1 Recognizing the Limits of Economists' Knowledge

> I prefer to use the term "theory" in a very narrow sense, to refer to an explicit dynamic system, something that can be put on a computer and *run*. This is what I mean by the "mechanics" of economic development—the construction of a mechanical, artificial world, populated by the interacting robots that economics typically studies.
>
> Robert E. Lucas, Jr.,
> *Lectures on Economic Growth,* p. 21

> Policymakers often have to act, or choose not to act, even though we may not fully understand the full range of possible outcomes, let alone each possible outcome's likelihood. As a result, . . . policymakers have needed to reach to broader, though less mathematically precise, hypotheses about how the world works.
>
> Alan Greenspan,
> "Risk and Uncertainty in Monetary Policy,"
> *American Economic Review,* p. 38

1.1. The Overreach of Contemporary Economics

On the occasion of his 1974 Nobel lecture, Friedrich Hayek appealed to fellow economists to resist the "pretence of exact knowledge" in economic analysis. Drawing on his prescient analysis of the inevitable failure of central planning, Hayek warned against the lure of predetermination: no economist's model would ever render *fully* intelligible the causes of market outcomes or the consequences of government policies. Decades later, experience as a Federal Reserve chief led Alan Greenspan to concur with Hayek. He told the economists assembled at a 2004 meeting of the American Economic Association that central banking requires creativity. Central bankers, just as all individuals, act in a world of imperfect knowledge; hence, they can comprehend neither "the full range of possible [market] outcomes" nor their likelihoods.

In contrast to these skeptical views, contemporary economists have been much less circumspect about the ability of economic analysis to uncover the causal mechanism that underpins market outcomes. In fact, over the past three decades, economists have come to believe that, to be worthy of scientific status, economic models should generate "sharp" predictions

that account for the full range of possible market outcomes and their likelihoods.[1] To construct such models, which we refer to as *fully predetermined,* contemporary economists must fully prespecify how market participants alter their decisions and how resulting aggregate outcomes unfold over time. By design, contemporary models rule out the importance of individual creativity in coping with inherently imperfect knowledge and unforeseen changes in the social context.

In modeling individual decision making and market outcomes, economists make use of a variety of assumptions and insights. The vast majority appeal to a set of a priori assumptions that putatively characterize how "rational" individuals make decisions. In contrast to these *conventional economists,* the increasingly influential *behavioral economists* appeal to empirical observations of how individuals "actually" behave. However different the conventional and behavioral approaches may appear, they share one key feature: both instruct economists to search for fully predetermined models of the causal mechanism that underpins change. Because of this common feature, we regard the conventional and behavioral approaches as branches of the *contemporary approach.*

Economists fully predetermine their models by first representing individual decision making in terms of causal variables, although they sometimes leave the particular set of causal variables unspecified. They also usually specify a set of qualitative conditions that restrict how the causal variables enter their representations of individual behavior at an arbitrary "initial" point in time.[2] While their representations at the initial point in time are qualitative, the insistence on sharp predictions of change leads economists to impose restrictions that relate exactly the properties of their representation at all points in time, past and future, to the properties of the representation at the initial point in time.

Contemporary models usually involve random error terms, the properties of which are also fully prespecified. These standard probabilistic representations imply a highly restricted view of uncertainty as mere random deviations from a fully predetermined model of behavior. Though they may appear to be different from their deterministic counterparts, contemporary probabilistic models represent market participants as "robots" who revise their behavior according to rules that are prespecified by an economist.

The insistence on models that fully prespecify change has led many economists to an extreme position concerning how policymaking should be

1. See chapters 3 and 4 for a formal discussion of the concept of sharp predictions in contemporary economics. For an early comprehensive treatment, see Sargent (1987).

2. For example, it is common for economists to assume that an individual's utility depends positively on her consumption of goods or that her forecast of a future market price depends positively on the current value of this price.

conducted. Academic economists have argued that discretion on the part of policymakers is likely to result in "inferior" (according to a given "social welfare" criterion) macroeconomic performance. The belief in the scientific status of such conclusions has been so strong that leading economists have advocated far-reaching institutional changes to eliminate all discretion on the part of policymakers.[3] In a seminal paper, for example, Finn Kydland and Edward Prescott advocate

> institutional arrangements which make it difficult and time-consuming to change the policy rules in all but emergency situations. One possible institutional arrangement is for Congress to legislate monetary and fiscal policy rules and these rules to become effective only after a 2-year delay. This would make discretionary policy all but impossible. (Kydland and Prescott, 1977, p. 487)

The trouble with such proposals is that, in reducing policymakers to passive executors of rules based on a fully predetermined economic model, they ignore the multifarious ways in which economies change over time. As Governor Mervyn King of the Bank of England once put it, "Our understanding of the economy is incomplete and constantly evolving, sometimes in small steps, sometimes in big leaps." Because neither economists nor policymakers can adequately prespecify all possible outcomes and their likelihoods, Governor King continued,

> Any monetary policy rule that is judged to be optimal today is likely to be superseded by a new and improved version tomorrow. . . . So learning about changes in the structure of the economy lies at the heart of the daily work of central banks. To describe monetary policy in terms of a constant rule derived from a known model of the economy is to ignore this process of learning. (King, 2005, pp. 8–10)

Although central bankers are always on guard for "changes in the structure of the economy," contemporary models presume that such changes are

3. To avoid misunderstanding, we should stress that what we question here is the scientific status of proposals for rules based on fully predetermined models. Nevertheless, some "rules" or guidelines, such as inflation or exchange rate targets, intended to anchor the decisions of market participants, may play a useful role in policymaking. For example, see Atkins (2006) for a report on how the Norwegian Central Bank uses guidelines and announces long-term forecasts in an attempt to influence market participants' decisions. However, as we discuss in section 1.7, to shed light on the consequences of such policy tools for individual decision making and aggregate outcomes, they would have to be analyzed in models that are not fully predetermined.

unimportant for understanding market outcomes and the consequences of government policies.

1.2. The Aim of This Book

This book arose from our conviction that the contemporary approach to economic analysis of market outcomes is fundamentally flawed. The practice of fully prespecifying the causal mechanism that underpins change leads to insuperable epistemological problems in modeling aggregate outcomes and lies at the root of contemporary models' failure to explain these outcomes in many markets. Our critique rests on the premise that the causal mechanism that underpins the way market participants alter their decisions is not fully intelligible to anyone, including economists or market participants themselves. We hope to persuade our colleagues that the exclusive pursuit of models that "can be put on the computer and *run*" has been misguided; the view that only such models are "scientific" has impeded economic research.

Our goal is to contribute to the development of a more insightful approach to modeling market outcomes and the consequences of government policies. As the first step toward such an approach, we place imperfect knowledge on the part of market participants and economists at the center of our analysis. Our proposed approach, which we call *Imperfect Knowledge Economics* (IKE), does not seek to explain exactly how market outcomes unfold over time. That is, we eschew the contemporary practice that relates change in outcomes precisely to a set of causal factors that has, in turn, been prespecified by an economist.

Following the tradition of early modern economics,[4] IKE constructs its models of aggregate outcomes by relating them to individual behavior. Like the contemporary approach, it represents this behavior mathematically. But IKE attempts to come to terms with early modern economists' justified modesty about how complete their representations of individual behavior could be. As in any scientific theory, IKE must presume that purposeful behavior exhibits regularities, even if these regularities are context-dependent.

4. For lack of a better term, we refer to Friedrich Hayek, John Maynard Keynes, and Frank Knight as *early modern* economists. These economists and some of their contemporaries were early modern in that they attempted to explain aggregate outcomes in reference to individual behavior. But their analyses were far more flexible than contemporary economists' because they understood that the causal factors underpinning individual behavior are often not fully intelligible to individuals themselves, let alone to outsiders such as economists. For further elaboration of this point, see the remainder of this chapter as well as the succeeding chapter.

However, IKE explores the possibility that these regularities, the ways in which market participants make and alter their decisions, may be formalized with *qualitative* conditions. In contrast to both conventional and behavioral models, these conditions only *partially* predetermine economists' representations of change.

IKE solves an intractable epistemological problem that is inherent to fully predetermined, microfounded models of market outcomes. These models, which aim to explain market outcomes on the basis of explicit representations of individual behavior, have become hallmarks of contemporary economics. Yet these models, in both their conventional and behavioral forms, are internally inconsistent in a world of imperfect knowledge: the aggregate outcomes that they predict deviate systematically from their representations of market participants' forecasts of those outcomes. Recognizing the imperfection of knowledge—the fact that no one, including economists, can fully prespecify change—is the key to solving the inconsistency problem that has plagued fully predetermined models. IKE begins with this premise.

In contemporary models, change in the composition of the set of causal variables and in their influences on outcomes is fully prespecified. By contrast, partially predetermined models do not fully specify which causal variables may become relevant in the future or how these variables may enter an economist's representation.[5] By design, IKE models do not imply sharp predictions of change, but they do generate qualitative implications. Moreover, we do not abandon the key aim of all scientific endeavor: IKE restricts its models sufficiently to enable an economist to distinguish empirically among alternative explanations of economic phenomena. At the same time, opening economics to models that generate *only* qualitative predictions is important to understanding salient features of the empirical record that extant approaches have found anomalous.

We use the foreign exchange market as a testing ground for the development of our alternative approach. We find that IKE sheds new light on features of the empirical record that have long resisted adequate explanations by fully predetermined models. We construct IKE models that deliver new, empirically relevant explanations of exchange rate dynamics, particularly their persistent and often large misalignments, as well as movements in the market premium (that is, excess return) on holding a speculative asset, such as foreign exchange.[6] Once we understand market outcomes with IKE

5. In our IKE model of exchange rate swings in chapter 14, we do not need to specify, even in a qualitative way, how a set of causal variables influences individual decision making at any point in time.

6. We develop IKE models of the market premium and long swings in the exchange rate in chapters 12 and 14, respectively.

models, some of the important "findings" that have been reported in the literature are rendered artifacts of a world viewed through the prism of fully predetermined models.

1.3. Contemporary Models in a World of Imperfect Knowledge

Our critique of the contemporary approach rests on the crucial premise that market participants and economists have only imperfect knowledge of the causal mechanism that underpins market outcomes. We recognize that, despite considerable effort by philosophers, the meaning of the term *knowledge,* let alone *imperfect knowledge,* cannot be encapsulated easily. In this book, we make use of a relatively narrow definition of imperfect knowledge that is closely tied to the idea of a fully predetermined model in contemporary economics. We refer to knowledge as *imperfect* if no one has access to a fully predetermined model that adequately represents, as judged by whatever criteria one chooses, the causal mechanism that underpins outcomes in all time periods, past and future. Because knowledge is imperfect, individuals are not constrained to view the world through the prism of a common model. Consequently, one of the main premises of our approach is that market participants, who act on the basis of different preferences, constraints, and causal factors, will likewise adopt different strategies in forecasting the future as well as the consequences of their decisions.

According to Hayek (1945), such a division of knowledge among market participants is the key feature that distinguishes the "rational economic order" from an "optimal" allocation of resources by a single individual:

> The economic problem of society is . . . not merely a problem of how to allocate "given" resources—if "given" is taken to mean given to a single mind which deliberately solves the [resource-allocation] problem. . . . It is rather a problem of how to secure the best use of resources known to any of the members of society, for ends whose relative importance only these individuals know. Or, to put it briefly, *it is a problem of the utilization of knowledge which is not given to anyone in its totality.* (Hayek, 1945, pp. 519–20, emphasis added)

An individual's forecasts of future market outcomes underpin her purposeful choices among alternative uses of her resources.[7] But, as Hayek indi-

7. The distinction between an economist and an individual whose behavior an economist is trying to explain plays a key role in our analysis. Thus, to facilitate our presentation, we have chosen

cated, market participants' choices and, hence, market outcomes, arise out of a division of knowledge whose totality remains opaque to any one individual.[8] As economic knowledge is diffuse and evolves in ways that cannot be fully foreseen, economists' fully predetermined models cannot adequately represent the causal mechanism that underpins purposeful actions, regardless of whether these actions are motivated by self-interest or other objectives.

