

# Heterogeneity and the Welfare Cost of Dynamic Factor Taxes

Zsolt Becsi

Federal Reserve Bank of Atlanta  
Working Paper 99-2  
March 1999

**Abstract:** The welfare costs of dynamic factor taxes are analyzed in a dynamic general equilibrium model with heterogeneous endowments, abilities, and tastes. Conventional functional form restrictions yield formulas for the transition effects and marginal welfare costs of factor taxes. Heterogeneity implies that taxes have feedback or distribution effects, beyond standard efficiency effects, that may lead to nonstandard aggregate dynamics. Also, marginal welfare costs vary systematically with initial distortions and agents' characteristics. Because factor taxes lower wealth inequality, equity gains offset efficiency losses with the offset weakening as initial distortions rise. However, distribution effects reinforce efficiency losses unless preexisting distortions are sufficiently high, in which case some types of heterogeneity yield offsetting distribution effects. Simulations suggest that, for labor taxes, distribution effects dominate dynamics, but not for capital taxes. Also, equity gains dominate efficiency losses and distribution effects for the marginal welfare cost of labor taxes, and vice versa for capital taxes.

JEL classification: D30, D6, H20, D90, D5

Key words: heterogeneity, welfare, factor taxation, dynamic general equilibrium

---

Much of this research is based on Becsi (1993), but insightful comments prompted significant extensions. The author is grateful to Robert Haveman, Martin David, and Peter Streufert for their support and invaluable suggestions. He also thanks Graziella Bertocchi, John Duca, Scott Freeman, Greg Huffman, Evan Koenig, Finn Kydland, Ed Prescott, Alan Stockman, Mark Wynne, Ping Wang, and seminar participants at the University of Texas, the University of Wisconsin, Pennsylvania State University, and Purdue University. The views expressed here are the author's and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the author's responsibility.

Please address questions regarding content to Zsolt Becsi, Research Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, NW, Atlanta, Georgia 30303-2713, 404/521-8785, 404/521-8058 (fax), [zsolt.becsi@atl.frb.org](mailto:zsolt.becsi@atl.frb.org).

To receive notification about new papers or to order copies of printed papers, contact the Public Affairs Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, NW, Atlanta, Georgia 30303-2713, 404/521-8020. The full text of this paper may be downloaded (in PDF format) from the Atlanta Fed's World-Wide Web site at [www.frbatlanta.org/publica/work\\_papers/](http://www.frbatlanta.org/publica/work_papers/).

# Heterogeneity and the Welfare Cost of Dynamic Factor Taxes

## Introduction

Because redistribution is an essential component of tax policy, a comprehensive welfare cost analysis of tax policies should encompass both efficiency and equity considerations. Formal welfare analysis of taxation was ushered in by Mirrlees (1971), who compared the static costs of distorting markets with the gains from redistribution. Later research demonstrated that when agents differ along one or more dimensions and differentiated lump-sum taxes are not feasible, welfare costs are characterized by equity-efficiency tradeoffs. However, this line of research has not been pursued in the dynamic taxation literature, which has generally eschewed distributional issues to concentrate instead on dynamic efficiency effects.<sup>1</sup> One reason why distributional issues have been neglected is that dynamic analyses have severely restricted heterogeneity for reasons of computational tractability, not for any theoretical or empirical reasons.<sup>2</sup> While this approach has yielded important insights, the price paid for tractability includes aggregation bias that may distort inferences about transitional dynamics and a focus on efficiency costs that limits the value of dynamic tax policy prescriptions.

Heterogeneity presents a key challenge to a full welfare-cost analysis of dynamic factor taxation. This paper shows that when three dimensions of heterogeneity are introduced into an otherwise standard neoclassical dynamic general equilibrium model, the resulting equilibrium will depend in a fundamental way on distributional considerations. Factor taxes will have transitional dynamic effects that depend critically on the distribution of agent characteristics, with non-standard dynamics a distinct possibility. Also, welfare costs will be characterized by an equity-efficiency tradeoff as factor taxes tend to lower wealth inequality. Still, whether welfare losses are less or greater than in a representative agent model depends on the exact distribution of characteristics.

---

<sup>1</sup> This is especially true for infinite horizon models, which are the focus of this paper. Instances of such models with representative agents are Hall (1971) and Judd (1987a and 1987b). Chamley (1986) analyzes optimal taxation with limited heterogeneity, while Judd (1985) provides non-optimal tax analysis with fewer restrictions on heterogeneity. Overlapping generations models, such as Blanchard (1985), Auerbach and Kotlikoff (1987), or Laitner (1990), analyze taxation with age-related heterogeneity but typically without intragenerational heterogeneity.

<sup>2</sup> See, for instance, the recent critical discussions by Kirman (1992), Stoker (1993), and Boadway (1994).

Three familiar forms of heterogeneity are considered here: endowments, abilities or market productivities, and tastes for leisure and consumption, where relatively strong preferences for leisure can be interpreted as high home productivities. These forms of heterogeneity are well known from microeconomic studies where observed differences in endowments and abilities, as well as unobserved differences in tastes, are shown to play an important role in determining labor (Killingsworth, 1983) and consumption (Blundell, Pashardes, and Weber, 1993). Research in theoretical public finance has also adopted these types of heterogeneity. For example, Mirrlees (1971) focused on endowments and abilities while Diamond's (1975) extension of Ramsey taxation to the multiagent case looked at taste and endowment differences.<sup>3</sup> By contrast, dynamic models have carefully restricted heterogeneity, as illustrated by Judd (1985), who allows endowments and sometimes tastes to vary but then limits the fallout.<sup>4</sup> Moving from a two-class model where heterogeneity has no effect on dynamics to a less restrictive setting where the dynamics and welfare could be affected by heterogeneity, Judd restricts the analysis to the long-run effects of factor taxes and prevents distributional concerns from mattering through differentiated lump-sum rebates.

Expanding the scope of dynamic models to include agent heterogeneity creates serious problems. Some forms of heterogeneity yield two-way interaction between the wealth distribution and aggregate dynamics, particularly if wealth distribution effects do not cancel in the aggregate and the wealth distribution in turn depends on the time path of aggregate prices. The interaction changes the nature of the aggregate dynamics, and this result of course has social welfare consequences apart from the direct inequality effects on welfare. Macroeconomists have typically dealt with the feedback problem by immunizing factor prices from distribution effects through restrictions on heterogeneity, making the wealth distribution exogenous, or else by dropping capital accumula-

<sup>3</sup> As recognized by many -- for example, Guesnerie (1995) -- Mirrlees's ability parameter can also be interpreted as a taste parameter because it enters as a parameter in the utility function.

