the infrastructure.[33] The critical point is that when one product sector industrializes, it increases the size of the market for the use of infrastructure services that would be used by other product sectors and so makes the provision of these services more profitable. But it is also possible that efficient industrialization

may not take place, even if the infrastructure is built, if other coordination problems are present.

4. *Training effects.* There is underinvestment in training facilities because entrepreneurs know that the workers they train may be enticed away with higher wages offered by rival firms that do not have to pay these training costs. There is also too little demand by workers for training because they do not know what skills to acquire. (In addition to not knowing whether firms will make investments requiring these skills, people are not born with perfect information about their comparative advantage; basic education helps workers discover it.) This is part of the economic case for mandatory public education. Note that in this case, openness to trade cannot resolve the coordination failure unless there is free mobility of labor across borders, which has yet to develop perfectly even within the European Union, where there are few formal barriers to such mobility, and is far from emerging for any developing country. In any case, relying on expatriate skilled workers is hardly an adequate solution to a country's own underdevelopment. Actually, infrastructure and trained workers are subsets of a general case of jointly used intermediate goods. Another example is joint research facilities for small firms in an "industrial district" (see Chapter 7).

Why the Problem Cannot Be Solved by a Super-Entrepreneur

Some readers may wonder, why can't one agent solve the coordination failure problems by capturing all the rent? In other words, why not have a super-entrepreneur who enters into all of the markets that need to be coordinated and receives the profits from all of them? For some types of coordination failures, this solution is ruled out in advance. For example, regarding education and skill development, there is a legal constraint on bonded labor. But in terms of our industrialization problem, why can't one agent become a super-entrepreneur in each of the N markets simultaneously? There are at least four significant theoretical answers and one decisive empirical answer.

First, there may be capital market failures. How could one agent assemble all the capital needed to play the super-entrepreneur role? Even if this were logistically imaginable, how would lenders have confidence in their investments? In particular, how could a penalty for default be imposed?

Agency costs Costs of monitoring managers and other employees and of designing and implementing schemes to ensure compliance or provide incentives to follow the wishes of the employer.

Second, there may be costs of monitoring managers and other agents and designing and implementing schemes to ensure compliance or provide incentives to follow the wishes of the employer; these are often referred to as **agency costs**. Monitoring is too expensive once the scale of a firm gets too large. Even if the plan is to sell off the industries, these industries must be developed simultaneously. The super-entrepreneur is likely to know more about the firms than the potential buyers do. In other words, if the firm is so profitable, why would its owners be selling? Thus, potential purchasers of the industries face a problem of **asymmetric information**, often known as the "lemons problem."[34]

Asymmetric information A situation in which one party to a potential transaction (often a buyer, seller, lender, or borrower) has more information than another party.

Third, there may be communication failures. Suppose someone says to you, "I am coordinating investments, so work with me." Should you do so?

How do you know this person will eventually be the coordinator? There is a potentially huge profit to be made by assuming the super-entrepreneur role, so many agents might wish to play it. If many try to claim the role, with which one should you coordinate? Even if each agent personally encounters only one pretender to the super-entrepreneur role, that pretender may still not be the right one (i.e., the coordinator with whom you can make money).

Fourth, there are limits to knowledge. Even if we stipulate that the economy as a whole has access to modern technological ideas, this does not mean that one individual can gain sufficient knowledge to industrialize (or even gain enough knowledge about whom to hire to industrialize).

Finally, there is the empirical reason that no private agent has been observed playing the role of super-entrepreneur. Whether because of problems of monitoring, knowledge, capital markets, or other diseconomies of scope, "solving" problems with ever-larger firms clearly provides no answer. For example, it is rare enough to find a firm producing steel and even a significant fraction of the products using steel, let alone one firm owning all the industries backwardly linked from steel or forwardly linked from steel-using industries to industries further down the production chain. Nor can the problem be solved by direct government production (at least without unacceptable cost), as the extreme case of the former Soviet Union demonstrates. Rather, public coordination of actions of private investors is generally needed to solve the problem, a common interpretation of the role of industrial policy in East Asia.

In a Nutshell Thus we have seen that under some conditions, pecuniary externalities associated with the development process can lead to multiple equilibria, which may create a case for a big push policy. Our main example (the moderate wage premium case) and each of the other examples have as a common feature a process by which an investing (industrializing) firm captures only part of the contribution of its investment to the profits of other investing firms. In these examples, firms adopting increasing-returns-to-scale technologies are having one or more of the following effects: raising total demand, shifting demand toward manufactured goods, redistributing demand toward the (later) periods in which other industrializing firms sell, reducing the fixed costs of later entrants, or helping defray the fixed costs of an essential infrastructure. Each of these has external beneficial effects on other industrializing firms.

4.4 Further Problems of Multiple Equilibria

Inefficient Advantages of Incumbency

The presence of increasing returns in modern industries can also create another kind of bad equilibrium. Once a modern firm has entered, it has an advantage over any rivals because its large output gives it low average costs. So if an even better modern technology becomes available to a potential rival, it may not be easy for the new technology to supplant the old. Even though the new technique has a lower per-unit cost for any given level of output, the firm with the old technique has an advantage because its large output lets it

produce at a lower per-unit cost than that of the new technique, which starts out with a small customer base and a large fixed cost. As a result, firms may need access to significant amounts of capital to cover losses while they build their customer base. If capital markets do not work well, as they often do not in developing countries (see Chapter 15), the economy may be stuck with backward, less cost-effective industries.[35]

Behavior and Norms

Movement to a better equilibrium is especially difficult when it involves many individuals changing their behavior from one of rent seeking or corruption to honesty and the value of building a reputation to reap the gains from cooperation (e.g., with business partners). Your choice of partner may determine much. If you naively cooperate with an opportunistic, predator type, you may be worse off than by going it alone. Only by cooperating with other good-willed cooperators may you reach the best outcome. Moreover, past experience may lead people to expect opportunistic behavior at least among certain groups of potential business partners, which in turn raises the incentives for the potential partners to actually act that way. If there is nothing to be gained and something to be lost by being honest, the incentives lie in being dishonest. On the other hand, in some settings, individuals take it on themselves to enforce norms rather than leaving this task to government. If many people work to enforce a norm such as honesty, each individual's enforcement burden is relatively low. You can have equilibria where most people resist corruption, and so corruption is rare; and you can have equilibria where few resist corruption, and corruption is common.

