

COPYRIGHT NOTICE:

Samuel Bowles: Microeconomics

is published by Princeton University Press and copyrighted, © 2003, by Princeton University Press. All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher, except for reading and browsing via the World Wide Web. Users are not permitted to mount this file on any network servers.

For COURSE PACK and other PERMISSIONS, refer to entry on previous page. For more information, send e-mail to permissions@pupress.princeton.edu

PROLOGUE

Economics and the Wealth of Nations and People

[Economics is the study of] human behavior as a relationship between given ends and scarce means.

—Lionel Robbins, *An Essay on the Nature and Significance of Economics* (1935)

An economic transaction is a solved political problem. . . . Economics has gained the title Queen of the Social Sciences by choosing solved political problems as its domain.

—Abba Lerner, “The Economics and Politics of Consumer Sovereignty” (1972)

TO ITS FOUNDERS, the subject of political economy was the wealth of nations and people.

In the fourteenth century, Ibn Battuta, one of the leading geographers and explorers of his age, traveled widely in Asia, Africa, the Middle East, Russia, and Spain. In 1347, he visited the land we now call Bangladesh. “This is a country . . . abounding in rice,” he wrote. He described traveling along its waterways, passing “between villages and orchards, just as if we were going through a bazaar.”¹ Six centuries later, a third of the people of Bangladesh are undernourished and the country is among the world’s poorest.

At the time of Ibn Battuta’s visit to Bangladesh, Europe was reeling under the impact of the bubonic plague, which took the lives of a quarter or more in many cities. Manual workers in London, probably among the better off anywhere on the continent, consumed fewer than 2000 calories per day.² The shortage of labor following the plague somewhat boosted real wages through the middle of the next century, but over the next four centuries, real wages of laborers did not rise in

The first epigraph comes from Robbins (1935:16), the second from Lerner (1972: 259).

¹ His account is published in Ibn Battuta (1929:267, 271). A second source (Yule 1886:457) quotes him as observing, “I have seen no region of the earth in which provisions are so plentiful,” but this may be a mistranslation due to Yule or to the French source on which he relied.

² This account follows Allen (2001). The wage series below can be found at <http://www.econ.ox.ac.uk/members/robert.allen/wagesprices.htm>.

any European city for which records exist; in most, wages fell by substantial amounts—in Northern Italy to half their earlier level. Over the past two centuries, however, real wages rose dramatically, first in England, where they increased *ten-fold*, and somewhat later but by even greater amounts in other European cities.

What accounts for these dramatic reversals of fortune? The most plausible answer, very briefly, runs as follows. The emergence and diffusion of a novel set of institutions that came to be called *capitalism* brought about a vast expansion in the productivity of human labor. This led to higher wages when workers' bargaining power was eventually augmented by the expansion of workers' political rights and by the drying up of the pool of new recruits from agriculture, household production, and other parts of the economy that were not organized according to these new institutions. This happened in Europe and not in Bangladesh.

What *did* happen in Bangladesh, as in much of the Mughal Empire and what became British India, was a growing entrenchment of the power and property rights of powerful landlords. Their influence was already substantial before the British, but during the Bengal Presidency it was greatly strengthened by the Permanent Settlement of 1793. This act of the colonial rulers conferred de facto governmental powers on the landlords by giving them the right to collect taxes (and to keep a substantial fraction for themselves). The fact that British taxation and land tenure policy was not uniform throughout the Raj provides a natural experiment to test the importance of these institutions for subsequent patterns of backwardness or development. Banerjee and Iyer (2002) compared the post-Independence economic performance and social indicators of districts of modern-day India in which landlords had been empowered by the colonial land tenure and taxation systems with other districts in which the landlords had been bypassed in favor of the village community or direct taxation of the individual cultivator. They found that the landlord-controlled districts had significantly lower rates of agricultural productivity growth stemming from lower rates of investment and lesser use of modern inputs. The landlord-controlled districts also lagged significantly in educational and health improvements.³ These findings suggest a remarkable persistence of the effects of an institutional innovation that occurred a century or more earlier.

³ The details of the causal connection between landlord control and these subsequent results remain to be explored. Because colonial practices changed over time in response to exogenous events (such as the revolt by Indian soldiers in 1857) and over space in response to the idiosyncracies of local administrators, Banerjee and Iyer were able to identify independent sources of variation in the land tenure and taxation policies not due to pre-existing conditions.

The effects of institutions on economic performance is further affirmed by a dramatic turn in land tenure in the Indian state of West Bengal.⁴ Following its election in 1977, the Left Front government of the state implemented a reform under which sharecroppers who registered with the Department of Land Revenue were guaranteed permanent and inheritable tenure in the plots they cultivated as long as they paid the landlord a quarter of the crop. Prior to the reform, the modal landlord's crop share had been one half, and landlord's had routinely used the threat of eviction to enhance their bargaining power with the sharecroppers. The cultivators' increased crop-share significantly increased the incentives to work the land productively. The security of tenure had two possibly offsetting effects: it enhanced the cultivators' incentive to invest in the land, while restricting the ability of the landlord to elicit high levels of output by threat of eviction. A further indirect effect may have also been at work. The increased economic security of the sharecroppers led to their more active participation in local politics; partly as a result, the local councils—the *panchayats*—became more effective advocates of the interests of the less well-off in the acquisition of agricultural inputs, credit, and schooling.

