12 On the Role of Social Security as a Means for Efficient Risk Sharing in an Economy Where Human Capital Is Not Tradable

Robert C. Merton

12.1 Introduction

In "A Framework for Social Security Analysis," Diamond (1977) describes a number of possible reasons why one should have a program similar to the present social security program. One reason given is market failure, and he goes on to analyze three such failures—the absence of a riskless real investment security, the absence of real annuities, and the problems in insuring the risk associated with varying length of working life. In this chapter, I examine another form of market failure—a failure, the nonmarketability of human capital—and show that, under certain conditions, a tax and transfer system not unlike the current social security system can reduce or eliminate the economic inefficiencies that result from such a failure.

Under the standard perfect market assumption used in the analysis of the optimal lifetime consumption–portfolio selection problem, an individual will, in general, prefer to use the private markets to design his own saving–retirement plan where benefits received are a function of the amounts he contributes and the investment experience from the portfolio allocation of these contributions. However, if there are assets of material significance which are not tradable, then this result need not obtain. Because an individual's opportunities to sell his future wage income are

Robert C. Merton is professor of finance at the Sloan School of Management, Massachusetts Institute of Technology, and a research associate of the National Bureau of Economic Research.

Aid from the National Bureau of Economic Research and the National Science Foundation is gratefully acknowledged. An earlier version of this paper was presented at the NBER Conference on Private and Public Pensions, Palo Alto, California, in January 1981, and I thank the participants for their comments. My thanks also to F. Black and P. Diamond for many helpful discussions. Any opinions expressed are mine and are not necessarily those of NBER, NSF, or, for that matter, F. Black or P. Diamond.
generally quite limited, a natural candidate for such a nontradable asset is human capital. As Diamond (1977) noted in his article, there are many possible types of market failures, and any such failure can affect individual welfare and behavior. However, the nontradability of human capital is an especially important market failure because human capital represents a significant fraction of national wealth, and because it is the major part of virtually everyone's initial endowment, its nontradability will affect all people in the economy. Indeed, even under the assumption of perfect certainty, significant welfare losses can occur from its nontradability, and these losses become still larger if this unrealistic assumption is relaxed.

It is well known that a major negative effect on individual welfare caused by the nontradability of human capital is that individuals cannot achieve their optimal life-cycle consumption program because early in life, when most of their wealth is in the form of human capital, they cannot consume as much as they would otherwise choose. This “forced-saving” distortion of the optimal program will obtain in both certainty and uncertainty models, although in certainty models it must be assumed that borrowing (against future wage income) is restricted. Otherwise human capital is tradable, because borrowing is a perfect substitute for sale. However, if future wage income is uncertain, then borrowing is not a perfect substitute for sale, and therefore the availability of credit will not eliminate the welfare loss from this market failure.

In addition to the distortion in the quantity of savings, the nontradability of human capital causes further welfare losses in an uncertainty model because investors cannot achieve an optimal portfolio allocation of their savings. This nonoptimality manifests itself in two ways. First, investors, especially the young, may be forced to bear more risk in their portfolios than they would choose in the absence of this market failure. Second, for any given level of risk, the portfolios held by investors will be inefficient. That is, in virtually every model of portfolio selection with perfect markets, an investor's optimal behavior is to invest part of his wealth (the risk part) in a well-diversified portfolio of all available risky assets and to invest the balance of his wealth in the riskless asset where the fractional allocation between the two is used to adjust the total risk level of his portfolio. However, if human capital is not tradable, then neither young nor old investors will be able to hold all available assets. Therefore, unless human and physical capital returns are perfectly positively correlated, the risk part of the investors' portfolios will be inefficiently diversified.

The focus of this chapter is on the elimination of this inefficiency in risk sharing, although in the particular model used in the analysis, the derived system of taxes and transfers also eliminates the distortions of savings. The framework for the analysis is an intertemporal general equilibrium
model of an economy with overlapping generations where people live for three periods: childhood, work, and retirement. Everyone is assumed to have the same utility function for lifetime consumption, which is of a very specific form. With the exception of the nontradability of human capital, all markets are assumed to be perfect and competitive. There is a single good, and there are two factors of production: homogeneous capital and labor. The stochastic production function is Cobb-Douglas. There is uncertainty about total output, factor shares, and the rate of population growth, and everyone agrees on the joint probability distributions for these random variables. Because labor is homogeneous, the wage rate is the same for all workers. Therefore, the model incorporates only the systematic, or aggregate, risk of human capital and not its individual-specific risks.

The analysis proceeds as follows: In section 12.2, the model is developed and the intertemporal general equilibrium path for the economy when human capital is tradable is derived as a benchmark for an efficient allocation. In section 12.3, under the assumption that labor is supplied inelastically, I derive a system of taxes and transfers which cause the economy when human capital is not tradable to replicate the efficient equilibrium path of section 12.2. This optimal system has constant proportional taxes on both wages and consumption, transfers to retirees equal to the contemporaneous revenues collected from the wage tax, and transfers to children equal to the contemporaneous revenues from the consumption tax. In section 12.4, the assumption of inelastically supplied labor is dropped, and the optimal system of the previous section is shown to distort the labor-leisure choice. However, it is further shown that if an eligibility requirement for retirement benefits (similar in spirit to the one currently used in the social security program) is imposed, then this distortion can be reduced and, under certain conditions, completely eliminated. In section 12.5, a brief summary of the analysis and its connection with some of the issues surrounding the current social security program are discussed.

Although the model used is relatively simple and highly aggregated, the formal derivation of the optimal tax and transfer system is long and somewhat complicated. Hence, before proceeding to the formal derivation, I briefly digress to provide an overview of how the derived system of taxes and transfers serves to correct the inefficiencies caused by the nontradability of human capital.

In general, the nontradability of human capital will cause a portfolio imbalance for younger people in the direction of forcing them to hold too much human capital relative to their holdings of physical capital. An extreme example would be a newborn person whose entire initial endowment is human capital. For older people, the imbalance goes in the opposite direction with too little human capital held relative to their
holdings of physical capital. Again, an extreme example would be a retired person who has no human capital.

To restore the proper portfolio balance for both young and old, it follows that the tax and transfer system should take away some of the human capital from the young and give it to the old, and take away some of the physical capital from the old and give it to the young. A wage tax, the proceeds of which are paid to current retirees, accomplishes the first part. That is, the tax takes some human capital away from the young, and, because the retirement benefits are a function of contemporaneous wage earnings, these benefits give older people an investment in human capital. A consumption tax, the proceeds of which are paid to current children, accomplishes the second part. Although the consumption tax takes away from all ages in the population, it takes proportionally more away from older people who as part of a standard life-cycle program will be currently consuming a larger fraction of their wealth. The transfers to children permit them to finance both current consumption and investment in physical capital. By choosing the proper tax rates, this package of taxes and transfers can correct the inefficiencies caused by the nontradability of human capital.

Of course, if human capital were tradable, there would be no need for such a tax and transfer system because the exchanges between young and old would take place directly in the private markets with the young selling claims on their future wage income to the old.

12.2 Equilibrium When Human Capital Is Tradable

12.2.1 Model Assumptions and Individual Optimal Behavior

In this section, an intertemporal general equilibrium model of an economy is developed under the assumption that financial markets are perfect. That is, we make the following assumptions:

ASSUMPTION 1: All assets (including human capital) are tradable.

ASSUMPTION 2: There are no transactions costs, taxes, or problems with indivisibilities of assets.

ASSUMPTION 3: There are enough investors with comparable wealth levels so that each investor believes that he can buy and sell as much of an asset as he wants at the market price.

ASSUMPTION 4: The capital market is always in equilibrium.

ASSUMPTION 5: There exists an exchange market for borrowing and lending at the same rate of interest.

ASSUMPTION 6: Shortsells of all assets, with full use of the proceeds, are allowed.

The aggregate production technology for the economy is described by the Cobb-Douglas-type production function
\[ Q(t) = A(t)[I(t-1)]^{\theta(t)}[L(t)]^{1-\theta(t)}, \]

where \( Q(t) \) is aggregate output produced at time \( t \) of the single good which can be used either as a capital good or for consumption; \( I(t-1) \) is the amount of capital which must be put in place at time \( (t-1) \) in order to be used in production at time \( t \); \( L(t) \) is the aggregate amount of labor used in production at time \( t \) and is chosen at that date; \( A(t) \) is a positive random variable assumed to be independent of the level of investment \( I(t-1) \); and \( \theta(t) \) is a random variable with range \( 0 \leq \theta(t) \leq 1 \). It is further assumed that the \( \{\theta(t)\}, t = 1, 2, \ldots \), are independent and identically distributed with \( E_{t-k}[\theta(t)] = \alpha \), for \( k = 1, 2, \ldots \), where \( E_{t-k} \) is the conditional expectation operator, conditional on knowing all relevant information available as of date \( t-k \).

