Swan, T.W. (1956) "Economic Growth and Capital Accumulation." *Economic Record*, 32, November, pp.334-361.
Whelan, K. (2000) "Computers, Obsolescence, and Productivity." *Federal Reserve Board, Finance and Economics Discussion Series,* Paper No.2000-6, February.
Young, A. (1995) "The Tyranny of Numbers: Confronting the Statistical Reality of the East Asian Growth Experience." *Quarterly Journal of Economics*, 110, August, pp.641-680.

## References

Abramovits, M. (1956) "Resource and Output Trends in the United States since 1870." *American Economic Revuew*, 46, May, pp.5-23.
Aghion, P. and Howitt, P. (1992) "A Model of Growth through Creative Destruction." *Econometrica*, 60, March, pp.323-351.
Aghion, P. and Howitt, P. (1998) *Endogenous Growth Theory*, Cambridge, MA: MIT Press.
Asquith P. and Mullins D.W. (1986) "Equity Issues and Stock Price Dilution." *Journal of Financial Economics*, No.13, pp.296-320.
Baily, M.N. and Gordon, R.J. (1988) "The Productivity Slowdown, Measurement Issues, and the Explosion of Computer Power." *Brookings Papers on Economic Activity*, No.2, pp.347-420.
Barro, R.J. and Sala-i-Martin, X. (1998) *Economic Growth*, Cambridge, MA: MIT Press.
Denison, E.F. (1967) *Why Growth Rates Differ*, Washington: Brookings Institution.
Denison, E.F. (1985) *Trends in American Economic Growth*, Washington: Brookings Institution.
Greenwald B.C., Stiglitz J.E. and Weiss A. (1984) "Informational Imperfections in the Capital Markets and Macro-Economic Fluctuations." *American Economic Review*, vol.74, No.1, May, pp.194-9.
Griliches, Z. (1988) "Productivity Puzzles and R&D: Another Nonexplanation." *Journal of Economic Perspectives*, 2, Fall, pp.9-21.
Grossman, G.M. and Helpman, E. (1991a) *Innovation and Growth in the Global Economy*. Cambridge, MA; MIT Press.
Hsieh, C-T. (1998a) "What Explains the Industrial Revolution in East Asia? Evidence from Factor Markets." Unpublished paper, University of California, Berkeley, January.
Jensen M. (1986) "Agency Costs of Free Cash Flow, Corporate Finance and Takeovers." *American Economic Review*, No.76, May, pp.323-9.
Jones, C. (1998) *Introduction to Economic Growth*, New York: W.W. Norton.
Jones, C. (1999a) "*Sources of U.S. Economic Growth in a World of Ideas*." Un published paper, Stanford University, Septermber.
Jorgenson, D.W. (1988) "Productivity and Postwar U.S. Economic Growth." *Journal of Economic Perspective*, 2, Fall, pp.23-41.
Jorgenson, D.W. and Stiroh K.J. (2000) "Raising the Speed Limit: U.S. Economic Growth in the Information Age." Unpublished paper, Harvard University, May. *Brookings Papers on Economic Activity*, forthcoming.
Lucas, R.E., Jr. (1990) "Why Doesn't Capital Flow from Rich to Poor Countries?" *American Economic Review*, 80, May, pp.92-96.
Myers S. and Majluf N. (1984) "Corporate Financing and Investment Decisions When Firms Have Information that Investors do not Have." *Journal of Financial Economics*, No.13, pp.187-221.
Oliner, S.D. and Sichel, D.E. (2000) "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?" Federal Reserve Board, Finance and Economics Discussion Series Paper No.2000-20, March. *Journal of Economic Perspectives*, forthcoming.
Romer, C.D. (1990) "The Great Crash and the Onset of the Great Depression." *Quarterly Journal of Economics*, 105, August, pp.597-624.
Romer, C.D. (2001) *Advanced Macroeconomics* 2nd ed., New York, McGraw-Hill
Solow, R.M. (1956) "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics*, 70, February, pp.65-94.
Solow, R.M. (1957) "Technical Change and the Aggregate Production Function." *Review of Economics and Statistics*, 39, pp.312-320.
Stiglitz, J.E. (1994) "Explaining Growth: Competition and Finance" in Baldassarri, M., Paganett, L. and Phelps, E.S. (eds.) International Differences in Growth Rates, St. Martin's Press.

## Exercise 1

Robert Solow's famous 1987 quip that "you can see the computer age every where but in the productivity statistics" provoked a series of research. Read Crafts' article (2001) "The Solow Productivity Paradox in Historical Perspective" (London School of Economics, http://www.j-bradford-delong.net/article-of-the-month/prof/Newsolow.pdf) and discuss how we can evaluate the Solow's productivity paradox.

assuming that the probabilities of success each period are independent. We can now see that lowering the interest rate charged the first period may increase net social product:

$$dS/di = [dS/dp^*][dp^*/dw][dw/di] < 0$$

Financial "repression" may increase net national product.

Similarly, we can show that lowering the wage (even below the market clearing level) may result in increase net national product.

## 8.5 Concluding Remarks

Much of growth policy is predicated on the economists' basic competitive model. Governments are advised simply to let markets work, or as the expression goes, "get the prices right". Yet many observed aspects of the growth process seem inconsistent with the competitive model. Almost all recent theorizing about growth processes – whether based on externalities, learning by doing, increasing returns, or financial market imperfections – identifies significant departures from the standard competitive paradigm as being central. It has become commonplace for economists, at this juncture, to sound the caveat that the existence of market failures does not in itself provide a justification for government intervention. One must show that the government can, and is likely, to intervene in ways which are welfare enhancing. In many of the East Asian countries governments seem to have taken an active role; they have intervened, not in the manner envisaged by the now thoroughly discredited central planning paradigm, but in more subtle, if no less pervasive ways. They have helped not only create markets, they have used markets: they have helped make them work in ways which may well have enhanced the success of these countries. How they have done this, and the extent to which the lessons we can learn from their experiences are replicable in other countries, is a matter for future research.[22,23]

---

22 A major World Bank project examining these questions is presently underway.

23 Amsden-Euh provide an interesting description of the interventions of the Korean government in their financial markets, interventions which are remarkably similar to those which the theories we have described above might suggest. The Korean government did a great deal to encourage the development of equity markets, including putting limits on the debt equity ratios; they kept interest rates charged low, with evidently no significant adverse effects on savings.

$\hat{p}$ :

$$\frac{wd\hat{p}}{\hat{p}dw} = -\frac{w\left(1-\frac{1}{H}\right)}{w+\hat{p}(1-w)h/H^2} > 0$$

where:

$$h \equiv H' = \frac{f(p)}{1-F}\{-\hat{p}+H(\hat{p})\} > 0$$

Hence:

$$\frac{di}{dw} = -\frac{(1+r)}{H^2}h\frac{d\hat{p}}{dw} < 0$$

Net social expected returns from the lending activity to those with wealth *w, S,* is:

$$S = NK(\hat{p})R - N(1-F(\hat{p}))(1+r) \quad (49)$$

differentiating with respect to $\hat{p}$, we obtain:

$$\frac{\partial S}{\partial \hat{p}} = Nf(\hat{p})[(1+r)-R\hat{p}] \quad (50)$$

But from (48):

$$\hat{p}R-(1+r) = (1-w)((1+i)\hat{p}-(1+r)) = (1-w)(1+i)(\hat{p}-H) < 0$$

Thus, an increase in w increases the net social return from the lending activities.

Consider now a two period model in which no one begins with any wealth; but in which those who are successful the first period accumulate a wealth of:

$$w = R-(1+i)$$

The first period, all potential entrepreneurs borrow $(\hat{p}=0)$, while the second period,

$$p^* = \Phi(R-(1+i(0)),1+r(R-(1+i(0)))$$

normalizztion). Lenders can observe the wealth, $w$, of entrepreneurs. They require entrepreneurs to invest all of their wealth in the project, and they lend the difference, $1-w$. The interest rate charged is such as to yield an expected return on all of those borrowing equal to the safe rate of interest $r$. Let $F(p)$ be the distribution function of $p$, $N$ the number of entrepreneurs, and define:

$$K(p) = \int_{\hat{p}}^{1} p dF(p) \tag{42}$$

$$H(p) = K(p)/(1-F(p)) \tag{43}$$

the mean probability of success of all projects with success probabilities greater than or equal to $\hat{p}$. An individual with wealth $w$ has a choice of borrowing at an interest rate $i$, or investing his wealth in the safe asset. Thus, he will borrow so long as:

$$p[R-(1+i)(1-w)] \geq (1+r)w \tag{44}$$

Define $\hat{p}$ as the marginal person borrowing. Then:

$$\hat{p}[R-(1+i)(1-w)] = (1+r)w \tag{45}$$

$\hat{p}$ can be expressed as a function of $w$ and $1+i$:

$$\hat{p} = \Phi(w, 1+i) \equiv \frac{(1+r)w}{R-(1+i)(1-w)} \tag{46}$$

$i$ is set so as yield the same expected return as the safe asset:

$$H(\hat{p})(1+i) = (1+r) \tag{47}$$

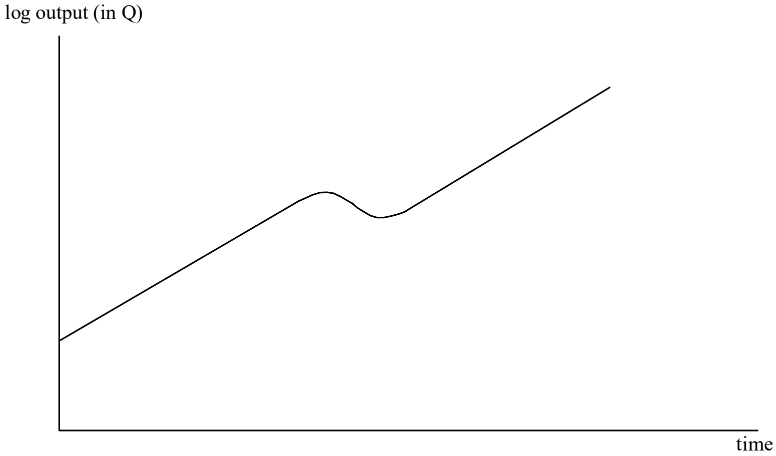
The above two equations can solved simultaneously for $i$ and $p$; e.g. substituting (47) into (45) we obtain:

$$\hat{p}R - \hat{p}(1-w)(1+r)/H(\hat{p}) = (1+r)w \tag{48}$$

We denote the solution for the rate of interest charged by $i(w)$.

It is apparent that an increase in $w$ lowers the nominal interest rate charged and increases on R&D expenditures, since it has an immediate positively will not be felt for some time. Greenwald, Salinger and Stiglitz ,1992, provide empirical evidence in support fo this contention, and Greenwald, Kohn and Stiglitz provide a theoretical model extending this analysis to the case of learning by doing.

**Figure 8 An Economic Downturn Has A Long Run Effect on Output**




Thus, in economies in which financial constraints are important, achieving macro-stability has distinct long-run benefits. Indeed, they are even greater than we have suggested, since one of the main sources of information imperfections concerns the ability of firms to withstand cyclical shocks. Reducing the magnitude of the cyclical shocks reduces, in this sense, the extent of information imperfections in the economy[20].

