Chapter 11

THE DISTRIBUTION OF WEALTH *

JAMES B. DAVIESa and ANTHONY F. SHORROCKSB

aDepartment of Economics, University of Western Ontario, London N6A 5C2, Canada; bDepartment of Economics, University of Essex, Colchester CO4 3SQ, UK and Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE, UK

Contents
1. Introduction 606
2. Theoretical approaches 608
   2.1. Simple models of wealth distribution 609
   2.2. Lifecycle accumulation 612
   2.3. Intergenerational issues 621
   2.4. Self-made fortunes 627
3. Empirical evidence on wealth inequality 628
   3.1. Household surveys 629
   3.2. Wealth tax data 635
   3.3. Estate multiplier estimates 637
   3.4. The investment income method 642
   3.5. Direct wealth estimates for named persons 642
   3.6. Portfolio composition 643
   3.7. Explaining national differences and changes over time 645
4. Applied work on the determinants of wealth distribution 648
   4.1. Simulation studies 648
   4.2. Intergenerational wealth mobility 651
   4.3. The aggregate importance of inherited versus lifecycle wealth 653
   4.4. Empirical studies of transfer behaviour 657
5. Summary and conclusions 663
References 667

* We should like to thank Andrea Brandolini, John Flemming, John Piggott, Edward Wolff and, most especially, Tony Atkinson and François Bourguignon, for their advice, comments and assistance. The first author also thanks the Social Sciences and Humanities Research Council of Canada for its support.

Handbook of Income Distribution: Volume 1. Edited by A. B. Atkinson and F. Bourguignon
© 1999 Elsevier Science B.V. All rights reserved.
Abstract

This chapter is concerned with the distribution of personal wealth, which usually refers to the material assets that can be sold in the marketplace, although on occasion pension rights are also included. We summarise the available evidence on wealth distribution for a number of countries. This confirms the well known fact that wealth is more unequally distributed than income, and points to a long term downward trend in wealth inequality over most of the twentieth century. We also review the various theories that help account for these features. Lifecycle accumulation is one popular explanation of wealth differences, but inheritance is also widely recognised as playing a major role, especially at the upper end of the wealth range. A recurrent theme in work on wealth distribution is the relative importance of these two sources of wealth differences. We discuss the results of studies that assess the contributions of inheritance and lifecycle factors, and give attention also to a variety of related issues, such as the link between wealth status across generations, and the possible motives for leaving bequests.

Keywords: Wealth distribution, wealth inequality, portfolio composition, lifecycle saving, inheritance, bequests, intergenerational transfers

JEL codes. main: C81, D31, D91, E21, G11; secondary: D11, D12, D63, D64, J14)

1. Introduction

This chapter surveys what is known about the distribution of personal wealth and its evolution over time. We review the descriptive evidence as well as theoretical and applied research that attempts to explain the main features of wealth-holdings and wealth inequality observed in the real world.

There are many reasons for interest in personal wealth, and many ways in which the concept of wealth may be defined. If we were concerned with the overall distribution of economic well-being or resources, it would be appropriate to examine the distribution of “total wealth”, that is, human plus nonhuman capital. But that is not our objective here. Instead, we exclude the human capital component and focus on material assets in the form of real property and financial claims. The term “wealth” will therefore usually refer to “net worth”—the value of nonhuman assets minus debts. Our aim is to examine the reasons for holding wealth, to document the observed differences in holdings across individuals and families, and to examine the causes of these observed differences.

One major reason for interest in wealth-holdings is that, unlike human capital, most real property and financial assets can be readily bought and sold. This allows nonhuman wealth to be used for consumption smoothing in periods when consumption is expected to be high (growing families) or income is expected to be low (retirement), and in periods of unanticipated shocks to either income or expenditure (the precautionary motive for
This consumption smoothing role is particularly important when individuals face capital market imperfections or borrowing constraints. Wealth may also be accumulated, or retained, for the purpose of making bequests. Additional noneconomic reasons for studying wealth include the power or social status which may be associated with certain types of assets such as privately-owned businesses. The pattern of wealth-holdings across individuals, families, and subgroups of the population, is therefore capable of revealing a great deal about both the type of economy in which people operate, and the kind of society in which they live.

The concept of net worth may appear to be straightforward, but should we deal with intangible assets which cannot be readily bought and sold? This category covers pension rights, life insurance, and entitlement to future government transfers (including "social security wealth"). Any attempt to include the rights to uncertain future benefits has to confront a variety of difficult valuation problems. For example, it is not obvious what discount rates should be used for these assets. Should they be risk-adjusted? Should a special adjustment be made for people who are borrowing constrained? Satisfactory answers to these questions require a considerable amount of painstaking work. It is therefore not surprising to discover that most applied work on wealth-holdings and wealth distribution confines itself to marketable wealth. When reviewing the empirical evidence, we use the term "augmented wealth" to refer to the broader concept which includes entitlements to future pension streams.

There are certain important "stylized facts" about the distribution of wealth which it is useful to highlight at the outset. These are:

1. Wealth is distributed less equally than labour income, total money income or consumption expenditure. While Gini coefficients in developed countries typically range between about 0.3 and 0.4 for income, they vary from about 0.5 to 0.9 for wealth. Other indicators reveal a similar picture. The estimated share of wealth held by the top 1% of individuals or families varies from about 15–35%, for example, whereas their income share is usually less than 10%.

2. Financial assets are less equally distributed than nonfinancial assets, at least when owner-occupied housing is the major component of nonfinancial assets. However, in countries where land value is especially important, the reverse may be true.

3. The distribution of inherited wealth is much more unequal than that of wealth in general.

4. In all age groups there is typically a group of individuals and families with very low net worth, and in a number of countries, including the US, the majority have surprisingly low financial assets at all ages.

5. Wealth inequality has, on the whole, trended downwards in the twentieth century, although there have been interruptions and reversals, for example in the US where wealth inequality has increased since the mid 1970s.

Possible explanations for these, and other, stylized facts will be investigated in this chapter. Section 2 begins with a review of simple models that try to account for the overall shape of the distribution of wealth. We also show how the accounting identity relating
James B. Davies and Anthony F. Shorrocks

Changes in wealth to earnings, rates of return, consumption and capital transfers provides a useful framework for investigating the proximate determinants of wealth inequality and its trend over time. Attention then turns to more detailed models of wealth-holdings, including numerous versions of the lifecycle saving model, and a variety of models concerned with bequest behaviour and the distributional impact of inheritance.

Section 3 reviews the descriptive evidence on wealth distribution in a number of countries, and discusses possible explanations for national differences and trends over time. The roles of asset prices, inheritance taxation, and other factors are examined. Then, in Section 4 we look at applied research which attempts to assess the contribution of different factors to the distribution and evolution of wealth-holding. These studies address a variety of questions, and use a variety of approaches, including decomposition procedures, simulation exercises, and conventional econometrics. But, in broad terms, all the research is linked by a common objective: to cast light on the reasons for wealth accumulation, and on the importance of inheritance vis-à-vis lifecycle saving as determinants of the level and distribution of wealth.

2. Theoretical approaches

Many different types of theories have been used to model aspects of wealth-holdings and wealth distribution. To a large extent they reflect the empirical evidence available at the time the theories were proposed. Prior to the 1960s, data on wealth were obtained primarily from estate tax and wealth tax records, with other evidence pieced together from small unrepresentative surveys and a variety of other sources. These tended to confirm the widely held view that wealth was distributed very unequally, and that material inheritance was both a major cause of wealth differences and an important vehicle for the transmission of wealth status between generations. There were also grounds for believing that wealth inequality was declining over time, and that the shape of the distributions exhibited certain statistical regularities which could not have arisen by coincidence. Early theoretical work on wealth distribution sought to explain these statistical regularities, and to understand the interplay of basic forces that could account for high wealth concentration and a declining trend over time.

More recently, research has shifted away from a concern with the overall distributional characteristics, focusing instead on the causes of individual differences in wealth-holdings. The change of emphasis was prompted in part by a recognition of the increasing importance of saving for retirement, and is reflected in the central role now assigned to the lifecycle saving model formulated by Modigliani and Brumberg (1954) and Ando and Modigliani (1963). The second major development has been the growth in the availability and sophistication of micro-data sets that offer not only estimates of the savings and asset holdings of individuals but also a range of other personal and household characteristics than can help account for differences in wealth.
This section reviews various models of wealth distribution and wealth differences, beginning with some of the early attempts to explain the overall distributional features and the evolution of wealth inequality. Section 2.2 discusses the simple lifecycle model and some of its many extensions, drawing out the implications for wealth accumulation over the course of the lifetime. Attention here is restricted to pure intra-generational models in which inheritances play no role. For the most part, Section 2.3 takes the opposite tack, suppressing interest in lifecycle variations and focusing instead on intergenerational links, especially those concerned with the motivation for bequests and the impact of inheritance.

The distinction between models of wealth accumulated for lifecycle purposes and models concentrating on the intergenerational connections reflects the theoretical literature on wealth distribution: very few contributions have attempted to deal simultaneously with both the lifecycle and inherited components of wealth. This is a major weakness of past theoretical work on wealth. While there may be some value in modelling lifecycle wealth in the absence of inheritance, it should be recognised that this exercise does not reveal the true pattern of lifecycle accumulation in the real world—since the real world also has inheritances. This rather obvious point is one which is often forgotten, particularly when attempts are made to assess the relative importance of accumulation and inheritance.

Current models of accumulation and bequests do not appear to capture the circumstances and motives of those who amass large fortunes in the course of their lifetimes. We look briefly at this issue in Section 2.4. Throughout, reference is made to relevant empirical evidence that has informed and influenced the development of the theories. Later, in Section 4, we consider in more detail some of the issues that have received most attention.

2.1. Simple models of wealth distribution

Early empirical work on personal wealth-holdings established two enduring features of the shape of the distribution of wealth. First, it is positively skewed, unlike the normal distribution, but roughly resembling a normal distribution when wealth is replaced by the logarithm of wealth. Second, the top tail is well approximated by a Pareto distribution, which yields a straight line graph when the logarithm of the number of persons with wealth above \( w \) is plotted against \( \log w \). These two statistical regularities were observed not only for wealth-holdings, but also for many other skewed distributions such as those for incomes, the turnover of firms, and the size of cities (see, for example, Steindl, 1965, Appendix B).

The fact that these size distributions are approximated by the lognormal suggests that, by appealing to the Central Limit Theorem, they can be generated by a random walk of the form:

\[
\ln W_t = \ln W_{t-1} + u_t, \tag{2.1}
\]
where $W_t$ is wealth at time $t$, and $u_t$ is a stochastic component (Aitchison and Brown, 1957). Equation (2.1) contains the implicit assumption that the factors influencing changes in wealth over time operate in a multiplicative fashion, rather than additively. This became known as Gibrat's "law of proportionate effect", following Gibrat (1931).

As a model of wealth distribution, the simple random walk Eq. (2.1) has one major technical disadvantage: it predicts that wealth dispersion, as measured by variance of log wealth, increases over time without bound. It cannot therefore apply when wealth inequality is stable or decreasing. To overcome this problem it is necessary to introduce a mechanism which offsets the inequality increasing impact of the stochastic component. One simple solution is the formulation proposed by Galton (1889) in his study of inherited genetic characteristics, which yields

$$\ln W_t = \beta \ln W_{t-1} + u_t, \quad 0 < \beta < 1,$$  \hspace{1cm} (2.2)

with $\beta$ indicating the degree of "regression towards the mean". Given appropriate constraints on the stochastic component $u_t$, this process can generate a stable lognormal distribution of wealth in which changes in wealth at the level of individuals exactly balance out to maintain equilibrium in the aggregate (see, for example, Creedy, 1985).

A number of other authors, including Champernowne (1953), Wold and Whittle (1957), Steindl (1972), Shorrocks (1973, 1975b), and Vaughan (1975, 1979), proposed alternative types of stochastic models capable of generating distributions of wealth with upper tails that are asymptotically Pareto. In broad terms, these models all assume that a variant of the law of proportionate effect applies at high wealth levels, where the expected change in wealth must be negative in order to prevent the top tail from drifting upwards over time. Equilibrium is maintained by some mechanism lower down the distribution—such as a pool of low wealth-holders in the case of Wold and Whittle (1957)—which stops wealth from converging to zero in the long run. As with the Galtonian specification Eq. (2.2), these models typically view the observed wealth distribution as the outcome of a stochastic process in which individual wealth movements net out to produce a stable aggregate configuration.

The simplest types of stochastic models lack an explicit behavioural foundation for the parameter values and are perhaps best viewed as reduced forms in which terms like the regression coefficient $\beta$ and the random component $u$ in Eq. (2.2) capture in some unspecified way the influence of factors such as the impact of wealth effects on consumption and the randomness of investment returns. Attempts to incorporate explicitly the relevant explanatory variables rapidly produced complex models that were difficult to solve analytically (Sargan, 1957; Vaughan, 1975, 1979; Laitner, 1979), and these technical obstacles have hampered further development of this line of research.

Meade (1964, 1975) offers another simple framework for analysing wealth distribution based on the accounting identity

$$W_t = W_{t-1} + E_t + r_t W_{t-1} - C_t + I_t, \quad \tag{2.3}$$
where $C_t$ and $E_t$ are, respectively, consumption and earned income in period $t$, net of taxes and transfers; $r_t$ is the average (net) rate of return on investments; and $I_t$ represents net “inheritances” (gifts and bequests) received in period $t$. In the Meade formulation, inheritances are suppressed (or, more accurately, absorbed into the initial wealth level $W_0$), and consumption is assumed to depend on both income and wealth. This yields

$$W_t = W_{t-1} + s_t(E_t + r_t W_{t-1}) - c_t W_{t-1}$$

$$= (1 + s_t r_t - c_t) W_{t-1} + s_t E_t$$

$$= (1 - \beta_t) W_{t-1} + s_t E_t,$$

where $s_t$ is the average rate of saving from current income, $c_t$ indicates the wealth effect on consumption, and $\beta_t = c_t - s_t r_t$ represents the “internal rate of decumulation out of wealth”.

Equation (2.4) provides a convenient framework for investigating the forces leading to greater or lesser wealth inequality. Assuming for the moment that $E_t$, $s_t$ and $\beta_t$ are constant over time and across individuals, Eq. (2.4) has the solution

$$W_t = W^* + (1 - \beta)^t (W_0 - W^*),$$

where $W^* = sE/\beta$. If $\beta$ is negative, the initial wealth differences expand over time and wealth inequality grows without bound. Conversely, if $\beta$ lies between 0 and 1, then wealth converges to the steady state level $W^*$, which depends only on savings from earned income and the internal rate of decumulation. Under the above assumptions, this implies that the distribution of wealth will be completely equalized in the long run. But once allowance is made for variations in earnings across individuals, the model suggests that long run differences in wealth will mirror the differences in earnings or income. As already noted, the tendency for the level of wealth inequality in the twentieth century to decline towards the lower level observed for incomes is one of the best documented findings of empirical studies of wealth distribution.1

The simple relationship between income inequality and the equilibrium level of wealth inequality is modified by variability in the values of $E_t$, $s_t$ and $\beta_t$ across individuals or over time. For example, as Meade (1964) points out, individual differences in the rates of return received on investments are likely to be a disequalizing influence, particularly so if the average rate of return increases with the level of wealth. Another possible consideration is the general equilibrium connections between the components contained in Eq. (2.4). Stiglitz (1969) bases his analysis on a similar model, but also links earnings and rates of return to average wealth via a simple neoclassical production

---

1 It is interesting to note that the recent rise in wealth inequality in the US and, to a lesser extent, Sweden (see Fig. 1 below) have occurred during periods in which income inequality has grown. Wolff (1992) reports a regression of wealth inequality against income inequality in which the coefficient is not significantly different from one.
function. His analysis shows that if the savings rate \( s_t \) is constant, \( \beta \) is positive, and earnings are the same for all workers, then \( E_t \) and \( W_t \) converge to their steady state values, and wealth distribution is again completely equalized in the long run. Assuming a stable balanced growth path for the economy, Stiglitz demonstrates that a variety of saving functions (linear, concave, depending on income, or depending on wealth) produce the same result, but other factors, such as wage differences and class saving behaviour, have disequalizing effects over time.

In principle, the framework proposed by Meade, and captured in Eq. (2.4), could absorb the insights gained from subsequent work on lifecycle saving behaviour, which suggests how savings depend on current income and wealth, and how the parameters may adjust to, say, increases in life expectancy. The rich potential of Eq. (2.4) is also evident in the fact that it can be interpreted equally well in an intergenerational context, as is done by Atkinson and Harrison (1978), with \( W_t \) referring to the “lifetime wealth” of generation \( t \), and the coefficient \( \beta \) capturing the impact of bequest splitting and estate taxation.

Returning to the accounting identity Eq. (2.3), and interpreting it in terms of the lifetime experience of a family (or individual) which begins at age 0 with zero wealth, implies that wealth at the end of age \( t \) is given by

\[
W_t = \sum_{k=1}^{t} (E_k - C_k + I_k) \prod_{j=k+1}^{t} (1 + r_j).
\]

From Eq. (2.6) it is clear that a family’s wealth is determined by: (i) its age, and its history of; (ii) earnings; (iii) saving rates; (iv) rates of return; and (v) inheritances. A complete economic theory of wealth distribution needs both to determine the distributional impact of each of these five elements and to account for differences in the components across individuals and families. In principle, even the distribution of the population across age groups requires explanation, since fertility and mortality are in part economic phenomena. Although a complete theory of wealth distribution remains a distant objective, considerable progress has been made on many of the specific components. In the following subsections we review many of the relevant contributions, concentrating in particular on questions concerned with savings rates and inheritances. Theories of earnings distribution are considered in detail by Neal and Rosen in Chapter 7 of the Handbook, while fertility, mortality, and family formation are discussed in Rosenzweig and Stark (1997).

2.2. Lifecycle accumulation

The paradigm for intra-generational accumulation is the lifecycle saving model (LCM) pioneered by Modigliani and Brumberg (1954).\(^2\) This now comes in several different

\(^2\) The life-cycle model of Modigliani and Brumberg (1954) was developed contemporaneously with the closely related permanent income hypothesis of Friedman (1957). Insights from both these studies continue to have an important impact on current work.
variants, reflecting the degree of realism introduced into assumptions about institutions and about the uncertainties faced by savers. All versions of the model share the basic assumptions that: (i) consumers are forward-looking; (ii) their preferences are defined over present and future consumption, and possibly leisure time; and (iii) life is expected to end with a period of retirement. Extensions of the basic LCM allow individuals to be interested in the consumption of their offspring or in the size of their planned bequest. These are sometimes known as the “bequest-augmented LCM”. The LCM can also be extended in many other ways—for example, by introducing capital market imperfections and/or borrowing constraints, and uncertainty in earnings, rates of return, or the length of life.

The simplest version of the LCM assumes there is no uncertainty; that everyone faces the same constant rate of return, \( r \), and has the same length of life, \( T \); and that there is no bequest motive. The consumer’s problem is then:

\[
\begin{align*}
\text{Max } U &= U(C_1, \ldots, C_T), \\
\text{subject to: }
\end{align*}
\]

\[
C^L_t = \sum_{t=1}^{T} \frac{C_t}{(1+r)^{t-1}} \leq \sum_{t=1}^{R} \frac{E_t}{(1+r)^{t-1}} = E^L,
\]

where \( C^L \) and \( E^L \) respectively denote lifetime consumption and lifetime earnings, and \( R \) is the retirement date. If there is a nonworking period at the end of life, minimal restrictions on the functional form of \( U(\cdot) \) will ensure that saving is undertaken for the purpose of financing consumption in retirement. This is the key explanation offered by the LCM for personal wealth accumulation.

