Rochester Center for

Economic Research

Capital Accumulation in the Theory of Long Run Growth

Romer, Paul M.

Working Paper No. 123
March 1988.

University of
Rochester

Capital Accumulation in the Theory of Long Run Growth

Paul M. Romer

Rochester Center for Economic Research
Working Paper No. 123

Capital Accumulation in the Theory of Long Run Growth*

Paul M. Romer

Working Paper No. 123

March 1988

*In the preparation of this chapter, I received helpful comments from Robert Barro, Michael Dotsey, Marvin Goodfriend, Robert King, Sergio Rebelo, Danyang Xie, and from students at the University of Rochester and Chicago. Support from National Science Foundation grant #SES8618325 is gratefully acknowledged. Forthcoming in Modern Business Cycle Theory, (ed.) Robert J. Barro (Cambridge: Harvard University Press).

After more than a decade of quiescence, growth theory may once again be entering a period of ferment. By the end of the 1960's, there was a general consensus about how growth should be modeled, and this formed the basis for a great deal of empirical work in growth accounting. This agreeable state of affairs was achieved by significantly narrowing the range of the questions that growth theory was expected to address. From the point of view of early theorists, the two most interesting questions about growth were dropped from consideration. How can one reconcile extraordinary, continuing, increases in average per capita income with the notion of diminishing returns? What determines the rate of growth of the population? Attention is once again turning to these issues.

During the 1960's the consensus view on these bothersome questions came to be that each should be assigned an exogenous, exponential trend term. Then the economic analysis of the other features of an economy could proceed. From a practical point of view, this finesse of the problems of endogenous per capita income growth and endogenous population growth was probably useful. Important theoretical progress in the understanding of dynamic models was made precisely because the most difficult questions about growth were set aside. But because it ignored the fundamental questions and concentrated on abstraction and formalism, growth theory came increasingly to be viewed as a sterile exercise. From the point of view of policy advice, growth theory had little to offer. In models with exogenous technological change and exogenous population growth, it never really mattered what the government did. Partly in reaction, development branched off as a separate endeavor, designed to offer the direct policy advice that growth theory could not.

The irony underlying the disrepute that growth theory fell into is that economists seem to have missed the self–referential nature of the theory. Economics, like any science, has a two sector technology. The final output good of economics, produced in one sector, is correct answers to questions that non–economists care about. A separate investment sector produces the intellectual capital that is the key input into the final output sector.

Development for the most part took 1960's vintage intellectual capital and has been producing policy advice with it ever since. In contrast, growth theory gave up any pretense of having anything to say about questions that a non-economist might care about and concentrated on intellectual capital accumulation.

This capital accumulation now shows signs of a significant payoff. One of the major themes of this essay is that the substantive contribution of growth theory has so far been quite small but the methodological impact has been far reaching and fundamental. The methodological advances have had their greatest impact in macroeconomics, where they can truly be said to have revolutionized accepted practice. To cite just one example from the other chapters in this volume, in the years between 1970 and 1980, the discussion of the theory of aggregate consumption moved from a point where it would have been impolite to mention Euler equations, to a point where it was impossible to carry on a discussion without them. By itself, this methodological impact has long justified attention to developments in this area, but now may be an especially good time to tune into growth theory. The second theme developed here is that the tools have developed to the point where growth theory is on the verge of having something interesting to say about growth.

To develop these themes, much of this chapter is devoted to a self–contained description of the methods used to study dynamic models. This description is a synthesis, not a survey.[1] It takes advantage of hindsight, and interprets all of these methods in the context of a unifying result from the mathematical theory of convex analysis. This result, the abstract Kuhn-Tucker theorem, lies at the very heart of all equilibrium models of growth, and for that matter of all of general equilibrium theory.

The plan of this chapter is as follows. Section II sets the stage with a description of some basic facts about growth that a complete theory should be able to address. Section

[1]Anyone interested in a more detailed overview of the literature on growth up through the 1970's can consult Hahn and Matthews (1964), Stiglitz and Uzawa (1969), Burmeister and Dobell (1970), Solow (1970), Hahn (1972), or Jones (1975).

III, the bulk of the chapter, lays out the tools that an economist needs to appreciate modern theories of economic dynamics, using illustrations along the way from models of growth. Section IV gives a description of the theories developed over the last three or four years that are the basis for the claim that we are now past a turning point and that theories of growth are returning to the fundamental questions in the field.

Section II: Data

In an influential article on growth written in 1961, Nicholas Kaldor stated his view that a theorist ought to start with a summary of the facts that are relevant for the problem of interest. This summary should be "stylized", he claimed, concentrating on broad tendencies. One could then construct hypotheses to explain these stylized facts. In the formative stages of a body of theory, this kind of informal treatment of the data can be quite useful, for without stylized facts to aim for, theorists would be shooting in the dark. When Kaldor wrote, the basic elements of a theory of growth seemed to be up for grabs in a developing debate between Cambridge, England and Cambridge, Massachusetts, and the facts he set out became the target for economists of on both sides.

If, as is claimed in the introduction, growth is entering a similar phase where the basic questions about growth are being re-examined, it may be useful to review and update Kaldor's list of facts. To do this in a way that does not bias the outcome, it is important not only to make sure that the facts have some connection with measured data, but also that the list be as inclusive as possible. Different theories can often explain different subsets of the facts. For example, depending on the set of countries one looks at, one can conclude either that per capita income across countries is converging rapidly or that no tendency towards convergence is present. (This point is discussed in greater detail below.) As another example, Solow (1970) observes that his model of growth can explain five of the six stylized facts described by Kaldor, but acknowledges that the sixth–the wide dispersion in growth rates across countries—is something of problem for him. Subsequent advocates of the neoclassical model have sometimes been less forthcoming, listing only five facts that a model of growth should explain.

These are Kaldor's six stylized facts:

1) Output per worker shows continuing growth "with no tendency for a falling rate of growth of productivity." (Emphasis in the original.)

2) Capital per worker shows continuing growth.

3) The rate of return on capital is steady.

4) The capital—output ratio is steady.

5) Labor and capital receive constant shares of total income.

6) There are wide differences in the rate of growth of productivity across countries.

It is readily seen that these statements are not all independent. Let Y, K, and L represent total output, capital and labor respectively. Let r denote the return on capital. If Y/L is growing and Y/K is constant, K/L must also be growing. Thus, fact 2) follows from facts 1) and 4). If Y/K is constant and rK/Y is constant, then r must also be constant. Thus, 4) and 5) imply 3). With no loss in generality, one can concentrate on 1), 4), 5) and 6). Based on the kind of data exhibited below, 1), 4), and 6) may still be reasonable stylized characterizations of the data. On the other hand, there is some evidence of a long run trend in factor shares.

In interests of being inclusive, this section adds five other features of the data that seem to stand out:

7) In cross-section, the mean growth rate shows no variation with the level of per capita income, while the variance in growth rates declines with the level of income.

8) Growth in the volume of trade is positively correlated with growth in output.

9) Population growth rates are negatively correlated with the level of income.
10) The rate of growth of factor inputs is not large enough to explain the rate of growth of output; i.e. growth accounting always finds a residual.
11) Both skilled and unskilled labor tends to migrate towards high income countries.

Observation 7) can be discerned from the wider array of data that are now available. Observation 8) has been noted in discussions of export lead development and 9) has been the focus of considerable study among demographers. However, since formal theories of growth have until recently been silent on the determinants of population growth and of international trade, they have not been considered relevant parts of the target that growth theories should aim at. Given the recent development of theoretical tools and models that can address these issues, this exclusion seems increasingly unjustified.

Observations 10) and 11) are not facts about growth that a statistician hired to report on growth would identify. The game here is to set facts that theories can aim for, but the theories themselves influence what we perceive to be relevant facts. The growth accounting result, arising ultimately from the work of Robert Solow, indicates how deeply the neoclassical model based on constant returns to scale has influenced the way we look at the world. Observation 11) concerning migration may not appear to have any direct bearing on theories of income growth; however, recent theoretical work by Robert Lucas suggests that this may be a crucial piece of evidence in distinguishing between theories of growth based on constant returns to scale and those based on increasing returns.

Tables 1 and 2 bear on Kaldor's first observation. Viewed from a long-run perspective, there is no question that cumulative growth in output per worker has been truly remarkable, and that the rate of growth increased over a long period of time. Table 1, taken from Maddison (1982) identifies the country with the highest level of output per hour worked during different historical epochs and estimates the rate of productivity growth for that country. The trend is clear, but the magnitudes may need some amplification. Using the fact that the natural logarithm of 2 is 0.69, it follows that a productivity growth rate of 2.3% per year for the United States leads to a doubling of output per worker every 30 years. Table 2 (using data from the same source), shows that growth rates of this magnitude are not unique to the United States. It lists the factor by which output per hour worked increased over the period 1870 to 1979 for 16 developed countries. These magnitudes speak for themselves. Table 2 also introduces symbols used to identify different countries in subsequent figures.

Figures 1 and 2 show the behavior of labor productivity in the United States in more detail and addresses the perception that growth rates are falling rather than increasing. Figure 1 shows the annual rate of change of output per man-hour for the private business sector in the post-war era. (Data are from the Bureau of Labor Statistics, published in the Monthly Labor Review.) Careful examination reveals a lower average rate of labor productivity growth in the period since 1969. This reduction in the rate of productivity growth has been the source of much concern and attention, and is indicative of the kind of evidence that has lead to concern that growth rates are slowing. Because these data are sensitive to business cycle variation that does not seem to impinge uniformly on the two halves of the sample, it is not clear that one should yet draw strong inferences about secular trends.

The data points in Figure 2 marked with boxes show the long run behavior of labor productivity.[2] For comparison, the figure also plots the growth of output per capita,

[2]The data for productivity used here attempt to track the private business sector data used in the post war sample as closely as possible. For the post war period, the source is the same as for Figure 1. From 1890 to 1950 data for the private business sector are taken from Kendrick (1961). Prior to 1890, the basic data on output come from the work of Robert Gallman, but must be retrieved from three different sources: summaries given in Gallman (1966), augmented with raw data from Gallman's work sheets that is published in Friedman and Schwartz (1982), and an estimate for growth during the decade of the 1860's reported in Kuznets (1971). Data on employment are from Lebergott (1966). In this early period, average hours worked per employee are assumed to have remained constant. To the extent that average hours in this period fell, as they did in subsequent periods, the reported rates of productivity growth are too high. Population data for the per capita series are taken from Maddison (1982).

with point denoted by plus signs. Year to year variation is smoothed by taking 20 year averages. Productivity alone does not tell the whole story. For example, the period 1919 to 1939 shows relatively strong growth in output per hour worked that masks a fall in employment and hours worked during the depression. Similarly, the fall in productivity growth at the end of the sample masks a sizable increase in the fraction of the population at work. Using either series, the recent decline in productivity must be judged against the background of a general upward trend. Judging from the variability evident in the data, it is probably too soon to conclude that this trend has permanently reversed itself.

These impressions are reinforced by an examination of other developed countries. The outstanding feature in the long run data is the unprecedented surge in growth after the Second World War. There has been some recent slowing of growth compared to the 1950's and 1960's, but only to levels that are still high by historical standards. Consequently, in a test for trend in per capita income for 11 developed countries described in Romer (1986), the evidence is in every case supportive of a positive trend; in most cases, the hypothesis of no trend can be rejected at conventional significance levels.

Overall, the data offer relatively strong support for Kaldor's first fact. Unless one is willing to draw very strong inferences from the few most recent observations on a relatively noisy time series and conclude that they represent a break with historical patterns, there is no reason for a theory of growth to aim for falling growth rates and stagnation.

Figure 3 bears on the constancy of the capital output ratio. Using data from Maddison (1982), it reports the growth rate of capital per hour worked and of output per hour worked for 3 time intervals and 7 countries. Each country is represented by a letter listed in Table 2. The numbers refer to different periods: 1 is the period 1870 to 1913; 2 is the period 1913 to 1950; 3 refers to the period 1950 to 1979. (Like the other data reported here, this sample includes all the countries and time periods for which data are reported in the specified source. Maddison's study covers 16 countries, but these seven, for the specified intervals, are the only ones for which capital stock data are reported.) For the capital-output ratio to be constant, capital and output must grow at the same rate, and a scatter plot of the growth rates should line up on the 45 degree line. Subtracting a constant from each pair–the growth rate of hours worked–should leave each pair on the 45 degree line. In the figure, they cluster around this line to a surprising extent.

Further evidence on this result can be offered. Let i denote the fraction of total income devoted to investment and let δ denote the depreciation rate on capital. Then the equation for the evolution of the capital stock is $\dot{K} = iY - \delta K$. Let g denote the rate of growth of output, $g = \dot{Y}/Y$. If g, i, and δ are constant, then the capital-output ratio will converge to the value

$$\frac{K}{Y} = \frac{i}{\delta + g},$$

with dynamics that behave like $e^{-(g+\delta)t}$. Table 3 reports average values for i, and g from national income accounts data for the seven countries in Figure 3 for the period 1950 to 1981. The data used here are from Summers and Heston (1984) and cover a slightly longer period than those from Maddison.[3] An estimate of the magnitude of δ can be derived as follows. Maddison (1987) reports an average ratio of total depreciation D to income Y for a similar sample of industrialized countries of between 11% and 12%.

[3]The basic data are from the published source, but the actual data used here are go beyond these in two senses. First, the published article gives data only through 1980, whereas the data used here came from a tape that was updated to include 1981. Second, the estimates of relative income for the African countries in the article and on the tape were recently found to have been overstated to a significant extent. Thus, the data used here for African countries are make use of a rough correction provided by Robert Summers. By the time this chapter appears in print, a much more comprehensive revision of the basic Summers and Heston data set will be available in the March 1988 issue of the Review of Income and Wealth.

Using this value for D/Y, one can solve for δ from

$$\delta = \frac{D}{Y}\frac{Y}{K} = \frac{D}{Y}\frac{g+\delta}{i}.$$

For the values of g and i for the countries reported in Table 3, this implies values of δ on the order of 3% or 4%. Using these two values for δ, the table reports estimates of the capital output ratio. For these values of δ and for a value of g on the order of 3% or 4%, the time for K/Y to converge half-way towards its steady state value is around 10 years; for the 31 year interval considered here, this steady state approximation may not be too misleading.

The interesting feature of the table is that the steady state capital-output ratios are relatively similar. There is a relatively large amount of variation in the investment share i and the growth rate g, but there is no systematic variation in the estimated value of K/Y across countries. Note that this is result is stronger than the finding from Figure 3 that K and Y increase in roughly equal proportions in a given country and time period. Equiproportionate increases in Y and K could arise in a world where output varies exogenously and the investment rate stays constant, but Table 3 shows that the investment share varies closely with the growth rate. The question suggested by these data is why the share of investment and the rate of growth of output move together in such a way that the implied capital-output ratio shows little systematic variation.

These statements must be qualified to some extent because the findings are weaker if one significantly increases the sample of countries considered. The conventional wisdom is that the data for developing countries do not show a strong correlation between growth and the share of output devoted to investment. Figure 4 shows why. It uses data from Summers and Heston (1984), plotting the average investment share and the average rate of growth of output for all 115 of the economies that they label (somewhat loosely) as market economies. For 50 of these countries—those denoted in the figure with an "x"—data are available only for the period 20 year interval 1960 to 1981. For the other 65 countries, data cover the interval 1950 to 1981. Seven of these countries, the ones from Figure 3 and Table 3, are denoted with squares and are plotted from right to left in the order in which they appear in Table 3. For clarity they are connected by a line. The remaining 58 countries with data for the period 1950 to 1981 are denoted with plus signs. The growth rate is the average annual (continuously compounded) growth rate of gross domestic product, valued at international prices. The investment measure includes investment by both the private and government sectors.

Judging solely from the countries denoted with an x, there is no strong evidence of a positive association between investment and growth. With the addition of the countries denoted by a plus sign, for which data are available only since 1960, a positive association is once again apparent, but it is not as tight as the relation observed among the seven developed countries connected by the line. Moreover, both the x countries and the + countries tend to lie systematically to the right of the locus for the developed countries. This suggests either that the process of growth for the developing countries differs fundamentally from that for the developed countries, that investment tends to be systematically underestimated in the less developed countries, or that the steady state assumption used here is misleading. For example, countries with large recent additions to their capital stock may have less depreciation than the steady state approximation would suggest. In this case, gross investment could be smaller, but net investment could be comparable to that in developed countries with similar growth rates. Overall, Kaldor's observation number 4), the constancy of the capital-output ratio, can still be judged to be a useful target for theories of growth, but so also might the apparent departures from this tendency for low income countries suggested by Figure 4.

Kaldor's fact number 5), the assertion that the share of capital in total income has remained constant, has increasingly been disputed. Attempts to measure this share in a consistent fashion over time tend to show a fall in capital's share. Table 4 reports estimates from different country studies that are collected in Maddison (1987). There are a number of judgmental issues that must be resolved in deciding what constitutes income to capital, and different authors have taken different positions on how to handle them. Consequently, the estimates are not comparable across countries and authors. Looking only within countries, the trend is for the share of capital to decline from around 0.4 to 0.3. Of course, realistic standard errors for these estimates might be on the same order of magnitude as this decline. The kind of problem that adds to the uncertainty is a systematic and sizable reduction over time in the fraction of self-employed workers and sole proprietorships, for whom it is particularly difficult to distinguish returns to capital from returns to labor. Another source of uncertainty is the somewhat arbitrary methods for imputing income on capital like housing that is outside of the corporate sector. Given this uncertainty, Maddison argues that for some purposes it may not be too serious a distortion of the data to use identical shares for different countries and to assume that the weights are constant over time. Nonetheless, after acknowledging the uncertainty involved, one must give some credence to the assertion that capital's share is falling.

Kaldor's fact 6), that growth rates differ substantially across countries, and the added fact 7), that the dispersion in the growth rates varies systematically with the level of income, are both clearly evident from Figure 5. This figure plots the data for all 115 market economies from Summers and Heston. The horizontal axis measures the ratio of per capita income in a country relative to that in the United States, with income in both countries measured in 1960. One of the major contributions made by Summers, Heston, and Kravis was to correct official exchange rates for departures from purchasing power parity so that this kind of comparison of levels is meaningful. The vertical axis measures the growth rate per capita income for each country in the subsequent interval 1960 to 1981.

The main result here is that the growth rate shows no systematic variation with the level of income. For countries with any initial level of income, the average growth rate is around 3% to 4%. The variance does seem to vary systematically, falling rapidly with per capita income, but this may at least partially refect the fact that the low income countries are much more heavily sampled than the high income countries. In a sample drawn from any distribution, the difference between the minimum and the maximium values will be monotonically increasing in the sample size.

It is perhaps worth emphasizing that the range of growth rates of over 10% is quite large. Over the span of a mere 21 years, the ratio of per capita income in the fastest and slowest growing countries has more than tripled. If even one tenth of this variation in growth rates is due to forces that government policy can influence, the potential long-term gains from better policy are sizable.

The absence of any negative slope in this scatter plot is evidence against the assertion that that low income countries tend to grow more rapidly than high income countries and that convergence in per capita income is taking place.[4] One can of course select a set of countries where convergence has taken place. Figure 6 plots data for the interval 1950 to 1981, the period when substantial convergence is typically alleged to have taken place. Overall, for this smaller sample of 65 countries with data extending back to 1950, one finds the same pattern as for the larger sample in Figure 5, a triangular shape that is roughly symmetric about a horizontal line. This figure also plots two lines that intersect at the point corresponding to Italy. These lines allow one to make income comparisons at the beginning and at the end of the period. The vertical line divides the sample into countries on the right, which had a higher per capita income than Italy in 1950, and those on the left, which had a lower per capita income. The downward sloping line divides the countries based on income comparisons at the end of the period. Countries that lie above this line had per capita income in 1981 that was higher than that in Italy in 1981. It is downward sloping because there are two ways to end up richer than Italy. If a

[4]Another way to address the question of convergence is to consider changes over time in the world distribution of income, an approach which was followed by Summers, Heston, and Kravis (1984). Their finding that world inequality income did not decrease over in the postwar years is another way to describe the result illustrated in Figures 5 and 6.

country starts out poorer, it must grow more rapidly. If it starts out richer, it can grow more slowly.

Although a selection criterion based on levels of per capita income at the end of the period seems suspect for purposes of testing for a downward slope, it is implicitly what one uses when one specifies a sample of industrialized countries that we now think of as being industrialized. The countries which lie above the downward sloping line are almost exactly the countries that Maddison (1982) studies and which Baumol (1986) subsequently uses in his analysis of convergence. The only difference is that New Zealand, Luxembourg, and Iceland have levels of income as high as those of Italy in 1981, but are omitted from Maddison's study, presumably because they are so small. If one had picked developed countries in 1950, Japan would not have made the list and countries like Argentina would have.

Ex post, it is always possible to tell stories about why Japan should have been included and why Argentina should not have been included, but this seems like a risky methodology. Judging from Figures 5 and 6, there is no obvious reason for treating some countries as being so different from the others that they must be excluded from the analysis of questions like convergence. Even if one did conclude that some truncation of the sample is called for, for example because of concern about data reliability, the way to truncate the sample without biasing the inferences is to use the the initial level of income rather than the terminal level. By inspection, it is clear that regardless of the initial level of income that is chosen (that is, regardless of where one chooses to draw a vertical line), the remaining points will not have a negative slope.[5]

[5]Using different data, this point was made in Romer (1986), and is dismissed as unimportant in a footmote in Baumol (1987). DeLong (1987) makes the same point, but rather than starting in 1950, starts in 1870 as Baumol does. This makes little difference, for all the action is in the post-war period. As is shown in Abramowitz (1986), even in Maddison's sample of countries, the convergence with the United States takes place only after 1950. From 1870 to 1950, the United States pulls steadily ahead of the other countries, and even within the reamining set of countries, there is no tendency for income to converge.

Fact 8), the correlation between growth and trade for developed countries is well summarized by the three panels of Figure 7. Over time and across countries, income growth and trade growth are positively correlated, with trade growth varying more than income growth. The data are drawn from Maddison (1982). Each panel represents a different time period. The variation across countries in a particular period suggests the kind of concern that is voiced in current trade disputes, that somehow increases in trade by some countries may increase their rate of growth at the expense of growth in other countries. In contrast, the variation over time suggests that in terms of growth rates, trade may not be a zero sum game. The rate of growth in all countries may be positively related to the rate of growth of world trade.