## A Simple Model

The complexity of the relationships that we have described cannot be fully captured in any single model[21]. The following simple model illustrates several of the themes we have emphasized. Assume that, as in Stiglitz and Weiss 1981, projects either are successful and yield a return of $R$, or are a failure, and earn a return of zero ($R$ may itself be a function of the wage rate). For simplicity, we assume all entrepreneurs have projects yielding the same $R$, but differ in the probability of success, $p$. Entrepreneurs know their probability of success, but lenders do not. We assume there are no functioning equity markets. Each project costs 1 (a

[20] This argument has to be qualified by the observation that in the presence of more stable macropolicies, firms may be induced to borrow more, thus exacerbating the effects of any economic downturn.

[21] For several models attempting to capture various aspects of the equity-growth relationship, see Greenwald-Kohn-Stiglitz.

infant industry arguments is that if, in the long run, a firm will gain a comparative advantage by producing, then it sould pay it to borrow, to sell below its current marginal costs (with learning, production should be related to the long run marginal cost; that is, the marginal return to producing more is not just today's increased profits, but the decreased costs of production in the future). If the interest rate is zero, then the relevant marginal cost is simply the asymptotic value (Spence). In the absence of externalities, it is socially profitable to enter an industry if and only if it is privately profitable to do so. But this analysis assumes that the firm can have easy access to capital. If it cannot, there may be high social returns to "investing" in learning, yet private firms simply cannot afford to do so.

**Collateralization**

This is particularly true because this form of investment cannot be collateralized; the "investment" is not like an investment in a building or a machine. Because the costs of information imperfections are greater for investments which cannot be collateralized, the market will have a "bias" towards collateralizable investments ,and against investments in "learning by doing", or R&D, which cannot be collateralized.

**Macro-Stability**

Theories emphasizing the importance of financial constraints also emphasize the importance of macro-economic stability for growth. In the older, neoclassical models, technical change was exogenous, so that any short term disturbance that might move the economy below its production possibilities curve would have only temporary effects; it might slow down the process of capital accumulation, and thus delay slightly the approach to the steady states, but that is all.

In models with endogenously determined expenditures on R&D, but with no financial constraints, cyclical fluctuations should have little effect on the pace of R&D is driven by long run considerations – say the savings from lower costs of production – and neither long run real wages, interest rates, or output is, in this perspective, likely to be affected much by a short term downturn in the economy, and therefore neither are incentives to invest in R&D or learning.

By contrast, models with financial constraints argue that short term macro-fluctuaitons have marked effects on R&D and learning, so that the growth path of the economy may be *permanently* lower, as illustrated in Figure 8. The reason for this is that stated earlier: with lower "net worth" firms are less willing or able to make these investments; with major downturns, in spite of the higher adjustment costs often associated with R&D ,firms cut back

These often serve an intermediary role; they receive capital (often from large institutional investors) which they reinvest in new ventures. They are specialized in screening and monitoring.

But intermediation is more pervasive: the nexus of production relations which characterized modern industrial economies gives rise to complicated patterns of information flows. A firm knows much about its suppliers and customers; if they are slow in delivering products, if product quality is variable, if customers are slow in making payments, questions about managerial or financial strengths of the firms are raised. Firms are, thus, often in a better position to monitor the firms with whom they have relationships than are banks, and it is accordingly not surprising that the nexus of production relationships is associated with a nexus of financial relationships, with firms supplying credit to each other. A large firm may simultaneously borrow funds from its bank and lend both to its suppliers and customers. It is acting as a financial intermediary.

Large conglomerates also can facilitate the flow of capital. Just as earlier literature stressed the importance of an internal labor market, there may be an internal capital market. A major lesson to emerge from the US experience with conglomerates is that these internal capital markets may not work very well; or at least that the diseconomies of scope may outweigh any gains from the improved reallocation of capital. This may not be surprising: the information flows among the disparate parts of conglomerates engage in relatively unrelated economic activities may not have been much, or any, better than those available to an unrelated bank.

What then accounts for the seeming success of the Korean conglomerates, or the analogous institutions in Japan? Answering this question would take us beyond the scope of this chapter, but a suggestion is that "better" government structures, which prevented or limited the abuse of managerial discretion, as occurred in the United States[19]: the closely held nature of the Korean firms, and the role of the main bank (which often held equity interests in the firms to which they lent) in Japan.

**Infant Industry Arguments**

We have seen that financial market imperfections (arising naturally out of the fact that information is imperfect and costly) mean that there is a discrepancy between private and social returns, a discrepancy which may differ across industries. While this in itself would provide a rationale for an industrial policy, it is worth noting that financial market constraints may, in particular, provide a rationale for "infant industry"arguments. The classical criticisms of

[19] For theoretical analyses explaining the rationale for these managerial "abuses" see Shleifer-Vishny or Edlin-Stiglitz.

suggested, has a positive effect on savings).

If they cannot pass on these lowered interest rate charges to their depositors (say because of competition from an unregulated financial intermediary), then the strength of financial institutions will be weakened, and this will have adverse effects on their ability and willingness to lend.

This analysis also suggests that government actions restricting competition in te banking sector may have more ambiguous effects than has previously been thought. Reduced competition may not lead to higher interest rates charged if interest rates charged are determined in a manner described by standard credit rationing models. Reduced competition may lead to lower interest rates paid to depositors, but the net effect on savings may be small; and the increased net worth of the financial institutions may result in increased lending activities (and may lead banks to be willing to lend to higher risk-higher expected returns projects).

Restrictions on competition always represent (at least) a two edged sward: isolated from competition, bank managers may become slack; rather than the increased profits being used to facilitate increased lending, the funds may be used to increase managerial perks.

**Regional Lending**

Earlier, we noted evidence of agglomeration economies, and posed the problem of identifying the sources and extent of these economies. Banks, we have argued, are essentially in the information business. Much of the information which they acquire is very particular, very localized, by-products of other activities. A lending officer can get often more reliable information about how a store is doing by randomly checking on the store, looking at inventory on the shelf, the number of customers, etc. than by looking at financial accounts. Localized hearsay information often yields important clues as to what is going on.

Similarly, venture capital firms tend to specialize not only with respect to the industries to which they provide finance, but also with respect to the locale. They want to be able, at low costs, to make on-site inspections.

It is thus not surprising that financial centers are often linked closely with production centers.

**Industrial Organization**

While we have emphasized the role of banks in financial intermediation, they are not the only institutions involved in that activity. We have already mentioned venture capital firms.

## Role of Government

This perspective has quite different implications with respect to the growth process and the consequences of a variety of forms of government intervention. In particular, policies which increase the financial strength of firms or of financial institutions may reap large dividends. By contrast, in neoclassical growth theory, the corporate veil is easily pierced: only real factors matter. Indeed, in the simplistic model of much of the new growth theory, there is a representative agent, who simply maximizes his intertemporal utility. The distribution of wealth among individuals, or between households and firms, is of absolutely no consequence.

Enhancing firm profits increases the potential for retained earnings, thus increasing firm's equity base, and their ability and willingness to invest and take risks. Since the social returns to investment may well exceed the private returns, government, by imposing high taxes on distributed profits, can encourage firms to retain a larger fraction of their profits (in rapidly growing economies ,the returns to retained earnings noted earlier are sufficiently high that no further encouragement from the government may be needed).

Profits will be high if wages remain low. Here we see an alternative mechanism for why surplus labor facilitates growth. Traditional theory has focused on the fact htat it enables growth to proceed without large increases in wages. Increases in wages, it was thought, slowed down the growth process presumably because savings rates out of wages were smaller than savings rates out of profits. In a sense, our analysis provides a theoretical rationale for these differences in savings rates, a rationale related to differences in returns to savings (since household saving is largely mediated through credit institutions, and the interest rates they pay depositors are typically far lower than the return to capital). Our analysis emphasizes that it is not just the amount of savings, but the form. If it were just the amount of savings, the deficiency could be made up through government borrowing abroad.

## Government Regulations

Lowering interest rates charged to borrowers has ambiguous effects. On the one hand, it increases the profitability of firms; it induces firms to borrow more; it increases their retained earnings; and through these channels has, in the long run, a possibly large multiplicative effect.

The effects on financial institutions depends on whether they can "pass on" the reduced interest rates in the form of reduced deposit rates. If they can, then there may a slight deleterious effect on household savings. But this is likely to be far outweighed by the positive effects from increased corporate savings (and indeed, the lower interest rates result in lowered default probabilities, again enhancing the financial stability of financial institutions, which, we

**Implications of Imperfect Capital markets**

These limitations on financial markets in turn have important consequences:

a) firms may be limited by their retained earnings in the amount which they can invest; the marginal return to firms' savings may be very high. More generally, the marginal returns to capital indifferent firms may differ[16];

b) with even small degrees of increasing returns (such as associated with learning by doing) markets will be dominated by a single firm in the absence of capital market imperfections. With capital market imperfections, this will not be true: financial market imperfections provide an alternative to imperfect competition as a resolution of the problem of increasing returns[17];

c) when firms have access to credit markets, retained earnings have a further effect in providing the collateral which gains them access to the credit market; on the other side, the more "equity (retained earnings) they have, the more willing they are to undertake the risks associated with borrowing, i.e. the lower the probability of default at any level of economic activity, and the lower the *marginal* probability; thus, the "marginal cost" of producing more and of investing more is reduced, leading to higher levels of these activities. Moreover, the greater their "wealth", the greater their willingness to engage in more risky activities, such as associated with investments in R&D[18]. Accordingly, there may be higher returns to savings; we can think of this equity as "high powered" capital;

d) the ability and willingness of bank and other financial institutions to lend depends too on their financial position, including their net worth; banks too are equity constrained. Their willingness and ability to borrow funds (recruit deposits), which they then lend out is affected by their net worth (a bank can simply be thought of as a firm whose activity is to make loans). A reduction in their net worth thus reduces the amount they are willing to lend.

---

16 For example, large firms may obtain a higher return because there are increasing returns associated with obtaining information; it pays them to spend more to screen among alternative projects. Moreover, larger firms may be able to diversify risks better, and thus be willing to undertake higher risk-higher return projects.

17 Actually, the two theories are, in many cases, complementary, with the financial constraints playing a more important role in some markets (particularly when there are many small firms, as in the computer industry) and imperfect competition in others.

18 In Greenwald-Stiglitz, we describe the portfolio theory of the firm, in which we show how the various actions of a firm can be thought of as a portfolio, the mix of which affects the probability distribution of final values of the firm.

**Table 1a. Gross Sources of Finance 1970-89**

(weighted average, undepreciated, revalued)

| | France | Germany | Japan | UK | US |
|---|---|---|---|---|---|
| Internal | 42.2 | 62.4 | 42.2 | 61.1 | 62.7 |
| Bank finance | 37.7 | 18.0 | 34.1 | 23.5 | 14.7 |
| Bonds | 2.7 | 0.9 | 3.3 | 1.4 | 12.8 |
| Equity | 14.1 | 2.3 | 3.1 | 7.4 | -4.9 |
| Trade credit | 2.9 | 1.8 | 14.9 | 1.9 | 8.8 |
| Capital transfers | 3.4 | 6.6 | --- | 2.3 | --- |
| Other | 0.3 | 8.0 | 2.4 | 2.4 | 5.9 |
| Notes | 1970-85 | 1970-89 | 1970-87 | 1970-89 | 1970-89 |

Source: Unpublished flow of funds figures from the CEPR: *International Study of the Financing of Industry*. Data courtesy of Tim Jenkinson and Colin Mayer.

