Version: Post-print

THE EQUITY/EFFICIENCY TRADE-OFF IN RETROSPECT

Lars Osberg
Department of Economics
Dalhousie University
Halifax, Nova Scotia

Revised April 26, 1995

Revised version of Keynote address to conference on "Economic Growth and Income Inequality", Laurentian University, Sudbury, Ontario, March 17, 1995. I would like to thank Mike Bradfield, Peter Burton, Shelley Phipps and Brian MacLean for their helpful comments and SSHRC for its financial support. I am solely responsible for any remaining errors.
Writing in 1975, Arthur Okun is generally credited with popularizing the idea of a "great tradeoff" between equity and efficiency. A generation of economists has been brought up thinking that in the allocation of scarce resources to competing ends, the tradeoff between "equity" and "efficiency" is just another one of the many tough choices that economically rational individuals have to make. Although "equity" and "efficiency" are rarely defined precisely and the mechanisms which link them are often only vaguely specified, the idea of a tradeoff is still a prominent feature of introductory economics texts (see Baumol, Blinder and Scarth, 1991 or McConnell et al, 1993:316).

However, a major theme of the recent literature on endogenous growth is that this tradeoff does not exist. The common theme of many recent articles, both theoretical and empirical, is that more equal countries grow faster, other things equal. The general presumption of recent growth literature is a positive association between equality and growth -- and the contrast between the 1990s analysis of growth and equality and the 1970s rhetoric on "tradeoffs" deserves some examination.

Section 1 of this essay surveys a sampling from the recent endogenous growth literature. Its objective is to demonstrate that a positive correlation between equality and growth is the new orthodoxy of the growth literature and to illustrate the diversity of empirical evidence and theoretical mechanisms which test and explain the positive relationship between equality and growth.

Section 2 then asks why this reversal in economic thinking has occurred. What was the original argument for an "equity/efficiency tradeoff"? Why are these original arguments no longer considered convincing? What new evidence has been brought forward on the connection between
"equity" and "efficiency"? Section 2.1 argues that part of the reason for the reversal of economic thinking lies in an inherent weakness in the original arguments for the existence of a "tradeoff" - its neglect of the intergenerational transmission of human capital and wealth. Section 2.2 discusses the influence of new empirical evidence on the costs to allocative efficiency of a progressive income tax system. Section 2.3 notes that the reversal of perspective on the equity/efficiency connection is partly due to a redefinition of what type of "efficiency" is really important - i.e. a shift in focus from static, allocative efficiency to a new concern with dynamic efficiency and the rate of growth of output over time. Section 2.4 considers some of the social pressures which may have influenced the debate.

Section 3 briefly considers the implications of economists' changing perspective on equality.

1. **The New Endogenous Growth Theory**

Hammond and Rodriguez-Clare (1993) are one among a number of recent survey articles summarizing the new growth theory.¹ As they note, there has been a long history of economic analysis which has stressed the reasons for divergent patterns of economic growth and the positive (or negative) feedback effects which may help (or hinder) economic growth. In addition to authors such as Tobin (1955) or Myrdal (1957) from the 1950s, one could equally cite Schumpeter, Marx or Adam Smith. However, in the 1960s and 1970s, the mainstream neoclassical approach to growth emphasized the determinants of savings behaviour and the

¹This section makes no claim to being a complete survey of recent growth literature -- the purpose is only to illustrate recent major trends.
diminishing marginal productivity of additional increments to the capital stock. Technical progress was seen as exogenously determined, driven by the internal logic of scientific discovery. Given these assumptions, without change in technology the rate of growth of output tends to zero in the limit (due to the diminishing marginal return to additional savings) but there is nothing that economic policy can do to affect the exogenous rate of scientific advance and technological change.

The explosion of endogenous growth models since 1986/87 has been driven by a perception that in the long run growth rates may not be equal and may depend both on the initial state of a national economy and on the policy measures adopted by national governments. In seeking to explain why some countries have grown so spectacularly in the post-war period while others have not, there is a new emphasis on the importance of externalities in production. Some models (e.g. Romer, 1986) emphasize the importance of the efficiency gains of learning by doing, and the creation of a greater stock of social knowledge of the production process (from which all firms benefit). Others (e.g Barro, 1990) argue that increasing efficiency in labour output is due to the presence of public services which enhance labour productivity. Alternatively, some contend that increasing returns to scale at the industry level are driven by the greater division of labour in intermediate production which is enabled by increased aggregate market size. Others have emphasized the role of feedback effects on the cost of learning or innovation.

Although the early endogenous growth literature (e.g. Romer, 1987; Stokey, 1988) often used a "representative agent" approach, which by construction rules out any consideration of inequality, more recently growth theory has drawn a connection between inequality and growth through a variety of mechanisms. One of the major themes of recent growth theory is the
importance of human capital in production and of the intergenerational transmission of human
capital -- which immediately raises the issue of inequalities of opportunity in inherited human
capital and in access to credit markets. A second major strand of the literature focuses on the
connection between economic inequality and the political economy of tax policy. A third strand
emphasizes the private incentives to social stratification in an unequal society and the externalities
and social inefficiencies of poverty ghettos. A fourth theme is that of the role of wage
compression and incomes policies in encouraging structural change and productivity growth.

A general theme of much of the endogenous growth literature is its emphasis on the
importance of human capital -- see Laitner (1993), Fagerberg (1994) or Sala-i-Martin (1994). In
concluding his survey article on the empirical evidence on economic growth, Sala-i-Martin puts it
very clearly: "Countries with better educated work force tend to grow faster" (1994:746).
Although the private return to greater education has been studied voluminously since the human
capital revolution of the early 1960s, historically there was a nagging doubt about the social
productivity of education. The "credentialist" argument (e.g. Arrow, 1973) maintained that
relative education operates as a signal of underlying ability, ranking individuals in a queue for
available positions, and thereby generating private, but not social, economic returns.

