

# 1

## INTRODUCTION

Currents of time swirling and eddying all about us, on the battlefields and in the military headquarters, in the factories and on the streets, in boardrooms and cabinet chambers, murkily at first, yet tending ever towards a moment of transfiguration in which pattern is born from chaos.

(Coetzee [1983] 1985, 158)

### I. THE APPROACH OF THE BOOK

#### 1. Order and disorder

The economic history of the developed capitalist world appears to be one of almost constant progress: inexorable growth, rising standards of living, rising productivity, and ever-improving health, well-being, and welfare. Seen from afar, it is the system's order, its internal coherence, which stands out.

Yet the closer one looks, the more haphazard it all seems. Individuals wander along entangled paths, propelled by obscure motivations toward some dimly imagined ends, crisscrossing and colliding as they act out their economic roles as buyers and sellers, bosses and workers, producers and speculators, employed and unemployed. Information, misinformation, and disinformation hold equal sway. Ignorance is as purposeful as knowledge. Private and public spheres are entwined throughout, as are wealth and poverty, development and underdevelopment, conquest and cooperation. And everywhere there appears a characteristic unevenness: across localities, regions, and nations; and across time, in the form of booms, busts, and breakdowns. Seen up close, it is the system's disorder that is most striking.

How does one address these two, equally real, aspects?

## 2. Neoclassical response to the real duality

Neoclassical economics, the present-day orthodoxy, provides one answer. It seizes on the first aspect, and purges, or least exiles to theoretical backwaters, the second. The perceived order of the system is recast as the supreme optimality of the market, of the ever-perfect invisible hand. This optimality is in turn projected back onto microscopic units, so-called representative agents, from whose superlatively rational choices it is said to derive. And so we arrive at a particular vision. In its perfectly ordered form, the system equalizes all prices for comparable goods, all wage rates for comparable labors, and all profit rates for comparable degree of risk. Moreover, it fully utilizes all available resources, including available plant, equipment, and labor. All of this without error, instability, or crisis. Only then, after this has been firmly established as the ruling conception, is potential disorder allowed into the story, *sotto voce*, in grudging concession to the obstinate indifference of the regrettably imperfect real.

## 3. Keynesian and post-Keynesian response to the real duality

Heterodox economics, most notably post-Keynesian economics, generally takes the opposite tack. It emphasizes the inefficiencies, inequalities, and imbalances generated by the system. In the place of perfect competition, we get imperfect competition; in the place of automatic full employment, we get persistent unemployment. Market outcomes now appear as conditional, on history, culture, politics, chance, and most of all, on power: oligopoly power, class power, and, of course, state power. From this point of view, what others may perceive as ordered economic patterns are really contingent paths, arising from historically specific constellations of forces. Desired social outcomes are not automatic, and automatic outcomes are not always desired. Unemployment is more probable than full employment, while inflation and crises are always possible. Hence, there exists an ever-present need for social and economic intervention to fill in the spaces between the actual and the desired. What neoclassical economics promises through the workings of the invisible hand of the market, Keynesian and post-Keynesian economics promises through the visible hand of the state.

The irony is that both sides end up viewing reality through an “imperfecionist” lens. Neoclassical economics begins from a perfectionist base and introduces imperfections as appropriate modifications to the underlying theory. Heterodox economics generally accepts the perfectionist vision as adequate to some earlier stage of capitalism but argues that imperfections rule the modern world. In either case, such approaches actually serve to protect and preserve the basic theoretical foundation, which remains the necessary point of departure and primary reference for an ever-accreting list of real-world deviations. After all, how can the basic theory ever be wrong if there is a particular ether for every troublesome result? This book follows a different path.

## 4. Different purpose of this book

To begin with, the very purpose of this book is different. Neoclassical economics investigates the workings of a deliberately idealized version of capitalism, from whose vantage point it seeks to characterize the world. Heterodox economics seizes on the

distance between this vision of perfection and the real world. Both sides attempt to bridge the resulting gaps by ladling various “imperfections” into the original mix. Both therefore remain forever off-balance, one foot in the ideal and the other in the real. The goal of this book is to develop a theoretical structure that is appropriate from the very start to the actual operation of existing developed capitalist countries. Its object of investigation is neither the perfect nor the imperfect but rather the real. For this reason, the theoretical arguments developed here, along with their main alternatives, are constantly confronted with empirical evidence.

Second, although the book attempts to demonstrate that the capitalist economic system generates powerful ordered patterns that transcend historical and regional particularities, the forces that shape these patterns are neither steely rails nor mere constellations of circumstance. They are, rather, moving limits whose gradients define what is easy and what is difficult at any moment of time. In this way they channel the temporal paths of key economic variables. Indeed, these shaping forces are themselves the results of certain immanent imperatives, such as the “gain-seeking behaviors” that define this particular social form in all of its historical expressions. It is not a matter of contrasting ahistorical laws to historically contingent outcomes. Agency and law coexist within a multidimensional structure of influences. But this structure is itself deeply hierarchical, with some forces (such as the profit motive) being far more powerful than others. *The stage on which history plays out is itself moving, driven by deeper currents.*

Third, the resulting systemic order is generated in-and-through continual disorder, the latter being its immanent mechanism. To attempt to theoretically separate order from disorder, or even to merely emphasize one over the other, is to lose sight of their intrinsic unity, and hence of the very factors that endow the system with its deep patterns. Yet order is not synonymous with optimality, nor is disorder synonymous with an absence of order. Order-in-and-through-disorder is of a piece, an insensitive force that tramples both expectations and preferences. This is precisely the source of the system’s vigor.

Fourth, if one is to demonstrate how order and disorder are intimately related in given circumstances, it is necessary to identify particular mechanisms. And here, the central goal of this book is to demonstrate that a great variety of phenomena can be explained by a very small set of operative principles that make actual outcomes gravitate around their ever-moving centers of gravity. This is the system’s mode of *turbulent regulation*, whose characteristic expression takes the form of *pattern recurrence*. The theoretical and empirical applications of these two notions are woven into the structure of this book.

Turbulent regulation and pattern recurrence apply to the system’s various gravitational tendencies. Of these, the first set consists of those that channel commodity prices, profit rates, wage rates, interest rates, equity prices, and exchange rates. These processes have two aspects. Equalizing tendencies driven by the restless search for monetary advantage, whose unintended outcome is to narrow the very differences that motivate them. And shaping tendencies which direct the path around which the equalizing tendencies operate. For example, equalization processes make individual wage and profit rates gravitate around the corresponding averages. Competition among workers and capitals plays a key role here. At the same time, the average wage rate itself depends on productivity, profitability, and the balance of power between employers

and employees, while the average profit rate depends on wages, productivity, and capital intensity. The averages emerge from individual (micro) economic interactions in which competition plays a central part. Both of these processes therefore fall within the domain of *real competition*, in which the profit motive plays the central role. As we shall see, the notion of real competition developed in this book is very different from that of perfect competition and its dual, imperfect competition. Real competition does not fit on some sliding scale between these two theoretical markers.

The second set of gravitational tendencies arises from the system's *turbulent macro-dynamics* with its characteristic expansionary processes, waves of growth and slowdown, persistent unemployment, and periodic bouts of depression including the global crisis that began, very much on schedule, in 2007. Once again, it is the profit motive that is the dominant factor in the regulation of investment, economic growth, employment, business cycles, and even inflation.

The centrality of the profit motive has several implications. First of all, the theory of profit, and hence of the theory of wages, takes on special significance. Second, it becomes important to delineate the precise role of profitability in the theory of real competition, because it affects all aspects of the behavior of the firm. This influence extends to the theory of competitive price setting and the theory of (endogenous) technical change. Third, the notion that (expected) profitability regulates both investment and growth implies a particular mode of interaction between aggregate demand and supply. We shall see that the resulting dynamic is neither neoclassical, nor Keynesian, Kaleckian, or Harrodian but rather fundamentally classical: profit regulates both supply and demand. Profitability also plays a critical role in the theory of persistent unemployment, through the channel of endogenous technical change and a correspondingly endogenous "natural" rate of growth. Finally, we will see that newly created purchasing power can pump up output and employment, just as Keynes argued, but that this can lead to a reduction in the rate of growth. Then while short-run output will be higher than it would otherwise have been, long-run output will be lower than it would otherwise have been.

Empirical evidence plays so large a part in this text that it is important to note that data is never just a collection of pre-existing facts. Theory always intervenes, not merely in the interpretation of events, but in their very representation (and occasionally in their suppression, as we know only too well). For instance, no analysis of unemployment can proceed very far without recognizing that in all official accounts in the advanced countries, a person is counted as "employed" if she/he "did any work at all for pay or profit" during the week.<sup>1</sup> It was only in the last three decades that US agencies have begun to publish measures of partially employed and discouraged workers, which, of course, reveal a much bleaker picture of the economy. We will see that a similar problem exists in official measures of the stock of capital, which have changed considerably as neoclassical constructs have supplanted classical and Keynesian ones in this field. Not just the levels and trends, but the very notion of capital itself, has been transformed. This is of some importance because the capital stock plays a critical role in the calculation of the rate of profit. Unlike *Candide*, data is never innocent.

<sup>1</sup> "How the Government Measures Unemployment" (Washington, DC: US Department of Labor, Bureau of Labor Statistics, 2001).

This book draws on a variety of sources. The principle of turbulent regulation has its roots in the method of Smith, Ricardo, and particularly of Marx, for whom “laws of motion” are regulative principles that exert themselves in-and-through various counter-tendencies. The theory of real competition has similar roots within the economics canon, but elements of it can also be found in the business literature. Most of the time, the patterns are directly visible, but sometimes formalization requires the tools of modern nonlinear dynamics and empirical testing requires the tools of modern econometrics.

The emphasis on growth as an immanent process also has roots in the classical and Marxian traditions, as well as in the works of Harrod and Robinson and others. As previously noted, the responsiveness of the business savings rate to investment needs opens up the way for the classical synthesis of Harrod and Robinson, in which growth is driven by profitability and yet capacity utilization gravitates around some normal rate. At the macroeconomic level, demand cannot be independent of supply, since the decision to produce leads to the purchases of materials and machines, and payment of wages to workers and rents, interest and dividends to landlords, creditors, and owners. Hence, supply is neither the imperial force of neoclassical economics, nor the ghostly presence of Keynesian and Kaleckian economics. Supply and demand are co-equals here, strutting on the stage in alternating splendor. But, as always, profit is pulling the strings.

The notion of persistent unemployment can be traced back to Marx’s theory of the reserve army of labor, to Harrod’s puzzle about the difference between warranted and natural rates of growth, and to Goodwin’s brilliant mathematical synthesis of these as a predator-prey cycle. In Harrod, Kaldor-Pasinetti, and Goodwin, the profit rate must adapt to make the warranted rate of growth equal to the natural rate. I argue that the natural rate, which is the sum of the productivity and labor force growth, is itself responsive to profitability: the rate of technical change depends on relative cost of labor, and labor force growth responds through changes in participation rates and the importation of labor, to profit incentives. Then profit-driven growth is capable of generating a persistent rate of unemployment, as in Marx and Goodwin. This can be shown to have major implications for the effects and limits of fiscal policy.

In all of these arguments, the goal is to weave a theoretical narrative that is internally consistent with regards to its logic, and externally consistent with regards to the empirical evidence. It should be noted that while the book’s focus is on the developed capitalist world, this is not due to a lack of interest in the developing world. On the contrary, it is strongly motivated by the belief that an analysis of capitalism in its most developed form is essential to an adequate understanding of the relations between the developed, developing, and underdeveloped arenas of the world. It is to this aim that my project has been dedicated.

## II. OUTLINE OF THE BOOK

This book is divided into three parts: Foundations of the Analysis, the theory of Real Competition, and the theory of Turbulent Macro-Dynamics. Excluding this introductory chapter and a brief concluding one, each part comprises five chapters. All theoretical arguments are contrasted to the corresponding neoclassical and Keynesian/post-Keynesian views and confronted with the empirical evidence. In the

following summary of the various chapters I leave out most citations and references since they appear in the relevant texts.

### 1. Part I: Foundations of the analysis (chapters 1–6)

Coming on the heels of this introductory chapter, chapter 2 sets the stage by presenting empirical evidence on characteristic long-run economic patterns in advanced capitalist countries. These include persistent growth in output, productivity, profits, and employment, all taking place in-and-through recurrent cycles and periodic Great Depressions; the socially influenced relation of real wages to productivity; the salutary impact of policy and institutions on unemployment; the surprising recurrence of golden long waves, even into the present day; the growth implications of the long-term path of profitability; the turbulent equalization of rates of return across industrial sectors; and the structural determination of industrial relative prices. Notions such as recurrence and turbulent regulation arise quite naturally from an empirical scrutiny of this sort. The chapter ends with a long view of the rise in global inequality that has led to the present state of affairs in the world, in which development exists alongside underdevelopment, growth alongside decline, extreme wealth alongside abysmal poverty. This purpose of this chapter is to make it clear that the object of investigation is capitalism itself.

Chapter 3 takes up the methodological questions raised by the very existence of persistent long-term patterns. It begins with the question of method. The conventional notion of equilibrium as a state of quietude is replaced by the notion of turbulent regulation in which balance is only achieved by recurrent over- and undershooting. This raises the question of the temporality of the processes involved, that is, the length of the “runs” over which various balances are supposed to be achieved. It also becomes necessary to address the inherent “lumpiness” of objects and social responses, because the thresholds imply *intrinsic nonlinearities* in the processes themselves.

The very persistence of long-term patterns raises yet another methodological issue: How is it possible for capitalist society to generate recurrent aggregate patterns across the ages, given that it is composed of mutable individuals embedded in evolving social structures and subject to ever changing fashions? One answer, currently favored by neoclassical economics, is to portray recurrent social outcomes as the hyper-rational choices of some unchanging “representative agent.” But the very notion of a representative agent suffers from intractable difficulties. First of all, in order to derive stable aggregate patterns across changing historical conditions, it is necessary to posit unchanging representative agents. Second, even under given social circumstances, the aggregate behavior of a group will not correspond to the underlying individual behaviors unless all agents within a group are identical. It is the general existence of nonlinearities arising from interactions among individuals that accounts for this result. Neoclassical economics simply ignores this problem and continues to plow ahead. Third, the assumption of hyper-rationality is not useful because it systematically misrepresents the underlying motivations and is not necessary because we can derive observed patterns without it.

The key is to recognize that aggregate outcomes have “emergent” properties due to interactions among individual elements: an organic whole is more than the sum of its parts. We still need to explain the persistent character of these emergent

properties in the face of changing historical conditions. And there, the secret lies in the *diversity* of individual agent behaviors. Lawful patterns can emerge from the interaction of heterogeneous units (individuals or firms) operating under shifting strategies and conflicting expectations because aggregate outcomes are “robustly indifferent” to microeconomic details. Diversity produces statistical distributions of outcomes whose averages and other features are shaped by social and cultural structures. This approach can be used to demonstrate that we can derive the major empirical laws of aggregate consumer behavior (downward sloping demand curves, characteristic income elasticities of necessities and luxuries, the nonlinearity of Engel’s curves, and the near-linearity of aggregate consumption functions) without reference to any particular model of consumer behavior. Agents do make choices, and choices are important. But the preceding general patterns do not depend on those details. It follows that one cannot simply compare competing micro models at the aggregate level. We must continue on to the micro level itself. Then it becomes clear that there is no reason to shy away from the complexity, whimsy, and occasional madness of actual human behavior. Diversity should be embraced, not suppressed.

Chapter 4 concerns itself with the structure of social production. On the surface, capitalism appears to be a system of generalized exchanges. Indeed, neoclassical economics presents the exchange of equivalents as the central organizing principle of capitalist society, only introducing production as a means of indirect exchange between the present and the future. The classical view is very different. Since production takes time, it precedes the exchange of products. And it is in production that we confront the constant struggles about wages and the length and intensity of the work. The first part of this chapter contrasts the classical emphasis on the importance of the time of production and on the active role of labor in production with the timeless and passive inputs-into-outputs methodology of most other economic traditions. The distinction between circulating investment and fixed investment is shown to have important implications for economic dynamics. Classical and conventional production accounts are shown to differ on measures of total output and value added, but surprisingly not on gross operating surplus. A formal mapping between the two schemas is developed in appendix 4.1. The second part of the chapter shows how the utilization of materials, fixed capital, and labor is linked to the length and intensity of the working day. These connections are used to deconstruct various standard representations of the production process ranging from fixed-coefficient to neoclassical production functions. Different potential combinations of shift work will switch back and forth along the production possibilities frontier at different levels of utilization. This new type of “re-switching” destroys any possibility of constructing a neoclassical *microeconomic* production function. On the other side, the social determination of shift lengths and work intensities imply that we cannot take either fixed or variable coefficient models of production as representing solely “technological” conditions. The third part shows that properly derived cost curves are very different from those posited in standard microeconomic texts. The notion of U-shaped cost curves gets hit particularly hard, because the normal cost changes from one shift to the next give rise to spikes in average variable costs and to corresponding sharp jumps in marginal costs. A direct implication of this spikiness is that a given price can intersect the marginal cost curve multiple times, so that the  $p = mc$  rule is of no use in determining the profit-maximizing point

of production. The fixed-coefficient approach can accommodate the results somewhat better, but only by treating each potential shift combination, operated at socially determined lengths and intensities of the working day, as a separate “technology.” The chapter ends with a survey of the empirical evidence on the length and intensity of the working day, on labor productivity, and on shapes of actual cost curves. We find that the classical treatment of production is quite consistent with the empirical evidence, and that the theoretical cost curves derived on this basis are similar to empirically observed curves and consistent with business experience. On the other hand, the ubiquitous neoclassical U-shaped cost curve is neither empirically grounded nor of much practical use.

Chapter 5 takes up the question of money. Production activities are undertaken by individual businesses concerned with their own profit, with no immediate regard for their fit with social needs. Each firm anticipates profit from sales of the planned product and anticipates buying other products for future inputs or personal consumption. The inevitable discrepancies between conflicting individual expectations and plans are resolved in the market. Neoclassical economics skirts the din of real markets by pretending that individual production plans mesh perfectly with social needs. This pretense is called general equilibrium. In point of fact, the turbulent order arising from real markets is achieved only in-and-through disorder, and money is its general agent.