Nevertheless, we suspect that some of our colleagues may find our critique of the contemporary approach uncompelling. They might argue that economics, like every other field of human inquiry, must abstract from many features of the real world, and that its fully predetermined models are simply particularly bold abstractions. In response to the claim that the assumptions of their models are unrealistic, economists often invoke Milton Friedman's argument:

> The relevant question to ask about the "assumptions" of a theory is not whether they are descriptively "realistic," for they never are, but whether they are sufficiently good approximations for the purpose in hand. And this question can be answered only by seeing whether the theory works[,] which means whether it yields sufficiently accurate predictions. (Friedman, 1953, p. 15)

Useful assumptions in science are therefore those that abstract from features of reality considered irrelevant for the problems under study. The hope is that the omitted considerations are relatively unimportant for one's understanding of a problem.

In general, the assumptions underlying the model of a phenomenon reflect a combination of extant knowledge, convention among scientists, and, at least in part, an investigator's luck and intuition as to what will "work" empirically. Thus, no one can prove on purely logical grounds that the contemporary approach will never succeed in explaining market outcomes. We can, however, appeal to the many epistemological and empirical failures of the contemporary approach and show how these failures stem precisely from contemporary economists' insistence that their models should fully prespecify change. Fully predetermined models are flawed not because they are abstract, but because they disregard a *key* feature that drives outcomes

---

to refer to an economist (or any other outsider) by different gender than that of an individual (a market participant). Of course, the specific choice of whom we refer to as "he" or "she" is without any significance.

8. Building on Hayek, Frydman (1982) formally shows that, in a world of imperfect knowledge, self-interest would lead to a division of knowledge. See chapter 2 for an extensive discussion and further references.

in real world markets: market participants must cope with imperfect knowledge in making decisions that underpin those outcomes.[9]

1.3.1. The Flawed Microfoundations of Fully Predetermined Models

The recognition of ever-imperfect knowledge requires a substantial revision of the contemporary approach to modeling aggregate outcomes on the basis of individual foundations:[10] fully predetermined models lack plausible microfoundations.

Just as any science, economics seeks to uncover and explain empirical regularities. This uncontroversial observation has an important implication for modern economic theory. On the one hand, models of aggregate outcomes are based on mathematical representations of individual behavior; on the other hand, individual behavior depends on forecasts of aggregate outcomes. This inherent two-way interdependence opens up the possibility of the inconsistency that we noted earlier: the forecasts of aggregate outcomes that an economist attributes to individuals may differ systematically from the predictions of an economist's aggregate model.

Lucas argued that models embodying such an inconsistency are "the wrong theory." For example, suppose an economist attributes to firms in each period the forecast that a given market price will remain constant at its current level, while the resulting aggregate model predicts that this price will rise period after period. "In such a model, you could *see* profit opportunities that firms were passing up. Why couldn't they see these opportunities too? But if they did, the model couldn't be right. If your theory reveals profit opportunities, you have the wrong theory" (Lucas, 2001, p. 13).

To rid economic models of this modeling inconsistency, Lucas (1972) and others embraced the Rational Expectations Hypothesis (REH). REH instructs an economist to choose only that representation of individual forecasting behavior that coincides exactly with the causal mechanism implied by the aggregate model that he himself constructs. As Lucas later put it, "John Muth's [REH] focused on this inconsistency . . . and showed how it can be removed" (Lucas, 1995, p. 255). In his Nobel lecture, Lucas recounted this revolutionary moment in the development of the contemporary approach to modeling aggregate outcomes:

> The prevailing strategy for macroeconomic modeling in the early 1960s held that the individual or sectoral models arising out of this

9. For an early critique of conventional models along similar lines, see Frydman (1982, 1983), Frydman and Phelps (1983), and Phelps (1983). Independently, Soros (1987), drawing on the concept of an open society (Popper, 1946), argued that the imperfection of knowledge on the part of market participants is the key to understanding financial markets.

10. See chapter 3 for a formal argument.

> intertemporal theorizing could then simply be combined in a single model. But models of individual decisions over time necessarily involve expected future prices. . . . However, . . . [aggregate] models assembled from such individual components implied behavior of actual prices . . . that bore no relation to, and were in general *grossly inconsistent* with, the price expectations that the theory imputed to individual agents. (Lucas, 1995, pp. 254–55, emphasis added)

In the wake of such statements, achieving consistency between representations on the individual and aggregate levels is widely perceived as the solution to the problem of modeling rational forecasting behavior. Moreover, because economists believe that REH delivers such consistency, it has become the key building block of most economic models of aggregate outcomes.

1.3.1.1. REH Models

Lucas's insight concerning the key theoretical importance of avoiding internal inconsistency in models of aggregate outcomes that are based on representations of individual behavior is compelling. But what he and other followers of REH seem to have overlooked is that the outcomes that their models try to explain stem from decisions by individuals who must all cope with ever-imperfect knowledge. Although Lucas formulated his argument against fully predetermined non-REH models, it also applies to REH models.

Building on Frydman (1982), we argue that, *in a world of imperfect knowledge,* REH models presume that individual market participants endlessly disregard systematic information in their forecast errors. REH supposes that an economist's fully predetermined model of aggregate outcomes adequately represents, at least in the aggregate, market participants' forecasting strategies. Recalling Lucas, if an REH model were to capture adequately empirical regularities in aggregate outcomes, and some individuals did not use this model to forecast, it would mean there were profit opportunities that individual market participants were seeing and endlessly passing up. In a world of imperfect knowledge, market participants make use of diverse forecasting strategies; to account for this diversity, therefore, an REH model has to presume gross irrationality. If Lucas's alarm at the prospect of inconsistent models was appropriate, as we think it was, we must conclude that REH models, too, are the "wrong theory" for modeling aggregate outcomes on the basis of individual foundations.[11]

11. For a more complete discussion of this issue, see chapter 3.

1.3.1.2. Behavioral Models

Behavioral economists, for their part, have uncovered many inconsistencies between the way market participants "actually" behave and conventional representations of rational behavior. They have not, however, interpreted their findings as evidence that the main problem with conventional representations is that they do not adequately represent rational behavior. Instead, they have concluded that market participants are irrational.

Behavioral economists justify this striking conclusion by arguing that market participants are, after all, human beings with inadequate foresight, unpredictable emotions, and limited computational abilities. But, despite their putative embrace of "psychological realism," behavioral economists emulate their conventional colleagues by disregarding the importance of individual creativity, which is arguably one of the defining features of human behavior. They do so by prespecifying "[h]ow *exactly* people deviate from the [conventional representation of rationality]" (Barberis and Thaler, 2003, p. 1056, emphasis added).

By design, a non-REH behavioral model embodies an inconsistency between its representations on the individual and aggregate levels: it represents an individual's forecasting strategy as one whose predictions are systematically inconsistent with the predictions of the aggregate model.[12] Lucas has argued forcefully that models, such as non-REH behavioral ones, in which market participants endlessly disregard systematic information in their forecast errors, posit the "wrong theory."

This reasoning leads us to conclude that there is an inherent conflict between the objective of searching for models of aggregate outcomes based on explicit microfoundations and the insistence of both conventional and behavioral approaches that these models be fully predetermined. The methodology raises an intractable epistemological problem: in real world markets, the "microfoundations" of fully predetermined models prove specious.

1.3.2. Structural Instability of Contemporary Models

Although fully predetermined representations of individual behavior cannot serve as the microfoundations of any theory of aggregate outcomes, this conclusion does not rule out the possibility that fully predetermined models of aggregate outcomes, even if they are based on explicit microfoundations, might be useful. With insightful selection of the causal variables and a bit of luck, fully predetermined models may capture adequately, according to

12. Puzzlingly, although Lucas argued that this inconsistency provided the main motivation for the REH revolution, behavioral economists, who developed their approach after REH, have disregarded Lucas's arguments.

statistical or other, less stringent criteria, the past relationship between the causal variables and the aggregate outcomes in a selected historical period.

But as time passes, market participants eventually alter the way that they make decisions and the social context changes in ways that cannot be fully foreseen by anyone. Disregarding these key determinants of change restricts the usefulness of such models for illuminating the historical record to limited periods of time. The well-known debacle of the hedge fund Long Term Capital Management suggests how fully predetermined models eventually become inadequate. After all, trading in financial markets cannot be reduced to mere financial engineering, even if it is based on the most recent advances in contemporary finance theory.

We are thus led to one of the key empirical implications of imperfect knowledge: even if a fully predetermined model's structure adequately represents outcomes in terms of a set of causal variables during a particular period of time, it will be inadequate during other periods. When such models are used to analyze time-series data—or as aids for policymakers, bankers, traders, and others—they should always be tested for possible structural changes. These test procedures should not require an economist to fully prespecify when structural change might occur or which causal variables might enter the structure of the postchange representation. In chapters 12 and 15, we develop such an approach in the context of modeling the premium on foreign exchange and the exchange rate, respectively.

1.4. The Non-Fully Intelligible Individual

The observation that extant models lack plausible microfoundations leads us to a re-examination of how individual decision making ought to be represented. Conventional economists often use the same fully predetermined representation to explain individual self-interested behavior and aggregate outcomes over many decades or in different economies or markets. They impute such generality to their "theory" because they believe that self-interest is a universal human trait and that they have found a way to represent it with fully predetermined rules. Behavioral economists share this view of rationality. This understanding leads them to diagnose the inconsistency between the actual behavior of market participants and conventional representations of "rational" behavior as a symptom of market participants' "irrationality."

IKE is compatible with—but does not necessarily require—the presumption that market participants in capitalist economies are motivated by purely self-interested concerns. The focus on self-interest in economic analysis has elided a central issue: even if self-interestedness were universal, such a presumption would not enable an outsider to fully prespecify his representations of self-interested behavior or their implications both for aggregate

outcomes and for the consequences of economic policies. The reasons for this assertion are already implicit in our foregoing discussion. Individual decisions depend on forecasts of future market outcomes. These outcomes are not only a result of the actions of many individuals, but they also depend on future economic policies, political developments, and institutional changes. Thus, even if individuals are presumed to be purely self-interested, the way that they deploy resources depends as much on the social context as it does on their personal motivations.[13]

An individual herself, let alone an outsider, cannot fully prespecify how she will form and revise her forecasts. Even if an economist were able to attribute clear objectives to a market participant, he would still be unable to assess the participant's rationality.[14] IKE supposes, therefore, that an economist cannot ascertain completely whether an individual behaves rationally or irrationally; that is, he cannot completely evaluate whether she pursues her objectives reasonably or unreasonably.

1.5. IKE Models

How can economic analysis recognize the centrality of imperfect knowledge while continuing to represent individual and aggregate behavior in mathematical terms? How can it acknowledge the importance of individual creativity and the inevitability of unpredictable changes in social contexts while still generating "predictions which . . . are of empirical significance"? The future relevance of economics to understanding real world markets and policy analysis lies in its ability to articulate answers to these questions; the IKE framework offers one response.

Like contemporary models, IKE models consist of representations of an individual's preferences, constraints, and forecasts of future outcomes that are relevant to her well-being. IKE also imposes qualitative conditions on its representations at an initial arbitrary point in time. But, in sharp contrast to the contemporary approach, IKE does not fully prespecify how its representations of preferences and forecasting behavior change between an initial point and all other points in time.

IKE recognizes that without some regularity in economic life, no economic theory that aimed for generality would be possible. IKE considers two types of regularities on the individual level. First, *an individual's preferences or forecasting strategy at different points in time may share certain qualitative features.*

13. Foley (2003) has advanced a related criticism of the notion of rationality invoked in economic analysis.

14. Kay (2004, p. 16) has called this fundamental difficulty "obliquity." As he quipped, "no one will ever be buried with the epitaph 'He maximized shareholder value,' . . . because even with hindsight there is no way of recognising whether the objective has been achieved."