<sup>4</sup> Rios-Rull (1995) surveys some recent dynamic models with heterogeneity, all without taxation. Recent work mainly focuses on ex ante homogeneity with ex post heterogeneity in production efficiency (Aiyagari, 1994) or tastes (Atkeson and Lucas, 1992) where idiosyncratic shocks average out with large numbers. Bencivenga (1992) interprets taste shocks as emanating from home production, while Blanchard and Fischer (1989, p. 291) suspect that aggregation is important, an idea that is considered here.

tion to compute equilibria (Rios-Rull, 1995). The current paper takes a different tack by providing a theoretical solution that overcomes the computational problem without compromising dynamics or welfare analysis. The device is functional form assumptions that have achieved a sort of benchmark status and thus readily afford comparability with existing results.<sup>5</sup> Not only do the assumptions ensure tractability, but they also yield an analytical solution to the equilibrium as well as explicit formulas for the transitional dynamics and the marginal welfare costs of labor and capital taxes.

The theoretical findings can be grouped according to different degrees of heterogeneity. Under weak heterogeneity, defined as differences in endowments and abilities only, the effect of any tax change on the paths of aggregate quantities and factor prices equals a “representative agent effect,” or the standard efficiency or dynamic substitution effects found by Hall (1971), Judd (1985, 1987a,b) and others.<sup>6</sup> The cumulative distortion of these representative agent effects explains the efficiency loss of factor taxes. However, factor taxes also have offsetting equity effects. They reduce wealth inequality through a positive human wealth effect that disproportionately benefits the poor. Thus, because there are equity gains from reduced inequality, the marginal welfare costs of factor taxes with weak heterogeneity are lower than under the representative agent case, with the difference narrowing with initial distortions. When agents are also allowed taste differences, denoted strong heterogeneity, a “distribution effect” comes into play that reflects feedback effects from differential wealth changes. Now, the total effect of any policy change on the paths of aggregate variables decomposes into a representative agent effect and a distribution effect, where the direction of the latter depends on whether wealth redistribution favors agents with high or low propensities to consume. The potential for non-standard dynamics is compounded because the distribution effect is persistent since even temporary policies have permanent wealth effects. Also, the distribution effect

---

<sup>5</sup> As Long and Plosser (1983) showed in a stochastic framework, assuming Cobb-Douglas production, logarithmic utility, and full capital depreciation yields a reduced-form solution for the infinite horizon model with capital accumulation. The assumptions are popular for dynamic models (King, Plosser, and Rebelo, 1988a) and can be readily extended to include overlapping generations (Blanchard, 1985).

<sup>6</sup> While King, Plosser, and Rebelo (1988b) show that skill differences do not affect dynamics, they do not look at the consequences for welfare analysis.

tends to aggravate the marginal welfare costs of factor taxes unless preexisting distortions are sufficiently high, in which case the exact specification of all individual characteristics becomes very important. These results can be thought of as extending Judd's (1987a) work on marginal welfare costs by including equity gains and distribution effects in a dynamic framework.

Numerical examples are computed to show how sensitive the dynamics and the marginal welfare cost formulas are to variations in the distribution of agent characteristics and a few critical parameters and also to different timing assumptions. The examples suggest that non-standard dynamics are less likely for capital taxes than for labor taxes, where distribution effects can easily dominate. Also, the examples reveal that labor taxes are better suited for redistribution than capital taxes, because labor taxes tend to have larger inequality effects and capital taxes have larger effects on capital and efficiency. In fact, equity gains for labor taxes are often so large relative to efficiency losses and distribution effects that the marginal welfare cost is usually positive under strong heterogeneity. Even though the equity gains for capital taxes are not insignificant, they pale next to the efficiency and distribution effects so that the marginal welfare cost usually remains negative and may even be quite large. The computed efficiency costs of labor and capital taxes come close to those of Judd (1987a), with the range for capital taxes being very sensitive to parameter and timing assumptions. The examples also demonstrate that the size of the initial government debt can be an important determinant of the disparity between the efficiency cost of capital taxes and the efficiency cost of labor taxes and can explain qualitative differences among studies such as Judd (1987a), Auerbach and Kotlikoff (1987), and McGrattan (1994). Finally, while empirical research gives little direct guidance on the joint distribution of agent characteristics, it is shown that wealthy agents are likely to have high abilities in home and market production and low propensities to consume. Other assumptions on the distribution of agent characteristics lead to outcomes that tend to be implausible or even counterfactual.

## I. The Model

The model consists of three sectors. In the production sector, perfectly competitive firms hire capital and efficiency-weighted labor to produce a single consumption good using a constant-returns-to-scale technology. The household sector has a constant population of infinitely lived agents with heterogeneous capital endowments, market productivities, and tastes (or non-market productivities). Agents choose feasible time paths for consumption, leisure, and savings in productive capital and government bonds to maximize intertemporally separable preferences. The government finances its expenditures and lump-sum transfers by levying proportional taxes on wage and interest income or by issuing bonds. All government instruments satisfy an intertemporal revenue constraint.

The production sector is fairly standard. In periods  $s \geq 1$ , competitive firms combine labor,  $h_s$ , and a predetermined stock of physical capital,  $k_{s-1}$ , to produce goods,  $y_s$ , using a Cobb-Douglas production technology:

$$(1) \quad y_s = k_{s-1}^\theta h_s^{1-\theta}$$

where  $\theta$  is the share of capital. Labor is measured in efficiency units, that is,  $h_s = \sum_i n^i a^i h_s^i$ , where  $n^i$  is the population share of group  $i$ ,  $a^i$  indexes individual market productivity, and  $h_s^i$  is individual labor time. Capital is given by  $k_{s-1} = \sum_i n^i k_{s-1}^i$  and depreciates fully within a period so that investment is the next period's capital stock. Under full depreciation, output can be interpreted as net of depreciation. Given interest rates  $r_s$ , individual wage rates  $a^i w_s$ , and average wage rates  $w_s$ , firms maximize profits by choosing inputs to equate marginal products with their costs:

$$(2) \quad a^i w_s = a^i (1-\theta) y_s / h_s \quad r_s = \theta y_s / k_{s-1}$$

