Chapter 5

Applying and extending the Diamond model

This chapter considers applications and extensions of the Diamond overlapping generations model. We start with an examination of *pension schemes* and their effect on aggregate saving and capital accumulation. The next sections introduce *endogenous labor supply* and *retirement* from the labor market. This provides a framework for an analysis of how a voluntary early retirement scheme affects aggregate labor supply and wealth accumulation in a small open economy. The last section considers multi-period endogenous labor supply problems with a focus on intertemporal substitution of labor supply.

5.1 Pension schemes and aggregate saving

By the *dependency ratio* is meant the number of retired people in proportion to the size of the working age population. In the Diamond model, with a constant population growth rate n, the dependency ratio is simply given by $1/(1+n)$. The demographic development after the second world war has entailed falling population growth and rising life expectancy. As a result there are now relatively fewer "young" (working age population) and a higher dependency ratio. This phenomenon is referred to as the "ageing society" or, with less piety, the "greying society". Many developed countries rely primarily on an unfunded national retirement pension scheme. There is increasing concern about how to finance *retirement pensions* in the future. A retirement pension is a stream of payments to an individual, starting at the time of retirement and continuing until death.

Overlapping generations models provide an appropriate framework for studying the macroeconomic effects and intergenerational distribution aspects of different retirement pension schemes. On the basis of a simple extension of the

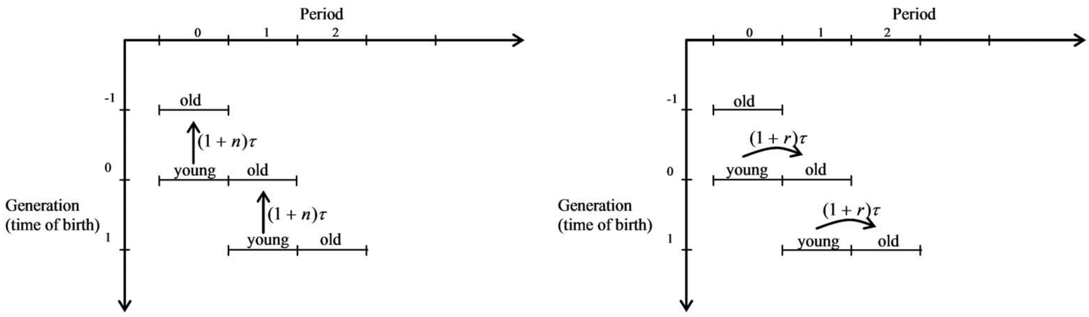

Figure 5.1: Pay-as-you-go (left panel) and funded system (right panel) ($\tau_t = \tau_{t+1} = \tau$).

Diamond OLG model we will compare the effects on aggregate saving and capital accumulation of different systems of national retirement pension provision: a funded pension system (i.e., a saving-based system, in Denmark for example the ATP system) and a tax-based pension system (sometimes called a pay-as-you-go system or just an unfunded system, in Denmark named "folkepension"). The benchmark case is the Diamond model without any national pension scheme at all, here named System 0.

We consider a closed economy described by the Diamond OLG model. For simplicity, technological progress is ignored. Let the pension received by an old person in period $t + 1$ be called p_{t+1} (p for pension) and let the mandatory (i.e., required by law) lump-sum contribution of a young person in period t be called τ_t. Otherwise, the notation is as in the previous chapter. The pension arrangements are as follows:

$$\text{Funded system:} \qquad p_{t+1} = (1 + r_{t+1})\tau_t, \tag{5.1}$$

$$\text{Tax-based system:} \qquad p_{t+1} = (1 + n)\tau_{t+1}. \tag{5.2}$$

Fig. 5.1 illustrates the two systems. In the funded system the mandatory contributions of the young are collectively *invested* and returned with interest in the *next* period. The system is a form of collective saving for old age. The tax-based system is different in that the contributions of the young are used to finance pensions in the *same* period. Thus the mandatory contribution is like a lump-sum tax on the young which finances current government pension expenditure. The system can be seen as a kind of social contract: the currently young pay the pension of the currently old and are "paid back" in the next period by a transfer from that period's young, each of them paying τ_{t+1}. Since for every old there are (on average) $1+n$ young, the pension to each old is $(1+n)\tau_{t+1}$. This immediately displays the tension generated by a decline in n.

To be more specific, let the utility function of the young born at the beginning of period t be

$$U(c_{1t}, c_{2t+1}) = u(c_{1t}) + (1+\rho)^{-1}u(c_{2t+1}),$$

where $\rho > 0$,[1] $u' > 0$, $u'' < 0$, and $\lim_{c\to 0} u'(c) = \infty$ (the No Fast Assumption). Assume for simplicity that there are no other taxes than the mandatory pension contribution and no other government expenditures than pensions. The young chooses saving, s_t, subject to the two period budget constraints, which are, respectively:

$$\text{No national pension scheme:} \quad c_{1t} + s_t = w_t, \quad c_{2t+1} = (1+r_{t+1})s_t.$$

$$\text{Funded system:} \quad c_{1t} + s_t = w_t - \tau_t, \quad c_{2t+1} = (1+r_{t+1})s_t + (1+r_{t+1})\tau_t.$$

$$\text{Tax-based system:} \quad c_{1t} + s_t = w_t - \tau_t, \quad c_{2t+1} = (1+r_{t+1})s_t + (1+n)\tau_{t+1}.$$

5.1.1 No national pension scheme

This is the case described by the original Diamond model of Chapter 3. Let the saving of the young individual in period t in case of no national pension scheme be called s_t. Then general equilibrium under this regime is described by the following four equations:

$$u'(w_t - s_t) = (1+\rho)^{-1}u'((1+r_{t+1})s_t)(1+r_{t+1}), \tag{5.3}$$

$$w_t = f(k_t) - k_t f'(k_t) \equiv w(k_t), \tag{5.4}$$

$$r_{t+1} = f'(k_{t+1}) - \delta \equiv r(k_{t+1}), \tag{5.5}$$

$$k_{t+1} = \frac{s_t}{1+n}. \tag{5.6}$$

The equation (5.3) is the first-order condition (the Euler equation) derived from the decision problem of the young. Equations (5.4) and (5.5) give the equilibrium real wage and interest rate, respectively. These are determined from firms' profit maximization under perfect competition and the assumption of market clearing, implying that the capital-labor ratio chosen by firms equals the capital-labor ratio from the supply side, $k_t = K_t/L_t$. Finally, equation (5.6) comes from the identities $K_{t+1} - K_t = S_t^N = S_{1t} + S_{2t}$, where S_{1t} and S_{2t} are the aggregate (net) saving

[1] Although the results to be derived are valid for any $\rho > -1$, we assume $\rho > 0$ to help intuition.

of the young and the old, respectively. Now $S_{1t} = L_t s_t$ and $S_{2t} = -K_t$ (the old generation enters period t with wealth K_t and leaves period t, and life, with zero wealth). Hence, $K_{t+1} - K_t = L_t s_t + (-K_t)$ so that $K_{t+1} = L_t s_t$. Combining this with the demographic assumption $L_{t+1} = L_t(1+n)$, we get

$$K_{t+1} = k_{t+1} L_t (1+n) = L_t s_t,$$

from which follows (5.6).

If the current period is period t, then k_t is predetermined by previous capital accumulation, whereas w_t, s_t, r_{t+1}, and k_{t+1} are endogenous. We have thus four endogenous variables and four equations. The causal structure is block-recursive. The real wage w_t is determined by equation (5.4), through the predetermined k_t, independently of the three other variables. But the value of w_t does affect these variables. Indeed, s_t, r_{t+1}, and k_{t+1} are determined simultaneously by the equations (5.5), (5.6), and (5.3), given the value of w_t. Therefore, there is a causal relationship going from k_t via w_t to s_t, r_{t+1}, and k_{t+1} simultaneously.[2]

Let the solution for s_t in the above situation (no national pension scheme) be denoted s_t^0.

5.1.2 Funded system

In a funded system the equations (5.4) and (5.5) are unchanged, but (5.3) and (5.6) are replaced by

$$u'(w_t - (s_t + \tau_t)) = (1+\rho)^{-1} u'((1 + r_{t+1})(s_t + \tau_t))(1 + r_{t+1}), \tag{5.7}$$

$$k_{t+1} = \frac{s_t + \tau_t}{1+n}. \tag{5.8}$$

The mandatory contribution τ_t per young is invested in capital by the government or the social security administration and gives the normal gross return $1 + r_{t+1}$ next period. There are two cases to consider:

Case a: $\tau_t \leq s_t^0$. In this case the young people can and will fully offset the savings which the social security administration does on their behalf. Indeed, since (w_t, s_t^0, r_{t+1}, k_{t+1}) satisfies the system (5.3), (5.4), (5.5), and (5.6), (w_t, $s_t + \tau_t$, r_{t+1}, k_{t+1}) with $s_t + \tau_t = s_t^0$ satisfies the system (5.7), (5.4), (5.5), and (5.8). Given that the funded system provides the same rate of return as private saving, the young just reduce their private saving by an amount equal to the mandatory contribution, that is, they set $s_t = s_t^0 - \tau_t$. In this way they end up

[2] In the language of *causal ordering,* w_t is determined at zero order in the causal structure, whereas s_t, r_{t+1}, and k_{t+1} are then simultaneously determined at first order in the "block" consisting of (5.5), (5.6), and (5.3).

with the same total return as before, namely $(1 + r_{t+1})(s_t + \tau_t) = (1 + r_{t+1})s_t^0$. Thus the funded pension system has no effect, neither on aggregate saving and capital accumulation nor on any individual's consumption over lifetime. The social security administration is just doing some of the saving for the young. The system is neutral.

Case b: $\tau_t > s_t^0$. This may not be a realistic case; yet, from a theoretical perspective it is worth pursuing its logic. Aggregate desired private saving of the young (the working) generation is now negative, $s_t L_t = (s_t^0 - \tau_t)L_t < 0$. That is, the young want to borrow. There are two sub-cases to consider.

Sub-case b1. Imagine the law says that the social security administration must invest all its funds in physical capital and rent it out to the firms. Then there are nobody a young person can borrow from. The other young also want to borrow and the old do not want to lend because they are not interested in postponing consumption until next period where they will be dead. So the desire of the young to borrow is frustrated and their actual saving, $\bar{s}_t L_t$, ends up equal to zero. Since in this case (5.8) becomes $k_{t+1} = \tau_t/(1 + n) > s_t^0/(1 + n)$, the funded pension system is no longer neutral. By forcing aggregate saving and investment in society to be above what it would be in the absence of social security, the pension system acts like a law of "forced saving".

Sub-case b2. Here we imagine the social security administration is allowed to place its means in interest-bearing deposits in banks as well as in physical capital, depending on where the return is highest. Then the equilibrium ends up the same as in case *a*. Out of the aggregate contribution, $\tau_t L_t$, the social security administration invests $s_t^0 L_t$ in physical capital at the end of period t and rents it out to the firms at the rental rate $r_{t+1} + \delta$. The remainder, $\tau_t L_t - s_t^0 L_t = -s_t L_t$, is lent to the banks and these lend it to the young who in the next period will repay the loan with interest r_{t+1}, again via the banks. Assuming the administrative costs of banking are vanishing, the net rate of return to the social security administration on its two kinds of placement is the same, r_{t+1}. The described allocation is an equilibrium (demand equals supply in all markets, all desired actions are realized) and this equilibrium is exactly the same as in case *a*.

The intuitive mechanism behind this equilibrium is the following. Initially the young individual who wishes to borrow the amount $-s_t$, but faces a binding constraint, is willing to pay interest $r_{t+1} + \varepsilon_1$, where ε_1 is some small positive number. The banks are eager to supply these loans and offer the pension administration the interest rate $r_{t+1} + \varepsilon_2$, where $0 < \varepsilon_2 < \varepsilon_1$. Now, competition among the banks drives ε_1 and ε_2 down to zero. The conclusion is that the funded system is neutral.

5.1.3 Tax-based system

In a tax-based system (5.3) is replaced by

$$u'(w_t - (s_t + \tau_t)) = (1+\rho)^{-1}u'((1+r_{t+1})s_t + (1+n)\tau_{t+1})(1+r_{t+1}), \qquad (5.9)$$

while (5.4), (5.5), and (5.6) from the system without a national pension scheme are maintained. The capital formation equation thus again reads

$$k_{t+1} = \frac{s_t}{1+n} \qquad (5.10)$$

in contrast to (5.8) of the funded system. This system is not neutral to aggregate saving. To show this, we assume for simplicity that $\tau_t = \tau > 0$ for all t.

Partial equilibrium effect of a rise in τ By the implicit function theorem, the equation (5.9) defines s_t as an implicit function of the three variables, w_t, r_{t+1}, and τ :

$$s_t = s(w_t, r_{t+1}, \tau).$$

Although in the end we are looking for the general equilibrium effect of a change in τ, we get some insight by first considering the partial effect of a change in τ. That is, we increase τ while keeping w_t and r_{t+1} unchanged. By substituting $\tau_t = \tau_{t+1} = \tau$ into (5.9) and using implicit differentiation with respect to τ on both sides, we get

$$u''(c_{1t})(-\frac{\partial s_t}{\partial \tau} - 1) = (1+\rho)^{-1}u''(c_{2t+1})[(1+r_{t+1})\frac{\partial s_t}{\partial \tau} + 1 + n](1+r_{t+1}).$$

By ordering, we find the partial derivative of the implicit saving function s,

$$s_\tau = \frac{\partial s_t}{\partial \tau} = -\frac{u''(c_{1t}) + (1+\rho)^{-1}u''(c_{2t+1})(1+n)(1+r_{t+1})}{u''(c_{1t}) + (1+\rho)^{-1}u''(c_{2t+1})(1+r_{t+1})^2} < 0. \qquad (5.11)$$

The negative sign comes from both the numerator and the denominator being negative.[3] The explanation of the negative sign is that a higher τ implies higher pension as old and therefore less need for your own saving.