Fact 9) refers to a negative correlation between per capita income and population growth. Using data from Summers and Heston for the years 1960 to 1981, Figure 8 shows a scatter plot of this relationship. A better test for the influence of per capita income on individual decisions would look at fertility rates, which correct for the age structure of the population and subtract out the effects of mortality and migration, but the gross correlation shown in the figure will almost surely survive any such refinement.

This cross-sectional variation has a time series counterpart that is referred to as the demographic transition. All developed countries have gone through a transition from high fertility and mortality rates to low rates. This transition can either be interpreted as the response of fertility to an exogenous change in mortality rates, as the common response of both mortality and fertility to increases in income, or both. The cross-sectional variation in population growth rates shown here is sometimes interpreted in these terms. The suggestion is that it reflects an exogenous fall in mortality that took place to recently and too rapidly for fertility to have yet responded, or that it is an example of a fall in mortality that has not been accompanied by an increase in income.

Fact 10) is an assertion about the growth accounting literature, which is far too vast to summarize here. A useful overview of the work of the three key participants, Edward Dennison, Dale Jorgenson, and John Kendrick, is given in Norsworthy (1984). A recent, particularly transparent application of the methodology is given in Maddison (1987). The basic reason for the persistent finding of a residual in growth accounting can be seen from Figure 3. Let $Y = F(K,L)$ denote the output from a constant returns to scale production function of aggregate capital and aggregate hours worked. Let $y = Y/L$ and $k = K/L$ denote output per hour worked and capital per hour worked. Then differentiating with respect to time and using the assumption of competitive markets so that $r = f'(k)$ gives

$$\frac{\dot{y}}{y} = \frac{rK}{Y}\frac{\dot{k}}{k}.$$

For values of capital's share $\frac{rK}{Y}$ on the order of 0.3 or even 0.4, there is no way to match the data in Figure 3. Fitting a regression line of $\dot{y}/y$ on $\dot{k}/k$ to this data gives a coefficient on $\dot{k}/k$ that is very close to one. For a country like Japan with capital and output growth rates of over 7%, a share parameter of even 0.4 will imply unexplained growth of over 4% per year for nearly 30 years.

This analysis also casts doubt on simple arguments that capital deepening in the neoclassical model explains why low capital countries like Japan and Germany grew faster and caught up with the leaders. This is typically identified as one of the great successes of the model, but the numbers do not fit the story. If the growth of capital per worker was 4 percentage points higher in Germany than in the United States, the model would predict a growth rate that is higher by 0.3 or 0.4 times 4 percentage points, or less than 2 percentage points. In fact, the growth rate is higher by the same 4 percentage points. The difference between these two numbers is of course the same residual identified above for Japan. One can argue about what accounts for this difference. A die-hard neoclassicist could claim that the exogenous rate of technological change was higher in countries like Germany and Japan, but this reduces the neoclassical model from a theory to a description of the data. Countries that grow fast are countries with fast exogenous growth in the technology. What ever it is, something other than neoclassical physical capital accumulation was taking place.

One possible other factor is accumulation of skills and education, that is of human capital as well as physical capital. To the extent that this kind of accumulation takes place, the correct measure of labor input is not man-hours, but man-hours adjusted for quality change due to better education, or on the job experience. There is sizable variation in education and experience across individuals in the labor force that is reflected in variation in earnings. In principle, this can be used to correct for quality change over time. One can use the cross-sectional variation in wages, education, and experience together with time series estimates of average experience and education in the labor force, to construct an estimate of growth in quality adjusted labor input. There is considerable latitude in how one goes about the details of this construction, with a corresponding variation in resulting estimate of the unexplained residual. The consensus view seems to be that in long-run data for the United States, there is still a sizable component of growth, on the order of 1% or more, that is not explained by growth in capital or quality adjusted labor input. Unless fast growing countries in Figure 4 like Japan, Germany, France, and Italy had substantially more rapid growth in the level of education and experience than that observed in the United States, they too will continue to have large residuals.

Evidence on fact 11) concerning migration flows is heavily influenced by the constraints imposed on these flows. Historical evidence suggests that the unconstrained flows into industrialized countries could be quite large. Greenwood and McDowell (1986) report that during the late 1960's and early 1970's, quotas on immigration into the United States favored migration of skilled workers and professionals. Attention then turned most naturally to consideration of the brain-drain. More recently, a policy shift in favor of applicants with refugee status, combined with legislative debates about illegal immigration have focused attention on unskilled migrants. Potential flows from either source are apparently large.

Section III. The Kuhn-Tucker Theorem and Dynamic Equilibrium Theory.

Growth is a general equilibrium process. All markets and all participants in an economy influence growth and are influenced by it. A growth theorist must therefore construct a dynamic general equilibrium model, starting with a specification of preferences and the technology, and specifying an equilibrium concept. To be able to say anything about the properties of the model beyond an assurance that some equilibrium exists, the theorist must be able to explicitly solve the model or at least give a qualitative description of the solution.

Either explicitly or implicitly, the central tool that is used in the characterization of dynamic competitive equilibrium models is the Kuhn-Tucker theorem. It offers a general procedure for reducing the problem of calculating a competitive equilibrium to the problem of solving a maximization problem. All of the theory of growth can be understood in terms of the application of this theorem to models with tractable functional forms for preferences and the technology. To develop these claims, it is best to start with a simple Irving Fisher type economy that has a finite number of choice variables and is assumed to have perfect markets. Next, the procedure is for studying perfect markets equilibria is extended to the kind of infinite dimensional space that arises in infinite horizon maximization problems, covering the cases of both discrete time and continuous time.

Section III.4 concludes by showing how these techniques can be extended to equilibria that do not satisfy all of the assumptions necessary for perfect competition.

Section III.1 The Kuhn-Tucker Theorem In $\mathbb{R}^n$

Recall that a function $f:\mathbb{R}^n \longrightarrow \mathbb{R}$ is concave if the cord connecting any two points on the graph of f lies on or below the graph of f. Concave functions are central to the theory of maximization because they allow a complete characterization of solutions to maximization problems. For example, if the function $f:\mathbb{R} \longrightarrow \mathbb{R}$ is differentiable, a point $x \in \mathbb{R}$ solves the problem of maximizing f(x) over all of $\mathbb{R}$ if and only if $f'(x) = 0$. That is, $f'(x) = 0$ is a necessary and a sufficient condition for x to be a solution. This maximization problem is unconstrained in the sense that x can be drawn from anywhere in $\mathbb{R}$. The Kuhn–Tucker theorem generalizes this complete characterization to concave maximization problems with constraints on the feasible choices of x. To be a concave problem, the objective function must be concave and the constraint set must be convex.

Consider a generic constrained maximization problem P:

P
$$\begin{aligned} &\max f_0(x) \\ \text{s.t.}\quad & x \in \Omega \\ & f_1(x) \geq 0, \\ & f_2(x) \geq 0, \\ & \vdots \\ & f_m(x) \geq 0. \end{aligned}$$

To make this a concave problem, assume that $f_0, f_1, \dots, f_m$ are differentiable, concave, real valued functions functions defined on a convex domain $\Omega \subset \mathbb{R}^n$. One further assumption is crucial in what follows, and it is convenient to give it a name. We will say that problem P satisfies the Slater condition if there is a point $\bar{x}$ in the interior of Ω

such that $f_i(\overline{x}) > 0$ for all $i = 1, \dots, m$. This is an example of an interiority condition. It says that there is at least one point in the interior of the set of feasible points.

Associated with the problem P, we can define a function $L:\Omega \times \mathbb{R}^m_+ \longrightarrow \mathbb{R}$ by the rule

$$L(x,\lambda) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x).$$

This function is typically called a Lagrangian. In the description of L, $\mathbb{R}^m_+$ denotes the non–negative orthant in $\mathbb{R}^m$, so $L(x,\lambda)$ is not defined for a vector λ which has a negative component. Components equal to zero are allowed. The key property of this function is that it is concave–convex. Holding λ fixed, the function $L_\lambda(x)$ which sends x to $L(x,\lambda)$ is a concave function. Holding x fixed, the function $L_x(\lambda)$ which sends λ to $L(x,\lambda)$ is a convex function.[6] (Recall that a function g is convex if $-g$ is concave.) A concave-convex function is sometimes referred to as a saddle function because if x and λ are numbers rather than vectors, the graph of such a function would look like a saddle. A point $(\hat{x},\hat{\lambda})$ is said to be a saddle point (or maxi–min point) of L if the following inequality holds:

$$L(x,\hat{\lambda}) \leq L(\hat{x},\hat{\lambda}) \leq L(\hat{x},\lambda).$$

This says that for fixed $\hat{\lambda}$, $\hat{x}$ maximizes $L_{\hat{\lambda}}(x)$ and that for fixed $\hat{x}$, $\hat{\lambda}$ minimizes $L_{\hat{x}}(\lambda)$.

For this kind of problem, it is possible to list three simple conditions that are equivalent to the statement that a point $(\hat{x},\hat{\lambda})$ is a saddle point of L with $\hat{x}$ in the

---

[6]For proofs of these and other assertions made in the text, see Rockafellar (1972).

interior of Ω:

C1) $$Df_0(\hat{x}) + \sum_{i=1}^{m} \lambda_i Df_i(\hat{x}) = 0$$ [7]

C2) $$f_i(\hat{x}) \geq 0, \ \hat{\lambda}_i \geq 0 \ \ i = 1, \dots, m$$

C3) $$\sum_{i=1}^{m} \hat{\lambda}_i f_i(\hat{x}) = 0$$

If $\hat{x}$ were actually on the boundary of Ω, the derivative condition in condition C1 might not hold, and would need to be replaced by a slight generalization. In practice, this is not a problem. If Ω imposed binding constraints on the choice of x, it would be better to make this constraint explicit by describing it in terms of an additional constraint function $f_{m+1}(x) \geq 0$ and giving it its own multiplier. The whole point of the Kuhn-Tucker theorem is to get all of the binding constraints attached to a multiplier in the Lagrangian so that one can take derivatives as in C1.

Condition C1 implies that $\hat{x}$ maximizes $L_{\hat{\lambda}}(\cdot)$. Condition C2 ensures that $\hat{x}$ satisfies the constraints and that $\hat{\lambda}$ is non-negative. Condition C3 implies that $\hat{\lambda}$ minimizes $L_{\hat{x}}(\cdot)$; since λ_i and $f_i(x)$ must be non-negative, the smallest value the summation involving λ can take on is 0. Condition C3 makes sure that this minimum is achieved. Given the non-negativity restrictions from C2, condition C3 can also be written as:

C3′) $$\hat{\lambda}_i f_i(\hat{x}) = 0 \text{ for } i = 1, \dots, m$$

---

[7]Throughout, $Df(x)$ will denote the derivative of a function defined on $\mathbb{R}^n$ and $D_i f(x)$ will denote the partial derivative of f with respect to its i'th argument. For functions defined on the real line, the usual prime notation for a derivative, $f'(x)$, will generally be used.

Stated in this form, these conditions are sometimes referred to as complementary slackness conditions.

These conditions are what one actually uses to solve problems like P, but for the purposes of economic theory, it is useful to work directly with the notion of a saddle point. The essence of the Kuhn–Tucker theorem is that saddle points of L are equivalent to solutions to P.

Theorem (Kuhn–Tucker): Assume that $f_0, f_1, \dots, f_m$ are concave, continuous functions from $\Omega \subset \mathbb{R}^n$ into $\mathbb{R}$.[8] Let the problem P and the function $L: \Omega \times \mathbb{R}^m_+ \longrightarrow \mathbb{R} \cup \{+\infty\}$ be defined as above.

i) <u>Sufficient Conditions for an Optimum</u>: If $(\hat{x}, \hat{\lambda}) \in \Omega \times \mathbb{R}^m_+$ is a saddle point of L, then $\hat{x}$ is a solution to P.

ii) <u>Necessary Conditions for an Optimum</u>: Assume that the Slater condition holds. Then if $\hat{x} \in \Omega$ is a solution to P, there exists a value $\hat{\lambda} \in \mathbb{R}^m_+$ such that $(\hat{x}, \hat{\lambda})$ is a saddle point of L.

The theorem stated here is a special case of a result that can be generalized. For a proof that uses a slightly weaker version of the Slater condition, see Theorems 28.2 and 28.3 in Rockafellar (1970).

The components of $\hat{\lambda}$ are referred to as Lagrangian multipliers, or more suggestively as shadow prices for the constraints. The notion of price is quite appropriate

[8]There is a technical point here. Concave functions are always continuous on the interior of the domain on which they are defined. The assumption of continuity here is necessary only to ensure that these functions do not have any jumps on any boundary points in Ω.

here, for the difference between the maximization problem P and the problem of maximizing $L_{\hat{\lambda}}(x)$ over x is precisely the difference between the problem faced by a social planner and that faced by a competitive price taking agent. In the problem P and in a social planning problem, the maximization problem must take explicit account of the resource constraints on the choice variables. In contrast, at the market prices, a competitive agent is assumed to act as if it is possible to purchase an unlimited quantity of all goods. Similarly, at the prices λ, someone solving the problem P is free to behave as if the constraints f_1 to f_m could be violated in the process of maximizing $L_{\hat{\lambda}}(x)$ over x. In each case, the prices convert a constrained maximization problem into an unconstrained problem, with prices that are determined separately.

This similarity is indicative of a much deeper connection between saddle–points for Lagrangians and equilibria for competitive systems. Formally, they are equivalent. A general proof of this assertion is given in Romer (1988). Here, it is sufficient to show this equivalence in a simple example that forms the basis for all of the growth models that follow. The arguments in the general case are essentially the same.

Consider a two period economy with a representative consumer. For ease of comparison with the dynamic models that follow, let $U:\mathbb{R}^2_{++} \longrightarrow \mathbb{R}$ take on the additively separable discounted form, $U(c_1,c_2) = u(c_1) + \beta u(c_2)$. The function u is assumed to be concave.[9] Let $e > 0$ denote the period 1 endowment in this economy and let $f:\mathbb{R}_+ \longrightarrow \mathbb{R}$ be a concave production function for converting forgone consumption in period 1 into consumption in period 2. Define an aggregate maximization problem for this economy

[9]Here $\mathbb{R}^2_{++}$ denotes the strictly positive orthant in $\mathbb{R}^2$. This is used instead of the non–negative orthant to accommodate functions like logarithmic utility that are widely used in practice. What follows could easily be modified to allow for consumption equal to zero.

(labeled P1 because it is the first in a series of similar problems) as follows:

P1
$$\max u(c_1) + \beta u(c_2)$$
$$\text{s.t } e - c_1 - k \geq 0$$
$$f(k) - c_2 \geq 0$$
$$c_1, c_2 > 0, k \geq 0.$$

This can be put in the form of the generic problem P by letting the choice vector x be a triple $(x_1,x_2,x_3) = (c_1,c_2,k)$, letting Ω be the domain $\mathbb{R}^2_{++} \times \mathbb{R}_+$ in $\mathbb{R}^3$, letting $f_0(x) = u(x_1) + \beta u(x_2)$, letting $f_1(x) = e - x_1 - x_3$, and letting $f_2(x) = f(x_3) - x_2$. In conformity with numbering scheme for problems, let $L1:\Omega \times \mathbb{R}^n \longrightarrow \mathbb{R}$ denote the Lagrangian associated with P1:

$$L1(x,\lambda) = u(x_1) + \beta u(x_2) + \lambda_1(e\text{-}x_1\text{-}x_3) + \lambda_2[f(x_3)\text{-}x_2].$$

Associated with the aggregate maximization problem P1 are the problems for a price taking consumer and a price taking firm. Let PF(p) denote the firm's problem for this economy when faced with given prices $p \in \mathbb{R}^2_+$, and let $\Pi(p)$ denote the profits it earns:

PF(p)
$$\Pi(p) = \max_{k \in \mathbb{R}_+} \; p_2 f(k) - p_1 k.$$

Let $PC(p,\pi)$ denote the problem of the price taking consumer who also takes as given the profits π received from ownership of the firm:

$$PC(p,\pi) \qquad \max_{c \in \mathbb{R}^2_+} u(c_1) + \beta u(c_2)$$
$$\text{s.t.} \quad \pi + p_1(e\text{-}c_1) - p_2c_2 \geq 0.$$

We will say that a pair $(\hat{x},\hat{p})$ is a competitive equilibrium if $\hat{x}_3$ solves $PF(\hat{p})$, if $(\hat{x}_1,\hat{x}_2)$ solves $PC(\hat{p},\pi)$ when $\pi = \Pi(\hat{p})$, and if supply is greater than demand: $e \geq \hat{x}_1+\hat{x}_3$, $f(\hat{x}_3) \geq \hat{x}_2$.

This is an example of an equilibrium with all trading in the initial period. The problems of the consumer and firm as described as if they meet and arrange all trades at time 1. This formulation is chosen only for the convenience of the theorist. There is an equivalent formulation of the equilibrium in terms of the spot prices and securities returns that people use on an everyday basis. For example, the interest rate in this two period example will simply be the ratio of the prices for the dated goods, $r = p_1/p_2$. Throughout this chapter, all of the equilibria will be described as if all trading took place at time zero, with the understanding that they can be converted into equilibria with spot prices and a large enough set of security returns.

The next proposition shows that if the Slater condition holds (which is equivalent here to assuming that f is productive and e is greater than zero), then saddle points of L1 are equivalent to competitive equilibria. As always, this is subject to the qualification that competitive prices are determined only up to a non-negative scaling factor.

PROPOSITION 1: Assume that u and f are continuous, concave functions. Suppose that $e > 0$ and that $f(y) \geq 0$ for some $y > 0$. If $(\hat{x},\hat{p})$ is a competitive equilibrium, then $(\hat{x},\hat{\gamma}\hat{p})$ is a saddle point of the Lagrangian L1 for some non–negative scale parameter $\hat{\gamma}$. Conversely, if $(\hat{x},\hat{p})$ is a saddle point of L1, then $(\hat{x},\hat{p})$ is a competitive equilibrium.

The essence of the proof of this proposition is to apply the Kuhn-Tucker theorem to all three of the problems P1, PC, PF. The necessary conditions for the maximization problem P1 imply the sufficient conditions for the maximization problems PC and PF, and vice versa. The details, which amount to keeping track of notation, are given in the appendix.

For concreteness, the statement and proof are given in the context of a particular economy, but they can both be extended immediately to a more general setting. Under the mild assumption that preferences can be represented by concave utility functions, this proof can be generalized to apply in any economy where the equilibrium is Pareto optimal. If there is more than one agent, the objective becomes a weighted sum of the individual utility functions, and the problem is transformed into one of calculating one of many possible Pareto optima.

In light of the equivalence between saddle points and competitive equilibria, it is possible to reinterpret the Kuhn-Tucker theorem. The sufficient conditions from the theorem embody the First Welfare Theorem; competitive equilibria are Pareto optimal. The necessary conditions embody the Second Welfare Theorem; for any Pareto optimal quantities, there exist prices that decentralize these quantities as a competitive equilibrium.

Section III.2 Discrete Time Extensions of the Kuhn-Tucker Theorem

In a dynamic application of this kind of procedure, the possibility of an infinite time horizon seems to remove a crucial upper end-point. Depending on whether time is discrete or continuous, the links between periods that arise from possibilities for intertemporal substitution show up in the first order conditions as difference equations or differential equations. In a finite horizon model, the upper end-point plays a special role in establishing boundary conditions for these equations.

To see how these issues arise in a simple setting, consider a multi-period extension of the growth model described in problem P1, but assume that the equilibrium quantities are exogenously specified so that we do not need to make any assumptions about the form of preferences. Let the technology be as in the two period model, but extended over more periods. To simplify the notation, let the initial endowment k_0 be given in terms of the capital stock so that the initial amount of resources available for use is $f(k_0)$. Then consumption in all periods $t \geq 0$ is related to the capital stock by

$$f(k_t) - c_t - k_{t+1} \geq 0.$$

The evolution of k_t is also limited by the restriction that it must be be non–negative for all t. Even though this constraint is often neglected, it will become clear that this restriction has important economic content.

Consider first a finite horizon problem so that t runs from 0 to T. Let the values for consumption in periods 1 to T take on exogenously specified values $\bar{c}_1$, $\bar{c}_2$, ... , $\bar{c}_T$ and consider the problem of maximizing consumption in time 0 subject to these constraints:

P2

$$\max \ c_0$$
$$\text{s.t.} \quad f(k_t)-c_t-k_{t+1} \geq 0 \quad \text{for } t=0,1, \ldots, T$$
$$k_t \geq 0 \quad \text{for } t=1, \ldots, T+1$$
$$c_t \geq \bar{c}_t \quad \text{for } t=1, \ldots, T.$$
$$k_0 = \bar{k}.$$

Solving this problem for the quantities is not of much interest. It usefulness will lie in what it can tell us about prices. Since the problem fits into the form of the generic problem P, its Lagrangian can be copied from the Lagrangian for P. Using boldfaced letters to denote sequences, let $\mathbf{c} = \{c_t\}_{t=0}^{T}$ and $\mathbf{k} = \{k_t\}_{t=1}^{T+1}$ denote the vectors of quantities that must be chosen. In a T period problem, the constraint on c_T involves k_{T+1}, which must also be specified. This is where the non-negativity condition on capital becomes relevant. Let $\boldsymbol{\lambda} = \{\lambda_t\}_{t=0}^{T}$ denote the vector of multipliers on the constraints linking periods, let $\boldsymbol{\gamma} = \{\gamma_t\}_{t=1}^{T+1}$ and $\boldsymbol{\omega} = \{\gamma_t\}_{t=1}^{T}$ denote the multipliers on the non–negativity constraints. Then L2 takes the form

$$L2(\mathbf{c},\mathbf{k},\boldsymbol{\lambda},\boldsymbol{\gamma},\boldsymbol{\omega}) = c_0 + \lambda_0[f(k_0)-c_0-k_1] + \sum_{t=1}^{T}\left\{\lambda_t[f(k_t)-c_t-k_{t+1}] + \gamma_t k_t + \omega_t(c_t-\bar{c}_t)\right\} + \gamma_{T+1}k_{T+1}.$$

The first order conditions for this problem are straightforward. First, hold the shadow prices fixed and maximize L2 over the quantities c and k. Differentiating with respect to c_0 gives $\lambda_0 = 1$, with respect to c_t gives $\omega_t = \lambda_t$. Differentiating with respect to k_t for $t \in \{1,2, ..., T\}$ gives

$$\lambda_{t-1} = f'(k_t)\lambda_t + \gamma_t. \tag{1}$$

The derivative with respect to k_{T+1} gives

$$\lambda_T = \gamma_{T+1}.$$

Consider next the complementary slackness conditions for minimizing L with respect to λ, γ, and ω. Assuming that $f(0) = 0$ and that $\bar{c}_t > 0$ for all t, k_t must be positive for all $t = 1, ..., T$. This implies that $\gamma_t = 0$ over the same range. Assuming that $f'(k)$ is positive for all k, the initial condition $\lambda_0 = 1$ together with the difference equation for λ then implies that λ_t is positive for all t. Then the complementary slackness condition associated with λ_t implies that k_{t+1} must equal output minus consumption,

$$k_{t+1} = f(k_t)-c_t. \tag{2}$$

Since λ_t equals ω_t, it follows that c_t must equal $\bar{c}_t$. Finally, the equality $\gamma_{T+1} = \lambda_T$ together with the complementary slackness condition $\gamma_{T+1}k_{T+1} = 0$ implies

$$\lambda_T k_{T+1} = 0.$$

The equations for λ and k are the only ones of any economic interest. The shadow prices γ and ω are superfluous. Equations (1) and (2) form a pair of coupled, first order difference equations. Generally, this kind of equation system requires two boundary conditions to pin down all of the values. They are the given initial value for k_0 and the terminal condition $\lambda_T k_{T+1} = 0$. The first is an example of an initial condition. For obscure reasons, the second is called a transversality condition. In this problem, one additional condition is need to pin down c_0. This is given by the condition $\hat{\lambda}_1 = 1$. Taken together, these conditions and the difference equations are just sufficient to determine the values for λ_t and k_t for all t.