Firms are assumed to be perfect competitors, and they make all production decisions so as to maximize the market value of the firm. When firms' managers make their investment decisions at time \( t-1 \), \( I(t-1) \), they do not know either \( A(t) \) or \( \theta(t) \). However, when they choose the amount of labor to employ at time \( t \), both \( A(t) \) and \( \theta(t) \) are known. Hence, the demand for labor at time \( t \) is determined by the solution to \( \max [Q(t) - \omega(t)L(t)] \), which from (1) leads to the first-order condition

\[ 0 = [1 - \theta(t)]A(t)[I(t-1)]^{\theta(t)}[L(t)]^{-\theta(t)} - \omega(t), \]

where \( \omega(t) \) is the wage rate at time \( t \). By multiplying (2) by \( L(t) \), we can rewrite (2) to express the aggregate demand for labor, \( L^d(t) \), as

\[ L^d(t) = [1 - \theta(t)]Q(t)/\omega(t). \]

By inspection of (3), labor's share of output at time \( t \) will be \( [1 - \theta(t)] \), and therefore the aggregate net revenues of firms available for distribution to shareholders at time \( t \) will be equal to \( \theta(t)Q(t) \). By the standard accounting identity, aggregate dividends paid (net of new financings), \( D(t) \), must satisfy

\[ D(t) = \theta(t)Q(t) - I(t), \]

where \( I(t) \) is the amount of physical investment chosen by managers at time \( t \) in preparation for production at time \( t+1 \).

Because, at time \( t \), \( \theta(t+1) \) is a random variable, both current stockholders and those who will be workers at time \( t+1 \) face uncertainty not only about aggregate output at time \( t+1 \) but also about the distribution of that output between the two factors of production. However, if all factors are tradable, it will be shown that this latter uncertainty can be eliminated in the sense that it has no effect on consumer-investor welfare. That is, through efficient risk sharing, factor-share risk can be diversified away.

At each point in time \( t \), there are three securities traded in the financial
markets. The first kind is shares of stock in the firms, which represent ownership of physical capital. The random variable return per dollar on these shares between \( t \) and \( t+1 \) is denoted by \( Z_1(t+1) \). Second is a human capital security which pays \( \omega(t+1) \) per share to its owner at time \( t+1 \) with no further payments thereafter, where \( \omega(t+1) \) is the wage rate at time \( t+1 \). The random variable return per dollar on this security is denoted by \( Z_2(t+1) \). Third is a riskless security which pays one dollar at time \( t+1 \) and whose return per dollar between \( t \) and \( t+1 \) is denoted by \( R(t) \).

Each person in the economy lives for three periods. At age 0, his initial endowment is equal to the market value of his human capital. In this childhood period, the individual chooses how much to consume and then allocates his savings in a portfolio decision. He finances these choices by selling part of his endowment in the private market. In the next, work period of his life, he chooses how much to consume of goods and leisure and then allocates his savings in a portfolio decision. In the last, retirement period, he consumes all his wealth because there is no bequest motive.

It is assumed that all people at birth have the same lifetime utility-of-consumption function which at time \( t \) is denoted by \( U_0(t) \) given by

\[
U_0(t) = \log [c_0(t)] + E_t[\Gamma \log \ell(t+1)] \\
+ \log [c_1(t+1)] + \log [c_2(t+2)]
\]

where \( c_k(t) \) is the consumption of a person of age \( k \) at time \( t \), \( k = 0, 1, 2 \); \( \ell(t+1) \) is the fraction of the person’s work period spent in leisure, \( 0 \leq \ell(t+1) \leq 1 \); and \( \Gamma \) is a nonnegative constant.

From assumption (5), it follows that each person of age 1 at time \( t \) will have a lifetime utility-of-consumption function, \( U_1(t) \), given by

\[
U_1(t) = \Gamma \log \ell(t) + \log [c_1(t)] + E_t[\log [c_2(t+1)]]
\]

and for each person of age 2 at time \( t \), the lifetime utility-of-consumption function, \( U_2(t) \), is given by

\[
U_2(t) = \log [c_2(t)].
\]

The solution of the individual’s optimal lifetime consumption program is derived in the appendix using the technique of stochastic dynamic programming. Because all assets are tradable, the problem can be formally expressed in the standard form used by Samuelson (1969) and Hakansson (1970) where all income derives from investment in securities. That is, if \( w_k(t) \) denotes the wealth of a person of age \( k \) at time \( t \), then the dynamic accumulation equation for wealth can be written as

\[
w_{k+1}(t+1) = s_k(t)x_{1k}(t)[Z_1(t+1) - R(t)] \\
+ x_{2k}(t)[Z_2(t+1) - R(t)] + R(t),
\]
where $s_k(t)$ is saving by an age $k$ person, $x_{1k}(t)$ is the fraction of his savings which is allocated to shares of firms, $x_{2k}(t)$ is the fraction allocated to the human capital security, and $[1 - x_{1k}(t) - x_{2k}(t)]$ is the fraction allocated to the riskless security.

To use this standard form of the accumulation equation, we adopt the convention of including as part of an investor's wealth the gross value of his human capital, which is defined as the current market price for the wage income he would earn if he were to work 100% of the time during the work period of his life. If we assume no bequests, each person's initial endowment is just his human capital. Therefore, if we include the gross value of each person's human capital in his wealth, it follows that the value of each person's initial endowment at time $t$ will satisfy

$$w_0(t) = p(t),$$

where $p(t)$ denotes the equilibrium price at time $t$ of one share of the human capital security.

It also follows from this convention that each person must buy back the amount of leisure time which he chooses to consume during the work period of his life. Hence, individual saving for age 0 and age 1 people is defined by

$$s_0(t) = w_0(t) - c_0(t)$$

and

$$s_1(t) = w_1(t) - c_1(t) - \omega(t) \ell(t),$$

where saving by age 1 people includes a deduction from wealth for both consumption of goods and leisure with the price per unit of leisure time equal to the wage rate, $\omega(t)$. The convention of including the gross value of human capital is adopted for analytical convenience only, and the same equilibrium quantities and prices will obtain in the alternative formulation which uses the net (of leisure spent) value of human capital with no separate deduction for leisure.

Because of the no-bequest assumption, the optimal consumption rule for an age 2 person is simply $c_2(t) = w_2(t)$. From the analysis in the appendix, the optimal consumption and saving rules for an age 0 person can be written as

$$c_0^*(t) = \frac{w_0(t)}{3 + \Gamma}$$

and

$$s_0^*(t) = \left(\frac{2 + \Gamma}{3 + \Gamma}\right) w_0(t).$$

The corresponding behavior rules for an age 1 person can be written as
As shown in the appendix, age 0 investors and age 1 investors will have the same fractional allocations in their optimal portfolios. That is,

$$x_{0}^{*}(t) = x_{0}^{*}(t) \neq x_{2}^{*}(t), j = 1, 2.$$ Exhibiting the well-known properties of the log utility function, these optimal portfolio weights, \([x_{1}^{*}(t), x_{2}^{*}(t)]\), do not depend on the level of the investor’s wealth and are given by the solution to the equation set

$$0 = E_t \left[ \frac{Z_{j}(t + 1) - R(t)}{Z^*(t + 1)} \right], j = 1, 2,$$

where \(Z^*(t + 1) = x_{1}^{*}(t)[Z_{1}(t + 1) - R(t)] + x_{2}^{*}(t)[Z_{2}(t + 1) - R(t)] + R(t)\) is the return per dollar on the optimal portfolio which is common to both age groups.

Equations (11)–(13) completely describe individual optimal consumption, saving, and portfolio selection behavior at each age and point in time. To determine the corresponding aggregate behavior necessary to derive the intertemporal equilibrium prices and quantities, I now turn to the assumed demographics for the economy.

Let \(N_{0}(t)\) denote the number of children born in the economy at time \(t\). Although \(N_{0}(t)\) is known at time \(t\), I assume it to be a random variable relative to times earlier than \(t\). I further assume that the stochastic process describing the evolution of \(N_{0}(t)\) is exogenous and independent of the level of economic activity. Because each person lives for three periods, it follows that the number of age 1 people in the economy at time \(t\), \(N_{1}(t)\), is given by

$$N_{1}(t) = N_{0}(t - 1)$$

and the number of age 2 people in the economy at time \(t\), \(N_{2}(t)\), is given by

$$N_{2}(t) = N_{1}(t - 1) = N_{0}(t - 2).$$

Hence, at time \(t\), the size of the work force for the economy at time \(t + 1\) will be known with certainty, and at that time the number of retirees in the economy at times \(t + 1\) and \(t + 2\) will also be known with certainty. If \(N(t)\) denotes the total population at time \(t\), then \(N(t) = N_{0}(t) + N_{1}(t) + N_{2}(t)\), and the dynamics for \(N(t)\) can be written as
If \( W_k(t) = N_k(t)w_k(t) \) denotes the aggregate wealth of all age \( k \) people in the economy at time \( t \), \( k = 0, 1, 2 \), then from (11) and (12), the corresponding aggregate consumption and savings for each age group can be written as

(17a) \[ C_0(t) = \frac{W_0(t)}{3 + \Gamma} \]

(17b) \[ S_0(t) = (\frac{2 + \Gamma}{3 + \Gamma})W_0(t) \]

(17c) \[ C_1(t) = \frac{W_1(t)}{2 + \Gamma} \]

(17d) \[ S_1(t) = \frac{\Gamma W_1(t)}{(2 + \Gamma)\omega(t)} \]

From the aggregation across all age groups in (17), national wealth, \( W(t) \), aggregate consumption, \( C(t) \), and aggregate saving, \( S(t) \), for the economy can be written as

(18a) \[ W(t) = W_0(t) + W_1(t) + W_2(t) \]