The household sector is also standard except for heterogeneous tastes and wealth. Each agent values consumption and leisure obtained in each period of an infinitely long life that starts in period one. Individuals' lifetime utility is defined as:

$$(3) \quad \sum_{s \geq 1} \rho^{s-1} u^i(c_s^i, l_s^i) = \sum_{s \geq 1} \rho^{s-1} [\alpha^i \ln(c_s^i) + (1-\alpha^i) \ln(l_s^i)]$$

where  $c_s^i$  and  $l_s^i = 1 - h_s^i \in (0, 1)$  are consumption and the fraction of time devoted to leisure or non-market hours. The rate of time preference,  $\rho \in (0, 1)$ , discounts utility in future periods,  $u^i(\cdot)$ , and reflects agents' impatience. Preference heterogeneity arises because individuals have different tastes for consumption and leisure. Compared with agents with low  $\alpha^i$ , high- $\alpha^i$  agents derive more pleasure from consumption and less from leisure, making them consumption-lovers and workaholics. Alternatively, for some household production formulations, a high taste for consumption can also be interpreted as low productivity in non-market or home production.<sup>7</sup>

Agents choose time streams of consumption and leisure that maximize (3) subject to an intertemporal budget. The budget constraint, derived by time-aggregating agents' intratemporal revenue constraints and ensuring that their transversality conditions hold, requires that the discounted stream of expenditures on consumption and leisure does not exceed full wealth,  $z^i$ :

$$(4) \quad \sum_{s \geq 1} \frac{\pi_s}{y_s} \left[ c_s^i + (1 - t_{ws}) w_s a^i l_s^i \right] = z^i$$

$$(5) \quad \frac{\pi_s}{y_s} \equiv \beta_s \equiv \prod_{u=1}^s \{(1 - t_{ru}) r_u\}^{-1}$$

with labor and capital tax rates of  $t_{ws}$  and  $t_{rs}$  and where  $\beta_s$  is the present value discount factor.

The discount factor and other variables to come are defined as fractions of output as an expositional shortcut because the corresponding aggregates are easily recovered in the analytical solution below.

Full wealth is defined as the endowment of capital and government bonds plus the present value of after-tax "full" labor earnings and lump-sum transfers:

$$(6) \quad z^i \equiv (k_0^i + b_0^i) + z_a^i \equiv (k_0^i + b_0^i) + \sum_{s \geq 1} \pi_s \frac{(1 - t_{ws}) w_s a^i + t_s}{y_s}$$

where  $t_s$  represents lump-sum transfers. Wealth is heterogeneous because human wealth,  $z_a^i$ , varies with individual labor productivity and also because non-human wealth varies with endowments of

---

<sup>7</sup> To see the formal equivalence, define home production as  $L_s^i \equiv (l_s^i)^{(1-\alpha^i)/\alpha^i}$  and redefine preferences in (3) over consumption of market and non-market goods so that  $u^i(\cdot) \equiv \ln(c_s^i) + \ln(L_s^i)$ . Also, see Greenwood, Rogerson, and Wright (1995), who survey the literature on household production in dynamic economies.

$k_0^i$  shares of productive capital and  $b_0^i$  units of government bonds. Both assets are perfect substitutes in the individual's savings portfolio, with  $k_{s-1}^i$  and  $b_{s-1}^i$  earning  $(1-t_{rs})r_s$ .

Individual households choose streams of consumption and leisure that maximize (3) subject to (4). Solving for the constrained utility maximization leads to familiar conditions:

$$(7) \quad \frac{1-\alpha^i}{\alpha^i} \frac{c_s^i}{l_s^i} = (1-t_{ws})w_s a^i, \quad \frac{1}{\rho} \frac{c_{s+1}^i}{c_s^i} = (1-t_{rs+1})r_{s+1}$$

where the marginal rate of substitution between consumption and leisure is equated to the relative price of leisure and the marginal rate of substitution between consumption in period  $s$  and  $s+1$  is equated to the period  $s+1$  return to savings.

The government finances expenditures and lump-sum transfers from tax revenues that satisfy an intertemporal revenue constraint described by:

$$(8) \quad b_0^g = \sum_{s \geq 1} \pi_s \left[ \gamma_s + \frac{t_s}{y_s} - \tau_s \right]$$

where  $b_0^g$  is the initial amount of government bonds outstanding and  $\gamma_s$  is the fraction of output devoted to government expenditures. Tax revenues from wage and interest income are

$$(9) \quad \tau_s y_s = t_{ws} w_s \sum_i n^i a^i h_s^i + t_{rs} r_s \sum_i n^i k_{s-1}^i + t_{rs} r_s (\sum_i n^i b_{s-1}^i + b_{s-1}^g)$$

Finally, goods, factor, and asset markets are assumed to clear in all periods. In particular, equilibrium in the goods and bond markets is given by:

$$(10) \quad c_s + k_s + \gamma_s y_s = y_s$$

$$(11) \quad \sum_i n^i b_s^i + b_s^g = 0$$

where aggregate consumption is  $c_s = \sum_i n^i c_s^i$ . Also, factor market equilibrium is attained when the factor prices seen by households and firms are equated. This concludes the description of the model, which has a perfect foresight equilibrium with standard properties:

**Definition.** (*Dynamic Equilibrium*) A perfect foresight equilibrium consists of sequences of optimal plans for household consumption, labor, and savings and sequences of optimal plans for

*firm output and inputs that perfectly forecast the time path of all prices and government variables and clear goods, factor, and asset markets.*

## II. Solution for the Perfect Foresight Equilibrium

As a first step toward a solution of the model, aggregate analogs of the individual optimality conditions are found by relating individual demands to their aggregate counterparts. Substitution of equation (7) into (3) yields individual demands for consumption and leisure that are linear in initial wealth. To aggregate the individual demands, define the aggregation or averaging operator by  $E_{xy} \equiv \sum_i n^i x^i y^i$  and the covariance operator by  $S_{xy} \equiv \sum_i n^i (x^i - E_x)(y^i - E_y) = E_{xy} - E_x E_y$ . Then if agent  $i$ 's wealth share is defined by  $\sigma^i \equiv z_1^i / E_z$  so that group  $i$ 's wealth share is  $n^i \sigma^i$ , individual demands are a time-invariant fraction of their aggregate counterparts:

$$(12) \quad c_s^i = \frac{\alpha^i \sigma^i}{E_{\alpha\sigma}} c_s, \quad \text{where} \quad \beta_s c_s = E_{\alpha\sigma} (1 - \rho) \rho^{s-1} E_z$$

$$(13) \quad a^i l_s^i = \frac{(1 - \alpha^i) \sigma^i}{E_{(1-\alpha)\sigma}} l_s, \quad \text{where} \quad \beta_s (1 - t_{ws}) w_s l_s = E_{(1-\alpha)\sigma} (1 - \rho) \rho^{s-1} E_z$$