In this particular problem, it is intuitively obvious that the condition $\lambda_T k_{T+1} = 0$ is satisfied by setting $k_{T+1} = 0$. There is no reason to leave anything after everyone is gone. In more general problems, for example a problem where it is costly to convert capital goods back into consumption goods, the case $\lambda_T = 0$ and $k_{T+1} > 0$ can also arise. The intuition for the zero price is clear. If capital is going to be abandoned anyway, having more of it can not be of any value.

This kind of analysis does not offer a very interesting theory of quantity determination, but conditional on the quantities, it determines the equilibrium prices $\hat{\lambda}_t$. By converting the prices here, which have the interpretation of time zero prices, into the kind of prices used by individuals in spot markets, it is straightforward to show that the gross rate of return on one–period bonds sold in period t is

$$R_t = \frac{\hat{\lambda}_t}{\hat{\lambda}_{t+1}} = f'(k_t).$$

This kind of problem has a natural extension to an infinite horizon. The difficulty that the infinite horizon poses is that the absence of an upper bound for time at first seems to remove the boundary condition $\hat{\lambda}_T \hat{k}_{T+1} = 0$. Because of the links between periods implicit in the difference equations, the loss of one of the boundary conditions would mean that none of the quantities and prices are determined. This indeterminacy, sometimes referred to as the Hahn problem, is only apparent. It arises only if one does not take account of the non-negativity condition on k. In the complete markets equilibrium concept used here, there is a second boundary condition and all prices and quantities are determined.

To see this, let P3 denote the extension of the problem P2 when the upper bound T is removed, so t runs from 0 to ∞. It takes some sophisticated mathematics to be precise about what the definition of a Lagrangian is in an infinite dimensional space and to prove a version of the Kuhn-Tucker theorem that applies in this context, but the result is intuitively appealing.[10] Subject an interiority or Slater condition, the Kuhn-Tucker theorem still says that solutions to maximization problems are equivalent to saddle points of a Lagrangian. In this case, the Lagrangian takes the form

$$L3(c,k,\lambda,\gamma,\omega) = c_0 + \lambda_0[f(k_0)-c_0-k_1] + \sum_{t=1}^{\infty}\left\{\lambda_t[f(k_t)-c_t-k_{t+1}] + \gamma_t k_t + \omega_t(c_t-\bar{c}_t)\right\}.$$

The difference equations for k and λ from this Lagrangian are exactly those from the the finite horizon problem. The initial condition is still given by k_0.

Provided that $f(0) = 0$ and $\bar{c}_t > 0$ for all t, it follows that k_t must be strictly positive for all t. This makes it appear that the non–negativity constraint on k_t is never binding and can be dropped from consideration, but this is not quite right. It is true that $\gamma_t = 0$ for all t. For any finite T, the non-negativity constraints on k_t could be dropped for all t between 1 and T, but this does not mean that the entire infinite sequence of constraints can be dropped. To see why, suppose that f(k) takes the form

[10] See Ekeland and Teman (1976) for an general treatment of convex analysis and Kuhn-Tucker theory in an infinite dimensional space. Araujo and Scheinkman (1983) apply this kind of framework to derive necessary conditions for a continuous time problem. Romer and Shinotsuka (1988) give an explicit derivation a Lagrangian for this kind of discrete time model.

$f(k) = rk$ for some constant r and for values of k which can be either positive or negative. This technology is like having a bank that offers deposits and loans at the rate r. In this case, if all the non–negativity constraints are dropped, the initial value of c_0 can be made arbitrarily large by letting k_t take on negative values that diverge to $-\infty$.

In a sense, the non–negativity constraint is binding at infinity. Corresponding to this constraint, there is a complementary slackness condition at infinity:

$$\lim_{t \to \infty} \hat{\lambda}_t \hat{k}_{t+1} = 0.$$

This is the transversality condition at infinity that serves as the second boundary condition for this problem. Intuitively it is just the limit of the transversality condition $\hat{\lambda}_T \hat{k}_{T+1} = 0$ for the finite horizon problem. It can be shown that this is a part of the Kuhn-Tucker necessary and sufficient conditions in the same sense that $\hat{\lambda}_T \hat{k}_{T+1} = 0$ is a part of these conditions in the finite horizon problem. In a problem satisfying the Salter condition this condition holds for any saddle point $(\hat{k},\hat{\lambda})$;[11] if a pair satisfies this condition plus the difference equations for k and λ, then it is a saddle point of L3.

Subject to the assumptions noted along the way (e.g. that $f'(k) > 0$ and that the Slater condition holds), this shows that for the specified technology and sequence of quantities for aggregate consumption starting at date 0, there exist prices that support an

---

[11]Specifying what the Slater is here is complicated, because one must define an interior point in an infinite dimensional space. For a general discussion of this issue, see Romer and Shinotsuka (1988). For this particular problem, let $(\hat{c}_0,\hat{k})$ denote the optimal quantities. If it is possible to construct a path $(\tilde{c}_0,\tilde{k})$ such that $\tilde{c}_0 < \hat{c}_0$ and $\tilde{k}_t > a\hat{k}_t$ for some $a > 1$, this problem can be shown to satisfy the Salter condition in the relevant sense. In the Solow-Swan model, this will be true as long as $\bar{c}_t$ is bounded below from the maximum sustainable value for consumption, $c_{max} = \max_k f(k) - k$.

efficient allocation as a complete markets competitive equilibrium. The allocation is efficient in the sense that it maximizes time zero consumption taking consumption in all subsequent periods as given. There are other equilibrium notions with the property that equilibria need not be efficient. The one used in Samuelson's overlapping generations economy (1958) is the canonical example. This kind of equilibrium cannot be computed using the technique used here because it is not a complete markets equilibrium. In Samuelson's model, private agents are assumed to be unable to trade in the entire set of markets for dated goods. Only the government can do this. Accordingly, only the government issues infinite lived, zero interest bearing bonds and earns a pure profit. Private individuals and firms in this model would also issue such bonds if they could.

Samuelson's example is a caution that the assumption of complete markets is not an empty assumption, even in the absence of uncertainty. It shows that there are other equilibrium concepts that may be of interest, with equilibria that cannot be calculated using the method described here. Interesting as they are at an abstract level, the issues raised by these alternative equilibrium concepts have no obvious bearing on the theory of long run growth. The Samuelsonian problem of inefficiency arising from chronic capital over accumulation has no apparent counterpart in the data cited in the beginning of this chapter.

The best known example of a theory of growth with exogenously specified quantities is the model identified with Solow and Swan. (Solow, 1956; Swan 1956) In the simplest form with constant population, these models, both posed in continuous time, assume that capital per capita k evolves according to the equation

$$\dot{k}(t) = sf(k(t)) - \delta k(t)$$

where s is the saving rate, δ is an exponential depreciation rate and $f(k)$ denotes output per worker $f(k) = F(k,1)$ before allowance is made for depreciation. The function $F(\cdot)$

is assumed to exhibit constant returns to scale. Consumption is the residual,

$$c(t) = (1-s)f(k(t)).$$

Under standard assumptions on f and F, the differential equation for k has a stable steady state. As is now quite familiar, exogenous population growth can be added to the model to generate growth in total income with constant per capita income. Exogenous technological change can be added to generate growth in per capita income.

Given this description of the evolution for the quantities, prices follow by exactly the kind of analysis given for the discrete model above. In particular, one can give a rigorous explanation for why $f'(k(t))$ is the instantaneous interest rate at time t. The analysis is incomplete in the sense that preferences are not fully specified, but conditional on the specification of the quantities, it is a perfectly reasonable general equilibrium model, one that can be readily taken to the data.

One of the key contributions of the Solow–Swan analysis was the renewed impetus it gave to the use of simple aggregate models along the lines suggested by Ramsey (1929). Once Ramsey's technology was recognized as a powerful tool, it was natural that his preferences should be adopted as well. The key hurdle here seems to have been the idea of discounting. Ramsey denounces discounting as "ethically indefensible" on the first page of his article, then proceeds to use discounting in the most interesting parts of his analysis. Samuelson and Solow (1956) reproduced and extended Ramsey's analysis, but they abided by his admonition not to discount, and their analysis seems to have had little impact. Koopmans (1965) and Cass (1965) are generally recognized as the key contributions in the process of legitimizing discounting and taking the analysis out of the realm of tricks and special cases needed to work with the undiscounted model.

Implicit in Ramsey's ambivalence about discounting is a sense that the objective function in such a problem should reflect the preferences of the economist rather than the economy. What can be called his positive analysis allows for discounting, and he speaks in this case of an infinite lived family, but he clearly had in mind a separate welfare analysis that did not respect the preferences of such families. Even after backsliding had set in and discounting was a firmly established practice, many papers in the 1960's seemed to retain the view that the objective function ought to reflect a higher authority than the mere consumer, God perhaps, or at least someone with a Ph.D. Macroeconomic applications of these tools played were crucial in moving the profession away from this normative interpretation of Ramsey models and establishing these models as positive models of equilibria.[12]

The mathematical treatment of a model with discounted Ramsey preferences is another application of the Kuhn-Tucker theorem. To see this in the discrete time case, add discounted Ramsey preferences to the technology from problem P3. Then the aggregate maximization problem for this economy is

P4
$$\max \sum_{t=0}^{\infty} \beta^t U(c_t)$$
$$\text{s.t.} \quad f(k_t)-k_{t+1}-c_t \geq 0 \quad t=0,1,2,\ldots$$
$$k_t \geq 0 \quad t=1,2,\ldots$$

[12]Many papers from macroeconomics have used Ramsey preferences to calculate a decentralized equilibrium that reflects the preferences of individuals. In a partial equilibrium setting with constant interest rates, Lucas and Prescott (1971) showed that these preferences could be interpreted as a description of a market demand curve without invoking a representative agent. Of the papers that invoke a representative agent, three of the most influential are Barro (1975), Hall (1978), and Lucas (1978).

The Lagrangian for this problem is exactly what one would expect:[13]

$$L4(\mathbf{k},\mathbf{c},\boldsymbol{\lambda},\boldsymbol{\gamma}) = \sum_{t=0}^{\infty} \beta^t u(c_t) + \lambda_t[f(k_t)-k_{t+1}-c_t] + \gamma_t k_t.$$

For sensible specifications of $u(\cdot)$ and $f(\cdot)$, the non-negativity constraint on k_t is not binding for any finite t. In this case the multipliers γ_t are equal to zero for all t and can therefore be ignored, provided one keeps in mind that the non-negativity constraint is still binding at infinity. This is what gives the transversality condition at infinity. For fixed shadow prices $\boldsymbol{\lambda}$, maximizing L with respect to $\mathbf{c}$ gives c_t as a function of λ_t and t, $c(\lambda_t,t) = u'^{-1}(\lambda_t/\beta^t)$. If $u(\cdot)$ is increasing, λ_t must be positive. Then the complementary slackness condition $\lambda_t[f(k_t)-k_{t+1}-c_t] = 0$ gives k_{t+1} in terms of λ_t and k_t. Differentiation of L with respect to k_t gives the same equation linking shadow prices in adjacent periods as did the problems P2 and P3. Taken together, these two equations, one for k and one for λ form a coupled system of first order difference equations in two variables:

$$\lambda_t = \lambda_{t-1}/\beta f'(k_t) \tag{3}$$
$$k_t = f(k_{t-1}) - c(\lambda_{t-1},t) \tag{4}$$

By comparison, this shows why the Solow-Swan model is easier. In that model, these equations have a triangular structure; the second equation does not depend on the first because c is exogenously specified in terms of k. It can be solved independently. An analogous case arises here if $f(k)$ is linear so that $f'(k)$ is constant. In this case, the first equation can be solved independently. If the utility function u is quadratic or of the

[13]For a derivation of this Lagrangian, see Romer and Shinotsuka (1988).

constant elasticity form $u(c) = \frac{c^{1-\sigma}-1}{1-\sigma}$ with $\sigma \in [0,\infty)$[14], the second equation can readily be solved.

The boundary conditions for these equations are the initial value for k_0 and the transversality condition at infinity,

$$\lim_{t\to\infty} \lambda_t k_{t+1} = 0. \tag{5}$$

Terminology in this area is not well established. Equations (3) and (4) are referred to variously as Lagrangian or Hamiltonian equations. In current usage, the term Euler equations is applied most often to a transformed version of these equations. Solving equation (4) for λ_t in terms of k_t and k_{t-1}, then substituting the result into the first equation yields a second order difference equation in k_t,

$$\beta U'(f(k_t)-k_{t+1})f'(k_t) = U'(f(k_{t-1})-k_t). \tag{6}$$

The Euler equation (6) can be derived directly by assuming that the evolution equation for k_t holds with equality, then substituting it into the objective function. This equation then follows by differentiation with respect to k_t. Treating the problem this way makes it seem as if all constraints can be transformed away so that the problem becomes an unconstrained maximization problem; no constraints or multipliers are in evidence. For finite horizon problems with fixed initial and terminal values for k–the kind of problem typically studied by physical scientists–this transformation into an unconstrained problem is possible. Hence, many treatments of dynamic maximization problems describe the

[14]The extra constant 1 in the numerator of this functional form has no effect on choices made by the consumer, but is useful since in this form, the function converges to $u(c) = \ln(c)$ when σ approaches 1. With this convention, this form makes sense for all values in the indicated range.

methods for solution in terms of the techniques from calculus for unconstrained maximization; just set all the derivatives in sight equal to zero. However, for infinite horizon problems, substitution cannot remove the binding non-negativity constraint on capital at infinity. To fully understand the transversality condition at infinity associated with this constraint, it is essential to have available the full machinery of the Kuhn-Tucker theorem for constrained problems. Viewed from the perspective of constrained maximization problems, it is an obvious generalization of a complementary slackness condition.

Actually proving that the transversality condition (5) is a necessary condition for the problem P4 requires checking the assumptions for the Kuhn-Tucker theorem. In particular, it requires checking that the infinite dimensional version of the Slater interiority condition holds. In models with no discounting, this condition can fail, and such models can be used to construct counter-examples to the transversality condition at infinity. The transversality condition will hold in any optimizing model where there is sufficient discounting relative to the maximum rate of growth of capital in the economy. In particular, if $\lim_{t\to\infty} \beta^t k_{t+1} = 0$ along any feasible path for capital, it will hold.[15]

Section III.3 Continuous Time Extensions of the Kuhn-Tucker Theorem

The extension of the infinite horizon model to continuous time is comparable to the extension of a model with discrete uncertainty to one with continuous random variables. Essentially, sums are replaced by integrals and difference equations become differential equations. Without covering the formal details, it is straightforward to give a heuristic derivation of the continuous time Lagrangian or Hamiltonian equations.

---

[15]For a complete discussion, see Romer and Shinotsuka (1988).

Write the continuous time maximization problem as

P5
$$\max \int_0^\infty U(c(t))\, e^{-\rho t} dt$$
$$\text{s.t.} \quad \dot{k}(t) = f(k(t)) - c(t) \quad \text{for all } t \geq 0$$
$$k(t) \geq 0 \qquad \text{for all } t \geq 0.$$

For simplicity, depreciation is hidden, but it could easily be made explicit. In the discrete model, f represented gross output. Here it represents net output. The choices are entirely a matter of convenience. In discrete time, one can convert from gross to net production by the substitution $g(k) = f(k) - k$. In continuous time, an explicit depreciation rate can be added by a substitution $g(k) = f(k) + \delta k$.

Letting bold face letters $\mathbf{c}$, $\mathbf{k}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\gamma}$ stand this time for functions defined on $[0,\infty)$, and using $c(t)$, $k(t)$, $\lambda(t)$, and $\gamma(t)$ or c, k, λ, and γ for their values at a point in time, the Lagrangian L5 takes the form

$$L5(\mathbf{c},\mathbf{k},\boldsymbol{\lambda},\boldsymbol{\gamma}) = \int_0^\infty \Big\{U(c(t))e^{-\rho t} + \lambda(t)[g(k(t)) - c(t) - \dot{k}(t)] + \gamma(t)k(t)\Big\}\, dt.$$

As in the discrete time problem, $\gamma(t)$ will be zero for all finite t in any reasonable model, so the term $\gamma(t)k(t)$ will henceforth be ignored. For fixed shadow prices $\boldsymbol{\lambda}$, the operation of maximizing L5 with respect to $\mathbf{c}$ can be passed through the integral, and the maximization can be done point by point for each fixed t. It is then useful to define the terms inside the integral other than $\lambda(t)\dot{k}(t)$ as a new function $H:\mathbb{R}^3_+ \longrightarrow \mathbb{R}$:

$$H(k,\lambda,t) = \max_c \; U(c)e^{-\rho t} + \lambda(g(k) - c)$$

Thus,

$$\max_{\mathbf{c}} \ L5(\mathbf{c},\mathbf{k},\boldsymbol{\lambda}) = \int_0^\infty \{H[k(t),\lambda(t),t] - \lambda(t)\dot{k}(t)\}dt.$$

To complete the calculation of a saddle point of L5, it remains to maximize L5 over **k**. Because H depends on k and $\dot{k}$, one cannot simply maximize point by point as one could with c. At each point one must trade-off the effect of increasing the level k with the effect of increasing its rate of change $\dot{k}$, which after all is the only way to increase the level. To do this kind of maximization and make this trade-off explicit, one needs to apply the tools of the calculus of variations. The basic result needed here is the first order condition for maximizing an integral of the form $\int M(k(t),\dot{k}(t),t)\ dt$ with respect to a path k(t). It is given as a differential equation,

$$D_1M(k(t),\dot{k}(t),t) - \frac{d}{dt}[D_2M(k(t),\dot{k}(t),t)] = 0.$$ [16]

To apply this result here, define M as follows:

$$M(k,\dot{k},t) = \{H(k,\lambda(t),t) - \lambda(t)\dot{k}\}.$$

The time dependence in M corresponds to the dependence of H on λ(t) and on the

[16]This is also known as an Euler equation, or sometimes even as an Euler–Lagrange equation. In classical notation this is often written as

$$\frac{\partial}{\partial k} M(k,\dot{k},t) + \frac{d}{dt}\left[\frac{\partial}{\partial \dot{k}} M(k,\dot{k},t)\right] = 0.$$

The notation used in the text is more explicit about what the time derivative refers to and why the meaning would be so different if $\frac{d}{dt}$ were replaced by $\frac{\partial}{\partial t}$. For a proof, see any text that covers dynamic optimization, for example Intriligator (1971) or Luenberger (1969).

exponential discounting. The partial derivative of M with respect to $\dot{k}$ is simply $\lambda(t)$. Then the differential equation becomes

$$\dot{\lambda}(t) = - D_1 H(k(t),\lambda(t),t). \tag{7}$$

To derive the other differential equation, note that the first order condition for the maximum over c in the definition of H is

$$u'(c)\, e^{-\rho t} = \lambda, \tag{8}$$

which is analogous to the expression for c derived for the discrete time model. This implicitly defines c as a function of λ and t, which we can denote $\hat{c}(\lambda,t)$. Substituting this into the definition of H, then differentiating with respect to λ and using the first order condition (8) gives

$$D_2 H(k(t),\lambda(t),t) = f(k(t)-c(\lambda(t),t))$$

If $u(\cdot)$ is strictly increasing, equation (8) implies that $\lambda(t)$ is positive. Then the complementary slackness condition $\lambda(t)[f(k(t)-c(t)-\dot{k}] = 0$ implies that

$$\dot{k} = D_2 H(k(t),\lambda(t),t). \tag{9}$$

Equations (7) and (9) form a coupled system of first order differential equations. The function H is called a Hamiltonian, and the equations are called Hamiltonian equations. Like the analogous difference equations for the discrete problem, they require two boundary conditions to completely specify the solution. As in the discrete case, one is given by the constraint on k at time zero, the other by the transversality condition at infinity, $\lim_{t\to\infty} \hat{\lambda}(t)\hat{k}(t) = 0$.

Compared to discrete time, which is conceptually simpler and lends itself more readily to uncertainty, continuous time is to be preferred only for the ease with which non-linear systems of differential equations can be characterized by geometrical means. In the present form, the equations for k and λ are non-autonomous; that is, they depend explicitly on time. However, by a change of variable, they can be transformed into an autonomous system of equations with no explicit time dependence that can be studied by drawing pictures.