We cannot rely on good organizations to prevail in competition if the rules of the game tend to reward the bad organizations. Rather, the critical importance of policies for developing or reforming institutions is highlighted, such as reform of the framework of property rights, antitrust, clean government rules, and other laws, regulations, and industry association norms that set the rules of the game for economic life. Once the new behavior assumes the status of a norm, it is much easier to maintain. Some neoclassical theorists have at times implied that good institutions would be developed through the market mechanism. Bad institutions would be outcompeted by good institutions. But reform of institutions aiding and abetting coordination failure—for example, by permitting or encouraging corruption—is itself subject to coordination failure.

Once cooperative relationships (e.g., in business) become a norm, more people may adopt cooperative behavior. But norms of all kinds are subject to inertia. Although norms may have been adaptive when they originated, they are hard to change, even when they become dysfunctional. An example is a value such as that to be a good citizen (or a good Hindu, Muslim, Christian, animist, etc.) one must have a large number of children. This value may have been adaptive at a premodern stage, but today it inhibits development. Another example may be to distrust anyone who is not a member of your family. This may be helpful in a tribal context, and caution is always advisable, but this extreme injunction hardly encourages the formation of successful business partnerships in a modern economy.

Linkages

There are several ways to undertake a big push, encouraging the simultaneous expansion of the modern sector in many industries. One strategy for solving coordination problems is to focus government policy on encouraging the development of industries with key backward or forward **linkages**. This could mean subsidies or quid pro quos for domestic industries to enter these key industries, as was done in South Korea; it could mean incentives for multinational firms to enter in key industries and provide advanced training, a policy followed in Singapore; or it could mean establishing a few key public enterprises to act as pioneers in an industry (that could later be sold), as was done in South Korea and Taiwan.[36] The theory of linkages stresses that when certain industries are developed first, their interconnections or linkages with other industries will induce or at least facilitate the development of new industries. Backward linkages raise demand for an activity, while forward linkages lower the costs of using an industry's output; both may involve interactions between the size of the market and increasing returns to scale and hence pecuniary externalities. In other words, linkages are especially significant for industrialization strategy when one or more of the industries involved have increasing returns to scale, of which a larger market may take advantage. For example, when the manufacture of power looms expands, enabling a reduction in the price of power looms, there are forward linkage effects due to increased output of woven cloth made by the power looms. When increased demand for chemicals used in textile manufacture causes expansion of the chemical industry that enables it to produce at a larger scale and hence lower cost, a backward linkage can occur. Both examples illustrate a pecuniary externality effect (a lowering of cost) when there are increasing returns in the linked industry.

Linkages Connections between firms based on sales. A backward linkage is one in which a firm buys a good from another firm to use as an input; a forward linkage is one in which a firm sells to another firm. Such linkages are especially significant for industrialization strategy when one or more of the industries (product areas) involved have increasing returns to scale that a larger market takes advantage of.

The linkage approach targets investment in a key linkage as a start to overcoming a coordination failure and generating positive feedback. Such a policy would select industries with a larger number of links to other industries and greater strength of those links. In choosing among industries with several strong links (and passing a cost-benefit test), one policy would generally select industries that have a smaller likelihood of private investment, because that is where the most intransigent bottlenecks are most likely to be found. If an investment is profitable, it is more likely that an entrepreneur will come along to fill that niche.[37] This observation provides a reason to interpret with some caution studies that show state-owned enterprises to be less efficient than private ones. If government systematically enters vital but less profitable industries because of their beneficial effects on development, it is unreasonable to hold these enterprises to the same profit standards as those of the private firms. This is certainly not to say that state-owned enterprises are generally as efficient as privately owned ones; in fact, there is much evidence to the contrary.[38] We can say, however, that a blanket statement, such as has often been made in publications from agencies such as the World Bank, that government should never be in the business of production, even temporarily or in any industry, is sometimes unreasonable in the light of linkages and other strategic complementarities that a developing economy needs to address.

Inequality, Multiple Equilibria, and Growth

Other important work being done on growth and multiple equilibria addresses the impact of inequality on growth. The traditional view has been that some inequality may enhance growth because the savings of the rich are higher than those of the poor. If at least some savings to be mobilized for investment purposes must come from within a country, then according to this view, too high a degree of equality could compromise growth. However, the poor save at much higher rates than previously believed, when savings are properly measured to include expenditures on health, children's education, and improvements on a home.

Moreover, where inequality is great, the poor may not be able to obtain loans because they lack collateral; indeed, one definition of what it means to be poor is to be entirely or mostly lacking in a source of collateral. Poor persons unable to get a loan to start a business due to such capital market imperfections may get stuck in subsistence or wage employment, although they (and perhaps potential employees) could do much better if they had access to financing or if there were a more even distribution of income. For example, Abhijit Banerjee and Andrew Newman show that multiple equilibria, including equilibria involving outcomes with virtually all citizens enjoying high incomes and outcomes with predominantly low-income people, can exist when imperfect credit markets provide too few people with the opportunity to become entrepreneurs.[39]

Similarly, if the poor lack access to credit, they may not be able to obtain loans to finance otherwise very productive schooling. If the poor are unable to bequeath much to their next generation, families can be trapped in poverty from generation to generation; however, if schooling could somehow be achieved, they could escape from this **poverty trap**. It is best to keep in mind a rather expansive definition of what is meant by a *transfer* from parents to be used for human capital accumulation by their children. It is more than tuition and more than forgone wages or work on the farm to help the family because it goes well beyond the cost of formal schooling and may be thought of as the building of a whole array of "capabilities" (see Chapter 1) that one acquires almost as a simple by-product of growing up in an affluent, educated family.