The effect on private saving is dampened or fortified by the wealth effect of the tax-based system, depending on the sign of this effect. To see this we write down

[3] The derivation of (5.11) could alternatively be based on "total differentiation" in terms of differentials. We use that method in connection with (5.14) and (5.15) below, just to prepare the reader for the diverse approaches to implicit differentiation applied in the literature.

the *intertemporal* budget constraint implied by the two period budget constraints given in (II). We find

$$c_{1t} + \frac{c_{2t+1}}{1+r_{t+1}} = w_t - (1 - \frac{1+n}{1+r_{t+1}})\tau = w_t - \frac{r_{t+1}-n}{1+r_{t+1}}\tau.$$

The right-hand side is the present value of a young's lifetime income evaluated at the end of period t. Since the young is born with no financial wealth, this equals the *total wealth* of the young. The total wealth is seen to decrease or increase with a rise in τ depending on whether $r_{t+1} > n$ or $r_{t+1} < n$, respectively. This explains why (5.11) gives $|\partial s_t/\partial\tau| \lesseqgtr 1$ for $r_{t+1} \gtreqless n$, respectively. Indeed, the marginal propensity to consume out of wealth is positive, so that $\partial c_{1t}/\partial\tau \lesseqgtr 0$ for $r_{t+1} \gtreqless n$, respectively. In view of $s_t = w_t - \tau - c_{1t}$, we have $\partial s_t/\partial\tau = -1 - \partial c_{1t}/\partial\tau$. Now, a rise in τ always decreases after-tax income as young. Along with this, when $r_{t+1} > n$, the rise in τ also decreases consumption as young, and so $-1 < \partial s_t/\partial\tau < 0$. When $r_{t+1} < n$, however, a rise in τ *increases* consumption as young; so saving as young is squeezed from both sides and we get $\partial s_t/\partial\tau < -1$.

This is only part of the story, however. There are *general equilibrium* effects on wages and interest rates as soon as lower saving by the young leads to less capital in the economy. Indeed, the immediate effect of a fall in s_t is that $k_{t+1} \downarrow \Longrightarrow r_{t+1} \uparrow$, the expectation of which has a feedback effect on s_t, which is positive (negative) if $s_r > 0$ ($s_r < 0$). Thus the feedback effect counteracts the partial equilibrium effect, if $s_r > 0$, and strengthens it, if $s_r < 0$.[4]

General equilibrium effect of a rise in τ To take all effects into account we need general equilibrium analysis. That is, we have to consider the fundamental difference equation of the model under the tax-based system. Thus we insert $s_t = s(w_t, r_{t+1}, \tau)$ into (5.10) and then substitute (5.4) and (5.5) to get

$$(1+n)k_{t+1} = s(w(k_t), r(k_{t+1}), \tau). \tag{5.12}$$

Presupposing the denominator in (5.14) and (5.15) below is not vanishing, this equation determines k_{t+1} as an implicit function of τ and k_t. A convenient approach to the derivation of the partial derivatives of this function is to first *take the differential* of each side of (5.12) with respect to k_{t+1}, k_t, and τ. This gives

$$(1+n)dk_{t+1} = s_w\ w'dk_t + s_r\ r'dk_{t+1} + s_\tau\ d\tau, \tag{5.13}$$

[4] Recall that we are concerned with a closed economy. In a small open economy with perfect mobility of financial capital and domestic and foreign financial claims being perfect substitutes, firms' choice of k will be determined by the real interest rate in the world financial market, which is not affected by a change of τ. Hence, in this case there is no feedback effect and the partial equilibrium analysis is the end of the story (see Exercise 5.??).

where $w' = -k_t\ f''(k_t) > 0$, $r' = f''(k_{t+1}) < 0$, and $s_\tau = \partial s_t/\partial\tau < 0$ from (5.11). By ordering, we find the general equilibrium "short-run multipliers",

$$\frac{\partial k_{t+1}}{\partial \tau} = \frac{s_\tau}{1+n-s_r f''(k_{t+1})} \quad \text{and} \tag{5.14}$$

$$\frac{\partial k_{t+1}}{\partial k_t} = \frac{-s_w k_t f''(k_t)}{1+n-s_r f''(k_{t+1})}. \tag{5.15}$$

The numerator in (5.14) is negative and that in (5.15) is positive. The denominator is the same in both expressions and is positive in the convenient case where

$$s_r(w(k_t), r(k_{t+1}), \tau) > (1+n)/f''(k_{t+1}) \tag{5.16}$$

for all pairs (k_t, k_{t+1}) consistent with an equilibrium path. This is the case where the transition curve is positively sloped everywhere, in Chapter 3 named The Positive Slope Assumption. It holds when the income effect on c_{1t} of a change in the interest rate does not dominate the substitution effect "too much". A sufficient, although not necessary, condition for this is that $s_r \geq 0$.

From now we assume (5.16) to hold. Then the slope, $\partial k_{t+1}/\partial k_t$, of the transition curve is positive everywhere, as in Fig. 5.2. And we have $\partial k_{t+1}/\partial\tau < 0$ so that a shift to a new pension contribution $\tau' > \tau$ implies a downward shift of the transition curve. To get an intuitive understanding of this downward shift, consider (5.13) under the condition $dk_t = 0$:[5]

$$(1+n)dk_{t+1} = s_r\ r' dk_{t+1} + s_\tau d\tau. \tag{5.17}$$

There are *two* influences on dk_{t+1} on the left-hand side of (5.17). First, there is the influence coming from the *direct effect* on s_t of the rise in the mandatory contribution τ. This effect, which according to (5.11), is negative, is on the right-hand side of (5.17) represented by the last term, $s_\tau d\tau$. Second, unless $s_r = 0$, there is an influence on dk_{t+1} coming from the *indirect effect* on s_t arising through the higher expected and actual r_{t+1}, caused by an incipient reduction in k_{t+1}. This influence is represented by the first term, $s_r r' dk_{t+1}$, on the right-hand side of (5.17).

The sign of this indirect effect on s_t is ambiguous because the sign of s_r is ambiguous.

Consider first the case $s_r < 0$. Combined with (5.16) this amounts to $(1+n)/f''(\cdot) < s_r < 0$. Since $r' = f''(\cdot) < 0$, the *tendency* to a negative value of dk_{t+1} will in this case make the feedback $s_r r' dk_{t+1}$ *negative,* thus reinforcing the tendency to a negative value of dk_{t+1} on the left-hand side of (5.17). Indeed, the

[5] The condition $dk_t = 0$ is relevant because we want to understand why, for *given* k_t, a rise in τ implies lower k_{t+1} than otherwise.

young will anticipate a rise in the interest rate as a result of the higher τ and respond to this by a further lowering of their saving, thereby confirming their anticipation.

Next, consider the opposite case, $s_r > 0$. The *tendency* to a negative value of dk_{t+1} will in this case make the feedback, $s_r r' dk_{t+1}$, *positive.* The direct negative influence on dk_{t+1} coming from $s_\tau d\tau$ on the right-hand side of (5.17) is thus partly offset by the positive feedback on the saving of the young when they face a higher interest rate. But only *partly* offset. This is seen by (5.14) where the numerator is $1 + n - s_r f''(k_{t+1}) > 1 + n > 0$ when $s_r > 0$. Intuitively, when $s_r > 0$, the positive response of saving to a higher interest rate can *mitigate* the tendency to a lower k_{t+1} which lies behind the higher interest rate, but not turn it into its opposite in the *same* period. Indeed, the positive feedback on the saving of the young will only be there *if* the interest rate rises in the first place. We cannot in the *same* period have both a *rise* in the interest rate that triggers higher saving *and* a *fall* in the interest rate *because* of the higher saving.

Dynamics In addition to the No Fast Assumption and the Positive Slope Assumption, cf. (5.16), we assume that our Diamond economy without any national pension scheme satisfies the Early Steepness Assumption, (A3), from Chapter 3. We further assume that the original mandatory contribution, τ, is not larger than to allow existence of at least two (non-trivial) steady states, as in Fig. 5.2.[6] The interesting steady state is the stable one, k^*. The rise in the mandatory contribution to τ' (still allowing existence of two steady states) shifts the steady-state value of k to $k^{*\prime} < k^*$. Hence, if the economy was initially in the old steady state, k^*, the shift to τ' implies a *decrease* in capital both in the short and the long run.

To fix ideas, suppose the shift from τ to $\tau' > \tau$ occurs in the beginning of period t_0. There are two reasons that the new steady state has lower k. First, the immediate effect of the upward shift in the mandatory contribution is to lower the saving, s_{t_0}, by the young as explained above. Second, the resulting lower k_{t_0+1} next period implies lower wage income, w_{t_0+1}, next period than otherwise and thereby a further reduction in the saving of the next young generation, s_{t_0+1}. Although the expected interest rate, $r^e_{t_0+2}$, has risen, this can at most *mitigate* the tendency to a lower k_{t_0+2} (not turn it around), as explained above. In the following periods the contraction process continues, but each further fall in k_t becomes smaller and smaller (the slope of the transition curve, although positive, is less than one). Ultimately the economy thus comes infinitely close to $k^{*\prime}$. Note that this fall in aggregate capital accumulation is really due to the reduction in the private incentive to save (a reduction which is also present in the funded

[6] Below we return to the issue why, for a fixed (and not too large) τ, there are, in general, at least two non-trivial steady states if any.

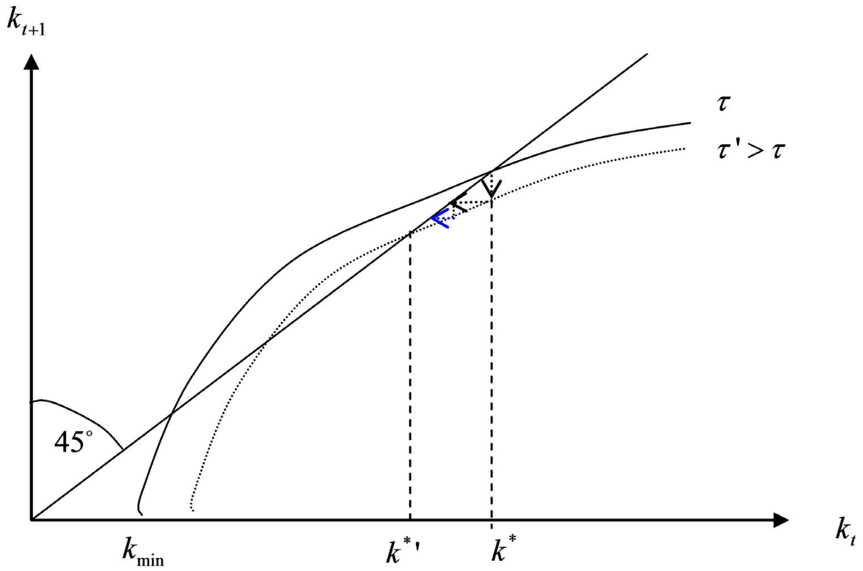

Figure 5.2: Shift in the transition curve associated with a shift from τ to τ' in the tax-based system (the case $f(0) = 0$).

system) *combined with* the property of the considered tax-based system that the mandatory contribution is not invested, but immediately transferred to the current period's old generation.

But why is it that, for a fixed and not too large τ, there is, besides the stable steady state, also an unstable one? Essentially, the reason is that the saving of the young is always less than the income of the young, i.e., $s_t < w_t - \tau = w(k_t) - \tau$. Since the real wage depends positively on the capital-labor ratio, it follows that for a given τ, a very small k_t implies that s_t (hence k_{t+1}) can not be positive.[7] In Fig. 5.2 this happens when $k_t \leq k_{\min}$, where $k_{\min}$ is defined by the condition $w(k_{\min}) = \tau$. That is, at this capital-labor ratio the given τ would imply that all wage income were confiscated. Thus the diagram in Fig. 5.2 reminds us of the fact that for a given technology in society there is a limit as to how high the mandatory contribution can be without generating sustained economic decline.

If we associate a low k with "the old days" (say half a century from now), then, at that time the actual mandatory contribution is likely to have been smaller than now. So we cannot read any *historical* evolution into the diagram (which also ignores technical progress). Rather, the diagram depicts the evolution that would result if a hypothetical economy with constant technology and constant τ has been close to its stable steady state for some time and then an upward shift

[7] Indeed, as the production function is neoclassical, $w(k) \to f(0) = 0$ for $k \to 0$, cf. equation (2.18) of Chapter 2.

in τ occurs.

The long-run effect To quantify the long-run effect one needs the long-run multipliers with respect to τ. Consider the fundamental difference equation in steady state:

$$(1+n)k^* = s(w(k^*), r(k^*), \tau). \tag{5.18}$$

Implicit differentiation on both sides gives $(1+n)dk^*/d\tau = s_w^* w'^* dk^*/d\tau + s_r^* r'^* dk^*/d\tau + s_\tau^*$. By ordering we find the "long-run multiplier"

$$\frac{dk^*}{d\tau} = \frac{s_\tau^*}{1+n-s_w^* w'^* - s_r^* r'^*} = \frac{s_\tau^*}{1+n-(s_r^* - s_w^* k^*) f''^*} < 0, \tag{5.19}$$

where we have used that $w' = -kf''$ (from (3.25) of Chapter 3) and where s_τ^* is the right-hand side of (5.11) evaluated in the stable steady state. The negative sign of $dk^*/d\tau$ is "demonstrated" by the graph in Fig. 5.2. But it is not immediately obvious that the formula in (5.19) necessarily gives a negative number. Why must it? The answer is that from (5.15), (5.16), and stability of the steady state follows that $0 < \frac{\partial k_{t+1}}{\partial k_t}(k^*) = \frac{-s_w^* k^* f''(k^*)}{1+n-s_r^* f''(k^*)} < 1$. This inequality implies $-s_w^* k^* f''(k^*) < 1+n-s_r^* f''(k^*)$ or $1+n-(s_r^* - s_w^* k^*) f''^* > 0$, which, since $s_\tau^* < 0$, makes the expression in (5.19) negative.