To do the change of variable, let $\theta(t) = e^{\rho t}\lambda(t)$ and define a new Hamiltonian $\tilde{H}$:

$$\tilde{H}(k,\lambda) = \max_{c} \; u(c) + \theta(f(k)-c).$$

The variable θ is called a current valued shadow price, as opposed to λ which is a present valued shadow price. $\tilde{H}$ is called a current valued Hamiltonian, as opposed to the present valued Hamiltonian H. It has no explicit time dependence. The equation that justifies the terminology current and present valued is

$$H(k(t),\lambda(t),t) = e^{-\rho t}\tilde{H}(k(t),\theta(t)).$$

H and λ are like discounted versions of $\tilde{H}$ and θ.

By the formula for a change of variables and a simple substitution, the

Hamiltonian equations can be restated in current value terms:

$$\dot{\theta}(t) = \rho\theta(t) - D_1\tilde{H}(k(t),\theta(t)),$$
$$\dot{k}(t) = D_2\tilde{H}(k(t),\theta(t))$$

This autonomous system can be represented in by a picture in a plane. Using language taken from physics, this is known as the phase plane. At each point in the k-θ plane, imagine an arrow that indicates the direction and speed of a point following these equations. This arrow will have components $\dot{k}$ and $\dot{\theta}$.

To characterize all of the arrows at all of the different points, it is useful to identify two different lines called isoclines ("iso" for "same", "cline" for "slope"). The first isocline is the locus of points such that $\dot{\theta} = 0$, the second the points such that $\dot{k} = 0$. If k is on the horizontal axis as in Figure 9, the $\dot{k} = 0$ isocline denotes those points where trajectories in the plane have a vertical tangent or slope. The $\dot{\theta} = 0$ isocline denotes points where trajectories have a horizontal tangent. Their intersection, if any, is a stationary point.

To illustrate how these can be used, it is useful to describe the kind of analysis that Cass (1965) gives to the problem P5. Suppose that both f(k) and u(c) satisfy Inada conditions, $u'(0) = f'(0) = \infty$, $u'(\infty) = f'(\infty) = 0$. From the first order condition for maximizing $\tilde{H}$ with respect to c, it follows that $u'(c) = \theta$, which can be inverted to give $c(\theta)$. From the properties of u, it follows that $c(\theta)$ goes to zero as θ goes to infinity and vice versa. Substituting this into $\dot{k} = f(k)-c(\theta)$ implies that the $\dot{k} = 0$ isocline is a downward sloping curve in the plane that must have both axes as asymptotes. The $\dot{\theta} = 0$ isocline is specified by the unique value of k such that $f'(k) = \rho$. From these properties, it follows that the two curves must intersect at a unique stationary point as shown in Figure 9. The arrows indicate possible directions for trajectories in the plane. For example, starting from a point on the $\dot{k} = 0$ isocline, an increase in k holding θ

constant will cause $\dot{k}$ to become positive. Thus, the horizontal arrows to the right of this locus point in the direction of increasing k, arrows to the left, of decreasing k.

Judging from the pattern of arrows, it is evident that there will be two paths that converge to the stationary point and two that diverge from it. The two convergent paths are sometimes referred to as branches of the stable manifold. They contain the points of economic interest. For any given initial value of k, there is a unique value of θ such that (k,θ) lies on the stable manifold. This determines the initial value $\theta(0) = \lambda(0)$. Along this trajectory, $k(t)$ and $\theta(t)$ stay bounded, so $e^{-\rho t}k(t)\theta(t) = k(t)\lambda(t)$ will converge to zero. Because the transversality condition at infinity, the Hamiltonian equations, and the initial condition characterize a saddle point of the Lagrangian L5, the stable manifold describes both the optimal quantities and the competitive equilibrium quantities.

It is the functions of time $k,\lambda:[0,\infty) \longrightarrow \mathbb{R}$ that form a saddle point of L5, but there is another sense in which the stationary point $(k^*,\theta^*) \in \mathbb{R}^2$ where both $\dot{\theta}$ and $\dot{k}$ are equal to zero is itself a saddle point. The dynamics around this point are described as saddle point dynamics, because they are suggestive of the dynamics one would observe if one rolled marbles down a saddle shaped surface. Formally, this can be captured by linearizing the differential equations around the stationary point. A linear differential equation system of 2 variables has the saddle point dynamics indicated here if one of its characteristic roots is positive and one is negative, properties that can readily be verified for this system.

Once the general methods have been set forth, this kind of analysis lets one say a surprising amount about the qualitative properties of a fairly complicated dynamic competitive equilibrium. Its evolution is described by a non-linear system of equations. The solution determines the path for interest rates and profits of the firm. Alternatively, if $f(k)$ is interpreted as output per worker, $f(k) = F(k,1)$ where F exhibits constant returns to scale, the model determines the path of wages rather than the path of profits. Savings and investment decisions by consumers and firms respectively are guided not just by current rates, but by expectations about the entire path of future rates. Stated in spot market terms, both investment and savings decisions at each point in time depend on the entire yield curve at that time. Moreover, the entire yield curve changes over time. If one were to contemplate calculating this kind of system by equating some specification of demand and supply curves for savings and investment at each point in time, this would seem to be a hopelessly complicated task. Here it is a matter of a little calculus and algebra, and remains so when the model is complicated. For example, it is a simple (and instructive) exercise to work out how the dynamics change when the degree of intertemporal substitution in consumption changes or when firms face adjustment costs so that $\dot{k}$ is a non-linear function of foregone consumption. Another useful exercise is to contemplate the effect of an announcement at time 0 that there will be an exogenous increase in the capital stock (funded from outside the economy) at a future date T. Just by using the fact that the price $\theta(t)$ must be a continuous function of time even if $k(t)$ is not continuous, it is possible to infer that the yield curve tilts, with yields on near term maturities increasing, and long term maturities falling, with the net effect that current investment falls relative to investment before the announcement.

III.4 Suboptimal Equilibria

If the approach described in the previous sections was applicable only to equilibria that are full information Pareto optimal, it would be of limited value. In fact, this method, or something very close to it, can be used in cases where the equilibrium is suboptimal because of some violation of the perfect markets assumptions. Even suboptimal equilibria maximize some kind of criterion. If one can add restrictions to the aggregate problem or change the objective in a suitable way, the solution to the modified problem may generate the outcomes observed in equilibrium.

This observation has emerged in the context of several different problems in economics. In contexts where agents have private information, the added restrictions take the form of incentive compatibility constraints. Many applications of this methodology are interested in contracts or mechanisms that are not competitive equilibria with price taking agents, but Prescott and Townsend (1984) show that the solution to this kind of constrained maximization problem can be decentralized as a price taking equilibrium if one extends the set of goods that are allowed in the model. In a model with differentiated commodities, Hart (1980) observed that an equilibrium might not generate the correct number of different types of goods, but taking the set of goods that are produced as a given constraint, the equilibrium will produce the optimal amount of each good that is produced. In this case, if one knew the set of goods that would be produced in equilibrium, one could calculate equilibrium quantities by writing down the problem of maximizing welfare subject to the given set of goods that can be produced and consumed. In dynamic models, this kind of approach was implicitly used by Arrow (1962) in a model with externalities, and by Brock (1975) in a model with money demand and inflation which acts like a tax on money holdings. Brock (1977) also considered a growth model with pollution externalities. No doubt, other examples of this kind could be cited.

What is important about this observation, is that it is operational. It describes a procedure that can actually be implemented to solve for equilibria in a way that is potentially as simple as the phase plane analysis given in the last section, and can ultimately form the basis for empirical work. To make this point, it is useful to start once again with a simple example, one that has some relevance for issues first raised by Alfred Marshall (1961) about increasing returns that were external to individual firms.

Extend the two period period model P1 to allow for an externality associated with the accumulation of capital. Let the total number of identical firms in the economy be $\bar{N}$, where N is assumed to be large. Let k_j be the capital held by firm j and let $\mathcal{K}$ be the aggregate stock of capital,

$$\mathcal{K} = \sum_{j=1}^{N} k_j.^{17}$$

To capture the external effect that total capital has on the production possibilities of firm j, write production as $f(k_j,\mathcal{K})$. The rationale for this formulation is based on the public goood character of knowledge. Suppose that new physical capital and new knowledge or inventions are produced in fixed proportions so that $\mathcal{K}$ is an index not only of the aggregate stock of physical capital, but also the aggregate stock of public knowledge that any firm can copy and take advantage of. Because each firm is only a vanishingly small part of the total economy, it chooses k_j to maximize profits taking the aggregate stock $\mathcal{K}$ as given.

As always, a competitive equilibrium with externalities for this economy can be stated in terms of the problem of a representative consumer and the problem of a representative firm. The problem of the consumer is exactly the same as the problem $PC(p,\pi)$ given previously:

$$PC(p,\pi) \qquad \max_{c \in \mathbb{R}^2_+} u(c_1) + \beta u(c_2)$$
$$\text{s.t.} \quad \pi + p_1(e-c_1) - p_2c_2 \geq 0.$$

The problem of the firm differs. Now $PF(p,\mathcal{K})$ depends parametrically on the aggregate stock $\mathcal{K}$ as well as on prices p:

---

[17] The use of a script letter $\mathcal{K}$ to distinguish aggregate quantities from firm quantities is unavoidable because upper case and lower case letters have already been used to indicate the distinction between total capital for a firm K and capital per worker k.

PF(p,$\mathcal{K}$) $$\Pi(p,\mathcal{K}) = \max_{k \in \mathbb{R}_+} \; p_2 f(k,\mathcal{K}) - p_1 k.$$

For notational convenience, we will continue to maintain the assumption that the number of consumers (and workers) is same as the number of firms. Under this assumption, the market clearing conditions for goods 1 and 2 are $e - c_1 - k \geq 0$, $f(k,\mathcal{K}) - c_2 \geq 0$, where e denotes the per capita endowment of period 1 resources. The equilibrium expression for the profits is $\pi = \Pi(p,\mathcal{K})$. To these conditions is added the equilibrium condition $\mathcal{K} = \sum_{j=1}^{N} k_j$. Formally, a competitive equilibrium with externalities will be a price quantity pair such that the quantities solve the maximization problems given the prices and the aggregate variable $\mathcal{K}$, and such that all the equilibrium conditions are satisfied.

This notion of an equilibrium has been described taking the set of firms as given and assuming that the firms earn profits. This simplifies the exposition, but is not essential. The production function $f(k,\mathcal{K})$ could represent the output per worker for a firm with an underlying technology $F(K,L,\mathcal{K})$ that exhibits constant returns to scale in the variables K and L holding $\mathcal{K}$ fixed. Then the profits can be reinterpreted as payments to workers, and free entry of firms is allowed. Since the scale and number of firms is indeterminate, it is harmless to assume that the number of firms is equal to the number of workers, and that they all produce at the same scale. Under this interpretation, $\mathcal{K}$ remains the total stock of capital in the economy.

Relying on the intuition described in the beginning of this section, consider an aggregate maximization problem that is restricted in the sense that a given level of aggregate capital (and knowledge) $\mathcal{K}$ is imposed. For each assumed level of $\mathcal{K}$, this defines a different problem, so the aggregate problem, like the problems of the consumer and firm, is actually a parametric family of problems:

P6($\mathcal{K}$)

$$\max_{c_1,\, c_2, k} \; u(c_1) + \beta u(c_2)$$
$$\text{s.t.} \quad e - c_1 - k \geq 0,$$
$$f(k,\mathcal{K}) - c_2 \geq 0.$$

Associated with P6($\mathcal{K}$) there is a Lagrangian $L6_{\mathcal{K}}$. As before, let x denote the triple $x = (c_1, c_2, k)$ so we can write:

$$L6_{\mathcal{K}}(x,\lambda) = u(c_1) + \beta u(c_2) + \lambda_1[e - c_1 - k] + \lambda_2[f(k,\mathcal{K}) - c_2].$$

As long as f is a concave function of its first argument for each fixed $\mathcal{K}$, this is a concave problem and the Kuhn-Tucker theorem will apply. By exactly the same argument as for Proposition 1, it will follow that any solution $\hat{x}$ to P6($\mathcal{K}$) will have associated with it prices $\hat{\lambda}$ such that $(\hat{x},\hat{\lambda})$ is a saddle point of $L6_{\mathcal{K}}$, and therefore solves the consumer's problem and the firm's problem at prices $\hat{\lambda}$. This does not ensure that it is a competitive equilibrium with externalities for this economy because the equilibrium condition $Nk = \mathcal{K}$ may not be satisfied. However, it is a simple matter to pick $\mathcal{K}$ so that this condition holds as well.

To see why this is so, consider the conditions C1, C2, C3 which characterize the saddle point of the Lagrangian associated with the problem $P6(\mathcal{K})$.

C1: i) $u'(x_1) = \lambda_1$

ii) $\beta u'(x_2) = \lambda_2$

iii) $\lambda_1 = \lambda_2 D_1 f(x_3, \mathcal{K})$

C2: $\lambda_i \geq 0, \; e - c_1 - k \geq 0, \; f(k, \mathcal{K}) - c_2 \geq 0$

C3: $\lambda_1 [e - c_1 - k] = 0, \; \lambda_2 [f(k, \mathcal{K}) - c_2] = 0.$

Assuming that the utility function $u(\cdot)$ is strictly increasing, the multipliers $\hat{\lambda}_i$ will be positive. Then the constraints on the quantities must be binding, and can therefore be substituted into C1:i) to yield

$$\frac{u'(e-k)}{\beta u'(f(k,\mathcal{K}))} = D_1 f(k,\mathcal{K}).$$

At this point, the equilibrium condition $\mathcal{K} = Nk$ can be imposed by a simple substitution. All that remains is to find a value $\hat{k}$ that solves the equation

$$\frac{u'(e-k)}{\beta u'(f(k,Nk))} = D_1 f(k,Nk). \tag{14}$$

Working back from the solution $\hat{k}$, it is easy to derive the quantities $\hat{c}_1$, $\hat{c}_2$, and $\mathcal{K} = N\hat{k}$, and the prices $\hat{\lambda}_1$ and $\hat{\lambda}_2$.

The fact that this does not give the socially optimal quantities can be clearly seen from equation (14). The social optimum would recognize the true marginal rate of transformation between periods $D_1 f(k,Nk) + ND_2 f(k,Nk)$. In the equilibrium, each firm has no incentive to take account of the second term. It reflects the positive effect that accumulation of its capital and knowledge has on the production possibilities of all other firms.

The general procedure that this illustrates is to start with a statement of a parametric family of restricted aggregate maximization problems. The parameters in this problem may be variables that are endogenously determined in the equilibrium, but for calculating the equilibrium, they are treated as fixed when taking first order conditions. Provided that this problem is concave for fixed values of the parameters, the Kuhn-Tucker theorem will apply. Derive the conditions from the theorem that characterize the quantities and shadow prices, and then substitute in the expression for the parameters of the problem in terms of the endogenous quantities. The order here is essential; one must take derivatives first, then substitute in for the parameters. By the arguments of Proposition 1, the quantities and prices that solve the resulting equations will form a sub-optimal competitive equilibrium.

The idea of inserting an equilibrium condition into a first order condition and then solving for quantities is an old trick in static tax analysis. What is surprising about the analysis here is the fact that it works even if f is not a concave function. For example, f could be an increasing returns function $f(k,\mathcal{K}) = k^{\alpha}\mathcal{K}^{\eta}$, with $\alpha+\eta > 1$. All that is required is $\alpha \leq 1$. Even more surprising, it is as easy to do in an infinite horizon dynamic model as in a static model.

To see this, consider problem P7, an extension of the Cass model with productive externalities. Suppose that production for each of N firms takes the form $F(K,L,\mathcal{K})$ where F exhibits constant returns to scale in own capital K and labor L taking the aggregate $\mathcal{K}$ as given. Normalizing by L in the usual way, define output per worker $f(k,\mathcal{K}) = F(k,1,\mathcal{K})$, and assume for convenience that the number of firms is equal to the number of workers, which is equal to the number of consumers and is held constant.

Let $\mathcal{K}:[0,\infty) \longrightarrow \mathbb{R}$ denote a path for aggregate capital and knowledge and define the restricted social planning problem as

P7($\mathcal{K}$)

$$\max \int_0^\infty u(c(t))\, e^{-\rho t} dt$$

$$\text{s.t.} \quad \dot{k}(t) = f[k(t),\mathcal{K}(t)] - c(t) \quad \text{for all } t \geq 0$$

$$k(t) \geq 0 \qquad \text{for all } t \geq 0.$$

The analysis here is exactly as for the two period problem. The conditions for the saddle point are the Hamiltonian equations. Using the current valued formulation, define

$$\tilde{H}(k,\theta,\mathcal{K}) = \max_{c} \; u(c) - \theta[f(k,\mathcal{K}) - c].$$

The differential equations are then

$$\dot{\theta}(t) = \rho\theta - D_1\tilde{H}(k(t),\theta(t),\mathcal{K}(t))$$

$$\dot{k}(t) = D_2\tilde{H}(k(t),\theta(t),\mathcal{K}(t)).$$

This is not yet an autonomous system that can be studied in the phase plane because of the dependence on the exogenously given path $\mathcal{K}(t)$. But after the substitution of $Nk(t)$ for $\mathcal{K}(t)$ it is autonomous.[18] For specific functional forms, the phase plane can immediately be characterized.

The easiest form of utility to work with is always $u(c) = \ln(c)$ since in this case the maximum over c in the definition of the current valued Hamiltonian yields the first

---

[18]Since the number of workers is assumed to be equal to the number of firms, there is one worker per firm and the equilibrium condition is still $\mathcal{K} = Nk$.

order condition $1/c = \theta$. Let F have the log-linear form $F(K,L,\mathcal{K}) = K^{\alpha}L^{1-\alpha}\mathcal{K}^{\eta}$ so f becomes $f(k,\mathcal{K}) = k^{\alpha}\mathcal{K}^{\eta}$. Inserting these into the Hamiltonian, differentiating to get the differential equations for k and θ, then substituting $\mathcal{K}(t) = Nk(t)$ gives

$$\frac{\dot{\theta}}{\theta} = \rho - \alpha N^{\eta} k^{\alpha+\eta-1},$$
$$\dot{k} = N^{\eta} k^{\alpha+\eta} - \frac{1}{\theta}.$$

If $\alpha+\eta$ is less than 1, this system has the same saddle point dynamics as the Cass model. As before, growth stops when the private marginal productivity of capital is equal to the discount rate; $D_1 f(k,Nk) = \rho$ is the equation of the $\dot{\theta} = 0$ isocline, which is analogous to the equation $f'(k) = \rho$ for the model with no externalities. This shows that increasing returns of the form $F(K,L,Nk) = K^{\alpha}L^{1-\alpha}(Nk)^{\eta}$ are not by themselves enough to sustain persistent growth. What is needed is that the private marginal product $D_1 f(k,\mathcal{K})$ not fall too rapidly as k grows. Accordingly, the dynamics change when $\alpha+\eta$ equals 1. When this is true, this economy will exhibit unceasing growth provided that αN^{η} is greater than ρ. In the phase plane, there is no $\dot{\theta} = 0$ isocline, hence no stationary point.

In fact, it is possible to explicitly solve the equations for this model and show that growth takes place at a constant rate. Since $-\dot{\theta}/\theta > 0$ is equal to the constant $g = \alpha N^{\eta} - \rho$, we can write $\theta(t) = \theta(0)e^{-gt}$. Then the expression for c is $c(t) = \theta(0)^{-1}e^{gt}$ and the equation for k becomes

$$\dot{k} = N^{\eta}k - \theta(0)^{-1}e^{gt}.$$

Since this is a linear differential equation, its solution can be found in any text book.

Using the fact that $k(0)$ is given, the solution is

$$k(t) = [\theta(0)(N^{\eta}-g)]^{-1} e^{gt} + \left\{k(0) - [\theta(0)(N^{\eta}-g)]^{-1}\right\} e^{(N^{\eta})t}.$$

The undetermined value $\theta(0)$ in this expression is determined by imposing the transversality condition at infinity, $\lim_{t\to\infty} e^{-\rho t}\theta(t)k(t) = 0$. Since $\theta(t)$ grows like e^{-gt}, $e^{-\rho t}\theta(t)$ times the first term in the expression for k goes to zero for any choice of $\theta(0)$. Since $g-\rho = \alpha N^{\eta}$, the second term goes to zero only if

$$\theta(0) = [k(0)(N^{\eta}-g)]^{-1}.$$

It follows that both k and c grow at the rate $g = D_1 f(k,Nk)-\rho = \alpha N^{\eta}-\rho$.

In the equilibrium represented by this solution, interest rates are constant, $r = D_1 f(k,Nk) = \rho+g$. If we normalize initial capital so that $\mathcal{K}(0) = 1$, exponential growth in $\mathcal{K}$ implies production possibilities for individual firms at each point in time of the form $e^{gt}K^{\alpha}L^{1-\alpha}$. The wage rate grows at the the rate g, and labor and capital receive a constant share $1-\alpha$ of national income. The share of capital could be made to fall, and the share of labor made to raise if the part of output due to private choices $K^{\alpha}L^{1-\alpha}$ were replaced by a constant elasticity production function.

It appears that this economy is observationally equivalent to one with constant exogenous technological change at the rate g, but they respond differently to interventions. It is best thought of as a model of endogenous technological change, and as such, the rate of technological change will be influenced by any intervention. If a proportional tax τ on output is introduced, this will change the incentives to invest, which will affect not only $k(t)$ but also the aggregate path $\mathcal{K}(t)$. In the exogenous technological change model, the analog of $\mathcal{K}(t)$ is exogenous and hence does not respond to the tax. The combined effect here is to reduce the rate of growth from $\alpha N^{\eta}-\rho$ to $(1-\tau)\alpha N^{\eta}-\rho$. For large enough values of the tax rate, growth can even stop or reverse.