For example, at each point in time, her utility may depend positively on changes in her wealth or her forecast of a future market price may depend on a changing subset of causal variables that is contained in an unchanging larger set.

Second, although the way that any of the causal variables affect an individual's preferences or forecasting behavior may change with time, these changes may share certain qualitative features. For example, although an individual might substantially revise her preferences or forecasting strategy, the effect of such revisions may be "conservative," that is, her pre- and postchange preferences or forecasts may not be "too" different.

IKE instructs an economist to search for regularities in individual behavior but presumes that they can at best be formalized with qualitative restrictions on its representations. Because IKE's restrictions only partially prespecify change in a model, we refer to them as *partially predetermining*. Just as in extant models, IKE represents future outcomes as uncertain. However, instead of the standard (fully predetermined) conditional probability distributions implied by a contemporary model, an IKE model relates the distribution of the outcome variables at a future time to its distribution at an initial time in only a qualitative way. Thus, IKE's partially predetermined probabilistic representations of change are compatible with Knight (1921) and Keynes's (1921, 1936) insight that economists cannot fully prespecify the consequences of individual decisions or future market outcomes and their chances of occurrence.[15]

In contrast to the fully predetermining restrictions of a contemporary model, the restrictions of an IKE model do not force an economist to take a position on how a set of causal variables or their influences on individual decision making may change between any two points in time. Because IKE only partially prespecifies its representations, a change in a causal variable back to its initial value does not lead an IKE model to devolve to its initial structure. Thus, historical change plays an essential role in an IKE model: as time passes, partially predetermined representations make allowance for the possibility that market participants' knowledge and their decisions concerning the use of resources will evolve in ways that cannot be fully foreseen.[16]

15. See chapter 4 for a formal treatment.

16. This point has an important implication for supply and demand analysis, which is basic to economics. Under IKE, the unique equilibrium implied by the usual fully predetermined supply and demand curves is replaced by a myriad of equilibria implied by partially predetermined *supply and demand paths*. Although these paths are not unique, they share common features: all demand paths are downward sloping and all supply paths are upward sloping. It is these common features that enable supply and demand analysis of aggregate outcomes under IKE to replace the usual analysis implied by fully predetermined models. See chapter 5.

1.5.1. Individual Preferences

Many studies have found that conventional representations of preferences, which usually involve expected utility theory and the assumption of risk aversion, are grossly inconsistent with the way individuals actually behave. Much of the evidence on how individuals make choices is based on laboratory experiments in which the structure of payoffs from various gambles is predetermined by the experimenter. This common experimental design allows the investigator to examine the nature of an individual's preferences without the confounding problem of having to represent her forecasts of the potential payoffs from gambling. The findings concerning the importance of loss aversion and the seminal formulation of prospect theory by Kahneman and Tversky (1979) and Tversky and Kahneman (1992) made use of such a setup.

Building on prospect theory, we develop an alternative representation of preferences for modeling decision making that is consistent with the experimental evidence. This representation, which we call *endogenous prospect theory,* supposes that an individual's preferences share certain qualitative features at every point in time. This utility ranking depends on her forecast of the outcomes of her decisions regarding the allocation of her resources, in particular, on her forecast of future returns and on her forecast of the size of the potential loss that she might incur. The representation also presumes that an individual's degree of loss aversion increases as her forecast of the size of the potential loss increases.[17] Because we represent this forecast with partially predetermining restrictions, the way in which an individual's degree of loss aversion changes between any two points in time is also partially predetermined in our models.

Although laboratory experiments have been the key to uncovering new ways to model preferences, their typical design effectively limits the economist's view of an individual's decision making; the economist is only able to observe the subject's responses to an experimenter's stimuli. This basic framework, which is used extensively in psychological research, sidesteps a key problem: participants in real world markets forecast payoffs, the experimenter's "stimuli," on the basis of imperfect knowledge. Moreover, these forecasts depend not only on the subject's creativity, her analytical abilities, and other personal characteristics, but also on the unfolding social context.[18] As a result, the basic type of model used in these psychological ex-

17. In chapter 9, we show that this assumption, which we call *endogenous loss aversion*, is needed to model an individual's decision about how much capital to gamble at any point in time solely on the basis of prospect theory.

18. Kahneman and Tversky (1979) recognized that, while laboratory experiments are useful in uncovering the properties of the utility function over single outcomes, they may be much less informative about an individual's choices over gambles with two or more uncertain outcomes in real world markets.

periments is grossly insufficient as a foundation for representing economic behavior.

1.5.2. Individual Forecasting Behavior

The premise that self-interested or, more broadly, purposeful behavior is to an important degree context dependent does not dispute the usefulness of insights from psychology in modeling individual behavior. Indeed, we make use of some of these insights in representing how an individual revises her forecasting strategy. For example, researchers have uncovered much evidence that individuals are conservative in how they revise their beliefs in the face of new evidence.[19] In our model of exchange rate swings, we formulate this finding in terms of partially predetermining restrictions that limit the change in a market participant's forecast that arises from the change in her forecasting strategy.

However, the importance of the social context in an individual's decision making implies that, in searching for empirical regularities that might be useful in modeling an individual's decisions, economists will need to look beyond laboratory experiments and insights from psychology. To represent individual behavior, an economist must search for—and attempt to formalize—the findings of other social sciences. Other social scientists have knowledge and intuitions concerning the social context within which individuals make decisions that may complement economists' work in modeling individual forecasting behavior.[20]

We make use of the insight that conventions among market participants play an important role in individual decision making.[21] We also draw on our understanding of the qualitative regularities that have characterized aggregate outcomes; we suppose that market participants must also be aware of these regularities when they are forming their forecasts. For example, the tendency of exchange rates to undergo long swings that revolve around historical benchmark levels plays a key role in our model of the premium on foreign exchange. Our representations of individual behavior, on the basis of which we construct our model for the market premium, involve specifications for bulls' and bears' forecasts of the potential loss from holding

19. See Edwards (1968) and Shleifer (2000).

20. The use of insights concerning the social context in modeling individual behavior has a venerable tradition in sociology. Perhaps best known is Weber's argument that Protestantism—an important aspect of the social context within which individuals made decisions—is key to understanding "rational" behavior at the time of the emergence of capitalism. See chapter 2 for a discussion and references.

21. For early insights on the role of social conventions in individual decision making and its implications for market outcomes, see Keynes (1936).

speculative positions.[22] These representations constrain, in a qualitative way, the revisions of market participants' forecasts of the potential loss to depend on the gap between the exchange rate and their assessments of its historical benchmark.[23]

The gap and conservative restrictions, because they only partially prespecify change in our models, are consistent with myriad possible ways that an individual's forecasting strategy may develop over time. Nevertheless, the possible ways that this strategy could unfold in a model share certain qualitative features. It is this partially predetermined aspect of IKE models that enables them to deliver testable implications.

The distinguishing feature of our IKE models of exchange rate swings and the market premium is that they do not require an economist to prespecify either the potential set of causal variables that underpin change in outcomes or the influences of these variables in his representation. This feature is important, as the presumption that an economist can prespecify, even partially, the set of causal variables and their influences is very bold.

Nevertheless, in addressing some problems, we acknowledge that an economist is required to represent these aspects of the causal mechanism. For example, to examine whether macroeconomic fundamentals matter for exchange rate movements, an economist must prespecify, at least partially, a representation of the causal mechanism that involves the set of potential fundamentals (potential causal factors) and how they influence the exchange rate. To this end, we consider the idea that the stock of extant economic models summarizes economists' insights concerning the causal factors that underpin market outcomes. Presumably, these insights are shared by market participants. This idea underlies the *Theories Consistent Expectations Hypothesis* (TCEH) proposed by Frydman and Phelps (1990). TCEH recognizes that a set of extant economic models at best indicates to a market participant, or to an economist attempting to represent her behavior, which causal variables may be important for forecasting market outcomes; it also suggests, in a qualitative way, how these variables may influence those outcomes.

In chapter 15, we propose a simple procedure that enables an economist to decipher qualitative features of the reduced forms of a set of models under imperfect knowledge. TCEH provides a way to take into account the qualitative features of more than one model in constructing its representa-

22. A bull (a bear) is a market participant who speculates on the belief that the asset price will rise (fall).

23. This idea, that the convention concerning the historical benchmark level plays an important role in understanding individual forecasting behavior in an asset market, was put forth by Keynes (1936). It was formalized by Tobin (1958) as a key component of his model of the speculative demand for money.

tions of market participants' forecasting strategies.[24] Although TCEH may seem to be a qualitative analog of REH, there are two fundamental differences. To account for the social context within which market participants act, TCEH recognizes that an economist cannot ignore the pluralism of models. And, furthermore, TCEH only partially prespecifies change.

1.5.3. How IKE Avoids Modeling Inconsistency

Lucas (2001) has argued that fully predetermined models that embody an inconsistency between their representations on the individual and aggregate levels do not constitute a theory of persistent regularities. Lucas's argument makes the general claim that, if a model predicts an unchanging feature of the causal mechanism underpinning aggregate outcomes regardless of whether the feature is qualitative or quantitative, then an economist's representation of individual behavior should not be inconsistent with this regularity.

Our argument that REH does not solve the inconsistency problem in a world of imperfect knowledge begs the question of how IKE avoids it. IKE does so by recognizing the limits of economists' knowledge, that is, by instructing economists only to partially prespecify their representations of change. IKE avoids the inconsistency in different ways, depending on whether an economist represents the same or distinct aspects of the causal mechanism on the individual and aggregate levels. If he represents the same features on the individual and aggregate levels, he must choose a representation of individual forecasting behavior that leads to the same qualitative predictions concerning the causal mechanism as that implied by his aggregate model.

In chapter 15, for example, we construct a model that generates qualitative predictions about the relationship between the exchange rate and macroeconomic fundamentals. We illustrate TCEH by using the qualitative features of three extant monetary models to represent an individual's forecasting strategy. If the models all agree on how a particular causal variable influences the future exchange rate, then our TCEH representation constrains the sign of the weight that is attached to this variable to be consistent with these models. Otherwise, TCEH leaves this sign unconstrained. Thus, although TCEH uses a variety of models to represent forecasting behavior, it avoids, by construction, the inconsistency between an economist's representations on the individual and aggregate levels.

Although TCEH offers a way to avoid an inconsistency in models that involve representations of the same aspect of the causal mechanism, some

24. For an early implementation of TCEH in modeling exchange rates, see Goldberg and Frydman (1996a).

features of the empirical record may only require that an economist represent different aspects of the causal mechanism in modeling individual and aggregate behavior. If an economist were to set out to construct a fully predetermined model of such phenomena, he could not avoid an internal inconsistency under imperfect knowledge. However, because IKE constrains its representations only partially, it enables an economist to model distinct aspects of the causal mechanism on the individual and aggregate levels, while avoiding the inconsistency problem.

For example, our IKE model of the premium implies that the excess return on foreign exchange depends positively on the gap between the exchange rate and market participants' assessments of its historical benchmark. Our representations on the individual level, however, involve predictions concerning a different aspect of the causal mechanism, namely, the potential loss from holding a speculative position. Moreover, because our model is partially predetermined, the qualitative prediction on the aggregate level places no constraints on how the gap should influence an individual's forecast of her potential loss. In this way, our IKE model of the premium avoids an internal inconsistency and, in contrast to fully predetermined models of the premium, it does not presume that market participants are grossly irrational.

1.6. IKE of Exchange Rates and Risk

In parts II and III of this book, we focus on many of the aspects of behavior in currency markets that are particularly difficult to reconcile with the conventional approach to exchange rate dynamics. The exchange rate modelers of the 1960s and 1970s believed that currency movements depended largely on macroeconomic fundamentals and that market participants acted in "rational" ways.[25] But the empirical failures of conventional exchange rate models have led economists to abandon these tenets; many now believe that macroeconomic fundamentals do not play an important role in currency movements and that market participants forgo obvious profit opportunities in making their speculative decisions.