Poverty trap A bad equilibrium for a family, community, or nation, involving a vicious circle in which poverty and underdevelopment lead to more poverty and underdevelopment, often from one generation to the next.

In a formal model of this problem, Oded Galor and Joseph Zeira examined the implications of missing credit markets for growth and the distribution of both income and human capital. They developed an endogenous growth model that points up the importance of both human capital and distribution, and of the interaction between the two, for economic growth and development as well as for more short-term macroeconomic adjustments. Their analysis contains two critical assumptions: (1) imperfect capital markets, which, as will be described in detail in Chapter 15, is a typical condition of these markets, and (2) indivisibilities in human capital investment, which means that markets treat investment in human capital as coming in discrete packages, such as a year of school, if not larger blocks, such as primary, secondary, and tertiary education. The second assumption does not seem unreasonable, both because of the nature of learning and because of the screening nature of markets for human capital. A threshold level of knowledge is necessary before an employer will be willing to pay for it. Further, because education acts as a screen for inherent ability, as will be discussed in Chapter 8, we have the well-known "sheepskin effect"; that is, there is a very large jump in the return

to human capital when an individual passes primary school and again when the person obtains a secondary school diploma and so on. This is not because the last course taken conveys so much more knowledge than the ones preceding it but because the degree itself is what enables the individual to prove that an entire regimen of requirements has been met. Note that indivisibilities in amounts of investment imply a region of increasing returns to scale, as in the fixed costs of the big push model. Once again, increasing returns play a key role in generating multiple equilibria.[40] Empirically, many studies have found a negative impact of inequality on growth, especially for the period after 1980.[41]

4.5 Michael Kremer's O-Ring Theory of Economic Development

Another innovative and influential model that provides important insights into low-level equilibrium traps was provided by Michael Kremer.[42] The notion is that modern production (especially in contrast to traditional crafts production) requires that many activities be done well together in order for any of them to amount to a high value. This is a form of strong complementarity and is a natural way of thinking about specialization and the division of labor, which along with economies of scale is another hallmark of developed economies in general and industrial production in particular. The name for Kremer's model is taken from the 1986 *Challenger* disaster, in which the failure of one small, inexpensive part caused the space shuttle to explode. The O-ring theory is interesting in part because it explains not only the existence of poverty traps but also the reasons that countries caught in such traps may have such exceptionally low incomes compared with high-income countries.

The O-Ring Model

The key feature of the O-ring model is the way it models production with strong complementarities among inputs. We start by thinking of the model as describing what is going on inside a firm, but as we will see, this model also provides valuable insights into the impact of complementarities across firms or industrial (product) sectors of the economy.

Suppose that a production process is broken down into n tasks. There are many ways of carrying out these tasks, which for simplicity we order strictly by level of skill, q, required, where $0 \leq q \leq 1$. The higher the skill is, the higher the probability that the task will be "successfully completed" (which may mean, for example, that the part created in this task will not fail). Kremer's concept of q is quite flexible. Other interpretations may include a quality index for characteristics of the good: Consumers would be willing to pay more for higher-quality characteristics. For example, suppose that $q = 0.95$. Among other interpretations, this can mean (1) that there is a 95% chance that the task is completed perfectly, so the product keeps maximum value, and a 5% chance that it is completed so poorly that it has no value; (2) that the task is always completed well enough that it keeps 95% of its maximum value; or (3) that the product has a 50% chance of having full value and a 50% chance of an error reducing the value of the product to 90%. For simplicity, assume that

the probability of mistakes by different workers is strictly independent. The production function assumed is a simple one: Output is given by multiplying the q values of each of the n tasks together, in turn multiplied by a term, say, B, that depends on the characteristics of the firm and is generally larger with a larger number of tasks. Suppose also that each firm hires only two workers. Then the **O-ring production function** looks like this:[43]

O-ring production function
A production function with strong complementarities among inputs, based on the products (i.e., multiplying) of the input qualities.

$$BF(q_i q_j) = q_i q_j \tag{4.1}$$

That is, to make things simple, for this exposition we let the multiplier, B, equal 1. In addition to the form of the production function, we make three other significant types of simplifying assumptions: (1) Firms are risk-neutral, (2) labor markets are competitive, and (3) workers supply labor inelastically (i.e., they work regardless of the wage). If we consider capital markets, we assume that they are competitive as well. For now, we also assume that the economy is closed.

One of the most prominent features of this type of production function is what is termed *positive assortative matching*. This means that workers with high skills will work together and workers with low skills will work together. When we use the model to compare economies, this type of matching means that high-value products will be concentrated in countries with high-value skills. In this model, everyone will like to work with the more productive workers, because if your efforts are multiplied by those of someone else, as they are in Equation 4.1, you will be more productive when working with a more productive person. In competitive markets, your pay is based on how productive you are. A firm with a higher-productivity worker can more afford to pay a higher wage and has the incentive to bid higher to do so, because the value of output will be higher with two productive workers, say, than with one low- and one high-productivity worker. As a result, there will be a strong tendency for the most productive workers to work together.