The effects of the reform have been estimated from a comparison of agricultural productivity between West Bengal and neighboring Bangladesh (a similar region in which no such reforms were implemented) and by exploiting the fact that the implementation of the reform (measured by the fraction of sharecroppers registering for its benefits) varied considerably within West Bengal. The resulting estimates are imprecise, and it remains difficult to determine which causal mechanisms were at work, but the effects of the reform appear to have been very substantial: rice yields per hectare on sharecropped land were increased by about fifty percent. Having lagged behind most Indian states prior to the reform, agricultural productivity growth in West Bengal has been among the most rapid since the reform.

The enduring importance of institutions is likewise suggested by the work of Sokoloff and Engerman (2000) concerning an analogous New World reversal of fortune. They estimate that in 1700 Mexico's per capita income was about that of the British colonies that were to become the United States, while Cuba and Barbados were at least half again richer. At the close of the eighteenth century Cuba had slightly higher per capita incomes than the United States, and Haiti was probably the richest society in the world. At the opening of the twenty-first century, however, the per capita income of Mexico was less than a third of the United States', and Haiti's was lower yet. In a series of papers, Sokoloff

⁴ This account is based on Banerjee, Gertler, and Ghatak (2002) and Bardhan (1984).

and Engerman provide the following explanation.⁵ In the parts of the New World in which sugar and other plantation crops could be grown (Cuba and Haiti) or in which minerals and indigenous labor were abundant (Mexico), economic elites relied on bonded labor or slaves, and consolidated their power and material privileges by means of highly exclusive institutions. These restricted access by the less well-off to schooling, public lands, patent protection, entrepreneurial opportunities, and political participation. As a result, over the ensuing centuries, even after the demise of slavery and other forms of coerced labor, opportunities for saving, innovation, and investment were monopolized by the well-to-do. Literacy remained low, and land holding highly concentrated. As the source of wealth shifted from natural resource extraction of manufacturing and services, these highly unequal economies stagnated while the far more inclusive economies of the United States and Canada grew rapidly. The ways their less exclusive institutions contributed to the success of these North American economies remains somewhat unclear, but a plausible hypothesis is that broader access to land, entrepreneurial opportunities, and human capital stimulated growth.

The source of the institutional divergence among the colonies of the New World appears to be their initial factor endowments, more than the distinct cultures or colonial policies of the European states that conquered them. British Belize and Guyana went the way of Spanish Honduras and Colombia; Barbados and Jamaica went the way of Cuba and Haiti. The Puritans who settled Providence Island off the coast of Nicaragua forsook their political ideals and became slave owners. Slaves on the island outnumbered the Puritans when it was overrun by the Spanish in 1641. According to its leading historian, “[T]he puritan settlement . . . with its economy fueled by privateering and slavery looked much like any other West Indian colony” (Kupperman 1993, p. 2). At the time of its demise, Providence Island was attracting migrants from the more famous Puritan colony far to the north; two boatloads of hapless Pilgrims arrived from Massachusetts just after the Spanish takeover.

A final example is provided by the precipitous collapse of Communist Party rule in the Soviet Union and its Eastern European allies around 1990 and the transition of the new states to market-based economies. Figure P.1, presenting the levels of gross domestic product (GDP) per capita relative to the year 1990 for fourteen of these nations, reveals dramatic differences in their trajectories. After a decade of transition,

⁵ See also Engerman, Sokoloff, and Mariscal (2002) and Acemoglu, Johnson, and Robinson (2002).