1.6.1. Exchange Rate Swings and Macroeconomic Fundamentals

It is clear from the past three decades of floating currencies that exchange rates have a tendency to undergo large and persistent swings away from historical benchmark levels, such as those based on purchasing power parity

25. The seminal open-economy model is from Mundell (1963) and Fleming (1962). Other early milestone studies include Dornbusch (1976), Frenkel (1976), and Kouri (1976).

(PPP). Economists have found this kind of behavior difficult to explain with models that use REH, such as those based on Dornbusch (1976).[26] It is striking that, by 1983, Rudiger Dornbusch himself had reached the conclusion that his own overshooting theory could not explain the long-swings behavior of exchange rates. He asserted, "A theory is needed that will explain why the dollar—real or nominal—is both high and stuck [away from parity]" (Dornbusch, 1983, p. 83). The failure to explain swings, to which Dornbusch alluded, led to the view that exchange rate fluctuations are driven by "irrational noise" traders who do not rely on macroeconomic fundamentals. This view was further reinforced by economists' unsuccessful search for a fully predetermined, mostly invariant, relationship between the exchange rate and macroeconomic fundamentals during the current period of floating rates.

However, there is much evidence in the literature that movements in macroeconomic fundamentals *do* influence exchange rates, but in different ways during different time periods. To account for this temporal instability, we replace REH with an IKE representation of forecasting behavior in the context of the Dornbusch model. Remarkably, swings can occur in our model even if all market participants form their forecasts solely on the basis of macroeconomic fundamentals. Indeed, it is the influence of persistent trends in such fundamentals on market participants' forecasts that cause them to bid the exchange rate away from PPP. Because we recognize that market participants must cope with imperfect knowledge, our model does not rely on the presumption that they are irrational.[27]

1.6.2. Returns on Foreign Exchange

Relying on invariant empirical relationships, many researchers report that future returns in currency markets co-vary negatively with the current value of the forward premium.[28] To explain this behavior, conventional economists have constructed exchange rate models in which risk-averse individuals require a positive return, a premium, to hold risky positions in currency markets. It is widely recognized, however, that this research effort has been unsuccessful.[29]

26. Conventional economists have recently generalized the sticky-price model of Dornbusch (1976) to include complete intertemporal microfoundations. We discuss this literature of "new open economy macroeconomics" in chapter 7.

27. For early analyses that also make use of qualitative assumptions about individual forecasting in explaining long swings in the exchange rate, see Schulmeister (1983, 1987) and Soros (1987). Soros (1998, 2006) uses a similar framework to analyze historical change more generally.

28. The forward premium depends on the difference between the forward and spot exchange rates.

29. See chapter 8 for a discussion of this literature.

The failure of REH risk-premium models has led many economists to the view that market participants forgo obvious profit opportunities. IKE does not rely on the presumption of irrationality. We thus explain returns as the compensation for the risk of capital loss from holding speculative positions in foreign exchange. To this end, we develop an IKE model of returns on foreign exchange that is based on endogenous prospect theory.[30]

Our alternative representation of preferences implies that all market participants require a minimum premium before they commit any capital to speculating in the foreign exchange market. This result, in turn, leads to a new equilibrium condition in the foreign exchange market, which we refer to as *uncertainty-adjusted uncovered interest parity*.

But experimental evidence on individuals' preferences is not enough to model their decisions about how much capital to hold in foreign exchange. We also need to represent individuals' forecasting behavior. We do so by replacing REH with partially predetermined IKE representations of forecasting strategies and their revisions.

Our IKE model of the premium is able to capture the fact that market participants differ in their forecasting strategies: bulls gamble on appreciation, while bears bet on depreciation. This difference between bulls and bears proves crucial to explaining the behavior of foreign exchange returns.

1.6.3. Is the Market Really Grossly Inefficient?

Unable to explain the negative co-variation between the return on foreign exchange and the forward premium that their studies report, economists have reached the startling conclusion that "one can make predictable profits by betting against the forward rate" (Obstfeld and Rogoff, 1996, p. 589). The apparent anomaly that these profits remain unexploited has become one of the major "puzzles" in the international finance literature.

There are several well-known studies in the literature that indicate that the relationship between the return on foreign exchange and the forward premium is temporally unstable. In part III, we add to this evidence and show that the correlation between the return on foreign exchange and the forward premium is sometimes negative, sometimes positive, and sometimes insignificantly different from zero.

30. Some economists have begun to recognize that the failure of conventional models to explain returns in asset markets stems from ignoring ever-imperfect knowledge on how the structure of the economy unfolds over time. In a recent study along these lines, Weitzman (2006) introduces "perennial" uncertainty concerning the variance of the future growth of consumption into a standard consumption-based asset-pricing model.

Acknowledging the importance of temporal instability goes a long way toward resolving the forward-rate "puzzle." A returns process that gives rise to both negative and positive correlations with the forward premium implies that betting against the forward rate will be profitable during some time periods but not in others. We show that a trading rule based on betting against the forward rate does not deliver significant profits over the modern period of floating in the major currency markets.

Because the contemporary approach has led economists to construct fully predetermined, mostly invariant models of foreign exchange returns that ignore temporal instability, the "finding" of a negative correlation between returns and the forward premium has led them to conclude that there is easy money to be made in the foreign exchange market. But, because the correlation is not always negative, fully predetermined trading rules based on the forward rate do not deliver profits. As in the case of the disjunction between the exchange rate and macroeconomic fundamentals, the forward-rate "puzzle" is another artifact of the epistemological flaws inherent in the contemporary approach. The existence of literally hundreds of studies attempting to explain this "puzzle" provides an example par excellence of how contemporary economics' insistence on sharp predictions has misdirected research and impeded its progress.

1.7. Imperfect Knowledge and Policy Analysis

As in the case of modeling aggregate outcomes, the validity of policy prescriptions based on fully predetermined models should be reexamined under imperfect knowledge. Although this task is outside the scope of this book, our analysis of exchange rate swings provides a strong indication that fully predetermined models are unsuitable for policy analysis in a world of imperfect knowledge.

The vast majority of contemporary economists presume that, once the policy environment is fixed, invariant models adequately represent market participants' behavior. However, we find that, if invariant representations were to represent individual behavior in a traditional monetary model under conditions of imperfect knowledge, then a policy rule that set money supply to grow at a fixed rate would imply an unbounded swing in the exchange rate away from PPP.[31] Stated succinctly, a fixed money-growth rule exacerbates, rather than limits, the magnitude of the exchange rate swing.

31. See chapter 14.

Our conclusion that fully predetermined models provide inadequate bases for policy prescriptions still leaves open the possibility that guidelines intended to anchor the forecasts of market participants, such as exchange rate or inflation targets or central bank forecasts, may play a useful role in a world of imperfect knowledge.[32] But this possibility will need to be analyzed in models that do not presume that economists can fully prespecify change.[33]

1.8. From Contemporary Economics to Imperfect Knowledge Economics

The research program of contemporary economics is predicated on the belief that it is possible to prespecify economic change over stretches of time as long as decades. The premise that seems, at least implicitly, to motivate this mechanistic way of modeling market outcomes is that there exists a fully predetermined causal mechanism that underpins actual behavior on the individual and aggregate levels.

But aggregate outcomes and individual forecasting behavior are not governed by an overarching causal mechanism in many, if not all, markets in capitalist economies. Creative forecasting behavior on the part of purposeful individuals alters the causal mechanism that underpins market outcomes in ways—and at points in time—that cannot be fully prespecified. Moreover, changes in the social context, including the evolution of institutions, values, and norms, are all important in engendering temporal instability in causal relationships in real world markets.

If change in capitalist economies is not governed by a fully predetermined causal mechanism, then attempting to explain individual behavior and aggregate outcomes on the basis of representations that presume the existence of such a mechanism is clearly misguided. It is not surprising, then, that the contemporary approach has had great difficulties in discovering the "mechanics of economic development" in many markets where

32. Indeed, a number of central banks have been using such tools in their policymaking. See footnote 3.

33. We have begun such a line of research in Frydman and Goldberg (2004) and Frydman, Goldberg, and Cavusoglu (2007), where we show that our IKE model of the premium leads to a new view of how policy officials can limit the magnitude of long swings in floating-rate regimes. In a manuscript under preparation (Frydman and Goldberg, 2008), we also examine other policy issues, such as inflation targeting, in the context of an IKE model.

profit-seeking inherently involves individual creativity in coping with ever-imperfect knowledge.

Economics calls for a new approach that represents individual behavior and aggregate outcomes mathematically, and that, at the same time, refrains from fully prespecifying economic change. Taking up this task, the following pages propose IKE as the beginnings of such a new approach.

2 A Tradition Interrupted

> The peculiar character of the problem of a rational economic order is determined precisely by the fact that the knowledge of the circumstances of which we must make use never exists in concentrated or integrated form but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess.
>
> Friedrich A. Hayek,
> "The Use of Knowledge in Society,"
> *American Economic Review,* p. 519

> **Evans and Honkapohja**: Do you think differences among people's models are important aspects of macroeconomic policy debates?
> **Sargent**: The fact is that you simply cannot talk about those differences within the typical rational expectations model. There is a communism of models. All agents inside the model, the econometrician, and God share the same model. The powerful and useful empirical implications of rational expectations . . . derive from that communism of models.
>
> George Evans and Seppo Honkapohja,
> "An Interview with Thomas J. Sargent,"
> *Macroeconomic Dynamics,* pp. 566–67

Relating aggregate outcomes to individual decision making has been a hallmark of modern economics. The largely narrative mode of analysis used by Friedrich Hayek, Frank Knight, John Maynard Keynes, and their contemporaries enabled these giants of early modern economics to examine the importance of individual creativity, the division of knowledge and its unfolding over time, and the roles of social norms and institutions for understanding how individual behavior and aggregate outcomes develop. Indeed, their great insight was to place at the center of economic analysis the inextricable connection between imperfect knowledge, non-routine behavior, and the pursuit of profits in capitalist economies. Although the early modern economists recognized the importance of explaining aggregate outcomes on the basis of individual behavior, they also pointed to a fundamental difficulty inherent in any attempt to do so: purposeful behavior, whether motivated by pure self-interest or other objectives, is not completely intelligible to outsiders, whether they are economists, policy officials, or social planners. Consequently, market outcomes that result from the decisions of many individuals are not completely intelligible either.

Post-1945 models of aggregate outcomes recognized the inherent tension between the early modern insights and the attempt to relate aggregate

outcomes precisely to individual behavior.[1] "For this and other reasons, [these] macro-economic models . . . were only loosely linked to optimizing behavior of individual agents" (Tobin, 1981, p. 14).

Phelps (1968a) and Phelps et al. (1970) are generally credited with pioneering the modeling of aggregate outcomes on the basis of explicit mathematical representations of individual behavior. These studies acknowledged the importance of the key early modern insight that the way an individual forecasts the future consequences of her decisions cannot be completely understood by an economist. However, it was not apparent how to incorporate this insight into mathematical representations of individual behavior and its implications for aggregate outcomes.[2] Evidently, the hope was that future research would continue to search for ways to resolve the fundamental tension between the non-fully intelligible individual and attempts to represent her behavior mathematically. In the event, subsequent researchers disregarded this vision and have undertaken an intensive effort to construct the "mechanics" of economic development.

2.1. The Stranglehold of the Contemporary Approach

Hayek's (1978) dismissal in his Nobel lecture of the search for "exact knowledge" carried a corollary warning against the common "belief that in order to be accepted as scientific it is necessary to achieve more" than qualitative predictions. Ignoring Hayek's warning, contemporary economists have embraced that belief and have undertaken an intensive empirical effort to uncover "exact knowledge" of aggregate relationships based on fully predetermined representations of individual behavior.