Exchange is not to be confused with gift-giving, even when the latter is reciprocal. A proper gift asks nothing in return, whereas a proper exchange asks nothing less. Potlatch is an example of a custom in which the social ranking of the participants was determined by how much they could give away. In reciprocal gift-giving, each side tries to give back something desirable to the other. In exchange, each side tries to get back something more desirable than it gives. In the same vein, a payment obligation should not be confused with a proper debt. For instance, tributes and taxes are one-sided payment obligations often enforced by a threat. These are generally one-sided, which is why we use terms like tribute and tax. A debt is a repayment obligation, so it involves a reflux in interest and amortization payments. This is different from time-separated exchange: tools can be exchanged on the spot for grain, or tools can be received now and the grain delivered later. Such differences are shown to play an important role in the theory of money and credit.

Barter is the earliest form of true exchange. It will establish multiple exchange ratios between any given commodity and all the others in its orbit. Money arises naturally as the reach of exchange is extended, in response to the intrinsic need to convert the many exchange ratios of a given commodity like grain with meat, salt, leather, tools, and so on into a single ratio between it and some given socially selected commodity like salt. Then salt is the local money commodity and all the commodities in its sphere acquire a salt price. *Price* is intimately connected to money: it is the *monetary* expression of a commodity’s quantitative worth.

The distinction between a mere commodity and a money commodity arises again and again in human history, with the latter taking various form such as salt, cattle, pigs, grain, shells, cocoa beans, beads, turmeric, red ochre, axe blades, arrows, spears, millstones, beetle legs, beeswax, metals, and tokens, with new forms constantly being invented. Monies start off as localized entities, and like royalty, most are deposed over history’s long march. Section II traces the evolution of money from its origins in



exchanges to private and state-issued coins, private and state-issued convertible and inconvertible tokens, state fiat money, and bank money. It ends with a statement of the three essential functions of money (medium of pricing, medium of circulation, and medium of safety) and a look at some striking long-term empirical patterns. Section III goes from classical theories of money and the price level to Marx's discussion of these same issues. Marx restricts himself to the case in which tokens directly or indirectly represent a money commodity (he promises to analyze pure fiat money and bank credit at a later date but does not live to do so). From this point of view, his treatment of commodity-based money applies up to 1939/40, which marks the end of gold standard. A central factor is his determination of the national price level as the product of two terms: the competitively determined relative price of commodities in terms of the historically chosen money commodity which in the West was gold; and the price of the money commodity determined by monetary and macroeconomic factors. Some striking empirical patterns come into view in the United Kingdom and United States when price levels are examined from this perspective. One of the benefits of this approach is the identification of a simple long wave indicator that continues to be valid to the present day (chapters 16 and 17).

Section IV links the classical treatment of fiat money in a commodity money (say gold) standard to the modern (Sraffian) treatments of relative prices of production, before moving to the key question: How does one address the case in which fiat money is no longer linked to any money commodity? It is argued that under modern fiat money the national price level is directly determined by monetary and macroeconomic factors, but in a manner different from Monetarist, Keynesian and post-Keynesian theories. Hence, this aspect is postponed to the analysis in chapters 12–14 of Part III of the book in which classical approaches to profitability, effective demand, growth, and inflation are developed and applied to macroeconomics. Modern theories of inflation and a classical alternative are then treated in chapter 15, along with a critical analysis of Chartalist and neo-Chartalist claims about the historical role and modern powers of the state.

Chapter 6 opens with an extended analysis of profit and capital. Two issues are paramount: the definition of capital and the determination of aggregate profit. Keynes cites Marx's notion of the circuit of capital  $M - C - M'$  as providing a particularly useful method for identifying capital. Over the life of its circuit, capital starts out as money  $M$ , is transformed into commodities ( $C$ ) comprising labor power, raw materials, and plant and equipment, and then hopefully recouped as more money ( $M'$ ). By contrast, the act of working for a living in order to earn an income falls within the circuit  $C - M - C$ . The two circuits interact, since wages received by employees are part of the capital expenditures of firms, while the consumer goods and financial assets purchased by employees are part of the profit-motivated sales of firms. So it is not a thing's qualities but rather the process within which it operates that turns it into capital. Capital is also not defined by its durability: circulating capital like a clay mold may last only part of a year, while fixed capital such as a machine may last decades. On the other side, durable goods such as household automobiles and dwellings are parts of personal wealth, not capital. Indeed, a car may be personal wealth for an individual owner while the same model may be capital for a car dealer waiting for it to be driven off the lot (at the right price). Neoclassical economics always conflates capital and durable wealth because it simply defines "capital" as wealth that lasts more than one year. Modern-day national

accounts often embody the neoclassical approach: for instance, private homeowners are treated as businesses renting their homes to themselves (appendix 6.7).

Section II demonstrates that there are two sources of aggregate profit, as originally argued by Sir James Steuart: the first arises from a transfer of wealth, the second from the production of new wealth in the form of a surplus product. This is the basis for the distinction between “buying cheap in order to sell dear” upon which merchant capital has been historically based, and the production of a surplus product on which industrial capital is based. Marx comments approvingly upon Steuart’s distinction between profit based on “unequal exchange” and profit based on the production of a surplus. Because he is most concerned with the latter, in Volume 1 of *Capital* Marx concentrates on the demonstration that positive industrial profit exists even when there is “exchange of equivalents.” He is careful to say that the other form which he calls “profit on alienation” plays an important role in various arenas, and says he will return to the issue at some later point (presumably in Volume 3 of *Capital*, which he does not live to complete). I demonstrate that the secret to Steuart’s first form of profit is a transfer into the circuit of capital and show that this plays a critical role in various “transformation problems” and in the unraveling of the mysteries of financial capital.

Section III concentrates on industrial profit. Harking back to the earlier discussion in chapter 4 on the relationship between the length and intensity of the working day and the total product, it is demonstrated that a surplus product only arises when the length of the working day exceeds the working time to reproduce the standard of living of the employed workers, that is, only when surplus labor is performed. Since both the evolution of technology and its operation are socially determined, this tells us that the existence of surplus labor is a social outcome, not a merely “technical” one. Several further results are derived. First, aggregate profit is zero when there is a zero surplus product, regardless of the prices adopted by individual industries. Even doubling all selling prices will not work, because this also doubles the reproduction costs of the same material inputs and labor power (assuming that the real wage is maintained): then what firms collectively gain as sellers they simultaneously lose as buyers, so aggregate real economic profits remain zero. Conversely, positive aggregate profit only exists when there is positive surplus labor time and a corresponding positive surplus product. Once again, doubling the absolute price level will not raise real aggregate profit because it also doubles all costs.

However, in the case of given positive surplus product a change in *relative* prices can change aggregate profit. Profit is still a reflection of the surplus labor, but now the mirror of circulation appears to be curved. The partial dependence of money profit on relative prices is completely general. It applies to neoclassical, Sraffian, and Marxian theories of price: in other words, *there is a “transformation problem” in all schools of thought*. In the Marxian case, aggregate profits vary when one moves from prices proportional to labor value to price of production. But the same can be said if one compares prices of production, which are after all purely theoretical constructs, to market prices or monopoly prices—a point that Sraffians have largely failed to note.

Section IV builds on Steuart’s insight that transfers of wealth and value can also affect aggregate profit. It is demonstrated that changes in the relative prices of commodities generally have different impacts on the circuits of capital and revenue, and can give rise to transfers between the two circuits even though the total money value of the product is unchanged. The sum of the transfers is always zero, but since one

circuit may gain what the other loses, or vice versa, aggregate profit can change. This is a completely general solution to “transformation” problems. It can be used to further explain why the particular set of output proportions associated with maximum balanced growth does not exhibit this phenomenon—that is, why its aggregate profit is invariant to relative prices in this case. Section V uses the general framework to address financial profit arising from realized capital gains and other transfers (let us never forget Ponzi or Madoff). Section VI shows that Smithian, Sraffian, Keynesian, and post-Keynesian theories of aggregate profit actually rely on the existence of a positive surplus product by implicitly or explicitly assuming that the real wage is less than the productivity of labor. Neoclassical theory is different, because it has a notion of profit due to transfer (emanating from a model of pure exchange) and a notion of profit on production (emanating from an aggregate production function). This is Steuart *re-dux*, but now the emphasis is on the justification of profit as a reward to abstinence and entrepreneurship. It is important to separate the explanation of profit from its justification. Smith and Ricardo explain profit and rent as a deduction from the net produce of labor, but do not dispute that capitalists or landlords have rights to these flows. Marx is equally clear that capitalists and landlords (like all ruling classes) have the socially constructed “right” to extract surplus labor—just as at some point workers gain the right to resist. All three authors are critical of capitalists whereas neoclassical and Austrian authors tend to celebrate them. Section VII addresses the literature on the effect of relative prices on aggregate profit, including Marx’s famous “transformation” discussion. The subsequent literature from Bortkiewicz to Samuelson and Sraffa is assessed for its strengths and weaknesses, as is the so-called “New Interpretation” of Foley and Duménil. The section ends by noting that in any case the empirical impact of relative prices on aggregate profit and on the profit rate is very small. Section VIII takes up the theory and empirical measurement of profit, capital, and the rate of profit. Most of the details are developed in the appendices. Appendix 6.1 provides a formal treatment of the relations between surplus labor and aggregate profit. Appendix 6.2 shows that if the rate of profit is measured as the money value of the total product and the current cost of materials, depreciation, and labor (as argued in chapter 6, section III.3), then it is also a real rate of profit: deflating the numerator and denominator by any common price index will not affect their ratio. On the other hand, deflating them by separate price indexes will not do because then the rate of profit will no longer be a pure number. Appendix 6.3 points out that the business notion of capital as gross stock is different from the neoclassical notion of capital as net stock, and appendix 6.4 shows that the treatment of fixed capital as a joint product then has two distinct forms: the one adopted by Marx which corresponds to gross stock and one adopted by Sraffa which corresponds to net stock. These two treatments turn out to have differing theoretical and empirical implications. Empirical measures of the capital stock present a new set of issues because of problems arising from the perpetual inventory method (PIM) through which investment flows are cumulated into capital stocks. Appendix 6.5 analyzes the meaning and impact of “quality adjustments” on price and quantity indexes and the apparently intractable aggregation problems arising from use of chain-weighted indexes which seem to make it impossible to generate capital stock measures based on less problematic assumptions. Section V of appendix 6.5 derives a new set of generalized PIM rules that apply even to chain-weighted aggregates, so that it becomes possible to construct new measures of the capital stock

and hence of the rate of profit. Capacity utilization poses yet another challenge, since we know that actual capacity utilization will generally fluctuate in response to various factors. Accordingly appendix 6.6 analyzes existing measures and develops a new simple and general methodology for estimating capacity and hence capacity utilization. This has the additional virtue of allowing us to judge the effect of technical change on the capacity–capital ratio. Appendix 6.7 details the sources and methods for all of the empirical measures and appendix 6.8 provides a spreadsheet with all the data tables corresponding to chapter 6 and appendices 6.1–6.7. The new measures are shown to give rise to patterns strikingly different from conventional measures: the corporate maximum rate of profit falls steadily from 1947 onward, providing strong evidence that technical change lowers the average “productivity” of capital in the neoclassical sense. The corporate net operating surplus, which is equivalent to the business measure of Earnings before Interest and Taxes (EBIT) is quite stable in relation to value added, falling modestly in the 1947–1982 “golden era” for labor then rising modestly thereafter as neoliberal policies erode the wage share (figures 6.2 and 6.5). As a result, the corporate average rate of profit falls steadily throughout the first era but stabilizes during the second in the face of a declining wage share. One could say that this was the whole point of the Reagan–Thatcher neoliberal era.

## 2. Part II: Real competition (chapters 7–11)

Chapter 7 presents the theory of real competition which is the theoretical foundation for the analysis in this book. The profit motive is inherently expansionary: investors try to recoup more money than they put in, and if successful, can do it again and again on a larger scale, colliding with others doing the same. Some succeed, some just survive, and some fail altogether. This is *real competition*, antagonistic by nature and turbulent in operation. It is the central regulating mechanism of capitalism and is as different from so-called perfect competition as war is from ballet. Competition within an industry compels individual producers to set prices that keep them in the game, just as it forces them to lower costs so that they can cut prices to compete effectively. Costs can be lowered by cutting wages and increasing the length or intensity of the working day, or at least by reducing wage growth relative to that of productivity. But these must contend with the reaction of labor, which is why technical change becomes the central means over the long run. In this context, individual capitals make their decisions based on judgments about an intrinsically indeterminate future. Competition pits seller against seller, seller against buyer, buyer against buyer, capital against capital, capital against labor, and labor against labor. *Bellum omnium contra omnes*.

Real competition generates specific patterns. Prices set by different sellers in the same industry are roughly equalized through the mobility of customers toward lower prices, and profit rates on new investments in different industries are roughly equalized through the mobility of capital toward higher profit rates. Both produce *distributions* around a corresponding common center. The classical notion of turbulent equilibration is very different from the conventional notion of equilibrium as a state-of-rest. Supply and demand play a role in the process but not in the final outcome, since both are affected by price-cutting and entry and exit. An important point is that price and profit rate equalizations are *quintessential emergent properties*, unintended outcomes of constant jockeying for greater profits.

The notion of competition as warfare has important implications. The competitive firm must be concerned with tactics, strategy, and prospects for growth. The relevant profit must be defensible in the medium term against all sorts of predation, which makes it very different from passive short-term maximum profit in neoclassical theory. In the battle of real competition, the mobility of capital is the movement from one terrain to another, the development and adoption of technology is the arms race, and the struggle for profit growth and market share is the battle itself. There are winners and losers, and places can be switched. No capital is assured of any profit at all, let alone the "normal" rate of profit, so it is completely illegitimate to count "normal profit" as part of operating costs as is conventionally done in orthodox economics. It is equally improper to count interest as part of operating cost. The division between debt and equity determines the division of net operating surplus into interest and profit. The interest rate also serves as an indication of the gap between rewards to active versus passive investment (chapters 10 and 16). Section II of chapter 7 develops the phenomena of price competition, section III those of profit rate competition, section IV unites the two through the notion of regulating capital. Section V summarizes the overall patterns associated of real competition.

Section VI turns to the evidence on the behavior of the firm, beginning with the finding of the Oxford Economic Research Group (OERG) that firms were price-setters forced by competition to keep their prices in line with those of the price-leader. Andrews and Brunner insisted that the OERG findings described the behavior of competitive, profit-driven, price-setting, and cost-cutting firms. Geroski shows that excess profit in an industry stimulates the adoption of best practice methods by insiders and outsiders, that new entrants tend to undercut existing prices, and that even the threat of entry may be sufficient to put downward pressure on prices and eliminate excess profits. Darlin reports that price-cutting behavior is characteristic of competition when there are substantial cost differences. Bryce and Dyer's study shows that more profitable industries had almost five times as many entrants as did the average industry and that challengers approach competition as a form of warfare. Salter's classic study notes that best practice techniques embodied in new plants generally have higher labor productivity, and that there is always a spectrum of techniques within any given industry because new methods are constantly coming into operation and old ones constantly being scrapped. Comparing 1924 to 1950 in the United Kingdom, Salter finds that most of the changes in industrial relative prices can be explained (in a purely statistical sense) by changes in relative labor productivity, the latter in turn being driven by ongoing technical change. Salter's relationships will be shown in chapter 9 to be an aspect of a powerful and more general explanation of relative prices. Megna and Mueller note that while persistent differences in profit rate are the norm, attempts to explain them in terms of market power, collusion, barriers to entry, differences in efficiency, and even alternate measures of profit and capital (including "intangible capital" associated with advertising and R&D) have generally been unsuccessful. Walton and Dhawan note that most business studies find that profit rates decline with firm size, but so do levels of risk and cost of capital. Tables 7.8 and 7.9 show that in a sample of 38,948 firms, the capital-sales ratio rises with firm size while the cost-sales ratio remains roughly constant. The latter is consistent with the observation that new entrants have larger scale and lower costs per unit output, which allows them to set lower selling prices. The data also indicates that the capital-cost

ratios unambiguously rises with firm size. This simple fact has major implications for the path of the profit rate under price-cutting behavior (section VII).

The last part of section IV examines empirical evidence on profit rate equalization. Classical theory expects that new investment is embodied in best practice plant and equipment. Even within a single firm, total capital will embody a variety of technologies and vintages, so we cannot treat the average rate of profit in a firm as a proxy for its regulating rate. The same problem exists at the level of an industry: the relevant measure is the rate of return on new investment. I show that this can be well approximated by the real incremental rate of return on capital, measured as the change in real profit (gross of interest, taxes, and depreciation) over real gross investment. Both variables are widely available across industries and across countries. I examine both average and incremental rates across OECD industries in 1970–1989, and across fifteen US manufacturing industries from 1960 to 1989, across thirty US industries from 1987 to 2005, and in more recent data for incremental rates of return in OECD industries. In every case, average rates of profit tend to remain distinct while incremental rates of profit are strongly equalized. Tsoulfidis and Tsaliki get the same results for twenty Greek manufacturing industries from 1962 to 1991. They also use Mueller’s econometric methodology to test for long-run profit rate equalization: for average rates of profit in fourteen out of twenty industries the estimated long-run deviations of industry profit rates from the overall mean are not statistically different from zero, but for the incremental rate of profit all twenty industries yield estimated long deviations not statistically different from zero. Similar results are shown for Turkey in an excellent paper by Bahçe and Eres. Such results provide considerable support for the classical hypothesis.

Section VII closes the chapter by addressing the all-important question of exactly how the regulating capital itself is selected in the competitive battle—that is, by addressing the “choice of technique.” Actual decisions are always in terms of current and expected market prices. The fact that market prices gravitate in a turbulent manner around prices of production does not imply that the two are close, so we cannot substitute the latter for the former. Second, in keeping with the price-setting and cost-cutting behavior of real competition, firms are forced to select the lowest cost reproducible conditions of production—costs being defined here in the usual business sense as the sum of unit depreciation, materials, and wage costs. Once we allow for fixed capital, the lowest unit cost technique may be different from the highest profit rate one. Moreover, given that real markets are always turbulent, all choices must be “robust” in the sense that they remain valid in the face of normal fluctuations in costs, prices, and profitability. Hence, the appropriate methodology for the choice of techniques is stochastic, not deterministic. If lower unit operating costs are generally achieved through higher unit capital cost (capital-biased technical change in which the capital–cost ratio rises), then the fact that price- and cost-cutting firms select lower cost methods will imply a falling average rate of profit even at a given real wage. By contrast, the conventional (Okishio) selection criterion of the highest profit rate at the “given” price relies on the assumption that firms are passive price-takers, as required in perfect competition, and this implies that the average profit rate rises at a given real wage.