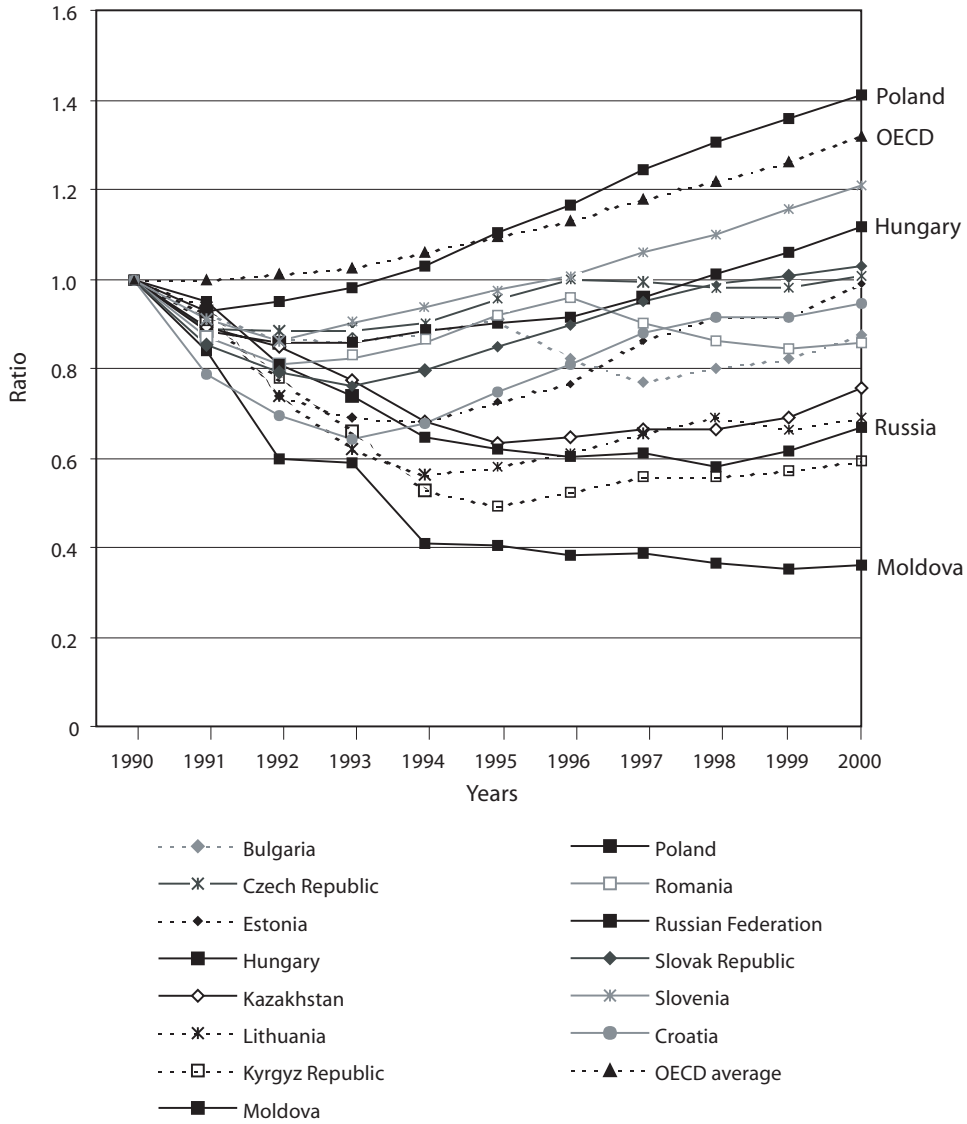


Figure P.1 Divergence of ex-Communist economies' real GDP per capita (relative to 1990). Source: World Bank (Statistical Information Management Analysis data base).

Poland's per capita income stood at 40 percent above the initial level, while Russia's had declined by a third and Moldavia's had fallen to less than 40 percent of the initial level. Over the same period China's per capita income more than doubled (not shown). Among these fourteen economies, only Poland out-performed the (unweighted) average of the OECD economies.

While the success of China's gradual reforms has been the subject of extensive study, the differences among the countries undertaking a rapid transition are poorly understood. A possible explanation is that, starting from quite similar institutions, small differences in the content or timing of reform packages or chance events resulted in large and cumulative differences in performance, because some countries (e.g., Hungary and Poland) were able to capture the synergistic effects of institutional complementarities while others were not (Hoff and Stiglitz 2002). Other explanations stress the substantial institutional differences among the countries or their differing levels of trust or other social norms. What is not controversial is that divergences in performance of this magnitude, emerging in less than a decade, suggest both the importance of economic institutions and the pervasive influence of positive feedback effects, whereby both success and failure are cumulative.

I have deliberately chosen cases that dramatize the central role of institutions. Other comparisons would suggest different, or at least less clear-cut conclusions. Over the period 1950–1990, for example, countries with democratic and authoritarian regimes appear to have differed surprisingly little in their overall economic performance (controlling for other influences) with major differences appearing only in their demographic record, with slower population growth in democracies (Przeworski, Alvarez, Cheibub, and Limongi 2000). Nonetheless, the examples above—the divergence of living standards in Europe from many parts of the world, the reversal in New World fortunes, and the heterogeneous consequences of economic liberalization in the once-Communist nations—are of immense importance in their own right and, as subsequent examples show, are hardly atypical.

What can modern economics say about the wealth and poverty of nations and people? No less important, what can it *do*?

CONTRARY to its conservative reputation, economics has always been about changing the way the world works. The earliest economists—the Mercantilists and the Physiocrats—were advisors to the absolute rulers of early modern Europe; today's macroeconomic managers, economic development advisors, and architects of the transition from Communism to market-based societies, continue this tradition of real world engagement. Economists have never been strangers to policy making

and constitution building. The hope that economics might assist in alleviating poverty and securing the conditions under which free people might flourish is at once its most inspiring calling and its greatest challenge.

Like many, I was drawn to economics by this hope. Having been a schoolboy in India and a secondary school teacher in Nigeria before turning to economics, I naturally came to the field expecting that it would address the enduring problem of global poverty and inequality. At age eleven I had noticed how very average I was among my classmates at the Delhi Public School—in sports, in school work, in just about everything. A question has haunted me since: how does it come about that Indians are so much poorer than Americans given that as people we are so similar in our capacities? And so I entered graduate school hoping that economics might explain, for example, why workers in the United States produce almost as much in a month as those in India produce in a year, and why the Indian population is correspondingly poor (Hall and Jones 1999). We now know that the conventional economic explanations fail: by any reasonable accounting, the differences in the capital-labor ratio and in the level of schooling of the U.S. and Indian workforces explain much less than half of the difference in productivity. It seems likely that much of the gap results from causes more difficult to measure and, until recently, less studied by economists: differences in historical experience, institutions, and conventional behaviors. These are the subject matter of this book.