However, by comparing the growth performance of countries with different levels of
average education, the new growth theory has emphasized the social productivity of greater
education. Furthermore, during the 1980s a huge literature has emerged on the increased
differential in annual earnings between university and high school educated workers in the U.S. --
an increased differential which is generally ascribed to structural change in production which has
increased the relative demand for highly educated workers (see Murphy and Welch (1993);
The greater relative increase in supply of highly educated workers in Canada -- the number of university graduates in Canada increased by 26.6% between 1986 and 1991 -- may be the reason why Morrisette, Myles and Picot (1993) concluded that the university/high school earnings differential did not widen in Canada during the 1980s.

Burtless et al (1990); Juhn, Murphy and Pierce (1993)). In the U.S., the increase in the average earnings differential between university educated and high school educated workers is responsible for a substantial part of the increase in total earnings in equality. Indeed, Burtless (1990) and others have noted that the average wage of U.S. high school graduates fell absolutely during the 1980s. Evidently, more education for the disadvantaged would substantially raise their own incomes, decrease economic inequality and increase the growth rate of average incomes -- but if education is so profitable, both socially and privately, why do some people not get it?

Galor and Zeira (1993) were among the first to draw an explicit connection between income distribution and macroeconomic growth. In a model of overlapping generations characterized by intergenerational altruism, even if all individuals are of equal ability and identical preferences, the bequest of the parent determines whether the child will be a skilled or an unskilled worker, which determines in turn the size of the child's bequest to the grandchild. Galor and Zeira assume that credit markets are imperfect, in the sense that the interest rate at which individuals can borrow from financial institutions is greater than the interest rate at which they can lend (due to the existence of monitoring costs) and they assume that there is an indivisible minimum size to human capital investment. As a result, the initial distribution of wealth determines the aggregate amount of human capital investment and the long-run growth rate. They conclude "Countries with a more equal initial distribution of wealth grow more rapidly and

\(^2\)The greater relative increase in supply of highly educated workers in Canada -- the number of university graduates in Canada increased by 26.6% between 1986 and 1991 -- may be the reason why Morrisette, Myles and Picot (1993) concluded that the university/high school earnings differential did not widen in Canada during the 1980s.
have a higher income level in the long run” (1993:48). They also conclude that more equal countries adjust better to technology and price shocks.

Torvik (1993) adopts a similar perspective on credit market imperfections and on intergenerational bequests, but adds the assumption of unequal ability across individuals and "learning by doing" between generations. He argues that there is an intergenerational externality, in that a greater average educational level in one generation will increase the productivity of future generations, because the education and learning by doing of one generation increases the stock of knowledge about production processes which future generations will have to work with. In this model, it is an individual's ability, plus the amount of wealth they inherit from their parents, which determines the personal benefits and costs of investing in education. The more unequal is the initial distribution of inheritances, the smaller is the relative importance of ability in determining an individual's choice of years of education (if inheritances were equal, the only thing that would matter in educational choice is the relative ability of an individual). As a consequence, greater inequality in initial income and bequests implies sub-optimal investment in education and a lower long-run growth rate.

Although a divergence between the rate of interest at which individuals can borrow from, or lend to, financial institutions is a natural consequence of the necessity for institutions to monitor (at some cost) borrowers in order to prevent default, Banerjee and Newman (1993, 1994) adopt a different perspective to explain capital market imperfections. They emphasize that borrowing, particularly to finance human capital acquisition, requires collateral. In their model of intergenerational bequest, individuals choose between the occupations of worker, self-employed artisan or entrepreneur. There is a random element in the returns to entrepreneurial or self-
employed status, but an individual's initial choice of occupation will be determined by their initial wealth -- i.e. their bequest from the previous generation. Since access to capital markets requires collateral, those with little initial wealth necessarily become workers, and the proportion who become entrepreneurs or self-employed depends on the distribution of wealth. Because of this imperfection of capital markets, the assumption of a random component in incomes does not produce the standard markov process result that the distribution of income converges to a unique ergodic distribution (unlike earlier models of intergenerational bequest such as Becker and Tomes, 1979). Banerjee and Newman conclude that a more unequal initial distribution of wealth may lead to stagnation of living standards.

In their 1994 paper, Banerjee and Newman emphasize that if poverty is conceived of as being close to the feasible lower bound on utility, the poor have little to lose by defaulting on a loan. If, in general, the poor have a higher probability of loan default, then lenders have an incentive to deny them access to credit. Barriers in access to capital imply less possibility for investment in human capital, and the intergenerational inheritance of poverty. From a macroeconomic perspective, the larger the percentage of the population which is poor, and therefore unable to invest in human capital, the lower is the aggregate growth rate.

Throughout the world, government intervenes in the human capital acquisition process by providing compulsory public schooling financed from tax revenue. The historic reason for this intervention was the principle of greater equality of opportunity, since there is no guarantee that in an unfettered capitalist market system, all parents will have enough resources and/or enough preference for the education of their children to provide an adequate level of education for all children. Eckstein and Zilcha (1994) note that, in addition, compulsory public schooling can
increase the rate of economic growth. They construct a model of overlapping generations, in which parents have differing levels of preference for their children's education. Since parents do not know their children's preferences for the education of the grandchildren, parents cannot completely internalize in their own decision-making the well-being of all subsequent generations of their lineage, in the manner suggested by Barro (1974). Compulsory public schooling then acts to decrease the dispersion in the schooling of each generation, and thereby decrease intragenerational inequality. By raising the average level of schooling, compulsory schooling improves the long-run growth rate. As Eckstein and Zilcha conclude "government intervention in providing compulsory schooling increases economic growth, while the intragenerational income distribution becomes more equal" (1994:339).

