Introduction

Economic growth, inflation, and unemployment are the "big three" topics of macroeconomics. Explicitly embodied in legislation in the United States and other countries are the goals of achieving rapid economic growth, a low rate of inflation, and a low rate of unemployment. When I teach lecture classes on elementary or intermediate macroeconomics to large auditoriums full of fresh-faced undergraduates, the semester begins with simple examples to show how much better off they will be in thirty years with fast rather than slow economic growth, how rapid inflation could erode their savings and that of their parents, and how much easier it will be to find a job for the summer or after graduation if the nation's overall unemployment rate is low rather than high.

THE LAY OF THE LAND

This book, then, is about the big topics of macroeconomics. It is divided into four parts, of which the first is undeniably the most important. Why was American economic growth faster between 1913 and 1972 than before or after? What caused productivity growth to slow down after 1972 and accelerate after 1995? In my view the driving forces of twentieth-century growth were the "great inventions" of the late nineteenth century, especially electricity and the internal combustion engine. The central theme of Part One is the role of these inventions in creating faster growth early in the twentieth century and then, as their influence waned around 1970, slower economic growth. Along the way, Part One asks whether the new internet economy of the late 1990s measures up to the great inventions, how we disentangle the role of technical progress from raw data on output and inputs, and how America's famous job machine that created 23 million new jobs between 1992 and 2000 may also be, in a subtle way, a source of slow productivity growth.

Part Two asks why productivity growth fluctuates over shorter intervals of a decade or so. While this question may seem to be of less than cosmic importance, its interpretation turned out, somewhat surprisingly, to be the key issue on which debates about macroeconomic doctrine were centered in the past two decades. This question also must be addressed in trying to figure out how much of the

post-1995 U.S. productivity growth revival was "structural" and how much was a temporary cyclical phenomenon caused by an unsustainable burst of output growth, especially in 1999–2000.

Then Part Three examines the theoretical relationship between output, inflation, and unemployment. Why cannot the central bank (Federal Reserve or "Fed") keep interest rates so low that the unemployment rate eventually declines to zero? The usual answer is that the Fed fears an acceleration of inflation if the unemployment rate is allowed to drop too low. But that answer presupposes that there is a negative "trade-off" between inflation and unemployment, a proposition that appeared to be effectively demolished in the late 1960s and early 1970s by two winners of the Nobel Prize in economics, Milton Friedman and independently by the equally perceptive Edmund S. Phelps.

The papers in Part Three constitute one of the most exciting developments in postwar macroeconomics, the introduction of a symmetric analysis of supply and demand shocks to replace the old-fashioned Keynesian analysis that was limited to the role of demand fluctuations. The traditional sources of demand shocks, investment cycles, wars, monetary policy, and fiscal policy boosted demand relative to supply and caused the same response that occurs in the elementary microanalysis of the supply and demand for corn or furniture – spurred by a positive demand shock, both aggregate output and the aggregate price level rise.

Starting in the mid-1970s, the analysis of business cycles was broadened to give an equal starring role to supply shocks, like the sharp increase in the price of farm products that occurred in 1972–3 or of oil in 1973–5. An adverse aggregate supply shock operates just like a crop failure in microeconomics – output declines but prices increase, moving in the opposite direction. The papers of Part Three show that when prices and wages in the economy outside of the "shocked" sector are slow to adjust (or "sticky"), the adverse supply shock creates a "macroeconomic externality." The total loss of output to the entire economy can be many times as large as the size of the crop shortfall or reduction in oil supplies that sets off the reaction. The teaching of macroeconomics today is much like it was in 1980. It was the five years between 1975 and 1980 that witnessed a revolutionary change in the development of the symmetric supply–demand analysis, as shown in Part Three, as well as its instantaneous introduction into undergraduate macro textbooks.