Chapter 8 consists of two main parts. Section I considers various alternative views of competition ranging from classical to post-Keynesian and section II examines the

empirical evidence on pricing and profitability. Section I opens with the classics. Smith and Ricardo (section I.1) and Marx (section I.2) all agree that competition tends to equalize wages rates and profit rates, so that market prices tend to gravitate around, but remain different from natural prices (prices of production). Marx in particular emphasizes the “anarchic” character of these gravitational fluctuations. He generalizes Ricardo’s argument that only certain conditions of production regulate the market price by extending the notion from agriculture to all industry. He also argues that competitive firms are active price-setters and aggressive cost-cutters (unlike the passive price-taking firms assumed in perfect competition), and that the creation of techniques with lower production costs generally requires greater investment in fixed capital per unit. This turns out to be important to his analysis of the choice of technique and the time path of the average rate of profit.

Section I.3 examines the post-classical move away from the analysis of capitalism into the analysis of its idealized form. The price-setting and cost-cutting firm is replaced by a passive price-taker and the anarchical movement of market prices around prices of production is replaced by exact equality obtaining within equilibrium-as-a-state. Competition is taken to prevail only if there is a multitude of small price-taking firms each of which pursues its own myopic interest. Jevons and Walras use this to build a story of a socially optimal and economically efficient market society, and this continues to dominate the profession. Section I.4 argues that the theory of perfect competition is internally inconsistent because it requires irrational expectations. If all firms are exactly alike any action undertaken by one of them must be undertaken by all. Any signal that causes one to increase output will cause all the others to do the same, so market supply will expand significantly and the price will drop. Given that perfectly competitive firms are also perfectly informed, it would be quite irrational for any individual firm to “expect” that it could sell as much as it wanted at any going price. Yet this is precisely what is required in the theory of perfect competition and in macroeconomics founded upon it. It follows that *the theory of rational expectations cannot be grounded in the theory of perfect competition*. Conversely, the theory of perfect competition collapses if firms are assumed to be sensible in their expectations, for even mildly informed firms would recognize that they face downward sloping demand curves under competitive conditions. This sheds an intriguing light on Sraffa’s (1926) critique of standard economics, on Keynes’s treatment of the firm (chapter 12), and even on Patinkin’s passing attempt to get around this difficulty.

Sections I.5 and I.6 examine the Schumpeterian and Austrian arguments. Schumpeter lauds Walras’s model of price-taking firms and maximizing agents but then also says that its static nature is incompatible with the constant creation of new methods and new commodities. He proposes to extend the perfectly competitive model by allowing for perturbations caused by innovations but has very little to say about the resultant patterns of prices and profits. Austrian economics rejects the notion of perfect competition because of its reliance on perfect knowledge, on competition as a state rather than a process, and on firms as passive price-takers rather than active innovators. The Austrian emphasis on competition as a process that bids away excess profits has many similarities to the classical theory of real competition, except for its explicit assumption of rapid profit rate equalization and the lack of a distinction between regulating and non-regulating capitals. Austrian economics also shares the

neoclassical vision that firms are efficient servants of consumers and that union activity and government intervention are unwarranted intrusions into market processes.

Sections I.7–I.9 examine the price theories of monopoly capital, imperfect competition, and Kaleckian and post-Keynesian schools. They all implicitly or explicitly associate competition with perfect competition and point to the historically rising scale and centralization of capitalist production as *prima facie* evidence of a rising degree of monopoly. Hilferding is the first to advance the claim, and Lenin's seal of approval subsequently makes this the official Marxist view. Monopolists are said to be driven to export capital abroad because the alternative of reinvesting their profits in their own sectors would expand supply and drive down prices and profit rates. Sweezy, Baran, Mandel, Bellamy Foster, and others argue that monopoly theory is more "reality based" than competition theory (which they typically conflate with perfect competition). Kalecki's monopoly markup price theory becomes the foundation for the Marxist monopoly school through Baran and Sweezy and for most of post-Keynesian economics. The orthodox theory of imperfect competition is also driven by the attempt to make standard theory more realistic, in this case by relaxing one or more of the assumptions of perfect competition: imperfect knowledge in order to focus on the uncertainty and indeterminacy of the future, non-negligible scale of production to ground the notion of barriers to entry, not very large numbers of consumers and firms to justify price-taking, diminishing returns to justify flat cost curves, and some consumption and production "externalities" arising from interactions of outcomes. Profit maximization is generally retained, but the condition  $p = mc$  is replaced by  $mr = mc$ . Sraffa (1926), Chamberlin, and Robinson are the key figures. Kalecki's central theme is that firms set prices, selling prices differ even for relatively homogeneous products, and lower cost firms charge lower prices. However, these same phenomena are also implied by the classical notion of real competition (chapter 7, section V). Then the distinguishing feature of Kalecki's formulation and of the subsequent post-Keynesian literature becomes the claim that prices are set through stable monopoly markups, in which case long-run profit rates differ even across price-leaders according to their respective degrees of monopoly power. As always, "competition" is generally taken to be the same as perfect competition, safely interred in some distant past.

Modern classical economics (section I.10) emphasizes the central role of competition and argues that market prices gravitate around prices of production, so that the two are not the same. One approach treats the two as close enough to take them as equal. A second position insists that market prices fluctuate considerably during their gravitation processes, so actual decisions are always in the context of fluctuating and uncertain market prices. A third position dispenses with price and profit rate equalization on the mistaken impression that competition requires their exact equalities. Prices and profit rates are then considered random variables and approached through statistical mechanics. I argue that the latter approach is more properly applied to the *deviations* of prices and profit rates from their regulating centers. The final issue concerns the behavior of the firm. Almost all modern classical economists treat the competitive firm in the same manner as neoclassical theory, as a price-taker. At one end, there are those who assume that market prices are close to prices of production and that firms are price-takers, so that competition is close to perfect competition, the choice of technique is based on the highest rate of profit available at some given



price, and that re-switching is a central issue. At the other end, there are those (including myself) who argue that competitive firms set prices and engage in price-cutting, that competition is an antagonistic and destructive process, that the choice of technique is based on the lowest cost, and that re-switching is not a particularly important phenomenon (chapter 9, section X).

Section II opens with a summary of the patterns expected by theories of perfect, imperfect, and real competition, respectively (table 8.1). Perfect competition assumes a very large number of very small firms, identical in scale and cost structure, and all facing the same horizontal demand curve. Firms are assumed to passively take prices and technology as “given,” and this uniform price is assumed to be supremely responsive to market demand and supply. Since firms are identical, they must all have the same profit margins and profit rates. Hence, there can be no correlation between the firm profitability and scale. Imperfect competition theory uses these patterns as benchmarks. Hence, industries in which the number of firms is not very large, the entry scale is not very small, prices are not very flexible, prices and costs are not uniform, and firms face downward sloping demand curves are all deemed uncompetitive. Similarly, price-setting and price-leadership by firms is viewed to be an indication of their monopoly power related to their scale, capital intensity, and relative market share (concentration ratio). By contrast, in real competition the intensity of the competitive struggle does not depend on the number of firms, their scale, or the industry concentration ratio. Price-setting, cost-cutting, and technology variations are viewed as intrinsic to competition. Market prices for a given product are expected to differ within limits, and firms are expected to respond to changes in demand and supply through periodic price adjustments. Newer firms will tend to have larger scale and lower costs, and tend to make room by cutting prices. Older firms will react as best as they can, but do not always fully match newer prices. Hence, in real competition one would expect to find a positive correlation between selling prices and unit costs, and a negative one between these and firm scale and/or capital intensity. Once we allow for price-cutting behavior, profit margins and profit rates can be the same or even lower for larger firms—precisely what most studies find (chapter 7, section VI). Given that more efficient firms tend to be larger and more capital-intensive, one would also expect concentration ratios to be correlated with so-called barriers to entry.

Perfect competition assumes that all firms are alike, so that each firm within a given industry is a regulating capital with a profit rate equal to its industry average. Since competition between industries equalizes profit rates, all firms everywhere must have the same rate of profit. Hence, a persistent difference in firm-level profit rates becomes evidence of imperfect competition, as does any correlation between profit margins and scale or capital intensity. In the theory of real competition, profit rate equalization implies that regulating firms with higher capital output ratios must have higher profit margins. Since capital intensity is linked to scale, one would expect that industries with higher entry scales will have higher profit margins. The distinguishing claim in real competition is that profit rates are equalized across regulating capitals in different industries. So the question becomes: Do industries with high concentration ratios and higher entry requirements have higher-than-normal profits?

Section II.3 examines the supposed nexus between price rigidity and monopoly power. Means attributes the relatively infrequent changes in prices of some firms to their monopoly power. Yet Tucker finds that profit rates are lower for larger businesses

(a common finding, see chapter 7, section VI.3). Eichner presents data in which the average price of concentrated industries is smoother than that of competitive industries. But he fails to note that the smoother prices do not increase any faster over time, and fails to provide evidence that concentrated industries have higher profit rates. Semmler shows that in various studies the degree of price flexibility does not correlate with concentration ratios. Section II.4 notes that if profit rates are equalized, they must be uncorrelated with industry capital intensity. Since the profit rate is the ratio of the profit margin to capital intensity, the former will then be positively correlated with the latter (and hence with scale). Hence, only a correlation of excess profit margins with capital intensity or firm size could be considered as support for the monopoly power hypothesis. Section II.5 addresses the “structure-performance” hypothesis that industries with higher concentration ratios have higher profit rates and/or profit margins, beginning with Bain’s original study and responses by Mann, Stigler, Brozen, Demsetz, and many others. In the end, neither hypothesis stands up in the face of the cumulating contrary evidence.

Chapter 9 focuses on the classical theory of relative prices and on a wealth of supporting evidence. Prices of production are competitive relative prices generated by three essential processes: selling prices equalized across sellers, labor incomes equalized across workers, and profit rate equalized across regulating capitals, all equalizations being turbulent. The classical tradition approaches the final outcome in several analytical steps because this helps identify the underlying structure of relative prices. Section II begins with self-employed producers who purchase their inputs and sell their product in competitive markets and move from one occupation to another in search of higher incomes (incomes not being wages yet since producers work for themselves). Then the mobility of producers across occupations will equalize hourly incomes and the corresponding prices will be proportional to the integrated labor time required to produce the commodities. Integrated labor time refers here to the labor required to produce the given commodity plus that required to produce its inputs and the inputs to its inputs, and so on. Now suppose that the producers have to share their proceeds with capitalists in such a way that each class gets a fraction of the value added, these fractions being the same across all industries (so that wage rates are equalized). Then there is no reason for relative prices to deviate from relative integrated labor times. Hence, neither capitalist relations nor positive profits need cause any such deviations. Furthermore, if capital–labor ratios happen to be the same in each industry, equal profit shares also imply equal profit rates at prices proportional to integrated labor times. This establishes that production price–labor time deviations do not arise *per se* from competition, private property in the means of production, equalization of labor incomes, capitalist relations of production, positive profits, or even from the equalization of profit rates: they arise solely from differences among industry capital–labor ratios. Then we are led to ask how the variation among capital–labor ratios is mapped into the price–labor time dispersion.

Section III follows on the last point by first demonstrating that the relevant dispersion of capital–labor ratios is not of the ones directly observed in each industry, but rather of the integrated ratios each of which is a weighted average of the capital–labor ratio of a given industry and that of its inputs and of the inputs of the inputs, and so on. Each industry’s production price is shown to be the product of two structural factors: its integrated unit labor time that links the industry to the production network in

which it is situated; and its integrated capital–labor ratio. Since the latter is a weighted average of the industry’s direct ratio and the direct ratios of all the industries that enter directly or indirectly into its means of production, the dispersion of integrated ratios is necessarily much smaller than that of direct ratios. This alerts us to the possibility that their contribution to the distance between relative prices of production and relative integrated labor times may be small (as it is shown to be in section IX). Section IV takes up the question of unit-independent and scale-free measure of such (vector) distances and shows that in addition to traditional unweighted root-mean-square type distance measures such as the coefficient of variation and the Euclidean distance, it is possible to develop a weighted distance measure based on the absolute values of deviations. The latter has the simple interpretation of representing the average absolute percentage deviation between any two sets of variables.

Sections V–VI present a great deal of evidence on the distance between market prices, direct prices (prices proportional to integrated labor times), and prices of production from 1947 to 1998. All three measures give roughly the same results. In terms of the weighted distance measure, the distance between market prices and direct prices is about 15%, that between prices of production at the observed rate of profit and integrated labor times is about 13%, and that between market prices and production prices at the observed rate of profit is once again about 15% (table 9.14). The fact that market prices are just as close to direct prices as they are to prices of production seems to be a puzzle given that market prices supposedly fluctuate around prices of production while the latter deviate systematically from direct prices. However, I show that even when market prices fluctuate randomly around production prices as the latter vary with the profit rate (and hence deviate systematically from direct prices) there are many points at which the distance between market prices and direct prices can be as great as, or even lower than, the distance between production price and direct price (figure 9.17). Temporal changes in normalized market, production and direct prices are similarly close. We can use statistical regressions in this case if we work with percentage deviations between sets of prices, because units and scaling factors then cancel out. The highest correlation and lowest distances occur over the smallest available time interval, which is four to five years, although the relations remain robust up to the (next available) interval of nine years: for instance, even over a nine-year interval the relation between changes in market prices and changes in direct prices yields  $R^2 = 0.82\text{--}0.87$  and weighted distance measures of 4%–6% (table 9.10). Comparisons of changes in prices of production at observed rates of profit and direct prices yield similar results: even over a nine-year interval  $R^2 = 0.89\text{--}0.90$  and the weighted deviations are 2%–5% (table 9.14). Finally, following a procedure developed by the eminent US mathematician Jacob Schwartz to address Ricardo’s famous estimate of the sensitivity of relative prices to changes in distribution, Claudio Puty shows that the change in market prices in going from peaks to troughs of successive business cycles averages 7%–8% (tables 9.11–9.12). *This is exactly Ricardo’s estimate!*

Sections VII–X examine the empirical properties of individual Sraffa standard prices, which turn out to be mildly curvilinear within a circulating capital model but entirely linear within a fixed capital one. In both cases, the corresponding wage–profit curves are near-linear (figures 9.8 and 9.12). Sraffa links the potential complexity of individual production prices to possibly complicated movements of industry output–capital ratios, but at an empirical level in the US data these ratios are

near-linear—which is precisely why standard prices and wage–profit curves are near-linear. For all practical purposes, Sraffa’s standard prices are integrated versions of Marx’s transformed values. If standard prices were linear throughout, the elasticity of distance between production and direct prices with respect to changes in the profit rate would be 1. At the empirical level, the elasticities are on the order of 1.10, that is, about 10% different from the linear case, at observed rates of profit (figure 9.14). This too is essentially what Ricardo hypothesized. Not surprisingly, empirical wage–profit curves turn out to be near-linear (figure 9.19). The overall results provide strong support for the classical theory of relative prices. The near-linearity of standard production prices greatly simplifies the analysis of the effects of changes in distribution and in technology, and their empirical strength gives them considerable practical value. They are consistent with the (slightly) curvilinear wage–profit curves we observe, so they do not exclude the logical possibility of re-switching or capital-reversals (although they do imply that such occurrences will be rare).

Section XI closes out chapter 9 with a history of the origins and development of the classical theory of relative prices: Smith, Ricardo, Marx, Sraffa, and the subsequent debates on re-switching and the possibility of aggregate production functions. The evidence in this chapter makes it clear that differences between various price forms are relatively small so that they wash out at the aggregate level and aggregate ratios are essentially the same whether we use market prices, prices of production, or integrated labor times (Marx’s labor values)—as Sraffa himself says.<sup>2</sup> Linear standard prices and wage–profit curves imply two apparently contradictory things: that Marx’s transformation procedure is essentially correct if recast in terms of integrated rather than direct “organic compositions of capital”; and that Samuelson’s aggregate pseudo-production function is basically correct because wage–profit curves are essentially linear. Hence, prices of production arising from the redistribution of surplus value give rise to an aggregate pseudo-marginal product of the capital (in money terms) which is equal to the profit rate at each switch point. This does not imply that the money value of capital determines the profit rate. Indeed, the classical causation is from individual wage struggles on the shop floor to the general rate of profit ( $r$ ) and the corresponding money values of capital  $K(r)$  and output  $Y(r)$ . Similarly, movement along a Samuelsonian wage–profit frontier does not reinstate the neoclassical theory of full employment. The neoclassical claim is that flexible real wages automatically lead to full employment, whereas Marx and Goodwin argue that flexible real wages serve to create and maintains a persistent pool of unemployed labor (chapter 14). There remains the fascinating issue of the properties of input–output tables that may account for the observed linearity of standard prices. Schefold has shown that exactly linear standard prices obtain if the subdominant eigenvalues of the integrated capital-coefficients matrix are all zero, and one possible explanation for this the hypothesis being that the subdominant eigenvalues of random matrices approach zero as the matrix size approaches infinity (appendix 9.1). This would, of course, constitute an advanced mathematical proof of what might be called “Marx’s Last Theorem.”

<sup>2</sup> In his notes, Sraffa says that the “the ratio between their aggregates (rate of surplus value, rate of profit) is approximately the same whether measured at ‘values’ or at the prices of production corresponding to any rate of surplus value. . . . This is obviously true” (Bellofiore 2001, 369).