**Table 1b. Net Sources of Finance 1970-89**

(weighted average, undepreciated, revalued)

| | France | Germany | Japan | UK | US |
|---|---|---|---|---|---|
| Internal | 66.3 | 80.6 | 71.7 | 98.0 | 91.3 |
| Bank finance | 51.5 | 11.0 | 28.0 | 19.8 | 16.6 |
| Bonds | 0.7 | 0.6 | 4.0 | 2.0 | 17.1 |
| Equity | -0.4 | 0.9 | 2.7 | -0.8 | -8.8 |
| Trade credit | -0.7 | -1.9 | -7.8 | -1.6 | -3.7 |
| Capital transfers | 2.6 | 8.5 | --- | 2.1 | --- |
| Other | -14.9 | 1.5 | 1.3 | -4.1 | -3.8 |
| Statistical adj. | -5.1 | 0.0 | 0.1 | -8.2 | -8.7 |
| Notes | 1970-85 | 1970-89 | 1970-87 | 1970-89 | 1970-89 |

Source: Unpublished flow of funds figures from the CEPR: *International Study of the Financing of Industry*. Data courtesy of Tim Jenkinson and Colin Mayer.

**Table 2. Sources of Funds by the Corporate Sector Korea**

(in %)

| | 1963-65 | 66-71 | 72-76 | 77-81 | 82 | 84 | 86 | 87 |
|---|---|---|---|---|---|---|---|---|
| Internal funds | 47.7 | 25.4 | 32.9 | 23.3 | 27.0 | 33.3 | 39.9 | --- |
| External funds | 52.3 | 74.6 | 67.1 | 76.7 | 73.0 | 66.7 | 60.1 | --- |
| Total | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| External funds | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| Borrowing from Monetary institutions | 48.4 | 41.8 | 51.1 | 53.7 | 55.4 | 60.5 | 44.8 | 46.4 |
| Banks | 33.5 | 32.8 | 34.3 | 32.6 | 30.0 | 21.6 | 33.6 | 21.2 |
| Non-banks | 15.0 | 9.0 | 16.8 | 21.1 | 25.5 | 38.9 | 11.2 | 25.2 |
| Securities | | | | | | | | |
| (direct finance) | 27.6 | 14.3 | 21.8 | 24.8 | 31.8 | 32.1 | 24.7 | 40.5 |
| Debts | 1.2 | 0.7 | 2.5 | 4.2 | 18.3 | 11.5 | 9.6 | 8.1 |
| Stocks | 21.4 | 11.8 | 18.1 | 14.4 | 10.4 | 20.6* | 15.1* | 32.4* |
| Capital paid in | 5.0 | 2.7 | 1.3 | 1.9 | 3.2 | --- | --- | --- |
| Corporate bills | --- | --- | 1.8 | 5.5 | 7.7 | --- | 100.0 | -2.3 |
| Government and curb market loans | 8.5 | 7.8 | -0.3 | 0.8 | 1.0 | 3.2** | 5.3** | 17.0** |
| Borrowing from abroad | 15.4 | 36.2 | 26.6 | 15.2 | 4.1 | -5.1 | 15.2 | -1.6 |

* Stocks and capital paid in.

** Others included.

Source: The Bank of Korea: Annual Report, various issues, Amsden-Euh[1].

financial institutions. Note that excessive competition in the banking industry may, without government regulation and support, actually increase the financial fragility of particular financial institutions, and thus have a deleterious effect on savings, at least savings intermediated through financial institutions. By the same token, reduced interest rates may increase the financial viability of these institutions, and this effect may more than offset the direct effect from the lower return[14].

The fact that so many countries have obtained low returns on their investment that incremental capital output ratios have been so high suggests that at least as important as the magnitude of the savings is the efficiency with which it is allocated. We now recognize that central planning bureaus simply did not get at the essence of the allocation problem: more important than determining the sector of the economy is determining the precise project and managing that project; and given that, the choice of who is to manage a project is crucial.

While one of the central functions of financial institutions is precisely that looking at particular projects and firms[15], and selecting which get loans they do this imperfectly: because information is costly, screening among projects is never done with perfect accuracy.

That is why they have to rely on "indirect control mechanisms." They know that the mix of applicants that they get, and the extent of risk taking is affected by the interest rates which they charge, and other non-price terms of the loan contract. Thus, expected returns may be reduced even when the firm increases the interest rate charged. As is by now well known, this may lead to credit rationing.

Credit rationing, in turn, has one consequence which is important for our purposes: interest rates do not reflect the marginal productivity of capital. The social return to savings (in forms which get lent in credit markets) exceeds the private return. Given that credit is rationed, it is natural to look to other forms of capital, in particular, to equity markets. But equity markets are, if anything, even more imperfect than credit markets. While equity has marked advantages over credit (from the borrowers perspective) in that it entails risk sharing and avoids the threat of bankruptcy which the fixed obligations associated with credit entail, equity remains a relatively unimportant source of funds, even in the more developed countries (Tables 1 and 2). Recent work has provided an explanation of this: on average, when firms issue shares, there is a marked decline in the price of each share (See, e.g. Asquith and Mullins (1986)). And we have good theoretical reasons for why this should be so, based on models of adverse selection (Greenwald, Stiglitz, and Weiss (1984), Myers and Majluf (1984) and agency (Jensen (1986)).

---

14 Obviously, maintaining a stable macro-economic policy is among the most important policies which governments can pursue to enhance the stability of the financial system.

15 The choice of a firm can be thought of as a decision about who should "manage" the funds.

### Imperfect Capital Markets

With learning by doing, in many cases it will be desirable for firms to produce a sufficiently large amount in earlier periods that price will be less than average costs (Dasgupta and Stiglitz (1980)). The losses can only be financed by borrowing or government grants. (The arguments began with the presumption that lump sum subsidies were precluded)[12].

But capital markets are notoriously imperfect. Work in the economics of information over the past fifteen years has explained why both equity and debt markets are imperfect: there is both credit and equity rationing[13].

But if firms cannot borrow, they will produce too little. And hence the level of technological progress will be less than desirable. It is the capital market imperfection which leads to too low a rate of technological progress. Elsewhere, Stiglitz develops a model in which this second intuition is explored, and confirmed. In doing so, we identify a new category of imperfect information – imperfect capital market failures, arising when lenders cannot commit themselves to borrow from the same lender in succeeding periods.

## 8.4 Financial Markets

The last two decades has seen large advances in our understanding the role of, and limitations in, financial markets.

### Role of Financial Markets

Financial markets, broadly defined to include an array of financial institutions, perform a number of functions, which we can briefly summarize as *a*) facilitating the accumulation of capital; *b*) reallocating capital; and *c*) monitoring the usage of capital. Much of the development literature has focused on *a*), and on the basis of this, governments which have restricted interest rates have been criticized for "financial repression". This, it is contended, has reduced savings rate. Yet, econometric studies have shown little evidence of a large interest elasticity of savings, and savings rates have been very high even in some countries with relatively low interest rates.

Equally, or perhaps more important than the interest rate *promised* is the security of the assets, avoiding the possibility of large negative returns, which result from bankruptcies of

[12] The motivation for this assumption was that if the subsidy was not related to output, every individual could claim to be a firm, and collect the lump sum subsidy; while, of course, if the subsidy were related to output, it is not lump sum, and distorts the level of production.

[13] Not only are these markets imperfect, but the market equilibrium is constrained Pareto inefficient.

Dixit-Stiglitz (1977) extended the Lerner result to the case where the number of products produced was endogenous and where, because of fixed costs, there could not be a perfectly competitive equilibrium without government intervention; they showed that the market solution was constrained Pareto efficient, where the government was restrained not to provide lump sum subsidies to firms.

We develop a set of models in which we ask, "are markets which are imperfectly competitive and in which there is learning by doing (as a concrete form of increasing returns) efficient, in some sense?".

The prevailing wisdom, that there will be too little production, is based on the similarity between R&D and learning by doing. Arrow (1962) argued that with imperfect competition that would be too little expenditure on R&D. First, he argued that competitive firms would spend too little on R&D, since competitive firms failed to capture the consumer surplus associated with improved technology on new products. Secondly, he argued that monopolists would produce less than competitors, and since the incentive to innovate increases with the level of production, the incentive to innovate would be lower with monopoly than under competition.

With learning by doing, at the margin, one can view the decision to produce as partially an investment in improved technology. Thus, if incentives to make conventional investments in improved technology are too low under imperfect competition, so too are those investments associated with production. In this view, with learning by doing there is underproduction with imperfect competition, for a new reason, beyond the traditional one of marginal revenue being less than price.

But the arguments given above suggest that Arrow's reasoning is too partial equilibrium in nature. In a general equilibrium model with all firms facing the same constant elasticity demand curve, production with monopoly is not less than under competition. Moreover, R&D is essentially like a fixed cost, and the Dixit-Stiglitz analysis suggests that if the government were constrained not to provide lump sum subsidies to firms, the market equilibrium might still be (constrained) Pareto efficient. In an earlier study, Stiglitz showed that these conjectures are in fact correct; that is, market equilibria with imperfect competition may entail a constrained Pareto efficient level of expenditures on R&D.

Thus, below we extend these results to an economy with learning by doing, establishing an analogous constrained Pareto efficiency result.

The results on constrained Pareto efficiency require, however, two rather stringent assumptions. The first is that credit markets are perfect, the second is that the labor supply elasticity is zero. If either assumption is violated, there is scope for government intervention.

$$p_i^m = w\left[\frac{1}{1-\frac{1}{\varphi}}\right] \bigg/ f_i^{'}[L_i^m] \tag{38}$$

marginal revenues equal marginal costs, where:
$p_i^m$ = price of ith good (produced by ith firm) in monopolistic equilibrium;
$L_i^m$ = employment in ith firm;
$\varphi$ = elasticity of demand;
$w$ = wage;
$f_i$ = the production function of the its firm.

Market equilibrium is defined by the solution to (38) and the market clearing equation

$$\sum L_i = L^* \tag{39}$$

where $L^*$ = total labor supply.

By constrast, the competitive market equilibrium is described by the solution to (39) and the equations:

$$p_i^c = w / f_i^{'}[L_i^c] \tag{40}$$

where $p_i^c$ = price of ith good in competitive equilibrium.

It is clear that if $L^*$ does not depend on real wages (and the composition of demand does not depend on the distribution of income), and the number of firms is fixed, the solutions to these sets of equations involve:

$$L_i^c = L_i^m \tag{41a}$$

and:

$$w / p^m = \left(1 - \frac{1}{\varphi}\right) w / p^c \tag{41b}$$

i.e. real wages are reduced as a result of imperfect competition, but the levels of output of each sector are unchanged.

There are two difficulties with this line of argument. First, it is not just population that has been treading up. The basic fact emphasized by Jones is that R&D's share and rates of investment in physical and human capital have been rising as well. Thus the failure of growth to rise is puzzling for these second-generation new growth models as well. Second, as Jones (1999b) point out, the parameter restrictions needed in these models to eliminate scale effects on growth are strong and appear arbitrary.