Alesina and Rodrick (1994) and Persson and Tabellini (1994) represent the "political economy" perspective on the connection between economic inequality and growth. Both emphasize the importance of tax policy and both rely upon a median voter model of the democratic process. In Alesina and Rodrick, output is produced by the combination of labour, capital (including human capital) and productive government services, which are financed by a tax on capital income. Since poorer individuals benefit from an increased marginal product of labour when expenditure on productive government services increases, but pay a relatively low share of the cost, the poor prefer a higher tax rate and a suboptimal level of aggregate capital accumulation. The greater the level of inequality, the greater is the divergence between the median voter's income and average incomes, and the farther is the tax rate from its optimal level. As they put it "Inequality is conducive to the adoption of growth retarding policies" (1994:465). In cross-country regressions using the Preston/Summers data base on international growth
performance for 1960-1985, they find support for the proposition that countries with lower levels of income inequality have higher rates of economic growth.

Similarly, Persson and Tabellini conclude that "inequality is harmful for growth" (1994:600). Their model, however, is somewhat different since they emphasize that growth comes from the accumulation of human capital and knowledge useful in production. In a general equilibrium, overlapping generations framework, they assume that the average capital accumulated by one generation (in, for example, technical knowledge) has positive externalities for the income of the subsequent generation. If higher levels of economic inequality produce greater redistribution of income through the tax system, a slower rate of accumulation of capital will diminish long-run growth. They test their model rather ingeniously by constructing 20-year intervals of economic growth performance going back as far as 1830 for a sample of nine countries.

The models of Durlauf (1994) and Benabou (1994) combine a consideration of human capital and political economy influences. Both focus on local spillover effects and the causes and long-run consequences of the creation of poverty ghettos. Durlauf starts from the proposition that when positive spillover effects mean that the marginal payoff of an activity by a given agent is an increasing function of the level of activity of other agents, there will, in general, be multiple equilibria. Spillover effects in education within a neighbourhood mean that the individual payoff to education differs by neighbourhood, and is a function of the average education of the neighbourhood. The mechanism for neighbourhood spillovers may be fiscal, in the sense that the tax rate required to finance good local schools will be lower in rich neighbourhoods with high property assessment than in poor neighbourhoods. Spillovers may also be social, arising from the
interaction of the aspirations and performance of students in affluent and less affluent
neighbourhoods. Either way, Durlauf argues that there is a tendency for stratification to imply
self-replicating neighbourhoods, and the segregation of the poor by housing prices or zoning
legislation. Benabou (1994) adds that stratification by neighbourhood implies increasing
inequality over time and a long-run lowering of the level and rate of growth of human capital and
income.

Although much of the endogenous growth literature emphasizes the importance of
intergenerational transmission of income and wealth, Agell and Lommerud (1993) are an
interesting exception. They note that the Scandinavian countries after World War II were
characterized by strongly egalitarian wage policies enforced by centralized labour unions, and by
low unemployment, rapid structural change and rapid growth in incomes and labour productivity.
They contend that the intellectual architects of the Scandinavian model were rather sceptical
about the effectiveness of wage differentials in motivating intersectoral labour mobility.³ If
expanding sectors have to pay a substantial wage premium to attract labour, their profitability,
retained earnings and growth rate is reduced. The "solidary" wage policy of the Scandinavian
union federations was intended to compress wage differentials between low productivity and high
productivity sectors, and to encourage growth by squeezing labour out of the low productivity
sector, and increasing the profitability of expansion of the high productivity sector.

To this historic prospective, Agell and Lommerud add the idea that learning by doing is an
increasing function of the size of the modern sector workforce. In this environment, laissez faire

³Interestingly, this perspective is confirmed by the empirical analysis of Canadian inter-industry
Levine and Zervos (1993) examine the robustness of cross-sectional regressions predicting income growth - notably they cannot find a negative association between inflation and growth. Implies a suboptimal level of labour mobility, because workers in the traditional sector demand a wage premium before they will move to the modern sector, and individuals have no reason to recognize the dynamic effects on labour productivity of the growth of the modern sector as a whole. In this model, "competitive wage premia therefore operate as a tax on learning formation in the modern sector" (1993:568). Wage compression through a solidarity wage policy is part of the answer, combined with a policy of setting the common wage at a high enough level to displace labour from the traditional sector, and an active labour market policy to speed the transition of labour between sectors. As they put it "by punishing expansive sectors of the economy, competitive wage premia impose an unwarranted drag on the rate of structural change. ...Egalitarian pay compression, combined with active labour market policies, works in the same way as an industrial policy of subsidizing sunrise industries" (1993:559).

A common feature of all the above articles is that they use best practice neoclassical methodology and have appeared in the mainline journals of the economics profession. They represent only a sampling from the recent endogenous growth literature, but hopefully it has been a sufficiently large sample to indicate the wide range of theories and empirical data now used to link the inequality of income distribution with the rate of macroeconomic growth. This literature has often combined both theory and empirical work (often using cross-sectional data on different nations) but to economists raised in the era of the "equity/efficiency tradeoff" one of the most notable features of the theoretical literature is that it often has the tone of "now that we all know that greater equality and faster growth go together, this article presents a theoretical mechanism

---

4Levine and Zervos (1993) examine the robustness of cross-sectional regressions predicting income growth - notably they cannot find a negative association between inflation and growth.
to explain why”. Eckstein and Zilcha, for example, refer to the evidence of Persson and Tabellini that "growth rates are positively associated with more equality in a cross-section of nations. This evidence is consistent with their model of income distribution and growth as well as with most other recent papers on this subject" (1994:342-emphasis added). This new orthodoxy represents quite a change from the tradeoff clichés of the 1970s. How can one explain the change?