To simplify the exposition, leisure has been neglected in the maximization problem Eq. (2.7). In other respects, the specification is too general to produce precise conclusions about patterns of saving behaviour. The solution to the problem does include, however, what Browning and Lusardi (1996) call the “central tenet” of the modern view of the LCM: that the consumer attempts to equalize the discounted marginal utility of expenditure in all periods. In order to achieve this equalization, given diminishing marginal utility, consumers engage in “consumption smoothing”. Retirement saving is one result. Another is that assets will fluctuate to keep consumption smooth. In addition, since earnings rise quickly in the initial working years, substantial net borrowing (i.e., negative net worth) is expected to be prevalent among young people. The fact that this is

3 We note below those occasions when endogenous labour-leisure choices have important implications for wealth-holding.

4 This goal can be achieved precisely in a world of certainty, but only in expectation when uncertainty is introduced.
not observed suggests either that individuals cannot readily borrow against future earnings (often referred to as a capital market imperfection), or that they have precautionary reasons for saving, as discussed below.

To make more concrete predictions about saving behaviour, intertemporal utility is typically assumed to be additively separable:

\[
U = \sum_{t=1}^{T} \frac{u_t(C_t)}{(1 + \rho)^{t-1}},
\]

(2.8)

where \( \rho \) is the rate of time preference. It is common to assume also that \( u_t = u \), although intuition suggests that the instantaneous utility function is likely to change with age. Constant time preference is a requirement for consistent consumption planning over the lifetime (Strotz, 1956).

Various functional forms have been assumed for \( u \). As discussed below, a quadratic function produces the certainty equivalent (CEQ) version of the LCM. Caballero (1991) and others use the constant absolute risk aversion (CARA) specification \( u(C) = -e^{-\delta C} \), which, like the quadratic, yields closed form solutions under uncertain earnings. But the most popular specification by far is the constant relative risk aversion (CRRA) form given by:

\[
u_t(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma}, \quad \gamma > 0, \quad \gamma \neq 1
\]

(2.9)

\[
u_t(C_t) = \log C_t, \quad \gamma = 1
\]

where \( \gamma \) is the coefficient of relative risk aversion. When Eqs. (2.8) and (2.9) are incorporated into Eq. (2.7), the optimal consumption path satisfies

\[
C_{t+1} = \left( \frac{1 + r}{1 + \rho} \right)^{1/\gamma} C_t = (1 + g)C_t,
\]

(2.10)

and planned consumption grows at a constant rate \( g \), which is approximately (or, under continuous time, exactly) equal to \( (r - \rho)/\gamma \). From Eq. (2.10) it is evident that the intertemporal elasticity of substitution is \( 1/\gamma \).
Consumption in each period is proportional to the sum of wealth left over from the previous period and the present value of earnings over the remaining lifetime. In the initial period this gives:

\[ C_1 = \frac{E^L}{\sum_{t=1}^{T} [(1 + g)/(1 + r)]^{t-1}}. \]  

The propensity to consume rises with age, and is lower for individuals whose preferences generate a higher desired growth of consumption. The impact of the interest rate on the propensity to consume is positive or negative according to whether \( \gamma \) is greater or less than unity respectively.\(^8\)

The prediction of a constant growth rate of consumption has a number of consequences for wealth-holding and wealth distribution. Consider, for instance, a situation in which earnings are constant up to retirement age and zero thereafter. Assume also that the interest rate and the planned consumption growth rate are both zero. Then the consumer will save a constant amount during each working year, and will dissave a constant amount during each retirement year, with accumulated savings falling to zero at the point of death. The net result is an age-profile of wealth which rises linearly with age until retirement, and then declines linearly to zero.

This very restrictive example highlights two important—and robust—implications of the LCM. First, the age-profile of wealth is expected to have a pronounced hump-shape, with the peak occurring at or near the date of retirement. Strictly speaking, this prediction is best tested against empirical estimates of wealth-holdings which include imputed pension rights. But a similar prediction holds for marketable wealth if pensions are treated as deferred earnings, as long as pension income is lower than earned income during working life.

The second implication is that substantial wealth inequality can arise between the richest members of society (those around retirement age) and the poorest (those just starting out on their working lives and those nearing death), even if everyone is completely equal in all respects other than age. Thus age differences alone are expected to account for a substantial proportion of observed wealth inequality. The significance of this point is that wealth variations due to age do not represent true differences in lifetime opportunities, and the resulting wealth inequality is therefore spurious from an equity perspective. The extent to which wealth inequality is attributable to age differences is a question which we consider in detail in Section 4.

More general consequences of the consumption path Eq. (2.10) are obtained if it is assumed that \( r > \rho \), so that consumption is planned to grow at a constant positive rate. If earnings have the typical hump-shaped age-profile, individuals should typically

---

\(^8\) Empirically, it is generally thought that \( \gamma \) exceeds unity by a significant amount. However, this does not necessarily mean that consumption is expected to decline with increases in the interest rate. \( E^L \) falls with \( r \), generating a “human wealth effect” which is found to dominate in simulation exercises: see, for example, Summers (1981).
dissave for some period when young, but then save fairly heavily in middle years and, of course, dissave in retirement. However, as many researchers have pointed out, in the real world we find that: (i) most individuals and families are net savers when young; (ii) consumption tends to track earnings or income, also having a hump-shaped age-profile rather than an exponential one; and (iii) when family size is held constant, consumption falls on retirement and then declines further over time, producing less dissaving than the model suggests (Attanasio, 1998). Davies (1979, 1982) shows that taking account of family size in Eqs. (2.7) and (2.9) deals to a large extent with problems (i) and (ii) over the working lifetime (see also Attanasio et al., 1995). However, once children have left the home the CRRA specification (with $r > \rho$) predicts rising consumption as long as family composition remains constant. This runs contrary to what is observed. The downward jump in consumption on retirement can be explained by incorporating labour-leisure choice if goods and leisure are substitutes, since there should be substitution away from goods and towards leisure on retirement (Davies, 1988). The continued downward trend throughout retirement would follow if $u_t$ varied with age in such a way that the marginal utility of consumption, at a given consumption level, decreased continuously over time.

Empirical evidence is not always in agreement on the main features of savings behaviour, a point well illustrated by the debate over whether individuals dissave in retirement, and if so to what extent. In the 1950s and 1960s a number of crossection datasets appeared to indicate that mean wealth did not fall in retirement, or that mean savings remained positive. This was viewed as a major challenge to the LCM, although Shorrocks (1975a) showed that the apparent lack of dissaving in the UK could be explained by compositional changes in an ageing population. More recent studies using panel data, for example the Retirement History Survey in the US, show significant dissaving in retirement (Hurd, 1997). This result is not always obtained. Alessie et al. (1997), for instance, find rising net worth in Dutch panel data between 1987 and 1991, and Jianakoplos et al. (1989) find that it is difficult to generalize about the results of going from cross section to panel data in the US. However, the broad consensus is that, after the first few years, private wealth (as well as social security wealth) declines in retirement. What remains in considerable doubt is the speed at which dissaving takes place: see, for example, Hurd (1997: table 9, p. 935).

As mentioned earlier, the simple LCM tends to predict that negative wealth-holdings will be common among the young. The fact that this is not borne out empirically suggests that liquidity constraints and/or precautionary saving behaviour should be taken into account. Other evidence points to the need to “augment” the LCM in other ways. White (1978), Darby (1979) and Kotlikoff and Summers (1981) all present calculations for the US which show that the aggregate savings rates and wealth-income ratios are much greater than those predicted by lifecycle saving alone. Although other researchers find the LCM to be more successful in this respect (see Hurd, 1997), most favour the

---

9 More recent evidence of this type is considered by Mirer (1979) for the US.
introduction of bequest motives into the LCM—which would reduce dissaving among the retired—and the explicit recognition of various forms of uncertainty, such as those relating to rates of return, earnings, medical costs, and the length of life.

The simplest way of incorporating bequest motivation is to make $U$ depend on terminal wealth $W_T$. This is sometimes referred to as the “warm glow” approach, since parents’ satisfaction from making a bequest depends only on the size of transfer, and not on the characteristics of heirs, such as their incomes and needs. This approach might be defensible in modelling aggregate behaviour, but it is too crude an assumption for distributional analysis. Alternative formulations of bequest motives are discussed in Section 2.3 below.

The effects of both uncertain rates of return and uncertain earnings received early analytical attention (e.g., Leland, 1968; Sandmo, 1970). However, more recent simulation-based work has tended to neglect uncertain rates of return, concentrating instead on earnings uncertainty and uncertain nondiscretionary expenditures like medical costs, which are formally equivalent to earnings uncertainty and can therefore be treated by similar methods.

The effects of uncertain rates of return are complex. On the one hand, introducing any form of uncertainty reduces the consumer’s welfare and real income. If consumption is a normal good, this tends to increase saving. On the other hand, by increasing saving the consumer exposes himself to greater risk when returns are uncertain. In general the net effect is ambiguous. However, Lippman and McCall (1981) show that there is a clear cut result in the CRRA case: if $\gamma < 1$, the introduction of risky rates of return raises consumption and reduces saving; when $\gamma > 1$, saving rises. In the lifecycle simulation literature, $\gamma$ is usually taken to be about 3 (Carroll, 1992; Hubbard et al., 1994), and some empirical work suggests even higher values (Mankiw and Zeldes, 1990). This suggests that introducing uncertain rates of return into the LCM might be a fruitful way of obtaining greater aggregate wealth, although, as mentioned earlier, this option has received little attention.

The simplest way to proceed when earnings are allowed to be uncertain is to use the certainty equivalent (CEQ) version of the LCM based on quadratic utility. This approach has been used recently to analyse the evolution of inequality in income, consumption, and wealth over the lifecycle. Deaton and Paxson (1994) show that the model predicts rising consumption inequality over the lifetime, as well as an increase in income inequality with age. These predictions are in accord with what is observed empirically. Davies (1996a) uses the CEQ approach to examine changes in wealth inequality with age.

---

10 Under this approach none of the observed heterogeneity in bequest behaviour can be captured. Models which can help to explain, for instance, why parents with similar incomes might desire very different levels of bequest are needed in order to understand the complex patterns of bequest and inheritance.
Some of the important implications of the CEQ model for the distribution of wealth are brought out clearly in the case \( r = \rho = 0 \), where consumption simply equals wealth plus expected future earnings, divided by the length of the remaining lifetime:

\[
W_{t-1} + E \left[ \sum_{k=t}^{R} E_k \right]
\]

\[
C_t = \frac{W_{t-1} + \left( R - t + 1 \right) E_t + \epsilon_t}{T - t + 1}. \tag{2.12}
\]

Earnings have a permanent and transitory component, and may be written as \( E_t = E_t^P + \epsilon_t \), where \( t \) has zero mean and a finite variance. Much recent work finds the permanent component of earnings to be highly persistent. Abstracting from the deterministic part of the age-earnings relationship, the earnings process in the US is approximated by a combination of white noise and a random walk (Carroll, 1992; Hubbard et al., 1994). This suggests

\[
E(E^P_k) = E_t^P, \quad k = t, \ldots, R.
\]

Under that assumption Eq. (2.12) yields:

\[
C_t = \frac{W_{t-1} + \left( R - t + 1 \right) E_t^P + \epsilon_t}{T - t + 1}, \tag{2.3}
\]

which indicates that the propensity to consume out of permanent earnings is much greater, and the propensity to save less, than out of current wealth or transitory earnings.

The result of an earnings process with highly persistent shocks is that earnings inequality tends to grow over the working lifetime. The wealth which has been built up by saving out of past permanent components of earnings will correspondingly become more unequal over time. However, as Davies (1996a) points out, empirical studies find that wealth inequality initially declines with age, and while it usually increases later, this increase does not begin until the later years of the working lifetime. The initial decline in wealth inequality can be mimicked by the CEQ model if transitory earnings are not merely measurement error. The high propensity to save out of transitory earnings, combined with the assumption that everyone begins life with zero wealth, means that wealth inequality will be dominated at low ages by saving (and dissaving) out of transitory earnings. Transitory earnings tend to "average out" over time, so that this source of wealth inequality declines with age.

While valuable insights are gained from the CEQ model, its prediction that earnings uncertainty does not affect saving is widely regarded as unrealistic and untenable. There is much empirical evidence that consumers save more in response to earnings uncertainty; in other words, they exhibit precautionary saving (Browning and Lusardi, 1996: pp. 1835–1838). As shown by Kimball (1990), any additive intertemporal utility function with a positive third derivative for \( u \) (such as a CRRA function), or what he terms "prudence", will produce precautionary saving.

The quantitative impact of earnings uncertainty on saving has been investigated by a number of authors. Skinner (1988) and Zeldes (1989) investigate the effects of assuming prudence, but do not include liquidity constraints or impatience—features which have
subsequently received much attention. While Skinner and Zeldes obtain higher—that is, more realistic—levels of aggregate saving, they do not predict the frequent incidence of low positive wealth-holding at all ages which is a prominent feature of wealth distribution in the US and many other countries. Deaton (1991) accounts for the latter with an infinite horizon model incorporating a high rate of time preference (impatience) and liquidity constraints, as well as prudence. In Deaton’s model, when their earnings are high, savers accumulate a buffer stock of assets which, like the observed assets of many actual consumers, may be quite small. These assets are run down when earnings are low.

Carroll (1992) shows that liquidity constraints are unnecessary to obtain buffer stock behaviour if consumers are impatient and prudent, and if there is a finite probability of complete earnings interruptions, which he claims is the case for some individuals in the US. Since “earnings” must be interpreted broadly (that is, to include transfer payments), this argument would not apply in countries which have more comprehensive income maintenance systems.

Hubbard et al. (1994, 1995) claim that while buffer stock models predict the observed low wealth-holding of many individuals, they cannot generate sufficiently high values for mean saving for retirement. Their model produces both these features, by retaining the assumptions of prudence and liquidity constraints, but also adding uncertain medical costs, uncertain lifetimes, and means-tested transfer payments and other social benefits. The latter strongly discourage low income individuals and families from saving since most assets must be exhausted before an individual qualifies for means-tested benefits.

Uncertain lifetime is readily incorporated in the LCM. Taking the CRRA approach, the objective function becomes:

\[
EU_t = \sum_{k=t}^{T} \frac{(1 - \phi_k)C_k^{1-\gamma}}{(1 + \rho)^{k-t}(1 - \gamma)},
\]

where \( T \) is now interpreted as the maximum length of life, and \( \phi_k \) is the probability of death before period \( k \).

In order to characterize behaviour when the individual has the objective function Eq. (2.14a), it is necessary to specify the nature of insurance markets. In a perfect insurance market with actuarially fair annuities, all individuals completely annuitize their wealth. In that case Eq. (2.14a) is maximized subject to the constraint:

\[
\sum_{k=t}^{T} \frac{(1 - \phi_k)C_k}{(1 + r)^{k-t}} \leq W_{t-1} + \sum_{k=1}^{R} \frac{(1 - \phi_k)E_k}{(1 + r)^{k-t}}.
\]
As shown by Yaari (1965), planned consumption again follows the consumption path Eq. (2.10), growing at the constant rate \( g \) (approximately equal to \( (r - \rho)/\gamma \)).\(^{12}\)

It is usually argued that annuity markets are sufficiently imperfect that the above solution does not apply.\(^{13}\) If imperfections are severe, annuities may have such low rates of return that rather than annuitizing all of their wealth, individuals may not buy annuities at all. It is widely believed that this is broadly the situation in practice (Friedman and Warshawsky, 1988). In this case the individual again maximizes Eq. (2.14a), but subject to Eq. (2.7b) and a nonnegative wealth constraint. The solution to this problem yields a nonconstant growth rate of consumption, \( g^u \), in periods of positive wealth-holding:

\[
g^u = \frac{r - \rho - \pi_t}{\gamma},
\]

(2.15)

where \( \pi_t = (\phi_{t+1} - \phi_t)/(1 - \phi_t) \) is the mortality hazard in period \( t \); that is, the probability of death conditional on having survived to age \( t \).

Under Eq. (2.15) the age-profile of consumption, and therefore the age-profiles of saving and wealth, are predicted to be quite different from those in a world of certainty. As is well-known, the mortality rate initially rises with age, and does so at an accelerating pace. Beyond some age (about 80 for males in the dataset used by Davies, 1981, for example) the mortality rate begins to decline towards a stable lower level. Assuming CRRA preferences and applying Eq. (2.15), this mortality pattern means that the growth rate of consumption will fall at an accelerating rate over the working lifetime and early retirement, but then rises at advanced ages. If \( r - \rho \) is sufficiently small, this can produce a realistic hump-shaped age-profile of consumption, or a profile which is predominately hump-shaped, but with a rising trend in later years of retirement.

The recognition of uncertain lifetimes also leads to more realistic age-profiles for saving and wealth-holding. In particular, the rate of dissaving in retirement may be lower than in the case with a fixed lifespan. In the absence of state or occupational pensions, no rational consumer would ever run assets to zero. This simple result suggests that introducing uncertain lifetimes will produce higher aggregate savings in simulation exercises, possibly eliminating the need to introduce a bequest motive. Davies (1981) shows that this is indeed the case with a wide range of reasonable values for the parameters \( r, \rho, \) and \( \gamma \), and realistic patterns of earnings and mortality. However, Davies' results also indicate

\(^{12}\) Although the growth rate is the same as under certainty, the overall consumption profile may shift up or down.

\(^{13}\) Both moral hazard and adverse selection have an effect on annuity markets, and these problems seem to be much more severe than for life insurance. Possible explanations include: (i) people take better care of themselves once they have an annuity, but are less inclined to take the opposite action when they are covered by life insurance (i.e., moral hazard is more severe under annuities); and (ii) the likelihood of unusually high life expectancy is largely private information, but poor health is often ascertainable and verifiable by an insurance company (i.e., adverse selection is more severe under annuities).
that it is wrong to disregard the effect of pensions, because if pensions are included and
if $\gamma$ is not too high, uncertainty about the length of life can cause faster decumulation.\footnote{Davies' pensions decline with age at an annual real rate of 2.25%, reflecting the imperfect indexation of pensions in the early 1970s in Canada. When pensions are constant in real terms throughout retirement, $\gamma$ must be quite high to yield the result that uncertain lifetime increases saving. This is one reason for the differences in the results obtained by Leung (1994) and Davies (1981).}

Over the course of the twentieth century, state and occupational pensions have tended to become more widespread and more generous. This does not necessarily mean that the need to save privately for retirement has declined, since retirement ages have also been falling and life expectancy has increased. However, these pension trends do imply that individuals may rationally run their private wealth to zero at some point during retirement.\footnote{Indeed, Leung (1994) shows that if they live long enough, people covered by pensions must exhaust their private wealth completely.} As Leung (1994) points out, this provides another possible explanation for the low marketable wealth of part of the population.

Despite the ambiguous effect of uncertain lifetime on the rate of decumulation in retirement in the presence of pensions, it is interesting to note that, with uncertain earnings, medical costs and lifetime, Hubbard et al. (1994, 1995) generate age- profiles of consumption and wealth-holding, aggregate saving rates, and overall wealth-income ratios which are all close to those observed for the US. Thus it is not necessary to include intentional saving for bequests in order to replicate the aspects of wealth-holding of interest to those concerned with just consumption behaviour. However, intentional bequest behaviour is likely to be required in order to explain the shape of the upper tail of the wealth distribution, intergenerational wealth mobility, and the size distribution of inheritances. These are important issues for those interested in the distribution of wealth.