Chapter 10 extends the classical approach to the theory of finance. The interest rate is the price of finance, financial firms exist to make profit, and competition makes the profit rate of the regulating financial capitals gravitate around the general rate of profit. From this point of view, the competitive interest rate is the “price of provision” of finance and is linked to the general rate of profit just like any other competitive price. For both financial and non-financial firms, the interest rate serves as a benchmark for investment. As both Marx and Keynes emphasize, investment is driven by the difference between the rate of profit and the rate of interest. In this chapter, I focus on the competitive determination of interest rates and bond and equity prices, leaving monetary policy issues to chapters 15 and 16. Section II begins by noting that the interest rate must generally be less than the profit rate if business borrowing is to be viable. The profit rate of a financial firm (bank) is the ratio of its profit (which is the difference between its interest revenue from loans and its costs of operation) to its capital stock (which is the sum of its reserves and its fixed capital). The equalization of the bank profit rate to the general rate of profit implies that for any given desired reserve-to-deposit and deposit-to-loan ratios the interest rate is determined by two things: the general rate of profit and the general price level that affects the costs of inputs such as paper, computers, office space, and labor time. Hence, the long-run competitive interest rate is not a “natural” rate because *there will be a different long-run rate at each different price level*. This provides a direct explanation for “Gibson’s Paradox” arising from the empirical finding that the nominal interest rate and the price level are positively correlated—in direct contradiction to Fisher’s hypothesis that the interest rate moves opposite to the rate of inflation. It also resolves an apparent contradiction within Marx’s argument, in which he vehemently opposed the notion of a natural rate of interest and yet says that financial capital, like all other capital, must participate in profit rate equalization. He is right on both counts. The approach is then extended to derive the yield curve in Hicksian fashion, starting with the bank or division that takes in demand (zero-period) deposits to make one-period loans, moving to the one that takes in one-period time-deposit to fund two-period loans, and so on. Longer loans have greater risks and therefore require higher reserve and deposit-to-loan ratios, so that the interest rates on longer term loans will have to be higher to achieve the same profit rate: the profit equalized yield curve will normally be upward sloping. Profit rate equalization therefore determines the long-run level of the base (one-period) interest rate and the long-run term structure of interest rates. In the short run, demand and supply for various types of loans determine interest rates, but in the long run, structural factors dominate.

Section III extends profit rate equalization to equity prices. Here competition equalizes the real rate of return on equities, which is sum of the rate of growth of real stock prices and the dividend yield (ratio of dividends per share to price per share), with the real incremental rate of profit. This determines the path of real stock prices in a dynamic context. Various standard hypotheses such as the dividend-discount and FED models of the equilibrium stock price are shown to obtain as improbable special cases of the general classical theory. Section IV analyzes bond prices. Arbitrage between financial instruments equalizes bond rates of return with bank interest rates of equivalent duration, and since these bank rates are generally smaller than the general rate of profit, bond rates of return will be below the profit rate. Since equity rates are equal to the profit rate, the bond rate of return will be lower than the equity rate.

This is a well-established empirical fact known in orthodox finance theory as the “equity premium puzzle” because it contradicts the hypothesis that bond and equity rates of return should be equal. Section V summarizes the classical theory and shows that in the stationary case it reduces to the standard dividend-discount model except that the “discount factor” is the profit rate, not the interest rate.

Section VI considers the empirical evidence. The current-cost (real) incremental rate of return on banking capital is shown to gravitate around the general incremental rate on all private capital, as expected by the hypothesis of profit rate equalization for banks. Bond yields are shown to equalize with bank loans of interest, and interest rates of different duration are shown to move together except in abnormal times such as the outbreak of the global crisis in 2007. On the other hand, the bank prime rate on business loans is shown to be generally below the profit rate, except during the last part of the Great Stagflation in which a combination of high inflation and bank and business failures drove up the interest rate. This leads directly to the empirical connection between the nominal interest rate and the price level which is visible from 1857 to 1982, after which monetary policy intervenes to drive the nominal rate ever downward (see chapter 16). By contrast, the Fisherian real interest rate (the nominal rate minus the inflation rate) is definitely not stable, contrary to the expectations of orthodox finance theory. In keeping with classical expectations, the equity rate of return and the corporate incremental rate are very similar, down to having essentially the same means and volatility. Also in keeping with the expectations of classical theory, the bond rate of return is only half of either of the other two rates (tables 10.1 and 10.2). Finally, in the classical argument the average rate of profit is not expected to equal the equity rate of return because the average rate is a mixture of rates of return on all vintages of capital. The appropriate measure is the real incremental rate of profit, which is shown to equalize with the real equity rate. Shiller’s critique of the Efficient Market Hypothesis (EMH) is addressed from this vantage point. The comovement between the equity return and the increment corporate rate of return is so close that there is no basis for Shiller’s claim that the stock market return is characterized by “excess volatility” due to the “irrational exuberance” of investors. Shiller arrives at his “excess volatility” conclusion because he takes the ruling Efficient Market Hypothesis (EMH) as the benchmark, and this requires the assumption that the expected stock market rate of return is constant through time. But the actual stock market rate is highly volatile, so any comparison between it and some constant rate of return is bound to signal “excess volatility” (figure 10.12). The difference between the classical and EMH hypotheses carries over to the definition of the long-run equilibrium (warranted) stock price: in the EMH case it is smooth and quite “out of touch” with the actual real price; in the classical case the actual and warranted prices cycle turbulently around each other in long swings consistent with the theory and particularly with Soros’s notion of reflexivity which is itself a critique of the EMH.

Section VII traces interest rate theories from Adam Smith to modern views. Smith, Ricardo, and Mill treat the long-run interest rate as proportional to the profit rate. Such a relation can be derived from the general argument in section II if one abstracts from operating costs and fixed capital in banks. But then there would be a “natural” rate of interest at each level of the profit rate—something which Marx rightly opposes because he was aware of Tooke’s finding that the interest rate is also related to the price level. At the same time, Marx argues that financial capital also enters into profit

rate equalization, and he even links financial profits to the difference between the interest rates at which they borrow and the rate they charge on their loans. In Volume 3, assembled by Engels long after Marx's death, there is no further treatment of the equalization of bank profit rates or of the term structure. On the neoclassical and Keynesian side, the striking thing is the treatment of finance as if it were a *non-capitalist activity* with neither operating costs nor capital advanced. Once costs and capital have been abolished from the picture, there is no possibility of a price of provision for finance. Then we can only anchor the interest rate in preference structures and expectations. Keynes turned to liquidity preference as the driver of his argument, and this quickly devolved into Hicks's IS-LM apparatus which was in turn suitably modified by neo-classicals to ensure full employment through the putative real balance effects. The neoclassical takeover of the IS-LM framework forced Keynes's followers in a variety of alternate directions. Wray insists on keeping liquidity preference as a foundation, while Panico argues that liquidity preference is insufficient to determine the interest rate because, in the end, this relies on "the common opinion" in the market. On the other hand, Rogers celebrates this conclusion by arguing that the interest rate is indeed purely conventional. Moore contends interest rates are set by central banks (which, however, says nothing about interest rates before central banks) through appropriate adjustments in the money supply. Lavoie and Wray confirm that this is now the consensus view in post-Keynesian economics. At the other end, Panico's path-breaking work recovers the classical analysis of the bank interest rate as a cost-based competitive price derived from the equalization profit rates. It is analyzed in some detail and provides the foundation for my own approach, albeit along somewhat different lines.

Section VIII concludes chapter 10 with a discussion of modern finance theory whose central hypothesis is that the mobility of capital equalizes risk-adjusted rates of return. This includes Markowitz's return-risk trade-off, the approximate equality of risk-adjusted returns in the Capital-Asset Pricing (CAPM) and Arbitrage Pricing Theory (APT) models, and the stochastic equality between expected and actual returns in Efficient Market Theory (EMT). The latter is based on the hypothesis that the price of an asset must reflect all available information because if it did not there would be a profit opportunity which would attract speculative capital. The ubiquitous dividend-discount model, in which the equilibrium price of a stock is said to be equal to the discounted present-value of the expected stream of dividends, is shown to derive from this same principle provided we assume that future rates of return are expected to be constant over time and that dividends per share grow at some constant rate lower than the rate of return. Outside of academia, most practitioners focus instead on earnings, not dividends. For instance, there are literally hundreds of models based on benchmark price-earnings ratios including the FED model derived in section V as a special case of the classical formulation. None of these models work well at an empirical level.

Chapter 11 closes Part II of this book by applying the classical argument to international competition, that is, international trade balances and terms of trade (real exchange rates).

The theory of international trade is a critical part of modern debates about the costs and benefits of the globalization of production and finance. Neoliberalism portrays markets as self-regulating social structures that optimally serve all economic needs,

efficiently utilize all economic resources, and automatically generate full employment for all persons who truly wish to work. Proponents of neoliberalism point to the indisputable fact that the rich countries are market-based economies that developed in the context of a world market. Critics of neoliberalism dispute all of these claims. They note that rich countries, from the old rich of the West to the new rich of Asia, relied heavily on trade protectionism and state intervention as they developed and that they continue to do so even now. They contend that the trade liberalization imposed on the developing world has actually led to slower growth, greater inequality, a rise in global poverty, and recurrent financial and economic crises in most countries. Most important, they generally argue that in any case orthodox free trade theory is irrelevant because free competition does not prevail even in the rich countries, let alone the poor ones—a standard trope among heterodox economists because they conflate competition with perfect competition (chapters 7 and 8). This chapter demonstrates that the theory of real competition has a very different set of implications for international trade. The conventional (Ricardian) theory of free trade does not follow in a competitive context and the very patterns to which heterodox economists point as evidence against (perfect) competition can be explained from real competition. From the latter perspective, globalization has worked as expected—favoring low-cost producers over the high-cost ones.

Section II examines two crucial premises of the theoretical foundations of orthodox trade theory: (1) that free trade is regulated by the principle of comparative costs; and (2) that free competition leads to full employment in every nation. The principle of comparative costs is eminently familiar, most often presented as the proposition that a “nation” would always stand to gain from trade if it were to export some portion of the goods it could produce comparatively more cheaply at home, in exchange for those it could get comparatively more cheaply abroad. It is implicit that trade will be balanced (i.e., that the value of imports will be equal to the value of exports). But this purely normative proposition has little significance unless it can be shown that free trade among market economies actually creates such outcomes. International trade is actually conducted by profit-driven exporting and importing firms. Therefore, whenever conventional trade theory seeks to appear more realistic, it switches to the positive claim that free trade will be regulated by comparative advantages and that the terms of trade will always arrive at a point which equates the values of exports and imports. No nation need fear trade due to some perceived lack of international competitiveness because, in the end, free trade will make each nation equally competitive in the world market. This conclusion requires that the terms of trade of any country will automatically and successfully move to eliminate trade deficits or surpluses. The assumption of universal full employment in rich and poor countries is equally critical: after all, who can say that trading exports for imports is a “gain” if that outcome is achieved at the expense of sustained job losses? The theory of comparative advantage then seeks to explain the determinants of comparative costs. For instance, on the dual assumptions that trade is ruled by comparative costs and that full employment always obtains, the Heckscher–Ohlin–Samuelson (HOS) model of comparative advantage claims that differences in national comparative costs are rooted in differences in national “endowments” of land, labor, and capital.

All three of the central propositions of orthodox trade theory have been vigorously disputed. The notion of universal full employment becomes a cruel jape in light of the



fact that there were a billion people in the world who were unemployed or underemployed even at the height of the global boom preceding the 2007 global crisis. The claim that a fall in the terms of trade will eventually improve the balance of trade has long been dogged by the infamous “elasticities problem.” And the claim that a trade deficit will automatically lower the terms of trade until the deficit is eliminated is bedeviled by the simple fact that balanced trade simply does not obtain anywhere, not in the developing world, not in the developed world, not under fixed exchange rates, not under flexible exchange rates. On the contrary, persistent trade imbalances are the rule.

Section III traces the two dominant reactions to the empirical problems of standard trade theory. The first type focuses on the fact that balanced trade and/or Purchasing Power Parity (PPP) are only meant to hold in the long run, so that existing post-war data (now spanning seventy years or so) may not be long enough. Others have shifted ground by focusing on a host of short-run models that contradict each another and many elements of the reality they intend to explain. Despite the fact that many mainstream economists “readily admit their failure,” the underlying notion of comparative cost advantage continues to dominate textbooks and models and economic policy itself. The other major reaction has been to modify one or more of the standard assumptions by incorporating oligopoly, economies of scale, and various concrete factors such as the composition of trade, differential elasticities of demand, and differences in technology and in accumulated and/or institutionalized human knowledge. All of these give rise to particular exceptions to the standard results, which in turn provide some (limited) room for state intervention in strategic sectors and strategic activities such as R&D. The resulting models are extremely complicated, encompass multiple possible outcomes and provide “few unambiguous conclusions.” I argue that the real problem lies at the very root of these models, which is the Ricardian principle of comparative cost.

Section IV re-examines Ricardo’s principle of comparative cost in light of the theory of real competition. In real competition within a nation, firms constantly seek to cut their costs in order to be able to cut their prices and displace their competitors. Firms with lower costs tend to emerge more often as winners while those with higher costs are more likely to end up as losers. This is the central selection mechanism of capitalist competition. Smith emphasizes that “private profit is the sole motive” in the application of capital to domestic or international trade. Ricardo begins from this same point, seeking to show how international trade patterns arise from the actions of individual profit-seeking capitals in different countries. In order to bring out the stark logic of his argument, Ricardo begins by assuming that Portuguese capitals initially have lower cost-based prices in all commodities, so that they dominate both English and Portuguese markets. But then, as money flows into Portugal from England, Portuguese costs and prices rise and English costs and prices fall. We can imagine that as Portuguese goods become progressively more expensive and English goods progressively cheaper, the Portuguese commodity with the smallest absolute cost advantage over its English counterpart will be the first to switch from the winner’s column to the loser’s. From the English point of view, this will be the commodity with the smallest cost disadvantage. But unless trade becomes balanced, the process will continue and the Portuguese commodity with the second smallest advantage (the English one with the second smallest disadvantage) will switch columns, and so on. All of this obtains

through the actions and reactions of individual profit-seeking producers in the two countries. When the Ricardian process comes to rest it will appear as if “Portugal” had chosen to specialize in producing the goods in which it had a “comparative cost advantage,” exchanging them for commodities of equal money value (since trade is balanced at the rest point) consisting of goods in which “England” had a comparative cost advantage. This allows Ricardo to jump from the argument that the behavior of individual profit-seeking firms will lead to the rule of comparative cost to the proclamation that countries should use comparative costs to determine their trade patterns. Neoclassical economics often skips the derivation altogether, resorting instead to the fictional representation of England and Portugal as individuals each of whom trades in order to “gain.” This has the ideological value of instilling the notion that the very purpose of free trade is to benefit all nations, rather than to make profits for their businesses. The section includes an extended treatment of the formal structure of the theory of comparative costs.

Ricardo’s conflation of the balance of trade with the balance of payments is extremely important to his construction. A country’s balance of payments is the sum of net inflows into the country: exports minus imports (the trade balance), direct investment in the country by foreigners minus investment abroad by domestic agents, short-term capital inflows such as private or business bonds purchased by foreigners (i.e., loans made by foreigners to domestic agents) minus similar financial transactions made in foreign countries by domestic agents, and so on. Ricardo proceeds as if commodity trade flows are completely separated from financial flows, so that a trade balance is synonymous with a payments balance. Money appears in his story as medium of circulation, but never as financial capital. This is extremely odd from a historical point of view, since the export and import of financial capital (international borrowing and lending) is intrinsically linked to the flow of funds arising from the export and import of commodities. It is equally odd from a theoretical point of view because it *implies that trade and finance flows are completely divorced from each other*. Both Marx and Harrod point to this as a critical weakness in Ricardo’s logic.

Section V develops the classical theory of absolute cost advantage. The Ricardian argument is really a story about the determination of international regulating capitals. When trade opens, Portugal and England each produce both wine and cloth, so there are two different regulating producers for each good, one in each country. Despite the fact that Portugal has the initially lower cost-based prices in both goods, the comparative costs argument says that international competition will end up selecting British firms as the regulating capitals for cloth leaving Portuguese firms with the regulating role for wine. In the theory of real competition, the price-leader (regulating capital) in any industry is the one with the lowest unit cost, the term “cost” now defined in the proper business sense as the sum of unit wages, materials, and depreciation. The first difficulty with the Ricardian story is that changes in the relative international prices of goods will also affect the relative costs of these same goods. This is the logical extension of Sraffa’s central point that prices and costs are inextricably intertwined (chapter 9, section XI). Then comparative costs may not change at all in response to any changes in the real exchange rate (nominal exchange rate and/or the relative national price level), leaving Portuguese capitals in charge of both industries and eliminating British ones. Even if comparative costs do respond to changes in real exchange rates, they may not respond sufficiently to displace Portuguese capitals,

so once again British capitals are doomed. To put it differently, sufficiently large absolute costs advantages will not be overturned by real exchange rate effects. Worst of all comparative costs may change in the “wrong” direction (i.e., they may make the absolute cost advantage of Portugal country even greater). This means that even if the real exchange rate did automatically vary with the trade balance, as Ricardo supposes, comparative costs will not move in the Ricardian manner as long as real costs (real wages and productivity) are determined at the national level. It is formally demonstrated that for given real wages and specific industry efficiencies in each country, in a two sector model the comparative cost in any industry is a ratio of two linear functions of the international relative price and may fall or rise with the relative price depending on the coefficients. Moreover, the extent of any such a movement is itself limited by the relative structures of production. In the end, international competitiveness will be tied to differences in efficiency, real wages, and technical proportions, and there is *nothing in free trade itself that will eliminate absolute cost advantages or disadvantages*.

The second problem with the Ricardian theory is that real exchange rates need not change at all in the face of trade imbalances. Marx comments that a country with a trade surplus will experience an increase of liquidity which will lower its interest rate, while a country with a trade deficit will experience a tightening of liquidity and an increase in the interest rate—all through the normal functions of capital markets. Harrod comes independently to the same conclusion. With capital flows offsetting trade imbalances, the net effect on the balance of payments will depend on the relative magnitude of these two effects: the exchange rate may not change at all, or if it does, it may change in the “wrong” direction (i.e., the exchange rate of the trade surplus country may depreciate rather than appreciate).

In international real competition, the regulating capitals will essentially be those with the lowest integrated real unit labor costs. Assuming that countries export the goods in which they have the lowest costs (for given quality), the terms of trade of any country will depend on the ratio of the integrated real costs of its exports relative to that of the producers from which it gets its imports. The key point is that the terms of trade are pinned by national real wages and structures of production, so that they cannot also move to endogenously balance trade as in the Ricardian theory. The classical formulation can be extended to cover nontradable goods, which will affect input costs insofar as they enter into production and affect the money wage insofar as they enter the wage basket. Then the classical argument implies that the terms of trade (real exchange rate) is driven by two components: relative real regulating costs and the ratio of tradable/nontradable goods. A similar expression is developed for the common currency ratio of any two national price indexes, which immediately tells us that this ratio will be constant only if the two had the same overall composition in the sense of having the same composition of goods and the same ratio of nontradable to tradable prices. The classical argument therefore implies that PPP will not generally hold.