2.1.1. REH

Initially, the contemporary approach to economic analysis of aggregate outcomes was based on representations of "rational" behavior, with its models of an individual's forecasting strategy and its revisions over time employing REH. Such representations presume that an economist can capture adequately what Hayek (1945) referred to as "the use of knowledge in society" with his own fully predetermined model of aggregate outcomes. "In rational

1. For a discussion of this issue, see Tobin (1981) and Hahn and Solow (1995).

2. In the absence of a better alternative, these early "microfounded" models represented individual forecasting behavior with a fixed error-correcting rule, called "adaptive expectations." See chapter 3.

expectations models, people's beliefs are among the outcomes of our theorizing. They are not inputs."[3]

Frydman (1982) argued that there is an inherent conflict between REH's presumption that people's beliefs can be adequately represented as one of the outcomes of an economist's theorizing and the premise that market participants are motivated by self-interest: purposeful individuals would not, in general, adhere to a single forecasting strategy.[4] In contrast, REH presupposes an agreement among market participants to rely on one common forecasting strategy. Phelps sharply criticized this "communism of models":

> In the theory of macroeconomic disturbances . . . to which the rational expectations hypothesis has frequently been applied, it is difficult to justify the premise that each agent presumes his expectations to be universal—as if some Jungian collective unconscious existed to bring expectations into an understood agreement. There is no nationwide expectation of Reagan's economic policy *the existence of which is public knowledge*. (Phelps, 1983, p. 32, emphasis in original)

REH is often believed to represent the way rational individuals use information to formulate and revise their forecasting strategies. As Sargent put it:

> The idea of rational expectations is sometimes explained informally by saying that it reflects a process in which individuals are inspecting and altering their forecasting records in ways to eliminate systematic forecast errors. . . . It is also sometimes said that [REH embodies] the idea that economists and the agents they are modeling should be placed on equal footing: the agents in the model should be able to forecast and profit-maximize and utility-maximize as well as . . . the econometrician who constructed the model. (Sargent, 1993, p. 21)

3. Thomas J. Sargent, in an interview with Evans and Honkapohja (2005, p. 566). All further quotes that appear in this section without explicit citation are taken from that interview. From the quotes in this section, it may appear that Thomas Sargent, one of the pioneers of the REH approach, has lately become one of its critics. We should acknowledge, however, that in personal communication with Frydman, Sargent emphasized that he does not regard himself as participating in any critique of REH.

4. Frydman showed that popular "learning" models that are often used to justify REH representations of forecasting behavior presume that market participants somehow agree to use the same fully predetermined learning rule. The subsequent literature has disregarded this difficulty and continued to prespecify common learning mechanisms to justify REH models. For a survey of the numerous studies that have attempted to do so, see Evans and Honkapohja (2001) and references therein.

Like Frydman (1980, 1983), and Frydman and Phelps (1983), however, he then pointed out that "these ways of explaining things are suggestive, but misleading, because they make [REH] sound less restrictive and more behavioral than it really is" (Sargent, 1993, p. 21).

Despite the fundamental flaws of REH as a representation of how rational individuals do or should behave, its proponents have been extraordinarily successful in persuading their colleagues to join the REH revolution. But, just as the number of economists embracing this approach and working out its implications has increased, so has empirical evidence of its predictive failures accumulated.

The failure of REH models is particularly apparent in financial markets. After reviewing many empirical studies, Maurice Obstfeld and Kenneth Rogoff concluded in their magisterial book on the REH approach to international economics that "the undeniable difficulties that international economists encounter in empirically explaining nominal exchange rate movements are an embarrassment, but one shared with virtually any other field that attempts to explain asset price data" (Obstfeld and Rogoff, 1996, p. 625).

The dismal performance of REH models in markets in which forecast revisions drive movements in prices and returns stands in sharp contrast to the widespread belief that these models have provided *the* solution to the problem of adequately representing the forecasting behavior of rational individuals. But, as Lucas (2003a,b) recently acknowledged, the problems with the conventional approach are not confined to financial markets. For example, he pointed out that little progress has been made in the three decades since Phelps et al. (1970) examined the persistent effects of changes in monetary policy on a nation's real output:

> New frameworks—contracts, monopolistic competition—are introduced, motivated by the inability of earlier theory to resolve [the] difficulty [of explaining persistent real responses to monetary shocks], but the problem of persistence has proved to be persistent itself. . . . Ever since the January, 1969, conference that Ned Phelps invited us to, the 14 authors of the Phelps volume have been apologetic about the fact that we couldn't resolve these issues. (Lucas, 2003b, p. 140)

Lucas is acknowledging that the conventional approach could not remedy its failures by modifying ancillary assumptions, such as the nature of contracts, market structure, and other institutional arrangements.

Behavioral economists, for their part, have focused on the lack of "psychological realism" of conventional models. However, instead of questioning whether conventional models adequately capture rational behavior, behavioral economists concluded from the empirical failure of these models that individuals are irrational. Indeed, they proceeded to prespecify

fully "irrational" behavior. This intellectual development illustrates par excellence the stranglehold the contemporary approach has maintained on the direction and methods of economic research.

2.1.2. The Retreat from Statistical Inference

Conventional economists have responded to the empirical failures of REH models in an even more startling way than their behavioral colleagues. They disregard the significance of the disappointing findings and continue to build their models on the basis of REH, while refining the microfoundations of their models.

In recent years, for example, international macroeconomists have been engaged in an intensive effort to salvage the older overshooting models of the exchange rate originated by Dornbusch (1976) by basing these models on representations of "intertemporally optimizing agents." This effort has required economists to prespecify fully the forecasting behavior of market participants over the indefinite future. As we discuss in chapter 7, this research effort, now called the *New Open Economy Macroeconomics* (NOEM), has been launched despite Dornbusch and Frankel's assessment that "the chief problem with the overshooting theory, indeed, with the more general rational expectations approach, is that it does not explain well the shorter-term [long-swings] dynamics in exchange rates" (Dornbusch and Frankel, 1995, p. 16).

To help rationalize this steadfast adherence to REH modeling in the face of its empirical failures, conventional economists have redefined the notion of empirical failure. Prior to the ascendancy of the conventional approach, economists, like *all* scientists, relied primarily on standard methods of statistical inference to confront their models with empirical data. By contrast, REH theorists have decided that these methods were too stringent for judging the adequacy of their models, so they have sought to develop an alternative methodology.

The structure of economic models embodies economists' hypotheses concerning individual behavior and its implications for aggregate outcomes. Assuming that an economist's model is an adequate representation of the causal mechanism underlying an economic outcome, such as the market price, the methods of statistical inference enable him to ascertain the likelihood that his model's restrictions are consistent with the observed data on the outcome and causal variables. The so-called likelihood ratio tests of restrictions that such hypotheses imply for the structure of economic models enable economists to distinguish between alternative explanations of economic phenomena.

As Sargent recounts, early REH theorists also relied on standard statistical methods to assess the adequacy of their fully predetermined models:

> My recollection is that Bob Lucas and Ed Prescott were initially very enthusiastic about rational expectations econometrics, [which relied on standard statistical methods to test its models.] After all, it simply involved imposing on ourselves the same high standards we had criticized the Keynesians for failing to live up to.

But,

> after about five years of doing likelihood ratio tests on rational expectations models I recall Bob Lucas and Ed Prescott both telling me that those tests were rejecting *too many good models.* (p. 568, emphasis added)

Instead of concluding that REH might not be as useful as originally believed, REH theorists responded to these failures "by lowering [their] standards relative to maximum likelihood," and began to rely on computer simulations—so-called "calibration"—to "match" the selected aspects of their models to the data.[5] According to Sargent, the decision to ignore the often gross inconsistency of the REH models with the data was "a sensible opinion that the time had come to . . . first devote resources to learning how to create a range of *compelling* [REH] equilibrium models to incorporate *interesting mechanisms. We will be careful about estimation in later years* when we have mastered modeling technology" (p. 569, emphasis added).

As intended, lowering the standards of what constitutes an adequate model has obscured many of the inconsistencies of REH models with the data, inevitably increasing the number of REH models that are not rejected. However, this does not necessarily mean that the nonrejected models capture "interesting mechanisms" underpinning real world outcomes. The calibration methodology may suggest to an economist that he has found an explanation of an economic phenomenon. Even so, however persuasive such an explanation may be to his colleagues, according to standard statistical criteria, it may nonetheless be grossly inconsistent with empirical evidence. For example, although NOEM models appear reasonable when viewed through the prism of calibration exercises, they are unable to explain even the most basic features of floating exchange rates, such as their tendency to undergo large and protracted swings away from historical benchmark levels.[6]

5. See Kydland and Prescott (1996) and references therein for an exposition and numerous examples of the calibration methodology.

6. See chapter 7 and references therein. See also chapter 12, where we show that the recent REH model of Barberis, Huang, and Santos (2001), which appears to provide an adequate characterization of equity returns when based on the calibration methodology, is strongly rejected by standard statistical tests when applied to the foreign exchange market.

Sims, who pioneered the search for empirical regularities in macroeconomic data using statistical methods, has commented on the use of the calibration methodology in testing economic models:

> Economics is not physics. Science in general does not consist of formulating theories, testing them against data and accepting or rejecting them. But we can recognize these points without losing sight of the qualitative difference between modern science and classical or medieval natural philosophy: . . . in scientific discourse certain types of apparently persuasive arguments are not legitimate.

Defending "the language of statistical inference" in order "to communicate about the central questions of the discipline," he then argues:

> The fact that economics is not physics does not mean that we should not aim for the same *fundamental* standards for what constitutes legitimate argument; we can insist that the ultimate criterion for judging economic ideas is the degree to which they help us order and summarize the data, that is, it is not legitimate to try to protect attractive theories from the data. (Sims, 1996, pp. 111–12, emphasis added)

Sims's call against protecting "attractive theories from the data" is important for the advancement of economic science. Yet, as Sargent reveals in his interview, the retreat from standard statistical methods was at least in part motivated by nonscientific considerations: the leaders of the REH revolution believed that these methods were rejecting "too many good [fully predetermined] models."

2.1.3. Lost Fundamentals and the Escape from Reality

The field of financial economics provides a particularly clear example of the detrimental effect of the belief that only models that generate sharp predictions are worthy of scientific status. Consider, for example, how the contemporary approach has impeded economists' thinking about the often wide price fluctuations in financial markets, such as those in the foreign exchange market. A key question for exchange rate modelers and policymakers is whether the long swings in exchange rates are linked to movements in fundamental macroeconomic variables, such as interest rates and current account balances.

There is much anecdotal evidence in the popular media, backed up by survey research, that participants in the foreign exchange market pay close attention to fundamental variables in forming their forecasts of future exchange rates. It is obvious, for example, that market participants hang on

every word that central bank officials utter, listening for hints of a change in monetary policy. Similarly, in the year or so preceding the writing of this book, market participants clearly responded to announcements of large and growing U.S. current account deficits by selling the dollar. Because individuals' forecasts drive their behavior in financial markets, we would expect fundamental variables to have considerable influence on exchange rate fluctuations.

But to build models on the foundation of individual rational behavior while remaining faithful to the contemporary approach, conventional exchange rate theorists modeled behavior, on the individual and aggregate levels, with fully predetermined representations.[7] These conventional models were thought to offer *the* way to understand how macroeconomic fundamentals and rational behavior affect the exchange rate.