Alfred Marshall's (1842–1924) *Principles* was the first great text in neoclassical economics. It opens with these lines:

Now at last we are setting ourselves seriously to inquire whether it is necessary that there should be any so called “lower classes” at all: that is whether there need be large numbers of people doomed from their birth to hard work in order to provide for others the requisites of a refined and cultured life, while they themselves are prevented by their poverty and toil from having any share or part in that life. . . . [T]he answer depends in a great measure upon facts and inferences, which are within the province of economics; and this is it which gives to economic studies their chief and their highest interest. (Marshall 1930:3–4)

Marshall wrote this in 1890. I suspect he would be disappointed by the progress economics made towards these lofty aims in the century that followed.

THE NEOCLASSICAL PARADIGM that Marshall helped found was ill-suited to the task he set. Its defining assumptions precluded analysis of many key aspects of economic progress and stagnation, among them the exer-

cise of power, the influence of experience and economic conditions on people's preferences and beliefs, out-of-equilibrium dynamics, and the process of institutional persistence and change.

Drawing on the contributions of many—economists and others—this book presents a theory of how individual behaviors and economic institutions interact to produce aggregate outcomes, and how both individuals and institutions change over time. It is based on assumptions that are quite different from those that define the neoclassical paradigm. In what follows, I use the term *Walrasian paradigm* (for Leon Walras [1834–1910], another of the founders of neoclassical economics) in preference to the more open-ended term “neoclassical.” By *Walrasian* I mean that approach to economics that assumes that individuals choose actions based on the far-sighted evaluation of their consequences based on preferences that are self-regarding and exogenously determined, that social interactions take the exclusive form of contractual exchanges, and that increasing returns to scale can be ignored in most applications. With some refinement, these assumptions account for the distinctive analytical successes and normative orientation of the Walrasian approach. The term *paradigm* refers to the core subject matter taught to students.

The approach developed here retains many of the fundamental tenets of the Walrasian paradigm and of the classical school that it superceded. Among these are a familiar triplet of ideas: that when individuals act they are trying to accomplish something; that intentional action is constrained by the effects of competition; and that the aggregate outcomes of large numbers of individuals interacting in this manner are typically unintended. These tenets have provided the foundation for the development of economics since its inception, and account for its many analytical insights. Other aspects of the Walrasian paradigm, however, are replaced.

The Walrasian approach represents economic behavior as the solution to a constrained optimization problem faced by a fully informed individual in a virtually institution-free environment. Robbins' celebrated definition of the subject (in the epigraph) reflects this equation of economics with constrained optimization. The passage of time is represented simply by a discount rate; people do not learn or acquire new preferences over time; institutions do not evolve. The actions of others are represented by nothing more complicated than a given vector of market-clearing prices, while proximity is captured by a cost of transportation. Property rights and other economic institutions are represented simply by a budget constraint. An economic actor in this model is roughly Robinson Crusoe, with prices standing in for nature. The economist's Crusoes inhabit a world in which goods are scarce, but whatever institutions are necessary to coordinate their activities in an

optimal manner are freely available. The “supply” of optimal institutions can thus be ignored for the same reason that Adam Smith used to explain why economists need not theorize about the value of water: they are free goods.

This description of the Walrasian paradigm is a caricature, of course, but a recognizable one, of the economics taught in leading doctoral programs as recently as the early 1980s. Since then a combination of new analytical tools—especially game theory and information economics—and the increasingly evident empirical inadequacies of the Walrasian model have combined to alter the way economics is taught and practiced. Economic agents no longer interact simply with nature or some other parametric environment, but also with each other, and strategically. Their interactions are no longer fully described by the prices of the goods they exchange because some aspects of their transactions are not expressed in enforceable contracts.

Nonetheless, in practice, even as some of the standard Walrasian assumptions are dropped, common tenets of the older paradigm are evident in many of the new approaches. Robert Solow expressed these as “equilibrium, greed, and rationality,” meaning that when economists “explain” something—say, unemployment—they mean that it can be represented as a unique stationary outcome in a model of interactions among self-interested individuals with advanced cognitive capacities and predispositions. Other ways of “explaining” unemployment may be entertained, but this is the default option. Solow’s concern about the adequacy of the trinity of core tenets is increasingly supported by both empirical and conceptual advances.

The approach I present here is based on the more modest, but perhaps more enduring, classical tenets of intentional action, competition, and unintended consequences. Just as the Walrasian paradigm assumes a particular kind of social interaction as the standard case—caricatured as Robinson Crusoe above—the approach here is designed to illuminate a generic situation based on the three empirically observed characteristics of structures of social interaction, individual behaviors, and technologies, introduced below. Here I simply outline the salient facts of these generic interactions and point to some important implications. I take up the task of modeling these interactions (and providing some relevant empirical evidence) in the subsequent chapters.