With decreasing returns, the lack of a trend in growth is not puzzling. In this case, a rise in, say, the saving rate or R&D's share leads to a temporary period of above-normal growth. As a result, repeated rises in these variables lead not to increasing growth, but to an extended period of above-normal growth. This suggests that despite the relative steadiness of growth, one should not think of the United States and other major economies as being on conventional balanced growth paths (Jones, 1999a).

Saving rates and R&D's share cannot continue rising indefinitely (though in the case of the R&D share, the current share is sufficiently low that it can continue to rise at a rapid rate for a substantial period). Thus one corollary of this analysis is that in the absence of countervailing forces, growth must slow at some point. Moreover, the calculations in Jones (1999a) suggest that the slowdown would be considerable.

## 8.3 Competition and Finance[11]

### Imperfect Competition

Though imperfect competition normally results in economic inefficiency, it does not necessarily result in too slow a rate of technical progress. Indeed, a central result of standard monopoly theory is that such firms are productively efficient. The monopolist can be viewed as producing output in different periods. Expenditures on R&D can be viewed as purely "technological" the firm decides how much to invest in capital goods, how much to spend on current inputs, and how much to spend on R&D; *given* the level of output, it makes all of these decisions efficiently.

Some years ago, Lerner argued that the levels of output produced in a monopolistic equilibrium by each firm would be optimal, provided only that there were no monopolists of intermediate goods, provided the elasticity of demand facing each monopolist was the same and that the elasticity of labor supply was zero. The argument was simple. Assume that there is a single input, labor; firms will hire labor to the point where:

[11] This section draws from Stiglitz (1994, pp.196-222)

Since all physical capital is used to produce goods, goods production is

$$Y(t) = K(t)^{\alpha} \left[(1-a_L)LA(t)\right]^{1-\alpha} \tag{37}$$

Our usual assumption of a constant saving rate $\left(\dot{K}(t) = sY(t)\right)$ completes the model. This is the case we have been considering with $\beta = 0, \theta = 1$ and $\gamma = 1$. To see the implications of this version of the model, note that (36) implies that $A$ grows steadily at rate $Ba_L L$. This means the model is identical to the Solow model with $n = \delta = 0$ and with the rate of technological progress equal to $Ba_L L$. Thus (since there is no population growth), the growth rates of output and capital on the balanced growth path are $Ba_L L$. This model provides an example of a situation where long-run growth is endogenous (and depends on parameters other than population growth), but is not affected by the saving rate.

**Scale Effects and Growth**

One important motivation for work on new growth theory is a desire to understand variations in long-run growth. As a result, early new growth models focused on constant or increasing returns to produced factors, where changes in saving rates and resources devoted to R&D permanently change growth. Jones (1995) points out an important problem with these models, however. Over the postwar period, the forces that the models suggest affect long-run growth have all been trending upward. Population has been rising steadily, saving rates have increased, the fraction of resources devoted to human-capital accumulation has risen considerably, and the fraction of resources devoted to R&D appears to have increased sharply. Thus new growth models with constant or increasing returns imply that growth should have increased considerably. But in fact, growth shows no discernible trend.

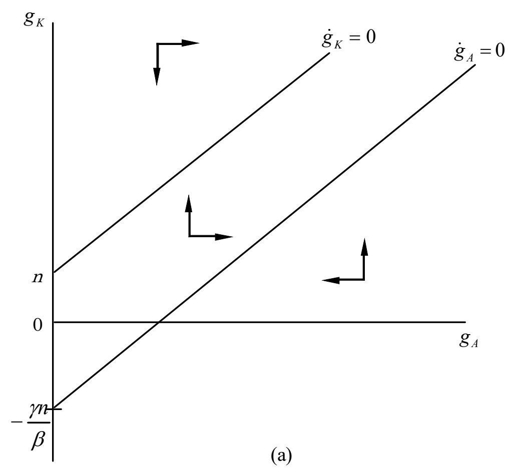
The simplest interpretation of Jones's results is that there are decreasing returns to produced factors; this is the interpretation proposed by Jones. Several recent papers suggest another possibility, however. They continue to assume constant or increasing returns to produced factors, but add a channel through which the overall expansion of the economy does not lead to faster growth. Specifically, they assume that it is the amount of R&D activity per sector that determines growth, and that the number of sectors grows with the economy. As a result, growth is steady despite the fact that population is rising. But because of the returns to produced factors, increases in the fraction of resources devoted to R&D permanently raise growth. Thus the models maintain the ability of early new growth models to potentially explain variations in long-run growth, but do not imply that worldwide population growth leads to ever-increasing growth.

to a balanced growth path. As in the case of $\theta = 1$ and $n$=0 in the model without capital, the phase diagram does not tell us what balanced growth path the economy converges to. One can show, however, that the economy has a unique balanced growth path, and that the economy's growth rate on that path is a complicated function of the parameters. Increases in the saving rate and in the size of the population increase this long-run growth rate; the intuition is essentially the same as the intuition for why increases in $a_L$ and $L$ increase long-run growth when there is no capital. And, as in Case 2, increases in $a_L$ and $a_K$ have ambiguous effects on long-run growth. Unfortunately, the derivation of the long-run growth rate is tedious and not particularly insightful. Thus we will not work through the details.

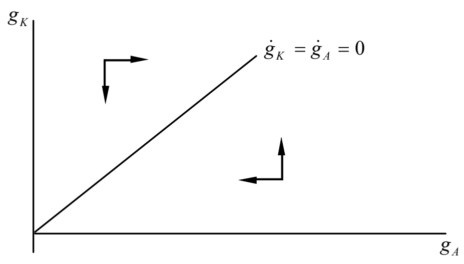
A Specific example of a model of knowledge accumulation and growth whose macroeconomic side fits into this framework is P. Romer's model of "endogenous technological change" (Romer, 1990). As here, population growth is zero, and there are constant returns to scale to the produced inputs in both sectors. In addition, R&D uses labor and the existing stock of knowledge, but not physical capital. Thus in our notation, the production function for new knowledge is

$$\dot{A}(t) = Ba_L LA(t) . \tag{36}$$

**Figure 6 The Dynamics of the Growth Rates of Capital and Knowledge When $\beta + \theta = 1$**




(a)

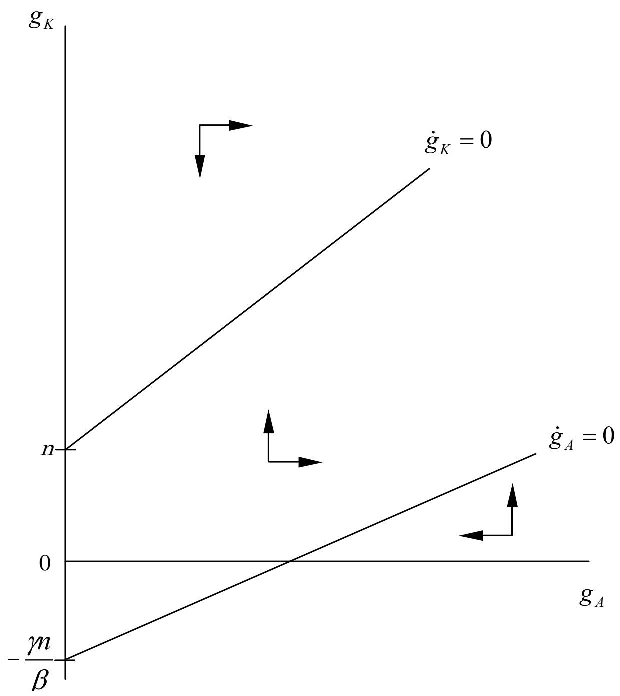



(b)

**Case 2:** $\beta + \theta > 1$

In this case, the loci where $g_A$ and $g_K$ are constant diverge, as shown in Figure 11. As the phase diagram shows, regardless of where the economy starts, it eventually enters the region between the two loci. Once this occurs, the growth rates of both *A* and *K*, and hence the growth rate of output, increase continually. One can show that increases in *s* and *n* cause output per worker to rise above its previous trajectory by an ever-increasing amount. The effects of changes in $a_L$ and $a_K$ are more complicated, however, since they involve shifts of resources between the two sectors. Thus this case is analogous to the case when $\theta$ exceeds 1 in the simple model.

**Figure 5 The Dynamics of the Growth Rates of Capital and Knowledge When** $\beta + \theta > 1$

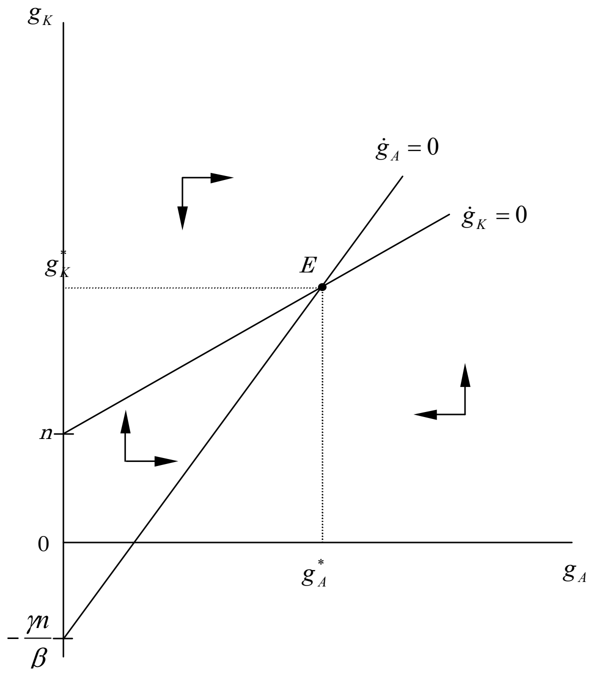



**Case 3:** $\beta + \theta = 1$

The final possibility is that $\beta + \theta$ equals 1. In this case, $(1-\theta)/\beta$ equals 1, and thus the $\dot{g}_A = 0$ and $\dot{g}_K = 0$ loci have the same slope. If *n* is positive, the $\dot{g}_K = 0$ line lies above the $\dot{g}_A = 0$ line, and the dynamics of the economy are similar to those when $\beta + \theta > 1$; this case is shown in Panel (a) of Figure 12.

If *n* is 0, on the other hand, the two loci lie directly on top of each other, as shown in Panel (b) of the figure. The figure shows that, regardless of where the economy begins, it converges

**Figure 4 The Dynamics of the Growth Rates of Capital and Knowledge When $\beta + \theta < 1$**




Rewriting (32) as $g_K^* = g_A^* + n$ and substituting into (33) yields

$$\beta g_A^* + (\beta + \gamma)n + (\theta - 1)g_A^* = 0, \tag{34}$$

or

$$g_A^* = \frac{\beta + \gamma}{1 - (\theta + \beta)} n. \tag{35}$$

From above, $g_K^*$ is simply $g_A^* + n$. Equation (23) then implies that when $A$ and $K$ are growing at these rates, output is growing at rate $g_K^*$. Output per worker is therefore growing at rate $g_A^*$.