In this informal analysis, no explicit account is taken of the tax revenue, as if it were simply thrown away. The results are the same if one does the more sensible balanced budget tax exercise where the proceeds are rebated to consumers. The requires only an additional parameter in the restricted aggregate maximization problem. Let $P8(\mathcal{K},T)$ denote a problem that depends on an exogenously specified path for aggregate capital $\mathcal{K}(t)$ and a path of per capita transfers $T(t)$ received by consumers:

$$P8(\mathcal{K},T) \qquad \max \int_0^{\infty} u[c(t)+T(t)]\, e^{-\rho t} dt$$

$$\text{s.t.} \quad \dot{k}(t) = (1-\tau)f[k(t),\mathcal{K}(t)]-c(t) \quad \text{for all } t\geq 0$$

$$k(t) \geq 0 \qquad \text{for all } t\geq 0.$$

In addition to the previous equilibrium condition $\mathcal{K}(t) = Nk(t)$, this problem requires the balanced budget condition $\tau f(k(t),Nk(t)) = T(t)$. Proceeding as above, define a current valued Hamiltonian, and take derivatives to get the differential equations

$$\frac{\dot{\theta}}{\theta} = \rho - \alpha(1-\tau)k^{\alpha-1}\mathcal{K}^{\eta}(t),$$
$$\dot{k} = (1-\tau)k^{\alpha}\mathcal{K}^{\eta} - \frac{1}{\theta} + T(t).$$

After imposing the equilibrium conditions, it follows as before that the growth rate is $\alpha(1-\tau)N^{\eta}-\rho$.

The difference between exogenous and endogenous technological change is of more than academic interest. For example, an analyses of the tax reform act of 1986 conducted in the context of a model with exogenous technological change, suggests that reductions in tax distortions between the household and corporate sectors and between short lived and long lived types of capital could lead to efficiency gains on the order of 1% of GNP per year. (Jorgenson, 1987) Since growth in this model is exogenous, removing the investment tax credit and increasing the capital gains tax can have no effect on the rate of growth of labor augmenting technological progress, estimated here to be on the order of 2% per year. It therefore has no long-run effect on the rate of growth of capital. For comparison, suppose that the rate of growth of the technology in this model is endogenous, and that the change in the investment tax credit or the capital gains rate leads to changes in research and development and to venture capital availability that cause a reduction of 0.1% in the rate of growth of technology from 2% per year to 1.9% per year. In an equilibrium where the real interest is 5% and aggregate GNP grows at 3% per year[19] (roughly the figures used by Jorgenson), the present discounted value of future GNP is $\frac{1}{0.020} = 50$ times current GNP. A increase in GNP of 1% in all future periods due to reduced tax distortions would give wealth of $\frac{1.01}{0.020} = 50.5$ times pre–reform GNP, an increase in wealth worth half a year's GNP. But at the pre-reform interest rates, a fall in the growth of output from 3.0% to 2.9% reduces the present value of future GNP to $\frac{1.01}{0.021} = 48.1$ times pre–reform GNP, a loss compared to pre–tax reform situation of nearly two years worth of GNP. As one should have suspected, even small growth effects can swamp large increases in levels.

A reduction in the growth rate of 0.1% is quite significant from an economic point of view, but it is quite small compared to range of variation in output growth rates observed in the data, even among apparently similar industrialized nations. Compare for example the behavior of the UK with that of Japan, or even of France, in the post-war period as indicated in Table 3 and Figure 4.) Something causes these differences, and if policy choices account for even a small fraction of the variation, the indirect effects of

---

[19]The 3% growth rate for output and capital follows from a growth rate of 2% per year in labor augmenting technological change and a growth rate of 1% per year in total quality adjusted labor.

policy on growth rates may completely dominate the direct effects that we can quantify in an exercise with an exogenous growth model.

One can argue about many of the specifics of this particular model of endogenous technology, but the point here is that methods are available for treating this kind of issue. At the risk of repetition, it is useful to emphasize how useful the methodology outlined here is. It gives price paths for interest rates and wages; allows firms and consumers to make their investment and savings decisions at every point based either on expectations about the future or equivalently based on the entire array of securities returns that can be observed; lets the rate of what looks like exogenous technological change actually depend on endogenous investment decisions, and assumes that individuals understand how this dependence operates; describes an equilibrium that is not Pareto optimal and therefore is not the solution to a simple Pareto optimization problem; gives an explicit dynamic formulation of one of Marshall's examples of external increasing returns; and permits a balanced budget analysis of taxes with both direct effects on individual choices and indirect effects on the rate of knowledge creation or technological change.

Finally, it is useful to point out why the full machinery of the Kuhn-Tucker theorem is important in this approach to sub–optimal equilibria, and why the specialized tools of dynamic programming and the Bellman equation do not readily apply. The essence of this approach is the ability to specify an aggregate maximization problem that depends on endogenous quantities. In dynamic models, quantities are naturally functions of time, so the statement of the problem must depend on an entire path like $\mathcal{K}(t)$ or $T(t)$. This induces a form of exogenous time dependence into the problem that dynamic programming is not well set up to handle.

Section IV. Recent Models of Growth

The models of growth that have been proposed in the last few years, like all general equilibrium models, can be characterized in terms of the assumptions they make on preferences, the technology, and the equilibrium concept. In all but one case, the equilibrium concept is a complete markets competitive equilibrium, or a complete markets competitive equilibrium with externalities. The exception, the Marshall-Romer model, uses a notion of monopolistic competition, but its dynamic behavior is identical with of a model of competitive equilibrium with externalities.

In terms of the technology, all of the models assume the existence of an aggregate production function $F(\cdot)$ that depends on a subset of the following list of inputs: services from physical capital K, labor services L from a person with a minimal level of schooling and training, services from additional human capital H, and measure of the technology or state of the art A. This production function can exhibit increasing returns or constant returns. In all cases except the Marshall-Romer model, $F(\cdot)$ can be thought of as a description of the technology available to a representative firm. In the exceptional case, $F(\cdot)$ is a reduced form that reflects elements of the technology and of the market structure relating final goods producers and intermediate goods producers.

With the exception of the Barro-Becker model, all of the models use standard discounted Ramsey preferences and assume that the population growth rate is exogenously given. Becker and Barro extend the specification to give parents preferences over both consumption per child and the number of children.

V.1 The Arrow-Romer Model

The model of endogenous technological change described in the last section is a special case of the model in Romer (1983) and (1986). The technology depends on physical capital K, physical labor L, and technology A. The production function F(K,L,A) exhibits constant returns in K and L taken alone, and therefore exhibits increasing returns when all three variables are taken together. Equilibrium is possible because K and L are the only factors that receive explicit compensation; A is like a public good. Movements in A are induced by assuming that private investments in capital induce increases in public knowledge A. For simplicity, the movements in A are assumed to take place one for one with movements in K, so the analysis can concentrate on a model with a single state variable.[20]

Previous attempts at making technological change endogenous were made during the 1960's, but the presence of increasing returns always limited progress in the theory. (See for example Shell 1967, Phelps, 1966, von Weizsacker, 1966.) The issue of aggregate increasing returns seems unavoidable in any discussion of endogenous technological change if one interprets technology in terms of the variable A: disembodied knowledge about things like mathematics, physics, chemistry, engineering, or manufacturing processes, that is contained in books, designs, blueprints, copyrights, patents, etc. These kinds of knowledge can be used repeatedly at essentially zero marginal cost, and in this sense, A is quite different from the skills H that are embodied in workers. A clear example of goods A and H that serve the same function is to think of H as the skills of an expert, A as a computer programmed expert system that makes the same decisions. The expert system is expensive to create, but essentially costless to replicate.

[20]The suggestion that innovation might move together with investment in physical capital is not entirely hypothetical. Schmoookler (1966) presents detailed evidence from a several industries that patents are closely correlated with investment in physical capital. The patents follow investment with a lag and the number of patents in technologically unrelated areas, e.g. track and non–track patents for railroads, show the same co–movement with physical investment. These pieces of evidence suggest that the causation may not merely run from exogenous discoveries to new invesment.

If one acknowledges the existence of this kind of input, increasing returns follows directly. One should be able to double output by doubling all tangible inputs and replicating all existing productive activities with no change in the underlying knowledge A. Once one allows A to vary as well, there must be increasing returns to scale. Whether or not it is actually possible to double all factors, this argument shows that it is mathematically impossible for all factors of production to be paid their marginal products. With increasing returns, this would more than exhaust total output. Models constructed during the 1960's resolved this by assuming that A came from the sky or perhaps from the National Science Foundation, and therefore did not need to be compensated in the market. The suggestion here is that A is a side effect of investment, but still does not receive direct compensation. Compared to the exogenous descriptions, this makes accumulation of A responsive to economic incentives. An alternative that accomplishes the same thing is explored in section IV.5 below; it supposes that A is compensated out of monopoly profits.

Arrow (1962) used the formulation considered here, assuming that improvements in the aggregate technology are the result of investment in physical capital.[21] He attributes the inspiration to Kaldor (1961). However, Arrow restricts attention to the case where the aggregate elasticity of output with respect to capital and knowledge ($\eta+\alpha$ in the example worked out above) is less than one. As a result, there is still a steady state. Arrow keeps growth going by adding exogenous population growth, but this is not completely satisfactory. Population growth becomes the only driving force in the model, and plays a role analogous to exogenous technical change in the Solow model. The growth rate of per capita income increases directly with the growth rate of the population and goes to zero if

[21]Arrow added an irrelevant fixed coefficients technology on top of his external effect. The result is a paper that is difficult to read, and easy to misunderstand. Many economists seem to have the mistaken impression that this paper is concerned with on the job learning-by-doing by workers. Levhari (1966,1967) and especially Sheshinski (1967) offered simpler versions of Arrow's analysis that captured the essentials, but these papers seem to have received relatively little attention.

population growth goes to zero. Neither savings rates, nor taxes can influence the growth rate. A permanent increase in the share of output devoted to investment, arising for whatever reason, has no permanent effect on the rate of growth.

Arrow's restriction to steady state analysis seems to have been made largely for technical reasons. One of the issues is how to make sure that the integral defined over $[0,\infty)$ in the objective function converges. If growth took place at a rate that was too fast, this could diverge. The second issue has to do with the equilibrium theory for the model. Arrow gives heuristic arguments about how the quantities he derives can be supported as a competitive equilibrium, but does so only for steady state growth paths. Thus, for example, the interest rate he derives is a number rather than a function of time. Brock (1975) and (1977) contain analyses of dynamic models with inflation and pollution distortions respectively, and these papers apply the equilibrium analysis to the entire path for the economy, not just to the steady state. The equilibrium analysis is heuristic and does not consider increasing returns and endogenous growth. Romer (1983) offered the first general formulation of the equilibrium theory behind this kind of model, with a statement and proof of a result like Proposition 1 above for dynamic models.

Romer (1983) and (1986) actually go beyond the suggestion given above that the exponents α and η in the aggregate production function $k^{\alpha}\chi^{\eta}$ can sum to 1. Under a modification of the technology for converting consumption goods into investment goods, these papers show how a sum greater than 1 can be accommodated. In this case, explicit solutions are not possible, but the equilibrium can be studied using phase plane analysis. The rate of growth can be monotonically increasing over time as opposed to decreasing, as it must ultimately be when $\alpha+\eta$ is less than 1, or constant, as is the tendency when the sum is equal to 1. In this sense, the model can capture even the long-run trend behavior of growth rates demonstrated in Tables 1 and 2. The analysis also shows that when the sum is greater than 1, the resulting increasing marginal productivity of capital can overturn standard convergence results. Capital and investment might flow from countries with low per capita income and capital to more developed countries.

V.2 The Uzawa-Lucas Model

Using a different model, Lucas (1988) makes a similar, but empirically more relevant point about the effects of increasing returns on flows between countries. Increasing returns can lead to pressure for migration even if there is full mobility of capital between countries.[22] The basic model depends on the variables K and H, and builds on an earlier model of Uzawa (1965).[23] This kind of model can be thought of as allowing for physical labor L as well as human capital H provided these two inputs are good substitutes. In particular, if they are perfect substitutes, H can simply be taken to be the sum of tangible and intangible human capital.

The input H resembles physical capital more closely than it does labor L or the technology A. In contrast to L, it is possible to increase H by investment, just as it is possible to increase K. In contrast to A, if one wants to replicate a productive activity, it is necessary to incur a cost to produce more H (i.e. train additional workers). The model is therefore a two capital good model, and it specifies two different sectors where investment can take place. The key sector for determining the rate of growth is the sector for producing new human capital. The specific technology assumed for the accumulation of H is linear in H, and this simplifies the analysis a great deal. If H_1 is the amount of human capital devoted to the production of consumption goods and H_2 is the amount of

---

[22]Lucas's paper actually considers two different models. The second is discussed below.

[23]Uzawa's original interpretation of his model was one with endogenous accumulation of technolgy A rather than of human capital H. Lucas proposed the interpretation used here in terms of human capital.

human capital devoted to the production of new human capital, then $\dot{H}$ is given by

$$\dot{H} = \delta H_2.$$

Consumption goods and physical capital are produced in the first sector according to a production function $F(K,H_1)$. In the Uzawa formulation, this function exhibits constant returns to scale. Lucas suggests that there are increasing returns to scale, with external effects that are associated with human capital H. By analogy with the previous formulation, output can be written in the form $F(K,H,\mathcal{H})$, where in equilibrium, the aggregate stock of human capital $\mathcal{H}$ is given by $\mathcal{H}(t) = NH(t)$. The externalities in the Arrow-Romer model arise from an indirect link between production of physical capital $\mathcal{K}$ and the technology A, which can be copied and used by all. In contrast, Lucas emphasizes direct interaction effects of human capital, of the kind that would arise from conversations between colleagues and co-workers. In his formulation of the model, Lucas makes the externality depend on the average level of human capital rather than the total amount as used here, but as long as the population is held constant, these specifications are equivalent.

To write down the aggregate maximization problem for this economy, it is notationally convenient to define a variable, u in Lucas's notation, that is equal to the fraction of total human capital that is devoted to the production of physical goods. Thus, $u(t) = H(t)_1/H(t)$. The relevant parametric maximization problem for calculating equilibria in this model is

P9($\mathcal{X}$)

$$\max \int_0^\infty u(c(t))\, e^{-\rho t} dt$$

$$\text{s.t.} \quad \dot{K}(t) = F(K(t), u(t)H(t), \mathcal{X}(t)) - c(t),$$

$$\dot{H}(t) = \delta u(t) H(t),$$

$(1 - u(t))$

$$H_1(t) + H(t)_2 \leq H(t),$$

$$K(t) \geq 0.$$

Calculating an equilibrium proceeds just as for the previous models. Taking the path for $\mathcal{X}(t)$ as given, write down the Hamiltonian for this system, which will depend on the two capital stocks or state variables K and H, and two multipliers or co-state variables. The maximum in the definition of the Hamiltonian will be over the choice of the control variables c(t) and u(t). Differentiate to get Hamiltonian equations, then substitute in the equilibrium condition $\mathcal{X}(t) = NH(t)$.

This model does generate unbounded growth, but does not rely on the aggregate increasing returns to do so. In Uzawa's original formulation of the model with no increasing returns and no external effects, there is also unbounded growth. What makes this possible are the assumptions of constant returns in the investment sector, combined with the assumption that all inputs can be accumulated so that there are no fixed factors. The form of constant returns in the human capital sector is particularly simple, $\dot{H} = \delta H_2$. This equation obviously exhibits no diminishing returns, and one accumulation equation of this type is enough to keep things going. If H grows without bound, its effect on the output equation is like the effect of exogenous technological change in the Solow model or exogenous population growth in the Arrow model; it raises the marginal productivity of physical capital over time, inducing physical capital accumulation. The asymptotic dynamics are essentially determined by the linearity of $\dot{H}$ in terms of H. Combined with constant elasticity utility functions, this makes the model behave like the other models we have seen with linear production and constant elasticity utility. If the model starts from the correct ratio of K to H, it will grow at a constant rate forever. Starting from a different ratio of K to H will give transitory dynamics associated with the adjustment of the ratio towards the value consistent with aggregate growth at a constant rate. Growth approaches the constant rate asymptotically.

All of these are features that the Lucas model shares in common with the Uzawa model. What the presence of increasing returns in the production of physical goods does is change the implications of the model for wages. In the Uzawa model, K and H grow at the same rate. There is no deepening of physical capital relative to human capital, so the rental rate on both types of capital is constant. Payments per worker increase because human capital per worker increases, but the quality adjusted wage, for example the wage of a high school educated male with no work force experience, will be constant over time. Once increasing returns are added, this is no longer true. The ratio of K to H will increase over time, with something resembling capital deepening taking place. Quality adjusted wages will increase over time.

The cross-sectional implication of this model is that wages will be higher in a more developed country even if there is free capital mobility. To see why, consider two countries, and suppose that aggregate output takes the form

$$F(K,H,\mathcal{K},\mathcal{H}) = \mathcal{K}^{\eta}\mathcal{H}^{\varphi}(K^{\alpha}H^{1-\alpha}).$$

Here, K^{α} and $H^{1-\alpha}$ represent the usual private marginal productivities of non–human and human capital; $\mathcal{H}^{\eta}$ and $\mathcal{H}^{\varphi}$ represent possible external effects. Lucas focuses on the case of human capital externalities, but the point here is symmetric in the arguments H and K.

Let K_b, H_b, K_s, H_s denote the quantities of inputs in a big country b and a small country s. If interest rates are equalized across the two countries, this implies that

$$\alpha K_b^{\eta+\alpha-1} H_b^{1-\alpha+\varphi} = \alpha K_s^{\eta+\alpha-1} H_s^{1-\alpha+\varphi}.$$

In a model with constant returns, i.e. where production is homogeneous of degree 1, the interest rate is homogeneous of degree 0; an increase in the scale of both arguments K and H leaves the interest rate unchanged. Since the production function here is homogeneous of degree $1+\eta+\varphi > 1$, the interest rate is homogeneous of degree $\eta+\varphi > 0$. It follows that the ratio of K to H cannot be the same in the two countries. If it were, the interest rate would be higher in the big country. Rather, interest rate equalization implies

$$\frac{K_b}{H_b} > \frac{K_s}{H_s}.$$

But from this, it follows that wages in the big country will be higher for two reasons. The scale effect increases wages because the wage rate will also be homogeneous of degree $\eta+\varphi > 0$. The higher ratio of non–human capital to human capital in the big country will raise the wage even more. All that matters for this argument is that increasing returns be present, that is, that one of η or φ be bigger than 0 and that $\eta+\alpha$ and $\varphi+1-\alpha$ be less than 1. It does not matter which good is associated with the positive externality.

In the neoclassical model, differences in income between different countries must be a reflection of differences in the capital–output ratio, which implies large differences in the rates of return to capital across countries. To emphasize just how large, suppose that the coefficient on capital is .4 and consider a country in Figure 5 with per capita income of one tenth that in the United States. Using the formula $y = k^{.4}$, it follows that capital in the United States must be larger by a factor $10^{2.5}$ and interest rates must be lower by a factor $10^{1.9}$. It would take an unbelievably high tax rate on foreign capital or probability of expropriation in the less developed country for investment in the United States to be sensible if this were the only reason for the income differential. Going outside of the model proper, the income differential might be a reflection of differences in the technology in use in the two countries, but in this case there would seem to be comparable profits from exporting the technology to the small country. In this kind of model, the persistence of large income differentials seems to imply that there are large, persistent, unexploited profit opportunities.

The Uzawa model explains income differences with no unexploited profit opportunities. Low income countries have less of both K and H, but in the same ratio as high income countries. In a sense this goes too far. There is little evidence of overwhelming barriers to capital flows, but there are binding constraints on migration. The Lucas model reaches an intermediate conclusion. Rates of return to capital can be equalized across countries, but if they are, payments to human capital will be higher in the high income country, and workers would be better off if they could move there.

If increasing returns are present and if barriers to flows of both capital and labor are removed, both capital and workers would move to the high income region. This may seem implausible at first, until one contemplates the distribution of capital and workers within a country where no barriers are present. In fact, they are not spread evenly across the available land, as a constant returns model would imply. As Lucas emphasizes, they are highly concentrated in a few locations, cities. In less developed countries where communication and transportation are costly, this process of concentration in just a few locations or even a single location is even more pronounced than it is in developed countries.

IV.3 The King-Rebelo Model

Following up on Lucas's interpretation of the Uzawa model in terms of human capital, King and Rebelo (1988) argue that the version of this model with no increasing returns or externalities is of interest on its own. They focus on the variables H and K and assume that F(H,K) exhibits increasing returns. This is useful, they argue, because a stochastic version provides a tractable framework for analyzing aggregate time series data that can accommodate both short-run business cycle variation and long-run trend behavior. Virtually all other theoretical frameworks for data analysis do not treat one of these sources of variation in the data seriously. Traditional business cycle models remove the trend behavior by some ad hoc means. Growth models average out the business cycle variation.

The full force of their argument is apparently directed at macroeconomics. For the study of growth theory, throwing out the high frequency variation in the data may be inefficient, but there is little evidence that how one does this will prejudice the conclusions one draws. Regardless of how one chooses to smooth the data, the long-trends are clear. In contrast, the inferences for business cycle frequencies are more delicate. The answers to important questions about business cycles are often highly sensitive to how one de-trends the time series. Moreover, most of the variation in aggregate data (measured in the sense of the magnitude of the matrix $X'X$ in a regression equation) comes from the trend behavior of the series. Throwing away the trends may sacrifice more information that throwing away the high frequencies.

One of the attractive properties that King and Rebelo emphasize about a stochastic version of the Uzawa model is that it can generate aggregate time series that have a unit root, i.e. are stationary in first differences. This is in contrast to the result for the Solow model with exogenous exponential technological change, in which aggregate series are stationary after exponential detrending. Now these two alternative methods for detrending data can be treated symmetrically from a theoretical point of view as well as from an econometric point of view.

Ultimately, it must be true that economists should aim for a single aggregate model that can explain both business cycles and growth, and Rebelo and King provide a useful challenge to those who believe that this completely infeasible given the existing tools. The issue is no longer whether this is possible, but rather whether it is yet time to mount an attack on the basic questions in macroeconomics and growth theory on a combined front or whether it will be more productive to continue the fights on two separate fronts. The costs and benefits are clear enough. From a rigorous point of view, the battles must ultimately converge. From a practical point of view, combining the two endeavors too soon risks diverting our attention away from growth models and macro models that have rich implications but are harder to work with. Since economics is fortunate enough to lack a central command structure, we will no doubt continue to see both styles of work.

IV.4 The Krugman-Lucas Model

The second model in the Lucas paper is very similar to a model that was first worked out by Paul Krugman (1988). These models place less emphasis on the details of growth and more on the interaction of growth with trade. The key point here is that increasing returns can dramatically overturn the usual presumptions about the positive and normative effects of trade.