Then Part Four provides empirical evidence to support the theories of Part Three. Did those supply shocks actually cause the "twin peaks" of unemployment and inflation in the 1970s? Why was inflation so low in the late 1990s? Why was unemployment in the United States in the late 1980s and 1990s so much lower than in most of the large European countries? Was the American economic miracle of the late1990s due to good luck, the emergence of a new paradigm that loosened previous constraints, and can it continue?

There is no reason for anyone to be interested in a collection of papers. Many authors of collected essays choose the papers and publish them, without the connecting threads that show how they emerged, whether they are still

interesting or valid, and why papers on the same topic written two decades apart reach the same or different conclusions. I take it as the job of the introductions in this book to make these papers interesting – to connect them to important themes and to each other and also to help the next generation of economists see where these ideas came from.

Each of the four parts of the book includes a substantive introduction to the main issues that, by and large, can be read independently of the papers themselves. The introduction to Part One is a new essay on twentieth-century growth that attempts to link together the themes of all six papers included there. The introductions to the other three parts are shorter and less ambitious but nevertheless provide a useful overview showing how the topic developed over the past twenty or thirty years.

HOW THE IDEAS AND PAPERS EMERGED

While the introductions are largely substantive, they do include a few remarks about the sources of ideas. Young economists may be interested in the circumstances that led to some of these papers. In many cases, an easy summary is that "events precede ideas." This is most obvious in the role of the inflation of the late 1960s and the inflationary recession of 1974–5 in revealing the inadequacy of then-current economic paradigms and pressing us to figure out what was wrong and how to fix it.

The initial catalyst for my interest in economic growth came during a two-year stay at Oxford, England, in 1962–4. There my previous interest in the microeconomic topic of industrial organization soon faded away. Britain in that era had finally recovered from the strains of postwar rationing and currency non-convertibility, but otherwise seemed to this outsider to be an economic basket case. The combative unions of "I'm All Right, Jack" held sway, the standard of living was far below that in the United States and had recently been overtaken by rapid recoveries in France and Germany, and history presented a dismal record in which the level of British productivity barely grew at all from 1895 to 1938. Clearly, the siren blared out that differences in the economic growth experience across nations and historical eras were the topic to study, and that is still true today. Perhaps the most important single piece of reading to which I was exposed at Oxford was Edward Denison's seminal 1962 study of the sources of growth (cited in Chapter Two), and especially his imaginative translation of dry data on educational attainment into implications for the sources of growth.

As described in the Introduction to Part One, much of my research on economic growth, and especially Chapter Two, can be traced to a summer job I had at MIT in 1965 as a new graduate student. Puzzles in then newly developed macro data on the history of U.S. economic growth led to my Ph.D. dissertation and to my interest in measurement errors of all types, but especially in important measurement problems that were big enough to skew the historical record over decades. From the beginning, my career developed along two parallel tracks. The first was to pursue these measurement puzzles and to make a

serious investment in creating new data, particularly on the prices of investment goods, both structures and equipment. Much of this work could be likened to working as a medieval scribe in the library, converting data from the Sears catalogue, *Consumer Reports*, and other sources into alternative price indexes for investment goods. This work took almost twenty years before a book emerged (see the references to Chapter One). Yet along the way there were lots of ideas that involved substantive problems rather than measurement, and these were the sources for the six papers in Part One.

The second research track was motivated by the central debate in macroeconomics that was boiling as I moved in 1968 from a pleasant life as graduate student at MIT to the intellectual cauldron of the University of Chicago. Just as inflation was accelerating in the late 1960s, Chicago's most famous macroeconomist, Milton Friedman, had delivered his perfectly timed presidential address launching the natural rate hypothesis, contending that in the long run inflation is independent of the unemployment rate. As a new assistant professor, I had to plunge into the hot water and figure out how, if at all, to reconcile my Keynesian MIT training with Friedman's distinction between the negatively sloped short-run Phillips curve (based on expectational errors) and the vertical long-run Phillips curve. From then on, my major topic in time-series macroeconomics was the inflation–unemployment tradeoff, which created the papers in Parts Three and Four of this book.