Non-contractual social interactions. When individuals interact, it is the exception, not the rule, that everything passing between them is regulated by a complete and readily enforced contract. Instead, noncontractual social interactions are ubiquitous in neighborhoods, firms, families, environmental commons, political projects, and markets. While many of these noncontractual social interactions take place in non-

market settings, they are also important in determining economic outcomes in highly competitive markets. Thus, in the pages that follow, I treat the grocery market with complete contracting—a staple of introductory economics textbooks—as a special case. The generic case is illustrated by labor markets and credit markets, where the promise to work hard or to repay the loan is unenforceable, or local environmental commons problems, where individual resource exploitation imposes noncontractible spillovers on others. A characteristic of markets with incomplete contracts is that one or both participants in a simple dyadic transaction typically receive rents, that is, payments above their next best alternative. In labor and credit markets, some workers and borrowers are unable to transact the quantities they prefer at the going terms of exchange; that is, they are quantity constrained, and the resulting markets do not clear in equilibrium, exhibiting excess supply (e.g., of labor) or excess demand (for loans).

If many aspects of economic interactions are not governed solely by contracts, how *are* they governed? The answer is that noncontractual aspects of interactions are governed by a combination of norms and power. An employment contract does not specify any particular level of effort, but the employee's work ethic or fear of job termination or peer pressure from workmates may accomplish what contractual enforcement cannot.

The idea that power is exercised in competitive market transactions will strike some readers as a commonplace; but to others it will appear a contradiction in terms. To neoclassical economists (like Abba Lerner, in the epigraph), “[A] transaction is a solved political problem.” It is “solved” by the device of complete contracts, so that everything of interest to all parties to a transaction can be enforced by the courts. With all the terms of a transaction contractually specified, nothing is left for the exercise of power to be *about*. For the same reason, norms are redundant: if the employee's contract were to specify a given amount of work for a given amount of pay and if work effort were readily verifiable, then the employer would care little about the work ethic of the employees. Relaxing the complete contracting assumption thus not only explains why many markets do not clear, it also reveals an important economic role for both power and norms, bringing the theory closer to the way observers and participants view real world exchanges.

Adaptive and other-regarding behaviors. Recent behavioral experiments by economists (confirming and extending earlier work by other social scientists) as well as observation in natural settings suggests a reconsideration of both the “rationality” and “greed” tenets in Solow's trinity. Individuals intentionally pursue their objectives, but they do this more often by drawing on a limited repertoire of behavioral responses

acquired by past experience than by engaging in the cognitively demanding forward-looking optimizing processes assumed alike by the Walrasian approach and by much of classical game theory. In many situations, emotions such as shame, disgust, or envy combine to produce a behavioral response. Moreover, while self-interest is a powerful motive, other-regarding motives are also important. In experiments and in real life, people frequently are willing to reduce their own material well-being not only to improve that of others but also to penalize others who have harmed them or others, or violated an ethical norm. These so-called *social preferences* help explain why people often cooperate toward common ends even when defection would yield higher material rewards, why incentive schemes based on self-interest sometimes backfire, and why firms do not sell jobs.

Thus models whose *dramatis personae* are simply identical individuals conforming to the self interest axioms of *Homo economicus* are often unilluminating. For many questions, adequate models must take account of the fact, confirmed in experiments and in natural settings, that people are both *heterogeneous*—some more self-interested, others more civic minded, for example—and *versatile*—actions adapting to situations rather than reflecting a single, all-purpose behavioral predisposition. As a result of both behavioral heterogeneity and versatility, small differences in institutions can make large differences in outcomes, some situations inducing selfish individuals to act cooperatively and others inducing selfish behaviors by those predisposed to cooperate.

Economists have commonly regarded behaviors that violate the stringent canons of formal rationality as idiosyncratic, unstable, or irrational, in short, not exhibiting the regularities that would allow scientific analysis. But the fact that experimental subjects consistently exhibit such “irrationalities” as intransitivity, loss aversion, inconsistency in temporal discounting, and the overvaluation of low probability events, suggests these behaviors are not only common but also susceptible to analysis.

People acquire their behavioral responses in part by copying the behaviors of those who, in similar situations, they perceive as successful by some standard or by acting to maximize one’s gains given one’s beliefs about how others will act. But other influences are also at work, including conformism and other types of frequency-dependent learning unrelated to the payoffs associated with behaviors. As a result, predictions of behavior based on forward-looking maximization of payoffs may be quite misleading. Moreover, behavioral responses acquired by individuals in one environment are unlikely to be acquired by the same individuals were they to be functioning in an entirely different environment. In this sense, not only individual beliefs (about the consequences

of their actions) but also individual preferences (their evaluations of the outcomes) are endogenous. The “given ends” invoked by Robbins is a useful simplification in many analytical tasks but is an arbitrary and misleading restriction in others.