Substituting these relations into the household optimality conditions and imposing factor market equilibrium yields the desired aggregate optimality conditions

$$(14) \quad \frac{1}{\varepsilon} \frac{c_s}{1 - h_s} = (1 - t_{ws}) \frac{(1 - \theta) y_s}{h_s}$$

$$(15) \quad \frac{1}{\rho} \frac{c_{s+1}}{c_s} = (1 - t_{rs+1}) \theta \frac{y_{s+1}}{k_s}$$

where heterogeneous tastes and wealth shares combine to form the ‘‘distribution term’’

$$(16) \quad \varepsilon \equiv \frac{E_{\alpha\sigma}}{E_{(1-\alpha)\sigma}} = \frac{E_\alpha + S_{\alpha\sigma}}{1 - (E_\alpha + S_{\alpha\sigma})}$$

The key component here is the covariance of tastes for consumption and wealth shares,  $S_{\alpha\sigma}$ , where an interior solution to (14) implies  $S_{\alpha\sigma} \in (-E_\alpha, 1 - E_\alpha)$  for any  $E_\alpha \in (0, 1)$ .

The distribution term represents the critical difference between this model and a representa-

tive agent model. In the latter case, where tastes or wealth shares are identical,  $S_{\alpha\sigma} = 0$  and  $\varepsilon$  is constant. Otherwise,  $\varepsilon$  depends on the wealth distribution through  $S_{\alpha\sigma}$ . Any change to the wealth distribution that raises  $S_{\alpha\sigma}$  and  $\varepsilon$  will tend to raise the aggregate consumption-leisure ratio and lower wages in all periods because  $\varepsilon$  enters (14) without a time subscript. Thus, in contrast to the representative agent case, the wealth distribution affects the path of the economy. Intuitively,  $S_{\alpha\sigma}$  rises with redistribution toward consumption lovers or agents with higher marginal propensities to consume. If  $S_{\alpha\sigma} > 0$ , such redistribution favors rich consumption lovers by raising their wealth share, but if  $S_{\alpha\sigma} < 0$  the wealth share of poor consumption lovers rises. In either case, the consumption of the high- $\alpha^i$  winners rises by more than the consumption of the losers falls and the winners' leisure rises by less than the losers' leisure falls, thus raising the aggregate consumption-leisure ratio and labor.

Returning to the solution of the model, note that, together with equations (1) and (10), equations (14) and (15) completely describe the evolution of aggregate labor, consumption, investment and output for any  $\varepsilon$ . To solve the system, find a division of output into consumption, investment, and government spending that fulfills the aggregate optimality conditions. If  $1 - \chi_s$  denotes consumption's share of output less government spending (or net output for short), so that

$$(17) \quad c_s \equiv (1 - \chi_s)(1 - \gamma_s)y_s$$

then goods market clearing implies

$$(18) \quad k_s = \chi_s(1 - \gamma_s)y_s$$

Inserting these expressions into the aggregate optimality conditions yields solutions for the consumption share of output and labor:

$$(19) \quad \frac{1}{1 - \chi_s} = 1 + (\theta\rho) \frac{1 - t_{rs+1}}{1 - \gamma_{s+1}} \frac{1}{1 - \chi_{s+1}} = 1 + \sum_{v \geq 1} (\theta\rho)^v \frac{1 - t_{rs+v}}{1 - \gamma_{s+v}}$$

$$(20) \quad h_s = \frac{1}{1 + \lambda_s}, \quad \lambda_s \equiv \frac{(1 - \chi_s)(1 - \gamma_s)}{(1 - t_{ws})(1 - \theta)\varepsilon}$$

Thus, the aggregate leisure-to-labor ratio,  $\lambda_s$ , rises with current labor taxes and the current consumption share but falls with  $\varepsilon$ . In contrast, the consumption share of current net output is forward-look-

ing, depending negatively on future capital taxes.

The last two equations can be used to recursively compute the equilibrium paths for output, prices, and other aggregate variables given  $\varepsilon$ . In particular, equilibrium capital can be derived by using the production function with the equilibrium expressions for labor and the capital share and then iterating backwards to the initial fixed capital stock:

$$(21) \quad k_s = \chi_s(1 - \gamma_s) h_s^{1-\theta} k_{s-1}^\theta = \prod_{v=1}^{s-1} \left( \frac{\chi_{s-v}(1 - \gamma_{s-v})}{(1 + \lambda_{s-v})^{1-\theta}} \right)^{\theta^v} k_0^{\theta^s}$$

This equation, together with (19), is used to compute a solution for output using (18) and consumption using (17). Finally, once equilibrium output and inputs have been determined, the firm's optimality conditions can be used to compute equilibrium factor prices. At this point in a representative agent model, equations (2) and (17) through (21) would completely characterize the equilibrium. However, with heterogeneous agents  $\varepsilon$  is endogenous with influence on aggregate quantities and prices. Thus, because  $\varepsilon$  depends on  $S_{\alpha\sigma}$ , a solution must characterize the wealth distribution.

#### *Solution for Distribution Term and Wealth Distribution*

To characterize  $\varepsilon$  one must solve for the components of  $S_{\alpha\sigma}$ . Once a solution has been found, it is a small matter to also characterize the wealth distribution, especially the variance of wealth shares,  $S_{\sigma\sigma}$ . This inequality measure is similar to  $S_{\alpha\sigma}$  as demonstrated by:

$$(22) \quad S_{x\sigma} \equiv \sum_i n^i (x^i - E_x)(\sigma^i - 1) = \sum_i n^i (x^i - E_x) \frac{z^i - E_z}{E_z} = \frac{S_{xz}}{E_z}, \quad x = \alpha, \sigma$$

There are two parts to (22): deviations from average wealth and average wealth. Purely differential policies such as a mean-preserving transfer will raise the covariance if transfer recipients have high  $\alpha^i$  and lower the variance if recipients are poor. Alternatively, because equal wealth changes affect the poor disproportionately, an increase in average wealth indicates an increase of the wealth share of the poor that lowers the variance but raises the covariance when the poor are consumption-lovers.