2.3. Intergenerational issues

Models of inheritance should ideally take account of demographic factors such as patterns of fertility and marriage, as well as factors which are more strictly economic. In contrast to lifecycle models, they should also analyse the evolution of the distribution over many generations. To do this it is helpful to begin with pure intergenerational models, which abstract from the relationship between inheritance and lifecycle saving, as well as the interaction between material inheritance and transfers in human form.

A useful point of departure is to consider how patterns of marriage, fertility, estate division, and taxation affect the evolution of wealth distribution in a world of pure accumulation, that is, one in which wealth grows at the rate of interest, and there is no consumption. The simplest case is that of a society in which all individuals marry, have children, and leave to their children (and no others) a total bequest equal to the amount they themselves inherited. In such a society the current distribution of inherited wealth depends only on the wealth distribution of the previous generation, together with patterns of marriage, fertility, estate division, and taxation. Blinder (1973) notes that in a world where all families have two children, of whom one is a boy and the other a
girl (i.e., in the absence of differential fertility and population growth), and where estate taxation is either absent or proportional, the distribution of (relative) inherited wealth remains unchanged over time if either (a) mating is completely assortative (children married partners from families with identical wealth) and all families at a given level of wealth divide estates in the same way between their son and daughter, or (b) the members of one gender inherit nothing. In all other cases the distribution becomes more equal over time, with the speed of equalization depending on the correlation of parental wealth between spouses and the way in which estates are divided. The most rapid equalization occurs with completely random mating and equal estate division.

More general assumptions have been explored by Meade (1964), Stiglitz (1969), Pryor (1973), Atkinson and Harrison (1978), and Atkinson (1980). These authors show that, although imperfect correlation of spouses’ inheritances has a powerful equalizing effect on the distribution of inherited wealth, other forces may have the opposite effect. For example, inherited wealth becomes steadily more concentrated over time under primogeniture (the practice of passing all wealth to the eldest son) if each family has more than two children, including at least one son, and all children marry (Atkinson, 1980: p. 48). In addition, if the wealthy consistently have fewer children, inherited wealth can become continuously more unequal even in the presence of some sharing of estates among siblings and imperfect sorting of marital partners by parental wealth.

As Blinder (1973) points out, evidence for the US suggests that differential fertility is not particularly marked, and the correlation between spouses’ family background characteristics, although positive, is low. Furthermore, unequal division of estates is relatively unusual in the US, and is also far from the norm in the UK, although it may be slightly more prevalent there. It is therefore likely that the patterns of marriage, fertility, and estate division in advanced Western societies tend toward equalization of the distribution of wealth over time. Complete equalization does not occur if rates of return on saving are positively correlated with the level of wealth, which will be the case if wealthier individuals choose riskier portfolios or have preferential access to advantageous investment opportunities. Against this, progressive tax systems tend to make after-tax rates of return lower for the wealthier. The net effect is difficult to judge.

The sensitivity of results to the pattern of lifetime saving becomes clear in the work of Pryor (1973) and Blinder (1976). Pryor simulated processes of marriage, fertility, saving over the lifetime, estate taxation and bequest division. In his model everyone marries and has two children. A normal distribution of earned income in each generation is assumed. Two different models of lifetime saving are considered: one in which the

16 Shorrocks (1979) examines how altruistic parents may condition their division of estates between sons and daughters on the anticipated bequest behaviour of in-laws. Where there is no gender preference, any fixed ratio of son’s to daughters’ bequests could be supported. In contrast, even mild gender preference could lead to an equilibrium with primogeniture.

17 The intuition is simple. Primogeniture maintains the same wealth distribution in each generation among those with positive wealth, but the proportion of positive wealth-holders is decreasing over time because of population growth.
bequest is proportional to income, and another in which bequests are proportional to income in excess of some threshold. Proportional saving out of income reduces the simulated inequality of income found after 30 generations have passed. The kinked bequest function, however, leads to an increase in income inequality after 30 generations, in the absence of primogeniture. (With primogeniture the effect is negligible.) The nature of lifetime saving therefore has important implications for the evolution of the distribution of wealth. This conclusion is echoed in the simulation exercise of Blinder (1976) who replaced the ad hoc lifetime saving models of previous authors with a bequest-augmented lifecycle model based on an explicit representation of preferences. Blinder found that, even with equal division of estates and random mating, the force of inequality coming from lifetime saving was strong enough to produce a rising trend in the inequality of inherited wealth over time.

In the last few decades considerable interest has been shown in models of intergenerational transfers rooted in an explicit specification of parental preferences and opportunities. The crucial distinction is between models in which transfers are made for reasons of altruism (see, for example, Becker and Tomes, 1976), and those in which they represent a quid pro quo for “attention” supplied by potential heirs. Models of the latter type are said to feature “exchange” or “strategic” bequests (Bernheim et al., 1985; Cox, 1987).

The Beckerian approach assumes that parental preferences are given by:

\[ v = v(c^p; c_1, \ldots, c_n), \]  

(2.16)

where \( n \) is the number of children, \( c^p \) is parental lifetime consumption, and \( c_i \) is the lifetime consumption of child \( i \).\(^{18}\) Parents spend \( x_i \) on human capital investment in child \( i \), and provide bequests \( b_i, i = 1, \ldots, n \). Children’s earnings equal the product of the human capital rental rate, \( w \), and their human capital, \( h_i \), which is generated using inputs of \( x_i \) according to a production function which depends on the child’s own characteristics:

\[ e_i = w h_i (x). \]  

(2.17)

Higher ability children have higher \( h_i \) for any given \( x_i \). If the first derivative of \( h_i \) is also higher, and if parents make efficient investments, it can be shown that in equilibrium more able children will have greater earnings than less able children.

With this structure, and assuming a zero interest rate for convenience, the consumption of parents and children is given by:

\[ c^p = e^p - \sum_{i=1}^{n} (x_i + b_i), \]  

(2.18a)

\(^{18}\) The amounts of “lifetime” consumption include only consumption in the adult phase of the lifecycle. Consumption of young children living at home with their parents is included in \( c^p \).
where $e^p$ is parental lifetime earnings. The parent’s problem may be written:

$$\max_{b, x} \left( e^p - \sum_{i=1}^{n} (x_i + b_i); e_1 + b_1, \ldots, e_n + b_n \right)$$

subject to $x_i, b_i \geq 0$, $i = 1, \ldots, n$.

Suppose that capital markets are perfect and that parents can borrow against their children’s earnings (relaxing the requirement that $b_i \geq 0$). In this case, Eq. (2.19) decomposes into two problems. The first is to choose the efficient level of investment in each child, $x_i$. The corresponding earnings level for the child is $e^*_i$. The second part of the problem is to maximize Eq. (2.19) with respect to the $b_i$’s, taking the $e^*_i$’s as given. If parental preferences treat children symmetrically, this produces the well-known result:

$$e^*_1 + b_1 = \cdots = e^*_n + b_n.$$  

In other words, under altruism parents plan to equalize completely the children’s incomes net of transfers.

As we discuss in Section 4 below, while some authors have found evidence of a negative impact of children’s earnings on bequests, nothing like fully compensatory bequest behaviour has been observed in practice, where bequest division has been studied.

Davies (1986) analyses the equilibrium impact of bequests and estate taxation on overall inequality in the Becker model with perfect capital markets. He shows that a linear redistributive tax on bequests increases long-run inequality. The reason is that it interferes with the intergenerational sharing of luck, which in equilibrium leads total incomes within a generation to be distributed more equally than earnings. In other words, the institution of bequest reduces equilibrium inequality in the simple altruistic model, and policies which interfere with the bequest process are likely to increase inequality. This conclusion has been widely regarded as unrealistic (see, for example, Bevan, 1979; Atkinson, 1988; Wilhelm, 1997), and can be avoided by abandoning the assumption that parents can borrow against their children’s future earnings (i.e., the perfect capital market assumption), or by incorporating nonaltruistic motives for bequests.

The Becker model produces more interesting insights, and no longer predicts fully compensatory bequests in all cases, if we recognise that parents may end up at a corner

19 Since there is no labour-leisure choice, with a perfect capital market parents choose $x_i$ to maximize $e_i - x_i$. This occurs when $\frac{\partial e_i}{\partial x_i} = 1$. With a positive interest rate, $r$, the corresponding condition is $\frac{\partial e_i}{\partial x_i} = 1 + r$.

20 Shorrocks (1979) points out that in a more realistic version of the model parents should take into account the bequests which their children’s spouses can be expected to receive. In general this will invalidate Eq. (2.20).

21 Bevan (1979) performs an intergenerational simulation in which negative bequests are not allowed. He finds that the existence of bequests may either raise or lower consumption inequality.
solution where \( b_i = 0 \). This will occur if parents do not find it optimal to set \( x_i \) high enough to achieve \( e_i = e^*_i \). If parental income is low, or if the required expenditure is too great, they may only be willing to fund the efficient level of investment if \( b_i \) is negative. But, while some children may be trusted to repay parents for investments in their human capital, in modern societies this cannot usually be enforced. Thus, it is possible that some children will fail to achieve their efficient level of earnings, and will receive zero bequest. In general, bequests will no longer perfectly compensate for differences in earnings. Also, since zero bequests will be more common in families with low income parents, one would expect the incidence of bequests to rise with parental income, and to rise in a nonlinear manner. For example, if parental income and children’s ability were perfectly correlated, bequests would be zero up to some threshold parental income, and increasing beyond that point. This prediction may help to explain why even many older people report not having received any significant bequest in their lifetime, and why Menchik and David (1983) found the estimated income elasticity of bequests to be less than one for the bottom 80% of the population, but considerably greater than one for the top 20%.

An important aspect of the Beckerian analysis is its simultaneous consideration of human and nonhuman transfers. Thinking about transfers in this way yields the following insights:

1. It provides an explanation for the empirical observation that many families, perhaps the majority, never receive significant inheritances, and that bequests are a luxury.
2. It suggests that inheritances may not be as disequalizing as one might have thought, despite their high concentration. If material transfers and human capital investments are substitutes to some extent, as the theory suggests, then within a family those children who get larger bequests will tend to receive smaller investments in the form of human capital.
3. It indicates how the actions of parents may reduce the degree of intergenerational income mobility which would occur “naturally” if all children received efficient investments in their human capital, and if there were no inheritances.
4. It highlights the fact that changes in educational systems, opportunities, and finance may have an important impact on saving and wealth accumulation.

The altruistic model does not perform well when we consider the optimal timing of intergenerational transfers. With perfect capital and insurance markets, both parents and children are indifferent between lifetime gifts and bequests. However, if children are often liquidity constrained as young adults, they will prefer to receive their transfers earlier rather than later. The fact that parents provide a large fraction of their total transfers in the form of bequests is therefore a challenge to the altruistic model.

22 Menchik and David (1983) used a longitudinal Wisconsin tax-filter dataset to construct estimates of parents’ lifetime earnings, and estimated a spline relationship between lifetime income and the size of estate. They found elasticities ranging from insignificant negative values to +0.7 for the bottom 80% of parents, but elasticities ranging from 2 to 4 for the top 20%.
It should also be noted that it is possible to extend the altruistic model to allow parents to care not only about their children’s consumption, but also about “attention” they may receive from children (see Cox, 1987; Davies, 1996b). Examples of attention include phone calls, letters, and visits, as well as the provision of more essential services during periods of ill health or infirmity. In an altruistic model, with the child behaving nonstrategically, the parent would pay the child for attention at a normal supply price—for instance, the child’s wage rate. The total transfer to the child would then be composed of expenditures on attention, plus a gratuitous component, $b_i^g$. This result has interesting implications. For example, in families where $b_i^g$ is relatively small, the behaviour of $b_i$ could be dominated by $b_i^p$, so that, for example, the observed correlation between $b_i$ and $e_i$ might be positive, holding parental income constant, despite parents being altruistic.

Bernheim et al. (1985) propose an alternative, strategic model of interactions between parents and their adult children. They set out a simple noncooperative game in which parents with two or more children could, it is claimed, credibly threaten to disinherit any child who did not accept a “forcing contract” specifying the amount of services they had to provide. This approach generates the same outcome as that obtained when altruistic parents purchase attention from children who act nonstrategically. In both cases, all rents from parent-child interaction accrue to the parents.

Cox (1987) suggests that it might be more realistic to consider a model in which both parents and children have some bargaining power. One way to do this is via a cooperative bargaining approach such as Nash bargaining. Another is to model the relationship between parents and their children as a repeated game with an uncertain horizon (the parents’ lifetime): see Kotlikoff and Spivak (1981) and Davies (1996b). Under the latter approach, parents show cooperation by holding some bequeathable wealth, and children cooperate by providing attention. Either side could defect, but would be punished. Under some, but not all, conditions, the threat of punishment can be sufficient to maintain cooperation.

The repeated game approach produces interesting insights. For example a parent-child pair may initially be in a region where cooperation cannot be supported because the actuarial value of the expected bequest is too small to induce attention from the child. However, as the parent ages, or perhaps as a result of parental ill health, the parent’s mortality probability may rise sufficiently to allow cooperation to be maintained. This provides a possible explanation for the empirical observation that attention increases when parental health declines, and does so more if parents’ bequeathable wealth is higher (Bernheim et al., 1985).

---

23 The observation that even altruistic parents may purchase attention from their children implies that an exchange of gifts and bequests for child services is not only to be obtained in a strategic model. There is a tendency in the literature to treat signs of exchange as evidence against the altruistic model and in favour of a strategic alternative.

24 It is questionable whether the threat to disinherit is credible. A parent who felt the slightest degree of altruism toward his children would be tempted to reinstate the disinherited child on his deathbed.
Regardless of whether or not exchange relationships between parents and adult children require a strategic explanation, one may ask why a higher incidence of bequests is not observed. Parental desire for attention is presumably widespread. Adding an altruistic motivation for transfers would seem to suggest that sizeable bequests should be quite common. This is not the case in practice. The explanation may lie in the fact that much exchange between parents and children is not recorded, because it takes place via small gifts or in nonpecuniary form. While these forms of exchange are more difficult to observe, they may nevertheless have important effects on the wealth-holdings of a large portion of the population.

Little attention has been given to the impact of exchange and strategic bequest behaviour on wealth inequality. However, some implications can be noted. First, if children are simply paid for attention at their usual wage rates then bequests should be treated as part of their lifetime earnings and the analysis, in lifetime terms, is equivalent to assuming a world in which inheritances do not exist. But the fact that bequests are received later than earnings, and are subject to greater uncertainty, may affect lifetime expenditure patterns, especially if children face capital market constraints. This will have complex effects on the evolution of wealth inequality over the lifecycle which would need to be studied by simulation. In addition, if attention is a luxury good, higher income children will receive a larger proportion of their lifetime income in the form of bequests. This is likely to mean that bequests have an apparently disequalizing effect on lifetime income and wealth. An even more disequalizing impact could be generated by a model in which children shared some of the rents from interaction with their parents, assuming again that these interactions are relatively more important in wealthier families.

2.4. Self-made fortunes

There is one category of people for whom the current theories of lifetime accumulation and bequest behaviour seem quite inadequate: those who have made sizeable fortunes during the course of their lifetimes, largely, or exclusively, as a result of their own efforts. Very often they continue working even though they have amassed enough wealth to live forever in extreme luxury. Moreover, bequests to children and grandchildren do not appear to be strong motives. Two prominent examples, Bill Gates and Warren Buffett, have both announced their intention to leave the bulk of their fortunes to charity. While such people are, by definition, rare, their fortunes are sufficiently large to have an impact on wealth inequality.

A number of authors have stressed the need to provide a separate model for this type of person. Atkinson (1980) draws attention to the different bequest intentions of those who have obtained substantial wealth through entrepreneurship. Masson and Pestieau (1996) and Arrondel and Lefèrere (1996) refer to “capitalist” bequest motives. Little has been done to develop a formal model of the bequest intentions of those with large

25 Even in their relatively high income sample of retired academic faculty and staff, Laitner and Juster (1996: p. 897) report that only 41% had received, or expected to receive, any substantial transfer.
self-made fortunes. But the fact that many past fortunes have ended up in charitable foundations named after their benefactor suggests that a prime motivation may be a desire to see the continuation of their name in the public arena. This might be expressed as a "desire for immortality." It seems less than adequate to capture this objective simply in terms of the utility derived from terminal wealth.

Current theories are also inadequate in explaining how self-made fortunes are generated. Casual empiricism suggests that they are linked inextricably with entrepreneurial activity, and that, although ability and ambition play a part, the size of the fortune depends largely on "being in the right place at the right time"—in other words, luck. In effect, social and technological developments create opportunities for fortunes to be made, which specific individuals exploit with varying degrees of success.

Shorrocks (1988) presents a simple model of this kind of situation. There are two types of opportunities to make fortunes. One requires entrepreneurial time, but no capital: for instance, panning for gold in a river. The other requires time and a substantial amount of capital: for example, sinking a shaft for a gold mine, or prospecting for off-shore oil. High-capital risks offer a greater expected return, but to gain access to them, individuals must first succeed as low-capital entrepreneurs. This drives down the expected return from low-capital risks to a point below the wage that entrepreneurs would receive if employed. It becomes unattractive to pool risks, and hence low-capital entrepreneurs gamble everything on one activity. If successful, however, individuals have an incentive to diversify in order to retain access to the advantageous high-capital opportunities.

This approach highlights several important influences on wealth holdings at the top of the distribution. First, it draws attention to the social forces and technological developments which create the opportunities for fortunes to be made. Recently in the UK, for instance, a number of large fortunes can be traced to the privatisation of publicly owned enterprises during the 1980s and 1990s. The approach also draws attention to the significance of capital market imperfections, an issue that plays a central role in the recent work on persistent inequalities reviewed by Piketty in Chapter 8 of the Handbook. The idea that prospective entrepreneurs with little initial capital must first establish a track record of success before being allowed access to finance for larger scale ventures is one that chimes well with real world practice, and one that could provide the basis of a model of self-made fortunes.

3. Empirical evidence on wealth inequality

Four sources of data have been used to estimate the distribution of wealth: household surveys of assets and debts, wealth and estate tax records, and investment income data. Other information on the very richest families is provided by independent estimates of the wealth of named individuals. We review each of these data sources in turn, discussing the problems that need to be addressed in order to produce reliable evidence on
the pattern of wealth-holdings. We also summarise the evidence on wealth inequality for a range of countries, using the share of wealth owned by the richest groups in the population, or the value of the Gini coefficient, to represent the degree of concentration. This is followed by a discussion of the portfolio composition of wealth-holdings. The section concludes with a review of the factors which may account for differences in wealth distribution across countries and for the observed variations in wealth inequality over time.

It should be stressed at the outset that comparisons of wealth distributions present numerous difficulties. To begin with, the precise concept of wealth is open to question. The most widely applied definition is net marketable wealth or net worth—the total value of all assets which the owner can sell in the market, less any debts. Because contributions into pension schemes are a substitute for other forms of lifecycle saving, this concept of wealth is often expanded to include the capitalized value of entitlements to pensions, including state pensions; following Wolff (1995), we refer to this as augmented wealth. Other empirical studies have applied narrower concepts such as gross assets or financial wealth. Generally speaking, the precise definition used in empirical work reflects the availability of data and a desire for comparability with other data sets, rather than a judgement about the ideal concept of wealth which, in any case, depends on the purpose for which the data are subsequently employed.