2. The Demise of the Equity/Efficiency Tradeoff

2.1 An Awareness of Intergenerational Transmission

The "trade-off" perspective recognized, from the first, two important exceptions: (1) more of both equality and income could be obtained if macroeconomic policy moved the economy to its full employment output potential and (2) policies to improve equality of opportunity could produce both more equality and a higher average income. However, in the 1970s there was a tendency to distinguish between "equality of opportunity" and "equality of result" and to argue that although policies to improve equality of opportunity might increase average incomes, redistributive policies to equalize economic outcomes had efficiency costs.

However, the distinction between equality of opportunity and equality of result only makes sense in a single generation model. If parents make decisions about the human capital, and other assets, which they bequeath to their children within the constraint of their own lifetime income, then the inequality of economic outcome within one generation predetermines the degree of inequality of opportunity within the next generation. In the early 1980's an emerging literature
within the human capital tradition explored the implications of parental bequest for the opportunities of children.\textsuperscript{5}

Around the same time, the importance of inherited wealth in the total capital stock became widely recognized,\textsuperscript{6} and it was necessary to find a theory which explained (a) the large aggregate size of bequests and (b) the fact that most households receive very little financial wealth by inheritance. The simple life cycle savings model had predicted that individuals would try to build up a stock of assets while working which they would consume

\textsuperscript{5}See Becker and Tomes (1979); Osberg (1984: pp. 195-217) surveys the literature.

\textsuperscript{6}Kotlikoff and Summers (1981), White (1978) and Brittain (1978).
while retired, but even though this model does predict that some bequests will occur (i.e. the unanticipated bequests of those who die before their expected time) it could not credibly predict the size or distribution of wealth holdings. Adding a bequest motive to savings, by presuming that parents care about the well being of their children, solved this problem. In the augmented life cycle savings/bequest model, parents face a budget constraint of their total lifetime resources (total earnings plus any inheritances or gifts received) which can be allocated either to their own consumption or to their gifts and bequests to their children. Diagrammatically, the top panel of Figure 2 represents the choices open to two sets of parents, when their lifetime consumption of is measured as $C$ (the vertical axis) and their total bequest is plotted as $B$ (the horizontal axis). Each generation is presumed to derive utility from their own consumption and from the well-being of the next generation. A family with more resources (i.e. own earnings $Y$ plus the bequest of the previous generation $B$) can attain a utility level of $U$, while a poorer family can reach only $U$. 

In saying that the present value of inheritances and gifts received plus lifetime earnings equals the discounted value of life time consumption plus bequests and gifts given, one is only really stating an accounting identity. However, the argument acquires predictive power if one presumes (as in the lower panel of Figure 2) that there are diminishing marginal returns to human capital, but that financial markets offer a fixed return. If so, it will be optimal to invest in human capital up to the point ($K$ in Figure 2) where the marginal return on human capital is equal to the rate of interest obtainable in financial markets. If the total bequest of a poor family is $b$, it is optimal for that family to leave a bequest entirely in the form of human capital. Rich families tend to pass on both human capital and financial wealth, as in bequest $b$. 
This model can explain the existence and concentration of inheritances, and it is also a model of unequal opportunity and inefficiency in human capital acquisition. Seen from the perspective of the children of the poor family, the bequest process generates a less than optimal amount of human capital ($b_0 < K$). The divergence in marginal rates of return ($r_0 > r_1$) between human and financial capital means that total output would increase if the poorer kids had access to human capital at the same rate of interest as the richer kids.

Figure 2 draws a picture of the implications of inequality of outcome in one generation for the inequality of opportunity and inefficiency of allocation of human capital of the next generation. Endogenous growth theorists have added three dimensions to this basic framework: (1) an explanation of why private credit markets cannot be relied on to solve the problem of underinvestment in the human capital of poor children; (2) a discussion of why the marginal productivity of human capital investment may be lower in poor neighbourhoods than in rich neighbourhoods (Durlauf, Berabou), thereby exacerbating the missallocation pictured in Figure 2 and (3) a discussion of the externalities of human capital acquisition which emphasizes the importance of human capital for long run growth, and not just for the level of current incomes, implying that sub-optimal investment in human capital can have large and long run consequences.

With hindsight, it can be seen that the economics literature was ready for endogenous growth models. Throughout the early 1980’s intergenerational/overlapping generations models of economic decision making became widely fashionable. Initially, overlapping generations models of the macro-economy often referred to an economy of “representative agents” in which, by construction, there can be no inequality and the growth rate converges to a unique steady state. Once, however, the heterogeneity of agents was recognized, the door was open to multiple
growth equilibria determined, in part, by the initial distribution of characteristics. Unequal outcomes in one generation then predetermine, if there is no state intervention, the inequalities of opportunity of the next generation - and the rate of growth of average incomes.

2.2 The Implications of a Progressive Income Tax

A major part of the "tradeoffs" story has always been its critique of redistributive taxation -- in particular, the progressive income tax. By driving a wedge between the price which firms are willing to pay for labour and the net wage which individuals will receive, the income tax distorts the incentives presented by market prices to the optimal allocation of resources, particularly the optimal choice of hours of work and hours of leisure. In the crude version of the "tradeoffs" critique, redistributive expenditures financed from income tax revenue necessitated "too high" a rate of income tax, which did not provide "enough" incentive for the supply of labour, thereby decreasing aggregate output. More sophisticated tradeoff critics emphasize the idea that the increasing marginal tax rates of a progressive income tax system can create a dead-weight loss in consumer surplus, whose size depends on the compensated wage elasticity of labour supply and the level of marginal tax rates.