To find average wealth, individual wealth is aggregated and transfers from the government's

intertemporal revenue constraint are substituted. These steps yield

$$E_z = k_0 + \sum_{s \geq 1} \pi_s \{(1-t_{ws})(1-\theta)\lambda_s + (\tau_s - \gamma_s)\}$$

where in equilibrium excess distortionary tax revenues are <sup>8</sup>

$$(23) \quad R_0 \equiv \sum_{s \geq 1} \pi_s \{\tau_s - \gamma_s\} = -k_0 + \sum_{s \geq 1} \pi_s \{-(1-t_{ws})(1-\theta) + (1-\chi_s)(1-\gamma_s)\}$$

Combining these two expressions and using (20), aggregate wealth can be simply expressed by

$$(24) \quad E_z = \sum_{s \geq 1} \pi_s (1-t_{ws})(1-\theta)\lambda_s (1+\varepsilon)$$

$$(25) \quad \pi_s = \beta_s y_s = \frac{k_0}{\chi_s (1-\gamma_s)} \prod_{u=1}^s \frac{\chi_u (1-\gamma_u)}{\theta (1-t_{ru})}$$

where the present value of output,  $\pi_s$ , combines (2), (5), and (18). Thus, aggregate wealth depends on the path of the discounted after-tax value of leisure,  $\pi_s (1-t_{ws})(1-\theta)\lambda_s = \beta_s (1-t_{ws})w_s l_s$ , and also on  $\varepsilon$  through changes in the time path of wages and human wealth. Both factor taxes lower discounted after-tax wages and raise the leisure-labor ratio with a net effect on average wealth that is positive for capital taxes and zero for labor taxes. Intuitively, budget-balancing lump-sum transfers of excess distortionary revenues from labor taxes exactly offset wealth losses from after-tax wages, while for capital taxes such transfers exceed discounted wage reductions.

Next, deviations from average wealth yield a simple covariance formula for  $x = \alpha, \sigma$ :

$$(26) \quad S_{xz} = S_{x(k+b)} + S_{xz_a} = S_{x(k+b)} + \sum_{s \geq 1} \pi_s (1-t_{ws})(1-\theta)(1+\lambda_s) S_{xa}$$

Differential wealth changes occur through the human-wealth component  $S_{xz_a}$  via the discounted after-tax wage, or  $\pi_s (1-t_{ws})(1-\theta)(1+\lambda_s) = \beta_s (1-t_{ws})w_s$ . If discounted wages and human wealth fall, individuals with high market abilities lose proportionately more than those with low productivities. Such a wealth-equalizing differential effect raises  $S_{\alpha z}$  if the rich also have lower propensities to consume or higher home productivity, or  $S_{\alpha a} < 0$ . However, a priori  $S_{\alpha a} > 0$  is also possible, in which case  $S_{\alpha z}$  falls. Also, because productive agents are hurt more by human wealth reductions,

---

<sup>8</sup> To derive (23), note  $-k_0 = \sum_{s \geq 1} \pi_s \{\chi_s (1-\gamma_s) - (1-t_{rs})\theta\}$ . While time-aggregating (13) directly implies (24), equation (23) is referenced because it will be used later to derive the marginal welfare costs of taxes.

there is a decline in human and overall wealth inequality when  $S_{\sigma_a} > 0$ .<sup>9</sup>

While no single inequality measure is completely satisfactory, the variance of wealth shares has other virtues apart from its similarity with  $S_{\alpha\sigma}$ . The variance equals the coefficient of variation squared and satisfies the Pigou-Dalton strong principle of transfers, just like the Herfindahl index,  $E_{\sigma\sigma} (= S_{\sigma\sigma} + 1)$ , to which it is cardinally equivalent (Cowell, 1995). While the measure has other desirable theoretical properties (Bourguignon, 1979, and Shorrocks, 1980), it is also very useful because it can be easily computed and does not require knowledge of specific wealth shares as would other inequality measures. On the downside, the variance is sensitive to the population shares and puts a disproportionate weight on rich groups (Cowell, 1980). To get around these problems, one can relate the variance to the familiar cumulative wealth share of the top or bottom  $x$ -percentile. If the population is split into two groups, where  $i = R, P$  indexes rich and poor so that  $\sigma^R > 1 > \sigma^P$ ,  $n^R \sigma^R + (1 - n^R) \sigma^P = 1$  and  $n^R + n^P = 1$ , one can show  $\sqrt{n^R(1 - n^R)S_{\sigma\sigma}} = n^R \sigma^R - n^R = n^P - n^P \sigma^P$ .<sup>10</sup>

Relating the variance to particular wealth shares will be very useful in describing social welfare, which is defined as the sum of individual utilities. Although alternative formulations exist with different degrees of inequality aversion, this welfare criterion is standard in the literature and also very tractable. Social welfare can be depicted using (12) and (13) and writing individual demands as a function of  $\varepsilon$ , or  $c_s^i = \alpha^i \sigma^i (1 + \varepsilon^{-1}) c_s$ , and  $l_s^i = (1 - \alpha^i) \sigma^i (1 + \varepsilon) l_s$ , where wealth shares could be found using (6) and (27). Substituting these relations into (3) and then aggregating and discarding nuisance parameters yields an expression for utilitarian social welfare in equilibrium:

$$(27) \quad U = \frac{\rho}{1 - \rho} \sum_i n^i \ln \sigma^i + \sum_{s \geq 1} \rho^{s-1} \left[ E_\alpha \ln(c_s(1 + \varepsilon^{-1})) + E_{1-\alpha} \ln(l_s(1 + \varepsilon)) \right]$$

Thus, social welfare depends on the entropy measure of wealth inequality,  $\sum_i n^i \ln \sigma^i$ , aggregate

---

<sup>9</sup> Note that although  $S_{\sigma\sigma} \geq 0$ ,  $E_z S_{\sigma(k+b)} = S_{(k+b)^2} + S_{(k+b)z_a}$  and  $E_z S_{\sigma z_a} = S_{z_a^2} + S_{(k+b)z_a}$ . Thus,  $S_{\sigma(k+b)}$  and  $S_{\sigma z_a}$  are guaranteed to be positive if  $S_{(k+b)a} \geq 0$ . However, when  $S_{(k+b)a} < 0$ , the relative size of  $S_{(k+b)^2}$  and  $S_{z_a^2}$  matters in determining which (if any) of the two covariances,  $S_{\sigma(k+b)}$  or  $S_{\sigma z_a}$ , is negative. To simplify the analysis it is assumed that  $S_{\sigma(k+b)} \geq 0$  and  $S_{\sigma z_a} \geq 0$  or equivalently that  $S_{(k+b)a}$  cannot be too negative.