A second major distinction is between the distribution of wealth across households (or families), and the distribution across individuals. Again, the choice is dictated primarily by the source of data rather than by an opinion about the most appropriate economic unit: as far as we are aware, there is no consensus on this issue. As will become apparent, other factors, such as the year in which the data are collected and the refinements applied to the raw data, may also significantly affect the estimates of wealth concentration. For all these reasons, wealth distribution data tend to be less homogeneous and more variable in quality than, say, income distribution data, and need to be viewed with more than the usual degree of caution.

3.1. Household surveys

Household surveys of assets and debts are a valuable source of data on wealth distribution, but the information obtained is typically less reliable than that for income. There are two main reasons for this. First, the distribution of wealth is so heavily skewed that, even with fairly large sample sizes, sampling error may significantly affect estimates of dispersion. Consider, for example, a country with 10 million families, 500 of which have wealth in excess of $20 million. With a random sample of 15,000 families, there is about an even chance of including one or more of these wealthy families. If none of them are captured in the survey, the upper tail of the sample distribution will be too short. In other cases, for instance when a “centi-millionaire” family is sampled, the tail will be too long. So the raw data from a sample survey which does not over-sample at the top end cannot
be expected to provide an accurate snapshot of the pattern of wealth-holdings among the very rich.

Nonsampling errors are also serious, and affect much more than the upper tail of the distribution. They fall into two main categories: differential response and misreporting. The problem of differential response is due to the refusal (or inability) of households to answer either all parts of the survey ("survey nonresponse"), or one or more specific questions ("item nonresponse"). Item nonresponse is more tractable than survey nonresponse since it can be combated by careful imputation of missing amounts (Juster and Kuester, 1991: pp. 51–54.). Survey nonresponse would not cause difficulties if the response rate was uncorrelated with wealth-holding, but the evidence strongly suggests that the response rate is lower for better-off families. Furthermore, the pattern of nonresponse in wealth surveys is more pronounced than in surveys of income. For example, Projector and Weiss (1966: Table 15, p. 52) reported an overall response rate of 83% in the 1962 Survey of Financial Characteristics of Consumers (SFCC) in the United States, but a response rate of just 37% for families with incomes above $100,000. In the high-income sampling frame of the 1977 Canadian Survey of Consumer Finances, the response rate was about 45% (Oja, 1980: p. 342). Kennickell et al. (1998: p. 23) report that the response rate to both the 1992 and 1995 SCF's differed considerably between the representative subsample and the "list sample", which focuses on those with high incomes. The response rates across the two surveys were 70% in the representative sample, but only 34% in the list sample. Differential response according to income or wealth can be partly overcome by standard weighting techniques; but if no family with wealth over $10 million ever consents to an interview, there is no way that this omission can be corrected by re-weighting the sample.

The second form of nonsampling error is misreporting, which most often takes the form of under-reporting both assets and debts, either by failing to report ownership of certain assets (which may not be known to the person completing the survey), or by undervaluing those items which are recorded. The extent of under-reporting varies widely across the different types of asset. At one extreme, US studies have found that respondents on average report the value of owner-occupied houses to within four percentage points of the appraisal by real estate experts (Kain and Quigley, 1972). At the other extreme, it has been estimated that respondents under-report bank balances by up to 50–60% (Davies, 1979: pp. 248–251). For other categories of assets, such as consumer durables and privately owned businesses, undervaluation may not be due to ignorance, forgetfulness or deliberate under-reporting, but may instead reflect the ambiguous nature of the assessment—for example, whether the asset is valued as a "going concern" or on a "sell-up" or "realization" basis.

The net impact of nonsampling errors is revealed by comparing the raw aggregate figures derived from surveys with independent estimates of the balance sheet totals for the household sector, a practice now widely adopted in order to assess the relia-
bility of wealth survey data. While the balance sheet figures are themselves prone to certain types of error, they are generally regarded as more accurate, and this has prompted refinements of the survey data designed to eliminate, or significantly reduce, the discrepancy with the balance sheets.

The difficulties that arise in deriving estimates of wealth distribution from sample surveys have been addressed in a number of ways. One method of dealing with the sampling error problem is to pool information for several years, assuming that suitable data are available. Another common strategy is to adopt a dual sampling frame at the survey design stage. A high-income frame consisting of families or individuals known to have income above some cutoff on the basis of income tax, census, or other data is added to the usual sampling frame. This approach was followed in wealth surveys in the US by both the 1963 SFCC and recent Surveys of Consumer Finances (SCF). The same method was used in the 1977 (but not the 1984) Canadian SCF.

While comparisons between aggregate survey figures and independent balance sheet estimates indicate the scale of the problems associated with response bias and under-valuation, the exact contribution of each of these factors is difficult to gauge. As a consequence, attempts to reconcile the two sets of data need to make assumptions about the source of the discrepancies. It may be thought, for example, that lower response rates for wealthier households account for the low average wealth values typically observed in survey data. This may be partially rectified by re-weighting the sample using whatever additional information is available, although the sampling problems discussed above limit the extent to which this problem can be solved completely. Alternatively, it might be assumed that the entire gap is due to under-reporting; in other words, that sampling problems can be neglected and there is no response bias. As the degree of underestimation varies considerably across asset and debt types, it is natural to assume—as a first approximation—that the percentage degree of under-reporting for any particular asset or debt is the same for all wealth-holders. In order to gauge the possible impact of the errors involved, it is clearly advisable to offer calculations based on a range of alternative assumptions. Exercises of this type are reported by Davies (1979) for Canada, by Wolff (1987b, 1994) and Wolff and Marley (1989) for the US, and by Leipziger et al. (1992) for Korea.

Numerous wealth surveys are in use around the world; here we mention only some of those that have attracted the greatest attention. Recent estimates of the overall wealth

---

26 See, for example, Avery et al. (1988) for the US; Davies (1979) and Oja (1986) for Canada; Brandolini and Cannari (1994) for Italy; and Leipziger et al. (1992) for Korea.

27 Davies (1979) explored a number of alternative procedures for aligning the 1970 Canadian SCF with independent household balance sheet figures, allowing for both under-reporting and differential response by wealth level. Wolff (1987b) adjusts the asset holdings of respondents in four different US wealth surveys to reconcile the figures with balance sheet information. In both studies, the net result was a small increase in the estimates of wealth concentration. Using 1988 household survey data collected by the Korean Development Institute, Leipziger et al. (1992) suggest that adjustments based on balance sheets would substantially increase estimates of wealth concentration in Korea, due to the fact that land is probably undervalued by a significant margin, and that land holdings are distributed very unequally.
distribution in the US are available from the short-panel Survey of Incomes and Program Participation (SIPP), the Panel Study of Income Dynamics (PSID) which has followed respondents over many years, and the Survey of Consumer Finances (SCF) which is now conducted every three years for the Federal Reserve Board.\(^{28}\) SIPP provides considerable asset detail but performs less well than the PSID and SCF in terms of comparisons with independent balance sheets, and in other respects (Curtin et al., 1989). Its relatively poor performance appears to be due in part to the fact that it does not over-sample high income families. In contrast, the PSID asks a small number of questions about broad asset categories, but comes closer to the balance sheet totals, perhaps achieving better responses because of the long relationship with its respondents. The PSID may be particularly useful for those who need a good wealth variable to use as a regressor in empirical studies. The SCF provides the best information on the upper tail of the distribution, and allows comparisons of changes in wealth distribution over time.

SCF’s yielding overall estimates of the distribution of wealth in the United States were conducted in 1983, and triennially since 1989. (See Avery et al., 1984, 1988; Weicher, 1995; Kennickell et al., 1998.) In addition, they can be compared with the 1962 SFCC, which employed a similar methodology. The SCF does not have a large sample—for example, the overall sample size in 1995 was 4299; but it includes a special high-income frame intended to represent roughly the top 1% of the population, and does well in comparisons with the Flow of Funds household balance totals (Curtin et al., 1989; Avery, 1989). A comprehensive range of marketable assets is covered and, for the 1983 SCF, estimates were also provided for the present value of social security benefits and private pensions for workers above age 40. Comparisons across years indicate that overall wealth inequality changed little between 1962 and 1983, but rose after 1983: see Wolff (1994, 1995) and Weicher (1995).\(^{29}\)

There has been much interest in the distribution of wealth in Japan, in part because of the belief that the stock market and real estate booms of the 1980s fuelled a considerable increase in wealth inequality—a trend causing concern in a country widely regarded as relatively equal. The concern about possible increases in wealth inequality was reinforced by a perception that bequests were becoming increasingly important as the population aged rapidly and wealthy cohorts began to die.

\(^{28}\) Other surveys provide longitudinal evidence on wealth of the elderly. These include the National Longitudinal Survey of Mature Men (NLS) which was conducted from 1967 to 1983, the Retirement History Survey (RHS) running from 1969 to 1979, and the Health and Retirement Survey (HRS) which began in 1992–1993. Juster and Kuester (1991) find that the estimates of wealth distribution from the NLS and RHS suffer from biases due to the lack of oversampling of high income families and other limitations. The HRS appears to provide more reliable estimates (Juster and Smith, 1997).

\(^{29}\) There is some disagreement about the magnitude of the increase in wealth inequality between 1983 and 1989. Wolff (1994) reports a 5 percentage point increase in the share of the top one half percent of wealth-holders, and a rise of 0.04 in the Gini coefficient. Weicher (1995) presents calculations based on alternative weighting schemes, and slightly different procedures. He does not report the share of the top groups, but claims that the increase in the Gini coefficient was no more than 0.027.
Although Japan has several alternative sources of wealth data, until recently they all suffered from limited coverage and other problems, as documented by Bauer and Mason (1992). The annual Family Saving Survey (FSS) excluded unattached individuals and nonfinancial assets. Unattached individuals were included in the National Survey of Family Income and Expenditure (NFIE), but this survey omitted those primarily engaged in agriculture, as well as merchants, private and corporate administrators, and professionals. Farm households were covered in the Farm Household Economic Survey (FHES). Full coverage of the population was achieved in the Financial Asset Choice Survey (FACS), conducted biannually since 1988 by the Institute for Posts and Telecommunications Policy (see Campbell and Watanabe, 1996a).

As Bauer and Mason point out, special difficulties affect the measurement of financial assets in Japan. In order to avoid high income and bequest taxes, many investors shelter their assets in holding companies, whose value tends to be underestimated due to the methods used to price shares in these companies. This problem needs to be borne in mind, since the main conclusions from Japanese survey data are that (a) wealth inequality is lower in Japan than in other countries such as the US or UK, and (b) most wealth inequality is due to inequality in holdings of land and housing (Mizoguchi and Takayama, 1984; Tachibanaki, 1989; Takayama, 1991). Takayama (1992) uses the NFIE, and finds a Gini coefficient of 0.52 and a share of the top 5% of 25%; Tachibanaki (1989) uses the FSS together with supplementary data to obtain a Gini of 0.58 (see Bauer and Mason, 1992: pp. 416–417). Campbell and Watanabe (1996b) find higher proportions of both wealthy and poor households, and higher mean net worth, in the 1994 FACS compared to the 1994 NFIE or FSS, but they do not report summary inequality measures.

Given the similarities between Japan and Korea, it is interesting to note that Leipziger et al. (1992) also report a Gini value of 0.58 for gross wealth inequality in Korea in 1988. This estimate was based on an income and wealth survey of 4291 households conducted by the Korean Development Institute (KDI), and fitted a truncated Pareto distribution to the upper tail in order to overcome the problems of sampling and differential response rates. Leipziger et al. review the wide range of supplementary information available for Korea, including inheritance and estate tax records, the distribution of financial assets across households, and, most importantly, detailed data on the distribution of land holdings. They even attempt to use share register information on named individuals to estimate the top tail of the distribution of equity in public companies. An aggregate balance sheet constructed from independent sources suggested that the raw KDI data underestimate average household wealth by at least 36%, and perhaps by as much as 73%, depending on the true aggregate value of land holdings. Reconciling the survey data with middle range estimates of land values, and assuming that nonresponse and undervaluation each accounted for roughly half of the missing wealth, produced estimates of 19, 37 and 48% respectively for the share of the top 1, 5 and 10% of wealth-holders. The single most important determinant of wealth concentration in Korea is undoubtedly land ownership: for the year 1988 it is estimated that the top 25% of landowners (com-
prising 11% of the adult population) owned over 90% of the total land value (Leipziger et al., 1992: p. 46).

Other countries which have had a variety of sample surveys of wealth-holding going back to the 1960s or earlier include Canada, Australia, France and Italy. Canada’s Survey of Consumer Finance provided information for 1964, 1970, 1977 and 1984. These surveys had reasonably large sample sizes (about 14,000 in 1984, for instance), but did not over-sample high income families except in 1977. The unadjusted 1984 results indicate shares of the top 1, 5, and 10% of families equal to 17, 38, and 51%, respectively, and a Gini coefficient of 0.69. Davies (1993) points out that this survey does not match the balance sheet totals well, and does not capture the upper tail of the Canadian wealth distribution. After reviewing various external sources of information, Davies concludes that the true shares of the top 1 and 5% in 1984 were approximately 22–27% and 41–46%, respectively.

Australia had a few early isolated wealth surveys, but they did not over-sample at the top end, and are not considered to have been reliable in the upper tail (Piggott, 1984). Dilnot (1990) supplemented the figures for housing wealth in the 1986 Income Distribution Survey (IDS) with estimates of nonhousing wealth generated by the investment income method (see Section 3.4 below). He obtained shares of the top 1, 5, and 10% of 20, 41, and 55% respectively. Baekgaard (1997) refines Dilnot’s methods and obtains results suggesting a much lower level of concentration. His shares of top 1, 5, and 10% for the 1986 IDS are just 13, 32, and 46%. For France, Kessler and Wolff (1991) report a Gini of 0.71 for gross wealth in the 1986 INSEE survey, and shares of 26 and 43% for the top 1 and 5% of households.

Information on wealth-holdings in Italy is available from the Bank of Italy’s Survey of Household Income and Wealth, conducted annually from 1965 to 1984, and biannually from 1987 onwards. Many researchers have used these data to investigate savings behaviour, intergenerational transfers, and the motives for bequests (see, for example, Ando et al., 1994), but only one unpublished study has looked at wealth inequality. Brandolini et al. (1994) report top wealth shares for 1987, 1989 and 1991, with the share of the top 1% declining over this period from 13 to 9%, and the share of the top 10% falling from 45 to 39%. These data include adjustments for nonreporting of assets and undervaluation, but sample selection bias means that about 30% of net worth, and about 50% of financial wealth, is excluded from the estimates. Brandolini et al. draw attention to the relatively low share of financial assets in Italy (especially holdings in pension and insurance funds), and to the almost negligible degree of indebtedness, both of which are attributed to the backwardness of the credit and insurance markets. Applying standard decomposition procedures, they suggest that real property accounts for around 80% of wealth inequality as measured by either the Gini coefficient or the variance.

---

30 There was no Canadian wealth survey between 1984 and 1999. Statistics Canada conducted a new Survey of Financial Security in April–June 1999. It is expected to increase asset coverage relative to the SCF, including pension equity for example.
In recent years, further evidence on wealth-holding has become available in new household surveys conducted in Germany, Sweden, the Netherlands, Britain and Ireland. These new data sources have many strengths including, in some cases, a panel design (see, for example, Alessie et al., 1997, for the Netherlands). However, the surveys have been used primarily to investigate poverty, savings behaviour, labour supply, income mobility and other income related household topics. Information on wealth has not been studied in detail, and the evidence has tended to be reported in the form of means, medians, inter-quartile ratios, and standard deviations, rather than the shares of the top wealth groups or standard measures of inequality. For these reasons it is often difficult to fit the new results into an overall picture of the degree of wealth concentration across countries.

The German Socio-Economic Panel (GSOEP) was used by Burkhauser et al. (1997) to contrast wealth inequality in Germany in 1988 with US data derived from the 1988 PSID. They report a Gini value of 0.69 for net worth in Germany compared to 0.76 for the US. Bager-Sjögren and Klevmarken (1996) make use of the panel characteristics of the Swedish HUS survey to estimate wealth inequality and wealth mobility over the period 1983–1992. They obtain Gini values for net wealth in the range 0.52 to 0.59, which is lower than the figure expected on the basis of wealth tax information. This may reflect the deficiencies of the HUS data, including a small sample size (just 1150 households in 1992), the omission of individuals above age 75, and the need to impute data for a high number of incomplete records (59% of the sample).

Nolan (1991) analyses data on 3294 Irish households obtained from the 1987 Survey of Income Distribution, Poverty and Usage of State Services, which included questions on property (including farms), and certain types of financial assets and debts, but not consumer durables or life insurance-based investments. Comparison with various sources of external information led Nolan to conclude that the survey data provide reasonable estimates of the wealth of the bottom two thirds of the distribution, whose holdings are not recorded in the estate duty statistics. However, the survey is not regarded as a reliable guide to asset holdings at the top end of the wealth scale. This may well account for the small share of the top 1% observed in the raw sample data, which is very low by international standards, and also low in comparison with earlier estate-based estimates of wealth concentration in Ireland. According to Nolan, making allowance for missing wealth is likely to raise the share of the top 1% from 10 to 13–14%, and the share of the top 10% from 42 to 48%. One special feature of the Irish data is the significance of farming land, which accounts for half the wealth of the richest 1% of households.

3.2. Wealth tax data

A number of European countries levy wealth taxes. Estimates of the distribution of wealth can be derived, and coverage of the population is sometimes greater than under estate taxes. However, some assets (e.g., consumer durables) are missing and others are
generally under-valued. Perhaps for these reasons, estimates of wealth distribution from this source have been made for relatively few countries.

In the 1975 data for Sweden used by Spant (1987), about 5000 families were selected at random from the wealth tax records, and a further 3000 added from a high wealth sampling frame. Spant made various adjustments in order to correct for missing and undervalued assets. His estimates of the shares of the top 1, 5, and 10% in net wealth at market prices were 17, 38, and 54%, respectively. Lindh and Ohlsson (1998) report corresponding estimates for more recent years which show that wealth inequality decreased between 1975 and 1984, then rose, with shares of the top 1, 5, and 10% reaching 20, 41, and 58% in 1992. On the basis of Danish tax records for 1975, Spant (1982) estimates the proportion of net wealth owned by the top 1, 5, and 10% in Denmark to be 25, 48 and 65%. This suggests that wealth inequality was higher in Denmark than in Sweden, at least during the 1970s. But the Danish survey did not over-sample the rich, and is considered less reliable than comparable Swedish data.

A major feature of the wealth tax information for Sweden is the length of time over which comparable estimates can be made for the distribution of wealth. The series reported in Spant (1987) show a continuous downward trend in net wealth inequality over the period 1920–1975, with the share of the top 1% falling from 50% in 1920 to 21% in 1975, and the share of the top 10% declining from 91 to 60%. However, these data refer to the wealth valuations declared in the tax returns, rather than true market values. Converting the data for the period 1920–1975 to the market value basis used since 1975 produces the long-term trend in the share of the top 1% shown in Figure 1 below. It is evident that wealth inequality in Sweden has declined significantly over the last century, with the rise after 1984 only restoring inequality to about its level in 1970.