As described by Krugman, the model is stated in terms of labor L and the technology A, and assumes that aggregate production $F(L,A)$ exhibits increasing returns. Like the Arrow-Romer model, the benefits of increases in A are enjoyed by all, so the increasing returns are external. Since there is no capital, A cannot be a function of investment. Here, it is assumed to depend on previous output, which is equivalent to making it depend on labor inputs. The twist on the model here is that there are many possible goods than can be produced, each with its own level of the technology A_i. For simplicity, output of good i is $y_i = A_i L_i$, and the technological coefficient A_i evolves according to a linear differential equation in previous production:

$$\dot{A}_i = -\delta_i y_i.$$ [24]

Lucas uses an equivalent formulation in terms of labor L and the human capital H. Output of good i depends on the amount of human capital $(\mathcal{H}_i/L_i)$ per worker in the i industry and multiplied by the amount of labor devoted to production of this good. Human capital $\mathcal{H}$ grows with previous output, but the effects are purely external. In either form, this captures what most economists think of when they describe learning by doing.

The implications for trade and welfare follow from the fact that the decisions about production made in the market will depend only on the relative magnitudes of the input coefficients, A_i or $\mathcal{H}_i/L_i$ in different countries, with no regard for the leaning or growth coefficients δ_i. Suppose for simplicity that there are two goods. The good in which a country has cost advantage relative to a trading partner may not be the good with the high rate of learning, i.e. a large δ_i. Under appropriate assumptions about preferences, it will be better to be in an industry with rapid learning and rapidly increasing output. In this case, opening a country to trade, which can lead to specialization in the industry with a slow rate of learning, can make a country worse off than it would be under autarky, where domestic production of both kinds of goods will take place.

---

[24]Krugman actually allows A_i in one country to depend on output in the foreign country as well as in the home country so there are international spillovers from the learning by doing. So long as domestic output has a stronger effect than foreign output, the qualitative results described here will hold.

This model has an obvious appeal at a time of increased tension over patterns of trade. It also suggests the direction that models of growth must pursue if they are to have any thing to say about growth and trade. There must be more than one good in a model, and there must be some reason to trade. The model of increasing returns, dynamics, and external effects used here is simple and effective for pointing out the potential for conflict that is present if there are are increasing returns, but it needs to be elaborated before it can address the kinds of questions about issues like private savings and investment.

IV.5 The Romer-Marshall Model

One model that tries to introduce many goods and maintain an explicit dynamic model of accumulation is Romer (1987). This model differs from the others that contemplate increasing returns because it does not rely on external effects to support a decentralized equilibrium. Rather, it uses a model of monopolistic competition. Nonetheless, the dynamic equilibrium can be computed using the same techniques as for the externality and tax distortion models.

This model is based on the second of two sources of external economies cited by Marshall (1961). He suggested that trade among different firms offering unique specialized goods causes a form of increasing returns that is external to individual firms. The degree of specialization, or equivalently, the number of different firms that are available at any point in time or location, is limited by the presence of fixed costs. In this model, these different goods are assumed to be intermediate inputs into production, and the technology is such that having more available goods is useful. Surely a large part of what distinguishes Silicon Valley from a cross-roads in Nebraska is the set of specialized goods and services immediately available for sale at each point. If you wanted to set up a business to produce new computer chips, land in Nebraska would be cheaper, but just try to find a firms nearby with the right equipment for baking, etching and testing silicon wafers. If there were no fixed costs, we could imagine little infinitesimal versions of such firms spread smoothly over the entire surface area of the United States (or the world).

Although Marshall choose to describe specialization in terms of a competitive equilibrium with externalities, it is now clear that a more rigorous way to capture the effects he had in mind is in a model with fixed costs. In an equilibrium with non-negative profits, price must exceed marginal cost to be able to recover these fixed costs, so the model must therefore contemplate some form of market power.

Models of price setting behavior are always more difficult to describe than models of price taking behavior, but the basic idea can be described heuristically. Suppose that output of aggregate consumption goods can be written as a function $Y(\mathbf{x},L)$ that depends labor L and on a list of intermediate inputs $\mathbf{x} = \{x_i\}_{i=0}^{\infty}$ that is potentially infinitely long. This input list describes all the inputs that could conceivably be used in production. One simple form for Y is

$$Y(\mathbf{x},L) = L^{1-\alpha}\sum_{i=1}^{\infty} x_i^{\alpha}.$$

This functional form is attractive primarily because of its simplicity.[25] At any point in time only a finite number of goods x_i will be available for use by such firms, but the list of potential goods is unbounded. Because Y is a constant returns to scale production function, the industry that buys goods x_i and labor and sells final output Y will be a conventional competitive industry. What is important about this form of production is that it captures the idea that there is a large and continually growing variety of inputs into production and that these inputs are not close substitutes. Increasing the quantity of one

[25]When this functional form is used to describe preferences, it is commonly referred to as the Dixit-Stiglitz preferences, based on their article of 1977. A continuous form of these preferences was used by Joseph Ostroy (1973). For the use of this form as a production function depending on intermediate inputs, see Ethier (1982).

input does not reduce the marginal productivity of the others. This implies that increasing the set of available inputs is always useful.

To avoid issues about integer constraints, it is easier to take a continuous version of this kind of production. Thus, suppose that the range of goods can be drawn from the entire real line, so $x(i)$ is now a function defined on $[0,\infty)$ and output Y takes the form

$$Y(\mathbf{x},L) = L^{1-\alpha}\int_0^{\infty} x(i)^{\alpha}di.$$

As before, the set of values for which $x(i)$ is greater than 0 will be finite at any point in time. For simplicity, we can denote this set as an interval $[0,M]$. To see how output increases with the range of inputs that are available, suppose that all intermediate inputs could be produced at a constant cost of 1 measured in terms of forgone resources. Then Z units of resources could be used to produce an input list $x_i = Z/M$ for all i between 0 and M. This would yield output

$$Y = L^{1-\alpha}M(Z/M)^{\alpha} = L^{1-\alpha}Z^{\alpha}M^{1-\alpha}.$$

Holding the amount of initial resources Z constant, output could be increased indefinitely by increasing the range of different specialized inputs that are used.

What keeps this from being relevant is the presence of fixed costs. Producing new goods is assumed to involve a fixed cost, and this limits the feasible range of goods that can be produced. Average cost curves will then be U-shaped, and for simplicity are assumed to be the same for all the different types of inputs. Ultimately, one would like a formulation that distinguishes new inputs from old inputs, but for the purposes here, it is easier to preserve the symmetry among the inputs. The fixed costs in the production of the goods x_i mean that the firms supplying these goods will not be price takers. A firm selling an intermediate input i will be the only firm producing that input, and will explicitly face a downward sloping derived demand for the input from the competitive firms that produce the final output goods. The equilibrium will be one in which entry of additional firms producing additional intermediate inputs continues until profits for all firms are zero. Given some level of the resources Z devoted to the production of intermediate inputs, the equilibrium will have positive output for a finite range of inputs of length M. By symmetry, the same quantity $\bar{x}$ of each of these inputs will be produced. Given an explicit functional form for the cost function for producing x, the values for M and $\bar{x}$ can be explicitly calculated in terms of Z by solving the profit maximization problem for each of the individual monopolists and then allowing entry until zero profit is achieved.

Suppose that we choose units so that 1 is the average cost of producing a unit of x from forgone consumption Z that is achieved in equilibrium. Then Z, $\bar{x}$, and M are related by $\bar{x} = x_i = \frac{Z}{M}$ for all i, and output still takes the form

$$Y = L^{1-\alpha}M(Z/M)^{\alpha} = L^{1-\alpha}Z^{\alpha}M^{1-\alpha}. \tag{15}$$

Now ask what would happen if the quantity of Z were to double. Consideration of the zero profit conditions shows that twice as many intermediate goods producing firms would enter, demanding twice as much of the primary resource, with the per firm quantity $\bar{x}$ and the average cost $\frac{Z}{M}$ left unchanged. (This simple result follows from the additive separability of the production function in the different intermediate goods.) Thus, the equilibrium quantity of goods M is proportional to Z. If we choose units such that $\bar{x}$ is also equal to 1, then we have M = Z. Thus, the reduced form expression for aggregate output in terms of the primary resource Z and labor L is

$$Y = L^{1-\alpha}Z. \tag{16}$$

This equation, describes a kind of reduced form production function. It relates final output to L and to the amount of resources Z devoted to production of intermediate goods. It closely resembles the previous descriptions of aggregate output as an increasing returns to scale function of capital and labor. This resemblance extends to the interpretation of this kind of function in terms of externalities.

Suppose that the good Z merely represents cumulative forgone consumption, i.e. resources that could have been consumed but were instead devoted to producing goods x_i. Thus, the evolution equation for Z is

$$\dot{Z} = Y - c. \tag{17}$$

Suppose preferences take the usual discounted form in continuous time.

$$\int_0^{\infty} u(c)e^{-\rho t}dt.$$

The social planning problem for this economy would be to maximize these preferences subject to the constraints imposed by equations (16) and (17), but the monopolistically competitive equilibrium described here will not support this optimum for reasons that appear to be very similar to those for the equilibrium with externalities. One can show (and a demonstration takes more than just the kind of hand waving offered here) that in this equilibrium, agents forgo current consumption in favor of future consumption as if they take account of the direct effect that this has on output–that is taking account of the term Z^{α} in equation (15)–but without taking account of the indirect effect this has on the range of goods produced M. In this sense, it is is just like the externality model where agents choose K(t) taking $\mathcal{K}(t)$ as given.

Using this intuition, it is possible to explicitly solve for the dynamic equilibrium with monopolistic competition. Mathematically, it turns out to be identical to problem P7 solved in section III.4. Formally, one chooses a path for $Z(t)$ to maximize utility taking a path for $M(t)$ as given. One then imposes the equilibrium condition that $M(t) = Z(t)$. Note that this is just what one would expect from the previously noted result in Hart (1980). The monopolistically competitive equilibrium may not get the set of goods that are produced right, but is optimal taking the set of goods as given.

In the equilibrium for this model, growth takes place at a constant exponential rate. This rate of growth is too low relative to the rate a social planner could achieve and increases with any intervention that increases savings. The main value of this model is that it offers a different interpretation from the previous one relying on growth with spillovers of knowledge. What keeps growth going and avoids the problem of diminishing returns to capital accumulation is the continual introduction of new goods. The models can be related in the sense that the fixed cost in the introduction of a new good could be research and development costs needed to produce the knowledge A required to make the physical good. Thus, for producers of final output, one can still think of knowledge as an input into his production, but now it comes embodied in new inputs, and can no longer be copied for free. In the aggregate, savings still has social benefits that are larger than the private benefits, but here the distortion arises because of departures from price taking, not from externalities or true spillovers. For comparisons across regions or countries, the relevant measure of the area over which one firm's actions affect other firms is no longer defined by how far knowledge can travel (the effect emphasized in the Arrow-Romer model and the Krugman version of the learning by doing model), nor by the necessity of direct contact between co-workers and colleagues (as in the Lucas models). Rather, it is determined by transportation costs and by how far goods can travel.

This model has immediate implications for trade. Removing barriers to trade increases total output in each country, and more importantly, raises the returns to savings and the amount of savings in each country and therefore raises growth rates in each country. Compared to the learning by doing models of Krugman and Lucas, it is more suggestive of the gains from trade than the potential for conflict. Nonetheless, if transportation costs are high, there is still some potential for conflict or rivalry. If all goods produced anywhere can be traded worldwide, all regions benefit equally from the introduction of goods in any location. As a result, advanced countries have no natural advantage over less advanced countries, and the latter should tend to catch up with the former. On the other hand, if there is a significant range of goods that are too expensive to transport and trade outside of a limited area, developed areas will tend to have a built in advantage over the less developed areas. Under these circumstances, convergence will fail. Starting from symmetric positions for two different countries, the country that can first take the lead may have a permanent advantage over the other.

V.6. Endogenous Population Growth and Preferences for Children

All of the models so far neglect population growth as an endogenous variable. The key issue in modeling population growth is how to value the tradeoff between more goods per person and more people. In the optimal planning literature of the 1960's and 1970's, three different strategies for dealing with the a growing population were suggested. Let C_t denote aggregate consumption at time t, and let N_t denote the population. A planner could maximize the discounted sum of a utility function depending on C_t, $U = \sum \beta^t u(C_t)$ ignoring the size of the population. Alternatively, the planner could maximize some notion of average utility, letting $u(\cdot)$ depend per capita consumption rather than aggregate consumption, $U = \sum \beta^t u(C_t/N_t)$. Finally, the planner could maximize some notion of total utility received by individuals, multiplying individual utility

$u(C_t/N_t)$ by the population N_t, $U = \sum \beta^t N_t u(C_t/N_t)$.

For the qualitative features of a model in which the path for N_t is taken as given, the choice does not matter very much. However, once population growth is allowed to be endogenous, these different specifications matter a great deal. Consider the thought experiment of holding aggregate consumption C_t constant and freely choosing N_t for each of these specifications of the objective function. In the first case, N_t has no effect. In the second, it is optimal to drive N_t to zero. In the third, it is optimal to drive N_t to infinity (so long as $u(\cdot)$ is strictly concave.) When production is added to the model, these conclusions are modified only slightly. The first case becomes like a model of a profit maximizing slave owner who sells output C_t. Additional bodies are valuable as long as they are net producers of output. In the second and third cases, N_t will generally tend towards 0 or ∞ asymptotically. As long as the question of how people should be valued is put in moral or ethical terms, it is hard to know how to come to any resolution of this question. Maybe population really "ought" to go to 0 or ∞.[26]

As macroeconomists pushed the equilibrium interpretation of Ramsey type models, emphasizing the idea that the objective function comes from the preferences of individuals, the treatment of population growth took on a more scientific flavor. Nerlove (1974) was one of the first economists to suggest that the emerging theory of household decision making be used to examine choices about fertility, and that this analysis could be applied to questions about long-run growth. Razin and Ben-Zion (1975) gave one of the first explicit treatments of this approach. To do this, they had to face the same questions about functional form raised above, but in doing so, they were guided by evidence on the preferences of parents. Their solution was to take the second form, so that what mattered was the average utility of future generations. By itself, this would imply that families would choose to have very few children, but they also added the idea that parents get

[26]See Pichford (1974) for a defense of the second form for the objective function. Meade (1955) was influential in convincing economists to use the third formulation.

direct utility from the presence of children. If n_t represents the reproduction rate of the family, the utility U_t of the parental decision making unit at generation t takes the form $U_t = v(c_t, n_t) + \beta U_{t+1}$.[27] U_{t+1} is the utility of each of the children in the next generation. Thus, consumption c_{t+1} per child (or really per parental decision making unit) rather than consumption C_{t+1} is what enters in v_{t+1}. Because of the recursive form used here, this says that parents care about the utility from per capita consumption of all of their descendants. Solving forward gives an implicit form of preferences for the head of a family at time 0,

$$U_0 = \sum \beta^t v(C_t/N_t, n_t).$$

Population does not go to zero because the direct value of children to parents from the term n_t in $v(c_t, n_t)$ offsets the increases in per capita quantities that reductions in family size would permit.

Robert Barro and Gary Becker extend this analysis to allow for the possibility that having more children has value that goes beyond the direct consumption value. (Barro and Becker, 1986; Becker and Barro, 1987). In effect, they argue that parents may care not only about whether or not their children are successful and happy (i.e. have C_t/N_t large), but also about how many happy and successful children they have. What they propose is an intermediate solution between case two above, which discounts $u(C_t/N_t)$, and case three, which discounts $N_t u(C_t/N_t)$. Suppose, they suggest, that a parent receives utility of the form

$$U_t = v(c_t) + \beta n_t^{\chi} U_{t+1},$$

[27] All of the analysis here abstracts from the fact that it takes two parents to raise a family. Treating this issue seriously requires more than multiplication by a factor of 2, because marriage leads to links between different families. For a exploration of the implications of these links, see Bernheim and Bagwell (1988).

where χ can lie between the values 0 and 1 which represent the two extreme forms of the objective function described above. In principle, their formulation can also accommodate a direct effect of n_t on current utility $v(\cdot)$, but much of their analysis ignores this effect since it is no longer needed to keep n_t from going to zero. When these preferences are solved forward, they imply infinite horizon preferences of the form

$$U_0 = \sum \beta^t N_t^{\chi} v(c_t),$$

where $c_t = C_t/N_t$ is per capita consumption. In the special case where the growth rate of the population n can be taken to be a constant, this reduces to

$$U_0 = \sum (\beta n^{\chi})^t v(c_t). \tag{18}$$

Both the Barro and Becker and the Razin and Ben-Zion formulations introduce a positive effect of n_t, but it makes a difference how this dependence enters. In the Barro-Becker preferences, changing the rate of reproduction is mathematically like changing the discount rate, and this can lead to important effects on growth that are not present in the Razin and Ben-Zion model. To see this, confront the preferences in equation (18) with a linear technology that depends only physical capital K_t. Thus,

$$C_t = \rho K_t - K_{t+1}.$$

Restated in per capita terms, this becomes

$$c_t = \rho k_t - n_t k_{t+1}. \tag{19}$$

If n is treated as a constant, it is a simple matter to use the techniques from section III to maximize the objective (18) subject to the constraint (19). Call this problem P9, and let L9 denote the associated Lagrangian:

$$L9(K,\lambda) = \sum_{t=0}^{\infty} (\beta n^{\chi})^t v(c_t) + \lambda_t[\rho k_t - nk_{t+1} - c_t].$$

Differentiating this expression with respect to k_t, it follows that the multiplier λ will grow at the rate n/ρ. The first order condition for per capita consumption becomes $v'(c_t) = \lambda_t(\beta n^{\chi})^{-t}$. Thus, $v'(\cdot)$ will fall over time and c_t will increase if the inequality

$$\frac{n^{1-\chi}}{\rho\beta} < 1$$

is satisfied. This suggests that whether or not there is accumulation and growth on a per capita basis depends on the size of the value n, that now must be determined endogenously. This effect can lead to an interesting connection between per capita income growth and population growth. A value of the population growth rate n that is low will lead to unlimited accumulation and growth in per capita terms. A value that is too high will cause dissavings.

This kind of possibility is exploited by Tamura (1988) in a paper that has a more complicated technology depending on human capital accumulation instead of physical capital accumulation, and which takes account of the fact that there are important time costs to raising children. Thus, the larger is the human capital of the parents, the higher is the cost of a child. In that model, the determination of the population growth rates $\{n_t\}_{t=0}^{\infty}$ depends on the initial stock of human capital per person. If it is too low, it is optimal to have high values of n_t and therefore and to dissave. If the initial stock of human capital is above some critical level, it is optimal to have the growth rates n_t be small and to accumulate more and more human capital per capita. Thus, depending on the initial conditions, one family or country might be stuck in a permanent state of low per capita income with no per capita income growth and high population growth, while another might be in an equilibrium with low population growth and high per capita income growth. This description is suggestive of the observed pattern of cross-sectional variation in population growth rates and the level of per income (recall Figure 8) and justifies further work with this form of preferences.

V. Conclusion

The facts described in section II do not exhaust the set of observations that are relevant for growth; nor do the models described in section IV exhaust the set theoretical issues that are relevant. For example, observations about the growth rates of individual firms and industrial organization are directly relevant for any model with increasing returns or spillovers of information; these observations are modeled in Prescott and Boyd (1988). Growth driven by the creation of new goods and invention are closely related to the legal status of patents, issues that are modeled in Judd (1985). Interactions between product innovation and population size are considered in Schmitz (1986). The fact that some goods disappear as others are introduced is unquestionably a feature of long run growth, one that is captured in Stokey (1986). The fact that goods can be ranked, with some of them introduced only after their prerequisites are available, is modeled in Vassilakis (1986).

Even within the restricted set of facts considered in section II, none of the models described in section IV is dominant. They emphasize different issues, and only after more experience with the models and the data will it become clear what the most important issues are and how they can be combined into a single model. As they stand, the models merely suggest how theory can begin to address the questions suggested by these facts: What explains growth rates that over the course of a century have increased per capita output by a factor of 10 or more in the most advanced economies in the world? Does a high rate of investment cause a high rate of growth or vice versa? Why are growth rates so different in different countries? What influence does international trade have on growth rates? Why has fertility fallen so dramatically in some countries but not in others? Why does labor try to move towards capital instead of vice versa? And most important of all, what policies influence the rate of per capita income growth in a country?

These are the kinds of questions that someone who is not an academic economist would like to have answers to, and if economists are to earn their keep, they must ultimately be able to address them. Twenty years ago, very little explicit attention was given to these questions, but this does not mean that the growth theory of this era was useless. As suggested in the introduction and in the *double entendre* in the title, growth theory was engaged in intellectual capital accumulation. Because of the insights developed into the connection between equilibrium theory, constrained optimization, and convex analysis, there now exist tractable general equilibrium models that economists can use for thinking about these questions and for analyzing data in an attempt to answer them. More and better models will no doubt follow, as will a consensus about what are the right answers.

REFERENCES

Abramowitz, Moses. "Catching Up, Forging Ahead, and Falling Behind." Journal of Economic History XLVI 2 (June 1986), pp. 385-406.

Arrow, Kenneth J. "The Economic Implications of Learning by Doing." Review of Economic Studies XXIX (June 1962), 155–173.

Araujo, A. and J. A. Scheinkman. "Maximum Principle and Transversality Condition for Concave Infinite Horizon Economic Models." Journal of Economic Theory 30, (June 1983), pp. 1–16.

Barro, Robert J. "Are Government Bonds Net Wealth?" Journal of Political Economy 82 (1974), 1095–1117.

__________________, and Gary S. Becker. "Fertility Choice in a Model of Economic Growth." Unpublished University of Chicago manuscript (1986).

Baumol, William J. "Productivity Growth, Convergence, and Welfare: What the Long–Run Data Show." American Economic Review 76 (December 1986), 1072–1085.

Becker, Gary S. and Robert J. Barro. "A Reformulation of the Economic Theory of Fertility." Rochester Center for Economic Research working paper no. 26 (March 1987).

Bernheim, Douglas B. and Kyle Bagwell. "Is Everything Neutral?" Forthcoming, Journal of Political Economy, (1988).