(18b) \[ C(t) = \frac{(2 + \Gamma)W_0(t) + (3 + \Gamma)W_1(t) + (6 + 5\Gamma + \Gamma^2)W_2(t)}{6 + 5\Gamma + \Gamma^2} \]

and

(18c) \[ S(t) = \frac{(4 + 4\Gamma + \Gamma^2)W_0(t) + (3 + \Gamma)W_1(t)}{6 + 5\Gamma + \Gamma^2} \]

12.2.2 Equilibrium in the Financial Markets

Because firms are competitive and, from assumption (1), the production technology exhibits constant returns to scale, the equilibrium ex-dividend aggregate market value of firms' shares, denoted by \( V(t) \), must satisfy

(19) \[ V(t) = I(t) \]

where \( I(t) \) is aggregate amount of physical investment made at time \( t \). From (19), (3), and (4), it follows that the aggregate value of firms' shares at time \( t \), prior to paying dividends or issuing new shares, will be equal to the net revenues available for distribution to shareholders, \( \theta(t)Q(t) \).
Hence, the equilibrium return per dollar on firms' shares between \( t \) and \( t + 1 \) can be written as

\[
Z_1(t + 1) = \frac{\theta(t + 1)Q(t + 1)}{V(t)} = \frac{\theta(t + 1)Q(t + 1)}{I(t)}.
\]

Because there are \( N_0(t) \) people of age 0, the equilibrium market value of gross human capital at time \( t \) is \( N_0(t)p(t) \), and this is the aggregate amount of the second security, the human capital security, which must be held in investors' portfolios. Because these securities are a claim on gross human capital, the aggregate dollar return on them at time \( t \) will be equal to total wages paid at time \( t \) plus the dollar amount of leisure time purchased at time \( t \), \( \mathcal{L}(t)\omega(t) \). It follows that the equilibrium return per dollar on the second security between \( t \) and \( t + 1 \) is given by

\[
Z_2(t + 1) = \frac{[1 - \theta(t + 1)]Q(t + 1) + \mathcal{L}(t + 1)\omega(t + 1)}{N_0(t)p(t)}.
\]

As the appendix shows, all investors will allocate their savings in the same relative proportions across the available securities, and hence it follows immediately that in equilibrium this common optimal portfolio must be the market portfolio (i.e., the portfolio which holds all securities in proportion to their market values). If \( \delta_1(t) \) denotes the fraction of the market portfolio held in the shares of firms and \( \delta_2(t) \) denotes the fraction held in human capital, then \( [1 - \delta_1(t) - \delta_2(t)] \) is the fraction held in the riskless security. Because there is no net supply of the riskless security, it follows that \( \delta_1(t) = 1 - \delta_2(t) \), and financial market equilibrium requires that

\[
\begin{align*}
\frac{x_1^*(t)}{x_2^*(t)} &= \frac{\delta_1(t)}{\delta_2(t)} = \frac{V(t)}{V(t) + N_0(t)p(t)}, \\
&\text{and} \\
\frac{x_2^*(t)}{x_2^*(t)} &= \frac{\delta_2(t)}{\delta_2(t)} = \frac{N_0(t)p(t)}{V(t) + N_0(t)p(t)}.
\end{align*}
\]

The total dollar return to the market at time \( t \) is equal to \( \theta(t)Q(t) + [1 - \theta(t)]Q(t) + \mathcal{L}(t)\omega(t) = Q(t) + \mathcal{L}(t)\omega(t) \). Therefore, the equilibrium return per dollar on the market portfolio between \( t \) and \( t + 1 \), \( Z_M(t + 1) \), is given by

\[
Z_M(t + 1) = [Q(t + 1) + \mathcal{L}(t + 1)\omega(t + 1)]/[V(t) + N_0(t)p(t)].
\]

Because in equilibrium \( Z^*(t) = Z_M(t) \), it follows from a straightforward manipulation of (13) that, in equilibrium,

\[
E_t[Z_1(t + 1)/Z_M(t + 1)] = 1.
\]

From (8), (9), (14), (17b), and (17d), it can be shown that
$S(t) = \Gamma N_0(t - 1)p(t - 1)Z_M(t)/(3 + \Gamma)$, and therefore (23) can be rewritten as

\begin{equation}
Z_M(t + 1) = Q(t + 1)/[V(t) + \left(\frac{3}{3 + \Gamma}\right)N_0(t)p(t)].
\end{equation}

Substituting for $Z_1(t + 1)$ from (20) and for $Z_M(t + 1)$ from (25) into (24), we have, as a condition for equilibrium, that

\begin{equation}
E_t[\theta(t + 1)] = \frac{(3 + \Gamma)V(t)}{[(3 + \Gamma)V(t) + 3N_0(t)p(t)]}.
\end{equation}

But, by assumption (1) $E_t[\theta(t + 1)] = \alpha$, a constant for all $t$, and hence, from (26), the equilibrium market value of human capital must satisfy

\begin{equation}
N_0(t)p(t) = \left[\frac{(1 - \alpha)(3 + \Gamma)}{3\alpha}\right]V(t).
\end{equation}

Substituting from (27) into (22) we have that the equilibrium portfolio weights in the market portfolio are constants over time and are given by

\begin{equation}
\delta_1 = \frac{3\alpha}{3 + (1 - \alpha)\Gamma}.
\end{equation}

and

\begin{equation}
\delta_2 = \frac{(3 + \Gamma)(1 - \alpha)}{3 + (1 - \alpha)\Gamma}.
\end{equation}

We can now summarize the derived conditions for equilibrium in the financial market. From (19) and (27) the equilibrium values of firms and human capital can be written as

\begin{equation}
V(t) = I(t)
\end{equation}

and

\begin{equation}
N_0(t)p(t) = \left[\frac{(1 - \alpha)(3 + \Gamma)}{3\alpha}\right]I(t).
\end{equation}

From (29) and the condition that aggregate financial saving must be equal to the market value of all securities, we have that

\begin{equation}
S(t) = \left[\frac{3 + (1 - \alpha)\Gamma}{3\alpha}\right]I(t).
\end{equation}

From (25), (27), and (29a), the equilibrium return on the market portfolio between $t - 1$ and $t$ can be written as

\begin{equation}
Z_M(t) = \frac{\alpha Q(t)}{I(t - 1)}.
\end{equation}
From (20), (21), (29), and (31), the returns on shares of firms and human capital between \( t - 1 \) and \( t \) can be written as

\[
(32a) \quad Z_1(t) = \frac{\theta(t)Q(t)}{I(t - 1)} = \left[ \frac{\theta(t)}{\alpha} \right] Z_M(t)
\]

and

\[
(32b) \quad Z_2(t) = \left[ \frac{(1 - \alpha)\Gamma + 3[1 - \theta(t)]}{(1 - \alpha)\Gamma + 3[1 - \alpha]} \right] \alpha Q(t) = \left[ \frac{(1 - \alpha)\Gamma + 3[1 - \theta(t)]}{(1 - \alpha)\Gamma + 3[1 - \alpha]} \right] Z_M(t).
\]

Finally, to determine the equilibrium return on the riskless security, we have from (13), (24), and (31) that

\[
(33) \quad R(t) = \frac{1}{E_t[1/Z_M(t + 1)]} = \frac{\alpha}{I(t)} / E_t[1/Q(t + 1)].
\]

Of course, equilibrium conditions (29)–(33) are not the proper reduced-form equations for these equilibrium prices, returns, and quantities because they contain endogenous variables on their right-hand side. To drive the proper reduced-form equations in terms of the exogenous and predetermined variables of the economy as well as the equilibrium quantities of physical investment, consumption, and output, it is necessary to examine the "real" sector of the economy.

12.2.3 Equilibrium in the Goods and Labor Markets

A necessary condition for equilibrium in the market for physical output is that

\[
(34) \quad Q(t) = C(t) + I(t).
\]

From (9), (17a), and (29b), we have that aggregate consumption by age 0 people at time \( t \) can be written as

\[
(35) \quad C_0(t) = \frac{(1 - \alpha)}{3\alpha} I(t).
\]

Because \( S_1(t) = S(t) - S_0(t) = S(t) - W_0(t) + C_0(t) \), it follows from (9), (29b), (30), and (35) that aggregate saving by age 1 people at time \( t \) can be written as

\[
(36) \quad S_1(t) = \left( \frac{1 + 2\alpha}{3\alpha} \right) I(t),
\]
and therefore, from (17c) and (17e), aggregate consumption by age 1 people must satisfy

\[(37) \quad C_1(t) = \left(\frac{1 + 2\alpha}{3\alpha}\right) I(t).\]

From the equilibrium condition that \(Z^*(t) = Z_M(t)\) and the accumulation equation (8), we know that aggregate consumption by age 2 people at time \(t\) can be written as

\[(38) \quad C_2(t) = S_1(t-1)Z_M(t) = S_1(t-1)\alpha Q(t)/l(t-1) \text{ from (31)},\]

\[= \left(\frac{1 + 2\alpha}{3}\right) Q(t) \text{ from (36)}.\]

Therefore, substituting for \(C(t)\) from (35), (37), and (38) into (34) and rearranging terms, we have that in equilibrium

\[(39) \quad \alpha(1 - \alpha)\frac{z(t)}{l(t)} = \left[\frac{\alpha(1 - \alpha)}{2\alpha + 1}\right] Q(t).\]