It is interesting to note that the 31 point drop in the share of the top 10% in Sweden between 1920 and 1975 is almost all accounted for by the 29 point decline in the share of the top 1%; indeed, over the full 55 year period, the share of those just below the top of the wealth distribution either remained constant or actually increased. The tendency for long run changes in wealth inequality in this century to be due almost entirely to a reduction in the share at the very top of the distribution is echoed in the estate-based figures for the UK (see Section 3.3 below), and was noted by Atkinson and Harrison (1978: chap. 6).

Table 1 summarises the evidence on household wealth inequality for a range of countries with survey or wealth tax data. To assist comparability across countries, the mid 1980s was chosen as the target time frame. Inequality appears to be lowest in Australia, Italy, Korea, Ireland, Japan and Sweden, with Gini values of 0.5–0.6 and shares of the top 1% of 20% or less. It has an intermediate value in Canada, Denmark, France, and Germany, where the share of the top 1% ranges up to 26%, and appears greatest in the US, where the Gini value is about 0.8 and the share of the top 1% of families exceeds 30%. Due to differences in definitions and procedures across countries, this ranking is necessarily rough. However, some confidence in the ranking is warranted, since it is broadly consistent with the careful assessments of international differences offered by a
Table 1
Selected survey and wealth tax data on the distribution of wealth across households

<table>
<thead>
<tr>
<th>Percentage wealth share of top x percent of households</th>
<th>Gini value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 5 10</td>
<td></td>
</tr>
<tr>
<td>USA 1983 (a)</td>
<td>35 56</td>
</tr>
<tr>
<td>USA 1983 (a)</td>
<td>33 55</td>
</tr>
<tr>
<td>France 1986 (b)</td>
<td>26 43</td>
</tr>
<tr>
<td>Denmark 1975 (c)</td>
<td>25 48 65</td>
</tr>
<tr>
<td>Germany 1983 (d1)</td>
<td>23</td>
</tr>
<tr>
<td>Germany 1988 (d2)</td>
<td>-</td>
</tr>
<tr>
<td>Canada 1984 (e)</td>
<td>17 38 51</td>
</tr>
<tr>
<td>Canada 1984 (e)</td>
<td>24 43</td>
</tr>
<tr>
<td>Australia 1986 (f1)</td>
<td>20 41 55</td>
</tr>
<tr>
<td>Australia 1986 (f2)</td>
<td>13 32 46</td>
</tr>
<tr>
<td>Italy 1987 (g)</td>
<td>13 32 45</td>
</tr>
<tr>
<td>Korea 1988 (h)</td>
<td>14 31 43</td>
</tr>
<tr>
<td>Korea 1988 (h)</td>
<td>19 36 48</td>
</tr>
<tr>
<td>Ireland 1987 (i)</td>
<td>10 29 43</td>
</tr>
<tr>
<td>Japan 1984 (j)</td>
<td>- 25 -</td>
</tr>
<tr>
<td>Sweden 1985 (k1)</td>
<td>16 37 53</td>
</tr>
<tr>
<td>Sweden 1985 (k2)</td>
<td>11 24 -</td>
</tr>
</tbody>
</table>


number of other researchers, including Harrison (1979), Kessler and Wolff (1991), and Wolff (1996).

3.3. Estate multiplier estimates

Estate tax records provide information similar to that recorded in wealth surveys, but the data refer to individuals rather than households, and to an unusual population: decedents sufficiently wealthy to pay (or to be assessed for) estate tax. Those who die comprise a nonrandom sample of the population. Most obviously, their age distribution differs from that of the general population. If mortality rates depended on gender and age alone, this problem could be overcome by re-weighting the sample by the reciprocal of the corresponding age-gender mortality rate (the “mortality multiplier”), so that, for
example, one decedent from a group with a 5% mortality risk is deemed to have been drawn from a sample of 20 similarly endowed living individuals. However, evidence on death rates by occupation and income suggests that mortality is higher for less wealthy individuals, substantially so at older ages. This means that the less wealthy are over-sampled in the estate tax records of any particular age group, and that different mortality multipliers need to be applied at different wealth levels. Atkinson and Harrison (1978: chap. 3) discuss in detail the construction and application of mortality multipliers, and investigate the theoretical impact of using multipliers graduated according to wealth level.

A related problem arises because those who die are more likely to have been in poor health prior to death. They may have been unable to work, or incurred larger than average expenditures for health and nursing care, or been more aware of the need to mitigate the impact of estate taxes. For all these reasons, the wealth of decedents may be lower than living persons with similar personal characteristics and life histories. However, it is difficult to assess the impact of this effect.

There are also a number of reasons why the estate tax records may be incomplete, and the reported values inaccurate. Those with estates below the tax threshold may not be required to file a tax return, and information on such individuals may be omitted entirely from the data. This is known as the problem of “missing persons”. The problem of “missing wealth” is principally due to tax concessions which cause certain types of property to be omitted from the reported valuations. Well-known examples are property settled on a surviving spouse (with no power to dispose of the capital) and discretionary trusts. Interestingly, for at least one asset category—life insurance—“missing wealth” can be negative. The valuation of life insurance policies in the estates of decedents corresponds to the amounts assured, which typically exceeds—sometimes by a large margin when young—the value of the same policies in the hands of the living. If no adjustments are made, this factor will tend to inflate the estate-based estimates of the relative wealth of smaller (and younger) wealth-holders (Shorrocks, 1981). As with survey data on wealth, independent balance sheet figures have been used routinely to assess the magnitude of the problems associated with missing persons and missing wealth: see in particular the detailed study by Revell (1967).

Considerable effort has been spent attempting to overcome these problems in order to produce reliable estimates of wealth distribution, one major attraction being the long period of time over which quality data are often available. Much of the early pioneering work in the field was carried out in Britain and the US, where wealth distribution figures have been provided intermittently over a long period of time (see, for example, the reviews of Atkinson and Harrison, 1978; Lampman, 1962). More recently, researchers have been able to build upon this earlier work to produce a consistent series of wealth concentration figures for most of the twentieth century. In Britain this series is continued on a broadly comparable basis in the annual “Series C” estimates by the Inland Revenue,
which incorporate several additional refinements, most notably adjustments to reconcile the raw data with personal sector balance sheets.\footnote{A summary of the results of the reconciliation exercise is published annually in \textit{Inland Revenue Statistics}.}

Atkinson and Harrison (1978) and Atkinson et al. (1989) report estimates of wealth concentration in England and Wales for the period 1923–1981. These show a pronounced downward trend in inequality, with the share of the top 1% declining, on average, by 0.7 percentage points per year (see Fig. 1).\footnote{The data series for the US refers to households and is taken from Table 1 of Wolff (1996). Data for Sweden are reported by Lindh and Olhsson (1998). The taxable wealth figures for the period 1920–1975 have been proportionately scaled to match the market value figures in 1975. Data for Britain refer to England and Wales during 1922–1938, to Great Britain for 1938–1977, and to the UK from 1977 onwards: see Atkinson et al. (1989), Good (1991) and Inland Revenue (1996).} The share of the top 1% of individuals in 1923 was $61\%$; by 1981 it had fallen to just $23\%$. Subsequent Inland Revenue “Series C” estimates for the UK suggest that wealth inequality in Britain has been relatively stable since 1980, with the shares of the top 1, 5, and $10\%$ of wealth-holders averaging $18$, $36$ and $49\%$, respectively.\footnote{Major changes to the estimation procedure were implemented in 1989, leading to revised estimates for the period 1976–1988. The net effect was to reduce the share of the top 1% by 2–3 points, and the share of the top 10% by 3–6 points, except for 1976 where the share of the top 10% was revised downwards from 60 to 50%. See Good (1991) for a detailed discussion of these changes.}

---

\textbf{Fig. 1.} Shares of net worth held by the top 1% of wealth-holders: 1920–1993.
The Inland Revenue also publish two additional series of figures on wealth concentration. “Series D” broadens the concept of wealth to include the capitalized value of private (mainly employment-related) pension rights, while “Series E” extends the coverage to state pension rights. As expected, pension rights are distributed more equally than marketable wealth, so the net effect is a significant reduction in the estimates of the top wealth shares. Figures for recent years suggest that the share of the top 1% of wealthholders is reduced from about 18 to 14% when private pension rights are included, and to about 11% when entitlements to state pensions are also added.

For the US, Wolff and Marley (1989) revise the earlier estimates of Lampman (1962) for the period 1922–1953, and those of Smith and Franklin (1974) and Smith (1984, 1987) for the period 1958-1976. Applying the narrow definition of wealth used by earlier authors, Wolff and Marley show that the share of the top 1% of individuals fell from 37% in 1922 to 17% in 1976, an average annual rate of decline of 0.36%—roughly half the rate of decline observed in UK. According to these estate multiplier estimates, by the 1970s the concentration of wealth in Britain and the United States was strikingly similar. The most rapid period of decline in the US was from 1939 to 1949; inequality increased a little from 1949 to 1972, and then fell sharply between 1972 and 1976 as a result of a drop in the value of corporate stock. Recent evidence, mostly from surveys, suggests that the period since 1976 has seen an increase in concentration that has taken the share of the top 1% in marketable wealth back to the level observed in the 1930s (see Fig. 1).

Wolff and Marley extend their concept of wealth to include trust funds and the capitalized value of pension rights. These modifications do not greatly affect the overall change in inequality from 1922 to 1976, but they do produce a hump in the mid 1960s (Wolff and Marley, 1989: Fig. 15.3, p. 787). Going further, and adding in an estimate of social security wealth, reduces the share of the top 1% in 1976 to just 14%, and almost eliminates the increase in inequality recorded previously for the period from 1949 through to the late 1960s. Adopting an augmented wealth concept also attenuates the increase in wealth inequality observed in recent years. When state and occupational pension rights are included, the share of the top 1% of households in 1989 is estimated to be 21%, exactly the same as the figure in 1969 (Wolff, 1995: p. 79).


Table 2 summarises the most recent data on wealth concentration obtained from careful analysis of information on estates. Comparability across countries is handicapped by different wealth concepts, estimation procedures and time periods. However, the broad conclusions of the studies cited above are that the distribution of wealth in New Zealand
Table 2

Estate multiplier estimates of the distribution of wealth across adult individuals

<table>
<thead>
<tr>
<th>Marketable wealth</th>
<th>Percentage wealth share of top x percent of individuals</th>
<th>Gini value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK 1966 (a)</td>
<td>33 56 69</td>
<td>0.81</td>
<td>Adjusted net worth</td>
</tr>
<tr>
<td>UK 1976 (a)</td>
<td>24 45 60</td>
<td>0.76</td>
<td>Adjusted net worth</td>
</tr>
<tr>
<td>UK 1985 (a)</td>
<td>18 36 49</td>
<td>0.65</td>
<td>Adjusted net worth</td>
</tr>
<tr>
<td>UK 1993 (a)</td>
<td>17 36 48</td>
<td>0.65</td>
<td>Adjusted net worth</td>
</tr>
<tr>
<td>US 1965 (b)</td>
<td>34 -- --</td>
<td>--</td>
<td>Adjusted net worth</td>
</tr>
<tr>
<td>US 1976 (b)</td>
<td>19 -- --</td>
<td>--</td>
<td>Adjusted net worth</td>
</tr>
<tr>
<td>US 1981 (b)</td>
<td>30 -- --</td>
<td>--</td>
<td>Adjusted net worth</td>
</tr>
<tr>
<td>New Zealand 1966 (c)</td>
<td>18 45 60</td>
<td>--</td>
<td>Net worth</td>
</tr>
<tr>
<td>Australia 1967–1972 (d)</td>
<td>22 45 58</td>
<td>--</td>
<td>Net worth</td>
</tr>
<tr>
<td>Ireland 1966 (e)</td>
<td>30 57 --</td>
<td>--</td>
<td>Net worth</td>
</tr>
<tr>
<td>France 1977 (f)</td>
<td>19 47 65</td>
<td>0.81</td>
<td>Net worth</td>
</tr>
<tr>
<td>Augmented wealth (including pension rights and social security entitlements)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK 1976 (a)</td>
<td>14 27 37</td>
<td>0.46–0.53</td>
<td>Adjusted net worth</td>
</tr>
<tr>
<td>UK 1985 (a)</td>
<td>11 25 36</td>
<td>0.48</td>
<td>Adjusted net worth</td>
</tr>
<tr>
<td>UK 1993 (a)</td>
<td>10 23 34</td>
<td>0.48</td>
<td>Adjusted net worth</td>
</tr>
<tr>
<td>US 1976 (b)</td>
<td>13 -- --</td>
<td>--2</td>
<td>Adjusted net worth</td>
</tr>
<tr>
<td>US 1981 (b)</td>
<td>20 -- --</td>
<td>--</td>
<td>Adjusted net worth</td>
</tr>
</tbody>
</table>

Note: "adjusted net worth" is net worth aligned to balance sheet totals.

and Australia has been more equal than that in either the US or UK throughout the twentieth century; that the degree of concentration in Ireland in the mid 1960s was similar to that in UK; that wealth concentration is lower in France than in UK or the US; and that wealth inequality in the US has recently overtaken that in the UK. These findings tally with the evidence from sample surveys, and are consistent with the conclusions drawn in the previous international comparisons undertaken by Harrison (1979) and Wolff (1996).34

34 Measured wealth inequality will decrease if all adults are paired up into family couples whose wealthholdings are not perfectly correlated. For this reason, the distribution of wealth across households is expected to be more equal than the distribution across adult individuals (see Royal Commission, 1975: pp. 95–96). However, comparison between the figures in Tables 1 and 2 for the US, France, Australia and Ireland provide little, if any, support for this prediction, except possibly for Ireland.
3.4. The investment income method

The fourth way of estimating the distribution of some types of assets is the investment income method, also known as the income capitalization method. As pointed out by Giffen (1913), if data are available on investment incomes, say from income tax records, then the corresponding asset values can be estimated by dividing the investment income figures by the appropriate rate of return. For some assets, such as savings accounts and government bonds, the rate of return can be measured with some degree of accuracy, and is also fairly uniform across wealth-holders. In other cases, however, for example corporate shares or real estate, rates of return vary considerably across investors, and perhaps between wealth ranges, reducing the reliability of the investment income method.

The investment income procedure has been applied in a number of countries. Atkinson and Harrison (1978) carefully examine its usefulness in the British context, and conclude that, although the results are less reliable than those of the estate multiplier approach, they provide a valuable check on the latter. Before World War II the method was quite popular in the US, as discussed by Harrison (1979: pp. 18–20). Walravens and Praet (1978) provided the first wealth inequality estimates for Belgium using this method, finding shares of the top 1, 5, and 10% in 1969 equal to 28, 47, and 57%, respectively. As noted earlier, Dilnot (1990) used the procedure to impute nonhousing wealth to the Australian 1986 Income Distribution Survey (IDS) data. His results indicate less wealth inequality than found for the 1967–1977 period by the estate multiplier method (see Piggott, 1984).

The sensitivity of wealth distribution estimates to differences in assumptions and procedures is illustrated by the revised figures for Australian wealth inequality obtained in the study by Baekgaard (1997). Like Dilnot, Baekgaard presents estimates based on the 1986 IDS. He adds business assets and superannuation (pension rights) to the assets considered by Dilnot, makes use of new estimates of balance sheet aggregates, and divides some assets into finer categories before applying the investment income multipliers. The result is a significantly lower level of concentration than that estimated by Dilnot. The share of the top 1%, for example, falls from 20% to 13% (excluding pensions). Baekgaard applies similar techniques to the 1993–1994 Household Expenditure Survey (HES), and finds shares of the top 1, 5, and 10% of 14, 32, and 46%, omitting pensions. Including pensions reduces these shares to 12, 29, and 43% respectively.

3.5. Direct wealth estimates for named persons

The final method of estimating wealth-holdings uses publicly available information on ownership of real estate, corporate equity, art treasures and other valuable items to assess the wealth of the very richest individuals and families. In the US for many years, Forbes Magazine and Fortune have conducted such an exercise in order to provide lists of the very wealthy, and to comment on the way that fortunes are made and dissipated. A similar exercise for residents of the UK has been carried out annually since 1989 by The Sunday Times (see Beresford, 1990).
Ch. 11: The Distribution of Wealth

This type of data presents special difficulties: the assets covered in the estimates are restricted to those most easily identified in public records; the asset valuations involve a considerable degree of guesswork; and the criteria used to decide which people should be included are somewhat arbitrary. Individuals can enter and exit the lists depending on the country chosen as their principal residence, or on whether their holdings are combined with those of related family groups outside their immediate family.

*The Sunday Times* data (available on the internet) illustrate these problems. The Queen of England was ranked in first or second place in the UK during each of the years 1989-1994, but then vanished from the top 10 wealth-holders when the royal art collection was excluded from her personal wealth. In 1998 she ranked equal 94th. Other individuals seem to have experienced spectacular year-on-year changes in their fortunes, mainly as a result of revaluing their holdings in private companies. The appearance of Robert Maxwell, the publisher, in the 1991 list of the top 10 wealth-holders in Britain, just months before events following his death revealed massive debts, is testimony to the difficulty of accurately estimating the wealth of named individuals.

While information obtained in this way is clearly limited in its population coverage, it is potentially a useful supplement to the standard sources of data, which invariably fail to capture the very top end of the wealth distribution. *The Sunday Times* figures, for instance, suggest a rapid rise in the wealth of the top 100 families during the decade to 1998, a trend not yet reflected in the UK estate data. The fact that the data form, in effect, an annual panel, and that the personal details of the individuals concerned are often well known, also offers opportunities to study wealth mobility from year to year, and the degree to which high wealth status is transmitted between generations. For instance, *The Sunday Times* of April 19th 1998 claims that 31% of the top wealth-holders in Britain inherited their wealth, compared to 57% a decade earlier. Whether this claim bears close scrutiny, and whether the finding has wider implications for wealth inequality in the UK, remains to be seen: academic researchers have been slow to exploit the potential of evidence on the wealth of named individuals.35

3.6. Portfolio composition

In addition to reporting estimates of overall wealth inequality, many studies have reported evidence on the concentration of holdings of specific types of assets, and on the way that asset portfolios vary by age, gender and wealth level. There are several reasons why this disaggregated evidence might be important. First, while certain assets like cash simply represent stores of potential consumption, others convey additional advantages. Private homes and extensive land holdings, for instance, often generate a sense of attachment and pride that goes beyond the satisfaction of owning a valuable

---

35 This kind of data has not been neglected entirely. Piggott (1984) mentions evidence on Australia's wealthiest “One Hundred” reported in *Business Review Weekly*, Leipziger et al. (1992) provides estimates of the wealth of the top shareholders in Korea; and Davies (1993) uses data on named individuals to produce the revised estimates of the top wealth shares in Canada shown here in Table 1.
asset. Similarly, direct equity holdings in stock market companies, and especially in private businesses with a small number of stock-holders, carry with them some degree of control over the lives of those associated with the firms, and hence some degree of social status not associated with bank balances, or indirect holdings in mutual funds. Information on the concentration of share-ownership in individual companies, and in all companies viewed collectively, is therefore likely to be relevant in assessing the type of society in which people live.

The second reason for interest in portfolio patterns is that observed differences across individuals are a potentially valuable source of information on household behaviour and personal preferences, for example with regard to risk. Furthermore, the rates of return on different types of assets have often diverged quite markedly over long periods of time. So, to the extent that individuals continue to hold the same physical assets and the same financial securities over time, portfolio data and asset price series can help account for changes in the distribution of the value of wealth-holdings.