Taking Harrod-neutral technological progress into account

It is straightforward to extend the above analysis to include Harrod-neutral technological progress. Let $Y_t = F(K_t, T_t L_t)$, where the production function F is neoclassical with CRS and $T_t = T_0(1+g)^t$ is the technology level growing at a constant rate $g > 0$. To help existence of a steady state, we introduce the Homotheticity Assumption from Chapter 4, saying that lifetime utility is *homothetic.* In the formulas above we then just have to replace k by $\tilde{k} \equiv k/T$, $1+n$ by $(1+g)(1+n)$, and τ by $\tilde{\tau} \equiv \tau_t/T$. And in Fig. 5.2, k and τ should of course be replaced by $\tilde{k}$ and $\tilde{\tau}$, respectively. Then the above results go through. Under the general assumptions of the model, a tax-based pension system reduces capital accumulation.

5.1.4 Discussion

The conclusion from the above analyses is that a funded pension system tends to be neutral to aggregate saving, whereas the tax-based system tends to diminish aggregate saving compared to what it would be in the absence of the system.[8]

[8]Although absent from the analysis above, also labor supply may be affected by which pension system is operative.

It would be wrong, however to conclude from this, without circumspection, that the funded system is therefore to be preferred. We emphasize this point not so much because of the theoretical possibility of having $1 + r < (1 + g)(1 + n)$ permanently (overaccumulation) so that *less* aggregate saving would be Pareto-improving. In Chapter 4 we referred to empirical evidence that overaccumulation is not known to have been a problem in any country in practice, including countries without a mandatory pension system. What we want to underline is that even a society which does not suffer from over-saving, is not well guided by the above analysis *alone*. This is because the analysis has not compared the two pension systems on an equal footing. When the tax-based retirement pension systems were first introduced *historically*, an old generation already existed. That generation was immediately taken care of by introduction of the tax-based system, but would not have been so by a funded system. In the start-up period the tax-based system incurred a "hidden debt" to that period's young generation who financed the transfers to the old. In the next period society pays back to that generation, but only by incurring a new hidden debt to the new young generation and so on. In this way the tax-based system implies a permanent rolling over of the hidden debt.

Any government considering a shift to a funded mandatory system will face the problem that there is no Pareto-improving way to do that. If the government decides in period t_0 that the mandatory contribution by the young should be invested rather than being immediately transferred to the old, those who are old in period t_0 would be left at the post. They paid *their* contribution in the previous period, but do no receive the expected "return" now.

Could this problem be solved by a policy of issuing government bonds at the beginning of period t_0 and using the proceeds to finance the pensions of the currently old? No, under "normal circumstances" such a policy would not be sustainable because of compound interest on the resulting government debt. To catch a glimpse of the reason, suppose that our Diamond economy is well-behaved and has until period t_0 been in a steady state with $1 + r^* \geq (1 + g)(1+ n)$.[9] Part of the saving of the young would then be placed in government bonds (as long as they pay the same rate of return as other investments) and so less of the saving would be available for financing capital investment. As a result capital accumulation, and thereby $\tilde{k}$, would fall so that the rate of return, r, would rise *above* r^*. Since, by assumption, there are no taxes to finance the debt service, the government will have to roll over the debt forever. So the debt would grow at the rate r which is higher than the upper bound, $(1 + g)(1 + n) - 1$, for sustained growth in income. As a consequence, the saving by the young would sooner or

[9] If the opposite inequality were true, there would be aggregate overaccumulation, a situation where society would hardly want to promote saving by shifting to a funded system.

later not suffice to buy the newly issued bonds – government debt default would be inevitable. The default is in fact likely to occur long before the saving by the young is exhausted because investors will foresee that bankruptcy is underway.[10]

We conclude that a shift to a funded system of social security is not a simple matter. Similarly, it would be false if one argued for a funded system in the following way. It is an empirical fact that the rate of return on the stock market tends to be higher than the growth rate of the economy. Thus, a funded pension system could give a higher rate of return on deposits than the tax-based system and might for this reason be claimed superior. The problem is again, however: who are going to pay off the hidden debt to the currently old?

There are many intricate aspects involved in social security reform. Different pension systems differ in the degree of risk sharing and redistribution and with respect to administration costs. An additional issue is whether a society aiming at a funded system, would prefer this in the form of a social security system or a privatized system.

5.2 Endogenous labor supply

An important ongoing demographic change in the more developed countries is the "ageing" of the population, due to lower fertility and higher life expectancy. This implies an increasing dependency ratio. In turn, this tends to increase the tax burden which can have undesired effects on incentives and may increase moonlighting. In order to obviate this challenge, governments in many developed countries try to find arrangements to increase the labor force, both on the *intensive margin* (more hours supplied per year per member of the labor force) and the *extensive margin,* also known as the participation margin. The latter margin may refer to enrolment into the labor force, temporary leave (perhaps due to parenthood), or permanent leave (retirement).

To prepare for discussion of these issues, in the first subsection below we give a refresher of the basics of endogenous labor supply at the intensive margin. In the second subsection we apply the concepts in a simple extension of Diamond's OLG model. Subsequently, endogenous retirement will be considered.

5.2.1 The intensive margin: A simple one-period model

Consider an individual with preferences represented by a utility function $u(c, 1-\ell)$, where c is consumption and ℓ is labor supply of this individual, taking the market wage as given. We let ℓ be measured in a time unit such that total time

[10] A formal account of this kind of explosive paths which cannot be sustained in general equilibrium is given in Chapter 28 which addresses the issue of "rational bubbles".

available is 1 per period. Then $1-\ell$ is leisure (perhaps including homework). We assume that $u(\cdot)$ is strictly quasi-concave and that marginal utilities are positive, but decreasing in own argument, i.e., $u_i > 0, u_{ii} < 0$ for $i = 1, 2$.[11] We take the consumption good as numeraire, i.e., its price is 1.

The decision problem is:

$$\begin{aligned} & \max_{c,\ell} u(c, 1-\ell) \quad \text{s.t.} \\ c &= a + w\ell, \\ c &\geq 0, 0 \leq \ell \leq 1. \end{aligned}$$

Here $w > 0$ is the real wage per unit of work and a is the value an of *exogenous* financial asset (a may be positive or negative, but we assume $a > -w$). In addition to the budget constraint, we have stated the definitional constraints on the control variables c and ℓ.

We substitute for c in the utility function so that it can be written as a function of only one control variable, $\tilde{u}(\ell) \equiv u(a+w\ell, 1-\ell)$. Assuming an interior solution, we get the first-order condition

$$\tilde{u}'(\ell) = u_1(a + w\ell, 1-\ell)w - u_2(a + w\ell, 1-\ell) = 0,$$

which can conveniently be written

$$u_2(c, 1-\ell) = u_1(c, 1-\ell)w. \tag{5.20}$$

This condition says that in the optimal plan, the utility cost of reducing leisure by one unit equals the utility benefit of having w more consumption units at one's disposal due to higher labor income.

In view of strict quasi-concavity of $u(\cdot)$, the first-order condition (5.20) together with the budget constraint determines labor supply, ℓ, uniquely as an implicit function of w and a, $\ell = \ell(w, a)$. By the budget constraint we then immediately get the consumption function $c = a + w\ell(w, a) \equiv c(w, a)$. It is natural to assume that consumption as a whole is a normal good (such that $c_a > 0$) and, perhaps, that for most people also leisure is a normal good (such that $\ell_a < 0$).

As is well-known, one cannot in general tell in what direction an increase in w affects labor supply. Indeed, the total effect on labor supply is the net result of partial effects going in opposite directions. In the literature there are different ways of decomposing the total effect into partial effects. Here we will use the Slutsky-decomposition, well-known from textbooks in microeconomics. This

[11] *Strict quasi-concavity* of $u(\cdot)$ is equivalent to the indifference curves being strictly convex to the origin. Given $u_i > 0, u_{ii} < 0$, a sufficient condition for this is that $(u_2)^2 u_{11} - 2u_1 u_2 u_{12} + (u_1)^2 u_{22} < 0$, which will be satisfied at least whenever $u_{12} \geq 0$.

decomposition, with its associated terminology, provides a unifying framework for studying a range of issues such as the choice between different consumption goods, between consumption and saving, or between work and leisure.

We distinguish between three partial effects of a rise in the real wage: a substitution effect, an income effect, and a wealth effect. In the present context the meaning of these effects is most easily seen if we rewrite the budget constraint as:

$$c + wx = a + w \cdot 1 \equiv \hat{a}, \tag{5.21}$$

where x denotes leisure, $1 - \ell$, and $\hat{a}$ is total wealth, which is positive. In this way we consider leisure as just another consumption good, with price w (the opportunity cost of leisure). On the right-hand side of the budget constraint we now have something which is exogenous to the individual, namely the sum of a and *potential* labor income, defined as the obtainable labor income if all available time, here equal to one, is used for work. This sum, the "total budget" or "total wealth", is used partly for consumption, partly for leisure. Denoting the demand for leisure $x(w, \hat{a})$, we have

$$x(w, \hat{a}) \equiv 1 - \ell(w, \hat{a} - w). \tag{5.22}$$

Then the *extended* Slutsky equation concerning the demand for leisure takes the form

$$\frac{dx(w, \hat{a})}{dw} = \frac{\partial x(w, \hat{a})}{\partial w}|_{u=u_0} + (-\frac{\partial x(w, \hat{a})}{\partial \hat{a}}x) + \frac{\partial x(w, \hat{a})}{\partial \hat{a}}\frac{d\hat{a}}{dw}, \tag{5.23}$$

where the term on the left-hand side is called the *total effect*, the first term on the right-hand side is the *substitution effect*, the second the *pure income effect,* and the third the *wealth effect* of a rise in w (see Appendix A). In words:

1. *Substitution effect.* This effect indicates how the individual "substitutes" one good for another when a price changes and at the same time the budget is adjusted so that the original utility level is just affordable. In the present case, a rise in the wage rate makes leisure more expensive. As long as we imagine the individual remains on the same indifference curve, the point of tangency between the new budget line and the indifference curve must be one with less demand for leisure ("it's worth working more now") and more for the consumption good whose price has not increased. Hence, the substitution effect of an increase in the wage rate is negative on leisure and positive on consumption.[12]

[12] Other names for the substitution effect are the "demand effect under a Hicksian wealth compensation" or just the "Hicks-compensated effect" (the individual is "compensated" for

2. *Pure income effect.* This effect indicates the demand effect on a good due to a change in the purchasing power of a *given* budget when a price changes. A rise in the wage rate (the price of leisure) implies that purchases of at least one of the goods in the "consumption basket" must be curtailed if the budget remains unchanged ("facing higher prices, a given budget can buy less"). Therefore, the pure income effect of an increase in the wage rate is negative on all consumption goods, including leisure, hence positive on labor supply, when all the goods are *normal*.

3. *Wealth effect.* This effect indicates the demand effect on a good when the budget changes. A rise in the wage rate implies that wealth, and therefore the budget, is increased, so that the individual can afford to buy more of all goods at the new set of prices compared to what could be bought with an unchanged budget. Hence, the wealth effect of an increase in the wage rate is positive on both leisure and consumption when both are normal goods, hence negative on labor supply ("you don't need to work so much any more").

In the *simple* Slutsky equation from *partial equilibrium* analysis, wealth is taken as an exogenous constant, i.e., independent of prices. So only the first two effects on the right-hand side in (5.23) are considered, the substitution effect and the pure income effect.

In *general equilibrium* analysis, however, we consider the extended Slutsky equation, (5.23), featuring also the wealth effect. We get a more compact version by noting that $d\hat{a}/dw = 1$ in view of (5.21), so that the right-hand side of (5.23) can be compressed to yield

$$\frac{dx(w,\hat{a})}{dw} = \frac{\partial x(w,\hat{a})}{\partial w}\Big|_{u=u_0} + \frac{\partial x(w,\hat{a})}{\partial \hat{a}}(1-x), \tag{5.24}$$

where $1 - x \equiv \ell$ is net supply of the good in question, here working hours. The second term on the right-hand side of (5.24), that is, the sum of the pure income effect and the wealth effect, is sometimes named the *total income effect*.[13] In spite of the pure income and wealth effects often being of opposite sign (as in the present problem), for a *normal good* the total income effect is of the same

the price change by an adjustment of the budget so that he or she is just able to stay on the same indifference curve). The appendix compares this compensation to another notion of compensated demand, called "the Slutsky wealth compensation". Considering only infinitesimal changes in prices, the Slutsky-compensated effect turns out to be exactly equal to the Hicks-compensated effect. Hence, for simplicity we call all three effects appearing on the right-hand side of (5.23) *Slutsky effects.*

[13] Colloquially, the total income effect is known as the "hammock effect" which may be contrasted with the "carrot effect", that is, the substitution effect.

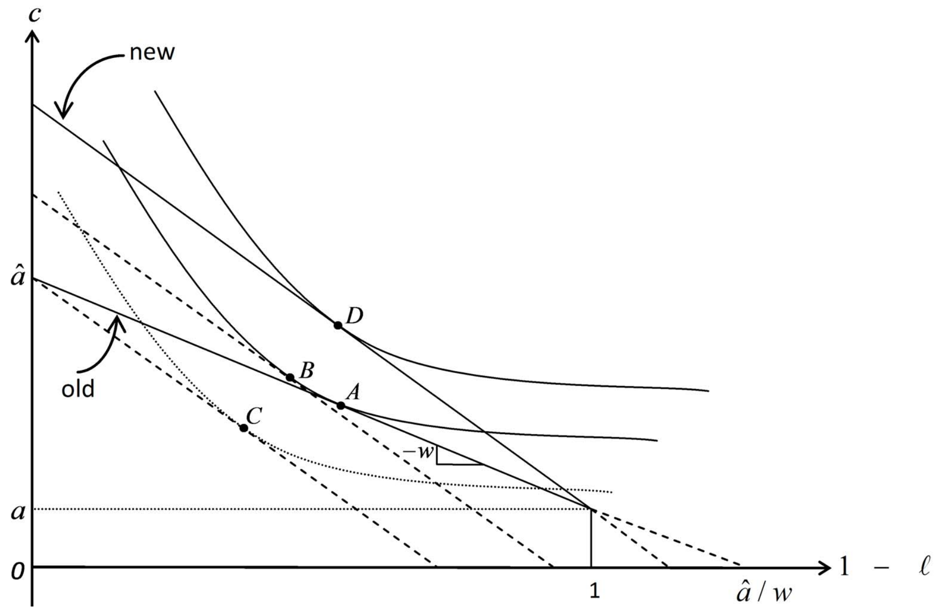

Figure 5.3: Substitution effect ($A \to B$), pure income effect ($B \to C$), and wealth effect ($C \to D$) on (x, c) of an increase in w.

sign as the net supply of the good. Given an interior solution to the labor supply problem, if leisure is a normal good, then the total income effect on leisure thus has the convenient property of being definitely of positive sign (i.e., dominated by the wealth effect). Moreover, the total income effect on leisure is proportional to the amount of labor supplied in the reference situation.