When the search failed to find an overarching relationship between the exchange rate and macroeconomic fundamentals, conventional economists concluded that swings in exchange rates away from benchmark levels were unconnected to changes in these fundamentals. Obstfeld and Rogoff (2000) have referred to this "anomalous" finding as the "exchange rate disconnect puzzle." Yet this "anomaly" disregards empirical evidence, much of it reported by conventional economists themselves, that, while macroeconomic fundamentals matter for exchange rate movements, the causal mechanism that underpins these movements is temporally unstable: not only do the coefficients of empirical models change from one sub-period of floating to another, but the sets of fundamentals that seem to matter for exchange rates also change.[8] Fully predetermined models cannot account for such structural change: the nature and timing of structural change depend on how market participants revise their forecasting strategies and on unforeseeable changes in the social context.[9]

As we noted in chapter 1, the constraint that economists should consider only explanations that are based on fully predetermined representations

7. Although the contemporary exchange rate models of the 1970s rely on REH, these models do not derive other individual decision variables, such as money demand and consumption, on the basis of an explicit mathematical model of individual choice. Instead, they use ad hoc functions to represent aggregate outcomes whose connection to individual choice is only implicit. Subsequent research added complete representations of individual behavior based on intertemporal optimizing decisions (for example, see Lucas, 1982; Obstfeld and Rogoff, 1996).

8. See chapters 7 and 15.

9. When and how market participants revise their forecasting strategies depend, in part, on historical events, such as the appointment of a new Federal Reserve chairman or the economic policies pursued by elected officials. Recent popular attempts to prespecify fully the nonlinearity between the exchange rate and macroeconomic fundamentals, therefore, are tantamount to a presumption that economists can prespecify fully the timing and consequences of such historical events.

has led many to presume not only that macroeconomic fundamentals are irrelevant to individual forecasting behavior—and thus, to long swings in the exchange rate—but that some or all market participants behave irrationally.

2.1.4. Misinterpreting the Failure of Conventional Theory

There is a breathtaking irony in the unfailing belief that contemporary economics' mechanistic approach is the only way to model change in individual decision making—whether rational or irrational—and aggregate outcomes, as well as to analyze the consequences of government policies. During the three decades in which economists increasingly embraced the contemporary approach, an experiment in economic planning that hoped to replace markets with mechanical instructions based on fully prespecified economic models was heading inexorably to its demise. As Hayek and Knight clearly understood, the dynamism of capitalist economies could not be captured adequately with fully predetermined models that "can be put on a computer and *run*."

Odd but true, the demise of central planning failed to diminish contemporary economists' reliance on fully predetermined representations of economic outcomes. Perhaps this failure is because many of our colleagues find Hayek's argument that central planning is impossible *in principle* to be too far removed from the practice of contemporary analysis of capitalist economies. Like central planners in socialist times, the failure to appreciate Hayek's argument has given rise to a belief among contemporary economists that their approach can be reformed.

Hence the split in contemporary economics between the conventional and behavioral approaches. Behavioral economists—the "reformers"—"built on the premise that mainstream economic *methods* are great, [and] so too are mainstream economic *assumptions*" (Rabin, 2002, p. 658, emphasis in original). We disagree with that premise. The conventional approach cannot be rescued by introducing more "psychological realism" into its models. Representing psychological insights with fully predetermined behavioral models merely replaces one mechanistic representation of individual behavior with another. Moreover, as early modern economists compellingly argued, understanding the social context within which an individual makes decisions is at least as important in modeling her behavior as insights from psychology.

2.2. The Non-Fully Intelligible Individual in Early Modern Economics

Economists have long understood that the way individuals decide on the allocation of resources available to them cannot be fully prespecified. In their

analyses of individual behavior and its implications for market outcomes, Adam Smith, Frank Knight, Friedrich Hayek, and John Maynard Keynes actually went so far as to argue that the economic importance of many decisions made by purposeful individuals stems from the fact that they are not fully intelligible to an outsider.[10]

2.2.1. Individual Preferences, Knowledge, and the Social Context

Prior to extolling the virtues of self-interest in *The Wealth of Nations,* Smith (1759) emphasized in *The Theory of Moral Sentiments* that even a purely selfish individual ranks the consequences of the alternative ways to deploy her resources differently in different social settings. For Smith, the social context, particularly the strength of common norms, was important to understanding the reasons underlying the choices of a self-interested individual regarding how to deploy her resources.[11]

Weber (1897) pursued a similar line of analysis. He devoted considerable attention to the question of whether and, if so, how an outsider could make sense of an individual's behavior, pointing out a number of difficulties in ascertaining whether an individual acts rationally in the sense of pursuing intelligible objectives. He argued, however, that an outsider might be able to gain some understanding of an individual's behavior if institutions, such as bureaucracies and organized religions, shape the values that an individual attaches to her actions.

In his analysis of the emergence of capitalist economic arrangements, Weber (1930) did precisely that, appealing to the social context to explain individual preferences. In his view, the post-Reformation emergence of Protestantism was associated with largely unanticipated and drastic changes in individuals' values and preferences. Protestantism, Weber argued, led individuals to eschew consumption, particularly of luxuries, and to attach great value to work and the accumulation of capital. According to Weber, once one understands Protestant ethics, an outsider can begin to make sense of the objectives and preferences that underpin individual behavior. He went so far as to argue that Protestantism had such a strong influence on values that it dominated all other factors shaping individual preferences: "In this case we are dealing with the connection of the spirit of

10. Frydman and Rapaczynski (1993, 1994) relate Knight's argument that an outsider cannot adequately imitate the creative aspects of self-interested behavior to the central importance of private ownership for economic performance. For econometric evidence on this point in the context of post-communist transition economies, see Frydman et al. (1999).

11. For many reasons why Smith's work on moral philosophy continues to be important for understanding contemporary capitalist arrangements see Dougherty (2002).

modern economic life [including individual preferences and values] with the *rational* ethics of ascetic Protestantism" (Weber, 1958, p. 27, emphasis added).

Weber's view that an outsider might be able to make sense of an individual's behavior by examining the social context has important implications for economic analysis of individual behavior. In the language of contemporary economics, institutions and other social factors help shape both an individual's assessment of the future consequences of her actions (her forecasts), as well as the ranking she attaches to these consequences (her preferences).

Smith and Weber did not attempt to pin down the precise strength of the effect of the social context or altruistic concerns on an individual's ranking of the consequences of alternative decisions for her well-being. They also did not attempt to pin down how the influence of social factors might change over time. The early modern thinkers nonetheless implicitly assumed that such influences arise through a complex process in which the interdependence of individual preferences and the social context unfold over time.

This view suggests that strongly shared values or dominant institutions are not enough for the social context to serve as a prism through which an outsider could make sense of an individual's behavior. For this to be the case, the social context would have to remain relatively stable. After all, Weber would certainly agree that if Protestantism were to diminish in importance, the analysis of Protestant attitudes—for example, attitudes toward saving and investment—would also become less informative for understanding individual behavior. Moreover, what actually matters are not changes in the environment per se, but rather an individual's perception of such changes and her forecasts of their future effects. Thus, an individual's preferences might be intimately linked to her knowledge of the social context within which she acts and her forecasts of its future changes.

Contemporary economists typically do not explicitly incorporate social factors in their representations of individual preferences.[12] This practice may be reasonable in some contexts: for some decision problems, the influence of social factors may not vary across individuals or may remain relatively stable over time. But such considerations as the state of knowledge, its division among individuals, and its unfolding over time cannot be ignored in constructing representations of individual forecasting behavior. Individual forecasts are ipso facto based on an individual's knowledge of the social setting and on how aggregate outcomes unfold over time.

12. For examples of ways in which behavioral economists have fully prespecified the effects of the social context, such as concern for the well-being of others or social norms, on individual preferences, see Charness and Rabin (2002) and references therein.

2.2.2. Individual Forecasting, Knowledge, and the Social Context

The giants of early modern economics cogently argued that no group of individuals, including economists or market participants, could represent fully and adequately the division of knowledge and its evolution in society. They would have been quite surprised, therefore, by the contemporary practice of fully prespecifying an individual's forecasting strategy and its revisions.

Knight (1921, pp. 231–32) argued that forecasts of returns from innovative entrepreneurial activities "deal with situations which are far *too unique* . . . for any sort of statistical tabulations to have any value for guidance." Knight's position points to a fundamental flaw in the contemporary attempt to fully prespecify change: only if creative behavior cannot be fully prespecified can individual creativity be truly indispensable.[13]

Like Knight, Keynes (1921, p. 34) also questioned the efficacy of fully prespecifying an individual's decision making. He noted that one of the main reasons why "[not] all probabilities are measurable" is that it is impossible to prespecify fully all the potential consequences of many economic and social decisions. Consequently, Keynes emphasized that many important individual decisions in market and broader societal settings necessarily involve forecasting strategies that mix rules and models based on probabilistic and statistical calculations with more informal procedures:

> We are merely reminding ourselves that human decisions affecting the future, whether personal or political or economic, *cannot* depend on strict mathematical expectation, since the basis for making such calculations does not exist; and . . . that our *rational selves* [are] choosing between alternatives as best as we are able, calculating where we can, but often falling back for our motive on whim or sentiment or chance. (Keynes, 1936, p. 162, emphasis added)

What is truly insightful about this description of individual decision making is that Keynes explicitly claims that rational individuals would adopt forecasting strategies that, in general, include factors, formal or informal, that cannot be adequately represented by standard probability theory.[14]

13. This interpretation of Knight draws on the analysis of the role of creative activities and private ownership in the postcommunist transition in Grosfeld and Roland (1997) and Frydman et al. (2006b).

14. For an early discussion of the fundamental difficulties in using probability theory to understand economic outcomes in a world of imperfect knowledge, see Peirce (1878). For a modern treatment of Peirce's ideas, see Wible (2007).

2.3. Jettisoning Insights from Early Modern Analysis

The foregoing discussion highlights two fundamental difficulties that severely limit the empirical relevance of contemporary models. The first stems from the uncontroversial observation that creative activities do not completely follow pre-existing rules and procedures. The second difficulty arises from ignoring the dependence of purposeful individual decisions on the relevant social context.

2.3.1. Ruling Out Autonomy in Forecasting Behavior

While, by design, both conventional and behavioral representations disregard the importance of individual creativity, contemporary models that use REH go even further: they rule out an autonomous role for market participants' forecasts in driving market outcomes. REH representations of an individual's forecasting strategy are derivative of an economist's specifications of preferences and constraints. Thus the causal variables that enter such a representation are limited to those that the economist uses in representing the other components of his model. Moreover, the parameters of his representation are restricted to be particular functions of the parameters of his specifications of preferences and constraints and the way that the policy and other causal variables unfold through time.

By design, REH precludes the introduction of so-called "free parameters"—those arising from the autonomous role of market participants' forecasting strategies and their revisions—in explaining outcomes. Proponents of the conventional approach proclaim this tightness as the greatest virtue of REH-based models. All graduate students of economics—and, increasingly, undergraduates, too—are taught that to capture rational, self-interested behavior in a scientific way, they must use REH.

By ruling out any autonomy for forecasting behavior, REH severely impedes economists' ability to develop empirically relevant explanations of market outcomes. For example, in attempting to explain the equity premium—the excess return of stocks over bonds—REH has led economists to an intensive search for alternative specifications of preferences.[15]

15. Mehra and Prescott (1985) and many others have found that the historical average return on stocks over risk-free bonds is too high to be consistent with conventional asset market models. As we discuss in chapter 8, these models are also unable to explain excess returns in the foreign exchange market. Much work has been invested in searching for an alternative specification of preferences that would improve the ability of conventional models to explain the data. For example, see Epstein and Zin (1989, 1991), Constantinides (1990), Campbell and Cochrane (1999), and Barberis et al. (2001).