Using UK estate data for 17 types of assets and debts, Shorrocks (1982) examines the way that age, gender and wealth level influence the composition of individual portfolios. Although gender differences were statistically significant, the age and wealth pattern was found to be qualitatively similar, and suggested that the 17 asset/debt components cluster naturally into five broad categories: (i) Cash and savings accounts, which increase with age as a percentage of total assets, but decrease with wealth level; (ii) Bonds and stock market equity, whose portfolio proportion increases with both wealth and age; (iii) Assets and debts associated with private businesses, which decline with age and increase with wealth; (iv) Illiquid personal assets, such as life insurance policies and household goods, whose share declines with age but shows little variation with wealth; and (v) Property assets and debts, whose share falls with age and has an inverse U-shape relationship with wealth.

The tendency for the portfolio share of corporate equity to increase with wealth, and for housing to be most important for the middle classes, is echoed in many other empirical studies: see, for example, Projector and Weiss (1966) and Wolff (1994) for data on the US; Beach et al (1981: Table 88) for Canada; and Nolan (1991: Table 9.3) for Ireland. The implications for the impact of asset price changes on wealth distribution are clear. A rise in equity prices will disproportionately benefit the higher wealth groups, and hence increase wealth inequality. In contrast, the impact of a uniform rise in property values on wealth concentration is ambiguous, since it increases the relative wealth of those in the middle of the distribution, reducing the gap with those at the top while simultaneously widening the differential over those at the bottom.

Simple theoretical models of portfolio behaviour tend to predict that individuals will hold nonzero quantities of all assets in their portfolio, but this is patently not true in practice. For example, King and Dicks-Mireaux (1982) report that only 18% of a sample

36 Arrondel et al. (1994) suggest that portfolio composition can help to identify bequest intentions: households with bequest motives will tend to hold less liquid assets and more real assets than households owning assets for precautionary purposes.
of Canadian households held more than 6 of the 12 identified types of assets, and less than 1% held more than 9.37 The composition of portfolios by value therefore depends to a large extent on the number and type of assets in which individuals choose to invest. King and Leape (1984) comment on the overall lack of diversification observed in a sample of wealthy Americans, over half of whom owned no corporate equity. They also draw attention to the pronounced age effect on the average number of assets held, which exhibits a classic hump-profile over the lifetime, peaking in the age range 40–60 and then declining significantly in retirement years.

Detailed investigation of individual portfolio behaviour poses many difficult econometric problems. Some progress is evident in the studies by Dicks-Mireaux and King (1984), Leape (1987), and King and Leape (1998), as well as recent work by Halliassos and Bertaut (1995), Blake (1996), Hochguertel and van Soest (1996), Hochguertel et al. (1997), Poterba and Samwick (1997), and Attanasio et al. (1998). Other studies, including Feldstein (1976a) and King and Leape (1986), have looked at the impact of taxation on the choice of portfolio. However, the implications of this work for the level and distribution of wealth-holdings remain largely unexplored.

3.7. Explaining national differences and changes over time

Wealth distribution figures for different countries, or different years, inevitably prompt speculation about the causes of the observed differences. Variations in wealth inequality across countries are often seen as resulting from different social traditions, general economic environments (including tax regimes and the incentives for entrepreneurial activity), and the relative importance of inheritance vis-à-vis life-cycle accumulation. For example, the incidence of owner-occupied housing differs considerably across countries. Until quite recently, one explanation offered for the high concentration in wealth in Britain was the large stock of public housing. Kessler and Wolff (1991) account for the lower concentration of wealth in France compared to the US partly in terms of the lesser importance of corporate share ownership in France, which is due to the greater share of productive capital in the hands of the public sector. As regards the impact of inheritance, faster growing economies are likely to have a higher share of aggregate wealth associated with lifetime accumulation and are therefore expected to show less wealth inequality. The greater egalitarianism of bequest practices in North America and Australia compared to Britain or Europe has also been seen in the past as a contributory factor. While some aristocratic families in Britain continue to practice primogeniture to this day, equal division of estates is the general rule in North America and Australia, even for wealthy families.

37 Sample surveys often reveal a surprisingly large number of households who claim to have little or no financial wealth of any kind. For example, 40% of a Korean sample fit into this category (Leipziger et al., 1992: p. 35). The corresponding proportion in the UK varied between 27 and 50% over the period 1973–1993 (Banks and Tanner, 1996: Table 9).
Several explanations have been proposed for variations in wealth concentration over time, in particular the downward trend in wealth inequality observed over the past century. The fact that most of the reduction in inequality has been due to the fall in the share of the top 1% suggests that estate and inheritance taxes could be influential. Wealth taxes have never raised significant amounts of revenue compared to other sources of government income, but they may have had an indirect effect by encouraging the spread of wealth ownership (or nominal wealth ownership) within the wealthiest families. For this reason, inequality in the distribution of wealth across families may not have declined to the same degree as across individuals. High rates of estate taxation have also led to the routine use of tax shelters, such as family trusts, by high wealth families. The growth of tax avoidance measures suggests that current estimates of wealth concentration may not be directly comparable with estimates produced for earlier years in this century, and that some of the observed downward trend may be attributable to this (spurious) factor.

Another likely explanation for the downward trend in wealth inequality is the growth in savings for lifecycle purposes, accumulated in the form of real estate as well as financial assets. In Britain, for instance, there has been a marked spread of “popular assets”, such as owner-occupied homes, consumer durables and, more recently, private pension fund assets, among middle and lower income households. This factor can be viewed as changing the wealth distribution via a net increase in the quantity of assets held by households. Changes in the wealth values can also result from changes in asset prices, and this will affect overall wealth inequality if the composition of individual portfolios varies systematically with wealth level, as the empirical evidence indicates. Asset price movements therefore offer a further explanation of changes in wealth concentration, particularly during stock market or housing booms or slumps. As already noted, a rise in stock market values is expected to increase wealth inequality, while a rise in house prices is likely to have a more ambiguous impact, reducing the wealth shares at both ends of the distribution.

The quantitative significance of the possible causes of changes in wealth inequality over time has been examined in most detail for Britain—by Harrison (1976), Atkinson and Harrison (1982), Harbury and Hitchens (1983), and Atkinson et al. (1989), all of whom use regression techniques. These studies show that a time trend explains much of the decline in shares of top groups in Britain, perhaps reflecting the underlying trends in lifecycle saving and inheritance practices discussed above. However, all of these studies also find that asset prices have a significant effect on the shares of the top wealth-holders. For example, Atkinson et al. (1989) decompose the substantial reduction from 31 to 21% in the share of the top 1% of individuals between 1972 to 1981. They find that the rise of share prices over this period boosted the top share by 5 percentage points. Acting in the opposite direction were a 2 point reduction due to the time trend and a 12 point decline due to the spread of popular assets. Atkinson and Harrison (1982) and Harbury

Ch. 11: The Distribution of Wealth

and Hitchens (1983) also report a regression, omitting the time trend, in which the rate of estate taxation proves to be highly significant (along with share prices). Inclusion of a time trend, however, robs the estate tax rate of its significance. While it is difficult to distinguish the effect of estate taxation from other secular factors reflected in the time trend, progressive estate and income taxation remains a plausible explanation for some of the decline in wealth inequality experienced in twentieth century Britain.

A different methodology was used by the Royal Commission on the Distribution of Income and Wealth (1979) to examine the change in UK wealth inequality from 1960 to 1972. They looked at each asset and debt component of the top wealth groups, decomposing the change in value into a price and quantity effect. In the period 1960–1972, the effects of a rise in house prices and in ordinary share values roughly cancel out, so asset price changes had little overall impact on wealth inequality. The 6 percentage point reduction in the share of the top 1% of wealth-holders was therefore attributed to lower than average growth in the quantities of assets held, in particular the quantity of life policies (which were commonly tied to mortgages, and used on maturity to repay housing loans). In contrast, almost all of the 1972–1976 reduction in top wealth shares was due to price effects resulting from a steep rise in house prices and a slump in the stock market.

Spant (1987) examines the influence of asset price changes on wealth inequality in Sweden between 1975 and 1983, a period during which real share prices rose considerably, but house prices declined. Overall, Spant estimates that price changes accounted for a 20% increase in wealth for the wealthiest 0.2% of Swedish households. In his view, this was the principal reason for the rise in the share of the top 1% of households from 17 to 20%.

Another episode during which changes in asset prices had a major impact on wealth distribution was the real estate and stock market boom of the late 1980s in Japan. In this case the price changes all contributed to an increase in inequality, and their impact was large. Using his 1984 data, Takayama (1991) simulated the effects up to 1987 and projected a rise in the Gini coefficient from 0.53 to 0.68.

The situation during the 1980s in the US appears more mixed. Weicher (1995: pp. 14–15) shows that the increase in the prices of stocks, bonds, investment real estate, and unincorporated businesses tended to raise the Gini coefficient for wealth by 0.03 between 1983 and 1989. However, over the same period there was a decline in the value of owner-occupied housing and farms, which exerted a slightly stronger, downward influence on inequality. Taken together, therefore, price changes would indicate a slight decline in the Gini coefficient over this period, although, as discussed earlier, the Gini value actually rose. Wolff (1992) uses a different, regression-based approach which suggests that growing income inequality and an increase in stock prices relative to house values contributed about equally to the rise in wealth inequality during the 1980s.

The overall conclusion is that long term factors such as lifetime accumulation behaviour, the cumulative impact of tax policies, and changing bequest practices appear to have important effects on the evolution of wealth inequality. However, the impact of
the individual factors is difficult to identify within the general downward time trend, and
further research will be required in order to assess with precision the significance of the
individual contributions. Over shorter periods it is evident that asset price changes can
cause sizeable deviations from trend.

4. Applied work on the determinants of wealth distribution

This section reviews further applied research that has improved our understanding of
the factors which combine to explain the distribution of wealth. A range of techniques
have been used in this work, including wealth-accounting or decomposition methods,
simulation exercises, and conventional econometrics. We look in turn at four categories
of contributions: (i) micro simulations of wealth accumulation and inheritance; (ii) em-
pirical work on intergenerational wealth mobility; (iii) assessments of the aggregate
significance of lifecycle saving and inherited wealth; and (iv) empirical research on
the determinants of bequests. Overall, (i)–(iii) agree in assigning an important role to
inheritance in determining the distribution of wealth, particularly its long upper tail.
This makes clear the relevance of the research reviewed under category (iv).

4.1. Simulation studies

All versions of the lifecycle saving model predict that wealth will vary with age. If
these age-related differences are quantitatively important, then a substantial portion of
observed wealth inequality may be due to the fact that people are sampled at different
points of their lifetimes. In an extreme scenario in which wealth depended only on age,
differences in wealth-holdings would not represent differences in lifetime opportunities,
and would be a purely demographic phenomenon, of no more interest or concern for
public policy than differences in age.

To determine whether this is a plausible approximation to the real world, a number
of authors have examined the distributions that would be observed in simple egalitarian
societies where all wealth differences are due to age. Atkinson (1971), for example,
considers a society in which everyone works for 40 years, retires for 10 years, and then
dies. In a benchmark case, population and earnings both grow at a constant rate of 2%
per annum, desired consumption grows at 3%, and the real interest rate is 4.5%. Under
these assumptions, the shares of the top 1, 5, and 10% of wealth-holders would be 2,
10, and 20% respectively. In contrast, Atkinson estimated the actual shares of wealth
(including state pensions) in Britain at the time to be 22, 41, and 52%.39 Uniform inher-
itations received by all at age 50 and passed on unchanged at death, and an allowance
for “propertyless women” raise the top shares to 3, 17, and 33% respectively, still far

39 These figures are very close to the “Series E” estimates for the UK in 1971 (see Inland Revenue, 1987),
but much higher than the Inland Revenue estimates for 1976 onwards (see Table 2 in Section 3). The share of
the top centile in 1993, for example, is estimated to be only 10%.
short of the estimated actual shares. Atkinson went on to establish that wealth inequality within age groups is similar to wealth inequality across the whole population, before concluding that “lifecycle factors are not a major explanation of the observed inequality of wealth-holding in Britain”.

Subsequent studies along similar lines examine more elaborate examples of pure lifecycle societies where other influences (apart from inheritance) are modelled. Oulton (1976) includes differences in earnings and rates of return, and derives expressions for the coefficient of variation (CV) and the variance of log wealth. He found that age and earnings differences together produce a CV equal to about 10% of the actual figures. Davies and Shorrocks (1978) extended Oulton’s analysis to other inequality measures and used an alternative estimate of actual wealth inequality. Age and earnings differences were found to account for 60 or 82% of actual wealth inequality as measured by the Gini coefficient or variance of logarithms, and from 45 to 89% of wealth inequality using Atkinson’s index with a range of parameter values. In their egalitarian society, the shares of the top 1, 5, and 10% of wealth-holders are 6, 17, and 27%, respectively. These figures can be raised to about 9, 28, and 45% by allowing uniform inheritance and propertyless women as in Atkinson (1971).

Flemming (1976) also introduces unequal earnings into Atkinson’s simple model, finding that the results can account reasonably well for the shape of the wealth distribution except in the upper tail. Interestingly, his top shares—5.5 and 16.5% for the top 1 and 5%—are almost exactly those obtained by Davies and Shorrocks (1978) using the Oulton model. This suggests that the differences between Oulton (1976) and Flemming (1976) are in large part a matter of how inequality is measured, and how the results are interpreted.

A similar exercise was performed for Canada by Wolfson (1977). In addition to differences in age, earnings, and savings propensities, his model includes unequal rates of return and a demographic sub-model determining family formation, divorce, remarriage, and mortality over the lifecycle of the cohort. Together, these factors appear to explain much of the observed wealth inequality, but not the concentration found in the extreme upper tail. Wolfson’s base run, for example, yields a Gini coefficient exceeding the estimated actual value, but the share of the top centile is only 11%, compared to an estimated actual figure of 20% from the 1970 Canadian SCF.

The general inability of pure egalitarian models to reproduce the extreme upper tail of wealth-holdings suggests the need to add a realistic distribution of bequests. Flemming (1979) examined the possible consequences of inheritance through illustrative calculations which suggested an important role for both accidental and voluntary bequests. Wolfson (1977, 1979) and Davies (1982) construct full simulation models to explore these issues. Beginning in 1970, Wolfson simulates the process of bequest over the next 30 years by modelling the mortality of all cohorts, and dividing the wealth of

---

40 It is interesting to note that the predicted share of the top 10% of wealth-holders in this version of an egalitarian society is actually above the recent Series E estimates by the Inland Revenue, which have ranged from 33–36% over the period 1979–1993 (Inland Revenue, 1996).
decendents among their heirs.\textsuperscript{41} Under various assumptions, the distribution of wealth predicted for the year 2000 was found to be very similar to that observed in 1970. If, instead, inheritances were outlawed, by the year 2000 the shares of the top 1 and 5% would fall from 20 and 41% to 14 and 36% respectively. Thus Wolfson's work attributes considerable wealth inequality to lifecycle factors, but also leaves substantial room for the role of inheritance.

Davies (1982) assumes that parental preferences depend on the expected lifetime incomes of children as well as on their own consumption and the number of children. Inheritances are derived for an initial generation by applying assumptions about mortality rates, estate division and other factors to an adjusted version of Statistics Canada's 1970 SCF wealth distribution. Parameter values were chosen to ensure that the predicted distribution of wealth among the living was similar to the observed Canadian distribution. The role of lifecycle factors was then assessed by running the model with its benchmark parameter values, but setting inheritances to zero. This exercise produced predictions of 9, 30, and 46%, respectively, for the shares of the top 1, 5, and 10%, compared to the actual figures for Canada of 20, 43, and 58%. In the absence of inheritances, a simulated Gini coefficient of 0.66 is obtained, compared to the estimated actual value of 0.75. The results of other simulation exercises suggested that the effect of earnings differences in Canada was somewhat less than in Britain in the 1970s, and was roughly the same magnitude as the impact of unequal rates of return or age differences. Finally, allowing a small variation across families in the rate of time preference had a strong impact on the degree of simulated wealth inequality.\textsuperscript{42}

As mentioned in Section 2, enhanced versions of the LCM have become popular vehicles for explaining observed patterns of consumption, saving, and wealth-holding. Laitner (1992) investigates a steady state model incorporating both lifecycle saving and altruistic bequests with liquidity constraints. He finds realistic wealth-income ratios and a high incidence of gifts inter vivos. Hubbard et al. (1994, 1995) and Huggett (1996) both model uncertain earnings and uncertain lifetime with imperfect annuity markets, and find that it is possible to achieve realistic overall saving rates in this way—that is, without including an explicit bequest motive.\textsuperscript{43} Huggett (1996) explores the implications for the overall distribution of wealth, as well as for the age-wealth profiles and wealth

\textsuperscript{41} An alternative approach was used by Blinder (1974), who used survey data collected by the Survey Research Center of the University of Michigan in 1960 to set up a distribution of inherited wealth. (Blinder did not report the results of his lifecycle simulation for the distribution of wealth, focusing instead on the distribution of lifetime income.) While the 1960 SRC data, and later survey data (see, e.g., Barlow et al., 1966) indicate a highly unequal distribution of inheritances, they are subject to problems of differential response and under-reporting.

\textsuperscript{42} The recent simulation study by Fuster (1997) incorporates both inheritances and lifecycle saving, and analyses the effect of social security on wealth distribution.

\textsuperscript{43} Gokhale et al. (1998) have explored a model with uncertain lifetime but certain earnings, in which the distribution of accidental bequests among heirs is modelled. Preliminary results indicate that considerable wealth inequality can be created by such a bequest process.
inequality within age groups. His simulation generates an overall Gini coefficient close to that estimated by Wolff (1987) on the basis of the 1983 SCF, but the simulated Gini coefficient values within age groups are less than the true figures, and the share of wealth held by the top centile is only about half the estimated actual figure. These deficiencies again suggest the need to pay further attention to modelling bequests. Huggett recognises the value of exploring the implications of “deeper models of the family”, but also recommends modelling entrepreneurship.

While all this work is of interest, there are several reasons why it falls short of conclusively establishing the quantitative impact of lifecycle factors on wealth inequality. First, little can be deduced from models that omit bequests and inheritances from consideration. Even if a model of a pure egalitarian society is capable of exactly replicating the wealth distribution observed in practice, this would have to be set against the fact that a pure inegalitarian society—with wealth determined entirely by inheritance—is also capable of generating the observed wealth distribution. The models which include both lifecycle saving and inheritance, and which succeed in reproducing the observed wealth distribution, address this problem. But it seems likely that more than one such model will be capable of achieving this objective, so there is an identification problem that can only be resolved by appealing to other evidence on wealth-holding. At the present time, there is no consensus on what that other evidence should be.

A further difficulty concerns the allocation of inequality contributions to factors which combine together in complex ways. One commonly used procedure is to compute the marginal impact of each factor, in other words the reduction in inequality that would result from eliminating the factor concerned. But it is very unlikely that these marginal contributions will sum to the observed degree of inequality. In the context of wealth distribution, such calculations could lead to the conclusion that lifecycle factors and inheritance each account for, say, 80% of wealth inequality. If that were the case, the results could well be interpreted as indicating that inheritance was relatively unimportant (because most of the observed wealth inequality can be “explained” by lifecycle factors), or that inheritance was the dominant influence, or that the two sources were equally significant. Whether any of these conclusions are correct remains open to question.