Sometimes in macroeconomics and labor market theory the total income effect is just called the *income effect.* For a student having just finished a basic course in *microeconomics* this labelling might lead to confusion with the *pure income effect,* i.e., the second term on the right-hand side of the Slutsky decomposition (5.23). Indeed, *that* term is in microeconomics often just called the *income effect.* Moreover, in many contexts, including the present one, the *pure* income effect is of sign *opposite* to that of the *total* income effect. To avoid confusion, we therefore add the prefix "pure" or "total" depending on which of the income effects is meant. Notwithstanding this terminological issue, there are many contexts, in particular when dynamics is considered, where it is expedient to use the full Slutsky decomposition given in (5.75).

The three Slutsky effects on consumption, leisure, and labor supply, respectively, are illustrated in Fig. 5.3. To understand the graph, rewrite the budget constraint as

$$c = a + w - wx \equiv \hat{a} - wx.$$

As the figure is drawn, financial wealth, a, is assumed positive. The budget line is represented by the line connecting the points $(1, a)$ and $(0, \hat{a})$. The budget constraint is kinked at the point $(1, a)$ because leisure cannot exceed total time available, which is 1. The figure shows three indifference curves. At the initial level of w, the bundle corresponding to the point A is optimal. Assuming both consumption and leisure are normal goods, the signs of the three partial effects on consumption, leisure, and labor supply of an increase in the real wage, w, are as indicated in Table 5.1 (where A, B, C, and D refer to points in Fig. 5.3).

Table 5.1. Sign of the Slutsky effects on c, x, and ℓ, respectively, of a rise in w, presupposing an interior solution and that the goods are normal.

		c	x	$\ell = 1 - x$
substitution effect:	$A \to B$	$+$	$-$	$+$
pure income effect:	$B \to C$	$-$	$-$	$+$
wealth effect:	$C \to D$	$+$	$+$	$-$
total effect:	$A \to D$	$+$	?	?

The sign of the total effect on leisure – and thereby also on labor supply – cannot be generally established. However, as indicated in Table 3.1, the sign of the total effect on consumption *can*. This difference may seem paradoxical. Indeed, given an interior solution, we have $x = 1 - \ell < 1$, and then from (5.21) follows that the increase in w raises the right-hand side of the budget constraint more than the left-hand side when evaluated at the old optimal bundle of the two "goods", consumption and leisure. The increase in w thus raises purchasing power, even taking the negative income effect into account. As both consumption and leisure are assumed to be normal goods, one might imagine that *both* were raised. Yet only consumption is unambiguously raised. The explanation is that whereas for consumption the substitution effect goes in the *same* direction as the *total* income effect (the consumption good has become relatively cheaper), for leisure the substitution effect goes in the *opposite* direction (leisure has become relatively more expensive). That is, for consumption both terms on the right-hand side of (5.24) are positive, but for leisure they are of opposite sign. Referring to Fig. 5.3, the point is that both goods being normal ensures that D is North-East of B. And the substitution effect ensures that B is North-West of A. Hence we can conclude only that D is North of A. We can not conclude whether D is North-East or North-West of A.

EXAMPLE 1 (the log utility case) Let $u(c, 1-\ell) = \ln c + \gamma \ln(1-\ell)$, where $\gamma > 0$ measures "taste" for leisure. Then (5.20) gives $\gamma/(1-\ell) = w/c$. Substituting this into (5.21) yields $c + \gamma c = w + a$ so that the solution for consumption is c

$= (1+\gamma)^{-1}(w+a)$ and the solution for labor supply is $\ell = 1-\gamma(1+\gamma)^{-1}(w+a)w^{-1} \equiv \ell(w,a)$. In this example, the *total effect* on labor supply of an increase in w is given by

$$\frac{\partial \ell}{\partial w} = \frac{\gamma}{1+\gamma}\frac{a}{w^2} \gtreqless 0 \text{ for } a \gtreqless 0,$$

respectively. Thus, in case of *positive* financial wealth the positive substitution effect on labor supply dominates the negative total income effect. In case of *negative* financial wealth, however, it is the negative total income effect that dominates (the labor supply curve will be negatively-sloped). The explanation is that when $a > 0$, total wealth, $\hat{a} \equiv a + w$, is "large" and thereby the demand for leisure is "large" when leisure is a normal good; so the total income effect is "small" (cf. (5.24)) and tends to be dominated by the positive substitution effect. It is opposite when $a < 0$. That the break-even point is at $a = 0$ is due to the logarithmic specification of the utility function. □

We may view labor supply as a function of w and $\hat{a}$ by defining $\tilde{\ell}(w,\hat{a}) \equiv \ell(w,\hat{a}-w) \equiv 1 - x(w,\hat{a})$, the latter identity being implied by (5.22). From (5.23) then follows:

$$\begin{aligned} \frac{d\tilde{\ell}(w,\hat{a})}{dw} &= \frac{\partial \tilde{\ell}(w,\hat{a})}{\partial w}|_{u=u_0} + \left(-\frac{\partial \tilde{\ell}(w,\hat{a})}{\partial \hat{a}}(1-\ell)\right) + \frac{\partial \tilde{\ell}(w,\hat{a})}{\partial \hat{a}}\frac{d\hat{a}}{dw} \\ &= \frac{\partial \tilde{\ell}(w,\hat{a})}{\partial w}|_{u=u_0} + \frac{\partial \tilde{\ell}(w,\hat{a})}{\partial \hat{a}}\ell, \end{aligned} \tag{5.25}$$

since $d\hat{a}/dw = 1$. Thus, also for labor supply do we have

total effect = substitution effect + pure income effect + wealth effect
= substitution effect + total income effect.

Typically, econometric studies attempt to estimate the

$$\textit{compensated labor supply elasticity} \equiv \frac{w}{\ell}\frac{\partial \tilde{\ell}(w,\hat{a})}{\partial w}|_{u=u_0},$$

and the

$$\textit{uncompensated labor supply elasticity} \equiv \frac{w}{\ell}\frac{d\tilde{\ell}(w,\hat{a})}{dw}.$$

The compensated labor supply elasticity, also known as the *Hicksian labor supply elasticity,* indicates the *substitution* effect of a rise in the real wage transformed into an elasticity. In contrast, the uncompensated labor supply elasticity, also known as the *Marshallian labor supply elasticity* is the *total* effect (including the effect via the increase in $\hat{a}$) transformed into an elasticity.

Measurement of these matters is complicated. There are many reasons for this. One is that in practice labor supply has also an intertemporal dimension and so wages, non-labor incomes, and interest rates in several periods have a bearing (cf. Section 5.4 below). Another complicating factor is the presence of involuntary unemployment, giving rise to a wedge between "true" labor supply and actual employment. Econometric estimates of the elasticities differ considerably. We get an indication of the order of magnitude from, e.g., Blundell and MaCurdy (1999) who survey empirical results about the intensive margin from the US and Western Europe. For men Blundell and MaCurdy report compensated hours elasticities in the interval $[0.01, 1.06]$ and uncompensated hours elasticities in the interval $[-0.25, 0.25]$ (though with an overweight of estimates in the positive range). For married women the estimates of the uncompensated elasticities are somewhat higher.[14] This corresponds well to the fact that the total income effect (which countervails the substitution effect) for women tends to be smaller than for men in view of ℓ being, on average, smaller for women, see (5.25).

Considering the external margin, fixed costs of entry and exit have a bearing. A survey of elasticities concerning the *extensive* margin reports estimates of magnitude... (, 1999).[15]

In the subsequent sections we extend the discussion to labor supply over a sequence of periods and also include a historical perspective on working hours.

5.2.2 Endogenous labor supply in an extended Diamond model

As in the standard Diamond OLG model we assume people live two periods, as young and as old. As young they chose to supply ℓ units of labor, $0 \leq \ell \leq 1$. As old they are unable to work. Uncertainty is ignored.

The problem of the young The decision or planning problem of the young in a given period is:

$$\max_{c_1,\ell,c_2} U(c_1, 1-\ell, c_2) = u(c_1, 1-\ell) + (1+\rho)^{-1} v(c_2) \quad \text{s.t.}$$

$$c_1 + s = w\ell, \tag{5.26}$$

$$c_2 = (1+r)s, \tag{5.27}$$

$$c_1 \geq 0, c_2 \geq 0,\ 0 \leq \ell \leq 1,$$

[14] Yet in recent years the distance to men's elasticities seem considerably narrowed down, both on the intensive and the extensive margin (Heim, 2007).

[15] Some authors and statistical agencies condition the exogenous margin of hours response on non-employment. Others condition on being outside the labor force (those available for employment).

where c_1 is consumption as young and c_2 is planned consumption as old, while s is saving, w is the real wage, and r is the real interest rate. The period utility function u is strictly quasi-concave and satisfies $u_i > 0$, $u_{ii} < 0$, $i = 1, 2$. The period utility function v satisfies $v' > 0$, $v'' < 0$. As in the standard Diamond model, initial financial wealth, a, is zero (no inheritance).

Substituting the constraints (5.26) and (5.27) into U, the problem is reduced to an unconstrained maximization problem with two choice variables, s and ℓ. Assuming an interior solution, the first-order conditions are $\partial U/\partial s = 0$ and $\partial U/\partial \ell = 0$, from which we get

$$u_1(c_1, 1-\ell) = (1+\rho)^{-1}v'(c_2)(1+r), \tag{5.28}$$

$$u_2(c_1, 1-\ell) = u_1(c_1, 1-\ell)w. \tag{5.29}$$

The condition (5.28) says that in the optimal plan, the opportunity cost (measured in current utility) of increasing saving by one unit equals the benefit (measured as discounted utility) of having $1 + r$ more units for consumption next period. The condition (5.29) says that in the optimal plan, the opportunity cost (measured in current utility) of reducing leisure by one unit equals the benefit (measured in current utility) of having w more units (the real wage) for consumption the same period.

Substituting the budget constraints into these two conditions, we can interpret them as defining saving and labor supply as implicit functions of w and r, namely $s(w, r)$ and $\ell(w, r)$.

The log utility case For concreteness we specify the period utility functions: $u(c_1, 1-\ell) = \ln c_1 + \gamma \ln(1-\ell)$ and $v(c_2) = \ln c_2$, where γ is a positive parameter (the relative weight given to utility from leisure). Condition (5.28) now gives

$$c_2 = \frac{1+r}{1+\rho}c_1. \tag{5.30}$$

Condition (5.29) gives

$$1-\ell = \frac{\gamma c_1}{w}. \tag{5.31}$$

It is useful to bring in the intertemporal budget constraint, IBC. Combining (5.26) and (5.27), we get $c_1 + (1+r)^{-1}c_2 = w\ell$. This can be written such that the opportunity cost of leisure (foregone earnings) appears on the left-hand side, parallel to the consumption components c_1 and c_2 :

$$c_1 + w(1-\ell) + \frac{c_2}{1+r} = w. \tag{IBC}$$

Note that the right-hand side of the budget constraint is now exogenous to the individual. The IBC says that the present discounted value, as seen from the end of the first period, of the consumption and leisure plan equals the total wealth, which in this model is just the *potential* wage income in the first period (no financial wealth is inherited from the parents and no labor income is earned in the second period). The IBC is useful for interpretation purposes and it also provides an expedient simple relation between c_1, c_2, and $1-\ell$. We shall use it to derive the consumption demand and labor supply functions. Substituting (5.30) and (5.31) into (IBC) gives $c_1 + \gamma c_1 + c_1(1+\rho)^{-1} = w$, that is,

$$c_1 = \frac{1+\rho}{2+\rho+(1+\rho)\gamma}w, \text{ hence,} \tag{5.32}$$

$$c_2 = \frac{1+r}{2+\rho+(1+\rho)\gamma}w, \tag{5.33}$$

from (5.30). Substituting (5.32) into (5.31) gives

$$\ell = \frac{2+\rho}{2+\rho+(1+\rho)\gamma}. \tag{5.34}$$

From (5.26) we have $s = w\ell - c_1$. In view of (5.34) and (5.32) this yields

$$s = \frac{1}{2+\rho+(1+\rho)\gamma}w. \tag{5.35}$$

Comments on the solution We see that with this specification neither the wage rate nor the interest rate plays any role in the determination of labor supply. Given the preference parameters, labor supply, ℓ, is a constant. Further, the saving, s, is independent of the interest rate. In this sense we are considering a benchmark case. The special character of this case is due to the combined effect of the time structure in the model (no inherited financial wealth and no labor income as old) and the log utility specification. Given this setup, the negative substitution and pure income effects on leisure of an increase in the wage rate are exactly offset by the positive wealth effect. This is roughly consistent with the econometric evidence referred to above, which tells us that labor supply, at least that of men, is rather inelastic.

As to the real interest rate, when it rises, consumption in period 2 becomes less expensive as seen from period 1, cf. (IBC). This induces a substitution effect away from consumption and leisure in period 1. Hence, the substitution effect on leisure as young of a rise in the interest rate in period 1 is negative. But (IBC) also shows that the higher interest rate makes any given consumption bundle $(c_1, 1-\ell, c_2)$ cheaper as seen from period 1. Therefore the pure income effect of a rise in the interest rate on leisure as young is positive (when leisure is a normal good). Finally, since wealth (in this model) is not affected by a rise in the interest rate (the right-hand side of (IBC) is unaffected), the wealth effect is absent. Because of the log specification of utility, the substitution and pure income effects exactly cancel each other. Thus, in the end the total effect of a rise in the interest rate on leisure as young, hence also on labor supply as young, is nil.