<sup>10</sup> One can have more than two groups and still retain the property that wealth shares are a function of  $S_{\sigma\sigma}$ . Adding groups requires successively more information or assumptions on shares without necessarily adding insight.

consumption and leisure, and the distribution term. The inequality measure isolates the effect on welfare of wealth share changes, while the other terms capture  $\sigma^i$ -constant changes in individual leisure and consumption. For the case of two groups above, one can relate the entropy measure to  $S_{\sigma\sigma}$ , which will come in handy when analyzing marginal welfare costs.

The results of this section can be summarized with the following proposition:

**Theorem 1.** *(Existence) There exists a perfect foresight equilibrium with heterogeneous agents that has a reduced-form solution with recursive structure.*

**Proof:** Given  $\varepsilon$ , equilibrium  $\chi_s$  and  $h_s$  are determined by (19) and (20). Then equilibrium  $k_s$  is determined by (21), equilibrium  $c_s$  by (17), and  $y_s$  by (18), and equilibrium  $w_s$  and  $r_s$  are determined subsequently by (2). Finally, equilibrium  $E_z$ ,  $S_{xz}$ , and  $S_{x\sigma}$  for  $x = \alpha, \sigma$  are determined by (24), (26), and (22), with  $\varepsilon$  determined subsequently by (16). When wealth shares can be expressed as a function of  $S_{\sigma\sigma}$ , equilibrium  $U$  is determined by (20) and (22) or by a solution for individual wealth shares that is found by evaluating (6) and (27) in equilibrium. Then individual outcomes can easily be derived using (12) and (13).

To distinguish between models where  $\varepsilon$  is exogenous, let a superscript *RA* denote variables evaluated under  $S_{\alpha\sigma} = 0$  so that, for instance,  $\varepsilon^{RA} \equiv E_\alpha / E_{1-\alpha}$ . Then

**Corollary.** *No Heterogeneity ( $S_{\alpha\sigma} = S_{\sigma\sigma} = 0$ ) implies  $\varepsilon = \varepsilon^{RA}$  and  $U$  is independent of inequality. Weak Heterogeneity ( $S_{\alpha\sigma} = 0$  and  $S_{\sigma\sigma} \neq 0$ ) implies only  $\varepsilon = \varepsilon^{RA}$ .*

Thus, the irrelevance of the wealth distribution for aggregate variables is due to identical tastes or wealth shares. However, wealth inequality alters social welfare once wealth shares are free to vary. Thus, any analysis of welfare costs without inequality either assumes no heterogeneity or implicitly

assumes that non-distortionary transfers are pegging wealth shares at preexisting levels.

### III. Transition Effects of Factor Taxation

To derive the dynamic response of the system to factor taxes, one totally differentiates and solves the equations describing the dynamic equilibrium in Theorem 1. It is convenient to assume that all policy shocks occur over a time interval,  $[S, T-1]$ , where  $1 \leq S \leq T-1$  and  $T \leq \infty$ . This assumption provides building blocks that could be combined in a variety of ways to form more complicated policy sequences. However, only simple policy experiments will be discussed here. Also, define an impulse function  $I_s$  that is unity over  $[S, T-1]$  and zero otherwise, and let  $dx_s = I_s dx$  for all policy variables  $x$ , let  $\hat{x}_s \equiv dx_s/x$  for all variables, and let  $x_s^{RA}$  denote a variable that has been evaluated under the assumption  $S_{\alpha\sigma} = 0$ . As will be seen shortly, the response of the aggregate system to changes in taxes separates into a distribution effect, depending purely on  $\hat{\epsilon}$  that results when  $S_{\alpha\sigma} \neq 0$  from differential wealth effects that redistribute to or from consumption-lovers, and on a representative agent effect that occurs under weak or no heterogeneity:<sup>11</sup>

$$(28) \quad \hat{\chi}_s = \hat{\chi}_s^{RA}, \quad \hat{\chi}_s^{RA} = -X_s \frac{dt_r}{1-t_r}$$

$$(29) \quad \hat{\lambda}_s = -\frac{\hat{h}_s}{l} = \frac{\hat{l}_s}{h} = \hat{\lambda}_s^{RA} - \hat{\epsilon}, \quad \hat{\lambda}_s^{RA} = I_s \frac{dt_w}{1-t_w} + \frac{\chi X_s}{1-\chi} \frac{dt_r}{1-t_r}$$

$$(30) \quad \hat{k}_s = \hat{k}_s^{RA} + l(1-\theta^s)\hat{\epsilon}, \quad \hat{k}_s^{RA} = -l(1-\theta)K_s^I \frac{dt_w}{1-t_w} - \frac{1-(1-l(1-\theta))\chi}{1-\chi} K_s^\chi \frac{dt_r}{1-t_r}$$

$$(31) \quad \hat{y}_s = \hat{y}_s^{RA} + l(1-\theta^s)\hat{\epsilon}, \quad \hat{y}_s^{RA} = \hat{k}_s^{RA} - \hat{\chi}_s^{RA} = \theta \hat{k}_{s-1}^{RA} + (1-\theta)\hat{h}_s^{RA}$$

$$(32) \quad \hat{c}_s = \hat{c}_s^{RA} + l(1-\theta^s)\hat{\epsilon}, \quad \hat{c}_s^{RA} = \hat{k}_s^{RA} - \frac{1}{1-\chi} \hat{\chi}_s^{RA}$$

$$(33) \quad \hat{w}_s = \hat{w}_s^{RA} - l\theta^s \hat{\epsilon}, \quad \hat{w}_s^{RA} = \hat{y}_s^{RA} - \hat{h}_s^{RA}$$

where  $X_s \equiv (1-\chi)\sum_{v \geq 1} \chi^{v-1} I_{s+v} \geq 0$ ,  $K_s^I \equiv \sum_{v=0}^{s-1} \theta^v I_{s-v} \geq 0$ , and  $K_s^\chi \equiv \sum_{v=0}^{s-1} \theta^v X_{s-v} \geq 0$  are discounted sums of the impulses.<sup>12</sup> Variables without time subscripts denote steady-state values, such as the

<sup>11</sup> An appendix provides derivations and formulas for the variables that follow and equations (28) through (39).

<sup>12</sup> For labor taxes,  $I_s > 0 \forall s \in [S, T-1]$  and  $K_s^I = 0 \forall s < S$  and  $K_s^I > 0 \forall s \geq S$ , except if  $T < \infty$ , then  $K_\infty^I = 0$ . For

investment share  $\chi = \theta\rho(1-t_r)/(1-\gamma)$  or the leisure-labor ratio  $\lambda = l/h = (1-\chi)(1-\gamma)/[(1-t_w)(1-\theta)\epsilon]$ .