The crude incentives argument draws its strength from the simple labour supply function, which generations of introductory economics instructors have drawn sloping gradually upwards to the right. Since an income tax decreases the net wage rate, it is therefore seen as decreasing aggregate labour supply. A weakness in this argument is that, as intermediate economics courses argue, any change in net wage rates has both income and substitution effects. Most labour economists have always recognized that at some wage rate the income effect must begin to
dominate the substitution effect, -- i.e. the labour supply function cannot continually slope upward
to the right, since if it did one would eventually run out of leisure time entirely. The neoclassical
model of labour supply predicts that labour supply curves must eventually bend backwards, if only
because the marginal utility of leisure increases as hours of work approach total time available.

Once it is recognized that labour supply functions must contain a backward bending
segment, the important empirical issue is the proportion of the population who will increase their
labour supply as their net wage rate decreases. Figure 3 illustrates the point, using a plot of the
Since the point of inflection of the female labour supply curve in Canada was at $13.09 per hour
in 1986, a substantial fraction of Canadian female workers are in the range of their labour supply
function in which increases in the hourly wage produce decreases in desired hours of labour
supply. For Canadian men, the desired hours of labour supply function is very inelastic, with a
slight backward slope throughout, as Figure 4 illustrates.

The consensus of a very large labour supply literature is that the wage elasticity of desired
labour supply is very small -- see Pencavel (1986), Killingsworth (1983), Osberg (1986), Phipps
(1993), or Heckman (1993) -- a common "best guess" from the literature is that the wage
elasticity of desired labour supply is about 0.1. However, even if desired hours of labour supply
did increase substantially with increases in the net wage, the increased incentives to desired labour
supply do not matter much to individuals who already cannot obtain all the hours of work which
they want. An increasingly large empirical literature (e.g. Kahn and Lang, 1988; Osberg and
Phipps, 1993) has emphasized the prevalence of constraints on available hours of work for many
workers -- even in the relatively buoyant labour markets of the mid 1980s.
There is not, therefore, much to the story that decreasing the marginal tax rate in the income tax system would substantially increase the aggregate supply of labour, but the more sophisticated exponents of the tradeoff perspective argued that even if the uncompensated wage elasticity of labour supply is rather small, it is the compensated wage elasticity of labour supply which matters for the calculation of the dead-weight loss in allocative efficiency imposed by the income tax system. The Slutsky equation implies that the uncompensated wage elasticity of labour supply (an observable) is equal to the compensated wage elasticity (an unobservable) plus the income elasticity of labour supply (an observable). Estimates of the compensated wage elasticity of labour supply are therefore always found by taking the difference of two observed estimates.

For example, Hausman (1983) got a rather large estimate of the compensated wage elasticity of labour supply, which was used to justify a large estimate of the dead-weight loss implied by the progressive income tax system. Since Hausman estimated the uncompensated wage elasticity of labour supply to be approximately zero, it was his unusually large estimate of the income elasticity of labour supply which mattered. Even at the time, Heckman (1983) argued that the income elasticity estimates were rather suspect, both because they depended on a priori constraints imposed in estimation and on data of questionable validity. Heckman (1993:118) summarizes the current thinking rather trenchantly: "These schemes have now fallen into disrepute. The econometric procedures used to produce the estimates were econometrically and economically inconsistent in part because it did not properly correct for the missing wage data for non-workers (see Heckman, 1983). Competent analysts have been unable to replicate the earlier findings even using the same data (see MaCurdy et al, 1990). When these models are re-
estimated using more robust schemes, weak wage and income effects of taxes are found for males in numerous countries."

Furthermore, Triest (1990) has pointed out that "when benefits are measured in terms of current consumer prices, then the marginal cost of public funds depends only on uncompensated price elasticities" (1990:564). Normally, evaluation of the costs and benefits of changes in government expenditure are in terms of current prices -- indeed, since lump sum non-distortionary taxes are a figment of the economist's imagination, the price vector which corresponds to such taxation is as well. As Triest points out, when the uncompensated wage elasticity of labour supply is negative, the marginal cost of public funds is less than unity, since an increase in the tax rate causes an increase in labour supply, leading to an indirect increase in government revenue.

To summarize, since the mid 1980s it has become more widely recognized that elasticities of labour supply are "closer to zero than one" (Heckman, 1993:118), and consensus estimates of the allocative efficiency costs of income taxation have become rather small. During the late 1970s and early 1980s when large estimates of the dead-weight loss in allocative efficiency due to progressive taxation were commonplace, redistributive expenditure programs to improve equality of opportunity were seen as having a relatively large cost in allocative efficiency. However, Stiglitz (1988:465), for example, now summarizes the dead-weight loss associated with the standard U.S. income tax rate as less than 2%, if the best estimate of labour supply elasticities is used, and a maximum of 12% of revenue raised, if the highest available (but now discredited) estimate of compensated labour supply elasticity is used. Both estimates assume a positive wage/hours relationship, but as Figures 3 and 4 indicate, the labour supply curve may be backward bending, for many groups of workers (implying that higher taxes on upper income
groups increase their labour supply). Although there continues to be a huge debate on the optimal structure of taxation, it is now framed in the realization that the efficiency costs of labour supply responses to higher taxation levels are much less than was imagined in the late 1970s.