The application of real competition to the theory of international trade leads to several distinct propositions. First, industry comparative costs and terms of trade are determined by relative real wages, relative productivities of regulating capitals, and the effect of tradable/nontradable goods. Second, the direction of a nation’s trade balance is determined by its absolute cost advantage or disadvantage (a classical channel) while its size will also depend on relative national incomes (a Keynesian channel).

Changes in the latter will affect the trade balance but will not permanently switch it from surplus to deficit unless they switch comparative costs. Third, trade imbalances will create payments imbalances which will affect interest rates and induce short-term international capital flows (a classical channel), and perhaps also change national income through their influence on investment (the Keynesian channel). The end result will be that countries with absolute cost advantages will recycle their trade surpluses as foreign loans while countries with absolute cost disadvantages will cover their trade deficits through foreign borrowing. All of this will arise through the workings of free trade and free financial flows, though, of course, policy measures may produce similar effects.

Section VI compares the standard and classical theories of free trade with the empirical evidence. The comparative cost hypothesis implies that the real exchange rate will vary so as to ensure that trade remains balanced while the international real competition implies that trade imbalances will be the norm. Trade data for fifteen major countries over the half-century from 1960 to 2009 makes it abundantly clear that trade does not generally balance. The orthodox PPP hypothesis posits that real exchange rates will be stationary over the long run, but the large empirical literature discussed in this section establishes that PPP does not hold. A chart of the real effective exchange rates in terms of producer prices for the United States and Japan shows both to be highly trended in opposite directions. The PPP argument can also be formulated as the hypothesis that nominal exchange rates will depreciate at the same rate as inflation (so as to maintain a constant real exchange rate). The US and Japan data makes it clear why this (relative) version of PPP is equally unsupported as a general empirical proposition. However, in the particular case of high inflation, (relative) PPP does appear to hold. The classical theory of trade predicts both the trended nature of real exchange rates evident in the US and Japan data and also the correlation between nominal exchange rates and inflation rates observed in the case of high relative inflation. The classical hypothesis is that the real exchange rate  $e_r \equiv p \cdot e/p_f = (p/p_f) e$ , where  $p$  = the domestic price level,  $e$  = the exchange rate (foreign/domestic currency), and  $p_f$  = the foreign price level, depends on relative real unit labor costs and the tradable/nontradable price ratio. Since the latter two terms change slowly from year to year, the real exchange must also change slowly (except for shocks). But the real exchange rate is the product of the domestic relative price level ( $p/p_f$ ) and the exchange rate. Hence, when the relative price level rises sharply in the face of rapid domestic inflation, the nominal exchange rate must depreciate at roughly the same rate.

The preceding argument also implies that the real exchange rate will be linked to corresponding integrated real unit labor costs adjusted for the ratio of tradable/nontradable prices. Direct unit labor costs were used in the absence of data on integrated costs to construct adjusted real unit labor for the two countries relative to their trading partners and the corresponding charts show that each country's real exchange rate does indeed track the classical fundamentals. On the econometric side, the actual and fundamental variables were found to be cointegrated with speeds of adjustment which are statistically significant and of the correct sign. Finally, it is shown that the deviations of the real exchange rates from adjusted relative real unit labor costs are stationary. Given the data limitations discussed and the large impact of the capital-flow and interest-rate shocks, it is remarkable how stable this actual/fundamental ratio is over the long run. Hence, the classical approach also provides us with a robust policy

rule-of-thumb for the competitively sustainable level of the real exchange rate—a rule which is clearly superior to the widely used PPP hypothesis.

### 3. Part III: Turbulent macro-dynamics (chapters 12–17)

Profit is central to both micro- and macroeconomics. The second part of this book elaborated on the microeconomic aspects: firms are active profit-seekers, price-setters, and cost-cutters operating under conditions of conflict and uncertainty created by their own actions. This is competition as it really exists, as the driving force in the determination of production decisions, technological change, relative prices, interest rates and asset prices, and exchange rates. Growth originates at the cellular level, and the measure of its success is the excess of the profit rate over the interest rate. This part of the book will draw out the linkages between real competition and effective demand.

Chapter 12 tracks the rise of modern macroeconomics beginning with Keynes's break with the prevailing orthodoxy and culminating with its recapture by neo-Walrasian economics. Chapter 3 had previously established that emergent macroeconomic properties that cannot be reduced to the desired outcomes of all-seeing representative agents. Hence, micro features do not necessarily carry to the macro level and any given macro pattern may be consistent with many different (even contradictory) micro foundations. In order to distinguish among competing hypotheses we must consider the validity of their microeconomic assumptions. The classical notion of equilibrium-as-a-turbulent-process implies that we must be explicit about the time of gravitation, while the fact that growth originates at the cellular level means that we must work with growth rates or ratios of variables. In real competition, firms face downward sloping demand curves, set prices, have different costs, and partition into price-leaders and price-followers (regulating and non-regulating capitals). Finally, money is endogenous and non-neutral, and aggregate demand and supply are both rooted in profitability so that macroeconomics cannot be reduced to either supply- or demand-side approaches.

At an aggregate level, we can express *ex ante* excess demand as  $ED \equiv D - \mathbb{Y} = [(C + I) - (Y - T)] + [G - T] + [EX - IM] = [I - S] + [G - T] + [EX - IM]$ , where  $D$  = aggregate demand for domestically available goods is the sum of consumption ( $C$ ), investment in desired stocks of fixed capital and inventories ( $I$ ), government ( $G$ ) and export ( $EX$ ) demands,  $T$  = total private sector taxes (households and business), and  $\mathbb{Y}$  = domestically available supply is the sum of domestic supply ( $Y$ ) and imports ( $IM$ ). This accounting relation identifies the sectoral sources of excess demand. In the most abstract case with no government or foreign sector, excess demand reduces to the familiar balance between investment and savings  $ED = I - S$  which plays a critical role in Keynes's break with the orthodoxy of his day. Since sales in excess of supply depletes inventories, we can also derive the corresponding *ex post* national accounts identity by substituting unplanned inventory change  $-\Delta INV_u$  for excess demand  $ED$  to get  $[(I + \Delta INV_u) - S] + [G - T] + [EX - IM] = 0$ . Neither of these two identities is a "budget constraint," since  $ED$  can take on positive or negative values. It is only by further assuming aggregate demand–supply equilibrium  $ED = -\Delta INV_u \approx 0$  that the three balance identities are converted into the constraint  $[I - S] + [G - T] + [EX - IM] \approx 0$ . The question is: How long does equilibration take? Neoclassical theory typically assumes instantaneous and continuous

equilibrium. Keynes usually focuses on comparative statics, so time disappears from view, but in some places he recognizes that production takes time—in which case the multiplier must be a temporal sequence. Modern Keynesian and post-Keynesian macroeconomic models characteristically avoid this issue by treating observed (annual or even quarterly) data as representing equilibrium outcomes. Given that excess demand is reflected in unplanned inventory changes, I would argue that it is more sensible to consider the three- to five-year (twelve- to twenty-quarter) inventory cycle (business cycle) as the equilibrating process for aggregate demand and supply. This certainly casts a different light on the political and social implications of macroeconomic policy. Finally, harking back to the discussions in chapters 4 and 7, normal capacity output  $Y_n$  is defined as the normal (potential) output corresponding to the lowest average cost (cost being defined in the business sense). This point is generally below the maximum (engineering) output, so firms typically have substantial desired reserve capacity—which is precisely why they can rapidly increase output in the short run. True excess capacity would only exist if output is persistently below the normal level. Since firms introduce new plant and equipment with some normal capacity in mind, normal capacity utilization exists when the actual output–capital ratio is equal to the desired output–capital ratio. This is a particularly important form of stock-flow consistency, so it is an irony that it is generally ignored or even denied in the Keynesian tradition.

Section II outlines the basis structure of the pre-Keynesian macroeconomics that had replaced the classical analysis of real capitalism with a postclassical analysis of a fictitious idealized system (chapters 7 and 8). Keynes took aim at certain core propositions which he attributed to the orthodoxy of his time, even though these notions were not fully formalized at the time of his attack: rational maximizing agents operating with perfect knowledge under perfect competition and stable expectations about the future; markets, including the labor market, that always “cleared” quickly and efficiently, so that full employment was the “normal state of affairs”; aggregate demand that adapted to full employment aggregate supply (Say’s Law) through automatic adjustments in the real interest rate in the market for loanable funds; and a general price level determined by the quantity of money. Real variables (including the real interest rate linked to the real profit rate) were determined in commodity and labor markets (the “classical dichotomy”) and nominal values were determined through the effects of the money supply on the general price level (the Quantity Theory of Money). Money was viewed as neutral on the grounds that it had no effect on the equilibrium values of real variables. Not surprisingly, government intervention was “neither necessary nor desirable.” An increase in supply of labor would lead an equal increase in employment but only at a lower real wage. Conversely, attempts by unions and the state to increase real wages above their market (presumed to be equilibrium) levels would only result in unemployment. To understand the logic of the basic neoclassical model, it is useful to recall that at the abstract level aggregate excess demand  $ED = [I - S]$ . Neoclassical theory assumes that private investment constituted a demand for loanable funds, that private savings provided the corresponding supply of loanable funds and that both responded solely to the real interest rate. Then equilibrium in the loanable funds market ensured that  $I = S$  and hence  $ED = 0$  (i.e., that aggregate demand would adjust to full employment aggregate supply). In the end, the system was supposed to quickly and efficiently produce an aggregate quantity of output that provides full employment and

simultaneously generates an aggregate demand sufficient to realize this same output. On this reasoning the widespread unemployment in the 1920s and subsequently in the Great Depression of the 1930s would soon be eliminated if the market was allowed to run its course. Government intervention would only be counterproductive, it was thought.

Section III takes up Keynes's break with the teachings of his day. Persistent mass unemployment following World War I convinced him that real markets did not work in the manner prescribed in the textbooks. Long before he wrote the *General Theory*, he was proposing that governments all over Europe engage in large-scale deficit-financed public expenditures. At the same time, he was struggling to identify the crucial theoretical flaws in the orthodox argument, ultimately zeroing in on two critical claims: that the real wage would move quickly to restore full employment; and that the real interest rate would automatically move to create the necessary amount of aggregate demand. His first step was to note that since production takes time, individual firms must hire workers and purchase inputs on the basis of profit anticipated from expected demand. On the other hand, actual aggregate demand arises from individual household consumption expenditures linked to income generated by current production; and individual business investment expenditures motivated by long-term profit expectations which were notoriously volatile, subject to "tides of irrational optimism and pessimism." He handled this aspect of his argument by taking investment as given in the short run but capable of rapid change from one short run to the other. There was no reason to believe that actual aggregate demand generated by the expenditures of many millions of consumers and firms would just match the expected demand that motivated the individual firms so that imbalance would be a normal state of affairs. Keynes leaves this aside in order to focus on the determinants of the equilibrium level of output and employment. With investment being "given" in the short run, savings must do the adjusting. But savings is the part of income which is not consumed and consumption is dependent on income created by production. So in the end production and hence employment must adjust to make savings equal to investment (i.e., to make aggregate supply equal to aggregate demand). This is Keynes's answer to Say's Law. A key assumption is that savings is a stable fraction of income: if investment rises by 100 and savings is one-fifth of income, output must rise by 500 to make savings return into balance with investment: the Keynesian multiplier. The same logic implies that a rise in the savings rate (greater thriftiness) will make aggregate savings exceed investment so that output and employment must fall in order to bring savings back into line with investment—hence, the Keynesian Paradox of Thrift.

All of this is predicated on investment being given in the short run, so it leads naturally to the question of how investment reacts. Like Marx, Keynes views investment as driven by its expected net profitability which is the difference between the expected profit rate (the marginal efficiency of investment) and the interest rate. It is plausible that a rise in unemployment would dampen profit expectations and raise the cost of borrowing in the face of increased risk, both of which would cause investment to fall and worsen matters. Keynes was clearly aware of the central neoclassical claim that unemployment would lower the real wage and thereby raise the normal-capacity rate of profit so that investment, output, and employment would eventually rise. He countered with a series of objections: wage-bargains are in terms of money, not real wages,

so a fall in aggregate demand that generated unemployment would also lower prices and initially raise the real wage, thereby making matters worse; lower wages would reduce cost and tend to reduce prices, so the real wage might even rise from this effect also; even if money wages were lowered this might make things worse by decreasing consumption and hence aggregate demand; and any reduction in prices might also undermine business confidence and further dampen profit expectations. On the side of the interest rate, he substituted his own liquidity preference theory for the neoclassical loanable funds argument. The interest rate, says Keynes, is determined by the demand and supply for money balances. Money supply is determined by the state. Money demand depends on income and the interest rate viewed as the reward for parting with the liquidity benefit of holding money, the latter being motivated by the need to hold money as insurance against rainy days, to facilitate transactions, and perhaps to invest some time later. All of these motivations depend on the state of confidence in the future, which is precisely why a collapse of confidence triggered by a crisis could precipitate a flight from financial assets into cash and provoke a *rise* in the interest rate at the very time that a fall was needed. Even if the state were to step in and reduce the interest rate, this might not override the fall in confidence. For all of these reasons, in a crisis it would be far better to use fiscal policy and have the state directly pump up aggregate demand through deficit spending, just as he had earlier advocated in the aftermath of World War I.

Keynes's argument got quickly trapped within the static confines of the Hicksian IS–LM framework. In Keynes's own argument, equilibrium output is determined by investment through the multiplier (IS), and investment depends on the excess of a volatile expected rate of profit over the interest rate. Hicks eliminates the expected profit rate so that investment is reduced to a simple passive function of the interest rate. It then becomes a mystery when in the face of bleak expectations (as in the current crisis) a reduction in the interest rate does not spur investment. His treatment of money demand (LM) similarly downplays volatility in money holding decisions so that money demand becomes to a stable positive function of the level of current (rather than expected) income and a negative function the interest rate (since a higher interest rate of financial assets will induce agents to hold less idle money balances). IS–LM equilibrium then requires a particular combination of income (output) and interest rate. The Hicksian formulation was extended to allow for government and export demand, in which case expansionary fiscal policy was supposed to raise the equilibrium level of output at the cost of a higher (nominal) interest rate. On the other hand, expansionary monetary policy would increase the money supply at a given price level and shift the LM curve outward thereby raising the equilibrium output but lowering the interest rate. It follows that the state could always exercise some combination of fiscal and monetary policy to bring output to the full employment level without affecting the interest rate or even the price level. The IS–LM framework also retains the Keynesian paradox of thrift in attenuated form because a reduction in the savings rate at a given level of investment raises the IS curve which raises the level of output (the paradox of thrift) but also raises the equilibrium interest rate which mitigates but does not overturn the initial effect.

At this level of abstraction, the price level rises only when aggregate demand exceeds full employment output. Robinson had already proposed that prices would start rising somewhat before this point, and by the early 1960s this idea was operationalized



by adding an inflation–unemployment curve to the basic Keynesian toolbox. Phillips had originally found that the rate of change of money wages rose in a nonlinear manner when unemployment fell below some critical level. This was restated as a stable inflation–unemployment curve along which Keynesian policymakers had the option of trading a higher inflation rate for a lower unemployment rate. Everything seemed manageable at first, but then things began to fall apart. A stable Phillips curve implied that inflation would fall as unemployment rose, yet by the 1970s, unemployment had risen and inflation had also risen. By the 1980s, the Phillips curve had disappeared in all major countries and “Hydraulic” Keynesianism was finished. We see in chapter 14 that there is indeed a clearly visible and stable Phillips-type curve, but it is not in terms of the rate of change of money wages or even prices. Knowledge of its existence might have enabled the Keynesian to provide a coherent defense against the monetarist and New Classical counter-revolutionaries.

Section IV analyzes the rise of neo-Walrasian economics which was set up in the 1950s and 1960s by Samuelson’s enormously influential mathematical restatement of (Marshallian) economics. Friedman’s revival of the Quantity Theory of Money (QTM) transformed Keynes’s money demand–supply relation into the hypothesis that velocity of circulation of money was stable in any given institutional configuration. His empirical work with Anna Schwartz concluded that an increase in per capita money supply would primarily lead to an increase in nominal income per capita. Given “long and variable” lags between the two, it was best to maintain stable growth of the former in order to maintain stable growth in the latter. He subsequently added the hypothesis that in a static economy real output “can be regarded as constant,” as in the “flex-price full employment” version of the IS–LM model in which equilibrium real output is determined by the labor supply and the equilibrium real interest rate is immune to monetary factors. Then an increase in the money supply translates solely into an increase in the price level and a money supply growing faster than output gives rise to steady price increases (i.e., inflation). The trouble was that by the 1970s, the supposed stable empirical relation between the money supply and the price level “had utterly fallen apart” all over the advanced world despite various efforts to rescue it by changing the definition of money. So in the end the new QTM lasted no longer than the Keynesian theory it sought to displace. By this time, all macroeconomic theories faced the difficulty of explaining rising unemployment occurring hand in hand with rising (rather than falling) inflation. Both Phelps and Friedman argued that observed unemployment was really the result of structural characteristics of actual labor and commodity markets, including market imperfections, stochastic variability in demands and supplies, the costs of mobility, and so on. The key point was that these real world characteristics led to a “natural rate” of unemployment dependent only on real factors as opposed to monetary ones. Both authors concluded that while unanticipated increases in aggregate demand would lead to temporary increases in real output and employment insofar as workers and firms initially failed to recognize that prices would rise, this stimulus would dissipate over time as prices rose so that unemployment would return to its natural level. Hence, Keynesian policies seeking to maintain an unemployment rate below the natural one would have to continually pump up the system through unexpected increases in aggregate demand whose cumulative effect would be an ever-increasing rate of inflation. This led to the Non-Accelerating-Inflation-Rate-of-Unemployment (NAIRU) argument that

the natural rate of unemployment is the only rate at which the inflation rate will be stable (see chapter 15).