My research on economic growth could have been carried out almost anywhere and did not require a particular university location. But the combination of graduate school at MIT and a first job at Chicago were crucial in making possible the research on the inflation–unemployment tradeoff. Being at Chicago with an MIT education was like watching two sticks rubbing together. The flame soon ignited, especially when I found that my students were teaching me more in my first Chicago graduate class than I was teaching them. But a second crucial piece of luck in timing and location came when Arthur M. Okun had the idea to start the Brookings Panel on Economic Activity as a triannual series of meetings, with the papers and discussions to be published almost immediately, within three months of the meeting rather than the two-year lag typical of conference volumes then and now.[1]

The original format of the Brookings panel (BPEA) was to have a core group of young macroeconomists who would become the resident expert on a particular topic, writing a major paper every year or two with "sector reports" in-between. While I was originally requested to cover consumption expenditures, I asked instead to handle the Phillips curve trade-off as my main topic. As a result, I wound up publishing no less than seven papers, most of them in print within four months after they were written, within the brief period 1970–3. Some

[1] The Brookings Panel held three meetings per year during 1970 to 1978 and two meetings per year since then, for a total of seventy-five meetings through the end of 2002. I have been to seventy-four of the seventy-five meetings, and coorganizer George Perry has been at every meeting.

of these papers had a lasting influence but none of them are included in this volume. Events, particularly the ongoing acceleration of inflation in 1969–70, the Nixon price controls, and the supply shocks of 1972–5, happened so fast that most of those papers were subject to rapid obsolescence. Nevertheless, the rapid publication schedule of BPEA gave me a rare and perhaps unfair advantage in always having the last word on the latest puzzle involving inflation. Another advantage conferred by BPEA also was of immeasurable value, the chance to interact with the top macroeconomists at the regular BPEA meetings, be exposed to the latest data that we could not explain, and start figuring out what to tackle next. In that sense, my two-track career was schizophrenic in nature, with the measurement research resembling a traditional cloistered ivory-tower experience, whereas the inflation research centered on the BPEA directed my communication and energy outside the confines of the local university campus.

Some acclaimed academics have succeeded by moving from topic to topic as the occasion emerged, often in collaboration with coauthors.[2] My professional hero, next to Bob Solow, was the late Zvi Griliches, whose prolific research career was exactly the opposite. Zvi "owned" the production function as a topic – when he moved to a new subtopic, whether hedonic price deflators, capital measurement, ability and human capital, patents, or research and development, it was part of a broad lifelong research plan to dig away at the outstanding puzzles related to the production process. My approach was similar but in a narrower area. I kept coming back to the same topics, whether price, output, and input measurement, cyclical productivity fluctuations, the Phillips curve, or aggregate supply shocks, both because macroeconomics is constantly creating new puzzles and new data, but also because I felt a responsibility to see whether my old theories and empirical results still worked. If they did not, I wanted to figure out why. There is no way to give advice to younger scholars on these different research strategies. To stick to the same topics over decades, those topics had better be important and longlasting in relevance. To move from topic to topic and have striking insights on a variety of unrelated topics, exactly the opposite of my own research approach, you had better be very smart.

CRITERIA FOR INCLUSION AND OMISSIONS

Selecting the seventeen papers for this volume was painful. The publisher set a page limit, which ruled out numerous long empirical time-series papers. Page limits also ruled out several long survey articles, including two from the *Journal of Economic Literature* on the sources of price and wage rigidities.