This can be seen easily if we imagine a four-person economy. Suppose that this economy has two high-skill q_H workers and two low-skill q_L workers. The four workers can be arranged either as matched skill pairs or unmatched skill pairs. Total output will always be higher under a matching scheme because

$$q_H^2 + q_L^2 > 2q_H q_L \tag{4.2}$$

Recall that $(x - y)^2 > 0$ for any x that is not the same as y, so let x stand for q_H and y stand for q_L. Then $x^2 + y^2 > 2xy$, the same as in Equation 4.2. (Or try this by plugging in any values $q_H > q_L$.) This generalizes to larger numbers of workers in the firms and the economy; the result is that workers sort out by skill level.[44]

Because total value is higher when skill matching rather than skill mixing takes place, the firm that starts with high-productivity workers can afford to bid more to get additional high-productivity workers, and it is profitable to do so. Of course, every firm would like to hire the most productive worker, but it would be in that worker's interest to team up with other high-productivity workers. Think of firms being formed while workers try to determine for which firm they want to work. After the high-productivity workers pair off, they are out of the picture. The less productive workers are then stuck with each other. If there are many classes of skill or productivity, first the highest-skill workers

get together, then the next highest, and so on, such that skill matching results as a cascading process. For example, a symphony orchestra will be adversely affected as a whole by hiring one single poor performer. So an otherwise excellent orchestra has every incentive to bid the most for an outstanding performer to replace the poor performer. Similarly, the best jazz performers play and record together rather than each leading a group of poorer players. The restaurant with the very best chef also hires mature, highly trained, full-time waiters, while a fast-food restaurant does not hire a famous chef.

This sorting process is perhaps most vividly easy to remember by analogy to Nobel laureate Gary Becker's famous "marriage market" model, which is a somewhat different case[45] but offers some additional intuition. If prospective spouses care only about attractiveness, every man wants to marry the most attractive woman, and every woman wants to marry the most attractive man, so the most attractive man and woman will marry. They are now out of the picture, so next, the second most attractive man and woman marry. This process continues until the least attractive man and woman marry. Of course, beauty is in the eye of the beholder, and most people care about things besides attractiveness in a mate such as kindness, intelligence, wealth, beliefs, interests, commitment, and sense of humor; but the marriage model serves as a memorable analogy. The result in the business world is that some firms and workers, even an entire low-income economy, can fall into a trap of low skill and low productivity, while others escape into higher productivity.

Although this model may seem abstract, a numerical example can show how the firms with high-skill workers can and will pay more to get other high-skill workers or will have more incentive to upgrade skills among existing workers. Suppose that there are six workers; three have $q = 0.4$ and are grouped together in equilibrium, while the other three have $q = 0.8$. Now suppose that the q of one of the workers in the first firm rises from 0.4 to 0.5 (perhaps due to training). Similarly, suppose the q of one worker in the second firm rises from 0.8 to 1.0. In each case, we have a 25% increase in the quality of one worker. As you may expect, a 25% increase in the quality of one worker leads to a 25% increase in output quality. But starting from a higher level of quality, that 25% clearly translates into a much larger point increase: In the example, the first firm goes from $(0.4)(0.4)(0.4) = 0.064$ to $(0.4)(0.4)(0.5) = 0.080$; this is a difference of $0.080 - 0.064$, which is a point change of 0.016; and $0.016/0.064 = 0.25$, which is a 25% increase. For the second firm, we move from $(0.8)(0.8)(0.8) = 0.512$ to $(0.8)(0.8)(1.0) = 0.640$; the change in this case is 0.128, which is again 25%. However, the point value of the increase is much greater—eight times greater—for a doubled point-value investment (0.2 in the second firm versus 0.1 in the first firm). If a firm can increase quality in percentage terms at constant marginal cost or even a not too quickly rising cost, there is a virtuous circle in that the more the firm upgrades overall, the more value it obtains by doing so. Accordingly, *wages will increase at an increasing rate as skill is steadily raised*. As Kremer shows, the O-ring model is consistent with competitive equilibrium.

The O-ring result of positive assortative matching relies on some rather strong assumptions. How important are each of these, and how much can they be relaxed? Two points are crucial: (1) Workers must be sufficiently imperfect substitutes for each other, and (2) we must have sufficient complementarity of tasks. As long as these conditions hold, the basic results will follow.

To see why workers must be imperfect substitutes, suppose they were perfect substitutes. Specifically, suppose there are two skill levels, q_L and $q_H = 2q_L$, so every q_H worker can be replaced by two q_L workers with no other change. Thus q_H workers will be paid twice the amount that q_L workers are paid. We can draw no predictions about what combination of worker skill levels a firm—or an economy—will use, so we can learn nothing about low-skill-level equilibrium traps. In fact, there is empirical evidence for imperfect substitutability across worker types in firms.

To see why we must have complementarity of tasks, suppose that there were two tasks indexed by g and h but with no complementarity between them. To be specific, suppose that our q_H worker is hired for the g task, and a q_L worker is hired for the h task; then

$$F(q_H q_L) = g(q_H) + h(q_L)$$

Here skills are imperfect substitutes for each other, because only one type of worker can be hired for each task (i.e., no two-for-one type of substitution is possible here). However, because tasks are not complementary, the optimal choice of skill for the g task is independent of that of the h task, and again no strategic complementarities are present.[46]

Implications of the O-Ring Theory

The analysis has several important implications:

- Firms tend to employ workers with similar skills for their various tasks.
- Workers performing the same task earn higher wages in a high-skill firm than in a low-skill firm.
- Because wages increase in q at an increasing rate, wages will be more than proportionally higher in developed countries than would be predicted from standard measures of skill.
- If workers can improve their skill level and make such investments, and if it is in their interests to do so, they will consider the level of human capital investments made by other workers as a component of their own decision about how much skill to acquire. Put differently, when those around you have higher average skills, you have a greater *incentive* to acquire more skills. This type of complementarity should by now be a familiar condition in which multiple equilibria can emerge; it parallels issues raised in our analysis of the big push model. Kremer shows that a graph like Figure 4.1 can apply to choices about how much skill to acquire.
- One can get caught in economy-wide, low-production-quality traps. This will occur when there are (quite plausibly) O-ring effects across firms as well as within firms. Because there is an externality at work, there could thus be a case for an industrial policy to encourage quality upgrading, as some East Asian countries have undertaken in the past (see Chapter 12, section 12.6, and its end-of-chapter case study of South Korea). This could be relevant for a country trying to escape the middle-income trap.