Similarly, due to the log specification of utility, the substitution and pure income effects on consumption as young of a rise in the interest rate offset each other, while there is no wealth effect. Hence, saving as young is independent of r, cf. (5.35).

The absence of a wealth effect of a rise in the interest rate is an artificial feature of the Diamond model and derives from the fact that labor is only supplied in the first period. If labor were also supplied when old, there would be a negative wealth effect on both leisure and consumption as young from an increase in r (the present discounted value of future labor income goes down, when r rises).

Dynamics at the aggregate level We must now distinguish between aggregate employment (here equal to aggregate labor supply), which we shall denote L, and the number of young, which we shall denote N. We have $L = \ell N$, where the individual labor supply, ℓ, is given by (5.34). It is assumed that $N_t = N_0(1+n)^t$, where $n > -1$ is the constant population growth rate.

Let the aggregate production function be $Y = F(K, L) = LF(k, 1) \equiv Lf(k)$, where $k \equiv K/L$ (the capital-labor ratio) and $f' > 0$, $f'' < 0$. For simplicity technological change is ignored. Then, in competitive equilibrium the real wage per unit of labor is

$$w = \frac{\partial Y}{\partial L} = \frac{\partial (Lf(k))}{\partial L} = f(k) - f'(k)k \equiv w(k).$$

Thus, in the log utility case aggregate capital accumulation is described by

$$K_{t+1} = s_t N_t = \frac{w(k_t)}{2+\rho+(1+\rho)\gamma} N_t,$$

where we have used (5.35). Dividing on both sides by $L_{t+1} = \ell N_{t+1} = \ell N_t(1+n)$ gives

$$k_{t+1} = \frac{w(k_t)}{[2+\rho+(1+\rho)\gamma]\,\ell(1+n)} = \frac{w(k_t)}{(2+\rho)(1+n)}, \tag{5.36}$$

in view of (5.34).

The capital-labor ratio can be written

$$k_t \equiv \frac{K_t}{L_t} = \frac{K_t}{\ell N_t} \equiv \frac{\bar{k}_t}{\ell}, \tag{5.37}$$

where $\bar{k}_t \equiv K_t/N_t$. Note that, contrary to k_t, capital per young, $\bar{k}_t$, is a predetermined variable (given by previous capital accumulation). It is therefore generally more useful to have the dynamics expressed in terms of $\bar{k}$ rather than k. We obtain this by multiplying by ℓ on both sides of (5.36) to get

$$\bar{k}_{t+1} = \frac{\ell w(\bar{k}_t/\ell)}{(2+\rho)(1+n)}, \tag{5.38}$$

where ℓ is given in (5.34). Thus, we end up with capital per young next period determined, through a simple transition function, by the predetermined level of capital per young in the current period. Since $w'(k) = -kf''(k) > 0$, the transition curve is positively sloped for all $\bar{k}_t > 0$. Fig. 5.4 illustrates.

Suppose there is a unique non-trivial stable steady state $\bar{k}^*$ (as when $F(\cdot)$ is Cobb-Douglas, for instance). An increase in impatience (ρ) or in the population growth rate (n) has a similar effect as in the simple Diamond model of Chapter 3. The new thing is that the extended model allows us to consider questions like: how does an upward shift in the relative preference for leisure, γ, affect the long-run level of capital per young, $\bar{k}^*$, and thereby the long-run interest rate? First, the effect on individual labor supply is negative as seen by (5.34). So, aggregate labor supply is permanently reduced. Most likely this will immediately result in less income to the young, hence less saving and therefore over time less capital per young and thus a higher long-run interest rate. Note that, apriori, the proviso "most likely" is necessary. This is because the real wage rises when labor becomes more scarce, due to the lower ℓ caused by the higher preference for leisure. At the theoretical level this might completely offset (or even more than offset) the lower ℓ. It can be shown, however, that this will only happen if the elasticity of substitution between K and L is lower than the gross capital income share (see Exercise 4.? in Chapter 4). The latter is generally estimated to be close to one third and the elasticity of substitution to be in the interval (0.5,1.0).[16] In this way our "most likely" statement is corroborated.

This finishes our presentation of a simple way to endogenize labor supply in the Diamond model. Some authors[17] even use this specification as a model of endogenous *retirement*. An alternative way of modelling endogenous retirement is considered in the next section.

5.3 Early retirement with transfer income

In this section we address macroeconomic effects of a voluntary early retirement scheme with transfer income (in Danish, "efterløn"). This relates to the political

[16]See, e.g., Antràs (2004).

[17]For example Heijdra and van der Ploeg (2002).

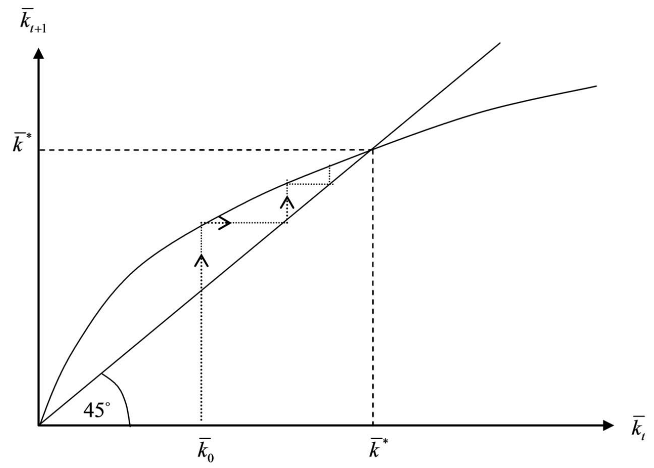

Figure 5.4: Movement over time of capital per young.

debate in several countries about whether the legislation concerning retirement is formed in a sound way from the perspective of fiscal sustainability. For simplicity, we base our analysis on a specific utility function (but the main points do not hinge on this special case).

Partial equilibrium analysis of an early retirement scheme

As in the Diamond model we consider an economy where people live for two periods, as young and as old. As young they supply inelastically one unit of labor. Their planned labor supply as old is ℓ, where ℓ depends on market circumstances. So ℓ is planned *senior working time* and we may interpret $1 - \ell$ as a measure of how early the individual retires from the labor market. The agents are price takers and there is no uncertainty. The planning problem of the young is:

$$\begin{aligned}
\max_{c_1,c_2,\ell} U &= \ln c_1 + (1+\rho)^{-1}\left[\ln c_2 + \gamma \ln(1-\ell)\right] \qquad \text{s.t.} \\
c_1 + s &= \hat{w}_1, && (5.39) \\
c_2 &= (1+r)s + \hat{w}_2\ell + \hat{m}(1-\ell). && (5.40) \\
c_1 &\geq 0, \; c_2 \geq 0, \, 0 \leq \ell < 1.
\end{aligned}$$

Here, $\hat{w}_1$ is the after-tax (real) wage in the first period, $\hat{w}_2$ is the expected after-tax (real) wage in the second period, and r is the expected real interest rate. For simplicity we ignore taxation of interest income. The parameter $\gamma > 0$ represents the individual's subjective relative weight on leisure as old. After retirement the individual receives from the government a net-of-tax transfer at rate $\hat{m}$ per time unit, where $0 \leq \hat{m} < \hat{w}_2$. We call $\hat{m}$ the *retirement pension rate*. Alongside the variables $\hat{w}_1$, $\hat{w}_2$, and r, also $\hat{m}$ is exogenous to the individual. But the total net-of-tax transfer, $\hat{m}(1-\ell)$, is of course endogenous. When the form of the tax scheme matters, we shall, for simplicity, assume proportional taxation with a time-independent tax rate, τ, such that $\hat{w}_1 = (1-\tau)w_1$, $\hat{w}_2 = (1-\tau)w_2$, and $\hat{m} = (1-\tau)m$, where w_1, w_2, and m are the corresponding pre-tax variables.[18]

By inserting (5.39) and (5.40) into U and considering s and ℓ as control variables, we get the first-order conditions:

$$\begin{aligned}
\frac{\partial U}{\partial s} &= -\frac{1}{c_1} + (1+\rho)^{-1}\frac{1}{c_2}(1+r) = 0 \\
&\Rightarrow \frac{c_2}{c_1} = \frac{1+r}{1+\rho}, \\
\frac{\partial U}{\partial \ell} &= (1+\rho)^{-1}\left[\frac{1}{c_2}(\hat{w}_2 - \hat{m}) - \gamma\frac{1}{1-\ell}\right] = 0 \\
&\Rightarrow \frac{\gamma c_2}{1-\ell}\left(= MRS_{c_2,1-\ell} = -\frac{dc_2}{d(1-\ell)}\,|_{U=\bar{U}}\right) = \frac{\hat{w}_2 - \hat{m}}{1}(= \text{price ratio}).
\end{aligned}$$

Observe that the marginal private opportunity cost of leisure is $\hat{w}_2 - \hat{m}$ and is smaller than the marginal *social* opportunity cost of leisure which is $w_2 = \hat{w}_2/(1-\tau) = \partial Y/\partial L$. The retirement pension rate, $\hat{m}$, implies a wedge between the two − in addition to the wedge generated by the tax rate τ.

Consider the individual's intertemporal budget constraint. Isolating s in (5.40), inserting into (5.39), and ordering we get

$$c_1 + \frac{c_2}{1+r} + \frac{\hat{w}_2 - \hat{m}}{1+r}(1-\ell) = \hat{w}_1 + \frac{\hat{w}_2}{1+r}. \tag{IBC}$$

The right-hand side is the present value of potential lifetime labor income (as seen from the end of the first period in life) and can be considered as the initial total wealth of the individual, who is born with no financial wealth. From the third term on the left-hand side we see that the opportunity cost of leisure (i.e., foregone earnings) is smaller the larger is the after-tax retirement pension rate, $\hat{m}$. Inserting the first-order conditions into (IBC) gives the solution

$$c_1 = \frac{1+\rho}{2+\rho+\gamma}\left(\hat{w}_1 + \frac{\hat{w}_2}{1+r}\right), \tag{5.41}$$

$$c_2 = \frac{1+r}{1+\rho}c_1 = \frac{1+r}{2+\rho+\gamma}\left(\hat{w}_1 + \frac{\hat{w}_2}{1+r}\right), \tag{5.42}$$

$$\ell = 1 - \frac{\gamma(1+r)}{(2+\rho+\gamma)(\hat{w}_2 - \hat{m})}\left(\hat{w}_1 + \frac{\hat{w}_2}{1+r}\right). \tag{5.43}$$

[18] A similar set-up was introduced by Hu (1979). Admittedly, the absence of uncertainty over such a long time horizon is not an appealing assumption. See Exercise 5.?.

How do c_1, c_2, and ℓ depend on the retirement pension rate $\hat{m}$? First, we see that $\partial c_1/\partial\hat{m} = 0$. This result is due to the log-utility specification, implying that the negative substitution effect and the positive pure income effect of a rise in $\hat{m}$ exactly offset each other, whereas there is no wealth effect, as seen from the right-hand side of (IBC). Second, $\partial c_2/\partial\hat{m} = 0$, that is, again the negative substitution effect and the positive pure income effect offset each other, while there is no wealth effect. Third, $\partial\ell/\partial\hat{m} = -\partial(1-\ell)/\partial\hat{m} < 0$, that is, *$\hat{m}$ affects planned senior working time* negatively. This unequivocal result comes about, because the substitution effect and the pure income effect on planned senior leisure time of a rise in $\hat{m}$ are both positive, while there is *no offsetting wealth effect.* In everyday language: if people like leisure and they are *paid* for taking leisure, it is no surprise they take more leisure. Note that a change in *income taxation* acts differently. Although such a tax change also entails a substitution effect and a pure income effect on leisure in the same direction, it brings about an offsetting wealth effect as well. Indeed, with the proportionate tax scheme described above, the tax rate τ cancels out in (5.43).[19]

As to the role of a rise in the interest rate, we find $\partial c_1/\partial r < 0$. So saving of the young, $s = \hat{w}_1 - c_1$, is affected positively by a higher interest rate. One the one hand, the negative substitution effect and the positive pure income effect on c_1 of a higher r exactly offset each other due to the log specification. On the other hand, on the right-hand side of (IBC) appears a negative wealth effect, which thus becomes decisive. Contrary to this, in the standard Diamond model labor is supplied only as young and therefore, unrealistically, a change in r has no wealth effect.

Besides, we see that $\partial c_2/\partial r > 0$. Here, the substitution and the pure income effects on planned consumption as old are both positive and dominate the negative wealth effect. Finally, $\partial\ell/\partial r = -\partial(1-\ell)/\partial r < 0$, reflecting that the substitution and pure income effects of a higher interest rate on planned senior leisure time are both positive and dominate the negative wealth effect.

What is the role of wages? Answering that, we take into account that in the real world the compensation m is likely to be related to the general wage level. Hence, let us assume

$$m = \mu\, w_2, \qquad 0 \leq \mu < 1, \tag{5.44}$$

where μ is an exogenous constant (the "degree of compensation"). Then, $\hat{m}$

[19] Thus, a government aiming at increasing labor supply through a tax reduction will have to design the change in the taxation scheme such that the wealth effect becomes small enough to not significantly offset the substitution and pure income effects. Many economists propose a combination of reduced labor income taxation and increased taxation of land and residental property.