This case is similar to the case when $\theta$ is less than 1 in the version of the model without capital. Here, as in that case, the long-run growth rate of the economy is endogenous, and again long-run growth is an increasing function of population growth and is zero if population growth is zero. The fractions of the labor force and the capital stock engaged R&D, $a_L$ and $a_K$, do not affect long-run growth; nor does the saving rate, $s$. The reason that these parameters do not affect long-run growth is essentially the same as the reason that $a_L$ does not affect long-run growth in the simple version of the model.

same as equation (29) in the simple version of the model. Taking logs and differentiating with respect to time gives

$$\frac{\dot{g}_A(t)}{g_A(t)} = \beta g_K(t) + \gamma n + (\theta - 1) g_A(t) \tag{31}$$

Thus $g_A$ is rising if $\beta g_K + \gamma n + (\theta - 1) g_A$ is positive, falling if it is negative, and constant if it is zero. This is shown in Figure 9. The set of points where $g_A$ is constant has an intercept of $-\gamma n / \beta$ and a slope of $(1-\theta)/\beta$ (the figure is drawn for the case of $\theta < 1$, so this slope is shown as positive). Above this locus, $g_A$ is rising; and below the locus, it is falling.

The production function for output (equation (23)) exhibits constant returns to scale in the two produced factors of production, capital and knowledge. Thus whether there are on net increasing, decreasing, or constant returns to scale to the produced factors depends on their returns to scale in the production function for knowledge, equation (24). As that equation shows, the degree of returns to scale to *K* and *A* in knowledge production is $\beta + \theta$ : increasing both *K* and *A* by a factor of *X* increases $\dot{A}$ by a factor of $X^{\beta+\theta}$. Thus the key determinant of the economy's behavior is now not how $\theta$ compares with 1, but how $\beta + \theta$ compares with 1. As before, we discuss each of the three possibilities.

**Case 1:** $\beta + \theta < 1$

If $\beta + \theta$ is less than 1, $(1-\theta)/\beta$ is greater than 1. Thus the locus of points where $\dot{g}_A = 0$ is steeper than the locus where $\dot{g}_K = 0$. This case is shown in Figure 10. The initial values of $g_A$ and $g_K$ are determined by the parameters of the model and by the initial values of *A*, *K*, and *L*. Their dynamics are then as shown in the figure.

The figure shows that regardless of where $g_A$ and $g_K$ begin, they converge to Point E in the diagram. Both $\dot{g}_A$ and $\dot{g}_K$ are zero at this point. Thus the values of $g_A$ and $g_K$ at Point E, which we denote $g_A^*$ and $g_K^*$, must satisfy

$$g_A^* + n - g_K^* = 0 \tag{32}$$

and

$$\beta g_K^* + \gamma n + (\theta - 1) g_A^* = 0 . \tag{33}$$

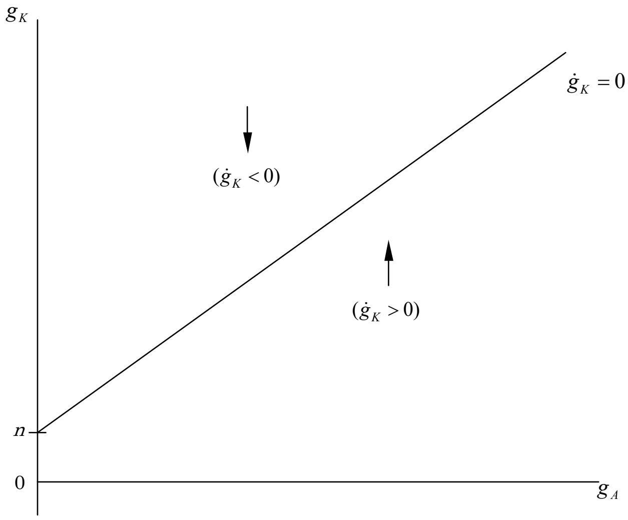
$$
\begin{aligned}
g_K(t) &\equiv \frac{\dot{K}(t)}{K(t)} \\
&= c_K\left[\frac{A(t)L(t)}{K(t)}\right]^{1-\alpha}
\end{aligned}
\quad . \tag{28}
$$

Taking logs of both sides and differentiating with respect to time yields

$$
\frac{\dot{g}_K(t)}{g_K(t)} = (1-\alpha)[g_A(t) + n - g_K(t)]. \tag{29}
$$

From (27), $g_K$ is always positive. Thus $g_K$ is rising if $g_A + n - g_K$ is positive, falling if this expression is negative, and constant if it is zero. This information is summarized in Figure 3. In ($g_A, g_K$) space, the locus of points where $g_K$ is constant has an intercept of $n$ and a slope of 1. Above the locus, $g_K$ is falling; below the locus, it is rising.

**Figure 3 The Dynamics of the Growth Rate of Capital in the General Version of the Model**




Similarly, dividing both sides of equation (24), $\dot{A} = B(a_K K)^{\beta}(a_L L)^{\gamma} A^{\theta}$, by $A$ yields an expression for the growth rate of $A$:

$$
g_A(t) = c_A K(t)^{\beta} L(t)^{\gamma} A(t)^{\theta-1} \tag{30}
$$

where $c_A \equiv B a_K^{\beta} a_L^{\gamma}$. Aside from the presence of the $K^{\beta}$ term, this is essentially the consider the possibility that it is negative. Thus,

$$\dot{L}(t) = nL(t), \qquad n \geq 0 . \tag{26}$$

This completes the description of the model.

**The Importance of Returns to Scale to Produced Factors**

The reason that these three cases have such different implications is that whether $\theta$ is less than, greater than, or equal to 1 determines whether there are decreasing, increasing, or constant returns to scale to *produced* factors of production. The growth of labor is exogenous, and we have eliminated capital from the model; thus knowledge is the only produced factor. There are constant returns to knowledge in goods production. Thus whether there are on the whole increasing, decreasing, or constant returns to knowledge in this economy is determined by the returns to scale to knowledge in knowledge production – that is, by $\theta$.

To see why the returns to the produced input are critical to the behavior of the economy, suppose that the economy is on some path, and suppose there is an exogenous increase in *A* of 1 percent. If $\theta$ is exactly equal to 1, $\dot{A}$ grows by 1 percent as well: knowledge is just productive enough in the production of new knowledge that the increase in *A* is self-sustaining. Thus the jump in *A* has no effect on its growth rate. If $\theta$ exceeds 1, the 1 percent increase in *A* causes more than a 1 percent increase in $\dot{A}$. Thus in this case the increase in *A* raises the growth rate of *A*. Finally, if $\theta$ is less than 1, the 1 percent increase in *A* results in an increase of less than 1 percent in $\dot{A}$, and so the growth rate of knowledge falls.

**The Dynamics of Knowledge and Capital**

When the model includes capital, there are two endogenous stock variables, *A* and *K*. Paralleling our analysis of the simple model, we focus on the dynamics of the growth rates of *A* and *K*. Substituting the production function, (23), into the expression for capital accumulation, (25), yields

$$\dot{K}(t) = s(1-a_K)^{\alpha}(1-a_L)^{1-\alpha} K(t)^{\alpha} A(t)^{1-\alpha} L(t)^{1-\alpha} \tag{27}$$

Dividing both sides by $K(t)$ and defining $c_K = s(1-a_K)^{\alpha}(1-a_L)^{1-\alpha}$ gives us

The quantity of output produced at time $t$ is thus

$$Y(t) = \left[(1 - a_K)K(t)\right]^{\alpha} \left[A(t)(1 - a_L)L(t)\right]^{1-\alpha}, \qquad 0 < \alpha < 1. \tag{23}$$

Aside from the $1 - a_K$ and $1 - a_L$ terms and the restriction to the Cobb-Douglas functional form, this production function is identical to those of our earlier models. Note that equation (23) implies constant returns to capital and labor: with a given technology, doubling the inputs doubles the amount that can be produced.

The production of new ideas depends on the quantities of capital and labor engaged in research and on the level of technology. Given our assumption of generalized Cobb-Douglas production, we therefore write

$$\dot{A}(t) = B\left[a_K K(t)\right]^{\beta} \left[a_L L(t)\right]^{\gamma} A(t)^{\theta}, \qquad B > 0, \quad \beta \geq 0, \quad \gamma \geq 0, \tag{24}$$

where $B$ is a shift parameter.

Notice that the production function for knowledge is not assumed to have constant returns to scale to capital and labor. The standard argument that there must be at least constant returns is a replication one: if the inputs double, the new inputs can do exactly what the old ones were doing, thereby doubling the amount produced. But in the case of knowledge production, exactly replicating what the existing inputs were doing would cause the same set of discoveries to be made twice, thereby leaving $A$ unchanged. Thus it is possible that there are diminishing returns in R&D. At the same time, interactions among researchers, fixed setup costs, and so on may be important enough in R&D that doubling capital and labor more than doubles output. We therefore also allow for the possibility of increasing returns.

The parameter $\theta$ reflects the effect of the existing stock of knowledge on the success of R&D. This effect can operate in either direction. On the one hand, past discoveries may provide ideas and tools that make future discoveries easier. In this case, $\theta$ is positive. On the other hand, the easiest discoveries may be made first. In this case, it is harder to make new discoveries when the stock of knowledge is greater; thus $\theta$ is negative. Because of these conflicting effects, no restriction is placed on $\theta$ in (24).

As in the Solow model, the saving rate is exogenous and constant. In addition, depreciation is set to zero for simplicity. Thus,

$$\dot{K}(t) = sY(t) \tag{25}$$

Finally, we continue to treat population growth as exogenous. For simplicity, we do not

## Framework and Assumptions

The view of growth that is most in keeping with the models we have seen is that the effectiveness of labor represents knowledge or technology. Certainly it is plausible that technological progress is the reason that more output can be produced today from a given quantity of capital and labor than could be produced a century or two ago. The natural extension is thus to model the growth of *A* rather than to take it as given.

To do this, we need to introduce an explicit *research and development* (or R&D) sector and then model the production of new technologies. We also need to model the allocation of resources between conventional goods production and R&D.

Since we are interested in growth over extended periods, modeling the randomness in technological progress would give little additional insight. And if we want to analyze the consequences of changes in other determinants of the success of R&D, we can introduce a shift parameter in the knowledge production function and examine the effects of changes in that parameter. The model provides no insight, however, concerning what those other determinants of the success of research activity are.

We make two other major simplifications. First, both the R&D and goods production functions are assumed to be generalized Cobb-Douglas function; that is, they are power functions, but the sum of the exponents on the inputs is not necessarily restricted to 1. Second, in the spirit of the Solow model, the model takes the fraction of output saved and the fractions of the labor force and the capital stock used in the R&D sector as exogenous and constant. These assumptions do not change the model's main implications.

## The Basic Model

The specific model we consider is a simplified version of the models of R&D and growth developed by Romer (1990), Grossman and Helpman (1991a), and Aghion and Howitt (1992). The model, like the others we have studied, involves four variables: labor ($L$), capital ($K$), technology ($A$), and output ($Y$). The model is set in continue time. There are two sectors, a goods-producing sector where output is produced and an R&D sector where additions to the stock of knowledge are made. Fraction $a_L$ of the labor force is used in the R&D sector and fraction $1-a_L$ in the goods-producing sector. Similarly, fraction $a_K$ of the capital stock is used in R&D and the rest in goods production. Both $a_L$ and $a_K$ are exogenous and constant. Because the use of an idea or a piece of knowledge in one place does not prevent it from being used elsewhere, both sectors use the full stock of knowledge, $A$.

factor returns rather than quantities. If rapid growth comes solely from capital accumulation, for example, we will see either a large fall in the return to capital or a large rise in capital's share (or a combination). Doing the growth accounting this way, Hsieh finds a much larger role for the residual.