- O-ring effects magnify the impact of local production bottlenecks because such bottlenecks have a multiplicative effect on other production.
- Bottlenecks also reduce the incentive for workers to invest in skills by lowering the expected return to these skills.

Following Kremer, consider a simple illustration of these bottleneck effects. Suppose that n tasks are required to produce a good. Let q be the standard skill level of these n tasks. But now let the actual skill level of two workers be cut in half in all firms. With an O-ring production function, output would fall by 75% (the result of cutting output in half once and then again). But then the marginal product of quality also falls by 75% for all the remaining $n - 2$ tasks, and thus so does the incentive to invest in increasing skill. The strong assumption of our simple O-ring production function may overstate the case, but the point that strategic complementarities can cause low-skill equilibria remains.

As workers reduce their planned skill investments, this further reduces the level of skill in the economy and thereby lowers further the incentive to invest in skill. To some extent, such bottlenecks could be ameliorated by international trade and investment, because foreign inputs and investors provide an alternative source of inputs from outside the bottlenecked economy. One explanation of why economies that have cut themselves off from the international economy, such as India or China before the 1980s, have not fared as well as those that are more integrated, such as South Korea, could well be their failure to take advantage of foreign inputs or investments; the O-ring analysis helps explain why the impact could be so great. Trade cannot solve all problems of industrialization, but the O-ring model helps explain why trade can play a key role as a part of an industrialization strategy.

The model also has implications for the choice of technology. When skill is scarce, a firm is less likely to choose a technique with higher value but complicated production technology with many tasks, because the costs of doing any one of those tasks poorly are magnified. In this way, the value of production is increasing in the complexity of the product, assuming that the product is completed successfully. Given positive assortative matching, firms producing products or using technologies that must be deployed at large scale or many steps will be induced to employ high-quality employees. Mistakes are costly to firms with large numbers of workers and production steps; therefore, such firms place exceptional value on high-quality, skilled workers who are unlikely to make mistakes.[47] This indicates one reason why rich countries with high-skill workers tend to have larger firms and specialize in more complex products; it also helps explain why firm size and wages are positively correlated within and across countries.

Finally, under some additional assumptions, the model can also help explain the international brain drain. It is often observed that when a worker of any given skill moves from a developing to a developed country, he or she immediately receives a higher wage for using those same skills. A version of the O-ring model is one way of explaining this.

Thus Kremer's O-ring model points out many of the implications of strong complementarities for economic development and the distribution of income across countries. As Kremer concludes, "If strategic complementarity is sufficiently strong, microeconomically identical nations or groups within nations could settle into equilibria with different levels of human capital."[48]

4.6 Economic Development as Self-Discovery

In simple models with perfect information, it is assumed that firms, and developing economies as a whole, already know their comparative advantage. But individuals must discover their own comparative advantage in labor markets; for example, no one is born knowing they are well suited to become an economist or international development specialist. Somewhat analogously, nations must learn what activities are most advantageous to specialize in. As Ricardo Hausmann and Dani Rodrik show, this is a complex task—and one prone to market failure.[49] It is not enough to tell a developing nation to specialize in "labor-intensive products," because even if this were always true, there are a vast number of such products in the world economy of today, and underlying costs of production of specific products can differ greatly from country to country. So it is socially valuable to discover that the true direct and indirect domestic costs of producing a particular product or service in a given country are low or can be brought down to a low level. It is valuable in part because once an activity is shown to be profitable, it can usually be imitated, at least after some lag, spawning a new domestic industry. An example is the ready-made garment industry in Bangladesh, which spread from the first pioneers as dozens of entrepreneurs entered the market. But as markets are eventually open to competing firms, they will take away potential profits from the original innovator. And since, due to this **information externality**, innovators do not reap the full returns generated by their search for profitable activities, there will be too little searching for the nation's comparative advantage—too much time carrying on with business as usual and too little time devoted to "self-discovery." The term *self-discovery* somewhat whimsically expresses the assumption that the products in question have already been discovered by someone else (either long ago, or recently in a developed economy); what remains to be discovered is which of these products a local economy is relatively good at making.

Information externality The spillover of information—such as knowledge of a production process—from one agent to another, without intermediation of a market transaction; reflects the public good characteristic of information (and susceptibility to free riding)—it is neither fully excludable from other uses, nor nonrival (one agent's use of information does not prevent others from using it).

Hausmann and Rodrik also point out another market failure: There can be too much diversification after the point where the nation discovers its most advantageous products to specialize in. This is because there may be an extended period in which entry into the new activity is limited. Hausmann and Rodrik conclude that in the face of these market failures, government policy should counteract the distortions by encouraging broad investments in the modern sector in the discovery phase. In fact, they also argue that policy should in some cases work to rationalize production afterward, encouraging movement out of higher-cost activities and into the lower-cost activities, paring down industries to the ones with the most potential for the economy. The authors draw parallels with some of the successful export and industrial policy experiences of East Asia, a topic to which we will return in Chapter 12.

The authors note three "building blocks" of their theory: There is uncertainty about what products a country can produce efficiently; there is a need for local adaptation of imported technology so that it cannot be used productively "off the shelf"; and once these two obstacles have been overcome, imitation is often rapid (reducing the profitability of pioneers). They present a number of case examples that show the reasonableness of each of these assumptions

in practice, such as the unexpected emergence of the information technology industry in India and the surprising differences in the exports from various countries with similar apparent comparative advantages, such as Bangladesh (hats but not bedsheets) and Pakistan (bedsheets but not hats); the history of local adaptations of various types of Western technology in East Asia (such as shipbuilding in South Korea); and the rapid diffusion of new products and techniques in the local economy (often facilitated by the movement of personnel across firms), as seen in the growth of the cut-flower export industry in Colombia.