The supply of labor at time \(t\), \(L^s(t)\), is equal to \(N_1(t) - L_d(t)\). From (17c), (17d), (37) and (39), we can rewrite this expression as

\[(40) \quad L^s(t) = N_1(t) - \frac{(1 - \alpha)Q(t)}{3\omega(t)}.\]

For the labor market to be in equilibrium, \(L^s(t) = L^d(t)\), and therefore, from (4) and (40), the equilibrium wage rate can be written as

\[(41) \quad \omega(t) = \left[1 - \theta(t) + \frac{(1 - \alpha)\Gamma}{3}\right] Q(t)/N_1(t).\]

Substituting for \(\omega(t)\) in (40) from (41), the equilibrium aggregate quantity of labor, \(L(t)\), can be written as

\[(42) \quad L(t) = \frac{3N_1(t)[1 - \theta(t)]}{3[1 - \theta(t)] + (1 - \alpha)\Gamma}.\]

Substituting for \(L(t)\) in assumption (1) from (42), we have that the equilibrium quantity of output at time \(t\) is given by

\[(43) \quad Q(t) = A(t)[l(t-1)]^{\theta(t)}\left[\frac{3N_1(t)[1 - \theta(t)]}{3[1 - \theta(t)] + (1 - \alpha)\Gamma}\right]^{1-\theta(t)}.\]

This completes the analysis of the intertemporal general equilibrium model of the economy with perfect markets. By substituting for \(Q(t)\) from (43) into each of the previously derived equilibrium conditions, the complete set of reduced-form equations for equilibrium prices, quanti-
ties, and returns can be written in terms of the exogenous variables $A(t), \theta(t), N_0(t), N_1(t)$, and $N_2(t)$ and the predetermined variable $I(t - 1)$. For convenience and ease of reference, these equations for the equilibrium prices, quantities, and returns are presented in tables 12.1, 12.2, and 12.3.

**12.3 A System of Taxes and Transfers Which Replicates the Perfect-Market Equilibrium Path When Human Capital Is Not Tradable and Labor Is Inelastically Supplied**

In this section, I assume that human capital is not tradable and derive a system of taxes and transfers which cause the economy to replicate the perfect-market equilibrium consumption and saving patterns derived in section 12.2. This system serves two functions: first, by providing transfers to people in the childhood period for consumption, it corrects the savings distortion caused by the nontradability of human capital. This transfer is financed by a proportional consumption tax where the tax rate

---

**Table 12.1 Equilibrium Prices and Quantities in the Real Sector**

<table>
<thead>
<tr>
<th>Aggregate Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q(t) = A(t)[I(t-1)]^{\gamma(t)} \left[ \frac{3N_1(t)[1-\theta(t)]}{3[1-\theta(t)] + (1-\alpha)\Gamma} \right]^{1-\delta(t)}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Capital Investment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I(t) = \left[ \frac{\alpha(1-\alpha)}{2\alpha + 1} \right] Q(t)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aggregate Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C(t) = \frac{1+\alpha+\alpha^2}{2\alpha + 1} Q(t)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aggregate Labor Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L(t) = \frac{3[1-\theta(t)]N_1(t)}{3[1-\theta(t)] + (1-\alpha)\Gamma}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Aggregate Leisure Time for Workers</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L(t) = \frac{(1-\alpha)\Gamma N_1(t)}{3[1-\theta(t)] + (1-\alpha)\Gamma}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wage Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega(t) = \left[ \frac{3[1-\theta(t)] + (1-\alpha)\Gamma}{3N_1(t)} \right] Q(t)$</td>
</tr>
</tbody>
</table>
Table 12.2  Equilibrium Values, Portfolio Allocations, and Returns in the Financial Sector

<table>
<thead>
<tr>
<th>Aggregate values and returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>National wealth</td>
</tr>
<tr>
<td>[ W(t) = \frac{[3(2 + \alpha^2) + \Gamma(2 - \alpha - \alpha^2)]}{3(2\alpha + 1)} Q(t) ]</td>
</tr>
<tr>
<td>Aggregate saving</td>
</tr>
<tr>
<td>[ S(t) = \frac{[3 + \Gamma(1 - \alpha)][1 - \alpha]}{3(2\alpha + 1)} Q(t) ]</td>
</tr>
<tr>
<td>Return per dollar</td>
</tr>
<tr>
<td>between ( t ) and ( t + 1 )</td>
</tr>
<tr>
<td>on market portfolio</td>
</tr>
<tr>
<td>[ Z_M(t + 1) = \left[ \frac{(2\alpha + 1)}{(1 - \alpha)} \right] \frac{Q(t + 1)}{Q(t)} ]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Components of the market portfolio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firms ((j = 1))</td>
</tr>
<tr>
<td>[ \delta_j = \frac{3\alpha}{3 + (1 - \alpha)\Gamma} ]</td>
</tr>
<tr>
<td>[ Z_M(t + 1) = \left[ \frac{\theta(t + 1)}{\alpha} \right] ]</td>
</tr>
<tr>
<td>Human capital ((j = 2))</td>
</tr>
<tr>
<td>[ \frac{(3 + \Gamma)(1 - \alpha)}{3 + (1 - \alpha)\Gamma} ]</td>
</tr>
<tr>
<td>[ Z_M(t + 1) = \left[ \frac{3[1 - \theta(t + 1)] + \Gamma(1 - \alpha)}{(3 + \Gamma)(1 - \alpha)} \right] ]</td>
</tr>
<tr>
<td>Age Group</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>Childhood</td>
</tr>
<tr>
<td>Work</td>
</tr>
<tr>
<td>Retirement</td>
</tr>
</tbody>
</table>
Social Security as a Means for Efficient Risk Sharing

\( \tau_c \) is a constant over time. Second, it provides more efficient risk sharing in the economy by eliminating the unnecessary, or diversifiable, factor-share risk which would otherwise be borne by all age groups in the economy. We eliminate the risk by making transfers to current retirees financed by a proportional tax on wages of current workers where the tax rate \( \tau_w \) is a constant over time. This pay-as-you-go tax and transfer system is similar to the retirement component of the present social security system.\(^7\)

To highlight both the benefits and sources of possible distortions to the economy from the system, the appropriate tax rates and transfers to replicate the perfect market equilibrium path of section 12.2 are derived first under the assumption that labor is supplied inelastically. That is, there is no demand for leisure time, and this is accomplished in the model by selling \( \Gamma \) equal to zero in workers' utility functions. In the next section, the effects of introducing demand for leisure into this system are examined. To determine optimal taxes and transfers, the individual lifetime consumption-saving problem is solved taking into account taxes, transfers, and the nontradability of human capital. For notational simplicity, I use the same variable symbols that were used in section 12.2, adding prime signs where necessary to distinguish them from their perfect-market counterparts.

The analysis begins with the examination of the behavior of those who are in the work period of their lives at time \( t \). Let \( \mu_1 \) denote the fraction of capital investment, \( I'(t-1) \), owned through the purchase of shares of firms by age 0 people at time \( t-1 \). The ownership of this capital is possible as the result of saving out of transfers received in childhood. Because all direct saving can be invested in shares of firms only, aggregate saving by age 1 people at time \( t \) can be written as

\[
(44) \quad S_1(t) = (1 - \tau_w)[1 - \theta(t)]Q'(t) + \mu_1\theta(t)Q'(t) - (1 + \tau_c)C_1(t).
\]

It follows, therefore, that aggregate retirement-period consumption at time \( t + 1 \) must satisfy

\[
(45) \quad C_2(t + 1) = \frac{1}{1 + \tau_c} \left\{ S_1(t) \left[ \frac{\theta(t + 1)Q'(t + 1)}{I'(t)} \right] + \tau_w[1 - \theta(t + 1)]Q'(t + 1) \right\},
\]

where the second term in the brackets is the retirement transfer payment which is equal to the total taxes on wages collected.

Under the assumption that \( \Gamma = 0 \), each age 1 person will choose current consumption, \( c'_1(t) \), so as to maximize \( \log[c'_1(t)] + E_t[\log[c'_2(t + 1)]] \) subject to his budget constraint. From (44) and (45), the first-order
condition for this maximization can be written in terms of age group aggregates as

\[
\frac{1}{C'_i(t)} = E_t \left[ \frac{\theta(t + 1)Q'(t + 1)/I'(t)}{C'^*_i(t + 1)} \right],
\]

where it is understood that \(C'_i\) and \(C'^*_i\) in (46) are the optimal consumption decisions.