$= (1-\tau)m = (1-\tau)\mu w_2 = \mu\hat{w}_2$ and (IBC) simplifies to

$$c_1 + \frac{c_2}{1+r} + \frac{(1-\mu)\hat{w}_2}{1+r}(1-\ell) = \hat{w}_1 + \frac{\hat{w}_2}{1+r}. \qquad \text{(IBC')}$$

The planned senior working time now becomes

$$\ell = 1 - \frac{\gamma}{2+\rho+\gamma}\frac{\frac{1+r}{\hat{w}_2/\hat{w}_1}+1}{1-\mu}, \qquad (5.45)$$

from (5.43). We see that, given our logarithmic utility function, planned senior working time is independent of a multiplicative wage change in both periods (by this is meant that $\hat{w}_1$ and $\hat{w}_2$ are multiplied by the same positive factor). Whereas the absolute wage level does not matter, *the growth rate in real wages,* $\hat{w}_2/\hat{w}_1 - 1$, *matters.*[20] An isolated increase in $\hat{w}_2$ generates negative substitution and pure income effects on senior leisure, which dominate the positive wealth effect. But if, in addition, $\hat{w}_1$ goes up proportionately, then this adds to the positive wealth effect just enough to offset the substitution and pure income effects. This corresponds to the empirical fact that annual working hours per worker, as well as the retirement age, have in most industrialized countries not risen over a century in spite of after-tax real wages being, in developed countries, about 4-5 times as large now as 100 years ago. The tendency goes in the opposite direction: less working hours (in the market) per year along with rising after-tax wages over time (Gali 2005, Huberman and Minns, 2007). This suggests that in the longer run the *wealth effect dominates.* For Denmark 1960-2002, the left-hand panel of Fig. 5.5 shows the evolution of the annual working hours per person in the labor force. With a view on the extensive margin the right-hand panel shows the evolution of the expected number of future years in the labor market that a "typical" sixteen-years-old person has.

General equilibrium analysis of an SOE with an early retirement scheme

We embed the above analysis in a Diamond-style OLG model of a small open economy (SOE for short) with government sector. The purpose is to study how a voluntary early retirement pension is likely to affect the economy as a whole.

Let the production function of the representative firm be

$$Y_t = F(K_t, T_t L_t),$$

where F is neoclassical with CRS and $Y_t, K_t,$ and L_t are output and input of capital and labor, respectively, while T_t is the technology level. Let T_t grow at a

[20] If also labor supply as young is endogenous, then a not unlikely additional effect of higher $\hat{w}_2/\hat{w}_1$ is a *lower* labor supply as young, cf. Section 5.4 below.

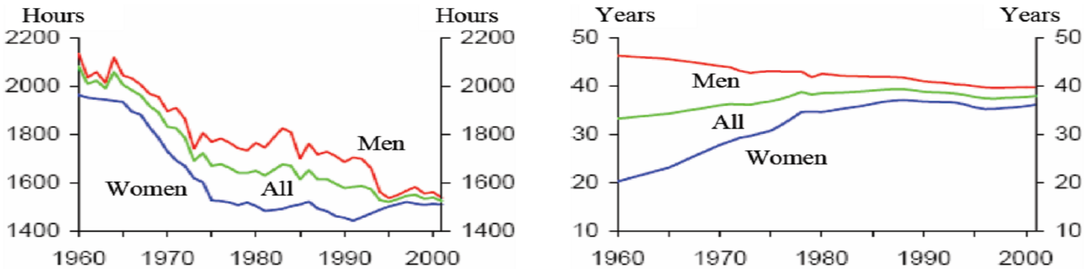

Figure 5.5: Left panel: evolution of annual working hours per person in the labor foce in Denmark over the period 1960-2002. Right panel: evolution of expected number of future years in the labor market for a sixteen-years-old in Denmark over the period 1960-2002. Source: Danish Welfare Commission (2004).

constant exogenous rate $g \geq 0$ so that

$$T_t = T_0(1+g)^t, \quad t = 0, 1, 2, ..., \tag{5.46}$$

We may write

$$\tilde{y}_t \equiv \frac{Y_t}{T_t L_t} = F(\frac{K_t}{T_t L_t}, 1) = F(\tilde{k}_t, 1) \equiv f(\tilde{k}_t), \qquad f' > 0, f'' < 0.$$

where $\tilde{k}_t \equiv K_t/(T_t L_t)$. We assume perfect competition in all markets. Then markets clear and L_t can be interpreted as labor supply as well as employment.

For simplicity we will assume:

(a) Perfect mobility of goods and financial capital across borders.

(b) No uncertainty; domestic and foreign financial claims are perfect substitutes.

(c) No need for means of payment, hence no foreign exchange market.

(d) No labor mobility across borders.

The assumptions (a) and (b) imply *real interest rate equality.* That is, in equilibrium the real interest rate in our SOE equals the real interest rate in the world financial market. We imagine that all countries just produce one and the same homogeneous output good. International trade will then only be *intertemporal* trade, i.e., international borrowing and lending of this good. Then, with negligible transport costs and no legal barriers to the international mobility of this good, real interest rate equality must hold. If it did not, there would be an arbitrage opportunity. One could borrow in one country at the lower interest rate and invest in another country at the higher interest rate. The resulting excess demand for loans in the first country and excess supply of loans in the second would quickly eliminate the interest rate difference.[21]

Let r denote the world market interest rate which is then exogenous to our SOE. The SOE is small enough not to affect r. Suppose r is positive and constant over time. Then, in the absence of corporate taxation, profit maximization leads to $f'(\tilde{k}_t) = r + \delta$, where δ is a constant capital depreciation rate ($0 \leq \delta \leq 1$). At least when f satisfies the Inada conditions, there is always a solution in $\tilde{k}$ to this equation and it is unique (since $f'' < 0$) and constant over time (as long as r and δ are constant). Thus,

$$\tilde{k}_t = f'^{-1}(r + \delta) \equiv \tilde{k}, \text{ for all } t. \tag{5.47}$$

In view of firms' profit maximization, the real wage before tax is

$$w_t = \frac{\partial Y_t}{\partial L_t} = \left[f(\tilde{k}) - f'(\tilde{k})\tilde{k} \right] T_t \equiv \tilde{w}(\tilde{k}) T_t. \tag{5.48}$$

Hence, the real wage grows over time at the same rate as technology, the rate g.

Labor income and transfers are taxed by the same constant tax rate τ; there is no taxation of interest income. With the assumed proportional taxation $\hat{w}_t = (1 - \tau)w_t$ and $\hat{w}_{t+1} = (1 - \tau)w_{t+1}$, so that the after-tax real wage also grows at the rate g. Then, from (5.45), the senior labor supply planned by a member of generation $t - 1$ is

$$\ell_t = 1 - \frac{\gamma}{2 + \rho + \gamma} \frac{1 + \frac{1+r}{1+g}}{1 - \mu} \equiv \ell. \tag{5.49}$$

As there is no uncertainty, planned actions are also realized ex post.[22] Note that for (5.49) to be consistent with positive labor supply as old, we need the parameter restriction

$$\gamma < \frac{(1 - \mu)(2 + \rho)}{\mu + \frac{1+r}{1+g}}, \tag{5.50}$$

i.e., the relative weight on leisure is not too high.

[21] In practice, real interest rate equality does not hold in the short run. As a description of an average tendency over longer time horizons, however, the hypothesis performs better (see, e.g., Homer and Sylla, 1991).

[22] If a shock changes the circumstances in period t compared to what was expected, then the old individual re-optimizes at the beginning of period t, facing the *one-period* problem $\max\left[\ln c_{2t} + \gamma \ln(1 - \ell_t)\right]$ s.t. $c_{2t} = (1+r)\bar{s}_{t-1} + \hat{w}_{2t}\ell_t + \hat{m}_t(1-\ell_t)$, where $\bar{s}_{t-1}$ denotes the now predetermined saving undertaken in the previous period by this person. Without the shock, the solution (c_{2t}, ℓ_t) to this one-period-problem coincides with the actions *planned* in period $t - 1$ by the young, cf. (5.42) and (5.43). See Exercise 5.X.

Similarly, in line with (5.41) and (5.42), respectively, we find

$$c_{1t} = \frac{1+\rho}{2+\rho+\gamma}(1+\frac{1+g}{1+r})(1-\tau)w_t, \tag{5.51}$$

$$c_{2t} = \frac{1+r}{1+\rho}c_{1t-1} = \frac{1+r}{2+\rho+\gamma}(1+\frac{1+g}{1+r})(1-\tau)w_{t-1}, \tag{5.52}$$

where $w_{t-1} = w_t/(1+g)$.

Aggregate employment equals aggregate labor supply which is $L_t = N_t + \ell_t N_{t-1}$, where N_t denotes the number of young people in period t. Ignoring other kinds of government expenses and taxes, we have

$$\text{retirement pay} = \mu w_t(1-\ell)N_{t-1}. \tag{5.53}$$

$$\text{tax revenue} = \tau\left[w_t L_t + \mu w_t(1-\ell)N_{t-1}\right], \tag{5.54}$$

Assuming a balanced government budget, we want to find the tax rate required to finance the early retirement pension. Given the constant population growth rate $n > -1$, we have $N_t = N_0(1+n)^t$. Hence, we can write aggregate employment as

$$L_t = N_t + \ell N_{t-1} = (1+n+\ell)\frac{N_t}{1+n}. \tag{5.55}$$

Equating (5.54) and (5.53), using (5.55), we find the required tax rate to be

$$\begin{aligned}\tau &= \frac{\mu(1-\ell)\frac{N_t}{1+n}}{L_t + \mu(1-\ell)\frac{N_t}{1+n}} = \frac{\mu(1-\ell)}{1+n+\ell+\mu(1-\ell)} \\ &= \frac{\mu\gamma\left(1+\frac{1+r}{1+g}\right)}{(1-\mu)\left[(2+n)(2+\rho)-\gamma\left(\frac{1+r}{1+g}-(1+n)\right)\right]},\end{aligned} \tag{5.56}$$

where the last equality is due to (5.49).[23]

The derived tax rate τ depends positively on both μ and $(1+r)/(1+g)$. In particular the role of the compensation rate μ is of interest. A higher μ implies a higher tax rate through two channels.[24] First there is a direct effect through the higher transfer to each retired worker. Second there is the indirect effect through the induced increase in the number of retired workers and the decrease

[23] In view of the parameter restriction (5.50), the formula (5.56) ensures $0 < \tau < 1$ for $0 < \mu < 1$. Indeed, the positivity of τ follows from positivity of the term in the square brackets of (5.56), and that $\tau < 1$ follows from positivity of the difference between denominator and numerator in (5.56).

[24] We are here performing *comparative* dynamic analysis. Studying effects of a change in μ in historical time is more complicated, see Exercise 5.?.

in employment and income (in view of $\partial\ell/\partial\mu = -\partial(1-\ell)/\partial\mu < 0$ from (5.49)). The interpretation of this indirect effect is facilitated if we rewrite (IBC') in the following way:

$$c_{1t} + \frac{c_{2t+1}}{1+r} + \frac{(1-\tau)(1-\mu)w_{t+1}}{1+r}(1-\ell) = (1-\tau)(1+\frac{1+g}{1+r})w_t. \qquad \text{(IBC")}$$

The result $\partial(1-\ell)/\partial\mu > 0$ comes about because both the substitution effect and the pure income effect on planned senior leisure are positive and their combined effect is *only partly offset by the negative wealth effect through the increase in* τ brought about. To recapitulate: a) *in itself* an increase in the early retirement pension rate affects planned senior leisure positively, because it reduces the opportunity cost of retiring *without* an opposing negative wealth effect: for a given τ the right-hand side of (IBC") is unaffected by μ; b) however, in general equilibrium we should take into account that higher μ implies higher τ and *thereby* a negative wealth effect on leisure; c) yet the higher τ decreases the opportunity cost of leisure *further* and this tends to offset the negative wealth effect. The net result is more leisure, that is, earlier retirement.[25]

The formula (5.56) shows that the required tax rate is lower, the larger are n and g. This is because higher n means a lower dependency ratio (#retired/#workers). And a higher g induces later retirement (to take advantage of the relatively high wage as old). On the other hand, the required tax rate is higher, the higher is the interest rate. This is because a higher r makes earlier retirement cheaper as seen from the young's perspective.

Although the main purpose of this stylized exercise is only to give a qualitative picture, let us on the basis of (5.56) make a rough numerical calculation of the tax rate τ required for financing the voluntary early retirement scheme. A "guesstimate" for the basic parameters (with West European countries in mind and one year as time unit) is: $\bar{n} = 0.005$, $\bar{g} = 0.02$, $\bar{r} = 0.05$, $\bar{\rho} = 0.03$, $\gamma = 0.2$, $\mu = 0.2$.[26] Transforming the first four parameter values to the period length of

[25] Considering the repercussions on c_1 and c_2, from (5.51) and (5.52) we find $\partial c_{1t}/\partial\mu = (\partial c_{1t}/\partial\tau)(\partial\tau/\partial\mu) < 0$ and $\partial c_{2t}/\partial\mu = (\partial c_{2t}/\partial\mu)(\partial\tau/\partial\mu) < 0$, respectively. The lower opportunity cost of early retirement due to an increase in μ (both directly and through the required higher taxation) induces a negative substitution effect on consumption in both periods and a positive pure income effect. But the higher taxation also results in a negative wealth effect, cf. (IBC"), and the total effect on consumption becomes negative.

[26] The (average) degree of compensation in the Danish system is around 0.5 (The Danish Welfare Commission, 2004). But since the voluntary early retirement scheme covers only a minor fraction of the "second period of life", we have adjusted the value for μ. Admittedly, the guesses on $\bar{\rho}$ and γ are shots in the dark.

the model, say, 30 years, we get:

$$\begin{aligned}1+n &= (1+\bar{n})^{30} = 1.005^{30} = 1.1614,\\ 1+g &= (1+\bar{g})^{30} = 1.02^{30} = 1.8114,\\ 1+r &= (1+\bar{r})^{30} = 1.05^{30} = 4.3219,\\ 1+\rho &= (1+\bar{\rho})^{30} = 1.03^{30} = 2.4273.\end{aligned}$$

Substituting into (5.56) yields a required tax rate, τ, slightly above 0.02.

It remains to see how capital and wealth accumulation are affected by the early retirement scheme. Notice that, given the degree of compensation μ, aggregate labor supply L_t is determined by (5.55) through (5.49). Hence, the endogenous stock of physical capital is in every period given by

$$K_t = \tilde{k}T_tL_t. \tag{5.57}$$

Consequently, $\partial K_t/\partial\mu = \tilde{k}T_t\partial L_t/\partial\mu < 0$. Since higher μ leads to earlier retirement and therefore lower aggregate labor supply, we end up with a lower capital stock being needed to equip the labor force.