New Classicals operate within this framework. They adhere to the notion of a natural rate of unemployment and to the notion that only surprises in economic policy can bring about temporary deviations from the natural rate of unemployment. But they transfer their allegiance from Marshall to Walras by explicitly assuming perfect competition, complete price, wage, and interest rate flexibility, perfect arbitrage, continuous market clearing, and the absence of money illusion (so that only relative prices matter for agent decisions). And they bring a new weapon to the fray: the concept of rational expectations in which theoretical agents populating a model universe must be presumed to “know” the structure of model in which they exist and to make use of this information in an efficient manner. Lucas combines the natural rate hypothesis with the notion of model-consistent expectations that are also hyper-rational. As with earlier arguments, only unexpected changes in policy (surprises) will change economic outcomes, but now there can be no extended effects because once the policy is in place hyper-rational agents immediately catch on so the economy jumps back to the natural rate of unemployment and prices shoot up. A further distinctive feature of the New Classical argument is the claim that the “structure” of the macro-economy is itself the result of dynamic optimization by representative agents so that the structure itself must change as agents adjust their behavior to new policies. This “Lucas critique” became highly popular at a theoretical level, although the empirical evidence was far less kind. I have already argued the contrary hypothesis that aggregates are generally “robustly indifferent” to the details of individual actions (chapter 3, section III). Given the New Classical assumptions of continuous market clearing and completely flexible wages and prices, temporary misperceptions in the face of surprises become crucial in explaining the positive correlations between demand, inflation, real output, and employment over the business cycle. But by the early 1980s the evidence against the monetary surprise and “informational confusion” hypotheses began to mount. Real Business Cycle Theory (RBCT) developed by retaining the hypotheses of rational expectations and continuous market clearing and adding random productivity shocks to generate aggregate fluctuations that mimicked business cycles. Agents were still assumed to have rational expectations, the aggregate economy was still treated as an interaction between a representative firm and a representative household and business cycles were taken to be strictly equilibrium phenomena. A technology shock was assumed to be propagated through the economy by the consumption smoothing response of households, the investment (“time to build”) responses of businesses, and by intertemporal substitution between labor and leisure. Full employment always obtains, so any drop in the employment is simply due to the fact that workers choose to substitute leisure for labor. In such a framework monetary policy is sidelined because it cannot influence real variables and there is no distinction between the short and long run (so that fluctuations are inseparable from trends) because the economy is continuously in equilibrium. RBCT theorists eschew econometric testing of their hypothesis in favor of simulations of “toy” models whose parameters are selected (calibrated) to make the model mimic (some) observed patterns and then changed to investigate the supposed impact of changes in policies and structure. Not surprisingly, there has been considerable criticism of the empirical significance of RBCT models. New Keynesian economists also begin from

standard micro foundations and the general equilibrium framework in which it is embedded, but they focus on introducing a plethora of “imperfections” such as costly price adjustments and imperfect competition in markets for commodities, labor, and credit. Given the inadequacy of the underlying theory there are a large number of potential imperfections from which to choose so New Keynesian economics now “consists of a ‘bewildering array’ of theories . . . [whose] ‘quasi religious’ adherence to microfoundations has become a disease” (Snowdon and Vane 2005, 343, 360–364, 429). New Behavioral Economics operates on the standard micro foundations themselves by incorporating asymmetric information, credit rationing, group norms of fairness, imperfect competition, rule-of-thumb behavior, and the weaknesses of certain cultures. The trouble is that each of these is meant as a single modification of the standard micro foundations, rather than starting from a different point altogether (chapter 3).

Sections V and VI examine the macroeconomics of the heterodox “imperfec-tionist” tradition whose micro foundations were previously analyzed in chapter 8. Kalecki’s macroeconomics is similar to Keynes’s in its short-run focus and its distinction between induced and autonomous components of aggregate demand. His original argument on effective demand was actually in terms of “free competition” which made it even more congruent to Keynes. Investment is given in the short run but over the longer run it responds positively to the gap between the prospective rate of profit and the rate of interest. The interest rate is determined by monetary factors and the profit rate is determined by the wage share and the rate of capacity utilization. Unlike Keynes, Kalecki incorporates class into his analysis by partitioning total income into that of workers and capitalists and assuming that each group has a fixed (marginal) propensity to save. The Kaleckian multiplier relation is therefore the same as the Keynesian one except that the aggregate propensity to save depends on the ratio of profits to wages, which is in turn determined by the monopoly markups that firms add to their prime costs. Markup pricing also implies that for given materials and labor coefficients, money prices are proportional to money wages. Then price inflation must be rooted in money wage increases. Kalecki’s argument further implies that for a given degree of monopoly the real wage and the wage share are not affected either by the unemployment rate or by worker struggles. Yet he was uncomfortable with the conclusion that the working class was powerless to change its own standard of living, so near the end of his life he modified his framework to allow for the possibility that the threat of labor militancy could induce businesses to reduce their markups. In that case a reduction in the unemployment rate that leads to a higher money wage might also lead to a higher real wage and wage share. From this point of view, Kalecki’s modified framework would be consistent with *three types of Phillips curves* (money wage, real wage, and wage share) whose theoretical and empirical foundations are examined in chapter 14. Like Keynes, Kalecki opposes the orthodox claim that an increase in real wages will reduce profitability and hence raise unemployment. His principle objection can be expressed as the proposition that an increase in real wages will have two opposing effects on the actual rate of profit: it will lower the normal rate of profit but will raise the rate of capacity utilization by increasing workers’ consumption demand. This highlights the key role of capacity utilization as a free variable. In the end, fiscal policy could be used to pump up output and employment while monetary policy could be used to mitigate any upward pressure on the interest rate. Kalecki was nonetheless

pessimistic about the *political* likelihood of maintaining full employment because it would threaten the power of the capitalist class.

The post-Keynesian tradition encompasses Keynesian and Kaleckian wings that share five central beliefs: aggregate demand drives output, money is endogenously created through the banking system, both persistent excess capacity and unemployment are the normal outcomes of market processes, and the state can achieve (effective) full employment with tolerable levels of inflation. Section VI analyzes the works of Paul Davidson, the leading representative of the Keynesian wing; Godley and Taylor representing the Kaleckian-Structuralist wing; and Lavoie representing the post-Keynesian wing. Several general points are identified as being important to the subsequent classical synthesis of real competition and effective demand (chapter 14). The notion that aggregate demand drives production requires that investment be independent of the supply of savings, which as both Keynes and Kalecki belatedly admitted, requires that it be initially financed entirely out of bank credit. The assumption that business savings be a fixed proportion of net income or profits implies that the business savings (retained earnings) are not linked to the needs of investment finance, which is contrary to business practice and empirical evidence. The idea that capacity utilization is a “free variable” even in the long run implies that firms are never able to eliminate genuine excess capacity, which makes no sense at the microeconomic level. Harrod’s own argument that capacity utilization hews to some normal level has been largely ignored by the post-Keynesian tradition, which is quite curious because it represents an important form of stock-flow consistency. Consider the post-Keynesian claim that wage-led and profit-led growth are alternative regimes rather than alternate phases of an adjustment process. A rise in real wages will have a positive impact on worker consumption at existing levels of employment and a negative impact on the normal-capacity profit rate. Even if the former effect outweighs the latter in the short run, as most post-Keynesian authors claim, the reestablishment of a normal rate of capacity utilization will lead to a fall in the actual rate of profit as it returns to the new lower normal rate and hence to a fall in the rate of growth. Then what is gained through a rise in the levels of output and employment is subsequently paid for through a slowdown in their rates of growth (chapter 13). Finally, the belief that persistent involuntary unemployment can be eliminated through appropriate fiscal and monetary policies runs up against the argument in Marx and Goodwin that capitalism generates and maintains a “normal” rate of involuntary unemployment—as opposed to the voluntary recusal from labor which is assumed in the neoclassical “natural” rate of unemployment. Chapter 14 is devoted to the analysis and implications of the normal rate of unemployment. We will see that attempts to maintain unemployment below the normal rate need not trigger inflation, let alone accelerating inflation (chapter 15).

Chapter 13 takes up the task of constructing a classical approach to macroeconomics founded on real competition. The central notion is that the rate of growth of capital is driven by the expected net rate of profit (i.e., by the difference between the expected rate of profit and the interest rate). Keynes’s and Kalecki’s theories of effective demand are founded on the very same proposition (chapter 12, section III). But in the classical tradition the expected rate of profit is itself tied to the actual rate of profit in the manner similar to Soros’s theory of reflexivity, whereas in Keynes’s theory the expected rate of profit is left “hanging in the air” perpetually out of reach of the short run on which he concentrates. Section II focuses on key elements of existing

theories of effective demand, beginning with micro foundations. Keynes's story is famously inconsistent on this issue. He explicitly favors "atomistic competition" over imperfect competition and even invokes the perfectly competitive condition  $p = mc$ . Elsewhere he says that since production takes time, firms must produce on the basis of expected proceeds and entrepreneurs have to try to forecast demand through trial and error. These are contradictory views since perfectly competitive firms are demand-indifferent (chapter 8). Some have suggested that Keynes could have resolved this difficulty by becoming a post-Keynesian. I would argue instead that he rejected imperfect competition because he was basing himself on a notion of competition similar to the classical one. Next consider the multiplier process. Both Keynes and Kalecki belatedly admitted that their claim that investment is independent of savings was predicated on the implicit assumption that any gap between desired investment and existing savings would be funded entirely out of new bank credit (and corresponding new business debt), so that the existing level of savings was not a constraint. Expressing the multiplier as a temporal process brings out two things: that it takes a permanent increase in the level of investment to produce a permanent increase in output; and that the standard multiplier story abstracts from debt payments and therefore implicitly assumes Ponzi finance for new investment. Conversely, as Ohlin long ago noted, allowance for debt payments implies a variable savings rate. The two poles can be encompassed by generalizing the multiplier process to make the savings rate responsive to the finance gap: then the standard multiplier holds if the savings rate is completely unresponsive while a fully responsive savings rate implies a zero multiplier (since new savings fully accommodate new investment). The responsiveness of the savings rate becomes crucial in the construction of a classical alternative in section III.

In the static Keynesian argument investment is a function of the difference between the expected profit rate and the interest rate, so that a given level of net profit rate implies a particular level of investment (chapter 12, section III.2). On the multiplier argument the investment level in turn implies a particular level of equilibrium output. But since investment increases the capital stock, capacity must be increasing. It follows that capacity utilization (the ratio of output to capacity) must be continually falling. The traditional multiplier story is therefore stock-flow inconsistent. One solution is to assume that it is the rate of accumulation ( $g_K \equiv I/K$ ) that responds to net profitability, as in the classical tradition. The trouble is that the resulting capacity utilization rate will generally be different from the normal rate. Only three years after the *GT*, Harrod had already demonstrated that only one "warranted" rate of accumulation is consistent with a normal rate of capacity utilization. So we arrive at a seeming impasse: if expected profitability drives accumulation, as in the classicals and Keynes, the capacity utilization rate will generally differ from the normal level; conversely if accumulation is to be consistent with normal capacity utilization, as in classicals and Harrod, the rate of accumulation must be driven by the savings rate. Section III shows that the real difficulty originates in the unwarranted assumption that business savings is independent of business investment.

Another set of issues arises from dynamic considerations. Investment is driven by expected net profitability which will generally be different from actual net profitability. I argue that the two are connected in the manner envisioned in Soros's theory of reflexivity: in a boom the expected rate rises above the actual rate and in a bust the former falls below the latter, so that the two fluctuate around one another in a

turbulent manner. That is clearly the general presumption in Marx and Keynes. In addition, it is necessary to situate the mutual adjustments of supply and demand in a growth context. This leads to the demonstration that the adjustment of actual capacity to the normal level is perfectly stable: *there is no "knife-edge" for the Harroddian warranted path*. As a subsidiary matter, it is demonstrated that there is no "Sraffian Supermultiplier" in a Harroddian context. Output growth is never demand-led here, and if it is demand-modified, then within the Harroddian framework any increase in the growth of exogenous demand will decrease the overall growth rate. Finally, a constant growth rate implies that the log of the level of the variable in question has a stochastic trend because it follows a unit root process. In the static Keynesian case, a temporary rise in expected net profitability has a temporary effect on the level of investment and hence on output and employment. But in the classical case, the temporary rise in expected net profitability raises the growth rate and permanently raises the level of output and employment. Section III.4 elaborates on the classical implications of this dramatic difference.

Three further points require attention. In the *GT*, Keynes assumes that the money supply is determined by the monetary authorities, and after the *GT*, he admits having assumed that any gap between savings and investment would be entirely funded by bank credit at any given interest rate. But then the money supply must vary directly with the demand for credit, which makes it endogenous. This contradicts the very foundation of his LM construction, because liquidity preference is no longer sufficient to determine the interest rate once the money supply is endogenous. The various post-Keynesian responses to this problem were discussed in chapter 10, along with the classical alternative in which the competitive profit rate equalization determines all interest rates, including even the base rate. One can view the classical argument as an alternate path to Keynes's conclusion that a competitive interest rate is not free to adjust aggregate demand to fit full employment supply (chapter 12, section III). Second, both neoclassicals and Keynesians assume that the price level only increases in the vicinity of full employment. The whole debate about the Phillips curve was about whether or not we could treat observed unemployment as effective full employment (chapter 12, section III.5). The classical approach implies that the growth rate is limited by the profit rate, and this provides an alternate explanation for inflation in a variety of countries (chapter 15). Lastly, Keynes's whole analysis rests on the assumption that appropriate policies can essentially eliminate unemployment. In chapter 15, I will argue that competitive capitalism operating under flexible real wages creates and maintains a certain rate of "normal" involuntary unemployment. As previously noted, this is not the same thing as the neoclassical "natural" rate of (voluntary) unemployment.

Section III begins the development of a classical approach to modern macroeconomics. The first point is that demand and supply are both regulated by profitability: production supply is based on profit, while consumption demand comes from wages, interest, and dividends funded out of profits and investment demand is regulated by expected profits. Classical macroeconomics is neither supply-side nor demand-side: *it is "profit-side."* The second point is that the savings rate is not independent of investment because business savings and business investment are undertaken by the same entity. If savings were to rise in such a way as fully finance any increase in investment, there would be no multiplier. But in general it is sufficient that the business savings

rate, and hence the overall savings rate, reacts in some degree to any gap between investment and current savings. The endogeneity of the savings rate is implicit in the classical tradition, plays a prominent role in the arguments of Godley and Cripps and Ruggles and Ruggles, and has recently been acknowledged within the post-Keynesian tradition by Blecker, Pollin, and others. At an empirical level in the United States, the business savings rate closely tracks investment rate (to which it is roughly equal). Yet theoretical models typically assume a savings rate that is completely independent of the needs for investment finance.

The simplest classical model is one in which the accumulation rate (growth rate of capital) responds to the expected net profit rate (expected profit rate minus the interest rate), and the savings rate responds to the relative finance gap between investment and savings. In the short run, the interest rate is likely to rise when the finance gap is positive, but in the long run the firms supplying the finance will be subject to profit rate equalization and the normal interest rate will end up being regulated by the normal profit rate and the price level (chapter 10, section II). Allowing for some interest rate sensitivity in (say) household savings rates, or for bond or equity issue, makes no fundamental difference to the classical dynamic because the endogeneity of the business savings rate is the key. However, bank credit does provide an internal mechanism through which current expenditures can exceed current incomes: banks can create new purchasing power which can permit investment to expand faster than savings and consumption to expand faster than income. Firms can always spend more than they make by drawing down money balances and extending the chain of credit, but bank credit greatly enhances this process. Similar considerations arise in the case of government deficits and trade surpluses. Insofar as we are concerned with effects on aggregate output and employment, what is important is the amount of credit directed toward expenditures on commodities rather than on financial markets, speculative activities, and in the case of central bank activities, to repairs of private and public sector balance sheets (chapter 15, section V).

The basic classical system embodies a set of reflexive relations between the expected and actual profit rates, demand and supply, output and capacity, and the actual and normal interest rate. The rate of profit is the linchpin of the whole system. In a growing system, the growth rate of nominal output rises when demand exceeds supply, the growth rate of the capital stock rises when output exceeds capacity, and capital flows more rapidly into the financial sector when the actual interest rate exceeds the normal one. This leads to the turbulent equalization of aggregate demand and supply over some short run process, of output and capacity and the actual and normal interest rate over longer runs, and of the actual and expected profit rates over some reflexive run. Notice that this synthesizes the Keynesian notion that demand may be relatively autonomous due to injections of new purchasing power, the classical and Keynesian notion that accumulation is driven by expected net profitability, the classical notion that expected profitability is regulated by normal profitability, and the Harroddian notion that the actual rate of capacity utilization is regulated by the normal rate. Equilibrium of demand and supply and of output and capacity determines particular ratios of savings and investment to output at a normal interest rate. This implies that the levels of savings and investment depend on both the interest rate and the level of output—as in traditional macroeconomic analysis—except here the interest rate is determined by the profit rate and the savings rate is linked to the investment

rate. In addition, because the actual growth rate fluctuates around its equilibrium rate, the (log of the) level of output will have both deterministic and stochastic trends so that the level of output will be path-dependent. Then even a temporary rise in the net profit rate and/or a temporary boost to net purchasing power will permanently raise the level of output and employment. This is the classical equivalent of the Keynesian multiplier. But, of course, the growth trend may also be affected. If the output path rises to a new level, unemployment may fall, real wages may rise, the profit rate may fall and output may grow more slowly. Then while animal spirits and excess demand can raise the level of the output path they can also lower its growth rate (chapters 14–16).

Chapter 14 derives the crucial linkages between unemployment, wages, profitability, and growth. The key conclusion is that competition under flexible real wages creates a sustained rate of involuntary unemployment. This is in sharp contrast to neoclassical theory in which flex-wage competition necessarily leads to full employment, and to Keynesian and post-Keynesian theory in which competition may or may not give rise to unemployment. Goodwin formalized Marx's argument that competition creates a persistent pool (Reserve Army) of unemployed labor and set the stage for modern heterodox approaches. A striking implication of both orthodox and heterodox approaches is that workers have no say in their own standard of living: in neoclassical theory the real wage is determined by the full employment condition; in post-Keynesian theory it is determined by productivity and the monopoly markup set by firms; and in Kaldor and Pasinetti's extension of Harrod, it is determined by productivity and the requirements for full employment. Even in Goodwin's formalization of Marx the real wage is determined by productivity and the requirement for the normal rate of unemployment (section II). But once it is recognized that labor force and productivity growth may themselves respond to accumulation through increases in labor force participation and/or immigration rates and through accelerated technical change (section III), then there is full room for the effects of labor struggles on the real wage and wage share.

Keynes who based himself on competitive markets specifically cites the role of wage bargains and labor struggles in determining money wages. He conceded that prolonged unemployment would erode real wages, but argued that in periods of high unemployment state intervention was preferable to a slow and socially devastating erosion of the living standards of workers. Keynes's views are consistent with the classical theory of real competition. In Kalecki's theory, the labor share in net output is entirely determined by the monopoly markups set by their employers. As previously noted, Kalecki struggled to incorporate some degree of worker agency into his story. These and other views in the post-Keynesian and post-Goodwin traditions are analyzed in considerable detail in the text.