Generalized increasing returns. Economic and other social interactions often lead to patterns of what Gunnar Myrdal (1956) termed “cumulative causation,” or what are now called “positive feedbacks.” *Positive feedbacks* include economies of scale in production, but the term refers more broadly to any situation in which the payoff to taking an action is increasing in the number of people taking the same action. More generic illustrations include, for instance, the fact that the payoff to learning a particular language depends on the number of speakers or that the payoff to engaging in a collective action depends on the number of participants. To distinguish this large class of positive feedback cases from the subset based on increasing returns to scale in production, I will use the term *generalized increasing returns* rather than increasing returns to scale. Institutional synergies may generate generalized increasing returns. For example, private ownership of property, competitive markets, and the rule of law often implement highly efficient solutions to allocational problems, but only if all three components are present and almost all members of the society adhere to these principles. Generalized increasing returns due to these institutional complementarities appears to be a source of divergence in the growth trajectories of the New World and ex-Communist economies mentioned above. Generalizing increasing returns may help to account for the increase in inequality among the peoples of the world over the past century and a half, despite the catching up of Japan, China, and other East Asian nations.⁶

These positive feedbacks create economic environments in which small chance events have durable consequences over very long time frames, and in which initial conditions may have persistent so-called “lock-in” effects. The “poverty traps” faced by peoples and nations as well as the “virtuous circles” of affluence enjoyed by others exhibit the effects of these influences. In the presence of generalized increasing returns, typically there exist more than one stationary outcome with the property that small deviations from that outcome are self-correcting. These multiple stable equilibria may be displaced by what appear in our models as exogenous shocks, mutations, or idiosyncratic play, but that in the real world take the form of wars, climatic changes, strikes, or other events not included in the model under examination.

A result may be infrequent but dramatic periods of change in institutions, behaviors, technologies, and the like as a population moves from

⁶ See Bourguignon and Morrison (2002) and the works cited there.

the neighborhood of one equilibrium to another, often followed by long periods of stability. Biologists use the term *punctuated equilibria* to refer to this alternating pattern of stasis and rapid change (Eldredge and Gould 1972). The collapse of Communism is an example. Another is the demise of foot binding of young women in China. This painful and disabling practice endured for a millennium, resisting attempts to end it over the centuries, yet it disappeared in the course of just a decade and a half in the early part the last century (Mackie 1996). The existence of multiple equilibria may also explain why seemingly similar populations may come to have quite different norms, tastes, and customs, often resulting in the widely observed pattern of *local homogeneity and global heterogeneity*, distinctive national cuisines and food tastes providing an example.

There is no reason and little evidence to suggest that the institutions and behaviors that result from processes in which generalized increasing returns are at work are in any sense optimal. Following the fall of Communism in the Soviet Union and Eastern Europe, for example, many economists confidently predicted that once state property was abolished, a workable configuration of capitalist institutions would spontaneously emerge. But in Russia and many of the other transitional economies, a decade of lawlessness and kleptocracy implemented a massive concentration of wealth under institutions providing few incentives for enhanced productivity or investment. The disappointing economic results of the end of Communist rule in these countries underlines the fallacy of the conventional view that good institutions are free in a world of material scarcity.

In the pages that follow, institutions, like goods, are taken to be scarce. The three basic assumptions outlined above—the noncontractual nature of social interactions, adaptive and other-regarding behaviors, and generalized increasing returns—define the generic case, my default option. The three are related. Relaxing the complete contracting assumption without modifying the behavioral assumptions of Walrasian economics is untenable, for the importance of other-regarding preferences, as we will see, is considerably enhanced when contractual incompleteness is taken into account. Similarly, the process by which preferences evolve exhibits strong generalized increasing returns. The reason is that norms generally take the form of conventions, adherence to which is in one's interest only as long as most others do the same. So relaxing the conventional behavioral assumptions raises doubts about nonincreasing returns. Finally, if generalized increasing returns are common, many different outcomes may be equilibria. Of these, the states most likely to be observed will depend critically on institutions governing the relevant dynamics, including such things as the exercise of power, collective action, and other forms of noncontractual social inter-

action. What is called *equilibrium selection* operates almost entirely through processes absent in the Walrasian model.

WHILE MOST of what follows is the result of recent research, virtually all of the models and ideas presented there were anticipated by writers half a century ago or more, sometimes much more. The importance of adaptive agents (with realistic cognitive capacities and predispositions) whose behaviors were based on local information was central to the work of Friedrich Hayek (1945) and Herbert Simon (1955). Simon's pioneering work on the incomplete nature of the employment contract (Simon 1951) and the role of authority in the functioning of firms formalizes the earlier work of Ronald Coase (1937) and long before Coase, Marx (1967). The basic concepts of game theory, bargaining, and other nonmarket social interactions were introduced in the early writings of John Nash (1950a), John von Neumann and Oskar Morgenstern (1944), Thomas Schelling (1960), and Duncan Luce and Howard Raiffa (1957). Nash even suggested the basic ideas of evolutionary game theory in his doctoral dissertation (Nash 1950b). Nash's famous solution to the bargaining problem was first proposed much earlier by F. Zeuthen (1930), in a work introduced glowingly by Joseph Schumpeter. Endogenous preferences were central to the work of James Duesenberry (1949) and Harvey Leibenstein (1950), both drawing on the much earlier work of Thorsten Veblen (1934 [1899]) and developing themes initially raised by Smith (1937) and Marx. The famous paradox of Maurice Allais (1953) pointed to problems with the expected utility hypothesis that have only recently attracted serious attention. The way that positive feedbacks support multiple equilibria was the key idea in Gunnar Myrdal's 1955 Cairo lectures (mentioned above). The application of biological reasoning to economics, now prominent in evolutionary game theory, was introduced a half-century ago by Armen Alchian (1950) and Gary Becker (1962).