4.2. Intergenerational wealth mobility

As well as investigating the determinants of cross-section wealth inequality, economists have also examined the processes governing wealth mobility. This sometimes refers to

---

44 In a related exercise, Huggett and Ventura (1998) investigate the determinants of the higher saving rate in the US among higher income households. They find that heterogeneity of preferences or temporary earnings shocks are not necessary to explain the difference in saving rates. The most robust explanation is provided by two factors: age and permanent income differences on the one hand, and the social security system on the other.

45 Venti and Wise (1998) report on a very recent empirical exercise using US data which raises the possibility of decomposing wealth inequality at retirement into contribution from earnings, inheritances, health shocks, and other lifetime influences.
mobility among members of the same generation over short periods of time (see, for example, Jianakoplos and Menchik, 1997; Bager-Sjögren and Klevmarken, 1996), but more often concerns mobility across successive generations. The topic is partly motivated by an independent interest in mobility as a social phenomenon both positively and normatively distinct from inequality. However, studies of intergenerational wealth mobility are also of interest for the light they shed on the likely importance of inheritance in explaining wealth inequality; it is generally thought that a higher degree of intergenerational wealth mobility indicates a lesser role for inheritance as a determinant of wealth.

Wedgwood (1928, 1929) pioneered a method of studying intergenerational wealth mobility. Taking a sample of 99 estates in 1924–1925 worth at least £200,000, and a further sample of 140 estates in 1926 worth between £10,000 and £200,000, Wedgwood tried to establish the extent to which the wealth of these rich decedents was due to inheritance, by tracing the wills of their parents and other close family relatives. This proved unsuccessful for more than half of the female decedents. Overall, however, he identified predecessors in 169 cases, including 83 of the 99 people with estates over £200,000. Analysis of the subsample of male decedents led him to conclude that:

of the men in the upper and middle classes at the present day, about one third owe their fortunes almost entirely to inheritance (including gifts inter vivos), another third to a combination of ability and luck with a considerable inheritance of wealth and business opportunity, and the remaining third largely to their own activities. (Wedgwood, 1929: p. 163)

In other words, one-third could be classed as "inheritors", and one third as "self-made". The remaining one-third were in an intermediate category.

Concentrating for the most part on the relationship between the estates of fathers and children, Wedgwood’s methods were later refined and reapplied to data from 1956–1957, 1965, and 1973 by Harbury (1962), Harbury and McMahon (1973), and Harbury and Hitchens (1976, 1979). The results largely echo Wedgwood’s earlier findings, although there is some evidence of a decline over time in the relative importance of inherited wealth among the top wealth-holders (Harbury and Hitchens, 1979: chap. 3).

Menchik (1979) reports the results of a similar study which used data on residents of Connecticut who left estates worth at least $40,000 when they died in the 1930s and 1940s. However, instead of tracing their predecessors, Menchik looked forwards, tracing the children of the initial decedents who had themselves died by 1976. He managed to identify the estates left by 300 children who died in Connecticut, and found that 75% of them were worth at least $50,000 (in 1967 dollars).

All these wealth mobility studies suffer from the fact that the starting samples of decedents are drawn from the upper tail of the wealth distribution. The degree of relative immobility they display is not necessarily an accurate reflection of the value for the population as a whole. Menchik (1979) tried to correct for the sample selection bias. He found that, while the intergenerational correlation of terminal wealth for his sample was 0.5, correcting for sample selection yielded a coefficient of 0.8. By way of comparison,
the best estimates of the intergenerational correlation coefficient for earnings in the US indicate a figure of 0.6 (Solon, 1992; Zimmerman, 1992).

Interest in intergenerational inheritance in the UK has resurfaced in connection with the growth of owner-occupation since the end of the war. House price inflation, particularly during the 1980s, meant that housing equity became the most important component of personal wealth, and prompted speculation about the economic and social consequences when this valuable and widely spread asset was passed on to the next generation. In a major public speech in 1988, Nigel Lawson, the Chancellor of the Exchequer, welcomed the fact that Britain was about to become “a nation of inheritors”. Others pointed to the distributional effects of the transfer of housing wealth if those at the lower end of the income and wealth spectrum failed to share in this bonanza (Hamnett, 1991).

Holmans and Frosztega (1994) analyse the long run effects of housing inheritance, and emphasise the significance of the age at which inheritance is received. If only two generations are alive at any given time, inheritances tend to be received early in life. In this case, bequests of housing stock effectively provide the younger generation with a debt-free home. In contrast, the “three-generation model” envisages houses being passed to the middle generation, many of whom will have already acquired significant housing equity. This would be likely to lead to the sale of the property, with consequent effects on the housing and financial asset markets, on debt repayment, and on consumer expenditure.

Results from a specially commissioned UK survey in 1989–1991 suggest that, while there is considerable variation in the age at which inheritances are received, the three-generation model conforms reasonably well with the evidence, which shows 80% of inheritors to be above age 30, an age at which most first-time house buyers have already bought their home. The second generation beneficiaries of past house price inflation will therefore tend to be those with substantial assets of their own, although Holmans and Frosztega claim that the impact on the distribution of wealth will be relatively modest. Their findings may contain lessons for Japan which, until recently at least, shared a similar concern towards the potential adverse distributional consequences for the next generation of past rises in property values.

4.3. The aggregate importance of inherited versus lifecycle wealth

In the 1970s, a number of simulation studies based on the life-cycle model found that the implied aggregate saving rate was substantially less than real-world levels (White, 1978; Darby, 1979). This is sometimes cited as evidence that inherited wealth must be quantitatively important in aggregate terms. However, the results of such simulations can be quite sensitive to assumptions on taste and other parameters, the earnings process, the length of the working and retirement periods, and the treatment of family size effects. As Davies and Shorrocks (1978) point out, very little can be deduced about the quantitative significance of the lifecycle component of saving until (at a minimum) one has first built a complete model capable of replicating the main features of saving and wealth behav-
our observed in the real world. Davies (1982), Hubbard et al. (1994, 1995) and Huggett (1996) provide examples of models which pass the latter test, and suggest that substantial amounts of aggregate wealth can be generated from pure life-cycle behaviour.

A separate line of research performs wealth decomposition exercises which measure the impact of inheritance on current wealth by summing either past inheritances or past life-cycle saving. To clarify some of the issues involved—and to highlight an important source of divergent opinions—we recall the accounting identity Eq. (2.6) for the wealth of a family aged $t$:

$$W_t = \sum_{k=1}^{t} (E_k - C_k + I_k) \prod_{j=k+1}^{t} (1 + r_j),$$

(4.1)

which can be rewritten as

$$W_t = \sum_{k=1}^{t} S_k R_k + \sum_{k=1}^{t} I_k R_k,$$

(4.2)

where $S_k = E_k - C_k$ denotes savings from earned income at age $k$, and $R_k = \prod_{j=k+1}^{t} (1 + r_j)$ indicates the returns to investments over a number of years.

Equation (4.2) offers a decomposition of wealth into self-accumulated (or "life-cycle") and inherited components. It is, however, somewhat arbitrary, since it implicitly assumes that all the investment income from inherited wealth is saved, and that all consumption occurs out of earnings. It is equally valid to write (4.1) in the form:

$$W_t = \sum_{k=1}^{t} (E_k - C_k + r_k W_{k-1}) + \sum_{k=1}^{t} I_k,$$

(4.3)

and to identify the first term, which cumulates forward all saving out of income, with lifecycle saving, and the second component with inheritance. Assuming positive rates of return to investment, the difference between the two decompositions is seen to lie in the smaller contribution of inheritance in Eq. (4.3) compared to Eq. (4.2).

Equation (4.3) might be thought to yield a lower bound to the contribution of inheritance, since it includes savings from returns on invested inheritances in the "life-cycle" component; but if one goes further, as in the Meade Eq. (2.4), and assumes that some portion of inherited wealth is consumed directly, and that this consumption exceeds the saving from interest on inheritances (i.e., $\beta > 0$ in Eq. (2.4)), then the contribution of inheritance to current wealth would be smaller still. The fact that one can generate different wealth decompositions depending on the way in which lifecycle saving is defined, and the way consumption depends on earnings, investment income, and wealth, illustrates again the need to consider the issue within the framework of a complete model of saving behaviour. The failure to recognise this fact is one of the inherent limitations of
pure accounting approaches to the decomposition of wealth into its various contributory influences.

The best known application of the accounting approach is the study by Kotlikoff and Summers (1981), who use Eq. (4.2) to decompose wealth into its lifecycle and inherited components. The present value of inherited wealth in the US was estimated in two different ways. One method sums the amounts received as gifts or bequests by the current generation. This yielded a figure for the present value of inheritances equal to 46% of current household wealth. A second exercise, which they considered more reliable and whose results have been more controversial, cumulates the difference between after-tax labour earnings and consumption expenditure in the US over time, in order to get an independent estimate of aggregate lifecycle wealth. This yielded a figure of 19% of current wealth due to past lifecycle saving. The residual, totalling 81% of current wealth, is attributed to inheritance.

The idea that 81% of current wealth is inherited is firmly dismissed by Modigliani (1988a, 1988b), who argues instead that about 80% is due to lifecycle saving, and only 20% to inheritance. Modigliani (1988a: p. 28) suggests that the 61% point gap between the Kotlikoff and Summers figure and his own estimate can be broken down as follows:

(i) 14% of wealth is accounted for by the fact that Kotlikoff and Summers treat household expenditure on durable goods as consumption rather than as saving. This has the effect of reducing the lifecycle saving component, thereby inflating the residual inheritance term.

(ii) 31.5% of wealth is due to capitalization of inheritances; in other words, to the fact that Kotlikoff and Summers use the decomposition Eq. (4.2) instead of Eq. (4.3).

(iii) 15.5% of wealth is due largely to parental support for dependent children over the age of 18, which Kotlikoff and Summers treat as a form of bequest, on the grounds that it represents investment in their human capital.

It seems reasonable to accept part of Modigliani’s criticisms, by agreeing not to treat all expenditures on dependents over 18 as bequests, and by recognizing lifecycle saving in the form of durables. This position is consistent with that of Blinder (1988), who carefully reviews both sides of the debate. On this basis, inherited wealth accounts for perhaps 20–30% of total wealth in the US if it is not capitalized, and for about 50–60% if it is. The earlier comments regarding Eqs. (4.2) and (4.3) indicate that the capitalization issue turns on the marginal propensity to save out of investment income (including capital gains). A broad brush figure of one half yields a final estimate of 35–45% for the contribution of inheritance to aggregate wealth, which we believe to be a reasonable rough estimate.

Kotlikoff and Summers (1981) reported a share of 52%, but this contains an error corrected in Kotlikoff and Summers (1988).

Ando and Kennickell (1987) also find a high share of lifecycle wealth in the US—between 80 and 85% of net worth. Like Kotlikoff and Summers, they cumulate estimates of past lifecycle saving. However, their lifecycle accumulation covers all saving out of disposable income, not just saving out of labour income.
The conclusion that Modigliani's estimate of the importance of inheritance needs to be revised upward receives support from a variety of other evidence:

1. Modigliani (1988a) and Hurd and Mundaca (1989) review sample survey evidence which indicates that inherited wealth accounts for 20% or less of total wealth (without capitalizing). But it is well-known that sample surveys are subject to nonresponse and under-reporting, and that these are particularly severe for the very rich (for whom inheritance is more important). Taking recall bias into account as well, it seems clear that the survey evidence yields a lower bound on the importance of inheritance.

2. The direct estimate by Kotlikoff and Summers (1981) of the contribution of inherited wealth (46%) is at the margin of the 35–45% range we consider most reasonable.

3. Wolff (1997) performs a simulation based on US cross-section data from the 1960s to the 1990s which generates estimates of life-cycle saving and inheritances (the latter obtained by applying age and gender specific mortality rates to the observed distribution of wealth among the living), and residual estimates of gifts inter vivos. He concludes that approximately one third of the wealth individuals accumulate by retirement (or, indeed, over the whole lifetime) comes from each of these three sources. That is, about two thirds of wealth comes from gifts or bequests.

4. Laitner (1992) presents a model incorporating both life-cycle saving and inheritances, based on an altruistic bequest motive and including liquidity constraints, which uses US data. He obtains shares of inherited wealth ranging from 58 to 67%.

5. Lord and Rangazas (1991) and Lord (1992) present simulation models for the US including both life-cycle saving and inheritance, modelling the latter as altruistic in the first article and as based on exchange in the second. While they find that the existence of bequests does not raise aggregate saving rates greatly, the bequest share in overall wealth centres around one half, looking across their various runs and taking the midpoint between the two capitalization procedures.

6. Studies for other countries indicate a greater share of inherited wealth (see, for example, Kessler and Masson, 1989, for a review of French studies). Canadian evidence is perhaps of particular interest, given the similarities between Canada and the US. Davies and St-Hilaire (1987: pp. 107–108) apply an accounting approach to Canadian data, finding a 35% share for inherited wealth without capitalizing, and a 53% share when inheritances are capitalized. In an approach which takes behavioural impacts into account, Davies (1982) finds that equilibrium wealth declines 42% in his simulations if bequests are outlawed. These figures are again consistent with our estimate of a 35–45% share of inherited wealth when account is taken of saving out of investment income.

7. Gale and Scholz (1994) used 1983 SCF data on gifts inter vivos, trust accumulations, and life insurance payments to children to estimate "intentional transfers" in the US. They estimate that such transfers plus bequests account for at least 51% of net worth. Overall, work on the aggregate importance of intergenerational transfers reinforces what has been learned through simulation studies of lifecycle accumulation and empiri-
Ch. 11: The Distribution of Wealth

cal studies of wealth mobility. The conclusion that, as economists and others have long believed, processes of inheritance play an important part in explaining the distribution of wealth, finds significant support in the work which has emerged from the debate sparked by Kotlikoff and Summers.

4.4. Empirical studies of transfer behaviour

There has now been a considerable amount of empirical work on the determinants of gifts and bequests at the micro level. Much of this work has focused on testing the altruistic model of transfers, and some of it has tested the exchange model as well. Discriminating between these models has important implications for the distribution of wealth. As we saw in Section 2, transfers have a strong equalizing aspect in the altruistic model which is absent under exchange. In addition, private transfers have a tendency to offset public transfers under altruism, which would suggest, for example, that the effects of the rise in social security wealth on the distribution of wealth in this century is more apparent than real. However, there is more to learn about transfer behaviour than simply whether it is based on altruism or exchange. For example, the degree to which bequests are typically divided equally or unequally, and the size of the (parental) income elasticity of demand for bequests, are important for long run analysis of the evolution of wealth distribution, as we saw earlier. The empirical literature on transfer behaviour helps to provide answers to these questions. In view of the emphasis placed on borrowing constraints and precautionary saving in the recent literature on consumption behaviour, it is also interesting to know whether transfers are timed to alleviate liquidity constraints on the part of recipients.

Some of the studies have looked at gifts, others at bequests. There are important differences in terms of possible transfer motivation and the significance of the behaviour for the distribution of wealth. For bequests, there is a question of whether transfers are intentional. Bequests can be “accidental”, that is merely the unintended consequence of the desire by the living to hold assets in bequeathable form (which could arise due to imperfect annuity markets, or a “capitalist” motive for accumulation). Gifts and bequests may also differ in their implications for saving and wealth-holding. Most obviously, in order to make a bequest a donor may hold substantial bequeathable wealth for a long period of time. A programme of gifts, in contrast, will likely give lower wealth for retirees and greater net worth for younger people, which could reduce wealth inequality relative to a no-transfer world even if gifts are not altruistically motivated.

Since gifts are clearly intentional, their relative size compared with bequests is itself an interesting question. It puts a lower bound on the importance of intentional transfers.

48 As shown by Cox (1987) these two models can be obtained as alternative regimes of a single model in which parents are altruistic towards their children. Such parents may be at a corner where they do not desire to make gratuitous transfers, but still make payments to children in exchange for services. This is the exchange regime. However, parents who are sufficiently well-off, or whose children need more help, will make additional transfers, yielding the altruistic regime.
Using survey data from the President's Commission on Pension Policy (PCPP), Cox and Raines (1985) find that three quarters of transfers are in the form of gifts. Even given the likelihood that bequests are underestimated in this survey, this is impressive evidence that much of the intergenerational transfer flow is intentional. Using the 1983 SCF data, Gale and Scholz (1994) estimate that in excess of 38% of transfers were in clearly intentional form (gifts, trust accumulation, and life insurance payments to children).

Hurd (1987) attempts to throw light on whether bequests are intentional by comparing rates of decumulation of bequeathable wealth in retirement by people with and without children. Using the Longitudinal Retirement History Survey (LRHS) over 1969–1979, he finds that rates of decumulation are not significantly lower for people with children. In part this might be explained by the fact that some childless retirees have reasons for making intentional bequests, for instance to friends, relatives, or charities. Hurd (1989) implemented a model using the same data (but restricted to single parents) in which the demand for bequests was estimated. Statistically significant desired positive bequests were only found in one specification, and were small.

Bernheim (1991) argues that evidence of a desire for intentional bequests can be found in patterns of demand for life insurance in the US. He looks at the life insurance and annuity holdings of individuals and families in the LRHS, and finds that demand for term life insurance is greater among those who are “over-annuitized” by high social security entitlements. This higher demand increases the value of wealth left at death, which is by definition an act of intentional bequest.

Other evidence also appears to suggest that a significant portion of the population has an intentional bequest motive. Laitner and Juster (1996) examine a sample of retirees covered by the US academic pension plan, TIAA-CREF. They find that 46% of those with children report that leaving an estate is “very important” or “quite important”. Interestingly, 23% of retirees without children also regard leaving an estate as important. Laitner and Juster also point out that 73% of joint annuity holders with children select a guaranteed annuity. This ensures that, if both parents die within a specified number of years (10, 15, or 20), some wealth will be left for heirs. Again, as in the case of Bernheim (1991), we have what are by definition intentional bequests.

While there is therefore evidence that some parents make intentional bequests, studies like Laitner and Juster (1996) which draw on attitudinal questioning find that the importance placed on bequests ranges greatly across families. That families may be quite heterogeneous in their bequest behaviour is also a conclusion from the literature which has tested alternative models of intentional bequests.

The two leading approaches to explaining intentional transfer behaviour are the altruistic and exchange-based models. As noted in Section 2, a simple model assuming equal

Gale and Scholz estimate that adding in gifts omitted from the SCF would raise intentional transfers by about one third.
parental altruism towards all siblings makes a number of strong testable predictions. These may be summarized as follows:

1. Transfers will not occur in all families or for all children. The probability that a child will receive a transfer should increase with parental income and decline with the child's earnings.

2. In families where transfers are made, a child's transfer should increase with parental income and decrease with child's earnings. The latter point implies that gifts and bequests should compensate if a child's earnings are low relative to the parent or siblings.

And, in view of these first two points:

3. The incidence of equal division of transfers between siblings should be low.

4. Across families, total transfers should decline with children's average earnings.

5. Total transfers should rise with parental income. The relationship is expected to be nonlinear, with most parents at low incomes leaving nothing and the incidence of positive bequests rising sharply at higher income levels. (This prediction rests on the plausible assumption that low income parents cannot force negative transfers on their children.)

Exchange-based models also predict that the likelihood of a transfer being received, and the amount, should be positively related to parental income. However, the exchange model differs in allowing a positive impact of child's earnings on the amount of transfer received. It also predicts that transfers should increase with the services provided by the child to the parent, which are generally referred to as "attention".