To give another example, growth accounting has been used extensively to study the productivity growth slowdown – the reduced growth rate of output per worker-hour in the United States and other industrialized countries that began in the early 1970s (see, for example, Denison, (1985) Baily and Gordon, (1988) Griliches, (1988) and Jorgenson, (1988). Some candidate explanations that have been proposed on the basis of this research include slower growth in workers' skills, the disruptions caused by the oil price increases of the 1970s, a slowdown in the rate of inventive activity, and the effects of government regulations.

In the mid-1990s, U.S. productivity growth returned to close to its level before the slowdown. Growth accounting has been used to study this rebound as well (Oliner and Sichel, (2000) Jorgenson and Stiroh, (2000) and Whelan, (2000). This research suggests that computers and other types of information technology are the main source of the rebound. Until the mid-1990s, the rapid technological progress in computers and their broad adoption appear to have had little impact on aggregate productivity (see exercise 1). Since then, however, their impact has been substantial.

Given that computer use is still spreading rapidly, this analysis suggests it is likely that the rapid productivity growth of the late 1990s will be sustained for at least a few more years. Even this is far from certain, however, and it is certainly far too soon to know whether the rebound will be long-lasting.

## 8.2 New Growth Theory[10]

This section investigates the fundamental questions of growth theory more deeply. The first part of the section examines the accumulation of knowledge. One can think of the models we will consider there as elaborations of the Solow model, given in the previous section. They treat capital accumulation and its role in production in ways that are similar to those earlier models. But they differ from the earlier models in explicitly interpreting the effectiveness of labor as knowledge and in formally modeling its evolution over time. We will analyze the dynamics of the economy when knowledge accumulation is endogenous and consider various views concerning how knowledge is produced and what determines the allocation of resources to knowledge production.

[10] This section is drawn from D. Romer (2001, Chapter 3, pp.98-101 and 107-115). In my later version, I would like to replace this section with my own materials.

sides by $Y(t)$ and rewriting the terms on the right-hand side yields

$$\begin{aligned}\frac{\dot{Y}(t)}{Y(t)} &= \frac{K(t)}{Y(t)}\frac{\partial Y(t)}{\partial K(t)}\frac{\dot{K}(t)}{K(t)} + \frac{L(t)}{Y(t)}\frac{\partial Y(t)}{\partial L(t)}\frac{\dot{L}(t)}{L(t)} + \frac{A(t)}{Y(t)}\frac{\partial Y(t)}{\partial A(t)}\frac{\dot{A}(t)}{A(t)} \\ &\equiv \alpha_K(t)\frac{\dot{K}(t)}{K(t)} + \alpha_L(t)\frac{\dot{L}(t)}{L(t)} + R(t)\end{aligned}. \quad (21)$$

Here $\alpha_L(t)$ is the elasticity of output with respect to labor at time $t$, $\alpha_K(t)$ is again the elasticity of output with respect to capital, and $R(t) \equiv [A(t)/Y(t)][\partial Y(t)/\partial A(t)][\dot{A}(t)/A(t)]$. Subtracting $\dot{L}(t)/L(t)$ from both sides and using the fact that $\alpha_L(t) + \alpha_K(t) = 1$ gives an expression for the growth rate of output per worker:

$$\frac{\dot{Y}(t)}{Y(t)} - \frac{\dot{L}(t)}{L(t)} = \alpha_K(t)\left[\frac{\dot{K}(t)}{K(t)} - \frac{\dot{L}(t)}{L(t)}\right] + R(t). \quad (22)$$

The growth rates of $Y$, $K$ and $L$ are straightforward to measure. And we know that if capital earns its marginal product, $\alpha_K$ can be measured using data on the share of income that goes to capital. $R(t)$ can then be measured as the residual in (22). Thus (22) provides a way of decomposing the growth of output per worker into the contribution of growth of capital per worker and a remaining term, the *Solow residual* (or total factor productivity: TFP). The Solow residual is sometimes interpreted as a measure of the contribution of technological progress. As the derivation shows, however, it reflects all sources of growth other than the contribution of capital accumulation via its private return.

This basic framework can be extended in many ways (see, for example, Denison, 1967). The most common extensions are to consider different types of capital and labor and to adjust for changes in the quality of inputs. But more complicated adjustments are also possible. For example, if there is evidence of imperfect competition, one can try to adjust the data on income shares to obtain a better estimate of the elasticity of output with respect to the different inputs.

Growth accounting has been applied to many issues. For example, it has played a major role in a recent debate concerning the exceptionally rapid growth of the newly industrializing countries of East Asia. Young (1995) uses detailed growth accounting to argue that the higher growth in these countries than in the rest of the world is almost entirely due to rising investment, increasing labor force participation, and improving labor quality (in terms of education), and not to rapid technological progress and other forces affecting the Solow residual. Hsieh (1998a), however, observes that one can do growth accounting by examining the behavior of

More fundamentally, the model does not identify what the "effectiveness of labor" is; it is just a catchall for factors other than labor and capital that affect output. To proceed, we must take a stand concerning what we mean by the effectiveness of labor and what causes it to vary. One natural possibility is that the effectiveness of labor corresponds to abstract knowledge. To understand worldwide growth, it would then be necessary to analyze the determinants of the stock of knowledge over time. To understand cross-country differences in real incomes, one would have to explain why firms in some countries have access to more knowledge than firms in other countries, and why that greater knowledge is not rapidly transferred to poorer countries.

There are other possible interpretations of $A$: the education and skills of the labor force, the strength of property rights, the quality of infrastructure, cultural attitudes toward entrepreneurship and work, and so on. Or $A$ may reflect a combination of forces. For any proposed view of what $A$ represents, one would again have to address the questions of how it affects output, how it evolves over time, and why it differs across parts of the world.

The other possible way to proceed is to consider the possibility that capital is more important than the Solow model implies. If capital encompasses more than just physical capital, or if physical capital has positive externalities, then the private return on physical capital is not an accurate guide to capital's importance in production. In this case, the calculations we have done may be misleading, and it may be possible to resuscitate the view that differences in capital are central to differences in incomes.

**Empirical Applications**

*Growth Accounting*

In the Solow model, long-run growth of output per worker depends only on technological progress. But short-run growth can result from either technological progress or capital accumulation. Thus the model implies that determining the sources of short-run growth is an empirical issue. *Growth accounting*, which was pioneered by Abramovits (1956) and Solow (1957), provides a way of tackling this subject.

To see how growth accounting works, consider again the production function $Y(t) = F\big(K(t), A(t)L(t)\big)$. This implies

$$\dot{Y}(t) = \frac{\partial Y(t)}{\partial K(t)}\dot{K}(t) + \frac{\partial Y(t)}{\partial L(t)}\dot{L}(t) + \frac{\partial Y(t)}{\partial A(t)}\dot{A}(t)\,. \tag{20}$$

$\partial Y/\partial L$ and $\partial Y/\partial A$ denote $[\partial Y/\partial(AL)]A$ and $[\partial Y/\partial(AL)]L$, respectively. Dividing both worker on the basis of differences in capital per worker is to notice that the required differences in capital imply enormous differences in the rate of return on capital (Lucas, 1990). If markets are competitive, the rate of return on capital equals its marginal product, $f(k)=k^{\alpha}$. With this production function, the elasticity of output with respect to capital is simply $\alpha$. The marginal product of capital is

$$\begin{aligned} f'(k) &= \alpha k^{\alpha-1} \\ &= \alpha y^{(\alpha-1)/\alpha} \end{aligned} \tag{19}$$

Equation (19) implies that the elasticity of the marginal product of capital with respect to output is $-(1-\alpha)/\alpha$. If $\alpha = 1/3$, a tenfold difference in output per worker arising from differences in capital per worker thus implies a hundredfold difference in the marginal product of capital. And since the return to capital is $f'(k)-\delta$, the difference in rates of return is even larger.

Again, there is no evidence of such differences in rates of return. Direct measurement of returns on financial assets, for example, suggests only moderate variation over time and across countries. More tellingly, we can learn much about cross-country differences simply by examining where the holders of capital want to invest. If rates of return were larger by a factor of 10 or 100 in poor countries than in rich countries, there would be immense incentives to invest in poor countries. Such differences in rates of return would swamp such considerations as capital-market imperfections, government tax policies, fear of expropriation, and so on, and we would observe immense flows of capital from rich to poor countries. We do not see such flows[9].

Thus differences in physical capital per worker cannot account for the differences in output per worker that we observe, at least if capital's contribution to output is roughly reflected by its private returns.

The other potential source of variation in output per worker in the Solow model is the effectiveness of labor. Attributing differences in standards of living to differences in the effectiveness of labor does not require huge differences in capital or in rates of return. Along a balanced growth path, for example, capital is growing at the same rate as output; and the marginal product of capital, $f'(k)$, is constant.

The Solow model's treatment of the effectiveness of labor is highly incomplete, however. Most obviously, the growth of the effectiveness of labor is exogenous: the model takes as given the behavior of the variable that it identifies as the driving force of growth. Thus it is only a small exaggeration to say that we have been modeling growth by assuming it.

---

[9] See Lucas (1990)

**The Central Questions of Growth Theory**

The Solow model identities two possible sources of variation – either over time or across parts of the world – in output per worker: differences in capital per worker ($K/L$) and differences in the effectiveness of labor (*A*). We have seen, however, that only growth in the effectiveness of labor can lead to permanent growth in output per worker, and that for reasonable cases the impact of changes in capital per worker on output per worker is modest. As a result, only differences in the effectiveness of labor have any reasonable hope of accounting for the vast differences in wealth across time and space. Specifically, the central conclusion of the Solow model is that if the returns that capital commands in the market are a rough guide to its contributions to output, then variations in the accumulation of physical capital do not account for a significant part of either worldwide economic growth or cross-country income differences.

There are two ways to see that the Solow model implies that differences in capital accumulation cannot account for large differences in incomes, one direct and the other indirect. The direct approach is to consider the required differences in capital per worker. Suppose we want to account for a difference of a factor of $X$ in output per worker between two economies on the basis of differences in capital per worker. If output per worker differs by a factor of $X$, the difference in log output per worker between the two economies is $\ln X$. Since the elasticity of output per worker with respect to capital per worker is $\alpha_K$, log capital per worker must differ by $(\ln X)/\alpha_K$. That is, capital per worker differs by a factor of $e^{(\ln X)/\alpha_K}$, or $X^{1/\alpha_K}$.

There is no evidence of such differences in capital stocks. Capital-output ratios are roughly constant over time. Similarly, although capital-output ratios vary somewhat across countries, the variation is not great. For example, the capital-output ratio appears to be 2 to 3 times larger in industrialized countries than in poor countries; thus capital per worker is "only" about 20 to 30 times larger. In sum, differences in capital per worker are far smaller than those needed to account for the differences in output per worker that we are trying to understand[8].