If taxes and transfers are to be chosen so as to replicate the equilibrium in section 12.2, then, as a necessary condition, \(I'(t) = I(t)\) and \(C'_k(t) = C_k(t), k = 0, 1, 2\). Therefore, it follows from tables 12.1 and 12.3 that

\[
I'(t) = \left[ \frac{\alpha(1 - \alpha)}{2\alpha + 1} \right] Q'(t),
\]

\[
C'_1(t) = \frac{(1 - \alpha)}{3} Q'(t),
\]

and

\[
C'_2(t + 1) = \frac{(2\alpha + 1)}{3} Q'(t + 1).
\]

Substituting these necessary conditions for \(C'_2(t + 1)\) and \(I'(t)\) into (44) and rearranging terms, we have that \(\mu_1, \tau_\omega, \) and \(\tau_c\) must be chosen so as to satisfy

\[
0 = \left[ \mu_1 - (1 - \tau_\omega) \right] \left[ \frac{2\alpha + 1}{\alpha(1 + \alpha)} \right] \theta(t) \theta(t + 1)
+ \left\{ \frac{2\alpha + 1}{\alpha(1 - \alpha)} \left( 1 - \tau_\omega \right) - \frac{(1 - \tau_c)(1 - \alpha)}{3} \right\} \theta(t)
+ \left[ \tau_\omega - \frac{(1 - \tau_c)(2\alpha + 1)}{3} \right].
\]

Because \(\tau_\omega, \tau_c,\) and \(\mu_1\) are assumed to be constants, (50) will be satisfied for all \(\theta(t)\) and \(\theta(t + 1)\) only if

\[
\mu_1 = 1 - \tau_\omega,
\]

\[
\tau_\omega = \frac{2\alpha + 1}{\alpha(1 - \alpha)} \left( 1 - \tau_\omega \right) \left( 1 + \tau_c \right) \left( 1 - \frac{1}{3} \right),
\]

and

\[
\tau_\omega = \frac{2\alpha + 1}{3} (1 + \tau_c).
\]

Solving the system of equations (51), we have that as a necessary condition for a replication of the perfect-market economy, \(\tau_\omega, \tau_c,\) and \(\mu_1\) must be chosen as follows:
(52a) \[ \mu_1 = \frac{(1 - \alpha^2)}{2 + 2\alpha - \alpha^2}, \]

(52b) \[ \tau_\omega = \frac{2\alpha + 1}{2 + 2\alpha - \alpha^2}, \]

and

(52c) \[ \tau_c = \frac{(1 - \alpha)^2}{2 + 2\alpha - \alpha^2}. \]

A further necessary condition for replication is that the substitution of these values of \( \mu_1, \tau_\omega, \) and \( \tau_c \) into (46) will lead to an optimal consumption choice \( C'_1(t) \) that satisfies condition (48). The reader may verify that \( C'_1(t) = (1 - \alpha)Q'(t)/3 \) does indeed satisfy (46) when \( \mu_1, \tau_\omega, \) and \( \tau_c \) take on the values given in (52).

Of course, unlike the tax rates, \( \tau_\omega \) and \( \tau_c \), the fraction of capital investment held by age 0 people, \( \mu_1 \), is not under direct government control. However, it can be controlled indirectly by choosing the appropriate amount of transfers made by the government to children. To determine this optimal level of transfers, we analyze the optimal consumption-saving decisions made by people who are in the childhood period of their lives at time \( t \).

If \( T_0(t) \) denotes aggregate transfers to age 0 people at time \( t \), then aggregate saving by this age group can be written as

(53) \[ S'_0(t) = T_0(t) - (1 - \tau_c)C'_0(t). \]

To satisfy the necessary conditions for replication, \( S'_0(t) = \mu_1I'(t) \) and, from table 12.3, \( C'_0(t) = [(1 - \alpha^2)/3(2\alpha + 1)]Q'(t) \). Substituting for \( I'(t) \) from (47), for \( \mu_1 \) from (52a), and for \( \tau_c \) from (52c), we have from (53) that these necessary conditions will be satisfied if

(54) \[ T_0(t) = \frac{(1 - \alpha)^2[\alpha(1 + \alpha) + 1]}{2\alpha + 1}(2 + 2\alpha - \alpha^2)Q'(t). \]

With taxes and transfers that satisfy (52) and (54), we have that \( C'_k(t)/Q'(t) = C_k(t)/Q(t) \) in the perfect-market equilibrium, \( k = 0, 1, 2 \). Therefore, \( C'(t)/Q'(t) = C(t)/Q(t) \), which implies that \( I'(t)/Q'(t) = I(t)/Q(t) \). By assumption, \( I = 0 \) and therefore the equilibrium quantities of labor will be equal, that is, \( L'(t) = L(t) = N_1(t) \), the number of age 1 people in the economy. It follows that equilibrium aggregate output will be the same, that is, \( Q'(t) = Q(t) \).

Hence, to ensure that this tax and transfer system will replicate the perfect-market economy, all that remains to be shown is that the government budget constraint is satisfied. Because transfer payments to retirees always equal wage taxes collected, these two cancel. Multiplying aggre-
gate consumption given in table 12.1 by the consumption tax rate from (52c), we have that consumption tax revenues can be written as

\[
\tau_c C(t) = \left[ \frac{(1 - \alpha)^2}{(2 + 2\alpha - \alpha^2)} \right] \left[ \frac{1 + \alpha + \alpha^2}{2\alpha + 1} \right] Q(t)
\]

\[
= T_0(t),
\]

from (54). Since consumption tax revenues just equal transfers to childhood-period people, the government budget constraint is satisfied, and the prescribed system of taxes and transfers will cause the economy with no trading in human capital to replicate the intertemporal equilibrium path for the economy when human capital is tradable.

12.4 A System of Taxes and Transfers Which Replicates the Perfect-Market Equilibrium Path When Human Capital Is Not Tradable and Labor Is Elastically Supplied

In the analysis of the previous section, a tax and transfer system was derived which caused the economy to replicate the perfect-market equilibrium path derived in section 12.2. This system eliminates completely the inefficiencies caused by the market failure of no trading in human capital, but it was derived for the special case where labor is inelastically supplied. In this section, we reinstate the labor-leisure choice and examine the effects of the system derived in section 12.3 on the equilibrium path of the economy when labor is not supplied inelastically.

As in section 12.3, we begin by examining the optimal behavior of those who are in their work period at time \( t \). At time \( t \), each age 1 person will choose current consumption and leisure time so as to maximize \( \Gamma \log [\ell'(t)] + \log [c'_1(t)] + E_t \log [C'_0(t + 1)] \) where aggregate saving by age 1 people at time \( t \) can be written as

\[
S'_1(t) = (1 - \tau_o)\omega'(t)[1 - \mathcal{L}'(t)] + \mu \theta(t)Q'(t) - (1 + \tau_c)C'_1(t)
\]

and \( C'_2(t + 1) \) is given by (45). Differentiating with respect to each of the choice variables, \( [\ell'(t), c'_1(t)] \), we have that the first-order conditions for a maximum, written in terms of aggregates, must satisfy

\[
\frac{1}{C'_1(t)} = E_t \left[ \frac{\theta(t + 1)Q'(t + 1)/I'(t)}{C'_2(t + 1)} \right],
\]

which is identical to (46) in section 12.3, and

\[
\frac{\Gamma}{\mathcal{L}'(t)} = E_t \left[ \frac{(1 - \tau_o)\omega'(t)\theta(t + 1)Q'(t + 1)/I'(t)}{(1 + \tau_c)C'_2(t + 1)} \right].
\]

Combining (57a) and (57b), we have that
It is straightforward to show that all other first-order conditions for age 0 and age 2 people will be identical to those deduced in section 12.3. Therefore, for the tax and transfer system given in (52) and (54), the optimal consumption and saving behavior per unit of current aggregate output, \( Q'(t) \), will be the same here as in the inelastic labor supply case of section 12.3. However, in the perfect-market equilibrium of section 12.2, the optimal aggregate amount of leisure time for workers satisfied the condition that

\[
\mathcal{L}'(t) = \frac{(1 + \tau_c)}{(1 - \tau_\omega)} \frac{\Gamma C_1(t)}{\omega(t)}.
\]

A comparison of (59) with (58) shows that both the consumption and wage taxes will cause a distortion of the labor-leisure decision in the direction of demanding more leisure than in the perfect-market case for the same wage rate and consumption.

As in sections 12.2 and 12.3, optimal aggregate work period consumption is given by \( C'(t) = (1 - \alpha)Q'(t)/3 \), and from (58) and (3), it follows that the wage rate which equilibrates the labor market is given by

\[
\omega'(t) = \left[ \frac{3[1 - \theta(t)] + (1 - \alpha)\Gamma(1 + \tau_c)}{1 - \tau_\omega} \right] Q'(t)/3N_1(t).
\]

A comparison of (60) with the perfect-market equilibrium wage rate given in table 12.1 will show that when the labor supply is elastic (i.e., \( \Gamma \neq 0 \)), the wage rate expressed as a fraction of current output will be higher with this system of taxes and transfers than in the perfect-market case. However, the equilibrium quantity of labor will be smaller. That is, from (58) and (60), the equilibrium quantity of labor can be written as

\[
L'(t) = \frac{3N_1(t)[1 - \theta(t)]}{3[1 - \theta(t)] + [(1 + \tau_c)/(1 - \tau_\omega)]\Gamma(1 - \alpha)},
\]

and from table 12.1 and (61), the ratio of the equilibrium quantity of labor in this section to the quantity in the perfect-market case is given by

\[
L'(t)/L(t) = \frac{3[1 - \theta(t)] + (1 - \alpha)\Gamma}{3[1 - \theta(t)] + (1 + \tau_\omega)(1 - \alpha)\Gamma/(1 - \tau_\omega)} < 1.
\]

For the same quantity of capital investment at time \( (t - 1) \), \( I(t - 1) \), the ratio of aggregate output at time \( t \) in the two cases is given by

\[
Q'(t)/Q(t) = \left[ L'(t)/L(t) \right]^{1 - \theta(t)},
\]

and therefore, from (62), equilibrium aggregate output will be lower with this system of taxes and transfers than in the perfect-market case.
Although the magnitude of the reduction in output caused by the distortion of the labor-leisure decision will, of course, depend on the magnitude of $\Gamma$, $\alpha$, and $\theta(t)$, the effect can be substantial. For example, given the same quantity of capital investment at time $(t - 1)$, if $\Gamma = 1$ (i.e., leisure time has the same utility weight as consumption), $\alpha = .25$, and $\theta(t) = \alpha$, the ratio of $Q'(t)/Q(t)$ will be approximately 0.72, which corresponds to a 28% reduction in output. Moreover, because $I'(t)/Q'(t) = I(t)/Q(t)$, $I'(t) < I(t)$, and therefore the reduction in output will become larger through time.