Section III constructs a framework in which labor struggles play a significant role in determining the real wage and normal-capacity accumulation maintains a persistent pool of unemployed labor. Shop-floor conflicts between labor and capital bring about a particular division of the money value-added in each firm. At an aggregate level this translates into a real wage linked to labor productivity through a term reflecting the average bargaining strength of labor. The labor strength term itself rises when unemployment falls below some critical level and falls in the opposite case. This implies that the rate of change of real wages relative to productivity (i.e., the rate of change of the wage share) is a negative function of the unemployment rate. I call this the classical



Curve. It is one of the two basic relations in the Goodwin model and can be shown to imply the aggregate log-linear “wage-curve” estimated empirically by Blanchflower and Oswald and many others. The unemployment rate in turn depends on the levels output, productivity, and the labor force. The rate of growth of output was previously shown in chapter 13 to depend on normal net profitability (which drives accumulation) and a driving term reflecting various factors including private, public, and foreign injections of new purchasing power. Productivity and labor force growth are assumed to respond to unit labor costs (the wage share) because a rise in the latter provides a strong incentive for firms to raise productivity and to increase the labor force by importing workers and/or raising the participation rate. The mutual adjustment between output and productivity growth creates a correlation known as Verdoorn’s Law.

The classical dynamical system yields a growing economy with a normal rate of unemployment and a stable wage share ( $\beta$ ) which reflects the social-historical strength of labor. When the stable wage share is combined with the national income identity that net output per worker ( $y$ ) equals the sum of the real wage and real profit per worker, the result is the relation  $y_t = A_t \cdot k_t^{1-\beta}$ , which looks just like an aggregate Cobb–Douglas production function even though it is explicitly derived from a labor-struggle theory of the real wage. Since growth is endogenous to the classical dynamic even a temporary rise in the growth of aggregate demand arising from state deficits, export booms, or from an acceleration in investment spending due to higher animal spirits, will permanently raise the levels of the growth paths of output, employment, productivity, and the real wage without affecting the wage share, the profit rate, or the rate of growth. On the other hand, persistent demand growth at a rate above that induced by net profitability will lead to a persistently higher wage share (and hence to a lower normal profit rate). Yet the growth rate would also be raised because the negative effect of lowered profitability is offset by a rising stimulus until some limits come into play (see the next section). A striking feature of the classical model is that the long-run wage share depends positively on the initial values of the wage share and unemployment rate, and negatively on the initial values of productivity and labor force growth. Hence, local actions that raise the existing wage share or employment rate will raise the long-term wage share, while local actions that raise productivity or labor force growth will lower the long-term wage share. Workers and employers are therefore justified in thinking that local actions do matter even in the long run. However, none of these will affect the equilibrium employment rate.

Section IV examines the further implications of a normal rate of involuntary unemployment. Pumping up aggregate demand can increase employment and output growth, but will not permanently eliminate unemployment because there are internal mechanisms that restore the normal rate. Therefore, it would take an increasing stimulus to maintain an unemployment rate below the normal rate. Even so, inflation is not an automatic outcome (chapter 15). On the other hand, the normal rate of unemployment can itself be lowered if the balance of power shifts against labor. Section V takes up the relation of the classical curve to various types of Phillips curves. Phillips’s original question was about the effect of unemployment on wages. His own answer was posed in terms of the rate of change of money wages, very much in keeping with a Keynesian money-wage perspective. Friedman and Phelps argued that workers’ struggle for a standard of living (i.e., for a real wage, not a money wage), so that the correct “Phillips-type” relation should be in terms of (expected) real wages. The classical

argument is that real wage struggle is conducted in relation to the general level of development (productivity), in which case the appropriate “Phillips-type” relation will be in terms of the rate of change of nominal wages relative to inflation and productivity growth. Given a stable classical curve, a stable real wage exists only if productivity growth is roughly constant, and a money wage curve exists only if inflation is also roughly constant.

Section VI presents the empirical evidence. As expected, from 1948 to 2011 the US wage share rose and fell broadly in line the growth rate of nominal output (a proxy for aggregate demand). The unemployment rate roughly doubled over this interval, and the unemployment duration quadrupled. The unemployment intensity, which is their product and as such a much better measure of the pressure on wage changes, rose to ten times its original value. As in the theoretical classical system, the actual wage share and the unemployment intensity trace out a clockwise three-dimensional spiral over time. Most importantly, a scatter diagram of the rate of change of the wage share versus unemployment intensity clearly displays a negative slope. Phillips’s original curve was based on cyclically adjusted data points in order to identify the underlying structural relations. I use the Hodrick–Prescott (HP) filter for the same purpose. The dramatic result is a stable classical curve from 1949 to 1982 which then shifts down in the face of subsequent neoliberal attacks on unionization and labor-support mechanisms. The new lower curve in turn reduces the critical unemployment intensity at which the wage share is stable. Also clearly visible are movements up the curve as the economy is pumped up during the Vietnam and dot.com booms and down the curve as the stimuli peter out. All such movements are fairly slow, as Keynes argued long ago. Finally, the lack of stable real or money wage Phillips curves is easily explained by the fact that inflation rose and productivity growth fell dramatically precisely in the era when Keynesian economics had to retreat from the price Phillips curve. Had Phillips answered his own question in classical rather than Keynesian terms (i.e., through a wage-share relation rather than a nominal-wage one), it might have been possible to avoid the theoretical crisis about the price “Phillips curve” during the Stagflation era of the 1970s and 1980s. Keynesian theory would still have required an explanation of inflation, and even if it had retained a markup theory of inflation based on nominal wage, the shifts in the underlying nominal wage curve would have been entirely comprehensible. Of course, the political attack aimed at weakening labor and raising the profit share might well have won the day in any case.

Chapter 15 tackles the theory of inflation under modern fiat money. It opens with a reminder that the historical path from private money to state money is long and torturous. The state did not invent money, coins, payment obligations, or debts. Once money has been established the state is impelled to expand its base beyond compulsory payments in labor and in kind to payments in money. Governments have typically imposed poll taxes, property taxes, and taxes on commodities, import, exports, tolls, and harbors, and more recently, on income. In addition, they have resorted to sales of public lands, the ransom of prisoners, and seizures of foreign ships, goods, and treasuries. At some late stage in history the state monopolizes the creation of coins and tokens. This is merely a takeover of a previously private function, and private banks continue to create the vast bulk of the medium of circulation and medium of payment. The state also comes to exercise some degree of control over banks—a control whose intrinsic limitations are periodically exposed during recurrent financial crises.

The general global crisis of the early twenty-first century is a stinging refutation of textbook fantasies of the Left and the Right, in which a wise and benevolent state supposedly controls money and finance for the common good. Fiat money, forced inconvertible token money, is the characteristic form of modern money. The history of money reminds us that private circulation gives rise to money tokens which are accepted as long as they are deemed able to perform certain functions as money and people accept inconvertible scrip for the same reason that they accept convertible scrip: because they believe that they can continue using them as money. While legal tender laws may be useful in establishing a currency, and legal restrictions on foreign currency and gold holding may impede recourse to alternatives, they cannot prevent private agents from seeking more secure monetary forms (chapter 5).

Section II provides a detailed survey of chartalist and neo-chartalist claims about money, beginning with the claims of Innes and Knapp who attribute great powers to the state and an extraordinary passivity to private agents. Keynes explicitly lauds Knapp for defining “State-Money” as anything which is accepted by the state, which means that gold coins, convertible tokens, and fiat money became State-Money when the state accepted them. This is perfectly consistent with the private invention and reinvention of monies to which the state periodically accedes. Unlike Knapp, Keynes only claims that the state invented fiat money. Neoclassical economics typically present money as a creature of the market and the state as an excrescence while Keynesian and post-Keynesians typically criticize the market and defend the state. Neo-Chartalists such as Goodhart, Wray, and Bell fall in the latter camp, and their views along with those of critics such as Merhling and Rochon are examined in some detail. No one disputes that modern fiat money can be created to any degree. So if one strips away the Chartalist claims about the origins of money and the passivity of money holders, their central argument becomes that under modern fiat money regimes government deficits in service of social programs need not cause inflation or raise interest rates.

Section III focuses on the effects and limits of fiat money. It frees the state from its direct budget constraint. It successfully fueled the American, French, Chinese, and other revolutions. And it has led to hyperinflation at various times in history (section VIII.4). As a result, the Treasuries of most advanced countries face legal prohibitions against directly creating money to finance deficits. The Treasury can only spend money available in its account which is replenished through the tax inflow, some part of which is a reflux, and through borrowing from (selling bonds to) the domestic public or to foreigners. But the modern central bank can create any mandated sum at the stroke of a key and transfer it to the Treasury by buying the latter’s newly issued bonds. Then the only restraint on this process would appear to be from the resistance of the central bankers and from a benighted view about the piling up of the government debt to itself—were it not for the possible effects on prices and interest rates. This is where the core neo-Chartalist propositions come back into the picture. As Keynesians, they believe that involuntary unemployment can be eliminated by deficit spending (as opposed to the classical view in chapter 14 that it cannot), and as post-Keynesians, they believe that the exchange rate can be set by the state at any desired level and that the price level is determined by monopoly markups ultimately resting on the money wage. On this basis, they propose a government (Employer of Last Resort, ELR) program to employ *at some fixed money wage* any labor that the private

sector is unable to absorb. The base wage rate would then provide a stable anchor for all other wages and, through stable markups, also all prices. Undesired effects of international interest rates on domestic ones could be negated through appropriate manipulation of the exchange rate. Undesired domestic income and interest rate effects could be avoided by having the state raise taxes in order to rein private spending and sell bonds to the public or foreigners so as to reduce the money supply. The neo-Chartalist core argument rests on several crucial propositions none of which obtain in the classical argument: (1) that unemployment can indeed be held at any desired level (chapter 14); (2) that the private sector money wage is determined by the base ELR wage, rather than through the ongoing struggles between workers and their bosses (chapters 4, 14); (3) that the price level is determined by the money wage in the private sector because of monopoly markup pricing (chapter 12, sections V–VI); (4) that the state can maintain the whole spectrum of bond rates at desired levels by fixing the base rate (chapter 10); and (5) that it can fix the nominal exchange rate at any desired level (chapter 11). The issue in each case is not about whether the state can carry out the prescribed acts but rather about their possible consequences, of which inflation is one.

Section IV constructs a classical theory of inflation. Competition only establishes relative prices through the equalization of profit rates. Under pure fiat money the price level is determined by aggregate demand and supply rather than the relative price of some money commodity. The growth in aggregate demand is fueled by new purchasing power (chapter 13, section III.3) and a modern credit system based on fiat money can fuel virtually unlimited growth in aggregate demand (chapter 5, section II.4). Then the limits to the growth of supply become crucial. It has already been established that the supply of labor cannot play this function because the system reverts to a persistent rate of unemployment (chapter 14, sections III–IV). The limit arises instead from the fact that no economy can sustain a rate of accumulation greater than that determined by the full reinvestment of the economic surplus (i.e., greater than the rate of profit). This is implicit in Ricardo's corn-corn model and Marx's Schemes of Expanding Reproduction and is explicit in Kaldor and von Neumann. The degree to which the actual rate of accumulation approaches its limiting value can therefore be viewed as a measure of the degree to which the maximum growth potential of the economy is being utilized—a "*growth-utilization*" index. The basic model is therefore one of demand-pull from newly created purchasing power and supply resistance from a tightening growth-utilization index. Since the profit rate is the ratio of profit to capital and the rate of accumulation is the ratio of investment to capital, the growth-utilization index is simply the share of investment in profit. The section ends with a discussion of the appropriate measurement of real average and incremental rates of profit which play a key role in the empirical analysis of section VIII.

Section V analyzes the demand-pull side. It was established in chapter 12 that aggregate excess demand in the commodity market can be expressed as three sectoral balances:  $ED = (I - S) + (G - T) + (EX - IM)$ . Once we consolidate inter-sectoral balances this leaves the portion of net new domestic credit from private and central banks and private businesses which goes into the purchases on new goods and services (as opposed into purchases of financial assets and existing homes, valuable objects, etc.), plus the current account balance (CA) of the trade sector and any part of net borrowing from abroad that fuels domestic commodity purchases. Over the interval

in which demand and supply roughly balance, an increase in commodity purchasing power will manifest itself in additional production and/or price increases, that is, in an increase in nominal gross output (defined in the sense of Leontief). Then the growth rate of nominal GDP will be some function of new purchasing power relative to GDP. This is consistent with both monetarist and Keynesian approaches. Section VI develops the supply-resistance side of the argument. The key point is that the response of real output growth becomes increasingly muted as the actual growth rate approaches the maximum growth rate (the profit rate). This is similar to Keynes's notion that as full employment is approached, less of new demand is absorbed by new output and more by price increases. Marx makes a similar point in a growth context and Pasinetti provides a formal analysis of the increasing prevalence of bottlenecks as the actual rate of growth approaches the theoretical maximum rate. The growth-utilization index is the strain-gauge of growth.

Section VI combines the demand-pull and supply-side arguments into a classical theory of inflation. The classical argument implies that real output growth responds positively to net purchasing power and net profitability as measured by the net real incremental rate of profit (chapter 13, section III) and responds negatively to the degree of growth utilization at least when the latter rises above some critical level. It seems likely that the interactions will be nonlinear. The unutilized growth-utilization potential plays the same role in classical inflation theory as the unemployment rate does in standard inflation theory. Then since the rate of inflation is equal to the difference between the rate of growths of nominal and real output, and since the former is a function of new relative purchasing power, we can say that inflation responds positively to new relative purchasing power, negatively to net profitability, and negatively to unutilized growth-utilization potential. When new purchasing power is growing sufficiently to offset the negative impact of falling profitability, we would have a Phillips-type inflation curve in terms of unutilized growth potential. From this point of view, we could view net new purchasing power and net profitability as shift factors of this basic curve. It is particularly important to note that since growth depends on net profitability and new purchasing power, it is possible that a fall in the former can be mitigated by a rise in the latter so that the growth rate would fall less than the profit rate and their ratio, the growth-utilization rate, would rise. The fall in the growth rate would increase the unemployment rate while rise in the growth-utilization rate would make the economy more inflation-prone. This is the secret of the dread "stagflation" that led to the overthrow of Keynesian theory (chapter 12, sections III–IV). The net rate of profit and the growth-utilization rate can only vary within certain limits, but there is no such constraint on new purchasing power in a fiat money system. Hence, when the rate of creation of new purchasing power is relatively low, one would not expect any direct relation between it and inflation because the other factors would be decisive. But as newly created purchasing power gets larger and larger, one would expect such a relation to emerge, and at very high rates one would expect the rate of inflation to be roughly equal to the rate of new purchasing power. This is similar to the theoretically expected nonlinear relation between a country's relative inflation rate and its nominal exchange previously derived in chapter 11, section VI. Finally, insofar as the net profit and the growth-utilization rates are positively correlated, it would be possible to treat the latter as a proxy for the former, which leads to a more restricted hypothesis in which inflation is a function of the growth-utilization rate in which the overall effect of

the latter is ambiguous because growth-utilization and net profit rates have opposite influences on inflation.

Section VIII considers the empirical evidence, starting with the United States. The strong graphical and statistical relation between nominal GDP growth and new relative purchasing power is consistent with the classical hypothesis that the former is a function of the latter. The second key hypothesis is that the growth rate of real output responds to purchasing power, net profitability, and the growth-utilization rate. The appropriate measure of net profitability is the real net rate of return on new investment as proxied by the real net incremental rate of profit developed in chapter 6, section VII. Real output growth is strongly positively correlated with this real net return on net investment. The two preceding hypotheses imply that the rate of inflation is a function of relative new credit, net profitability, and the degree of unutilized growth capacity, the latter taking the place of the unemployment rate in conventional theory. Scatter plots of the inflation rate versus unutilized growth potential are compared to standard ones using the unemployment rate instead, for the whole postwar period 1951–2010 and for sub-periods 1951–1981 and 1982–2010. The differences are striking. In every case, the classical inflation “Phillips” curve displays a clear downward slope, whereas the conventional curve does not (as we already know from chapter 12, section III.5). Given that the net profit rate and new relative purchasing power act as shift factors in the classical inflation curve, the observed differences in the patterns exhibited in two sub-periods can be explained by the changes in the levels of those two variables.

Handfas tests my inflation hypothesis on seven OECD countries (Canada, France, Germany, Japan, South Korea, the United Kingdom, and the United States) and three developing ones (Brazil, Mexico, and South Africa), the latter being tentative because of small sample sizes. On the assumption that the net profit and the growth-utilization rates are positively correlated and that the latter is likely to have an inhibiting effect only when it reaches as sufficiently high level, he posits that there will exist a nonlinear long-run relation between inflation and net purchasing power and the growth-utilization rate. He tests this using an error-correction representation of an autoregressive distributed lag (ARDL) model from which he can estimate the long-run coefficients. In all OECD countries, the long-run relations are significant and have relatively good fits, but less so in Brazil and South Africa and are not satisfactory in Mexico. A striking result is that in all countries the coefficients of the nonlinear function of the growth-utilization rate have the expected signs suggesting a U-shaped functional form with a negative region for some values of the rate. The average rate of the United States puts it in the positive (inflationary region) of its estimated curve, but the rate in Japan falls with the negative (deflationary) region of its curve. As previously noted in the summary of section VI, the classical argument also implies that a direct relation between inflation and new purchasing power will only emerge when the latter is high. A 1988 study by Harberger covering twenty-nine countries over 1972–1988 exhibits exactly this property, as does an extended sample produced by Ramamurthy covering forty-six countries over 1988–2011. Argentina in 1982–1984 appears at the high end of Harberger’s sample with an average inflation rate of 255% and an average growth of total credit of 312%, but even this is modest compared to Argentina in 1989 when inflation was 5,380%. Despite the absence of current account data, one can see extremely strong relations between total (public and private) credit growth and nominal GDP growth, inflation, and currency depreciation. At their peaks, nominal GDP

and the price level grow substantially less than total credit, which could be accounted for by purchasing power going into asset price inflation and into currency flight—both well-known phenomena in such circumstances. In addition, at the peak the exchange rate depreciates even faster than prices increase—as expected by the combination the equilibrium classical effect of inflation on exchange rates (chapter 11, section VI, and table 11.4) and currency flight.