Brock, William A. "A Polluted Golden Age." In Economics of Natural and Environmental Resources, edited by V. Smith. New York: Gordon and Breach (1977).

____________________. "A Simple Perfect Foresight Monetary Model." Journal of Monetary Economics 1 (1975), 133–50.

Burmeister, E. and A. R. Dobell. Mathematical Theories of Economic Growth. New York: Collier-Macmillian (1970).

Cass, David. "Optimum Growth in an Aggregative Model of Capital Accumulation." Review of Economic Studies 32 (1965), 233–240.

Delong, Bradford. "Have Productivity Levels Converged?" Unpublished MIT working paper. (1987)

Dixit, Avinash K. and Joseph P. Stiglitz. "Monopolistic Competition and Optimum Product Diversity." American Economic Review 67, (June 1977), pp. 297–208.

Either, Wilfred J. "National and International Returns to Scale in the Modern Theory of International Trade." American Economic Review 72 (June 1982), pp. 389–405.

Ekeland, Ivar and Roger Temem. Convex Analysis and Variational Problems.

Amsterdam: North Holland (1976).

Friedman, Milton, and Anna J. Schwartz. Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices, and Interest Rates, 1867–1975. Chicago and London: The University of Chicago Press (1982).

Gallman, Robert E. "Gross National Product in the United States, 1834–1909," in Output, Employment, and Productivity in the United States After 1800. National Bureau of Economic Research, Conference on Research in Income and Wealth 30 (1966), 3–90.

Greenwood, Michael J., and John M. McDowell. "The Factor Market Consequences of U.S. Immigration." Journal of Economic Literature XXIV (December 1986), 1738–1772.

Hahn, Frank H. Readings in the Theory of Economic Growth. London: Macmillan, (1970).

________________ and R. C. O. Matthews. "The theory of Economic Growth: A Survey." Economic Journal LXXIV (December, 1964), p. 779–902.

Hall, Robert E. "Stochastic Implications of the Life Cyle-Permanent Income Hypothesis: Theory and Evidence." Journal of Political Economy 86, (December 1978).

Hart, Olivier D. "Perfect Competition and Optimum Product Differentiation." Journal of Economic Theory 22 (1980), 279–312.

Intriligator, Michael D. Mathematical Optimization and Economic Theory. Englewood Cliffs NJ: Prentice Hall, (1971).

Jorgenson, Dale W. "The Economic Impact of Tax Reform." Unpublished Harvard University working paper (1987).

Jones, H. G. An Introduction to Modern Theories of Economic Growth. New YOrk: McGraw Hill (1975).

Judd, Kenneth L. "On the Performance of Patents." Econometrica 53, (May 1985), pp. 567–585.

Kaldor, N. "Capital Accumulation and Economic Growth," in F. A. Lutz and D. C. Hague (eds.), The Theory of Capital, New York: St. martin's Press (1961), 177–222.

Kendrick, John W. Productivity Trends in the United States. Princeton: Princeton University Press (1961).

______________________. Postwar Productivity Trends in the United States, 1948–1969. New York, NY: NBER and Columbia University Press (1973).

King Robert G., and Sergio Rebelo. "Business Cycles with Endogenous Growth." Unpublished working paper, University of Rochester (1988).

Koopmans, Tjalling C. "On the Concept of Optimal Economic Growth." In The Econometric Approach to Economic Planning. Amsterdam: North Holland (for Pontific Acad. Sci.)

Consequences of Mrs. Thatcher: Notes on Trade in the Presence of Dynamic Scale Economies." Forthcoming, Journal of Development Economics, (1988).

Kuhn, H. W., and A. W. Tucker. "Non–linear Programming," in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, ed. by J. Neyman, Berkeley, CA., University California Press (1951).

Kuznets, Simon. "Notes on the Pattern of U.S. Economic Growth." In The Reinterpretation of American Economic History, edited by Robert W. Fogel and Stanley L. Engerman. New York: Harper and Row (1971).

Kydland, F. and E. C. Prescott. "Time to Build and Aggregate Fluctuations." Econometrica 50, (Septermber, 1980), pp. 1345–1370.

Lebergott, Stanley. "Labor Force and Employment, 1800–1960," in Output, Employment, and Productivity in the United States After 1800. National Bureau of Research in Income and Wealth 30 (1966), 117–204.

Levhari, David. "Further Implications of 'Learning by Doing'." Review of Economic Studies 33 (January 1966), 31–9

__________. "Extensions of Arrow's 'Learning by Doing'." Review of Economic Studies 33 (April 1966), 117–32.

Luenberger, David G. Optimization by Vector Space Methods. New York: John Wiley and Sons. (1976).

Lucas, Robert E. "On the Mechanics of Economic Development." Forthcoming, Journal of Monetary Economics, (1988).

__________ and Edward C. Prescott. "Investment Under Uncertainty." Econometrica 39 (1971), 659–81.

__________ and Nancy L. Stokey. "Optimal Growth with Many Consumers." Journal of Economic Theory 32 (1984).

__________. "Money and Interest in a Cash–In–Advance Economy." Econometrica 55 (May 1987), 491–51.

Maddison, Angus. "Growth and Slowdown in Advanced Capitalist Economies: Techniques of Quantitative Assessment." Journal of Economic Literature XXV (June 1987), 649–698.

__________. Phases of Capitalist Development. Oxford: Oxford University Press (1982).

Marshall, Alfred. Prinicples of Economics, Ninth Edition. London: Macmillian (1961).

Matthews, R. C. O., Charles H. Feinstein, and John C. Odling-Smee. British Economic Growth 1856-1973. Oxford: Clarendon Press (1982).

Meade, J. E. Trade and Welfare. Oxford: Oxford University Press (1955).

Nerlove, Marc. "Household and Economy: Toward a New Theory of Population and

Economic Growth." Journal of Political Economy 82, (March/April 1974), pp. S200–S218.

Norsworthy, J. R. "Growth Accounting and Productivity Measurement." Review of Income and Wealth (1984), 309–329.

Ohkawa, Kazushi and Henry Rosvsky. Japanese Economic Growth. Stanford, CA: Stanford University Press (1973).

Ostroy, Joseph. "Representation of Large Economies: The Equivalence Theorem." Presented at the Winter Meetings of the Econometric Society, (1973).

Pichford, J. D. Population in Economic Growth. Amsterdam: North Holland (1974).

Phelps, Edmond S. "Models of Technical Progress and the Golden Rule of Research." Review of Economic Studies 33 (April 1966), 133–45.

Prescott, Edward C., and Robert M. Townsend. "General Competitive Analysis in an Economy with Private Information." International Economic Review 25 (February 1984), 1–20.

_______________, and John H. Boyd. "Dynamic Coalitions, Growth, and the Firm." Minnesota Studies in Macroecnomics 1, E. C. Prescott and Neil Wallace (eds.) Minnesota: University of Minnesota Press, (1987).

Ramsey, F. P. "A Mathematical Theory of Savings." The Economic Journal XXXVIII (December 1928), 543–559.

Razin, Assaf and Uri Ben–Zion. "An Intergenerational Model of Population Growth." American Economic Review 65, (December 1975), 923–933.

Rockafellar, R. T. Convex Analysis. Princeton: Princeton University Press (1972).

Romer, Paul M. "Dynamic Competitive Equilibria with Externalities, Increasing Returns and Unbounded Growth." Unpublished doctoral dissertation, University of Chicago (1983).

_______________. "Increasing Returns and Long–Run Growth." Journal of Political Economy 94 (October 1986), 1002–1037.

_______________. "Growth Based on Increasing Returns Due to Specialization." American Economic Review (May 1987).

_______________. "General Equilibrium Theory as a Special Case of the Kuhn-Tucker Theorem." Unpublished working paper, (1988).

_______________ and T. Shinotsuka. "The Kuhn Tucker Theorem and Transversality Conditions at Infinity." Unpublished working paper, (1988).

Samuelson, Paul A., and R. M. Solow. "A Complete Capital Model Involving Heterogeneous Capital Goods." The Quarterly Journal of Economics LXX (November 1956), 537–562.

Samuelson, Paul A. "An Extract Consumption–Loan Model of Interest with or without the Social Contrivance of Money." Journal of Political Economy 66 (December 1958), 467–82.

Schmitz, James. "Optimal Growth and Product Innovation." University of Wisconsin working paper, (1986).

Schmookler, Jacob. Invention and Economic Growth. Cambridge MA: Harvard University Press, (1966).

Shell, Karl. "A Model of Inventive Activity and Capital Accumulation." in Essays in the Theory of Optimal Economic Growth, Karl Shell, ed. Cambridge, Mass.: MIT Press (1967).

Sheshinski, Eytan. "Optimal Accumulation with Learning by Doing." In Essays on the Theory of Optimal Growth, edited by Karl Shell. Cambridge, Mass.: MIT Press (1967).

Solow, Robert M. "A Contribution to the Theory of Economic Growth." The Quarterly Journal of Economics LXX (February 1956), 65–94.

——————. Growth Theory: An Exposition. New York and Oxford: Oxford University Press (1970).

Stiglitz, Joseph E. and Hirofumi Uzawa. Readings in the Modern Theory of Economic Growth. Cambridge MA.: M.I.T. Press (1969).

Stokey, Nancy L. Learning By Doing and the Introduction of New Goods." Northwestern University working paper (1986).

Summers, Robert, and Alan Heston. "Improved International Comparisons of Real Product and its Composition: 1950–1980." Review of Income and Wealth (June 1984), 207–262.

——————, and Alan Kravis. "Changes in the World Income Distribution." Journal of Policy Modeling 6, (1984).

Swan, T. W. "Economic Growth and Capital Accumulation." The Economic Record XXXII (November 1956), 334–361.

Tamura, Robert. "Fertility, Human Capital and the 'Wealth of Nations'." Unpublished paper, University of Chicago. (1988).

Uzawa, Hirfumi. "Optimum Technical Change in an Aggregative Model of Economic Growth." International Economic Review 6 (January 1965). pp. 18-31.

von Weizsacker, Carl C. "Tentative Notes on a Two Sector Model with Induced Technological Progress." Review of Economic Studies 33 (July 1966), 245–51.

Vassilakis, Spyros. "Increasing Returns and Strategic Behavior." Unpublished doctoral dissertation, Johns Hopkins University, (1986).

Appendix: Proof of Proposition 1.

i) Competitive equilibrium implies saddle point.

Suppose that $(\hat{x},\hat{p})$ is a competitive equilibrium. Since $\hat{x}$ is a solution to $PC(\hat{p},\pi)$ and since the Slater condition holds for this problem, we can apply the necessary conditions from the Kuhn–Tucker theorem to conclude that there exists some $\hat{\gamma} \geq 0$ such that $(\hat{x},\hat{\gamma})$ is a saddle point of the Lagrangian LC for $PC(\hat{p},\pi)$:

$$LC1(x,\gamma) = U(x_1,x_2) + \gamma[\pi + \hat{p}_1(e-x_1) - p_2x_2].^{28}$$

Thus, $(\hat{x}_1,\hat{x}_2)$ maximizes $LC_{\hat{\gamma}}(x)$. Similarly, we can (trivially) invoke the Kuhn-Tucker theorem for the problem of the firm to conclude that $\hat{x}_3$ maximizes the Lagrangian $LF(\hat{p})$, where

$$LF(x) = \hat{p}_2f(x_3)-\hat{p}_1x_3.$$

This will still be true if we replace $\hat{p}$ by $\hat{\gamma}\hat{p}$. Since $L1(x,\gamma\hat{p}) = LC(x,\gamma) + \gamma LF(x)$, it follows that $\hat{x}$ maximizes $L1_{\hat{p}}(\cdot)$. Since $\hat{\gamma}$ minimizes $LC_{\hat{x}}(\gamma)$ over non–negative scalars, it follows that

$$\hat{\gamma}[\pi + \hat{p}_1(e-\hat{x}_1) - \hat{p}_2\hat{x}_2] = 0.$$

[28] Implicitly, the function $LC1(x,\gamma)$ depends on the value of p that is used in the definition of the problem $PC1(p)$. For notational simplicity, this dependence is suppressed.

Combined with the fact that

$$\pi = \Pi(\hat{p}) = \hat{p}_2 f(\hat{x}_3) - \hat{p}_1 \hat{x}_3,$$

this implies that

$$\hat{\gamma}[\hat{p}_2(f(\hat{x}_3) - \hat{x}_2) + \hat{p}_1(e - \hat{x}_1 - \hat{x}_3)] = 0$$

Since the expression on the left hand contains all the terms from $L1(\hat{x}, \hat{\gamma}\hat{p})$ containing $\hat{\gamma}\hat{p}$, $\hat{\gamma}\hat{p}$ minimizes $L1_{\hat{x}}(\cdot)$. Then $(\hat{x}, \hat{\gamma}\hat{p})$ is a saddle point of $L1(x, \lambda)$.

ii) Saddle point implies competitive equilibrium

Now suppose that $(\hat{x}, \hat{p})$ is a saddle point of $L1(x, \lambda)$. From the definition of L1, the constraints on the problem must be satisfied. Thus, $e - \hat{x}_1 - \hat{x}_3 \geq 0$, $f(\hat{x}_3) - \hat{x}_2 \geq 0$. This implies that supply in each market will be greater than demand. It remains to show that $\hat{x}$ solves $PF(\hat{p})$ and $PC(\hat{p}, \pi)$, where $\pi = \Pi(\hat{p})$. By the sufficient conditions from the Kuhn-Tucker theorem it is sufficient to show that there is a value γ such that $(\hat{x}, \gamma)$ is a saddle point of $LC(x, \gamma)$ and that $\hat{x}$ maximizes $LF(\cdot)$. Since $LC(x,1) + LF(x)$ is equal to $L1(x, \hat{p})$, and since $LC_{\gamma}(\cdot)$ depends on only the first two components of x, whereas $LF(\cdot)$ depends on only the third component of x, it follows immediately that $\hat{x}$ maximizes $LC_{\gamma=1}(\cdot)$ and that $\hat{x}$ maximizes $LF(\cdot)$. Since $\hat{p}$ minimizes $L1_{\hat{x}}(\cdot)$,

$$\pi + \hat{p}_1(e - \hat{x}_1) - \hat{p}_2 \hat{x}_2 = \hat{p}_2(f(\hat{x}_3) - \hat{x}_2) + \hat{p}_1(e - \hat{x}_1 - \hat{x}_3) = 0,$$

so $\gamma = 1$ minimizes $LC_{\hat{x}}(\cdot)$. Thus, $(\hat{x}, 1)$ is a saddle point for LC. □

Table 1

Productivity Growth Rates for Leading Countries

Leading Country	Interval	Average Annual Growth Rate of GDP per Man–Hour (%)
Netherlands	1700-1785	–.07
United Kingdom	1785-1820	.5
United Kingdom	1820-90	1.4
United States	1890-1970	2.3

Source: Maddison (1982)

Table 2

Increases in Output Per Man-Hour

Country	Symbol	Output Per Man-Hour 1870	1979	Ratio
Australia	A	1.30	6.5	5
Austria	T	.43	5.9	14
Belgium	B	.74	7.3	10
Canada	C	.64	7.0	11
Denmark	D	.44	5.3	12
Finland	L	.29	5.3	18
France	F	.42	7.1	17
Germany	G	.43	6.9	16
Italy	I	.44	5.8	13
Japan	J	.17	4.4	26
Netherlands	N	.74	7.5	10
Norway	W	.40	6.7	17
Sweden	S	.31	6.7	22
Switzerland	Z	.55	5.1	9
United Kingdom	K	.80	5.5	7
United States	E	.70	8.3	12

Source: Maddison (1982). Country symbols are used in Figures 3 and 7.

Table 3

Investment, GDP Growth, and the Capital-Output Ratio

Country	Investment Share (%)	GDP Growth (%)	Capital-Output Ratio Depreciation Rate	
			$\delta = .03$	$\delta = .04$
Japan	31	7.4	3.0	2.8
Germany	28	4.7	3.7	3.3
Canada	28	4.2	3.9	3.4
Italy	26	4.4	3.6	3.1
France	25	4.2	3.5	3.1
United States	24	3.0	3.9	3.4
United Kingdom	17	2.1	3.4	2.8

Source: Summers and Heston (1984).

Table 4

Estimates of the Share of Capital In Total Income

Country and Authors	Interval	Share of Capital (%)
Japan:		
(Ohkawa and Rosovsky, 1973)	1913–38	40
	1954–64	31
United Kingdom:		
(Matthews, Feinstein, Odling-Smee, 1982)	1856–73	41
	1873–1913	43
	1913–51	33
	1951–73	27
United States:		
(Kendrick, 1961)	1899–1919	35
	1919–53	25
(Kendrick, 1973)	1929–53	29

Results collected in Maddison (1987).

Post War Productivity Growth
8.0%
7.0%
6.0%
5.0%
4.0%
3.0%
2.0%
1.0%
0.0%
−1.0%
−2.0%
−3.0%
1949
1954
1959
1964
1969
1974
1979
1984

Fig. 1

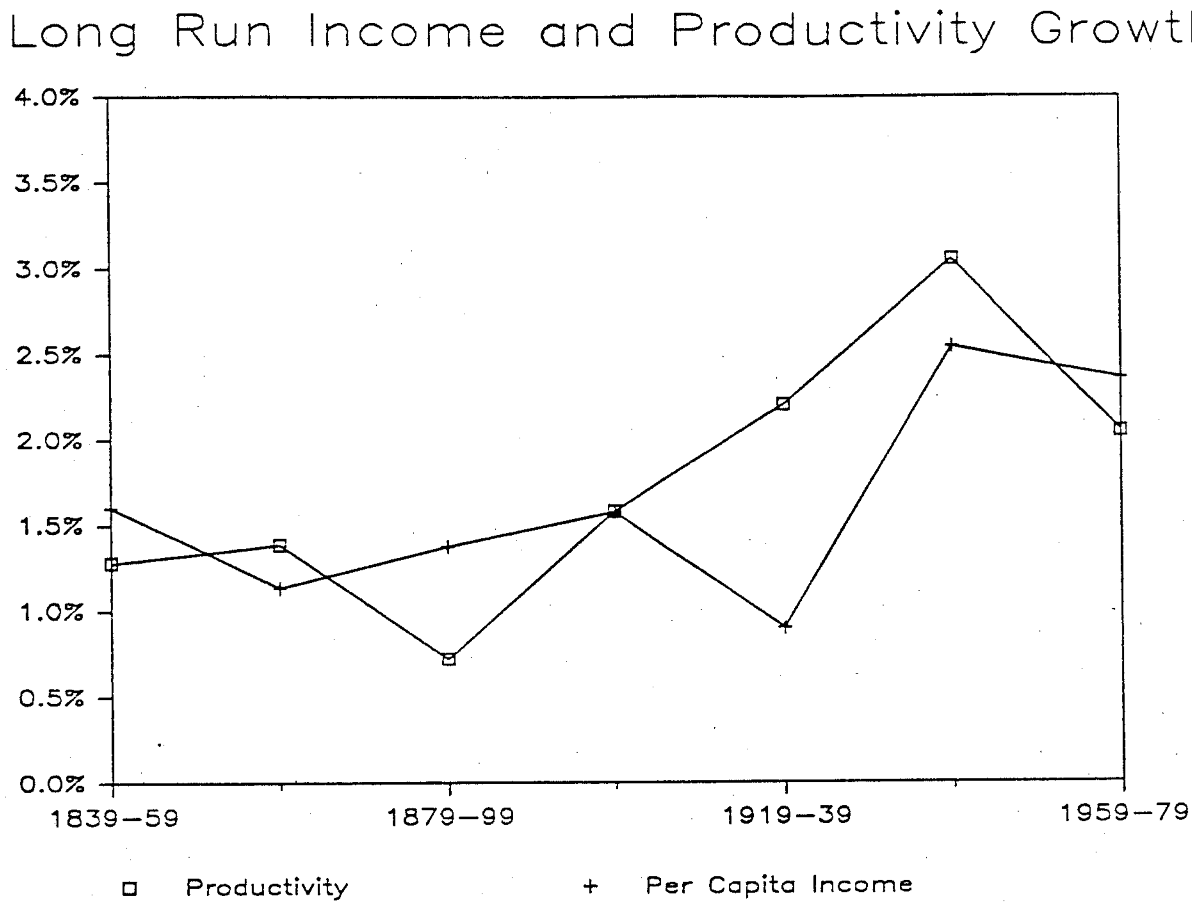

Long Run Income and Productivity Growtl
4.0%
3.5%
3.0%
2.5%
2.0%
1.5%
1.0%
0.5%
0.0%
1839–59
1879–99
1919–39
1959–79
Productivity
Per Capita Income