The fact that most of the key ideas presented in the pages that follow were anticipated during the 1950s or before but ignored in subsequent decades poses an intriguing question. Why did the Walrasian paradigm become virtually synonymous with economics in the third quarter of the twentieth century only to be displaced at the century's end by a set of ideas most of which had been articulated by well-placed academics just prior to the rise to prominence of the Walrasian paradigm? Herbert Gintis and I (Bowles and Gintis 2000) have attempted to answer the question, but to address it here would be a diversion.

RELAXING the canonical Walrasian assumptions to take account of non-contractual social interactions, adaptive other-regarding behaviors, and

generalized increasing returns will require a method more empirically grounded and less deductive than the usual Walrasian approach. Making little reference to the specifics of time, or place, or indeed any empirical facts, the Walrasian paradigm deduced a few rather strong predictions concerning the outcomes likely to be observed in the economy. The expansion of the domain of economics to include the family, the organization of production, and political activity such as the voluntary provision of public goods, lobbying, and voting, produced valuable insights unattainable using the conventional methods of sociology and political science. But research in these areas, as well as the return to prominence of the classical economists' concern with long-term economic growth and distribution, have cast doubt on the generality of the standard assumptions. Responding to the malaise now felt among economists, the American Economic Association's *Journal of Economic Perspectives* devotes a regular column to "anomalies," which they define as follows:

Economics can be distinguished from other social sciences by the belief that most (all?) behavior can be explained by assuming that rational agents with stable well defined preferences interact in markets that (eventually) clear. An empirical result qualifies as an anomaly if it is difficult to "rationalize" or if implausible assumptions are necessary to explain it within the paradigm. (Thaler 2001)

Readers responded avidly to the invitation to write in with their favorite examples.

In place of deduction from a few (once) uncontroversial behavioral and institutional axioms, economics has increasingly (if unknowingly for the most part) moved toward an approach that combines the mathematical advances of the last century with three of the methods of the classical economists. From Adam Smith to John Stuart Mill and Karl Marx (and excepting David Ricardo), the classical economists were *nondisciplinary* (the disciplines had not been invented), *concerned about the empirical details* of the social problems of their day, and *modest about the degree of generality to which their theories aspired*.

First, the study of the economy must draw upon the insights of all of the behavioral sciences, including ecology and biology. The Walrasian assumptions provided a rationale for a rigid division of labor among the disciplines. Its defining assumptions allowed Walrasian economists to disavow an interest in other-regarding behaviors, norms, the exercise of power, or history as some other discipline's concern and in any case not pertinent to the workings of the (Walrasian) economy. While the traffic across the disciplinary boundaries has in the last half-century consisted primarily in the export of economic methods to the other behavioral

sciences, there is much to be imported if the role of power, norms, emotions, and adaptive behaviors in the economy are to be understood. Core economic phenomena such as the workings of competition, incentives, and contracts cannot be understood without the insights of the other behavioral sciences.

Second, relaxation of the Walrasian assumptions confronts us with an embarrassment of riches. In the absence of some empirical restrictions or theoretical refinements, the price of generality will be vacuousness. This was the conclusion of Hugo Sonnenschein (1973b:405) concerning Walrasian theory of market demand: “The moral . . . is simply this: if you put very little in, you get very little out.” But the same applies to any post-Walrasian paradigm. Few empirical predictions will be forthcoming if individuals may be self-interested or not depending on the person and the situation, if some interactions are governed by contracts, others by handshakes, and others by brute force, and if there exist multiple stable equilibria.

The need for empirical grounding of assumptions is nowhere clearer than in the analysis of individual behavior, where the process of enriching the conventional assumptions about cognition and preferences can easily descend into ad hoc explanation unless disciplined by reference to facts about what real people do. It is not enough to know that self-interest is not the only motive; we need to know which other motives are important under what conditions. These restrictions are most likely to come from one of the sources that undermined the Walrasian paradigm, namely, the great advances in empirical social science stemming from new techniques in econometrics, the improvement in computational capabilities and data availability, experimental techniques, and continuing progress in quantitative history.

Theory, too, can provide useful restrictions on the set of plausible assumptions and outcomes. The modeling of genetic and cultural evolution, for example, can help restrict the range of plausible behavioral assumptions by distinguishing between those emotions, cognitive capacities, and other influences on behaviors whose emergence and diffusion can plausibly be accounted for over the relevant periods of human history, and those that cannot. Similarly, while generalized increasing returns may support a large number of equilibria, some of these equilibria are extremely inaccessible under any plausible dynamic process. By contrast, other equilibria may be both accessible and robust. In this case, specification of an explicit dynamic process—for example, an account of how individuals adapt their behaviors in light of their recent experiences and the experiences of those whom they observe—may allow the elimination of what may be termed *evolutionarily irrelevant equilibria*. Making explicit the dynamics governing a system gives us an account of its out-of-equilibrium behavior and

thus not only helps in the process of equilibrium selection but also in studying the response to shocks and other problems for which the standard comparative static method is ill-suited.