The indirect way of seeing that the model cannot account for large variations in output per

---

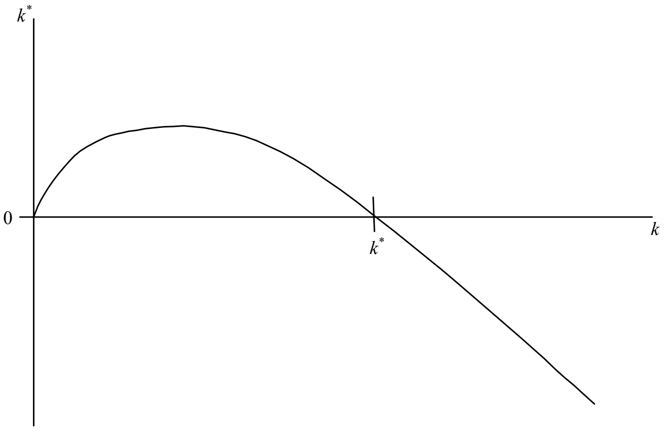
(1999a) argues, however, that the underlying determinants of the level of income on the balanced growth path have in fact been far from constant in these economies, and thus that the resemblance between these economies and the balanced growth path of the Solow model is misleading.

[8] One can make the same point in terms of the rates of saving, population growth, and so on that determine capital per worker. For example, the elasticity of $y^*$ with respect to $s$ is $\alpha_K/(1-\alpha_K)$. Thus accounting for a difference of a factor of 10 in output per worker on the basis of differences in $s$ requires a difference of a factor of 100 in $s$ if $\alpha_K = 1/3$ and a difference of a factor of 10 if $\alpha_K = 1/2$. Variations in actual saving rates are much smaller than this.

two must eventually cross. Finally, the fact that $f''(k) < 0$ implies that the two lines intersect only once for $k > 0$. We let $k^*$ denote the value of $k$ where actual investment and break-even investment are equal.

Figure 2 summarizes this information in the form of a *phase diagram*, which shows $\dot{k}$ as a function of $k$. If $k$ is initially less than $k^*$, actual investment exceeds break-even investment, and so $\dot{k}$ is positive – that is, $k$ is rising. If $k$ exceeds $k^*$, $\dot{k}$ is negative. Finally, if $k$ equals $k^*$, then $\dot{k}$ is zero. Thus, regardless of where $k$ starts, it converges to $k^*$.

**Figure 2 The Phase Diagram for k in the Solow Model**



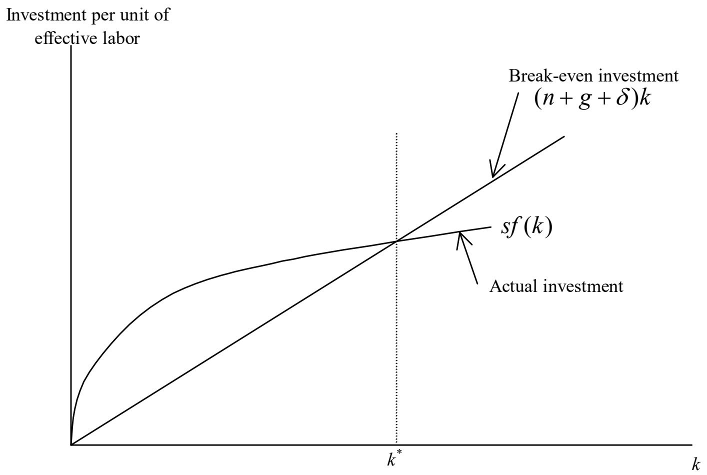

*The Balanced Growth Path*

Since $k$ converges to $k^*$, it is natural to ask how the variables of the model behave when $k$ equals $k^*$. By assumption, labor and knowledge are growing at rates *n* and *g*, respectively. The capital stock, *K*, equals *ALK*; since $k$ is constant at $k^*$, *K* is growing at rate $n+g$ (that is, $K/K$ equals $n+g$). With both capital and effective labor growing at rate $n+g$, the assumption of constant returns implies that output, *Y*, is also growing at that rate. Finally, capital per worker, $K/L$, and output per worker, $Y/L$, are growing at rate *g*.

Thus the Solow model implies that, regardless of its starting oint, the economy converges to a *balanced growth* path – a situation where each variable of the model is growing at a constant rate. On the balanced growth path, the growth rate of output per worker is determined solely by the rate of technological progress[7].

[7] The broad behavior or the U.S. economy and many other major industrialized economies over the last century or more is described reasonably well by the balanced growth path of the Solow model. The growth rates of labor, capital, and output have each been roughly constant. The growth rates of output and capital have been about equal (so that the capital-output ratio has been approximately constant) and have been larger than the growth rate of labor (so that output per worker and capital per worker have been rising). This is often taken as evidence that it is reasonable to think of these economies as Solow-model economies on their balanced growth paths. Jones

**Figure 1 Actual and Break-even Investment**




Equation (18) is the key equation of the Solow model. It states that the rate of change of the capital stock per unit of effective labor is the difference between two terms. The first, $sf(k)$, is actual investment per unit of effective labor: output per unit of effective labor is $f(k)$, and the fraction of that output that is invested is *s*. The second term, $(n+g+\delta)k$, is *break-even investment*, the amount of investment that must be done just to keep *k* at its existing level. There are two reasons that some investment is needed to prevent *k* from falling. First, existing capital is depreciating; this capital must be replaced to keep the capital stock from falling. This is the $\delta k$ term in (18). Second, the quantity of effective labor is growing. Thus doing enough investment to keep the capital stock (*K*) constant is not enough to keep the capital stock per unit of effective labor (*k*) constant. Instead, since the quantity of effective labor is growing at rate $n+g$, the capital stock must grow at rate $n+g$ to hold *k* steady. This is the $(n+g)k$ term in (18).

When actual investment per unit of effective labor exceeds the investment needed to break even, *k* is rising. When actual investment falls short of break-even investment, *k* is falling. And when the two are equal, *k* is constant.

Figure 1 plots the two terms of the expression for $\dot{k}$ as function of *k*. Break-even investment, $(n+g+\delta)k$, is proportional to *k*. Actual investment, $sf(k)$, is a constant times output per unit of effective labor.

Since $f(0)=0$, actual investment and break-even investment are equal at $k=0$. The Inada conditions imply that at $k=0$, $f'(k)$ is large, and thus that the $sf(k)$ line is steeper than the $(n+g+\delta)k$ line. Thus for small values of *k*, actual investment is larger than break-even investment. The Inada conditions also imply that $f'(k)$ falls toward zero as *k* becomes large. At some point, the slope of the actual investment line falls below the slope of the break-even investment line. With the $sf(k)$ line flatter than the $(n+g+\delta)k$ line, the

**The Dynamics of the Model**

We want to determine the behavior of the economy we have just described. The evolution of two of the three inputs into production, labor and knowledge, is exogenous. Thus to characterize the behavior of the economy, we must analyze the behavior of the third input, capital.

*The Dynamics of k*

Because the economy may be growing over time, it turns out to be much easier to focus on the capital stock per unit of effective labor, *k*, than on the unadjusted capital stock, *K*. Since $k = K/AL$, we can use the chain rule to find

$$\begin{aligned} \dot{k}(t) &= \frac{\dot{K}(t)}{A(t)L(t)} - \frac{K(t)}{[A(t)L(t)]^2}\left[A(t)\dot{L}(t) + L(t)\dot{A}(t)\right] \\ &= \frac{\dot{K}(t)}{A(t)L(t)} - \frac{K(t)}{A(t)L(t)}\frac{\dot{L}(t)}{L(t)} - \frac{K(t)}{A(t)L(t)}\frac{\dot{A}(t)}{A(t)}. \end{aligned} \tag{16}$$

$K/AL$ is simply *k*. From equation (8) and (9), $\dot{L}/L$ and $\dot{A}/A$ are *n* and *g*, respectively. $\dot{K}$ is given by (15). Substituting these facts into (16) yields

$$\begin{aligned} \dot{k}(t) &= \frac{sY(t) - \delta K(t)}{A(t)L(t)} - k(t)n - k(t)g \\ &= s\frac{Y(t)}{A(t)L(t)} - \delta k(t) - nk(t) - gk(t). \end{aligned} \tag{17}$$

Finally, using the fact that $Y/AL$ is given by $f(k)$, we have

$$\dot{k}(t) = sf(k(t)) - (n + g + \delta)k(t)\,. \tag{18}$$

constant and equal to $g$.

A key fact about growth rates is that the growth rate of a variable equals the rate of change of its natural log. That is, $\dot{X}(t)/X(t) = d\ln X(t)/dt$. To see this, note that since $\ln X$ is a function of $X$ and $X$ is a function of $t$, we can use the chain rule to write

$$\begin{aligned}\frac{d\ln X(t)}{dt} &= \frac{d\ln X(t)}{dX(t)}\frac{dX(t)}{dt} \\ &= \frac{1}{X(t)}\dot{X}(t).\end{aligned} \tag{10}$$

Applying the result that a variable's growth rate equals the rate of change of its log to (8) and (9) tells us that the rates of change of the logs of $L$ and $A$ are constant and that they equal $n$ and $g$, respectively. Thus,

$$\ln L(t) = [\ln L(0)] + nt \tag{11}$$
$$\ln A(t) = [\ln A(0)] + gt \tag{12}$$

where $L(0)$ and $A(0)$ are the values of $L$ and $A$ at time 0. Exponentiating both sides of these equations gives us

$$L(t) = L(0)e^{nt} \tag{13}$$
$$A(t) = A(0)e^{nt} \tag{14}$$

Thus, our assumption is that $L$ and $A$ each grow exponentially.

Output is divided between consumption and investment. The fraction of output devoted to investment, $s$, is exogenous and constant. One unit of output devoted to investment yields one unit of new capital. In addtion, existing capital depreciates at rate $\delta$. Thus

$$\dot{K}(t) = sY(t) - \delta K(t). \tag{15}$$

Although no restrictions are placed on $n$, $g$, and $\delta$ individually, their sum is assumed to be positive. This completes the description of the model.

The Solow model is grossly simplified in a host of ways. To give just a few examples, there is only a single good; government is absent; fluctuations in employment are ignored; production is described by an aggregate production function with just three inputs; and the rates of saving, depreciation, population growth, and technological progress are constant.

$$
\begin{aligned}
F(cK,cAL) &= (cK)^{\alpha}(cAL)^{1-\alpha} \\
&= c^{\alpha}c^{1-\alpha}K^{\alpha}(AL)^{1-\alpha} \\
&= cF(K,AL).
\end{aligned} \tag{6}
$$

To find the intensive form of the production function, divide both inputs by $AL$; this yields

$$
\begin{aligned}
f(k) &\equiv F\left(\frac{K}{AL},1\right) \\
&= \left(\frac{K}{AL}\right)^{\alpha} \\
&= k^{\alpha}
\end{aligned} \tag{7}
$$

Equation (7) implies that $f'(k) = \alpha k^{\alpha-1}$. It is straightforward to check that this expression is positive, that it approaches infinity as $k$ approaches zero, and that it approaches zero as $k$ approaches infinity. Finally, $f''(k) = -(1-\alpha)\alpha k^{\alpha-2}$, which is negative[5].

**The Evolution of the Inputs into Production**

The remaining assumptions of the model concern how the stocks of labor, knowledge, and capital change over time. The model is set in continuous time; that is, the variables of the model are defined at every point in time[6].

The initial levels of capital, labor, and knowledge are taken as given. Labor and knowledge grow at constant rates:

$$\dot{L}(t) = nL(t) \tag{8}$$
$$\dot{A}(t) = gA(t) \tag{9}$$

where $n$ and $g$ are exogenous parameters and where a dot over a variable denotes a derivative with respect to time (that is, $\dot{X}(t)$ is shorthand for $dX(t)/dt$).