This volume includes papers that fit together tightly into the four themes, and three of the seventeen are papers have not been published before. Many of my other papers are ruled out on the basis of length or topics that do not fit within

[2] In Michael Szenberg's fascinating book *Passion and Craft: Economists at Work* (Michigan, 1998), I was struck at the ease with which such subjects as Greg Mankiw, Avinash Dixit, and others could move from topic to topic, reflecting their innate brilliance and analytical ability.

the framework of this volume, whether on corporate tax shifting, government investment during World War II, or remeasuring the volatility of U.S. GDP before 1929. Also, I have leaned toward including papers from lesser-known sources rather than my articles from well-known journals like the *American Economic Review* or the *Journal of Political Economy*.

The topics of the four parts of this book have the virtue that all are still relevant as the economy inches its way into the twenty-first century. We wonder when reading Part One whether productivity growth in the next decade will look more like the ebullient five years after 1995 or the dismal twenty-three years before 1995. Part Two makes us wait eagerly for the next quarterly productivity report, to see in retrospect how much of the post-1995 revival was cyclical rather than structural, and whether the 2001 recession and 2002–3 recovery adhered to old cyclical patterns or created a new one. Parts Three and Four help us to understand explicitly why inflation was so quiescent in the late 1990s and why, for the first time in decades, there was no sharp upward spike of interest rates as previously had been necessary to quell a significant acceleration of inflation. Since tight monetary policy did not cause the 2001 recession, its causes are to be found elsewhere, especially in the collapse of the New Economy investment boom.

A final word is needed on the ground rules for reproducing these papers. All of them have been newly typeset. Every paper is reproduced exactly in its original form, including any forecasts that were made long ago, no matter whether they turned out to be right or wrong (the introductions provide hindsight retrospective on several of these forecasts). No changes were made except to provide the final published reference for items of the reference lists that were originally "forthcoming," i.e., not yet published at the time the paper went to press, and also to clarify references to dates (e.g., "four years ago" in a 1982 paper is changed here to "in 1978").

PART ONE

THE HISTORY, THEORY, AND MEASUREMENT OF PRODUCTIVITY GROWTH

INTRODUCTION

STUDYING GROWTH AT THE FRONTIER

The gains to human welfare resulting from even minor increases in the rate of economic growth are enormous. In the oft-quoted words of Nobel Prize-winning University of Chicago economist Robert E. Lucas, Jr., "the consequences for human welfare are simply staggering. Once one starts thinking about them, it is hard to think of anything else" (Lucas, 1988, p. 5). Changes in the rate of economic growth in the history of the United States over long intervals have been sufficiently large to create the enormous differences that Lucas was thinking about. A slow growth rate of income per person, say 1 percent per year, causes the standard of living to double in seventy years, roughly every three generations. But a more rapid growth rate, say 3 percent per year, will cause the standard of living to double in a mere twenty-three years, making each generation twice as well off as that of its parents.

The consequences of differences in growth rates are most obvious in comparing rich and poor countries. In the past century, there has been little if any improvement in standards of living in some African countries and among rural populations in some Asian and Latin American nations. Yet the standard of living in the United States, at least as conventionally measured, has increased by a factor of about 8, and that of Japan by a factor of perhaps 25. While the average American is choosing among home entertainment systems and sport-utility vehicles, living in a suburb with clean water, reliable electricity, and constant propaganda to buy broadband cable access for the one or more family personal computers, life in the poorest countries involves a very different set of choices about obtaining sufficient food to survive and avoiding deadly diseases for which no medicines or cures are readily available.

My work on productivity and economic growth has focused mainly on the United States over the period since 1870, and particularly on the explanation of accelerations and decelerations in the rate of productivity growth. While a study of epochs in U.S. history may perhaps seem of less cosmic importance than asking why some nations are rich and some are poor, the U.S. experience has strong appeal as a research topic. First, starting sometime in the late nineteenth

century, when the per-capita income of the United States surpassed that of the United Kingdom, the level of productivity in the United States has defined the "frontier" of what is possible for the world's developed nations. Accelerations and decelerations in the growth rate at the frontier call out for explanations – are the sources of speed up and slow down unique to the United States, in which case Japan and the richer nations of Europe might overtake the U.S. frontier position, or are those sources of deceleration universal and without implication for the ranking of income per capita across nations?