2.3.2. Reintroducing Autonomous Representations of Individual Behavior

Despite its behavioral implausibility and empirical failures, some behavioral economists continue to rely on REH to represent individual forecasting. By retaining REH, however, they have ruled out any autonomous role for forecasting behavior in driving asset prices.[16]

In contrast to the conventional approach, however, the behavioral approach does not oblige an economist to use REH. This freedom enables a behavioral economist to specify his representation of forecasting behavior as an autonomous component of his model alongside, but not derivative of, his representation of preferences and constraints. Yet their "reformist" objective—"behavioral economics is not meant to be a separate approach [of contemporary economics] in the long run" (Camerer and Loewenstein, 2004, p. 42)—has led behavioral economists to fully prespecify forecasting behavior and thereby contradict one of their core beliefs: economic models require "greater psychological realism."[17]

2.3.3. Individual Behavior and the Social Context

The opportunities, incentives, and institutional arrangements that propel individuals to engage in creative activities vary, often substantially, among different capitalist economies, and even among different markets within the same economy. Capitalist economies are particularly effective in tying an individual's creativity in coping with inherently imperfect knowledge of the unfolding social context to her self-interest. The intrinsic links between creativity, the social context, individual incentives, and the resulting aggregate outcomes are, of course, not unique to capitalist economies. A revealing and relatively transparent historical example of these links comes from the former centrally planned socialist economies. As is well known, these economies repudiated private ownership of productive assets and other institutional arrangements that motivate the economic activities of self-interested individuals in capitalist economies. The results are equally

16. For a widely cited model that mixes preferences motivated by behavioral and conventional considerations with REH, see Barberis et al. (2001). As we discuss in chapter 9, because it disregards some of the key experimental findings of Kahneman and Tversky (1979) and others, even the specification of preferences used in this model suffers from shortcomings on purely behavioral grounds.

17. Some economists, most notably Phelps (2006b), have emphasized that the motivation for being creative stems not only from the profit motive, but from a basic human desire to make a unique personal contribution to the world. For earlier arguments on the importance of creativity for understanding the implications of population growth, see Phelps (1968b).

well known: individuals allocated as little effort as possible to creative activities aimed at producing goods and services in the official, state sector, which resulted in chronic shortages, narrow assortments, and execrable quality. Nevertheless, although individuals in these economies pursued mostly routine activities intended to signal compliance with state directives, there was abundant evidence of individual creativity and entrepreneurship. For example, faced with the lack and poor assortment of goods and services provided by the state sector, individuals innovated by diverting resources from state-controlled firms to private use. In turn, this allocation of individual creative effort away from the state-controlled sector compounded shortages and poor assortments, providing further stimulus to creative activities in the unofficial economy. Despite central planners' efforts, prespecified instructions for individual behavior and aggregate outcomes, such as output targets and quotas in the Soviet Union, could never capture the existing links among individual creativity, social context, individual behavior, and aggregate outcomes. This failure may explain why reading contemporary economics often reminds one of nothing so much as a model of some fictitious planned economy.

3 Flawed Foundations

The Gross Irrationality of "Rational Expectations" and Behavioral Models

> [M]any . . . values toward which experience shows that human action may be oriented . . . often cannot be understood completely.
>
> MAX WEBER,
> *Economy and Society,* p. 5

> Beware of theorists bearing free parameters.
>
> ATTRIBUTED TO ROBERT E. LUCAS, JR., BY THOMAS J. SARGENT,
> in *The Conquest of American Inflation,* p. 73

Modern economics constructs models of market outcomes on the basis of representations of individual decision making. This *methodological individualism* is inherently in conflict with the contemporary economists' insistence that their models should imply sharp predictions. In this chapter, we illustrate this conflict in the context of a simple algebraic model of a market outcome—the equilibrium price that equates the supply and demand for a single good. In specifying the microfoundations of this model, we sketch how conventional and behavioral economists construct their fully predetermined representations of "rational" and "irrational" behavior. This example enables us to show how fully predetermined models presume that participants in real world markets endlessly disregard obvious systematic information in their forecast errors. This presumption of gross irrationality holds true regardless of whether contemporary models are based on REH or behavioral representations of forecasting behavior.

All graduate students of economics—and, increasingly, undergraduates, too—are taught that to capture rational, self-interested behavior in a scientific way, they must use REH. In an attempt to shed light on how economists came to espouse such extreme views, we use our algebraic example to highlight the milestones in the development of contemporary macroeconomics over the past four decades. In doing so, we illustrate Lucas's argument that fully predetermined models, such as non-REH behavioral models, suffer from glaring internal inconsistencies. Our algebraic example also reveals a simple, but important, point: in real world markets, where knowledge about the future course of market outcomes is imperfect, REH,

too, implies that individual market participants ignore obvious systematic information in their forecast errors and thus, to paraphrase Lucas's own argument, REH models are also the "wrong theory."

We also illustrate another fundamental difficulty with using REH to construct the microfoundations of aggregate models. REH instructs an economist to determine his individual and aggregate representations jointly. While an REH theorist would specify individual preferences and constraints autonomously from his aggregate model, his representation of an individual's forecasting behavior is derivative of these other components. Thus, explanations of aggregate outcomes that rely on REH are not based on bona fide microfoundations.

By contrast, a behavioral economist who does not rely on REH specifies his representation of forecasting strategies and their revisions autonomously. This autonomy allows him to complete his specification of individual behavior without having to construct his aggregate model. However, because these models are fully predetermined, they entail, by design, an inconsistency between their representations on the individual and aggregate levels. As Lucas compellingly argued, such models presume gross irrationality on the part of market participants and thus cannot serve as plausible microfoundations of models of aggregate outcomes.

3.1. Conventional and Behavioral Representations of Preferences with Uncertain Outcomes

Economists represent an individual's decision as a choice among alternative ways to deploy her resources. This opportunity set includes options that an economist may be able to infer from his understanding of the social setting (institutional constraints, legislation, and the like). In general, however, some of the options that an individual contemplates are creative, and these are not, ipso facto, completely intelligible to an economist. Nevertheless, the contemporary approach leads an economist to disregard the importance of such alternatives. To represent the opportunity set, an economist specifies a set of outcome variables that represent the consequences of each of these alternatives. For example, the variable that contemporary finance models typically use to represent the consequences of an individual's decisions is the level of her wealth or consumption. In contrast, as we discuss below and in chapter 9, behavioral finance models sometimes use the change in an individual's wealth or consumption relative to a reference level to represent individual behavior.

Table 3.1 Representation of Uncertainty Faced by an Individual

	Option 1		*Option 2*	
Value	y_{t+1}^{11}	y_{t+1}^{12}	y_{t+1}^{21}	y_{t+1}^{22}
Probability	p_{t+1}^{11}	p_{t+1}^{12}	p_{t+1}^{21}	p_{t+1}^{22}

3.1.1. Fully Predetermined Probabilistic Representations of Uncertain Outcomes

The key feature of the contemporary approach is that, to represent uncertainty concerning the future consequences of an individual's decisions, it uses conditional probability distributions whose change over time is fully prespecified. In keeping with the simplicity of our presentation, we suppose that, at some time t, an individual contemplates only two alternative deployments of her resources. We refer to these alternative deployments as "option 1" and "option 2." We also suppose that an economist represents the uncertainty of the future consequences of each of these options by the probability distribution of the future, time $t+1$, values of some outcome variable, y_{t+1} shown in table 3.1.

In the table, y_{t+1}^{ij} ($i = 1, 2$ and $j = 1, 2$) is the jth outcome implied by the ith option to deploy resources and p_{t+1}^{ij} is the probability associated with that outcome, so that $p_{t+1}^{11} + p_{t+1}^{12} = 1$ and $p_{t+1}^{21} + p_{t+1}^{22} = 1$.

3.1.2. Preferences of a "Rational" Individual: The Expected Utility Hypothesis

To determine whether his representation implies that an individual would choose option 1 or 2, an economist must specify a preference ranking for options whose outcomes are uncertain. To this end, economists specify a parametric utility function $u(\cdot)$, which converts the values of an outcome variable into utilities. For example, if y_{t+1} is the level of consumption at $t+1$, $u(y_{t+1})$ associates each potential level of consumption with a utility number; a higher utility number implies a higher level of well-being.

To select and justify the particular parametric functions that they use to represent rational preferences, economists appeal to a set of a priori assumptions that are supposed to characterize how a rational individual chooses among the consequences of her decisions. These assumptions postulate that

an individual's choices among the available options follow a consistent pattern. For example, one of these conditions is that preferences are transitive: if an individual prefers outcome A to outcome B and outcome B to outcome C, then she is assumed to always prefer A to C. Such consistency is thought to capture the way rational individuals choose among the alternatives available to them. The a priori assumptions concerning individual preferences that are adopted by the contemporary approach are thus often thought of as axioms of rational choice.

To represent the choice of a particular option, an economist typically picks the option that yields the greatest utility. However, an economist's representation of the ranking of the outcomes in terms of his parametric utility function is insufficient to determine which option yields the highest utility. Although the utility numbers generated by $u(\cdot)$ imply a ranking of the outcomes if they were certain to occur, the consequences of each of the options are uncertain. Option 1, for example, can result in one of the two outcomes, y_{t+1}^{11} or y_{t+1}^{12}.

Thus, an economist must construct a specification of preferences that ranks options whose outcomes are uncertain. To this end, conventional economists have relied on the expected utility hypothesis (von Neumann and Morgenstern, 1944): the utility of option $i = 1, 2$, which we denote by U^i, is equal to the expected value of the utilities of the outcomes that are associated with the option:

$$U_{t+1}^1 = p_{t+1}^{11} u(y_{t+1}^{11}) + p_{t+1}^{12} u(y_{t+1}^{12}) \tag{3.1}$$

$$U_{t+1}^2 = p_{t+1}^{21} u(y_{t+1}^{21}) + p_{t+1}^{22} u(y_{t+1}^{22}). \tag{3.2}$$

In addition to the axioms of rational choice, conventional representations of preferences are often based on the assumption of risk aversion: an individual is risk averse if replacing an uncertain final wealth by its expected value makes her better off.[1] In the appendix to chapter 6, we make use of a typical functional form for $u(\cdot)$ to represent the well-being of risk-averse rational individuals. That function relates an individual's utility to the level of her consumption.

3.1.3. Behavioral Representations of Preferences: Prospect Theory

Kahneman and Tversky and many others have used experiments to examine the adequacy of the axioms of rational choice and the assumption of risk aversion. They present numerous findings that the a priori assumptions

1. For a recent exposition of the expected utility hypothesis and risk aversion, see Gollier (2001).

underpinning conventional representations of preferences are grossly inconsistent with observed behavior.

The empirical failure of the axioms of rational choice led Kahneman and Tversky (1979) and Tversky and Kahneman (1992) to develop a seminal alternative to modeling individual preferences, dubbed "prospect theory." Prospect theory represents well-being in terms of *changes* in the relevant outcome variables, called losses and gains, relative to a reference value, rather than their postchoice *levels*. Moreover, a utility function based on prospect theory has a different functional form than its conventional risk-averse counterparts.[2]

As with the expected utility hypothesis, prospect theory represents preferences over gambles with weighted sums of utilities of the single outcomes that are implied by an option. Kahneman and Tversky (1979) present experimental evidence that these weights, which they call "decision-weights," are nonlinear functions of probabilities. Prospect theory, therefore, leads to the following counterparts to the expected utilities in equations (3.1) and (3.2), which are called "prospective utilities":

$$V^1_{t+1} = \pi(p^{11}_{t+1})v(y^{11}_{t+1}) + \pi(p^{12}_{t+1})v(y^{12}_{t+1}) \tag{3.3}$$

$$V^2_{t+1} = \pi(p^{21}_{t+1})v(y^{21}_{t+1}) + \pi(p^{22}_{t+1})v(y^{22}_{t+1}), \tag{3.4}$$

where $\pi(\cdot)$ denotes decision weight, $v(\cdot)$ is the prospect–theory based counterpart of $u(\cdot)$ in equations (3.1) and (3.2), and y is the change in the level of wealth or consumption relative to some reference value.

3.2. Self-Interest, Social Context, and Individual Decisions

Equations (3.1)–(3.4) make clear that specifying the utilities of the outcomes, $u(\cdot)$ or $v(\cdot)$, is insufficient to model individual decisions in real world markets. The use of the expected utility hypothesis or prospect theory also requires that an economist represent the way individuals form forecasts of future outcomes and their associated probabilities.