4.7 The Hausmann-Rodrik-Velasco Growth Diagnostics Framework

Encouraging efficient investment and widespread entrepreneurship plays a prominent role in accelerating growth and promoting development more broadly. But the once popular idea of finding a "one size fits all" policy for economic development is now generally recognized as a myth. Different countries face different binding constraints on achieving faster rates of growth and economic development. A key mission for economic development specialists is to help determine the nature of the constraints for each country. Ricardo Hausmann, Dani Rodrik, and Andrés Velasco (HRV) propose a **growth diagnostics** decision tree framework for zeroing in on a country's most binding constraints on economic growth. HRV explain that targeting the most binding constraint has important advantages over other approaches to policy selection.[50]

Growth diagnostics A decision tree framework for identifying a country's most binding constraints on economic growth.

If a developing nation experiences a relatively low level of private investment and entrepreneurship, what steps should it take? The basic decision tree for addressing this question is seen in Figure 4.3, with arrows leading to the ten bottom boxes (that is, the boxes from which no arrows extend further). At the first stage of the tree, the analyst seeks to divide countries between those for which the main problem is a low underlying rate of return and those for which the problem is an abnormally high cost of finance. Let us consider the former case first, following the left arrow pointing to *Low return to economic activity.*

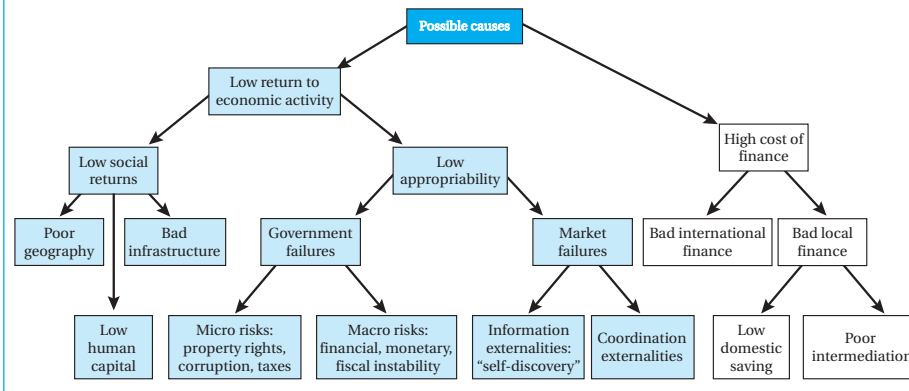
Low returns to investors may be due to the fact that there are intrinsically low underlying **social returns** to economic activities. Alternatively, low returns may be caused by what is termed *low private appropriability,* meaning limited ability of investors to reap an adequate share of the rewards of their otherwise profitable investments. Considering these cases in turn, *low social returns* may be caused by one of three factors.

Social returns The profitability of an investment in which both costs and benefits are accounted for from the perspective of the society as a whole.

First, as noted in Chapter 2, *poor geography* such as tropical pests, mountains, and other physical barriers, distance to world markets, and landlocked status (which may render port access politically dubious or economically costly) may limit the ability of a low-income country to initiate and sustain economic development, especially when other compounding factors are present. When these constraints are most binding, development policy must initially focus on strategies for overcoming them. Second, *low human capital*—skills and education as well as health of workers—are complementary with other factors in production, affecting the returns to economic activity. For

FIGURE 4.3 Hausmann-Rodrik-Velasco Growth Diagnostics Decision Tree

Problem: Low levels of private investment and entrepreneurship



Source: Ricardo Hausmann, Dani Rodrik, and Andrés Velasco, "Getting the diagnosis right," *Finance and Development* 43 (2006), available at http://www.inf.org/external/pubs/ft/fandd/2006/03/hausmann.htm. Reprinted with permission.

example, if economic returns are most affected by lack of literacy and numeracy, this becomes a development policy priority. (The importance of health and education was also stressed in Chapter 2, and this will be examined in depth in Chapter 8.) Third, every developing nation must provide the vital infrastructure needed to achieve and sustain a modern economy, beginning with basic physical structures such as roads, bridges, railroads, ports, telecommunications, and other utilities. With *bad infrastructure,* otherwise high-return economic activities may prove unprofitable. In some countries, inadequate and imbalanced infrastructure is the main factor preventing an acceleration of growth, and in such cases, policies focusing on providing it would boost investment and growth the most.

But the problem may lie not with the underlying social return to economic activities but with *low appropriability*, meaning that investors cannot reap an adequate share of returns to investment. We get to low appropriability from the right arrow emanating from Low return to economic activity. In turn, appropriability problems can be due to either *government failures* or *market failures.* In the HRV diagram, government failures are divided between *micro risks* and *macro risks*. Micro risks address fundamental institutional weaknesses such as inadequacy of property rights, government corruption, and excessively high effective taxation. That is, the return to economic activity may be high enough, but elites rather than investors may capture a large fraction of the returns and make investments unattractive. Despite the difficulty of effectively reforming institutions when reform threatens the interests of elites (see Chapter 2), such reform must become the development priority when micro risks are binding. As the case study of China at the end of this chapter

demonstrates, reform can sometimes be accomplished in stages through transitional institutions. Appropriability may also be limited by macro risks—the failure of government to provide financial, monetary, and fiscal stability.

The fundamental problem may also be large-scale market failures of the type stressed in this chapter. These may include the *self-discovery* problems pointed up by Hausmann and Rodrik and reviewed in section 4.6. They may also take the form of *coordination externalities*, such as seen in the big push model of underdevelopment, examined in section 4.3. Other types of market failure and government failure are examined in Chapter 11.

In yet other cases, the main problem may not be underlying low rates of return but rather an abnormally *high cost of finance*. The possibilities are outlined following the right arrow from the top box in Figure 4.3 to *High cost of finance*. Here the problem may be *bad international finance*—inadequate access to foreign sources of capital or problems with debt, examined in Chapter 13; or the problem may reside in *bad local finance*, due either to low availability of loanable funds through domestic financial markets, traced to low *domestic saving*, or to *poor intermediation* owing to an inadequate or overregulated banking system that is unable or unwilling to channel funds to the economic activities with high returns. These also lead to other policy challenges, examined in Chapter 15.