Fig. 2

Output and Capital Per Hour Worked

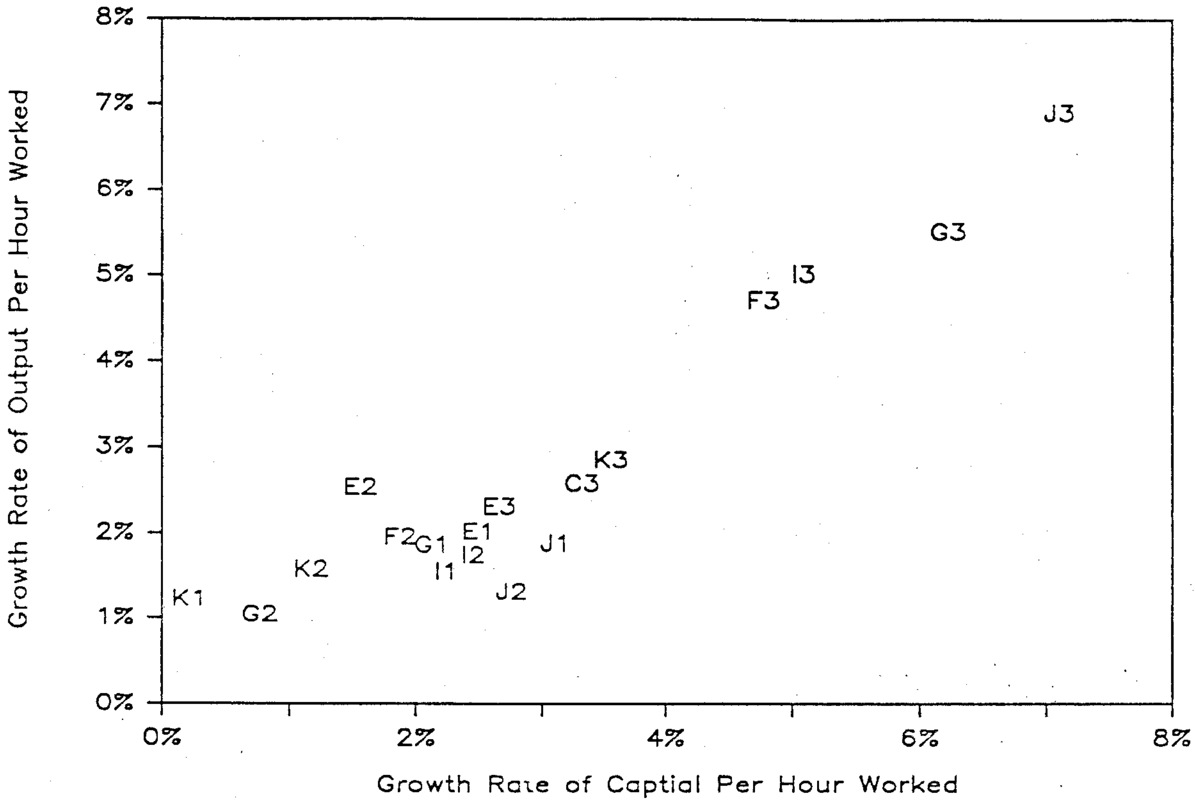


Fig. 3

Investment Share and GNP Growth

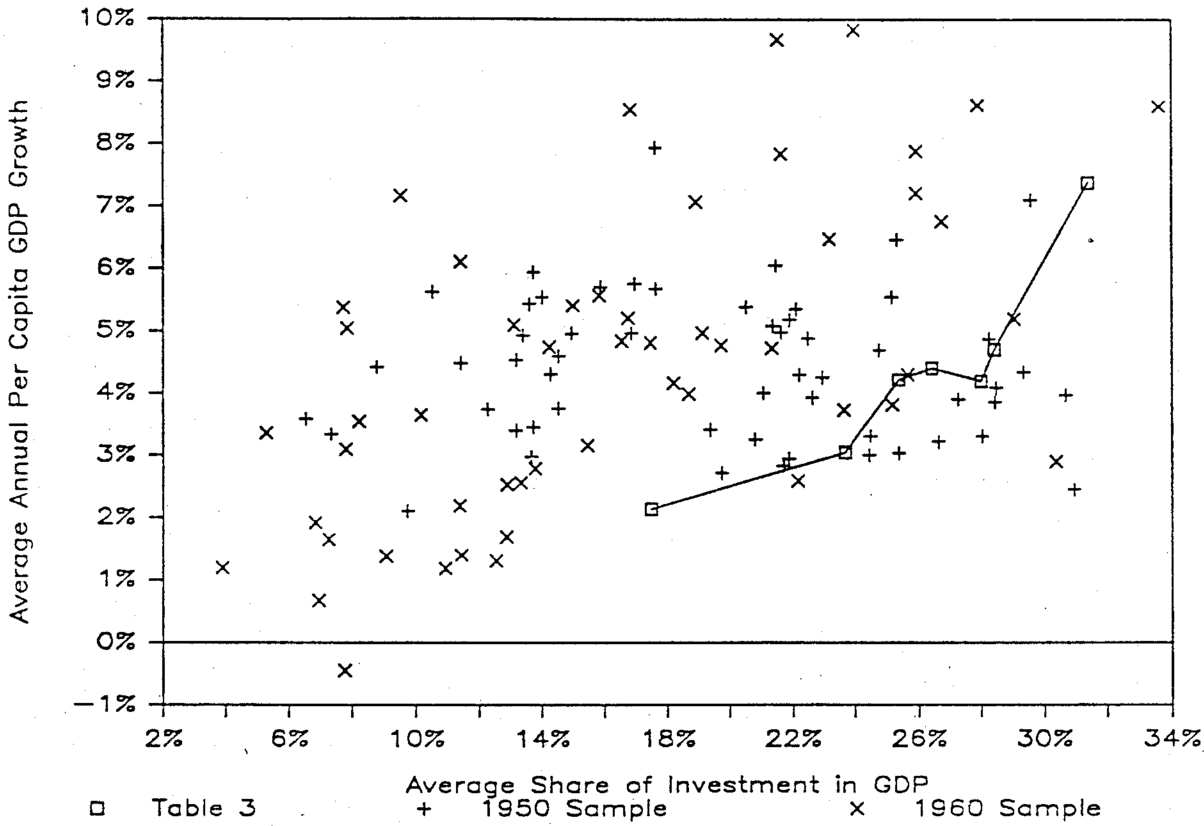


Fig. 4

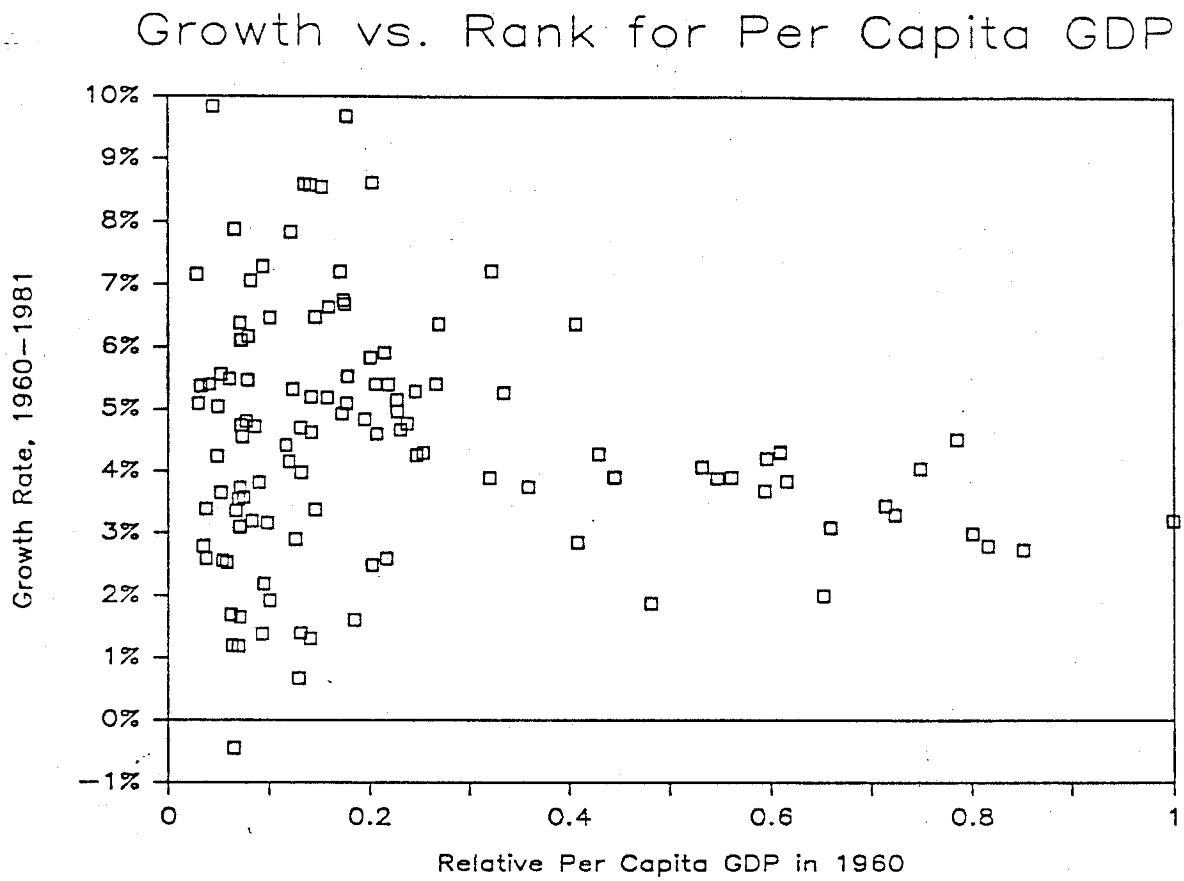
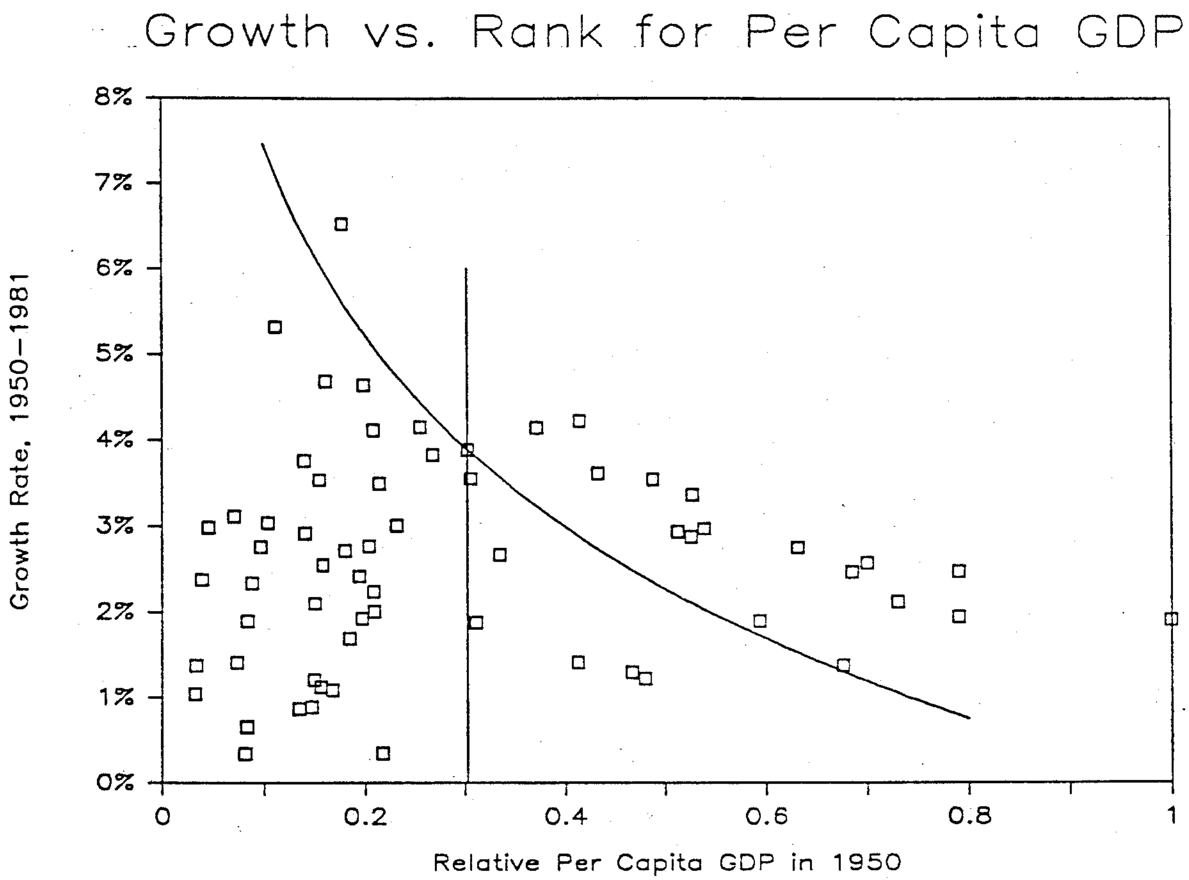

Growth vs. Rank for Per Capita GDP
10%
9%
8%
7%
6%
5%
4%
3%
2%
1%
0%
−1%
Growth Rate, 1960−1981
0
0.2
0.4
0.6
0.8
1
Relative Per Capita GDP in 1960

Fig. 5



Growth vs. Rank for Per Capita GDP
8%
7%
6%
5%
4%
3%
2%
1%
0%
Growth Rate, 1950–1981
0
0.2
0.4
0.6
0.8
1
Relative Per Capita GDP in 1950

Fig. 6

Growth of GDP and Exports

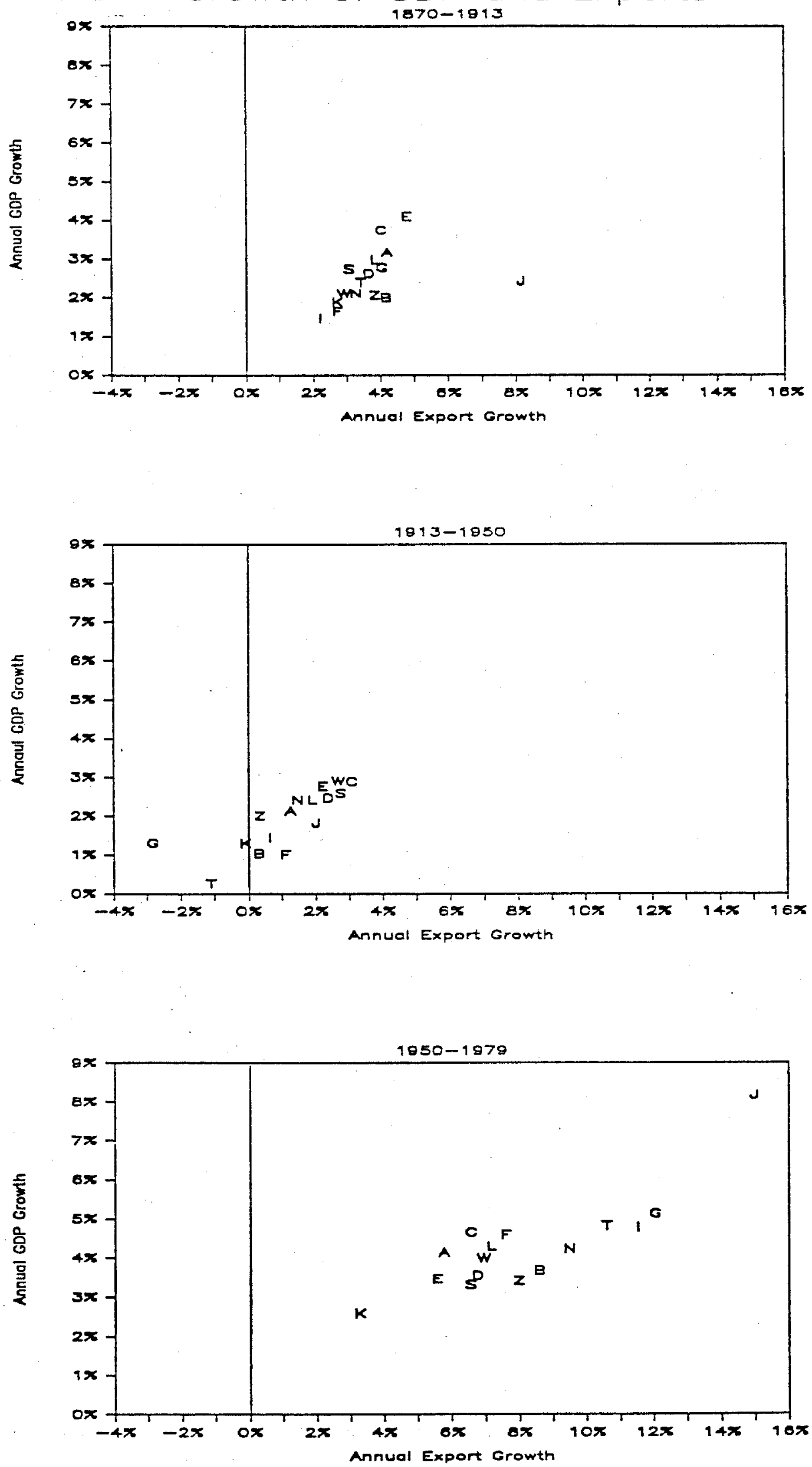


Fig. 7

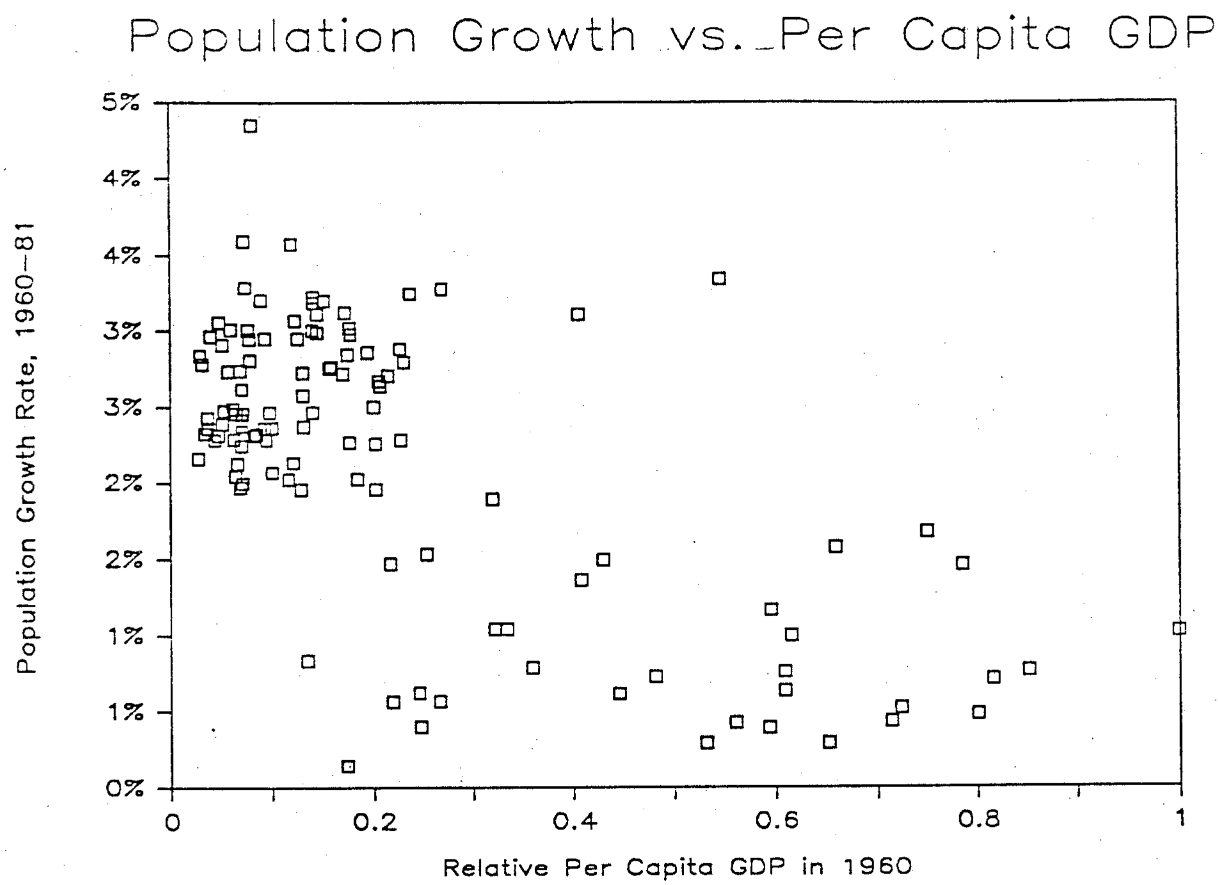

Population Growth vs. Per Capita GDP
5%
4%
4%
3%
3%
2%
2%
1%
1%
0%
Population Growth Rate, 1960–81
0
0.2
0.4
0.6
0.8
1
Relative Per Capita GDP in 1960

Fig. 8

Phase Plane for the Cass Model

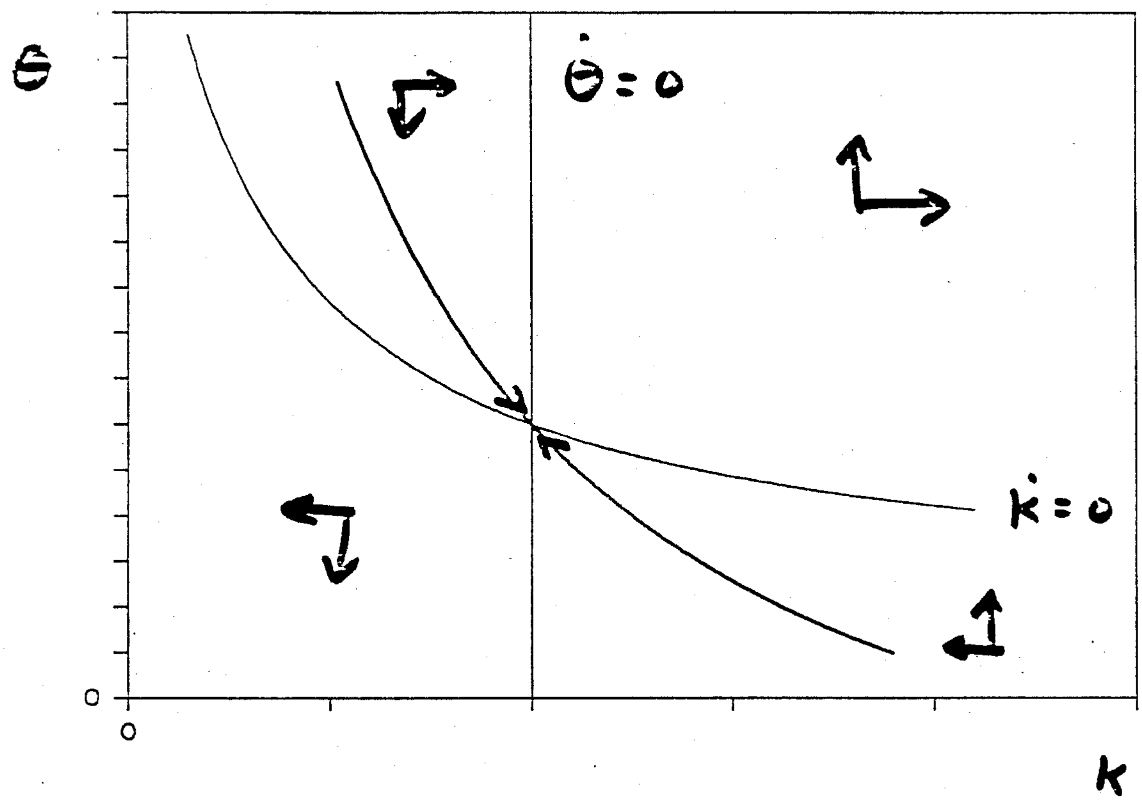


Fig. 9