### 2.3 Static and Dynamic Efficiency

The assertion of a tradeoff between "equity" and "efficiency" is little more than an attitude of mind, until one defines "equity" and "efficiency". In the public finance literature, "equity" has a number of possible definitions. Head (1993) distinguishes between horizontal equity in the sense of similar individuals being treated in a similar fashion, vertical equity in the sense of taxation in accordance with ability to pay and the "benefit principle" of equity, that taxpayers should pay for public services in the same proportion that they use them. Osberg (1993) adds that public policy must also consider intergenerational equity and the legal principles of procedural equity.

It is unlikely that all these concepts of equity ever implied that greater equity means less efficiency. And in practice, the "tradeoff" perspective did not really mean "equity", but "equality", and had a very simplistic conception of "equality". Among many others, Osberg (1984) and Jenkins (1991) have emphasized that the measurement of inequality depends crucially on what concept of income is being measured (e.g. lifetime or annual), since lifetime income is usually more equally distributed than annual income. Measurement also depends on what population is being examined (individuals or households) since income pooling within households means that household income is more equally distributed than personal income. International comparisons of inequality are affected because countries differ in average household size, and the percentage of single person and multiple earner households.
Measurement of inequality also depends on what measure of dispersion (Gini, Theil, or Atkinson Index, coefficient of variation or decile shares) is being used. Some measures (like the coefficient of variation) are more sensitive to the top end of the distribution and others (like the Gini Index) are more sensitive to differences in the middle-range. Since Lorenz curves often cross, unambiguous statements about relative inequality are often not possible and the choice of inequality index matters. It can make a great deal of difference whether one is considering inequality in the dispersion of individual earnings, or inequality in the consumption possibilities of households, because transfer payments and demographic change can maintain stability in the distribution of household income, despite rising inequality in the distribution of individual earnings -- as happened in Canada in the 1980s. Finally, it matters whether one wants to emphasize the inequality between the middle class and the rich or between the poor and the middle class because some of the phenomena one is trying to address -- such as the formation of poverty ghettos -- are really about particular segments of the income distribution, not the distribution as a whole. However, although the proponents of the "tradeoff" perspective of the 1970s were lamentably imprecise in specifying what exactly they meant by "inequality", the endogenous growth theorists of the 1990s are no better.

The definition of "efficiency" is more important to the changing perspectives of economists, because the 1970s and 1990s discussions are concerned with differing conceptions of "efficiency". The "tradeoff" perspective was concerned with movement along a given production possibility frontier. It argued that the pursuit of greater equality in the distribution of income comes at the expense of a lower average income. In Okun's famous "leaky bucket" analogy, there is a dead-weight loss in redistributional transfers from rich to poor, both because of the direct
administrative cost of transfer programmes and because of the distortions which redistributional policies build into the signals of the price mechanism (the dead-weight loss produced by labour supply responses to progressive income tax system was seen as especially important -- see Section 2.2).

In Figure 5, points A and B represent alternative combinations of average income and income equality open to a given society, at a given point in time, $T_0$. The assumption that there is an empirical tradeoff between average incomes and income equality was often combined with the assumption that the preferences of society could be expressed in a "social welfare function" in which average incomes and income equality appear as arguments. The social cost of choosing "too much equality" could then be represented as the choice of point B in Figure 2 rather than the optimum A. At point A, society is at its maximum possible level of well-being ($SWF_1$) while at point B society is on a lower indifference curve ($SWF_0$). The social cost of "too much equality" can be represented as the distance $(Y_1 - Y_0)$, or the amount of the increase in average income which would be required to move society to a level of well-being equivalent to its highest possible level of well-being, holding equality constant.

It is useful to distinguish between "allocative efficiency" in the choice of the welfare maximizing combination of desired goods, given that more of one good can only be achieved at the expense of less of another good, and "productive efficiency" in the sense of producing the maximum possible amount of goods from the inputs available at a particular time. Okun is, after all, probably more famous as a popularizer of the "Okun's Law" estimation of the output gap, (i.e., the difference between actual output and potential output) than as the originator of the trade-off cliché. Excessively high unemployment due to misguided macroeconomic policy, or
technical inefficiency in production at the firm level, can mean that a society at point C could potentially have more of all desired goods.

Both "allocative" and "productive" efficiency refer to the production and consumption possibilities open to society at a particular point in time. In terms of Figure 5, productive efficiency is maximized by moving to the production possibility frontier, while allocative efficiency is maximized by moving along the productive possibility frontier to the point of maximum social welfare -- but in both cases the production possibility frontier is thought of as being the maximum possible amount of output that can be achieved from the inputs of capital, labour and technology that are available at a given point in time.

The focus of the endogenous growth literature is on "dynamic efficiency", in the sense of maximizing the rate of growth of productive possibilities over time. In Figure 5, the connection between inequality and growth is represented by the fact that the production possibility frontier shifts out at a different rate, for differing initial choices of economic equality. Given that small differences in growth rates cumulate within a generation or two to large differences in average living standards, society can have both more equality and a higher average standard of living, if equality encourages growth. The deadweight loss of the income tax system may be in the 2% range - but eliminating this impediment to allocative efficiency would produce a one time gain in income - improving growth prospects from 2% per annum to 2.5% would compound to a difference of 15% of original income after 20 years, 48% of original income after 40 years, 112% after 60 years, etc.

2.4 Social Pressures
A social science such as economics evolves because of both the internal logic of the development of theoretical arguments and the sequence of discovery of empirical regularities and because of the social pressures to which economists are sensitive. Economists' perceptions of "reasonable" assumptions to make or "interesting" questions to examine change as social conditions change. Part of the reason why the growth theorists of the 1990s have diametrically opposite expectations about equality/efficiency compared to the attitudes prevalent in the 1970s can surely be found in the economic history of the intervening period.