In sum, one size does not fit all in development policy. Economic development strategies focusing on resource mobilization through foreign assistance and other capital flows, along with increased domestic national saving, can be most effective when domestic returns are both high *and* privately appropriable. In contrast, strategies focusing on market liberalization and opening up the economy can be most effective when social returns are high and the most serious obstacle to private appropriation are government-imposed excessive taxes and restrictions. Finally, strategies focusing on industrial policy (elaborated on in Chapter 12) can be most effective when private returns are low, not because of what a government does (errors of commission), but because of what a government does not do (errors of omission).

HRV illustrate their approach with case studies of El Salvador, Brazil, and the Dominican Republic. They argue that each case exhibits a different "diagnostic signal" of the most binding constraint, as seen in Box 4.3. HRV stress that an approach to development strategy that determines one or two policy priorities on this diagnostic basis will be more effective than pursuing a long laundry list of institutional and governance reforms that may not be targeted toward the most binding constraints.

It is often difficult to observe a binding constraint directly. In practice, growth diagnostics usually involves some economic detective work. To evaluate whether a proposed constraint is binding, a growth diagnostician looks for evidence on its implications. If the constraint is excessive taxation, we can expect to see high movement into the informal sector or underground economy. If the constraint is infrastructure, we can expect to see significant congestion. If the constraint is education, we can expect to see high rates of return to education. In general, the analyst looks for economic behavior consistent with agents trying to get around a constraint.

Growth diagnostics is also subject to some limitations and criticisms. One implicit assumption is that development can be equated with growth, which

BOX 4.3 FINDINGS Three Country Case Study Applications of Growth Diagnostics

El Salvador

HRV argue that this economy is constrained by a lack of productive ideas. The binding constraint is a lack of innovation and demand for investment to replace the traditional cotton, coffee, and sugar sectors, or low "self-discovery." So the best strategy focus for El Salvador would be to encourage more entrepreneurship and development of new business opportunities.

Brazil

HRV identify the country's binding constraint as lack of sufficient funds to invest despite an abundance of productive ideas. They argued that private returns in Brazil are high, and therefore other flaws (inadequate business environment, a low supply of infrastructure, high taxes, high prices for public services, weak contract enforcement and property rights, and inadequate education) are not as binding in Brazil. So investment is instead constrained by Brazil's inability to mobilize sufficient domestic and foreign savings to finance needed investments at reasonable interest rates. Although Brazil could increase national savings to a degree by reducing government expenditure, this might not be politically feasible. If so, HRV suggest that higher taxes and user fees and lower infrastructure and human capital subsidies might work. "If the country can move to a faster growth path and if waste does not grow with GDP, it may outgrow its burdens and gradually improve its tax and spending system as fiscal resources become more abundant." In subsequent work, Hausmann has emphasized the importance of "creating a financially viable state that does not over-borrow, over-tax or under-invest" to successfully raise domestic savings.

Dominican Republic

HRV conclude that the Dominican Republic is constrained by core public goods in product sectors key for growth. The country began a new reform sequence during the 1980s, after it could no longer rely on sugar and gold exports. It followed a narrow strategy of investing in needed public goods for two emerging product (or service) sectors with high potential, tourism and *maquila* assembly manufacturing. The keys were security and infrastructure near the main tourist destinations and special trade policy benefits for the light manufacturing assembly (*maquila*) sector. As the economy grew from these sources, other constraints were hit, notably in the financial sector; getting past them (particularly a costly financial crisis) was bumpy, but the binding constraints stayed or became visible, so policymakers could focus on relaxing them to keep growth going.

Sources: Ricardo Hausmann, Dani Rodrik, and Andrés Velasco, "Growth diagnostics," in *One Economics, Many Recipes: Globalization, Institutions, and Economic Growth*, by Dani Rodrik (Princeton, N.J.: Princeton University Press, 2007), ch. 2; Ricardo Hausmann, "In search of the chains that hold Brazil back," October 31, 2008, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1338262. An excellent practicum is found in "Doing Growth Diagnostics in Practice: A 'Mindbook.'" See http://www.cid.harvard.edu/cidwp/177.html. The World Bank offers a set of growth diagnostics exercises at its Web site, http://web.worldbank.org/.

in turn is held back by investment. This is a useful analytical assumption for this and a range of other purposes, but it does not and cannot provide a complete understanding of development purposes, mechanisms, and constraints. And of course, it is often not a simple matter to find a single binding constraint. There can be uncertainty about the "position" of each constraint in the economy, so we can only make a probabilistic assessment of which one is binding. If there are important complementarities between two investments,

combining them (in some sense) should be considered. Further, the fact that one constraint is not binding today does not mean that we can neglect it when there are long gestation periods before current investments become productive. For example, consider investments in education: Students require several years of schooling followed by experience before these investments become productive. So although education may not be binding for a particular country such as Bolivia at a particular point in time, this does not mean that it will not become binding at a later time; in response, we may need to make investments today. Clearly, identifying and addressing constraints that are likely to become binding in the future is even more challenging than targeting today's more visible bottlenecks.

Growth diagnostics has already had an effect on the work of development agencies. For example, the Inter-American Development Bank (IDB), the regional development bank for the western hemisphere, has been commissioning growth diagnostic studies of many member economies while training staff and nationals in the skills needed to conduct their own growth diagnostics. World Bank economists have applied the method in a dozen country pilot studies in Africa, Asia, and Latin America. And developing country scholars have applied the approach to their own countries. Although growth diagnostics might be criticized as "more art than science," at the very least this new approach forces the analyst to focus on country-specific circumstances and thus to get to know the individual country very well. This is one of the reasons that growth diagnostics offers a valuable complement to econometric studies.