Second, the focus on U.S. history helps to establish links between the pace of economic growth and the timing of the invention of new products and technologies.[1] More than in other countries, the development of the United States was spurred by economic opportunities based on a high ratio of land and natural resources to labor – these outsized potential returns not only rewarded investment in existing technologies, but they also fostered prolific inventions from McCormick's reaper to Edison's electric light and motor – and in turn the inventions created further opportunities for profitable investment. The "Great Inventions" emphasized in the first essay in this book were disproportionately American creations. The United States was directly involved in the invention of electricity through the work of Thomas A. Edison, as well as Alexander Graham Bell's telephone, George Eastman's roll film, and Lewis Waterman's fountain pen. The Germans Nikolaus Otto and Karl Benz played the major role in the development of the internal combustion engine and automobile. Nevertheless, America soon dominated the development and exploitation of motor transport, not to mention the Wright Brothers' first flight in 1903 that led two decades later to commercial air transport. So completely did the United States take over the development of the internal combustion engine that in 1929 the United States had 26 million registered automobiles and almost 5 million registered trucks, doubtless more than the rest of the world combined.[2]

Third, the lightbulb of guidance provided by economic data is not "turned off" by wartime destruction, as in Japan and most European countries. Students of U.S. economic growth do not have to be concerned with the destruction of substantial portions of the capital stock in World Wars I and II, and the loss of American lives and skills in both wars was sufficiently minor as to be safely ignored.[3] Nevertheless, as we shall see, the combined experiences of the Great

[1] My two favorite books on the interplay between inventions and economic growth, both written from a global rather than U.S. perspective, are Mokyr (1990) and Rosenberg-Birdzell (1986).

[2] Perhaps most important was that, judging from 1941 production figures, at the outbreak of World War II the U.S. had the capacity to manufacture 4 million automobiles and 1 million trucks, swamping the capacity of all other nations in the world combined. In one interpretation, World War II was a "war of engines" and, thanks to Lend-Lease, the Russian Army rode to victory in over 350,000 American trucks, which if delivered over the three years 1942–4 represented a mere 10 percent of American productive capacity to manufacture trucks (Overy, 1995, p. 214).

[3] U.S. military deaths in World War II were only 0.3 percent of the 1940 population of 131.7 million.

Depression and World War II create significant distortions in U.S. data that are important enough to alter the long-term growth record in a misleading way.

LABOR PRODUCTIVITY AND MULTIFACTOR PRODUCTIVITY

Whether talking about differences in economic growth across countries or across eras within one country, it matters whether one is talking about "average labor productivity" (ALP) or "multifactor productivity" (MFP). This distinction turns out to be central to an understanding of U.S. economic growth and of the role of technical change in creating that growth. ALP is output per labor hour, whereas MFP is output divided by a weighted average of labor input and capital input – the weights are usually the income shares of labor and capital, say 75 and 25 percent, respectively. The latter concept (MFP) almost always registers slower growth than the former (ALP) – the difference between them is the effect of "capital deepening," the boost to ALP that comes from the fact that capital input almost always grows faster than labor input.[4]

Another equivalent definition is that the growth rate of MFP is a weighted average of the growth in labor productivity or ALP (weighted by labor's share) and the growth rate of "capital productivity," that is, the growth rate of output minus the growth rate of capital input. This turns out to be crucial for understanding the "big wave" interpretation of U.S. economic growth in the twentieth century examined in Chapter Two. If labor productivity growth is steady but there is a big bulge in the growth of capital productivity, then MFP growth will soar relative to that of labor productivity. That is part of the story of the U.S. economy in the mid-twentieth century and, as we shall see below, much of the bulge in the growth of capital productivity reflects measurement errors rather than historical fact.[5] MFP is now often called "Solow's residual" as a tribute to Robert M. Solow's pathbreaking (1957) work that provided a