In the early 1970s, it was still possible to take growth for granted in much of the developed world. It was not yet apparent that the 1973/74 period would be seen as a watershed separating the steady growth, relatively low unemployment and stable inflation years of the 1950s and 1960s from the stagnation in real earnings and high unemployment which has characterized the last twenty years. In the early 1970s the United States was, moreover, just beginning to digest the lessons of the expansion of social expenditures of "the war on poverty" of the 1960s. It was clear that although some programs had been successful, others had been wasteful. At that point in time, it seemed reasonable to concentrate attention on the "leaky bucket" of redistributive social expenditures and the allocative inefficiencies of redistribution of a given stock of resources, (since total output was thought to be growing at an exogenously given rate).

In the 1990s, North Americans look back on approximately twenty years of stagnation in average real hourly wages and have serious worries about the previously unthinkable -- that future generations may face declining living standards. It is clear that growth in output, employment and living standards cannot simply be assumed to occur exogenously, and it is equally clear that some other regions of the world have continued to grow dramatically.
The interest in endogenous growth models is driven in part by the slowing of economic growth in North America, compared to previous periods and compared to other countries. The old neo-classical growth theory never could explain much of total growth with reference to increments in the capital and labour stock. Hammond and Rodriguez-Clare (1993:420) note that the "Solow residual" in traditional neo-classical models of economic growth can account for approximately 7/8 of aggregate growth. Hence, there is real social pressure to come up with a new explanation of economic growth, which can generate differing long-run rates of economic growth and is consistent with the "stylized facts" of slow growth and high growth economies in the 1980s.

One of the factors which makes a presumption of a positive association between equality and growth credible in the 1990s is the fact the U.S.A. tried the inequality option so spectacularly in the 1980s, with so little payoff in average incomes. The Reagan administration was explicitly anti-egalitarian in its rhetoric and the promise was that increased incentives and decreased government regulation would produce rapid economic growth whose benefits would "trickle down" to the working class. Both in the statistical evidence and in the day-to-day reality of urban life, academics can see the social costs and the increased poverty which this strategy produced.

At the same time, some economies have grown spectacularly throughout the 1980s. As Fields (1995:93) notes: "The newly industrializing economies of East Asia (Hong Kong, Republic of Korea, Singapore, and Taipei, China) have had sustained records of full employment, improving job mix, rising real wages, falling absolute poverty, and low to moderate levels of income inequality. Their rates of improvement are the envy of the rest of the world."
Poverty and inequality affect the growth rate through more avenues than just human capital acquisition. A considerable literature has examined the links between poverty and health, crime and inner city decay - all of which are extremely costly. These social costs of poverty and inequality could also be used to motivate a focus on the link between equality and growth, but it appears that such links are, as of now, not explicitly drawn - the bibliographies of endogenous growth articles do not contain references to the sociological or social work literature on the social costs of poverty and inequality. This literature could, however, easily be used to justify an additional argument for the equality/growth link.

A large part of what the state does is guarantee property rights and maintain law and order. If crime is higher in a more unequal society, the state will have to spend more on police and prisons and the private sector will also spend more on security services. These activities are best viewed as "defensive necessities" rather than productive investment or enjoyable consumption. Employing an ever larger percentage of the labour force in these activities will necessarily crowd out expenditure on new investment in physical or human capital. Similarly, expenditure on health care can be viewed as a defensive necessity to maintain the labour force at a given level of productivity. The public health literature has increasingly recognized a link between economic inequality, poverty and illness. The costs of the health care to deal with such illness partially crowd out investment. The endogenous growth literature reminds us that such decreased investment has long run implications for the growth rate. Societies with a sicker population and local governments which have to spend more on penitentiaries than on schools pay now, in higher taxes, and they also pay later, in lower growth.

Surveyed in Osberg, 1990.
3. **Conclusion**

Some of the endogenous growth literature is purely theoretical, and it takes a leap of faith to generalize from such highly stylized arguments\(^8\) to the real world. Some (e.g. Hum, 1993) would argue that such stylized discourse is simply the rhetoric in which academic economists must frame their arguments, if they want to be taken seriously by their peers. The interesting point, however, is that the language of individual utility maximization and the grammar of mathematics is now being used to argue that more equal societies grow faster.

One must also underline that the empirical results of the endogenous growth literature on the equality/growth correlation are those observed within the historical range of observation. Nobody in this literature is arguing that one can extrapolate outside this range to extreme conclusions -- that, for example, absolute equalization of all incomes would produce infinite growth. All the authors cited above recognize that zero inequality would mean zero incentives -- but they also recognize that the relevant issue for real societies is to make choices about marginal changes in inequality, from where they are now. In the context of Canadian society, the important fact is that Canada is not a particularly egalitarian society, as measured by the summary statistics of inequality among developed nations (see Smeeding, 1991:45).

What are we to conclude from the fact that the new endogenous growth economics literature argues that more equal societies grow faster, other things equal, and the fact that Canada has substantial room for improvement in equality (and in growth)?