The distorting effects exhibited here of the wage and consumption taxes on the labor-leisure decision are well known and have been discussed at length in the public finance literature. To my knowledge, no general method for eliminating these distortions has been derived. However, for the problem examined here, it is possible to reduce the magnitude of the distortion without affecting the basic functional purposes of the taxes by adding eligibility requirements for retirement benefits to the system of taxes and transfers. Indeed, for the specific model analyzed here, it is possible to eliminate the distortion entirely by such an addition.

Suppose that, in addition to (52) and (54), the system of taxes and transfers is augmented with a schedule of individual retirement benefits which depend in a progressive way on the relative amount contributed by the individual. For example, let the schedule announced at time $t$ for the individual's retirement benefit at time $t + 1$, $b(t + 1)$, be given by

$$b(t + 1) = \max \{\lambda(t), 1 + \gamma(t)[\chi(t) - 1]\} B(t + 1)/N_2(t + 1),$$

where $B(t + 1)$ is the aggregate amount of retirement benefits paid at time $t + 1$, $N_2(t + 1)$ is the number of retirees at time $t + 1$, $\chi(t) = \tau_w[1 - \ell(t)] \omega(t)/\tau_w[1 - \theta(t)] Q'(t)/N_1(t)$ is the ratio of the individual worker's contribution to the average contribution of all workers at time $t$, and $\lambda(t) \geq 0$ and $\gamma(t) > 0$ are policy variables to be chosen. By inspection of (64), we see that the minimum individual retirement benefit paid, $b_{\min}(t + 1)$, is given by $\lambda(t)[B(t + 1)/N_2(t + 1)]$, and that for the schedule to be feasible, $\lambda(t) \leq 1$.

All those retirees whose contributions relative to the average were less than $\chi_{\min}(t) = 1 - [1 - \lambda(t)]/\gamma(t)$ will receive the same minimum retirement payment independent of the specific amount contributed. Provided that $\lambda(t) < 1$, at least some of the retirees will have contributed more than $\chi_{\min}(t)$, and for them the amount of retirement benefits received will be an increasing function of their relative contributions.

Because in the model studied here all work-period people are identical, it follows that in equilibrium $\chi(t) = 1$, and all retirees will receive the same individual retirement benefits which are given by $b(t + 1) = B(t + 1)/N_2(t + 1) = B(t + 1)/N_1(t)$. Therefore, provided that we choose
Social Security as a Means for Efficient Risk Sharing

In equilibrium, \( \chi(t) > \chi_{\min}(t) \) for all workers. Hence, when a worker determines his optimal quantity of leisure time, he will take into account not only the loss of after-tax wages but also the loss of retirement benefits in evaluating the marginal cost of consuming leisure time. Because the magnitude of the marginal loss of benefits depends on the policy variable \( \gamma(t) \), the strategy here is to find that value of \( \gamma(t) \) which will eliminate the distorting effects of the wage and consumption taxes on the labor-leisure decision.

In the relevant region where \( \chi(t) > \chi_{\min}(t) \), aggregate consumption by retirement age people at time \( t + 1 \) can be written as

\[
C'(t+1) = \begin{cases} 
(1 - \tau_{\omega})\omega'(t)[N(t) - L'(t)] + \mu \theta(t)Q'(t) \\
\quad - (1 + \tau_{e})C_1(t) \frac{\theta(t+1)Q'(t+1)}{I'(t)} \\
\quad + \left[ 1 + \gamma(t) \left( \frac{\tau_{\omega}\omega'(t)[N(t) - L'(t)]}{[1 - \theta(t)]Q'(t)\tau_{\omega}} - 1 \right) \right] \\
\quad \tau_{\omega}[1 - \theta(t+1)]Q'(t+1) \end{cases} / (1 + \tau_{e}).
\]

For age 1 people at time \( t \), the first-order condition with respect to current consumption will be the same as in (57a). However, the first-order condition with respect to leisure corresponding to (57b) is now given by

\[
\Gamma = \begin{cases} 
(1 - \tau_{\omega})\omega'(t)\theta(t+1)Q'(t+1)/I'(t) \\
\quad + \gamma(t)\omega'(t)\tau_{\omega}[1 - \theta(t+1)]Q'(t+1)/(1 - \theta(t)]Q'(t) \\
\quad (1 + \tau_{e})C_2(t+1)
\end{cases}
\]

Combining (66) with (57a), we have that

\[
\Gamma = \begin{cases} 
(1 - \tau_{\omega})\omega'(t)/(1 + \tau_{e})C_1(t) + \gamma(t)\omega'(t)\tau_{\omega} \\
\quad [1 - \theta(t)]Q'(t)(1 + \tau_{e}) \\
\quad E_t \left[ [1 - \theta(t+1)]Q'(t+1) \right] \\
\quad C_2(t+1)
\end{cases}
\]

Note that this relation between \( \omega'(t) \) and \( C_1(t) \) differs from the one given in (58) because of the extra cost of leisure time caused by the loss of retirement benefits. As already noted, in equilibrium, \( C_1(t)/Q'(t) = C_1(t)/Q(t) = (1 - \alpha)/3; \ I'(t)/Q'(t) = I(t)/Q(t) = \alpha(1 - \alpha)/(2\alpha + 1); \) and \( \chi(t) = 1 \). Substituting these equilibrium conditions along with the tax parameter values given in (52) into (65), we have that in equilibrium, aggregate retirement-period consumption at \( t + 1 \) can be written as
\[ C_2'(t+1) = \frac{\tau_w}{1 + \tau_c} Q'(t+1). \]

Substituting for \( C_2'(t+1) \) from (68) into (67) and noting that \( E_i[1 - \theta(t+1)] = 1 - \alpha \), (67) can be rewritten as

\[ \frac{\Gamma}{E'(t)} = \frac{\omega'(t)}{C_1'(t)} \left[ \frac{(1 - \tau_w)}{1 + \tau_c} + \frac{\gamma(t)(1 - \alpha)^2}{3[1 - \theta(t)]} \right]. \]

If the policy variable \( \gamma(t) \) is chosen such that the term in brackets in (69) is equal to one, then \( \mathcal{L}'(t) = \Gamma C_1'(t) / \omega'(t) \), and, by comparison with (59), the distortion of the labor-leisure decision by the wage and consumption taxes will be eliminated. Therefore, the optimal value for \( \gamma(t) \) is given by

\[ \gamma(t) = \left[ \frac{3(\tau_w + \tau_c)}{(1 - \alpha)^2(1 - \tau_c)} \right] [1 - \theta(t)] \]

\[ = \left[ \frac{2 + \alpha^2}{(1 - \alpha)^2} \right] [1 - \theta(t)]. \]

A system of taxes and transfers has been derived which causes the economy when human capital is not tradable to replicate the equilibrium path of the corresponding perfect-market economy when human capital is tradable. As summarized in table 12.4, none of the tax or transfer parameters in this optimal system depends on the utility parameter \( \Gamma \), which determines the individual trade-off between labor and leisure time. Hence, essentially the same system will be optimal in the more general case when \( \Gamma \) is permitted to differ across individuals. However, while in the case examined here any value of \( \lambda(t) \) less than one is permissible, care must be taken to choose \( \lambda(t) \) not to be too large in the more general case. Otherwise, the labor-leisure choice for some individuals may be distorted. While a sufficient condition to ensure no such distortion would be to choose \( \lambda(t) = 0 \), there may be other reasons in a more general model why a positive minimum retirement benefit which is independent of individual wage tax contributions would be appropriate.

Finally, it should be pointed out that for the wage tax-retirement benefit part of the system to work, it cannot be a voluntary system. That is, for most values of \( \theta(t) \), the present value of aggregate retirement benefits to current workers will be less than the aggregate wages taxes paid by these workers. Because the economy with this system will replicate the perfect-market equilibrium path, it is straightforward to show that the equilibrium shadow value at time \( t \) of aggregate retirement benefits to be paid at time \( t+1 \) will be equal to \( (1 - \alpha)^2 Q(t) / (2 + 2\alpha - \alpha^2) \). If \( \eta(t) \) denotes the ratio of the shadow value of these aggregate benefits to aggregate current wage tax contributions, then \( \eta(t) \) can be written as

\[ \eta(t) = (1 - \alpha)^2 / [(2\alpha + 1)(1 - \theta(t))]. \]
Table 12.4 An Optimal System of Taxes and Transfers to Correct the Market Failure of No Trading in Human Capital

**Taxes**

A tax on wages with a constant proportional tax rate given by

\[ \tau_w = \frac{2\alpha + 1}{2 + 2\alpha - \alpha^2}. \]

A tax on consumption with a constant proportional tax rate given by

\[ \tau_c = \frac{(1 - \alpha)^2}{2 + 2\alpha - \alpha^2}. \]

**Transfers**

Aggregate transfers to children equal to the total revenues collected by the consumption tax

\[ T_0(t) = \left[ \frac{(1 - \alpha)^2(1 + \alpha + \alpha^2)}{(2\alpha + 1)(2 + 2\alpha - \alpha^2)} \right] Q(t). \]

and individual transfers to each child given by

\[ T_0(t)/N_0(t). \]

Aggregate transfers to retirees equal to the total revenues collected by the wage tax

\[ B(t) = \left[ \frac{(2\alpha + 1)}{(2 + 2\alpha - \alpha^2)} \right] [1 - \theta(t)]Q(t). \]

with individual transfers to each retiree given by

\[ b(t) = \max \left\{ \lambda(t - 1), 1 + \frac{(2 + \alpha^2)}{(1 - \alpha)^2}[1 - \theta(t - 1)]\left[\chi(t - 1) - 1\right] \right\} \frac{B(t)}{N_2(t)}, \]

where \( \chi(t - 1) \) is the dollar wage taxes paid by the individual at time \( t - 1 \) divided by the average dollar wage taxes paid by all workers at time \( t - 1 \), and \( \lambda(t - 1) < 1 \).