Representative agent effects in equations (28)-(33) are the standard dynamic substitution effects of factor taxes that occur under weak or no heterogeneity when  $S_{\alpha\sigma} = 0 = \hat{\epsilon}$ . Briefly, representative agent effects arise because labor taxes reduce the return to labor and the relative price of leisure during  $[S, T-1]$ , while capital taxes reduce the return to capital during  $[S, T-1]$  and raise the relative price of consumption and leisure in the interval  $[1, T-1]$ . Thus, labor taxes cause households to substitute toward leisure during  $[S, T-1]$  with savings reductions to smooth the impact on consumption, and capital taxes cause households to lower consumption and savings. Generally, factor taxes have a negative representative agent effect on aggregate labor and capital and, thus, also on output during  $[S, T-1]$ . Where the factor taxes primarily differ is with regard to their anticipation effects as captured by changes of the savings-consumption ratio, or  $\hat{\chi}_s$ , and with regard to the endogeneity of the initial tax base, which for capital taxes is fixed. Labor taxes do not cause a consumption-savings tradeoff, nor do they have anticipation effects; thus, all aggregate variables are unchanged during the anticipation phase and consumption declines thereafter.<sup>13</sup> By contrast, capital taxes have negative anticipation effects on aggregate variables and cause consumption to rise relative to savings prior to  $T$ . This latter effect may dominate in (32) and raise consumption over  $[S, T-1]$  if the interval is short, but as  $T-S$  grows the intertemporal substitution effect tends to dominate and agents will tend to reduce consumption. Finally, the representative agent effect on wages or the capital-labor ratio in (33) is tied down by the initial fixity of the capital stock and the terminal fixity of the after tax rate of return. Specifically, because the initial capital stock is fixed and initial labor falls, factor taxes will raise initial wages unless labor taxes are expected or capital taxes unexpected and short-term, in which case there is no effect. But over time the capital-labor ratio moves toward its long-run steady-state level, which is only affected (negatively) by permanent capital taxes.

capital taxes  $X_s > 0 \forall s < T-1$ ,  $X_s = 0 \forall s \geq T-1$ . Also,  $K_s^\chi > 0 \forall s$  except if  $S = 1 = T-1$ , then  $K_s^\chi = 0 \forall s$ , or if  $T < \infty$ , then  $K_\infty^\chi = 0$ . The exception where  $T < \infty$  and  $s \rightarrow \infty$  is that of a temporary tax that is neutral in the long run, reflecting the tendency of infinite horizon models to return to their original state. Also, because an unexpected one-period rise in capital taxes is a lump-sum tax, this policy is neutral in models characterized by Ricardian neutrality.

<sup>13</sup> As discussed by King, Plosser, and Rebelo (1988a), this property is a result of complete depreciation and logarithmic utility, which together imply that anticipation effects have offsetting income and substitution effects.

Several aspects of the distribution effect are worth noting before proceeding to the characterization of  $\hat{\epsilon}$ . First, adding the distribution effect of factor taxes to the representative agent effect creates the potential for non-standard dynamics, because depending on the sign of  $\hat{\epsilon}$  such redistribution can be reinforcing or offsetting. Specifically, a positive distribution effect,  $\hat{\epsilon} > 0$ , will raise aggregate quantities in all periods, thereby offsetting the representative agent effect on aggregate inputs and production. A positive distribution effect also offsets the representative agent effect of a labor tax on consumption. However, it can reinforce or offset the representative agent effect of a capital tax on consumption depending on whether  $[S, T-1]$  is short or long.<sup>14</sup> Second, the distribution effect tends to be stronger compared with the representative agent effect outside  $[S, T-1]$ . For example, consider the long run effects of temporary shocks. As  $s \rightarrow \infty$ ,  $\hat{k}_\infty = \hat{c}_\infty = \hat{h}_\infty = \hat{y}_\infty = (1-h)\hat{\epsilon}$  in (29) to (32), but the capital-labor ratio reverts to its original state. When  $\hat{\epsilon} = 0$ , stability conditions dictate that temporary shocks are neutral in the long-run as in Judd (1985). However, when  $\hat{\epsilon} \neq 0$ , temporary factor tax shocks have long-run effects on aggregate quantities because temporary policies have permanent effects on wealth and with heterogeneous tastes differential wealth effects do not cancel in the aggregate. Also, distribution effects can dominate during the anticipation phase where representative agent effects either do not exist for labor taxes or still are small for capital taxes. Third, while the distribution effect on labor is time invariant, it grows over time for other aggregate quantities. Thus, strong heterogeneity introduces a degree of persistence and permanence to the dynamics of aggregate quantities not seen in representative agent models but found in real-world data (Cogley and Nason, 1995). Thus, the model illustrates Stoker's (1986) thesis that distribution effects explain serial correlation in aggregate data.

Next, to compute the distribution effect  $\hat{\epsilon} = (E_{\alpha\sigma}E_{(1-\alpha)\sigma})^{-1}dS_{\alpha\sigma}$  (as well as the effect on the variance of wealth shares  $dS_{\sigma\sigma}$ ), one must find a solution for the differential of (22), or

$$dS_{x\sigma} = \frac{1}{E_z} \sum_i n^i (x^i - E_x)(dz^i - dE_z) - S_{x\sigma} \hat{E}_z \quad \text{for } x = \alpha, \sigma$$

which adds together a differential and an average wealth effect where the first reflects mean-preser-

---

<sup>14</sup> Capital taxes create distribution effects that differ for the various aggregate variables. Because  $\hat{\chi}_s^{RA} < 0$  yields  $|\hat{c}_s^{RA}| < |\hat{y}_s^{RA}| < |\hat{k}_s^{RA}|$ , capital taxes have stronger representative agent effects on capital than on consumption.

ving wealth changes and the second reflects equal absolute wealth changes that have disproportionate individual effects. The strength of the differential effect relative to the average wealth effect is determined by variation of the human wealth component. When human wealth does not vary, there is no differential effect and only an average wealth effect because human wealth changes equally for all. By contrast, without endowment variation only differential human wealth effects matter. Thus, if  $V_x \equiv S_{xz_a} / (E_z S_{x\sigma})$  is the share of  $S_{x\sigma}$  for  $x = \alpha, \sigma$  due to human wealth variation and  $1 - V_x$  is the share from endowment variation, then differential effects are zero if  $V_x = 0$  and average wealth effects are zero if  $V_x = 1$ , while for any  $V_x \in (0, 1)$  both effects appear.