Keynes concluded his *General Theory* with the observation:

---

\(^8\)Galor and Zeira (1993), for example, discusses a one good world with two technical production possibilities and two types of labour.
... the ideas of economists and political philosophers, both when they are right and when they are wrong, are more powerful than is commonly understood. Indeed the world is ruled by little else. Practical men, who believe themselves to be quite exempt from any intellectual influences, are usually the slaves of some defunct economist. Madmen in authority, who hear voices in the air, are distilling their frenzy from some academic scribbler of a few years back. I am sure that the power of vested interests is vastly exaggerated compared with the gradual encroachment of ideas. Not, indeed, immediately, but after a certain interval; for in the field of economic and political philosophy there are not many who are influenced by new theories after they are twenty-five or thirty years of age, so that the ideas which civil servants and politicians and even agitators apply to current events are not likely to be the newest. But, soon or late, it is ideas, not vested interests, which are dangerous for good or evil. (1964:383)

I think that the "trade-off" analysis of the 1970s had an influence on the economic policy of the 1980s and continues to influence policy in the 1990's. If Keynes is right, the endogenous growth theorists of the 1990's will also eventually influence economic policy - but we will probably have to wait a while in Canada, since the current generation of policy makers are deeply entrenched and not easily influenced by outside ideas.

The economic policy-makers of the 1980s inherited from the 1970s a set of economic ideas, and a set of social institutions, as well as a series of economic problems. In dealing with the problem that has dominated macroeconomic policy -- inflation -- the intellectual framework that perceived a tradeoff between equity and efficiency also perceived no long-run cost to higher unemployment, since the NAIRU (non-accelerating inflation rate of unemployment) was seen as stable. Hence, greater inequality produced by higher unemployment was seen as the unfortunate but necessary price to pay for getting inflation under control and for improving macroeconomic efficiency in the long-run. These ideas remain influential in the 1990s, as does the perception that
the social institutions established during the 1960s and 1970s (such as an extensive transfer
system or public involvement in post-secondary education) may serve desirable equity goals but
are, generally speaking, impediments to greater economic efficiency.

In the 1980s, macroeconomic policy in Canada deserted its historic concern with
unemployment and focussed on the control of inflation, but structural reforms to social policy
were much more limited. As a result, the social institutions established earlier did more or less
what they were initially designed to do. Public subsidies kept tuition fees low and maintained the
accessibility of the Canadian post-secondary education system, with the result that the supply of
university graduates increased substantially. Although the 1980s saw an increase in earnings
inequality in Canada, unlike the United States there was no trend to a greater wage premium for
university graduates, because greater supply of university graduates kept up with increased
demand. In Canada, increased earnings inequality has been driven by increased inequality in hours
worked per year, not by greater inequality in hourly wages (see Morissette, Myles and Picot,
1993). Furthermore, the availability of unemployment insurance payments to those affected by
the rising unemployment of the 1980s substantially limited the increase in income inequality (see
Osberg, Erksoy and Phipps, 1995). Fritzell (1992) argues that Canada largely escaped the 1980s
trend to increased inequality of total family income observed in other developed countries.

If the endogenous growth literature is correct, limiting the domain of inequality in these
ways has a long-run positive impact on economic growth. However, policy-makers who retain

---

9For example, Audas (1994) using the Labour Market Activity Survey, found that the
probability that youth would drop out of high school was higher if one or both of the parents had
experienced unemployment in the past year, but that these effects were somewhat offset if
someone in the household had received UI benefits. The positive impact of family UI benefits on
youth school attendance can be explained as an "income effect" -- by partially maintaining income
the equity/efficiency tradeoff perspective of the 1970s continue to see these programs as fulfilling equity objectives that now have to give way to the need to improve economic efficiency -- as "nice things that we can no longer afford". As a result, the social programs that maintained a degree of equality of opportunity and of outcome in Canada are under concerted attack.

In the larger society, perspectives are changing. When a publication like Business Week devotes a feature article (August, 1994) to the adverse impacts of inequality on U.S. growth, it is clear that the endogenous growth literature has somewhat penetrated the world outside academia. As well, the fiscal crisis of the modern welfare state impels many governments to look for new sources of additional revenue. As the IMF has noted, there was a world-wide trend to decreasing the progressivity of income tax rates in the early 1980's, spurred on partly by the hope that such tax cuts would increase growth, but the results have been disappointing in almost all countries (see IMF, 1994). Many countries are now considering and implementing higher marginal rates of tax. The idea of a tax on inherited wealth has also been discussed in Canada in recent years (see Davies and Duff, 1994). If set at a level comparable to U.S. inheritance taxes (which yield about 0.29% of GDP), such a tax would yield about $2.2 billion per year (which happens to be approximately the amount the federal government now spends on post-secondary education, but is about to cut).

In short, a changed political climate and a need for additional revenue could swing the pendulum of policy back to a more egalitarian focus - but if cuts to post-secondary education financing produce higher tuition fees and lower accessibility while cuts to transfer programmes in poorer families, UI benefits reduced the necessity for youth to quit school and try to get a job to augment family income.
widen the inequality of family incomes, Canada may also see greater inequality of opportunity and of outcome. If the perspectives of the 1970s are as influential in influencing social policy in the 1990s as they were in dominating macroeconomic policy in the 1980s, Canada will pay the price for many years, in slower growth, of the neglect of the importance of economic equality.
BIBLIOGRAPHY


Figure 3
Female Labour Supply Function
Ages 25 - 54, 1986 LMAS

Annual Hours of Supply

Hourly Wages

$0.00 $5.00 $10.00 $15.00 $20.00 $25.00 $30.00 $35.00 $40.00

$0.00 $5.00 $10.00 $15.00 $20.00 $25.00 $30.00 $35.00 $40.00

300 400 500 600 700 800 900
Figure 4
Male Labour Supply Function
Ages 25 - 54, 1986 LMAS

Hourly Wages

Annual Hours of Supply
Social Welfare Function (SWF)

\[ \text{SWF} = F(\text{average income, income equality}) \]