The *growth rate* of a variable refers to its proportional rate of change. That is, the phrase the *growth rate of X* refers to the quantity $\dot{X}(t)/X(t)$. Thus equation (8) implies that the growth rate of $L$ is constant and equal to $n$, and equation (9) implies that A's growth rate is

---

5 Note that with Cobb-Douglas production, labor-augmenting, capital-augmenting, and Hicks-neutral technological progress are all essentially the same. For example, to rewrite equation (5) so that technological progress is Hicks-neutral, simply define $\widetilde{A} = A^{1-\alpha}$; then $Y = \widetilde{A}\left(K^{\alpha}L^{1-\alpha}\right)$.

6 The alternative is discrete time, where the variables are defined only at specific dates (usually $t$=0,1,2,…). The choice between continuous and discrete time is usually based on convenience. For example, the Solow model has essentially the same implications in discrete as in continuous time, but is easier to analyze in continuous time.

$$y = f(k) \qquad (4)$$

That is, we can write output per unit of effective labor as a function of capital per unit of effective labor.

These new variables, $k$ and $y$, are not of interest in their own right. Rather, they are tools for learning about the variables we are interested in. As we will see, the easiest way to analyze the model is to focus on the behavior of $k$, rather than to consider directly the behavior of the two arguments of the production function, $K$ and $AL$. For example, we will determine the behavior of output per worker, $Y/L$, by writing it as $A(Y/AL)$, or $Af(k)$, and determining the behavior of $A$ and $k$.

To see the intuition behind equation (4), think of dividing the economy into $AL$ small economies, each with 1 unit of effective labor and $K/AL$ units of capital. Since the production function has constant returns, each of these small economies produces $1/AL$ as much as is produced in the large, undivided economy. Thus the amount of output per unit of effective labor depends only on the quantity of capital per unit of effective labor, and not on the overall size of the economy. This is expressed mathematically in equation (4).

The intensive-form production function, $f(k)$, is assumed to satisfy $f(0) = 0$, $f'(k) > 0$, $f''(k) < 0$. Since $F(K, AL) = ALf(K/AL)$, it follows that the marginal product of capital, $\partial F(K, AL)/\partial K = ALf'(K/AL)(1/AL) = f'(k)$. Thus the assumptions that $f'(k)$ is positive and $f''(k)$ is negative imply that the marginal product of capital is positive, but that it declines as capital (per unit of effective labor) rises. In addition, $f(\bullet)$ is assumed to satisfy the *Inada conditions*: $\lim_{k \to 0} f'(k) = \infty$, $\lim_{k \to \infty} f'(k) = 0$. These conditions (which are stronger than needed for the model's central results) state that the marginal product of capital is very large when the capital stock is sufficiently small and that it becomes very small as the capital stock becomes large; their role is to ensure that the path of the economy does not diverge.

A specific example of a production function is the Cobb-Douglas function,

$$F(K, AL) = K^{\alpha}(AL)^{1-\alpha}, \quad 0 < \alpha < 1. \qquad (5)$$

This production function is easy to analyze, and it appears to be a good first approximation to actual production functions.

It is easy to check that the Cobb-Douglas function has constant returns. Multiplying both inputs by $c$ gives us

The central assumptions of the Solow model concern the properties of the production function and the evolution of the three inputs into production (capital, labor, and knowledge) over time. We discuss each in turn.

*Assumptions Concerning the Production Function*

The model's critical assumption concerning the production function is that it has constant returns to scale in its two arguments, capital and effective labor. That is, doubling the quantities of capital and effective labor (for example, by doubling *K* and *L* with *A* held fixed) doubles the amount produced. More generally, multiplying both arguments by any nonnegative constant *c* causes output to change by the same factor:

$$F(cK, cAL) = cF(K, AL) \qquad \text{for all } c \geq 0. \tag{2}$$

The assumption of constant returns can be thought of as combining two assumptions. The first is that the economy is big enough that the gains from specialization have been exhausted. In a very small economy, there are probably enough possibilities for further specialization that doubling the amounts of capital and labor more than doubles output. The Solow model assumes, however, that the economy is sufficiently large that, if capital and labor double, the new inputs are used in essentially the same way as the existing inputs, and thus that output doubles.

The second assumption is that inputs other than capital, labor, and knowledge are relatively unimportant. In particular, the model neglects land and other natural resources. If natural resources are important, doubling capital and labor could less than double output. In practice, the availability of natural resources does not appear to be a major constraint on growth. Assuming constant returns to capital and labor alone therefore appears to be a reasonable approximation.

The assumption of constant returns allows us to work with the production function in *intensive form*. Setting $c = 1/AL$ in equation (2) yields

$$F\left(\frac{K}{AL}, 1\right) = \frac{1}{AL} F(K, AL) \tag{3}$$

Here $K/AL$ is the amount of capital per unit of effective labor, and $F(K, AL)/AL$ is $Y/AL$, output per unit of effective labor. Define $k = K/AL$, $y = Y/AL$, and $f(k) = F(k,1)$. Then we can rewrite equation (3) as cannot account for either the vast growth over time in output per person or the vast geographic differences in output per person. Specifically, suppose that capital accumulation affects output through the conventional channel that capital makes a direct contribution to production, for which it is paid its marginal product. Then the Solow model implies that the differences in real incomes that we are trying to understand are far too large to be accounted for by differences in capital inputs. The model treats other potential sources of differences in real incomes as either exogenous and thus not explained by the model (in the case of technological progress, for example) or absent altogether (in case of positive externalities from capital, for example). Thus to address the central questions of growth theory, we must move beyond the Solow model.

**Assumptions**

*Inputs and Output*

The Solow model focuses on four variables: output ($Y$), capital ($K$), labor ($L$), and "knowledge" or the "effectiveness of labor" ($A$). At any time, the economy has some amounts of capital, labor and knowledge, and these are combined to produce output. The production function takes the form

$$Y(t) = F(K(t), A(t)L(t)) \tag{1}$$

where $t$ denotes time.

Two features of the production function should be noted. First, time does not enter the production function directly, but only through $K$, $L$, and $A$. That is, output changes over time only if the inputs to production change. In particular, the amount of output obtained from given quantities of capital and labor rises over time – there is technological progress – only if the amount of knowledge increases.

Second, $A$ and $L$ enter multiplicatively. $AL$ is referred to as *effective labor*, and technological progress that enters in this fashion is known as *labor-augmenting* or *Harrod-neutral*[4]. This way of specifying how $A$ enters, together with the other assumptions of the model, will imply that the ratio of capital to output, $K/Y$, eventually settles down. In practice, capital-output ratios do not show any clear upward or downward trend over extended periods. In addition, building the model so that the ratio is eventually constant makes the analysis much simpler. Assuming that $A$ multiplies $L$ is therefore very convenient.

[4] If knowledge enter in the form $Y = F(AK, L)$, technological progress is *capital-augmenting*. If it enters in the form $Y = AF(K, L)$, technological progress is *Hicks-neutral*.

growth performance over most of the twentieth century was dismal, and it is now near the middle of the world income distribution. Sub-saharan African countries such as Chad, Ghana, and Mozambique have been extremely poor throughout their histories and have been unable to obtain any sustained growth in average incomes. As a result, their average incomes have remained close to subsistence levels while average world income has been rising steadily.

Other countries exhibit more complicated growth patterns. The Ivory Coast was held up as the growth model for Africa through the 1970s. From 1960 to 1978, real income per person grew at an average annual rate of almost 4 percent. But in the next decade, average income fell in half. To take another example, average growth in Mexico was extremely high in the 1960s and 1970s, negative in most of the 1980s, and again very high – with a brief but severe interruption in the mid-1990s – since then.

Over the whole of the modern era, cross-country income differences have widened on average. The fact that average incomes in the richest countries at the beginning of the Industrial Revolution were not far above subsistence means that the overall dispersion of average incomes across parts of the world must have been much smaller than it is today (Pritchett, 1997). Over the past few decades, however, there has been no strong tendency either toward continued divergence or toward convergence.

The implications of the vast differences in standards of living over time and across countries for human welfare are enormous. The differences are associated with large differences in nutrition, literacy, infant mortality, life expectancy, and other direct measures of well-being. And the welfare consequences of long-run growth swamp any possible effects of the short-run fluctuations that macroeconomics traditionally focuses on. If real income per person in Bangladesh continues to grow at its postwar average rate of 1.4 percent, it will take close to 200 years for it to reach the current U.S. level. If Bangladech achieves 5 percent growth, as some countries have done, the time will be reduced to only 50 years. To quote Robert Lucas (1988), "Once one starts to think about [economic growth], it is hard to think about anything else."

In the following, we investigate several models of growth. The ultimate objective of research on economic growth is to determine whether there are possibilities for raising overall growth or bringing standards of living in poor countries closer to those in the world leaders[2].

This section focuses on the model that economists have traditionally used to study these issues, the Solow growth model[3]. The Solow model is the starting point for almost all analyses of growth.

The principal conclusion of the Solow model is that the accumulation of physical capital

[2] Jones (1998) provides a treatment of economic growth at a level slightly less advanced than this book's. Barro and Sala-i-Martin (1998) and Aghion and Howitt (1998) provide treatments at a slightly more advanced level.

[3] The Solow model (which is sometimes known as the Solow-Swan model) was developed by Robert Solow (Solow, 1956) and T.W. Swan (Swan, 1956).

# Chapter 8 Finance and Development

## 8.1 The Solow Growth Model[1]

### Some Basic Facts about Economic Growth

Over the past few centuries, standards of living in industrialized countries have reached levels almost unimaginable to our ancestors. Although comparisons are difficult, the best available evidence suggests that average real incomes today in the United States and Western Europe are between 10 and 30 times larger than a century ago, and between 50 and 300 times larger than two centuries ago.

One important exception to this general pattern of increasing growth is the *productivity growth slowdown*. Average annual growth in output per person in the United States and other industrialized countries since the early 1970s has been about a percentage point below its earlier level. The data from the late 1990s suggest a rebound in productivity growth; whether this represents a temporary spurt or the end of the slowdown is not clear, however.

There are also enormous differences in standards of living across parts of the world. Average real incomes in such countries as the United States, Germany, and Japan appear to exceed those in such countries as Bangladesh and Kenya by a factor of between 10 and 20. As with worldwide growth, cross-country income differences are not immutable. Growth in individual countries often differs considerably from average worldwide growth; that is, there are often large changes in countries' relative incomes.

The most striking examples of large changes in relative incomes are *growth miracles* and *growth disasters*. Growth miracles are episodes where growth in a country far exceeds the world average over an extended period, with the result that the country moves rapidly up the world income distribution. Some prominent growth miracles are Japan and the newly industrializing countries (NICs) of East Asia – South Korea, Taiwan, Singapore, and Hong Kong.

Growth disasters are episodes where a country's growth falls far short of the world average. Two very different examples of growth disasters are Argentina and many of the countries of sub-Saharan Africa. In 1990, Argentina's average income was only slightly behind those of the world's leaders, and it appeared poised to become a major industrialized country. But its

[1] This part draws heavily from Romer (2001, Chapter 1, pp.5-17 and 26-30). In my later version, I would like to replace by my materials.