By inspection of (71), \( \eta(t) < 1 \) whenever \( \theta(t) < \alpha(4 - \alpha)/(2\alpha + 1) \). Therefore, unless \( \theta(t) \) is approximately two to four times larger than its expected value, \( \alpha, \eta(t) < 1 \) and workers would not, at the margin, voluntarily stay in the system.

### 12.5 Summary, Conclusions, and Extensions

In the model analyzed here, optimal individual life-cycle behavior in perfect markets calls for saving to be invested in the market portfolio. That is, investors would prefer to hold portfolio allocations of physical and human capital which are in proportion to their respective market values. By investing in this way, investors can eliminate factor-share risk. However, when human capital is not tradable, younger members of the economy will have too much of their savings invested in human capital while older members will have too little. Therefore, each will be exposed to factor-share risk. A constrained Pareto-optimal system of taxes and
transfers was derived which corrects this portfolio imbalance and provides to all age groups more efficient risk positions by, in effect, causing their savings to be invested in the market portfolio.

Echoing Samuelson’s (1975) comment about his own model of optimal social security, obviously the severe idealizations of the model presented here will have to be qualified before applying the results. The degree of real-world success this system may have in overcoming the efficiency losses from such a market failure will certainly depend on the reasons for the failure, and because such a failure is simply postulated to exist without any explanation for its cause, the analysis presented here does not deal with this issue. Both the magnitude of the efficiency loss and the detailed specification of the optimal system to correct it are, of course, sensitive to the specific general equilibrium model used to analyze the problem. For example, without the assumption of homothetic and logarithmic utility, all individual investors’ optimal portfolios would not have been identical, and it is therefore unlikely that the simple system of taxes and transfers derived here would have eliminated all the inefficiencies caused by this market failure.¹ Hence, the analysis should be viewed in terms of the qualitative insights it provides for dealing with the inefficiencies caused by this market failure rather than as a quantitative prescription for policy.

With this purpose in mind, I chose to make two further extreme assumptions to both highlight the effect of this market failure and place the heaviest burden on the system derived to correct it. First, I assumed that the market failure was total, that is, that individuals could neither sell their human capital nor borrow against it in any amount. Second, I assumed that there was no intergenerational utility dependence, that is, that parents made no bequests to children and children had no concern for their parents’ welfare. Since certain forms of interpersonal cooperation that are not legally binding can serve as substitutes for either markets or government intervention if there is positive interpersonal utility dependence among the participants,¹⁰ this assumption rules out the possibility of such alternative forms being used to offset the effects of the market failure. While both some amount of marketability of human capital and the existence of intergenerational cooperating family units will tend to soften the impact of this market failure on economic efficiency, and thereby reduce the need for corrective government intervention, further research (possibly along the lines of Barro [1974]) is needed to determine whether or not they would be adequate to eliminate the need for such intervention altogether.

As I noted in the introduction, this model can be used to analyze some of the issues surrounding the present social security system. The wage tax and retirement benefit component of the optimal system derived here bears certain similarities to the funding and retirement benefit part of the
present social security system. In both systems, current wages are taxed to pay current retirement benefits. Moreover, the schedule for determining individual retirement benefits in the social security system is similar to the one presented in equation (64) for the system derived here. As was shown in (71) for the optimal system, and as is alleged by some for the present social security system, the present value of retirement benefits for current workers will generally be less than the wage tax contributions made by current workers. Hence, both pay-as-you-go systems require compulsory participation.

Of course, there are also differences between the structures of the two systems. The optimal system derived here requires that aggregate benefits always be equal to current aggregate tax revenues. Under the present social security system, benefits and tax rates are determined separately by law and, with the exception of the financial solvency constraint, the existing law does not require that current benefits be equal to current tax revenues. Because the wage tax rate is constant over time, aggregate benefits in the optimal system will change in a fashion perfectly correlated with changes in aggregate wage income. Therefore, individual benefits in the optimal system are, de facto, indexed to aggregate wage income divided by the number of retirees. In contrast, under present law, individual social security benefits are indexed to the Consumer Price Index.

By construction, the optimal system can never become insolvent in the sense that revenues raised, both currently and in the future, will never be insufficient to pay promised benefits, both currently and in the future. In contrast, the present social security system can become insolvent if promised future benefits and tax rates are defined to be equal to the current schedule of benefits and tax rates. However, since Congress can and has changed existing law with respect to both benefits and tax rates, the only strictly vested benefits in the system are the current ones, and even these are limited by current tax revenues if the available trust funds should become exhausted. Although in principle Congress could keep the schedule of benefits fixed and correct any deficits or surpluses in the system by changing tax rates, normal congressional behavior appears to be to make changes in both benefits and tax rates. Therefore, aggregate social security benefits are likely to be strongly correlated with aggregate wage income, especially in the intermediate to long run. Hence, as long as such benefits are funded solely by a wage tax, the pattern of retirement benefits from social security may be a reasonable approximation to the pattern of retirement benefits generated by the optimal system derived here.

These derived similarities between the two systems may cast some light on the widely discussed issue of whether social security should be a pay-as-you-go or fully funded system. These derived similarities suggest
that a latent function served, at least in part, by the present pay-as-you-go social security system is that of the optimal system derived here, namely, to improve the efficiency of risk bearing in the economy when human capital is not tradable. Indeed, the returns from a fully funded system which invests its contributions in traded securities cannot possibly replicate the returns from investing in a nontraded asset except in the singular case where a traded security exists whose returns are perfectly correlated with those of the nontraded asset. Hence, any system which attempts to replicate the returns from such a nontraded asset must at least appear to be a pay-as-you-go system. When there are significant nontraded assets in the economy, the creation of such a system will cause changes in equilibrium consumption, private saving, and portfolio allocation behavior, although the direction of these changes is, in general, ambiguous. However, as demonstrated in the model analyzed here, the effect of introducing such a system can be to increase economic efficiency whichever direction these changes take.

This analysis does not imply that the present social security system is optimal. Even if it exactly replicated the optimal system derived here, the present social security system would only be optimal if the economic objectives of the system were the same as those of the optimal system. The optimal system presented here is not designed to be the sole, or even the major, source of retirement benefits. Rather it is designed to complement private saving by providing only benefits which (by hypothesis) cannot be purchased in the private market. If the system were to be a general substitute for private saving for retirement (as at least some have suggested is the purpose of social security), then the benefits should also be linked to the returns on physical capital, and these benefits should be funded by additional taxes with such revenues invested in physical capital. That is, the optimal system for this purpose would be partially funded.

As I noted in the introduction, the absence of riskless real annuities was one of the three market failures explicitly discussed by Diamond (1977) as possible reasons for social security. Although Diamond (1977, p. 277) claims that "someone reaching retirement age with a capital sum might reasonably want to purchase a real annuity," only in very singular cases would a person at retirement optimally choose a lifetime annuity whose payments are riskless (even in real terms). As was true for the optimal consumption and portfolio decisions during the accumulation period of their lives, people will generally prefer to bear some amount of risk with respect to their retirement payments in return for a higher level of expected returns. In the model analyzed here, the optimal choice for retirees would be a life annuity whose payments depended on the returns on the market portfolio, and these are certainly not riskless. Thus, the
result derived here, that benefits received by retirees are uncertain, should not be viewed as somehow suboptimal.

To reinforce this point at a somewhat more applied level of analysis, note that for a 15-year expected life and a 12% nominal interest rate an actuarially fair, nominally fixed annuity would generate an annual nominal cash flow equal to about 15% of the initial capital sum. For a 4% real interest rate (a number which at times has been suggested to be the long-run average real rate of growth in the economy), an annuity, fixed in real terms, would generate a first-year cash flow of about 9% of the initial capital sum. However, such an annuity would not be actuarially fair because to earn that average or expected real rate of 4%, the provider of the annuity would have to bear the aggregate risks of the economy, which are not diversifiable and certainly not zero. A more appropriate indicator of the proper rate to be applied to a real riskless annuity would be the historical average real return from rolling over short-term Treasury bills which, on a pretax basis, is approximately zero. Hence, an actuarially fair riskless real annuity would generate a first-year cash flow of about 6.7% of the initial capital sum. Therefore, to provide a rather modest first-year retirement income of $13,300, the capital accumulation would have to be $200,000, a considerable sum.