Third, the quest for ever more general theories will continue to engage students of the economy, and there is still much to be learned by studying such topics as markets in general. But for the foreseeable future it seems likely that insights will come from models that take account of the specific institutional and other aspects of particular types of economic interaction. For the classical economists it was evident that labor markets differ in fundamental ways from credit markets, which in turn differ from shirt markets or foreign exchange markets, and so on. Models may be more specific with respect to time and place, as a way of capturing the importance of time-varying institutions or different cultures. If the exciting novelties of the Walrasian era were highly abstract theorems of surprising generality, the excitement in the coming years may come from compelling answers to such questions as are raised by the empirical puzzles concerning the wealth of nations and people, with which I began.

It would be salutary for economists to focus more on answering such questions and less on demonstrating the use of our increasingly sophisticated tools. But it seems that a more problem-driven and less tool-driven approach will require yet more sophisticated tools. The mathematical demands of the theoretical framework I am proposing will be greater, not less, than that of the Walrasian paradigm. The reason is that models that represent noncontractual social exchanges among individuals who are both heterogeneous and versatile in their behaviors and who interact in the presence of generalized increasing returns do not allow the standard simplifications such as price-taking behavior and convex production sets that made Walrasian models tractable. As has long been recognized in physics and biology, many important problems do not yield simple closed form solutions, or indeed any solutions at all that are susceptible to simple interpretation. In these cases—some of which you will encounter in chapters 11 through 13—computer simulations of the relevant social interactions will prove insightful as a complement (not a substitute) for more traditional analytical methods. Simulations have been extensively used in developing the ideas on which this book draws. Simulations do not yield theorems or propositions that are generally true; rather, like experiments, they yield a wealth of data that may point to unambiguous conclusions but often do not.

THOUGH MOTIVATED by an interest in the impact of economic institutions on human well-being, I have adopted an evolutionary rather than a social engineering approach. Like the idea of “selfish genes” seeking

to maximize their replication or an auctioneer presiding over a general equilibrium exchange process, the omniscient and omnipotent social engineer seeking to maximize social welfare is a fiction whose usefulness depends on keeping in mind its fictive character. Social outcomes—even those involving states and other powerful bodies—are the combined result of actions taken by large numbers of people acting singly. Such devices as fictive auctioneers, social engineers, or anthropomorphic genes cannot substitute for an understanding of how real individuals behave and the ways that distinct institutions generate population-level dynamics that aggregate these behaviors to produce social outcomes. The evolutionary character of the analysis will become evident in the way that individual behaviors are modeled, the kinds of population-level dynamics studied, the ways that behaviors and institutions coevolve, and the absence of any grand blueprints for human betterment. The evolutionary approach is modest about what interventions can accomplish, but it does not restrict the economist to purely contemplative pursuits. I take up questions of good governance and policy in the concluding chapter.

The first part of the book introduces a variety of models applied to what I have just called the generic social interaction, namely, noncontractual social interactions among adaptive agents in the presence of generalized increasing returns. I begin with two chapters on institutions and the evolution of structures of social interactions before turning to preferences and beliefs. The unconventional ordering of these topics—most microeconomics texts start with preferences—reflects the importance of institutions as influences on the norms, tastes, and understandings that individuals bring to the situations in which they act. I then investigate allocational inefficiencies that occur in noncontractual interactions, and the problem of dividing the gains to cooperation that arises when these inefficiencies can be overcome. The middle part of the book concerns the institutions of capitalism, and especially markets, lending institutions, and firms. I give particular attention to the way that the incomplete nature of most contracts gives rise both to a well-defined political structure of the economy and to an important role for social preferences. The last part concerns the process of cultural and institutional change; I emphasize the role of technical change, collective action, and intergroup conflict as constituent parts of the process by which the rules governing social interactions and individual behaviors coevolve. Here I address the evolution of familiar institutions such as private property and customary rules of division, as well as the puzzling evolutionary success of other-regarding individual behaviors. The concluding chapter compares three structures governing economic interactions—markets, states, and communities—and explores ways that they

might serve as complementary approaches to handling problems of allocation and distribution.

In 1848, John Stuart Mill (1965) published *Principles of Political Economy*, the first great textbook in microeconomics; it was the staple of instruction in the English-speaking world until displaced by Marshall's *Principles* a half-century later. Mill's readers may have been reassured to read, "Happily, there is nothing in the laws of Value which remains for the present writer or any future writer to clear up; the theory of the subject is complete" (p. 420). When I studied economics in the 1960s during the heyday of the Walrasian paradigm, a similar complacency reigned. This book conveys no such reassurance. Our understanding of microeconomics is fundamentally in flux. Little is settled. Nothing is complete.