In an open economy national wealth generally differs from the capital stock. Ignoring land (as usual in simple macro models), national wealth is $A_t \equiv K_t - D_t$, where D_t is net foreign debt. Ignoring public debt, national wealth is the same as private financial wealth. In our Diamond-style model we have

$$A_{t+1} = s_tN_t, \tag{5.58}$$

in view of (5.51). Hence, $\partial A_{t+1}/\partial\mu = (\partial A_{t+1}/\partial\tau)(\partial\tau/\partial\mu) = N_t(\partial s_t/\partial\tau)(\partial\tau/\partial\mu)$, which is negative since $\partial s_t/\partial\tau < 0 < \partial\tau/\partial\mu$ as long as $s_t > 0$.[27] This negative sign of $\partial A_{t+1}/\partial\mu$ is explained by the reduced after-tax income and saving by the young, due to the higher taxation required by increased μ. Though c_{1t} goes down, after-tax income goes more down so that also saving of the young, $s_t = (1-\tau)w_t - c_{1t}$, goes down.

In order to encourage later retirement and more saving, the government might consider reducing the degree of compensation or introducing a tax discount for late retirement. If the government, for some reason, say intra-generational distributional considerations (outside the model), does not want to do that, other policies might be considered. One could think of a senior policy improving education and work conditions for elderly people, thereby decreasing γ; better transport and health conditions could have the same effect. Given τ, labor supply and saving are likely to be stimulated by such means. Besides, by (5.56), τ would

[27] This requires $(1+r)/(1+g) > (1+\rho)/(1+\gamma)$, which holds when r and γ are large and g and ρ small.

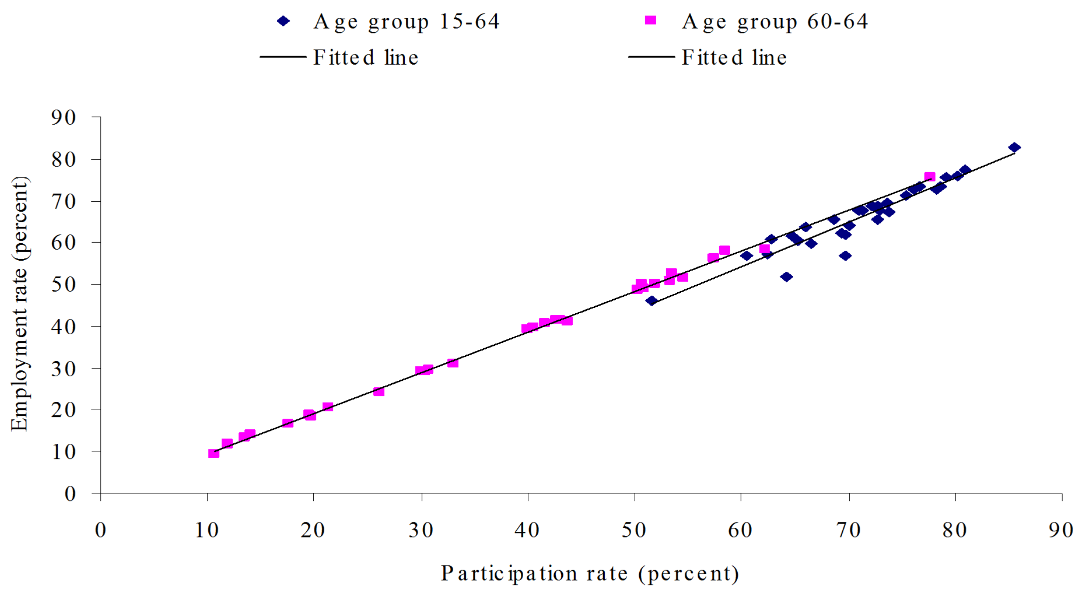

Figure 5.6: Employment rate and participation rate across OECD countries in 2004. Note: Employment and participation (labor supply) are measured as percentage of population in age 15-64 and 60-64, respectively. Source: OECD Labour Market Statistics (2004).

decrease (see Exercise 5.?), thus stimulating saving further. To target an early retirement scheme more precisely to the problem of employee attrition, the right to early retirement can be conditional on having participated in the labor market a sufficient number of years..

Owing to the assumptions of full employment and homogenous agents within generations, this stylized model does not capture all aspects of a voluntary early retirement scheme. In any case, as Fig. 5.6 shows, countries with later retirement – and thereby higher participation rate in the age group 60-64 years – do have correspondingly higher employment rate within this age group.

On the causal structure of an SOE model with perfect competition Notice that an SOE tends to be more simple to analyze than a closed economy. At least when we ignore monetary matters, less mutual dependency is present in an SOE. Indeed, a *recursive* causal structure is displayed. Thus, in the above SOE model, first, the exogenous r determines $\tilde{k}$ independently of anything else except the production function. Then $\tilde{k}$ together with the exogenous level of technology, T_t, determines w_t. Senior working time, ℓ, is also determined by r independently of anything else except the degree of compensation, μ, and the rate of technological progress, g. Then we immediately find aggregate labor supply, L, through (5.55). Finally, aggregate capital, K, is determined through (5.57) and aggregate national wealth next period through (5.58).

5.4 Intertemporal substitution of labor supply

Until now individual labor supply has been considered endogenous only in one period of life. A more satisfactory treatment requires that we allow for elastic individual labor supply in several periods, thus taking the phenomenon of intertemporal substitution in labor supply into account. We first illustrate this phenomenon by considering an easily tractable two-period problem. Subsequently, a more general problem with many periods is studied.

A two-period example with additive period utility

Consider an individual who is a price taker and works and consumes over two periods. Assume there is no concern about subsequent periods and ignore possible uncertainty about getting fully employed. In compact form the decision problem is:

$$\max_{c_1,\ell_1,c_2,\ell_2} U = \ln c_1 - \gamma\frac{1}{1+\varepsilon}\ell_1^{1+\varepsilon} + \frac{1}{1+\rho}\left[\ln c_2 - \gamma\frac{1}{1+\varepsilon}\ell_2^{1+\varepsilon}\right] \quad \text{s.t.} \tag{5.59}$$

$$c_1 + \frac{c_2}{1+r} = a + w_1\ell_1 + \frac{w_2\ell_2}{1+r}, \tag{5.60}$$

$$c_1 \geq 0, \quad c_2 \geq 0, \quad 0 \leq \ell_1 \leq 1, \quad 0 \leq \ell_2 \leq 1.$$

Here, c_i and ℓ_i are planned consumption and labor supply, respectively, in period i, $i = 1, 2$, w_i is the (after-tax) real wage in the same period, a is a given initial financial wealth (which may be positive or negative), and r is the real interest rate. The parameter $\gamma > 0$ indicates the weight the individual attaches to the "disutility" of labor. For simplicity we ignore that this weight might depend on time, for instance being age dependent. The parameter $\varepsilon > 0$ measures the elasticity of disutility of labor. The inverse of ε will turn out to have *two* interesting interpretations, as we shall see. For simplicity, we restrict the analysis to the case where parameters and exogenous variables are such that the constraints $0 \leq \ell_i \leq 1$, $i = 1, 2$, are not binding.

To solve the problem we use the *substitution method.* Consider c_2, ℓ_1, and ℓ_2 as decision variables and let c_1 be determined by these and the intertemporal budget constraint (5.60):

$$c_1 = a + w_1\ell_1 + \frac{w_2\ell_2}{1+r} - \frac{c_2}{1+r}.$$

Substituting this into U, we find the first-order conditions:

$$c_1^{-1} = \frac{1+r}{1+\rho}c_2^{-1}, \tag{5.61}$$

$$\ell_1 = \left(\frac{w_1}{\gamma c_1}\right)^{1/\varepsilon} \equiv \left(\frac{w_1}{\gamma c_1}\right)^{\sigma}, \tag{5.62}$$

$$\ell_2 = \left(\frac{w_2(1+\rho)}{\gamma c_1(1+r)}\right)^{1/\varepsilon} \equiv \left(\frac{w_2}{\gamma c_2}\right)^{\sigma}, \tag{5.63}$$

where we have defined $\sigma \equiv 1/\varepsilon > 0$. Substituting the first-order conditions into (5.60) yields

$$\frac{2+\rho}{1+\rho}c_1 - \gamma^{-\sigma}\left[w_1^{1+\sigma} + (1+\rho)^{\sigma}(\frac{w_2}{1+r})^{1+\sigma}\right]c_1^{-\sigma} = a. \tag{5.64}$$

We can write this as $bc_1 - a = hc_1^{-\sigma}$, where b and h are positive constants determined by w_1, w_2, and r. A graph of the left-hand side of (5.64) as a function of c_1 will quickly convince the reader that this equation has a unique positive solution, c_1^*, which we may write as an implicit function of a, w_1, w_2, and r, $c_1^* = c_1(a, w_1, w_2, r)$. Inserting this into (5.61), (5.62), and (5.63) gives the unique implicit solution for c_2, ℓ_1, and ℓ_2, respectively.

The first-order conditions (5.62) and (5.63) are interesting. We see that the constant $\sigma \equiv 1/\varepsilon$ enters as an exponent on w_1/c_1 and w_2/c_2, respectively. Hence σ measures, for fixed current consumption, the elasticity of labor supply in each period with respect to the wage in the same period. It follows that this elasticity is in our example the same at any point $(c_1, \ell_1, c_2, \ell_2)$.

A key special feature of the period-utility function in our example is the additive separability in its arguments. Then, holding current consumption fixed is equivalent to holding the marginal utility of current consumption fixed, which is in turn – again due to the separability – equivalent to holding current marginal utility of wealth fixed. Hence, also what is known as the *Frisch elasticity of labor supply*, see next sub-section, will in our example equal the constant $\sigma \equiv 1/\varepsilon$ and is thus the same at any point $(c_1, \ell_1, c_2, \ell_2)$.

Another elasticity concept is the *elasticity of intertemporal substitution* for labor supply. In the present example this is identical to the elasticity of ℓ_1/ℓ_2 with respect to the corresponding price ratio in present-value terms, which is $w_1/\left[w_2/(1+r)\right]$, when we move along a given indifference curve in the (ℓ_1, ℓ_2) plane and so keep total discounted utility, U, fixed. Combining (5.62) and (5.63) gives the labor ratio

$$\frac{\ell_1}{\ell_2} = (1+\rho)^{-\sigma}\left(\frac{w_1}{w_2/(1+r)}\right)^{\sigma}. \tag{5.65}$$

As σ here enters as an exponent on the relevant "price ratio", $w_1/\left[w_2/(1+r)\right]$, it follows that σ measures not only the c_t-fixed labor supply elasticity with respect to the wage in a given period but also the elasticity of intertemporal substitution.[28] So also this elasticity is in our example the same at any point $(c_1, \ell_1, c_2, \ell_2)$.

Because $\sigma > 0$, (5.65) indicates that a "temporary" rise in the wage ratio (w_1 goes up or w_2 goes down or both) elicits a rise in ℓ_1/ℓ_2, i.e., relatively more labor now and relatively less in the next period where leisure has become relatively cheaper than before. Thus σ measures how sensitive the relative allotment of labor to the two periods is to a temporary change in the relative wage.

So far we have not found an explicit solution for each endogenous variable separately. To find an explicit solution, we simplify by assuming $a = 0$. Then (5.64) gives the explicit solution

$$c_1^* = \left(\frac{1+\rho}{(2+\rho)\gamma^\sigma}\right)^{\frac{1}{1+\sigma}} \left[1 + (1+\rho)^\sigma (\frac{w_2}{w_1(1+r)})^{1+\sigma}\right]^{\frac{1}{1+\sigma}} w_1.$$

Substituting this into the above formula for ℓ_1 gives

$$\ell_1 = \left(\gamma \frac{1+\rho}{2+\rho}\right)^{\frac{-\sigma}{1+\sigma}} \left[1 + (1+\rho)^\sigma \left(\frac{w_2}{w_1(1+r)}\right)^{1+\sigma}\right]^{\frac{-\sigma}{1+\sigma}}.$$

And this combined with (5.65) gives

$$\ell_2 = \left(\gamma \frac{1}{2+\rho}\right)^{\frac{-\sigma}{1+\sigma}} (1+\rho)^{\frac{\sigma^2}{1+\sigma}} \left[\left(\frac{w_1(1+r)}{w_2}\right)^{1+\sigma} + (1+\rho)^\sigma\right]^{\frac{-\sigma}{1+\sigma}}.$$

We see that $w_1 \uparrow \Rightarrow \ell_1 \uparrow$ and $\ell_2 \downarrow$. That is, a temporary wage increase leads to substitution of labor for leisure in the current period and the opposite in the next period.

We also see that a "permanent wage change" (the wage in both periods is multiplied by the same positive factor) leaves labor supply unchanged in both

[28] Conceptually, an elasticity of substitution is defined with reference only to preferences, i.e., to the utility function as such (or, as in Chapter 4.5, to a production function). A general definition of the elasticity of intertemporal substitution is given in Chapter 3.3 for consumption and does not involve any price ratio but only the marginal rate of substitution, MRS, between the two consumption goods in question. For labor supply (or leisure) the definition is analogue. Now, if the considered individual is a price taker in the markets involved, then, at the individual optimum, MRS between working in period 1 and period 2 equals the relevant price ratio, $w_1/(w_2(1+r))$. This is why (5.65) indicates that σ measures the elasticity of intertemporal substitution in labor supply. In our specification (5.59) the elasticity of intertemporal substitution for consumption is also a constant but equal to 1.

periods. For interpretation, it is as usual convenient to rewrite (5.60) such that the exogenous total "endowment" (wealth) appears alone on the right-hand side:

$$c_1 + w_1(1-\ell_1) + \frac{c_2}{1+r} + \frac{w_2}{1+r}(1-\ell_2) = w_1 + \frac{w_2}{1+r}.$$

When both w_1 and w_2 increase, there is a substantial positive wealth effect on leisure in the two periods, $(1-\ell_1)$ and $(1-\ell_2)$. This fully offsets the negative substitution and pure income effects on leisure. Furthermore, there is no change in the relative price of leisure across the two periods. Therefore there is no intertemporal substitution of leisure. These clear-cut results are of course dependent on the particular preference specification in (5.59) combined with the no-initial-financial-wealth assumption. The analyzed case should be thought of as a benchmark case.