Because it is assumed that the durations of each person's work and retirement periods are exogenous and known with certainty, the model in its present form cannot be used to analyze the other problems of market failures discussed in Diamond (1977) where these durations are uncertain. However, the model can be extended along the lines of the Sheshinski and Weiss (1981) analysis of failure in the annuities market, to take into account durations which are exogenously stochastic. Moreover, because the present model does include a labor-leisure choice in the work period, it should be straightforward to adapt the model to the case where the length of the work period is endogenous, that is, where workers can voluntarily choose early retirement.

As a closing note, the analysis presented here indicates that the nontradability of human capital will, in general, make the solution of distortion problems caused by taxes more difficult. For example, if a proportional consumption tax were proposed to raise revenues for general government expenditures, then, by inspection of (58) and (59), it appears that a wage subsidy (i.e., a negative tax) of $\tau_\omega = -\tau$ would eliminate the distortion of the labor-leisure choice. However, such a negative wage tax will only make worse the problems of efficient risk bearing when human capital cannot be traded because the young will now find themselves forced to hold an even larger proportion of their savings in human capital.
Appendix: Optimal Consumption and Portfolio Decisions

Using the method of stochastic dynamic program, the individual optimal consumption and portfolio rules given in the text are derived.

Define the derived or indirect utility function for a person of age $k$ at time $t$ by

\begin{equation}
J_k[w_k(t), t] = \max \{U_k(t)\}, k = 0, 1, 2
\end{equation}

where $w_k(t)$ is the wealth of the person at age $k$ and $U_k$ is the lifetime utility of consumption function defined in assumption (5), (6), and (7). In the usual fashion of dynamic programming, the optimal solution is derived by working backward. From (7), it follows immediately that optimal consumption at retirement, $c_2(t)$, is simply given by

\begin{equation}
c_2(t) = w_2(t),
\end{equation}

and therefore, from (A1),

\begin{equation}
J_2[w_2(t), t] = \log[w_2(t)].
\end{equation}

At age 1, the derived utility of wealth function is given by

\begin{equation}
J_1[w_1(t), t] = \max \{r \log[w(t)] + \log[c_1(t)]
+ E_t\{J_2[w_2(t+1), t + 1]\}\.
\end{equation}

From (8) and (10b), we have that

\begin{equation}
w_2(t + 1) = [w_1(t) - c_1(t) - \omega(t)\ell(t)]x_{11}(t)[Z_1(t + 1)
- R(t)] + x_{21}(t)[Z_2(t + 1) - R(t)] + R(t).
\end{equation}

Substituting for $J_2$ from (A3) and for $w_2(t + 1)$ from (A5) into (A4), and maximizing with respect to the choice variables $\{c_1(t), \ell(t), x_{11}(t), x_{21}(t)\}$, we have the following first-order conditions:

\begin{equation}
c_1^*(t) = [w_1(t) - \omega(t)\ell^*(t)]/2,
\end{equation}

\begin{equation}\nonumber
\ell^*(t) = \Gamma[w_1(t) - c_1^*(t)]/(1 + \Gamma)\omega(t),
\end{equation}

and

\begin{equation}
0 = E_t\left\{\frac{Z_j(t + 1) - R(t)}{x_{1j}(t)[Z_1(t + 1) - R(t)]
+ x_{2j}(t)[Z_2(t + 1) - R(t)] + R(t)}\right\}, j = 1, 2.
\end{equation}

From (A6) and (A7), it follows that

\begin{equation}
c_1^*(t) = \frac{w_1(t)}{2 + \Gamma}.
\end{equation}
(A10) \[ \ell^*(t) = \frac{\Gamma w_1(t)}{(2 + \Gamma)\omega(t)}, \]

and

(A11) \[ s^*_1(t) = \frac{w_1(t)}{2 + \Gamma}. \]

These optimal rules are reported in (12) in the text.

Substituting for these optimal rules in (A4), we have that

\[ J_i[w_1(t), t] = (2 + \Gamma)\log[w_1(t)] - \Gamma\log[\omega(t)] + \Gamma\log\Gamma - (2 + \Gamma)\log\Gamma + E_t[\log[Z^*_1(t + 1)]], \]

where \( Z^*_1(t + 1) \) is the return per dollar on the optimal portfolio which satisfies (A8).

At age 0, the derived utility of wealth function is given by

(A13) \[ J_0[w_0(t), t] = \max \left( \log[c_0(t)] + E_t[J_i[w_1(t + 1), t + 1]] \right), \]

where, from (8) and (10a) in the text, we have that

(A14) \[ w_1(t + 1) = [w_0(t) - c_0(t)]x_{10}(t)[Z_1(t + 1) - R(t)] + x_{20}(t)[Z_2(t + 1) - R(t)] + R(t). \]

Substituting for \( J_i \) from (A12) and \( w_1(t + 1) \) from (A14) into (A13), and maximizing with respect to the choice variables \( \{c_0(t), x_{10}(t), x_{20}(t)\} \), we have the following first-order conditions:

(A15) \[ c^*_0(t) = \frac{w_0(t)}{3 + \Gamma}, \]

(A16) \[ 0 = E_t \left\{ \frac{Z_j(t + 1) - R(t)}{x^*_{10}(t)[Z_1(t + 1) - R(t)] + x^*_{20}(t)[Z_2(t + 1) - R(t)] + R(t)} \right\}, j = 1, 2, \]

and from (10a) and (A15), it follows that

(A17) \[ s^*_0(t) = \left( \frac{2 + \Gamma}{3 + \Gamma} \right)w_0(t). \]

These optimal age 0 consumption and saving rules are reported in (11).

By inspection of (A8) and (A16), the fractional allocations in the optimal portfolios of age 0 and age 1 people are identical. Hence, all investors will optimally hold the same relative proportions of securities, and these common optimal portfolio weights are given in the text by (13).
Notes

1. While this characterization of optimal portfolio choice is usually identified with the mean-variance model of portfolio selection, such broad diversification is, indeed, a property of most optimal strategies for risk averters. See Merton (1982), sections III and IV, and especially proposition 4.2.

2. For another overlapping-generations model with the same three-period life, see the appendix to Sheshinski and Weiss (1981). As will be apparent later in my analysis, in the absence of bequests three is the minimum number of periods required in order for trading in human capital (i.e., future wage income) to take place.

3. Of course, this does not imply that such cross-sectional risks among workers are believed to be unimportant.

4. It is, of course, reasonable to expect that population growth will be influenced by the level of aggregate economic activity. For further discussion and a simple model which incorporates such population dependencies, see Merton (1969). However, unlike in the social insurance models of Green (1977) and Smith (1981), demographic uncertainties play no essential role in this analysis.

5. This result is most closely associated with the equilibrium conditions in a mean-variance portfolio model with homogeneous beliefs. However, as seen here, it does hold in other cases, including every model with a representative man. As noted in Merton (1982, sec. IV), “Indeed, if there were one best investment strategy, and if this ‘best’ strategy were widely known, then whatever the original statement of the strategy, it must lead to simply this imperative: ‘hold the market portfolio.’”

6. Indeed, under the extreme assumptions of no borrowing against or sale of human capital and no bequests, children, and therefore society, could not survive without such correction.

7. The terms “pay as you go” and “fully funded” have been used in a variety of ways in the literature. I use them as they are defined in Sheshinski and Weiss (1981, p. 189): a system is “fully funded” if contributions to the system are invested at the market rate of interest and a system is “pay as you go” if taxes on the currently working population are used to finance benefits to the retired population. For a brief description of the present social security system, see Diamond (1977).

8. See Atkinson and Stiglitz (1980) for a general discussion of the consumption and wage taxes and their distortion of the labor-leisure choice.

9. However, a similar system will work to eliminate the inefficiencies in the somewhat more general case where lifetime utility is given by

\[ U_d(t) = \log[c_0(t) + a_0] + E_t[\Gamma \log[F(t+1) + \beta] + \log[c_1(t+1) + a_1] + \log[c_2(t+2) + a_2]. \]

Note: \( U_d(t) \) is not homothetic. For further discussion of the properties of this utility function, see Rubinstein (1976).

10. See Kotlikoff and Spivak (1981) for an example of such substitution. Kurz (1981) provides some empirical evidence which rejects this extreme no-bequest life-cycle assumption of behavior.

11. For the formula used in the present social security system to determine individual benefits, see Diamond (1977), p. 276, Eq. (1) and n. 8.

12. As mentioned in note 4 above, demographics do not significantly affect the analysis of inefficiencies in this model. Indeed, for a given level of aggregate output, the population size or its age distribution has no effect on aggregate consumption and saving or their distributions among age groups. However, since these aggregates, including aggregate retirement benefits, have this property, per capita consumption and therefore individual welfare are significantly affected. It should be noted that these effects on individual welfare
caused by demographics are identical in both the perfect market and optimal tax and transfer economies.

13. Diamond (1977, p.277) reports that "Congressional attitude appears to be that it is appropriate to increase benefits whenever the system can finance such an increase over the following 75 years . . . ." Since 1977, there have been increases in payroll taxes voted by Congress, and at the current time serious consideration is being given to the reduction of benefits in response to the belief that revenues will not be adequate to fund future benefits at the current levels.


References

———. 1982. On the microeconomic theory of investment under uncer-