[4] An example clarifies these concepts. In the long run output growth and capital growth tend to be the same, say 4 percent per year. Growth in population or labor input is, say, 1 percent per year. Then ALP (average labor product) growth would be 4 minus 1, or 3 percent per year. If, as stated in the text, the weight on capital input is 25 percent, then input growth would be 1.75 percent (25 percent times capital growth of 4 percent plus 75 percent plus labor growth of 1 percent). MFP growth is 4 minus 1.75, or 2.25 percent per year. The "capital deepening" effect is the capital weight of 25 percent times the difference between capital and labor growth (4 – 1), or 0.75 percent. Thus MFP growth (2.25) equals ALP growth (3.0) minus the capital deepening effect (0.75).

[5] Continuing with the same example as in the previous footnote, MFP growth of 2.25 percent is a weighted average of ALP growth of 3 percent (4 minus 1) and capital productivity of zero percent (4 minus 4) with weights of 75 and 25 percent, respectively. Imagine that capital growth dropped from 4 to 0 while output growth remained at 4 and labor input growth remained at 1. MFP growth would accelerate from 2.25 to 3.25. MFP growth would exceed ALP growth, because the capital deepening effect is negative. To the extent that the decline in capital growth was a measurement error, so would be the acceleration of MFP growth in the opposite direction.

rigorous theoretical rationale for the near-universal practice since then of measuring the elasticity of output to changes of inputs by the income share of that input.[6]

INTERPRETING THE ACCELERATIONS AND SLOWDOWN

The recorded annals of U.S. productivity growth since the late nineteenth century have been marked by several notable turning points, which are examined and interpreted in several of the papers in the first section of this book. When did these turning points occur? The easiest way to remember the chronology of productivity growth is "slow" 1870–1913, "fast" 1913–72, "slow" 1972–95, and a recovery of debatable size and duration to somewhere between "fast" and "slow" during 1995–2000. These long intervals are chosen to span several business cycles and, in particular, to leap over the economic distortions of the Great Depression and World War II by treating 1928–1950 as a single period within the longer 1913–72 interval.

A marked acceleration distinguished the post-World War I half century from the four decades prior to World War I and was one of the most important features of the historical record, noted by Solomon Fabricant in his introduction to Kendrick's pathbreaking 1961 volume, that for the first time had set out the historical record in annual data.[7] It took much less time for the post-1972 productivity growth slowdown to be recognized, and indeed it was analyzed by William Nordhaus (1972) just as it was beginning by today's standard chronology.[8] The post-1995 productivity growth revival was preannounced by the perceptive economics staff of *Business Week* just as it was commencing,[9] but several years elapsed before it was recognized by academic economists.

The first two chapters of this book, Chapter One on the new economy and the "great inventions," and Chapter Two on interpreting the "big wave" of U.S. economic growth, directly address the three turning points in U.S. growth history in 1913, 1972, and 1995. They share a common theme, that the "great inventions" of the late nineteenth century (especially electricity and the internal combustion engine) were so powerful in their economic influence that they propelled a fifty-year-long boom in productivity growth between 1913 and 1972, and that the post-1972 slowdown could be interpreted as caused by diminishing returns,

[6] A short, revealing, and unique history of the "residual" is provided by Griliches (2000, Chapter One).

[7] "The change in trend that came after World War I is one of the most interesting facts before us. There is little question about it. . . . Some readers of the charts might prefer to see in them not a sharp alteration of trend, but rather a gradual speeding up of the rate of growth over the period as a whole. The latter reading is not entirely out of the question, but it seems to fit the facts less well than the former" (Solomon Fabricant, p. xliii in the introduction to Kendrick, 1961).

[8] Nordhaus (1972) examined the slowdown that occurred in the late 1960s relative to the preceding portion of the postwar era.

[9] See the cover banner in *Business Week*, October 9, 1995, "Productivity to the Rescue."