According to one of Kaldor's stylized facts, the real wage is in the long run growing at the same rate as productivity. This is to say that apart from minor temporary fluctuations, w_{t+1}/w_t tends to be constant. In turn, given the considered preference specification, there should thus be no trend in annual labor supply per employed person. Consequently, the predicted uncompensated labor supply elasticity at the intensive margin is *nil.* A quick look back in economic history of industrialized countries over more than a century, with systematic growth in the after-tax real wage along with systematic decline in annual hours per employed person,[29] suggests that the uncompensated labor supply elasticity (at least in some lagged form) at the intensive margin is in fact *negative,* reflecting that the *wealth effect dominates* in the long run. This raises doubt about the realism of the preference specification (5.59) and similar specifications that have for a long time been popular in macroeconomics.

A more general problem with many periods. The Frisch labor supply elasticity*

Let $u(c, 1-x)$ be a concave period-utility function with positive but diminishing marginal utilities with respect to consumption and *leisure* $x \equiv 1-\ell \in [0,1]$. Let the utility discount factor be $\beta \equiv (1+\rho)^{-1} \in (0,1)$. Consider the problem: choose a sequence, $\{(c_0,\ell_0),(c_1,\ell_1),\ldots,(c_{T-1},\ell_{T-1})\}$ so as to

$$\begin{aligned} \max U &= \sum_{t=0}^{T-1} \beta^t u(c_t,\ell_t) \qquad \text{s.t.} \\ a_{t+1} &= (1+r_t)a_t + w_t\ell_t - c_t, \qquad a_0 \text{ given}, \qquad (5.66) \\ a_T &\geq 0, \qquad \text{(solvency)} \end{aligned}$$

[29] See, e.g., Gali (2005) and Huberman and Mins (2007).

where a_t is real financial wealth at the beginning of period t. The individual can, within limits, freely lend and borrow at the interest rate r_t. The limit is indicated by a terminal solvency constraint, saying that at the end of the last period all debt must be settled.[30]

To solve the problem we apply the Lagrange method, with one multiplier for each period. This approach is convenient because the multipliers involved are of theoretical as well as econometric interest. Since any plan with $a_T > 0$ can be improved upon by consuming more, it cannot be optimal. Hence a solution must have $a_T = 0$ and this allows us to concentrate on the Lagrange function

$$L = \sum_{t=0}^{T-1} \beta^t u(c_t, \ell_t) - \sum_{t=0}^{T-1} \mu_t \left[c_t - w_t\ell_t - (1 + r_t)a_t + a_{t+1}\right]$$

where $a_T = 0$ and μ_t is the (discounted) Lagrange multiplier for period t, $t = 0, 1,\ldots, T-1$. It is convenient to rewrite the Lagrange function as

$$L = \sum_{t=0}^{T-1} \beta^t \left[u(c_t, \ell_t) - \lambda_t(c_t - w_t\ell_t - (1 + r_t)a_t + a_{t+1})\right], \tag{5.67}$$

where $a_T = 0$, and $\lambda_t \equiv \mu_t\beta^{-t}$ is the undiscounted Lagrange multiplier for period t. Following the standard procedure, we then partially differentiate L with respect to the $2 \cdot T + T - 1$ endogenous variables $c_0,\ldots,c_{T-1}$, $\ell_0,\ldots,\ell_{T-1}$, and $a_1,\ldots,a_{T-1}$ and next equate these partials to zero. For c_t, ℓ_t, and a_{t+1}, respectively, we get

$$\frac{\partial L}{\partial c_t} = 0 \Rightarrow u_1(c_t, \ell_t) = \lambda_t, \qquad t = 0, 1, \ldots, T-1, \tag{5.68}$$

$$\frac{\partial L}{\partial \ell_t} = 0 \Rightarrow u_2(c_t, \ell_t) = -\lambda_t w_t, \qquad t = 0, 1, \ldots, T-1, \tag{5.69}$$

and

$$\frac{\partial L}{\partial a_{t+1}} = 0 \Rightarrow \lambda_t = \beta(1 + r_t)\lambda_{t+1}, \qquad t = 0, 1, \ldots, T-2. \tag{5.70}$$

The latter first-order equations are the Euler equations expressed in terms of the lagrange multipliers.

Suppose interior solutions for the c's and ℓ's exist. Then, from the $2 \cdot T + T - 1$ first-order conditions, together with the T dynamic budget identities of form (5.66), we can in general, solve for the $2 \cdot T + T - 1 + T$ unknowns (the c's and

[30] For simplicity we ignore that for a realistic life-cycle perspective, the period-utility function u should have a third argument, a vector of individual characteristics such as age, household composition, area of residence etc.

ℓ's, the a's, and the λ's), at least numerically. The resulting solution vectors, $(c_0, \ldots, c_{T-1})$, $(\ell_0, \ldots, \ell_{T-1})$, and $(a_1, \ldots, a_{T-1})$, will then make up a solution to the optimization problem.

In case one is not interested in the λ's, the equation structure allows their elimination before solving for the c's, ℓ's, and a's.

Now, however, we keep the λ's and focus on the first-order conditions (5.68) and (5.69). The purpose is to define the concept of a *Frisch elasticity of labor supply* with respect to the wage. Suppose $u(c, 1-x)$ is *strictly* concave in consumption and leisure. Given the utility function u, we can then (explicitly or perhaps only implicitly) solve the equations (5.68) and (5.69) uniquely for c_t and ℓ_t to get the consumption and labor supply functions:[31]

$$c_t = c(w_t, \lambda_t), \tag{5.71}$$
$$\ell_t = \ell(w_t, \lambda_t). \tag{5.72}$$

What is in the literature known as the *Frisch (or λ_t-fixed) elasticity of labor supply* with respect to the (after-tax) wage in period t is the partial elasticity:

$$\frac{\ell_t}{w_t} \frac{\partial \ell(w_t, \lambda_t)}{\partial w_t} \equiv \eta(w_t, \lambda_t). \tag{5.73}$$

In this expression λ_t can, in optimum, be interpreted as the current *marginal utility of wealth,* wealth being defined as the sum of initial financial wealth, a_0, and potential human wealth, $h_0 \equiv \sum_{t=0}^{T-1} w_t \cdot 1/\Pi_{i=0}^{t}(1+r_i)$. To understand this interpretation of λ_t, imagine an exogenous inflow of one unit of real wealth ("manna from heaven") in period t. Everything else equal, this increases the factor $(c_t - w_t\ell_t - (1+r_t)a_t + a_{t+1})$ in (5.67) by one unit over and above the level zero in the original optimum. This allows a new optimum with $u(c_t, \ell_t)$ approximately λ_t units larger than before (only approximately because we consider finite changes, not infinitesimal changes). The formula (5.73) gives the percentage increase in ℓ in response to a one-percentage rise in w_t, conditional on a fixed λ_t. This conditioning means keeping the marginal utility of wealth fixed. So the formula (5.73) deals with a thought experiment requiring a "compensating" reduction of another income component (either the wage in another period or some non-labor income) so as to leave the marginal utility of wealth unchanged in spite of the initial rise in w_t.

The general point is that when we want to study the effect on labor supply of a change in the (after-tax) wage, we may want to condition on other endogenous

[31] For analytical characterization of the resulting functions, a method analogue to that used in Chapter 2.4.1 for a profit maximization problem with strictly concave production function can be applied.

variables in order to *isolate* different causal channels. In the static one-period setup of Section 5.2 we considered not only the total labor supply effect of a rise in the wage but also the Hicks-compensated labor supply effect. That is the case where the individual is "compensated" for the pay change by an adjustment of the budget so that he or she is just able to stay on the same indifference curve. In this way we separated the substitution effect from the total income effect.

In the simple problem of the previous sub-section where the period-utility function is additively separable in consumption and leisure, the conditioning variable is current consumption. Given the additive separability, holding current consumption fixed is equivalent to holding the marginal utility of current consumption fixed, which is in turn equivalent to holding current marginal utility of wealth fixed. In that specific case the *Frisch (or* λ_t-fixed) *labor supply elasticity* is the same as the c_t*-fixed labor supply elasticity.* Generally, they are not the same although both deal with intertemporal substitution.

For theoretical and empirical analysis, it seems in general more convenient to condition on λ_t, the marginal utility of wealth. This approach nicely disentangles the within-period trade-offs as in (5.68) and (5.69) from the trade-off over time, as described by the Euler equations (5.70). Hence a sizable empirical literature has attempted to estimate parameters of the functions c and ℓ and in so doing assess the Frisch labor supply wage elasticity. The evidence summarized by Browning et al. (1999) suggests a Frisch elasticity between 0.1 and 0.4 for annual hours by men, but higher for women.

Here the manuscript breaks..

5.5 Concluding remarks

5.6 Literature notes

(Incomplete)

Section 5.1 focused on macroeconomic aspects of social security and did not, for example, discuss the different reasons for having mandatory pension schemes, such as myopia problems, social security politics, imperfections on life insurance markets and issues related to income distribution within and between generations. Regarding this kind of matters, the reader is referred to, e.g., Diamond (2003) and, in Danish, Velfaerdskommissionen (2004). Shiller (2005) and Bovenberg et al. (2008) discuss advantages and disadvantages of mandatory individual savings accounts for social insurance. The symposium in *Journal of Economic Perspectives,* vol. 19, no. 2, 2005, contains a series of different views on social security and social security reform.

The claim in Section 5.1 that there is no Pareto-improving way to set up a transition from tax-based pension to a funded system is explored further in Breyer (1989). It is another matter that if the means to compensate the last generation that paid a contribution as young can be obtained by removing some inefficiency in the system, then *many* Pareto-improving re-allocations exist, including some that are compatible with a transition to a funded system; about this, see de la Croix and Michel (2002).

Labor supply elasticities, theory and empirics: Stern (1986), Browning et al. (1999), Blundell and MaCurdy (1999). On the econometric difficulties involved in reaching sharp conclusions about labor supply responses to changes in taxes, see e.g. Manski (2012).

There is debate between Prescott (2003) and Blanchard (2004??) about the explanation of the difference between American and European employment ratios.

5.7 Appendix: The extended Slutsky equation

Consider a simple two-goods consumer problem. Let $U(x_1,x_2)$ be a strictly quasi-concave utility function, where x_1 is consumption of good 1 and x_2 is consumption of good 2. The budget constraint is

$$p_1x_1 + p_2x_2 = y,$$

where p_1 and p_2 are the given prices on the two goods and y is the "wealth" or "budget" of the consumer. Let $x_i = x_i(p_1, p_2, y)$, $i = 1, 2$, be the resulting Walrasian demand function. Then the total derivative of x_i with respect to p_j, taking into account that also the budget may change when p_j changes, can be written

$$\begin{aligned} \frac{dx_i(p_1,p_2,y)}{dp_j} &= \frac{\partial x_i}{\partial p_j}\Big|_{U=U_0} + \left(-\frac{\partial x_i(p_1,p_2,y)}{\partial y}x_j\right) + \frac{\partial x_i(p_1,p_2,y)}{\partial y}\frac{dy}{dp_j}, \\ \text{for } i,j &= 1,2, \end{aligned} \tag{5.74}$$

This equation is known as the *extended Slutsky equation.* The first term on the right-hand side is the partial derivative of the Hicksian demand function with respect to p_j, evaluated at the original price-utility combination (the Hicksian demand function gives the demand (x_1,x_2) as a function of prices (p_1,p_2) if y is adjusted to keep the level of utility constant at the original level, U_0).[32] In words:

$$\text{total effect} = \text{substitution effect} + \text{pure income effect} + \text{wealth effect}. \tag{5.75}$$

[32]For derivation of (5.74), see for example Varian (1992).

This is the terminology we generally apply in this book. In the literature sometimes the substitution effect is called the "demand effect under a Hicksian wealth compensation" or just the "Hicks-compensated effect". Under the Hicksian wealth compensation the individual is "compensated" for the price change by an adjustment of the budget so that he or she is just able to stay on the same indifference curve which exactly corresponds to the first term on the right-hand side of (5.74). What Slutsky himself considered was a situation where the individual is compensated for the price change by an adjustment of the budget so that he or she is just able to still buy the original optimal bundle of consumption goods (the "Slutsky wealth compensation"). Fortunately, considering only infinitesimal changes in prices, the Hicks-compensated effect turns out to be exactly equal to the Slutsky-compensated effect (Varian, 1992, p. 135).

The "given budget" serving as reference budget when the pure income effect is calculated is under a Hicksian wealth compensation the budget that corresponds to unchanged utility after the price change; under a Slutsky compensation it is a budget such that the original optimal bundle can be just afforded. Again, when considering only infinitesimal changes in prices, the difference is inconsequential.

The decomposition (5.74) is applicable to a wide range of issues, such as how changes in the wage affects the labor supply/leisure choice or the consumption/saving choice or how a change in the interest rate affect these choices. In the *simple* or ordinary Slutsky equation from partial equilibrium microeconomics, y is considered as exogenous, so that the last term on the right-hand side of (5.74) drops out.

It is sometimes convenient to compress the pure income effect and the wealth effect into one term. Then equation (5.74) simplifies to

$$\frac{dx_i(p_1, p_2, y)}{dp_j} = \frac{\partial x_i}{\partial p_j}\Big|_{U=U_0} + \frac{\partial x_i(p_1, p_2, y)}{\partial y}\left(\frac{dy}{dp_j} - x_j\right), \tag{5.76}$$

for i, $j = 1, 2$. If good i is a normal good, $\partial x_i/\partial y$ is positive and then the sum of the pure income and wealth effects has the same sign as $dy/dp_j - x_j$. The sum of the pure income and wealth effects is called the *total income effect.* As noted in the text, sometimes in the macroeconomic and labor market literature the prefix "total" is dropped. This is unfortunate since it may lead to confusion with the pure income effect which is in *microeconomics* usually just called the income effect. Whenever there is a risk of confusion between the two concepts, we shall therefore in this book add the prefix "pure" or "total" when speaking of income effects. Notwithstanding the terminological issues, there are many contexts, in particular when dynamics is considered, where it is expedient to use the full Slutsky decomposition given in (5.75).