To fix ideas, we consider a typical conventional problem and suppose that the outcome variable y is the quantity of a composite good that an individual consumes.[3] We assume that the decision problem facing an individual is to choose in the current period how much of her current and future real

2. Behavioral economists have developed other alternatives to risk-averse preferences. For example, see Epstein and Zin (1990) and Gul (1991).

3. We focus here on conventional preferences because their use is simpler to present.

income, x_t and X_{t+1}, she would like to spend on the consumption good in the current and future periods. She has the option of saving some of her current income to consume more in the future and less in the present, or she can borrow from her future income to consume more in the present and less in the future. In so doing, she locks in a nominal rate of interest in the current period. Her real rate of interest, then, and thus her future real income, depends on the future price level P_{t+1}. Consequently, in making her decision about y_t and y_{t+1}, the individual must forecast her future real income and the future price level, which are uncertain in the current period t.[4]

This typical setup illustrates the key point emphasized by early modern economists: the assumption of self-interested behavior—that an individual acts to maximize her well-being—is far from sufficient to represent her decision concerning how much of her income she should spend on consumption today. How an individual chooses to deploy her resources and forecast payoffs from those choices depends at least as much on the social context within which she forms forecasts of the future consequences of her decisions as it does on her personal motivations. In our example, the social context is represented by X_{t+1}, P_{t+1} and the aggregate forecast of the future market price. The future real income is sometimes related to the way monetary authorities set money supply. The future equilibrium price depends on decisions of all market participants, which, in turn, stem from their forecasting strategies.

The utility maximization problem in our example gives rise to a representation for quantity demanded by a self-interested individual in the current period t, which we denote by Q_t^{D}. It depends on the price and real income in the current period, p_t and x_t, respectively, as well as on an individual's point forecasts of the price and real income in the future period, $\widehat{P}_{t|t+1}$ and $\widehat{X}_{t|t+1}$, respectively.[5] For simplicity, we use the following linear form to represent quantity demanded at time t:

$$Q_t^{\mathrm{D}}\left(p_t, \widehat{P}_{t|t+1}, x_t, \widehat{X}_{t|t+1}\right) = \alpha_{0_t} + \beta_t\left[\widehat{P}_{t|t+1} - p_t\right] + \gamma_{1_t} x_t + \gamma_{2_t}\widehat{X}_{t|t+1}, \quad (3.5)$$

where $\beta_t > 0$. The variables $\widehat{P}_{t|t+1}$ and p_t enter with coefficients that are opposite in sign, because a rise in the current (future) market price, ceteris paribus, raises (lowers) the return from saving current income. Moreover, because both current and future consumption raise utility, higher current

4. This decision problem of how best to allocate current and future income over an individual's lifetime is common in contemporary macroeconomics. See, for example, Blanchard and Fischer (1989) and Obstfeld and Rogoff (1996). See also chapter 6.

5. The subscript $t|t+1$ indicates that a forecast of the future period's price or income is formed on the basis of the current period's information and the forecasting strategy that an individual uses in that period.

or future income, ceteris paribus, is associated with an increase in current consumption, that is, $\gamma_1 > 0$, and $\gamma_2 > 0$. We note that, in conventional applications, the parameters α_{0_t}, β_t, γ_{1_t}, and γ_{2_t} are functions of the parameters of the utility function $u(\cdot)$ and of the parameters of the processes driving the causal variables (X in the present case) that are used to represent the causal factors.[6]

Contemporary models typically assume that the structure of the utility function is unchanging over time.[7] Moreover, most economists not only fully prespecify change in the process underpinning the causal variables that represent the social context, but they also do not allow for any change in institutions, government policies, the state of knowledge, and other determinants of change in modern societies.

In this chapter, we adopt a particularly simple invariant representation for real income:

$$X_{t+1} = \rho x_t + \mu + \epsilon_{t+1}, \tag{3.6}$$

where ρ and μ are constant parameters, ϵ_{t+1} is an error term whose distribution is unchanging over time, and $E(\epsilon_{t+1}|X_t) = 0$, $E(\epsilon_{t+1}\epsilon_\tau) = 0$ for all t and τ.

We also follow the usual practice and constrain the structure of the utility function to be invariant over time, which results in the following representation for the quantity demanded on the individual level:

$$Q_t^{\mathrm{D}}(p_t, \widehat{P}_{t|t+1}, x_t) = \alpha + \beta\left[\widehat{P}_{t|t+1} - p_t\right] + \gamma x_t, \tag{3.7}$$

where $\alpha = \alpha_0 + \gamma_2\mu$, and $\gamma = \gamma_1 + \gamma_2\rho$.

To represent an individual's forecast of the future market price $\widehat{P}_{t|t+1}$, a contemporary economist either invokes REH, which attributes to individuals the conditional probability distribution of the aggregate outcome implied by his own model, or he makes use of a specification that is motivated by empirical observations on how individuals forecast. Before we consider the implications of both approaches, we need to sketch how an economist would construct his representation of an aggregate outcome—the price that would equate the aggregates of individual demand and supply—on the basis of his representation of individual behavior.

6. See the appendix to chapter 6 for an example of how the parameters in equation (3.5) are related to the parameters of the utility function and the representations of the causal variables.

7. In general, an individual's future preferences differ from her current preferences, as she experiences different levels of outcome variables and as her social context changes. But even if the contemporary representations allow for such change, they fully prespecify it. A widely used representation that fully prespecifies change is the so-called "habit formation" model of preferences. See Constantinides (1990).

3.3. Individual Behavior and Aggregate Outcomes

In real world markets, there is a diversity of preferences, forecasting strategies, and decision rules among participants. In moving from the individual to the aggregate level, such diversity might be important for explaining aggregate outcomes. But even if a contemporary economist allows for diversity among the ways market participants make decisions, he fully prespecifies how this diversity unfolds over time.

Contemporary economists often disregard the importance of heterogeneity: they attempt to capture an aggregate (some weighted sum) of market participants' decisions with a representation of the decisions of an average, so-called "representative," individual. With this assumption, we interpret the specification of an individual's demand in equation (3.7) and her forecast of the next period's price as also representing aggregate (market) demand and the average of the forecasts across market participants.

As with an individuals' demand decision, economists sometimes represent an individual's supply decision to depend on her forecast of the future price.[8] But to simplify our presentation, we assume that the supply decision Q_t^s depends only on the current market price. Moreover, we adopt the following (invariant) linear specification:

$$Q_t^s(p_t) = \delta + \lambda p_t, \tag{3.8}$$

where $\delta \leq 0$ and $\lambda > 0$ are constant parameters. We invoke the representative agent assumption and interpret the specification in equation (3.8) as representing the total supply in the market.

Equating aggregate demand in equation (3.7) with aggregate supply in equation (3.8) yields the following representation for the equilibrium price:

$$p_t = a + b\widehat{P}_{t|t+1} + cx_t, \tag{3.9}$$

where

$$a = \frac{\alpha - \delta}{\lambda + \beta},\ b = \frac{\beta}{\lambda + \beta} > 0, \text{ and } c = \frac{\gamma}{\lambda + \beta} > 0. \tag{3.10}$$

The positive signs of b and c are implied from the considerations highlighted above that set $\beta > 0$, $\gamma > 0$, and $\lambda > 0$.

8. For example, in the seminal paper that introduced rational expectations into economic analysis, Muth (1961) relates supply for period t to the forecast of P_t formed by suppliers at time $t - 1$.

3.4. From Early Modern to Phelps's Microfoundations

To express his model of the market price in equation (3.9) solely in terms of a set of causal variables, an economist must show how a representative agent forms and revises her forecasts. In a series of papers, Phelps (1968a) and Phelps et al. (1970) pioneered an approach to modeling aggregate outcomes on the basis of explicit mathematical microfoundations. Phelps accorded "a crucial [autonomous] role" to individuals' expectations in his explanation of aggregate outcomes (Phelps et al., 1970, p. 5). To capture the idea that individuals do not form the same expectation, he formulated his well-known island model involving informationally isolated labor markets. As he put it, individuals on each island "have to cope ignorant of the future and even much of the present. Isolated and apprehensive, these Pinteresque figures construct expectations of the state of the economy . . . and maximize relative to that *imagined* world" (Phelps, 1970, p. 22, emphasis added).

Because the island model specified that a market participant maximizes utility relative to the world she "imagines," it was not apparent how to represent an individual's forecasting behavior. In the absence of a better alternative, the early "microfounded" models represented individual forecasting behavior with the following fixed error-correcting rule, called "adaptive expectations":[9]

$$\widehat{P}_{t|t+1} - \widehat{P}_{t-1|t} = \lambda \left[\widehat{P}_{t-1|t} - p_t\right]. \tag{3.11}$$

The early proponents of the explicit microfoundations approach used their models to examine particular historical episodes, such as the looming inflation at the end of the 1960s. They did not claim their fixed representations of individual behavior and aggregate outcomes, including the adaptive expectations rule, were general enough to apply to other historical episodes.[10]

3.5. "Rational Expectations": Abandoning the Modern Research Program

Lucas set out to construct a theory of aggregate outcomes that would apply across social settings and for long stretches of time. His early work followed Phelps closely in the use of the island model involving individuals acting in the context of informationally isolated markets. But, in a momentous break,

9. Modeling expectations with an error-correcting mechanism, such as in equation (3.11), was originally proposed by Cagan (1956) and Friedman (1956).

10. See Phelps (1972).

Lucas replaced Phelps's metaphor of "Pinteresque figures" on each island with a mechanistic image of "rational" robots. Lucas's research program instructed economists to search for the "'mechanics' of economic development" (Lucas, 2002, p. 21) based on fully predetermined representations of rational behavior.

3.5.1. Imposing Consistency in Fully Predetermined Models

Lucas argued that economic theory that is based on rational individual behavior should not involve an inconsistency between its representations on the individual and aggregate levels. But, beyond ruling out inconsistency, Lucas and others also believed that economic theory should imply sharp predictions. Aiming to develop economic theory that would meet both of these objectives, Lucas and others embraced REH introduced by Muth (1961). REH imposes complete consistency within a fully predetermined model: it instructs an economist to choose only that representation of individual behavior whose probabilistic predictions coincide exactly with the probabilistic predictions implied by the aggregate model that he himself constructs. REH became the main building block of contemporary models based on putatively rational behavior.

To illustrate how REH removes the inconsistency within a fully predetermined model, we write the REH representation of an individual's forecasting strategy as a linear function in x_t:

$$\widehat{P}_{t|t+1} = \hat{a} + \hat{c}x_t. \tag{3.12}$$

To solve for the REH representation, we determine the coefficients $\hat{a}$ and $\hat{c}$ to be specific functions of the parameters a, b, and c in equation (3.9) so as to ensure the required consistency between the individual and aggregate representations:

$$\widehat{P}^{\text{RE}}_{t|t+1} = E[P^{\text{EM}}_{t+1}|x_t] \text{ for all } x_t, \tag{3.13}$$

where the superscripts RE and EM denote REH representation and the economist's model, respectively.

It readily follows that the solutions for $\hat{a}$ and $\hat{c}$ that imply that equation (3.13) holds for any x_t are:

$$\hat{c} = \frac{c\rho}{1 - b\rho}, \text{ and } \hat{a} = \frac{a(1 - b\rho) + c\mu_0}{(1 - b)(1 - b\rho)}. \tag{3.14}$$

Substituting equations (3.12) and (3.14) into equation (3.9) implies the following REH representation for the aggregate outcome:

$$P^{\text{RE}}_t = a + \frac{ba(1 - b\rho) + bc\mu_0}{(1 - b)\,(1 - b\rho)} + \frac{c(1 - b\rho) + bc\rho}{(1 - b)\,(1 - b\rho)}x_t. \tag{3.15}$$