Section XI concludes the chapter by comparing the classical hypothesis to the Non-Accelerating-Inflation-Rate-of-Unemployment (NAIRU) hypothesis which dominates modern discussions of inflation. The classical proposition can be expressed as the hypothesis that the level of inflation is a positive function of the extent to which the unutilized growth capacity falls below some critical rate, subject to shift factors stemming from net profitability and new purchasing power. The simplest form of the NAIRU hypothesis is that the change in inflation (the acceleration of the price level) is a positive function of the extent to which the unemployment rate is below the “natural rate of unemployment.” Both hypotheses link inflation to departures from the critical values of their respective driving variables. In addition, both expect the system to return to some normal level of unemployment. However, in the classical case, this is a rate of involuntary unemployment not directly related to the inflation rate (chapter 14), whereas in the NAIRU it is in effect a full employment rate. From a classical perspective, it is possible to lower the normal rate of unemployment by reducing wages relative to productivity, either through neoliberal attacks that seek to lower the growth rate of real wages by weakening labor or through “Swedish” policies that stimulate productivity growth in excess of real wage growth (chapter 14, section VII). Furthermore, the critical growth-utilization rate is not an equilibrium rate because there is no presumption that the economy sticks at this rate, whereas under the NAIRU hypothesis the natural rate of unemployment is exactly the rate to which the economy returns in the absence of sustained efforts to prevent that. In the classical case, inflation can be zero as long as the growth-utilization rate and the rate of creation of new purchasing power are not too high. Inflation can even be negative (i.e., there can be deflation) under appropriate circumstances. In the classical case, the inflation rate is determinate but the corresponding price level will be path-dependent, while in the NAIRU case the rate of change of inflation is zero at the natural rate of unemployment, but the particular value of inflation will be path-dependent—precisely the basis for the policy conclusion that unemployment must be maintained above the natural rate of inflation rate for some time so that inflation can be “wrung out.” In the NAIRU argument, hyperinflation comes about from persistent attempts by the state to maintain unemployment below the natural rate because this sets up an unstable expectational spiral. In the classical case, the proximate causes of inflation are an increase in the growth rate relative to the profit rate and/or an increase in the creation of new purchasing power, with hyperinflation arising only if the state takes the latter to extremes. Lastly, the classical theory of inflation is rooted in the operation of real competition, whereas the NAIRU hypothesis, like much of modern macroeconomics on both neoclassical and post-Keynesian sides, is typically based on imperfect competition.

Chapter 16 provides a classical reading of the economic crisis that swept across the world in 2007. This is the first Great Depression of the twenty-first century, and like its predecessors, its first manifestation was a financial collapse—in this case, of the

subprime mortgage sector in the United States. But that was not its cause. Recurrent crises are an absolutely normal part of capitalist history as long booms give way to long downturns and the health of the economy goes from good to bad. In the latter phase, a shock can trigger a crisis, as was the case in the 1820s, 1870s, 1930s, and 1970s. Those who choose to see each such episode as a singular event conveniently forget that it is the very logic of profit which drives the system to repeat these patterns. I have argued throughout this book that capitalist processes are inherently turbulent with powerful built-in rhythms modulated by conjunctural factors and affected by specific historical events. Capitalist accumulation is no different. Business cycles are the most visible elements of its intrinsic dynamics, including a fast (three- to five-year) inventory cycle, a medium term (seven- to ten-year) fixed capital and possibly longer structures cycles. Underlying all of these is a still slower rhythm consisting of alternating long phases of accelerating and decelerating accumulation. Capitalist history is played out on a moving stage.

The Great Depression of the 1930s had very high unemployment and falling prices, while the Stagflation Crisis of the 1970s had half the unemployment rate but high inflation. The difference is both a tribute to Keynesian policy and a warning about its limitations (chapter 12). A new boom began in the 1980s in all major capitalist countries, greatly enhanced by a sharp drop in interest rates which raised the net rate of return on capital (i.e., raised the net difference between the profit rate and the interest rate). Falling interest rates also lubricated the spread of capital across the globe, promoted a huge rise in consumer debt, and fueled international bubbles in finance and real estate. Deregulation of financial activities in many countries was eagerly sought by financial businesses themselves, and except for a few countries such as Canada, this effort was largely successful. At the same time, in countries like the United States and the United Kingdom, there was an unprecedented attack on labor which led to a slowdown of real wages relative to productivity. The drop in interest rates and in relative real wages greatly boosted to the net rate of profit. The normal side effect to a wage deceleration would have been a stagnation of real consumer spending. But with interest rates falling and credit being made ever easier, consumer and other spending continued to rise, buoyed on a rising tide of debt. And then it all came crashing down, triggered by the mortgage crisis in the United States. The crisis is still unfolding. Massive amounts of money have been created in all major advanced countries and funneled into the business sector to shore up its assets. But unemployment intensity is still high (chapter 14). It is striking that so little has been done to expand employment through government-created work, as was done through public works and/or war preparations during the 1930s. The fundamental question is: How can a system whose institutions, regulations, and political structures have changed so significantly over the course of its evolution still exhibit recurrent economic patterns? The answer lies in the fact that the profit motive always remains the central regulator of the system because both supply and demand are ultimately rooted in profitability (chapter 13). In what follows I will focus largely on the United States as the hegemonic power of the capitalist world. Of course, the real toll is global, falling most of all on large numbers of already suffering women, children, and unemployed of this world.

Figure 16.1 displays Kondratieff long waves from 1790 to 2010 in the United States and United Kingdom that are clearly visible when one expresses the price level in each country in terms of its gold equivalent (chapter 5, figures 5.5–5.6) and we see that



general crises typically begin roughly in the middle of long downturns. *The Great Depression of 2007 was quite on schedule*. Orthodox economics typically insists that each crisis is unique and will not be repeated because the problem has been resolved. Ricardo, Fisher, Samuelson, and Bernanke are some of the names associated with such proclamations. And, of course, the economic orthodoxy continues to exalt the virtues of the market and downplay (or even ignore) the current crisis. Figure 16.2 shows that the normal maximum rate of profit falls steadily throughout the postwar period, that is, *technical change is consistently capital-biased* (chapter 7, section VII). The normal profit share is stable in US labor's "golden age" from 1947 to 1968, falls during the Stagflation Crisis of 1969 to 1982, rises considerably during the neoliberal era starting in the 1980s and then retains its high level during the Global Crisis that begins in 2007. This is consistent with the previous finding in chapter 14, figure 14.14, of a downward shift in the wage share Phillips curve and the continued downward movement along this new curve. The combination of a continuously falling wage share and fiscal deficits dramatically raises the profit share even during the crisis. The normal profit rate being the product of the normal profit share and the normal maximum rate of profit, it falls faster during the Stagflation Crisis but then stabilizes during the neoliberal era right up to the current crisis. In effect, technical change steadily erodes the level of the normal profit rate in all three periods but in the neoliberal era an induced decrease in wage share is able to offset the steady fall in the normal maximum profit rate. Actual profit measures are evidently subject to many fluctuations, such as the big run-up during the 1960s in reflection of the deficit financed escalation of the Vietnam War. However, over the long term, structural factors predominate. The net average and incremental rates of profit are combinations of profit and interest rate paths. We can see that the Stagflation Crisis of the late 1960s was precipitated as both net rates sank to historic lows, after which the whole behavior of the system changed: growth slowed, bankruptcies and business failures soared, unemployment rose sharply, real wages fell relative to productivity, and the stock market fell by over 56% in real terms—as it did in the worst part of the Great Depression. In Keynesian response, the federal budget deficit rose fortyfold and inflation shot up but so did the unemployment rate and intensity (chapter 14, section VI). The historical solution to the Stagflation Crisis was a reduction in the wage share and a great reduction in the interest rate, both of which worked their magic on net profit rates. This is the real secret of the great boom that began in the 1980s. The trouble was that the induced boom was inherently contradictory. Cheapening finance set off a spree of borrowing and sectoral debt burdens grew dramatically. Households compensated for their slowing wage incomes by taking on more debt, so consumer spending was maintained until the subsequent collapse of the subprime mortgage sector in 2007 triggered a general crisis that spread rapidly across an already fragile global economy.

Section II examines the general consequences of the crisis. Given the bent of orthodox economic theory, it is not surprising that the crisis shocked most academic economists and central bank officials. The US Federal kept banks, big businesses, and financial markets afloat by flooding the markets with money and US financial firms have essentially returned to their old ways. Norway and Canada were more circumspect in their treatment of financial markets and have therefore avoided many difficulties despite having to suffer the impact of a contraction in world exports. Iceland was hit very hard by the global financial crisis when the three largest banks and

the currency collapsed, bringing down the whole economy. But it sharply devalued its currency to try to make itself more competitive (which reduced real wages sharply) and let its banks default to make foreign creditors absorb large losses, thereby faring comparatively well. By contrast, the Irish government stepped in to protect its banks, shifted their debt to the state and then imposed its repayment burden on the population through job and wage cuts. Unemployment and poverty rose sharply. Unlike Iceland, Ireland was already in the eurozone, so it was blocked from undertaking currency devaluation. Greece, Spain, and Cyprus experienced equally severe economic problems, and Britain is now in a slump more severe than the Great Depression of the 1930s. India and China shot into view in the first decade of the twenty-first century with extremely high growth rates but are now experiencing inflation, real estate bubbles, and slowed growth. Cheap finance became a way to expand employment and pump up financial markets in the neoliberal era, but the crisis has severely undermined that tactic. It is estimated that there are now almost 200 million people in the world without jobs. Youth unemployment is particularly high, comprising almost 74 million young people at an unemployment rate that stood at 12.6% in 2014 and is expected to increase. These are official unemployment rates, which greatly understate the true state of affairs, since they do not properly account for part-time employment and the discouraged. On a global scale almost 900 million workers live in dire poverty.

Section III considers the policy debate on austerity versus stimulus. While governments all over the world have scrambled to save failing banks and businesses, they have been far less concerned with expanding employment. At the heart of this is a debate between those who push austerity in order to make workers more docile and labor markets more “competitive,” and those who push for measures to increase employment and maintain wages. We know from history and from theory (chapter 13) that increased government spending can stimulate an economy for a considerable length of time. This was evident in the Great Depression of the 1930s in which the Work Projects Administration (WPA) in the United States employed millions of people while in Germany Hitler’s large rearmament program quickly attained full unemployment. In times of war, these activities are often accompanied by massive deficit financing. In World War II, from 1943 to 1945, the US budget deficits averaged 25% of GDP, whereas its level in 2014 was under 3%. War is only one form of social mobilization and there is no practical reason why the same mode could not be employed during a crisis. In either case, it becomes necessary to subordinate the profit motive to the perceived social good which is, of course, politically far easier with a war as cover. Normal times are different, because then stimulus operations are limited by the return of capacity utilization to normal levels and by the inverse relation between the wage share and the profit rate. Section IV returns to the central proposition that theory is crucial to economic analysis and policy. Orthodox economics starts from perfect competition, Say’s Law, and full employment and then arrives at effects that mimic some aspects of reality by “throwing a bucketful of grits” into the machinery of perfect competition. Post-Keynesian economics starts directly from imperfect competition in order to build its macroeconomic theory and policy. I argue throughout this book that the theory of real competition is the appropriate theory of competition and also the appropriate ground for Keynes’s own theory of effective demand. In both demand and supply, profitability plays the dominant role.

The final chapter of this book summarizes its structure and addresses further implications. The purpose of the book is to demonstrate that the central propositions of economic analysis can be derived without any reference to hyper-rationality, optimization, perfect competition, perfect information, representative agents, or so-called rational expectations. These include the laws of demand and supply, the determination of wage and profit rates, technological change, relative prices, interest rates, bond and equity prices, exchange rates, terms and balance of trade, growth, unemployment, inflation, and long booms culminating in recurrent general crises. In every case, the theory developed in the book is applied to modern empirical patterns and compared with neoclassical, Keynesian, and post-Keynesian approaches to the same issues. Economic thought is assessed in the light of economic laws of the object of investigation, which is capitalism itself. I argue that this is the essence of the classical, Keynesian, and Kaleckian approaches.

A central finding is that lawful patterns can emerge from the interaction of heterogeneous units (individuals or firms) operating under shifting strategies and conflicting expectations because aggregate outcomes are “robustly indifferent” to microeconomic details. Hyper-rationality is not necessary since one can derive observed patterns without it, nor useful because it does not capture the underlying motivations. The classical approach is grounded in the observation of actual patterns and outcomes. The neoclassical tradition is grounded in their idealization. Abstraction plays a different role in each: abstraction-as-typification in the first, abstraction-as-idealization in the second. In the former, the goal is to get back to actual patterns by successively introducing more concrete factors. All Newtonian masses fall at the same rate in a vacuum, but in a fluid such as air they fall at different rates depending on their shapes, masses, and material compositions. The introduction of these influences is a necessary scientific step toward the concrete. The “ideal vacuum” is in no sense a desired state, at least for living beings.

The chapter goes on to consider various important patterns that could be further investigated. General crises, including the present global crisis that broke out in 2007, are shown to occur in the downturn phases of successive long waves. The further task is to link profit-driven accumulation to recurrent long wave patterns. Turbulent equalization of prices and profit rates in the face of ongoing technical change creates persistent distributions for each variable. The analysis of wage rates follows a similar logic, with the additional elements that labor is an active subject in the division of value added, and wages will differ by occupations if even they are equalized within each. These considerations lead us to consider the shapes and forms of wage distributions. The econophysics “two-class” theory of income distribution (EPTC) shows that labor incomes tend to follow an exponential probability distribution (which has a Gini coefficient = 0.50) and property incomes follow a power law (Pareto) distribution. I demonstrate that the framework developed in chapter 14 in order to analyze the aggregate relation between wages and value can be extended to account for differences between firms arising by competition and occupational differences. This is used to show how and why exponential or near-exponential distributions of labor incomes can arise. At the same time, the overall degree of inequality as measured by the Gini coefficient can be shown to depend solely on the ratio of property income to labor income and the degree of financialization of income flows. This implies that the dramatic rise in the ratio of profits to wages beginning in the 1980s (chapters 14 and 16)

can be viewed as the material basis of the corresponding sharp rise in observed overall income inequality.

The state adds another dimension to the analysis of income distribution: it can intervene directly in the balance of power between capital and labor as in the neoliberal era (chapter 14) and affect growth and employment through fiscal and monetary policy. Both interventions can change the distribution of income by altering the absolute and relative levels of profits and wages. It can also levy taxes and transfers to change the post-tax distribution of income. But then one must also account for the effects of social expenditures on health, education, and general welfare. A surprising finding is that net social wage, which is the difference between taxes and social expenditure, is quite small across major countries, averaging only 1.8% of GDP and 2.2% of Employee Compensation. The market wage is the central determinant of labor's overall standard of living and even the best welfare states largely serve to redistribute this.

This leads to a consideration of Piketty's influential bestseller *Capital in the Twenty-First Century*, which is a welcome return to the tradition of grounding economic analysis in actual patterns. His central claim is that capitalism has a tendency toward increasing inequality only occasionally interrupted by great shocks such as World Wars, Revolutions, and Depressions, because the rate of profit tends to exceed the rate of growth ( $r > g$ ) so that those who live off income from wealth are able to accumulate faster than wage and salary earners. His theoretical explanation relies on orthodox economic theory, including the notion of an aggregate production function and its generic properties. On empirical side, I note that the previously discussed EPTC approach can explain the overall degree of inequality solely through the ratio of property income to labor income which is itself grounded in the division of value added into wages and profits, and the degree of financialization of resulting income flows. On the theoretical side, in the classical argument the wage share is determined by the degree of unemployment and the balance of power between labor and capital to the profit share (chapter 14); the capital–capacity ratio is determined by the choice of technique arising from the cost-cutting imperative imposed on individual firms by competition (chapter 7, section VII); and the rate of profit is jointly determined by the two. Aggregate production functions and pseudo-marginal products, insofar as they appear to exist, are mere statistical artifacts (chapter 3, section II.2). Moreover, the normal rate of profit is always greater than the normal rate of growth, since the former is the ratio of the surplus to the capital stock and the latter the ratio of the reinvested portion of the surplus (investment) to the capital stock (chapter 15, sections IV, VI). Lastly, I argue that Piketty's own measure of the rate of profit is completely inconsistent: the capital stock used as its denominator includes not only plant and equipment but also land, residential real estate, and net financial assets, while the profit measure in the numerator excludes rents, interest, capital gains, and other items that make up the return on the secondary assets. This is why his rate of profit rises in the Great Depression and falls in the booms of the latter half of the twentieth century, which is a most contrary finding.

On an international scale, one must account for the enhanced influence of concrete factors such transportation costs, taxes and tariffs, and the far greater role of history, culture, and national restrictions in channeling the mobility of labor. The economic orthodoxy offers visions of perfect competition and ideal macroeconomic outcomes to justify a greater reliance on markets, increased “flexibility” in labor markets created

by increasing the powers of employers, greater privatization of state enterprises so that their assets and employees will be available to foreign and domestic capital, and the opening up of domestic markets to foreign capital and foreign goods. The heterodox tradition generally argues against these measures on the grounds that competition no longer prevails. I argue that the patterns we find on a global scale are expected from the theory of real competition: competitive advantage goes to those nations whose costs are lower either because they have been able to block or destroy lower cost rivals, or because they have benefitted from some historically achieved combination of state intervention and natural advantages. None of this would be necessary without the competitive pressure emanating from the gravitational field of global competition. Failure to understand the concrete manifestations of these capitalist universals can lead to serious misunderstandings of the development process.

The second major divide in the development literature is between orthodox and heterodox theories of macroeconomics. Faced with the absurdities of full-employment rational-expectations models, it seems sensible to turn to monopoly-markup models of demand-constrained unemployment. In post-Keynesian theory, firms are insulated from competition and individual demand pressures can create the profits they desire through an appropriate markup. The aggregate corollary is that appropriate fiscal and monetary policies can enable the state to create something close to full employment. Yet we have seen that even in the advanced countries such policies failed (chapter 12). The classical argument is that competition creates and maintains a “normal” pool of unemployed workers, so that efforts to pump up the economy in order to eliminate unemployment will not succeed unless they are accompanied by policies that raise productivity faster than the real wage so as to offset any negative effects on profitability, that is, *unless they prevent real unit labor costs from rising* (chapter 14). The criterion for international competitiveness is the same, except that here unit labor costs must generally be reduced fast enough to stay ahead of international competitors—precisely as past and present successful development has demonstrated. In the end, capitalism remains constrained by the laws of real competition on which it rests.