4.8 Conclusions

The important point is not that people keep doing inefficient things. This is not in itself very surprising. The deeper point is that people keep doing inefficient things because it is rational to keep doing them, and it will remain rational as long as others keep doing inefficient things. This leads to a fundamental problem of coordination failure. Sometimes firms and other economic agents will be able to coordinate to achieve a better equilibrium on their own. But in many cases, government policy and aid will be necessary to overcome the resulting vicious circles of underdevelopment.

The purpose of economic development theory is not only to understand underdevelopment but also to devise effective policies to redress it. The analysis of coordination failure problems in this chapter affirmed that early development theorists such as Paul Rosenstein-Rodan identified important potential problems that are ignored in conventional competitive equilibrium models.[51] The new perspectives offer some important overall lessons for policy, but they are not simple lessons with easy applicability, and indeed they present something of a two-edged sword. On one side, the analysis shows that the potential for market failure, especially as it affects the prospects for economic development, is broader and deeper than had been fully appreciated in the past. Rather than the small "deadweight triangle losses" of conventional economic

analysis of monopoly, pollution externalities, and other market failures, coordination failure problems can have more far-reaching effects and consequently much greater costs.[52] For example, the interactions of slightly distorted behaviors by potential investors failing to consider the income effects of the wages they pay may produce very large distortions, such as the outright failure to industrialize. This makes the potential benefit of an active role for government larger in the context of multiple equilibria.

The coordination failures that may arise in the presence of complementarities highlight potential policies for deep interventions that move the economy to a preferred equilibrium or even to a higher permanent rate of growth that can then be self-sustaining. For example, once a big push has been undertaken, government coordination may no longer be necessary. The unaided market can often maintain industrialization once it is achieved, even when it cannot initiate or complete the process of industrialization. For another example, we will see in Chapter 8 that in some cases, the presence of child labor represents a kind of bad equilibrium among the families with children who work, one that might be fixed with appropriate policy. After successfully abolishing child labor, it is possible that the regulations will not have to be actively enforced to keep child labor from making a resurgence (because most parents send their children to work only because they have to). If there is no incentive to go back to the behavior associated with the bad equilibrium, government has no need to continue the interventions. Instead, government can concentrate its efforts on other crucial problems in which it has an essential role (e.g., problems of public health). This onetime-fix character of some multiple-equilibria problems makes them worthy of special focus because they can make government policy much more powerful in addressing problems of economic development. Among other implications, the prospect of deep interventions can mean that the costs of implementing policy can be reduced and that carefully targeted development assistance could have more effective results.

The other edge of the sword, however, is that with deep interventions, the potential costs of a public role become much larger. Policy choices are more momentous because a bad policy today could push an economy into a bad equilibrium for years to come. This is because government can be a major part of the problem, playing a key role in perpetuating a bad equilibrium such as a high-corruption regime, in part because some government officials and politicians may benefit personally from it. Bad policy can even initiate a move to a worse equilibrium than a country began with. To expect government to be the source of reform that moves the economy to a better equilibrium in countries where government has been part of the complex nexus of a bad equilibrium can be naive. For example, as the 2001 Nobel laureate Joseph Stiglitz pointed out, development officials should have been more suspicious of corrupt government officials' embracing of the World Bank's doctrine of thoroughgoing privatization in the late 1980s and early 1990s. Why would corrupt officials have done so if they benefited from a stream of rents captured from public enterprises? The answer, Stiglitz suggests, is that these officials found that by corrupting the process of privatization, they could get not only a stream of corrupt rents from the annual operations of the enterprise but also

a share of the present discounted value of the whole future operations of the enterprise.[53] The results of corrupt privatization in Russia in particular have been devastating for its economy, preventing it from enjoying the benefits of the market and potentially keeping it in a suboptimal equilibrium for many years to come. Even when a government is not corrupt, the potential impact of a well-intentioned but flawed government policy is much greater when it can push the economy to a fundamentally different equilibrium, which may be difficult to reverse. This is all the more problematic in the many cases in which "history matters" in a developing economy—that is, when past conditions determine what is possible today.

Both government failure and market failure (including coordination problems and information externalities) are real, but public- and private-sector contributions to development are also vital. Therefore, we need to work toward the development of institutions in which actors in the public and private sectors have incentives to work productively together (directly and indirectly) in such a way as to create the conditions necessary to break out of poverty traps. In achieving this goal, the international community also has a vital role to play, providing ideas and models and serving as a catalyst for change, as well as providing some of the necessary funding.

The growth diagnostics approach is a valuable tool for domestic and international analysts who start with a detailed understanding of a developing country; it can be helpful in identifying binding constraints on national growth and the policy priorities to address them.

In sum, the contributions of the new theories of development reviewed in this chapter include a better understanding of the causes and effects of poverty traps, achieved by more precisely pinning down roles of different types of strategic complementarities, explaining the role of expectations, clarifying the importance of externalities, illuminating the potential scope for deep interventions, and improving our understanding of both the potential role of government and the constraints on the effectiveness of that role—when government itself becomes a player in an underdevelopment trap. Finally, the new approaches point out more clearly the real potential contributions of outside development assistance that extend beyond provision of capital to modeling new ways of doing things.

As democratic government spreads in the developing world, the new understandings of underdevelopment traps can make for a more effective guide to policy design than was available even a few years ago. As Karla Hoff has aptly summarized, "Governments fail, even in democracies, just as markets do. But a positive development of recent years is to try more limited interventions to harness the spillovers among agents, and to try to sequence policy reforms in a way that makes it more likely for good equilibria to emerge."[54]

In Parts Two and Three, as we consider pressing issues affecting developing countries today, we will be using the insights provided by both the classic theories and the new models of development and underdevelopment to inform our understanding of both the nature of the problems faced and the potential benefits and pitfalls of policies designed to help overcome them.