The simplest way to test the altruistic model is to look at patterns of estate division. Assuming that compensatory transfers are not exclusively undertaken by means of gifts, altruism predicts unequal division of estates. The results of Menchik (1977, 1988), Wilhelm (1996) and others, however, indicate that equal division is the dominant practice, at least in the United States. 50 Menchik (1988) reported exactly equal division in 84% of estates in a study based on probate records from Cleveland, Ohio. 51 In his data based on federal estate tax returns, Wilhelm (1996) finds equal division in 69% of estates with two or more heirs. Approximately equal division (within 2%) occurs in 77% of the estates.

While the incidence of equal division of estates is high, it is interesting to know what determines the division of bequests in the 15–25% of unequal cases. Tomes (1981) found a significant negative impact of recipient's income on bequest. Further, he found that impacts of parental schooling on child's schooling and bequests were also in line

50 Wedgwood (1928, p. 48) reported contrary evidence for the wealthiest families in Britain during the early part of the twentieth century. More recent UK data analysed by the Royal Commission (1977: chap. 8) showed unequal division in half the cases examined, but the degree of unequal division was not systematically related to the size of the estate, as might be expected if primogeniture was widely practiced by the wealthy. Horsman (1978) found a higher frequency involving roughly equal division (about 75%). None of his cases of unequal division was clearly linked to primogeniture, although almost all favoured male heirs.

51 Menchik's sample was drawn from the same frame as that used by Tomes (1981). Tomes relied on reports by heirs of the amounts they had received rather than using the probate records. His finding of equal division in only 21% of cases is a reflection of the degree of reporting error by respondents.
with predictions from the altruistic model. Wilhelm (1996) finds a significant (although small) negative impact of recipient’s income on bequest in cases where estates were divided unequally. A small compensatory effect could be explained either by a minority of parents behaving in line with the simple altruistic model, or by altruistic forces being present in most families but largely offset by other influences—for example the need to avoid work disincentive effects on heirs.

Frequent equal division of estates would be consistent with altruism if compensatory behaviour were accomplished instead via gifts inter vivos. Considerable attention has focused on the study of gifts. Until recently, this work appeared to show that altruism was eclipsed by exchange (Cox, 1987; Cox and Rank, 1992). However, more recent studies have found evidence suggesting some altruistic motivation for gifts.

Cox (1987) studied the determinants of gifts using data collected by the President’s Commission on Pension Policy (PCPP) in the US. He found that while the probability of receiving a gift was negatively related to recipient’s income, conditional on a gift occurring the amount received increased with the donee’s income. This rejects the altruistic model but is consistent with exchange-based transfers. It could be, for example, that higher income children charge a higher “price” for their services, and that parents’ demand for these services is inelastic. Cox also found that, ceteris paribus, a family with more children receives less transfers from other families. This runs counter to altruism—since more children spell greater needs—but agrees with the exchange model, where more children imply the likelihood of supplying fewer services to other families. Finally, female headed families received larger transfers, and families with a married head received smaller transfers, ceteris paribus. The latter findings are consistent with the notion that unmarried children and women supply more attention.

Direct observations of parental income are not provided in the PCPP data, so Cox (1987) relied on mean income in the recipient’s neighbourhood as a proxy for parental income. Since this is an imperfect measure, it is possible that the positive impact of recipient’s income on the gift amount is spurious. It could reflect correlation between recipient’s income and unmeasured components of parental income. Cox and Rank (1992) provide a check on this using the National Survey of Families and Households (NSFH), which does have a direct measure of parental income. Results are similar to those of Cox (1987). However, independent studies using other datasets have disputed these findings.

Altonji et al. (1992) studied the degree of altruistic linkage between parents and their “split-off” adult offspring in the Michigan Panel Study of Income Dynamics (PSID). They looked at the extent to which the food consumption of either parents or children moved in sympathy with income shocks to the other party. The altruistic model predicts a strong sympathetic movement, reflecting adjustments in underlying transfers. Altonji et al. found that although the induced changes in consumption were in the predicted direction, they were small. This may indicate that only a minority of parents operate in the altruistic regime, or that other influences (for instance, the need to avoid work disincentive effects on recipients) limit altruistic behaviour.
Altonji et al. (1997) use the transfer supplement to the 1988 PSID to examine the determinants of attention and transfers between parents and children. Like Cox (1987) and Cox and Rank (1992), they find that parental income has a positive impact on the probability and amount of transfer. However, in contrast to Cox, the child’s income is negatively related to the transfer amount. Further evidence of a negative effect is provided by Rosenzweig and Wolpin (1993, 1994), who find that parental assistance to young people attending college is substantial, and increases as the child’s income declines. McGarry and Schoeni (1995), McGarry (1996), Dunn and Phillips (1997), and Dunn (1997) all report similar results.

Dunn (1997) is able to control for a variety of biases that may have affected earlier work by using panel data from the National Longitudinal Survey of Labor Market Experience (NLS). Not only does this dataset provide direct measures of parental income, but its panel aspect makes it possible to control for fixed effects. An investigation of the response of transfers to transitory as well as permanent incomes is performed. As set out by Cox (1990), both altruism and exchange suggest that the likelihood and amount of transfers should be positively related to both the permanent and current income of parents. Also, the likelihood of transfer should be negatively related to the child’s permanent and current income. Only with respect to the amount of transfer is there a difference in prediction, with altruism predicting a negative impact of permanent income, and exchange allowing a positive effect. Under both altruism or exchange, the transfer amount should be negatively affected by current income, since it is efficient to make transfers to the child when liquidity constraints are binding for the child, which is more likely to occur when the child’s current income is low.

Like Cox (1990), Dunn (1997) finds that parents time their transfers to correspond with periods when the child’s current income is low. Thus gifts are likely to have a significant impact on relieving the liquidity constraints faced by the young. This effect has interesting implications for saving behaviour, implying less need for precautionary saving, for example. Dunn (1997) finds that transfer amounts are negatively related to the child’s permanent income, as well as current income, which is consistent with altruism.

Returning to bequests, it may be noted that altruism has implications for wealth distribution even if transfers on death are never used to achieve inter-sibling compensation. Compensation relative to the parent, rather than other siblings, could still be a goal pursued via bequests.

There is evidence that total transfers are positively related to parental income and negatively related to children’s average earnings across families. In other words, bequests do seem to be used to achieve some smoothing of resources between parents and children. Wilhelm (1996) finds such effects in his tax return data, although the significant negative impact of children’s average earnings in his results is not large. As mentioned earlier, Laitner and Juster (1996) present results based on a survey of retirees covered by the US academic pension plan, TIAA-CREF. While the sample is not representative of the US population, it is of special interest since it focuses on a group which would be likely to make voluntary nonhuman transfers on the basis of the altruistic model.
Lower income families will often be at a corner solution where all transfers are made in human form.) For the sample as a whole, the altruistic model does not predict behaviour well. But for the half of households who intend to leave estates, behaviour is in accordance with some degree of altruism. In these households, bequeathable wealth is highest for those with the lowest assessment of their children's likely earnings, and levels of retirement-age net worth are higher by several hundred thousand dollars per household than for families where bequest is not considered important.

Further evidence on bequest behaviour which is consistent with a role for altruism is provided by Menchik and David (1983) who studied a sample of decedents in Wisconsin from 1960–1964. Inheritance records were matched with income tax returns from the period 1946–1964, making possible a very good study of the relationship between lifetime income and estate left on death. (Unfortunately, there was no information on heirs' incomes.) Menchik and David, who used a spline relationship, found that the elasticity of demand for bequests increased with income. In fact, it was below unity for the bottom 80% of the population, but in the range 2–4 for the top 20%. Such a pattern is predicted by the altruistic model, as noted earlier.

Exchange models predict that gifts and bequests should be positively related to the amount of "attention" supplied by children, as measured by the frequency of visits, telephone calls, letters, etc. However, Tomes (1981) found that the attention provided had a significant negative effect on the size of bequest received by a child. He also estimated an OLS "supply of visits" equation, finding an insignificant negative coefficient on inheritance received.

Bernheim et al. (1985) argue that Tomes' equations are misspecified because they fail to take into account the simultaneous determination of attention and bequests. They find that parents' bequeathable wealth has a significant positive effect on attention supplied by children in two stage least squares estimates using the LRHS. Nonbequeathable wealth has an insignificant impact on attention.

In summary, empirical work on transfer behaviour has found that at least half of intergenerational transfers are intentional since they are made in the form of gifts, life insurance payments and the like. There is also evidence that bequests are to some extent intentional. It is therefore interesting to ask why parents make intentional transfers. Broadly speaking, the evidence indicates that both altruism and exchange may contribute to the explanation, although the models investigated so far do not explain the bulk of the variation in transfers. Parents make larger gifts to children with lower earnings, and appear to save more for bequest if their children are on average lower earners. These observations suggest that altruism has some role. On the other hand, both gifts and bequests appear to increase with the amount of attention children give their parents, providing evidence of exchange. There is a high frequency of equal division of bequests, and the income elasticity of demand for bequests on the part of parents is low for most of the population, but in excess of unity for about the top fifth. Finally, we learn that gifts are more likely to be given, and are larger, when children are suffering temporarily low incomes or are investing in education. Whatever the underlying purpose of transfers,
they appear timed for the maximum effect, providing reassurance that the economic approach to the analysis of transfer behaviour is worth pursuing.

5. Summary and conclusions

This chapter began by noting some of the most important facts about personal wealth-holdings: the distribution of wealth is more unequal than income, and has a long upper tail; financial assets and inherited wealth are both more concentrated than wealth as a whole; many households never accumulate much private wealth, even in rich countries; and wealth inequality has been on a downward trend for most of the past century. These are the kinds of evidence on wealth-holdings which we have tried to document and explain in the course of our survey.

Progress in explaining wealth distribution was first made via stochastic models, which accounted for the fact that the distribution of wealth is approximately lognormal, and its upper tail is approximately Pareto. Other early theoretical contributions used simple models of savings and inheritance behaviour to understand the interplay of basic forces that lead to high wealth concentration and a declining trend over time. This concern with the overall features of wealth distribution has been largely abandoned in recent research, which focuses instead on the causes of individual differences in wealth-holdings.

One reason for the change of emphasis is a recognition of the growing importance of wealth-holdings acquired for retirement purposes. This is reflected in the central role now assigned to the lifecycle saving model (LCM), and the numerous ways in which the LCM has been refined and extended. When allowance is made for uncertainty, liquidity constraints, precautionary saving, and the effects of social security and means-tested benefits, the LCM is able to account for such observed features as the low positive saving of many households and the slow rate of dissaving in retirement. In simulation exercises, the model is able to generate overall wealth inequality similar to that observed, but not the high concentration of wealth in the upper tail. Differences in ages, earnings, returns to investments, and saving rates, all appear to make significant contributions to inequality in lifecycle wealth.

Inheritances and other forms of intergenerational transfers can be incorporated within the LCM framework by including a bequest motive, although most of the behavioural models of intergenerational transfers abstract from the detailed pattern of lifetime consumption. Much work has been done in an altruistic framework, which captures not only gifts and bequests, but also parental investments in the human capital of children. The altruistic model correctly predicts that gifts and bequests will be relatively more important in wealthier families, but its prediction that intergenerational transfers will strongly compensate for earnings differences between siblings is not borne out in practice. Alternative models, emphasizing exchange between the generations and strategic behaviour, are also consistent with money transfers being relatively more important.
James B. Davies and Anthony F. Shorrocks

at higher income levels, but avoid the unrealistic prediction of strongly compensatory transfers. Other evidence showing that gifts are strongly (and positively) correlated with periods when children's earnings are temporarily low, supports the view that intentional transfers are important, and suggest that, whatever their motivation, transfers help to alleviate liquidity constraints on the young.

The implications of behavioural models of intergenerational transfers for the distribution of wealth have received little attention, although it seems likely that altruistic motives cause inheritances to have an equalizing impact on wealth inequality, while the reverse is true for exchange or strategic motives. Other theoretical studies with less explicit behavioural foundations have examined the effect of differential fertility, assortative mating, and the pattern of estate division on long-run wealth inequality, and have shown that equal division of estates, and low correlation in the wealth of husbands and wives, can be powerful forces acting towards wealth equalization in the longer run.

Empirical evidence on wealth-holdings is sometimes obtained from estate or wealth tax records, although these sources are becoming less important. Survey information on wealth-holdings is now available in many countries, and wealth details are being recorded in an increasing number of panel datasets. Surveys are very valuable sources of data, but they frequently suffer from under-reporting of assets and higher nonresponse rates among the wealthy. For this reason, estate and wealth tax data probably yield more reliable information on the upper tail of the distribution. Estimates of the wealth of named individuals provides another potentially valuable source of data on the very richest families, but they have not been given much attention to date by academic researchers.

The deficiencies of survey data on wealth mean that the results need to be treated with caution. But the weight of evidence strongly supports many long-standing beliefs. All studies agree that wealth is more highly concentrated than income, even when pension entitlements are included in the definition of wealth (augmented wealth in our terminology). Data constructed for the US, UK and Sweden confirm the long run decline in wealth inequality over the period 1920–1975. After 1975 the evidence is more mixed, and wealth inequality in the US seems to have moved upwards along with income inequality. Another widespread empirical finding is the high proportion of families that claim to have little or no financial wealth of any kind. It also appears that asset portfolios, even those of the wealthiest households, are not very diversified.

Corporate stock is held disproportionately by the wealthiest families, so wealth inequality is expected to rise during stock market booms. In contrast, middle class families are expected to benefit most from increases in house prices. Both of these predictions are confirmed by the data, and may account for part of the recent rise in wealth concentration in the US. The reasons for the general downward trend in wealth inequality in the twentieth century are more difficult to determine with precision, but the spread of "popular" assets (especially houses and consumer durables) seems to have been an important factor. More equal estate division and progressive taxation may also have been influential.
A recurring theme in wealth distribution analysis concerns the importance of self-accumulated wealth relative to inherited wealth. We have reviewed a number of sources of evidence on this issue, including survey data on inheritances; simulations of wealth distribution which incorporate both life-cycle saving and inheritance; and studies of intergenerational wealth mobility. On the basis of this evidence, we conclude that inheritance probably accounts for about 35–45% of aggregate wealth in the US. Simulation studies indicate that a version of the lifecycle saving model which incorporates uncertain earnings, uncertain lifetime and liquidity constraints (but not intentional bequests) can reproduce the observed shape of the wealth distribution over most of its range, but accounts for only about half the share of wealth held by the top centile. This suggests that intentional bequests remain an essential ingredient in any credible explanation of the long upper tail of the wealth distribution, a conclusion corroborated by other evidence such as the finding that about half of the wealthiest people in the UK and US owe a great deal to inheritance.

While we have attempted to cover most of the important issues concerning the distribution of personal wealth, a number of topics have been omitted or given short shrift. Some of these relate to subject matter covered in detail elsewhere in the Handbook. Others reflect gaps in the literature. For example, we have touched upon estimates of the distribution of wealth which incorporate private pensions and social security wealth, but have not given this topic the attention it deserves. Estimates of private pension wealth require detailed information about pension entitlements which typically differ across, and even within, employers. Information on contributions and benefits is more readily available for state pension schemes, but estimating social security wealth presents its own special difficulties. Not the least of these is the fact that longevity rises significantly with income and wealth. Thus while it often appears that social security benefits are strongly redistributive, this is offset by the lower expected benefit period for low income retirees. Detailed analysis of the distributional implications of social security wealth would also need to embrace the large and controversial literature on the impact of social security schemes on private saving which began with Feldstein (1974). Taking account of the induced change in intergenerational transfers would complicate the analysis still further, so this topic presents a major challenge for researchers in the future.

Another area we have tended to neglect is the economic links between wealth distribution and demographic factors relating to mating, fertility, mortality, and family formation and dissolution. The individual chapters in Volume 1A of the Handbook of Population and Family Economics (Rosenzweig and Stark, 1997) provide excellent introductions to these topics, although here again the implications for wealth distribution have only been studied in a fairly cursory manner. One interesting line of future enquiry is the evolutionary modelling of bequest behaviour which has been provided by some economists (e.g., Chu, 1991; Bergstrom, 1994), building on the insights of anthropologists and biologists. This work points out that male offspring of higher income groups may get higher bequests if sons on average produce more grandchildren than daughters. This difference in reproductive fitness may occur where there is formal or
informal polygyny, conditions which have prevailed in almost all human societies, with the notable exception of the West in the last two centuries. Primogeniture may well be the optimal estate division policy when there are increasing returns in the process which converts bequests into grandchildren.

There is clearly a need for much more research on most of the topics we have covered. This applies to description, theory, simulations, and conventional empirical work. Measuring the distribution of wealth will always be subject to error and uncertainty, but this can be reduced in the case of surveys by research on the determinants of nonsampling error, by over-sampling the wealthy, and by careful alignment with independent balance sheets. Adoption of a common framework in different countries, along the lines that have been developed for income distributions, would improve the scope for comparative studies. Where possible, we also strongly encourage official agencies to provide consistent annual estate multiplier estimates of wealth inequality. The availability of such data in the UK has facilitated econometric and other studies which cannot be undertaken in the US, for example, where estate multiplier estimates are only available for isolated years. Another underutilised resource is the lists of the wealthy provided by journalists and others, which could be used more regularly to check and, if necessary, revise estimates of wealth-holdings in the extreme upper tail.

Although the lifecycle model has been extensively developed, and some progress has been made in linking the LCM with intergenerational models, important weaknesses remain. Most models assume riskless investment and a uniform rate of return. The riskiness of investment needs to be recognised, and its contribution to wealth inequality assessed more carefully. The acquisition of great fortunes by entrepreneurs and investors is another feature which needs to be modelled, since such explosions of individual wealth are clearly an important source of renewal for wealth inequality.

Some important recent secular trends pose challenges for research on the distribution of wealth. Up to the mid 1970s the downward trend in wealth inequality seemed relentless, and was mainly attributable to a decline in the observed share of the very rich. More recently, however, the shares of the richest groups in the population appear to have stabilised (Sweden, UK) or began to rise (US). We do not as yet have a good understanding of the reasons for this, although growing income inequality, rising share prices, the impact of social programs, and increasing use of tax avoidance measures, are all candidate explanations. The contributions of important demographic trends such as rising divorce rates, declining retirement ages, and rising longevity are also poorly understood, although the last two factors are certain to become more significant in future years. It seems likely that interest in wealth-holdings and wealth distribution will increase as the consequences of rising demand for assets for lifecycle saving purposes is seen more clearly in asset markets and changing asset prices.
References


Dunn, T.A. (1997), The distribution of intergenerational income transfers across and within families, Economics Department and Center for Policy Research, Syracuse University, mimeo.


Fuster, L. (1997), Is altruism important for understanding the long-run effects of Social Security? Universitat Pompeu Fabra, mimeo.


Gibrat, R. (1931), Les Inégalités Économiques (Sirey, Paris).


Huggett, M. and G. Ventura (1998), Understanding why high income households save more than low income households, University of Western Ontario, mimeo.
Ch. 11: The Distribution of Wealth

Kimball, M.S. (1990), Precautionary saving in the small and the large, Econometrica 58: 53–73.


Lyons, P.M. (1972), The distribution of personal wealth in Ireland, in A.A. Tait and J.A. Bristow, eds., Ireland: Some Problems of a Developing Economy (Gill and Macmillan, Dublin).


McGarry, K. (1996), Inter vivos transfers and intended bequests, University of California at Los Angeles, mimeo.