The above equation is solved by differentiating (24) and (26). Straightforward but lengthy derivations yield a formula for the distribution effect:

$$(34) \quad \hat{\varepsilon} = -P_0 \Psi_w^\varepsilon \frac{dt_w}{1-t_w} - (1+\lambda^{-1}) \frac{X_0}{1-\chi} \Psi_r^\varepsilon \frac{dt_r}{1-t_r}$$

$$\text{where } X_0 = \chi^{S-1}(1-\chi^{T-S}), P_0 \equiv \rho^{S-1}(1-\rho^{T-S}), \Psi_w^\varepsilon \equiv \frac{V_\alpha}{\frac{1+\lambda^{-1}}{1+\varepsilon} \frac{E_\alpha}{S_{\alpha\sigma}} + V_\alpha}, \Psi_r^\varepsilon \equiv \frac{1-V_\alpha + h\phi_0 V_\alpha}{V_\alpha} \Psi_w^\varepsilon,$$

$$\text{and } \phi_0 \equiv \frac{\chi(1-\rho)}{\rho-\chi} \left( \frac{P_0}{X_0} \frac{1-\chi}{1-\rho} - 1 \right).$$

There are three notable parts to the coefficients in (34). First,  $P_0$  and  $X_0$  embody the effect on wealth of the timing of tax policies. Thus, less-anticipated (lower  $S$ ) or longer-lasting (larger  $T-S$ ) factor taxes have greater effects on wealth, implying a stronger distribution effect. Second, the strength of the differential effect relative to the average wealth effect depends on  $\phi_0 \geq 0$  and the timing of tax policies. The term is a weight on  $V_\alpha \in [0, 1]$  in the numerator of  $\Psi_r^\varepsilon$ , which identifies the differential effect; the average wealth effect  $1 - V_\alpha$  remains unweighted. The numerators of  $\Psi_w^\varepsilon$  and  $\Psi_r^\varepsilon$  show that the distribution effect for capital taxes balances differential and average wealth effects, but the distribution effect for labor taxes operates only through differential wealth effects (as explained previously for (24)). Thus, the distribution effect for labor taxes is strongest when  $V_\alpha = 1$ .

By contrast, whether the distribution effect for capital taxes is stronger when  $V_\alpha = 0$  or when  $V_\alpha = 1$  depends on conditions that are sensitive to the sign and size of  $E_\alpha/S_{\alpha\sigma}$  and on the timing of taxes through  $\varphi_0$ . For instance, less-anticipated or longer-lasting capital taxes weaken the differential effect relative to the average wealth effect.<sup>15</sup> Rather than characterize all possibilities here, the numerical examples below suggest that the issue may not be too serious.

Third, the crucial terms are the  $\Psi_i^\varepsilon$ , which depend on  $S_{\alpha\sigma}$  and  $V_\alpha$ . While the numerator of  $\Psi_i^\varepsilon$  is non-negative, the denominator rises with  $S_{\alpha\sigma}$  and takes on its sign. To see this last point, note that a given steady-state  $\lambda$  pegs  $\varepsilon$  but allows  $S_{\alpha\sigma}$  and  $E_\alpha$  to vary according to  $S_{\alpha\sigma} = \varepsilon(1+\varepsilon)^{-1} - E_\alpha$  in (16). Thus, for  $E_\alpha \in (0, 1)$ ,  $S_{\alpha\sigma} \in (\varepsilon(1+\varepsilon)^{-1}, -(1+\varepsilon)^{-1})$ . As  $S_{\alpha\sigma}$  rises over its range and  $E_\alpha$  falls, the denominator of  $\Psi_i^\varepsilon$  rises from a negative lower limit of  $V_\alpha - (1+\lambda^{-1})$  to a positive upper limit of  $V_\alpha$ . In other words, the  $-\Psi_i^\varepsilon$  terms rise with  $S_{\alpha\sigma}$  from a lower negative limit and to an upper positive limit, with both terms equal to zero when  $S_{\alpha\sigma} = 0$ , the representative agent case. Consequently, the  $\hat{\varepsilon}$  has a sign opposite that of  $S_{\alpha\sigma}$ , but  $\hat{\varepsilon} = 0$  if  $V_\alpha = 0$  for labor taxes or if  $V_\alpha = 1$  and  $\varphi_0 = 0$  for capital taxes. Furthermore,  $\hat{\varepsilon}$  grows in proportion to how much  $S_{\alpha\sigma}$  deviates from zero.

Intuitively, when  $S_{\alpha\sigma} < 0$ , factor taxes raise the wealth share of poor consumption lovers, thus causing a positive distribution effect on consumption and labor because the consumption of the high- $\alpha^i$  winners rises by more than the consumption of the losers falls and the winners' leisure rises by less than the losers' leisure falls. By contrast, when  $S_{\alpha\sigma} > 0$ , factor taxes redistribute toward poor leisure lovers which causes a negative distribution effect. Thus, because factor taxes tend to redistribute towards the poor, they will offset (reinforce) the representative agent effects on aggregate inputs and output when the poor have stronger tastes for consumption (leisure) or lower (higher) home productivities. When  $S_{\alpha\sigma}$  is sufficiently negative, positive distribution effects could dominate the representative effect on inputs or output, in which case factor taxes will have a positive net effect.

The next result summarizes the discussion of this section:

---

<sup>15</sup> Notice that  $\varphi_0 > 0$  for  $T \rightarrow \infty$  or  $S = T-1 > 1$ , where  $\varphi_0$  rises when  $T-S$  falls, or, given  $T-S$ , when  $S$  rises. Also,  $\varphi_0 = 0$  for  $S = T-1 = 1$  describes an unexpected temporary capital tax, which effectively is a lump-sum tax.

**Theorem 2.** (*Comparative Dynamics*) *Labor taxes have no distribution effect if  $S_{\alpha\sigma} = 0$  or  $V_{\alpha} = 0$ . Capital taxes have no distribution effect if  $S_{\alpha\sigma} = 0$  or  $V_{\alpha} = 1$  and  $\phi_0 = 0$ . Otherwise, factor taxes yield a positive (negative) distribution effect, or  $\hat{\epsilon} > (<) 0$ , if  $S_{\alpha\sigma} < (>) 0$  and  $V_{\alpha} \in [0, 1]$ , whereby  $|\hat{\epsilon}|$  increases with  $|S_{\alpha\sigma}|$ . A positive (negative) distribution effect offsets (reinforces) the negative representative agent effect on aggregate inputs and output during  $[S, T-1]$ . For labor taxes, a positive (negative) distribution effect is also offsetting (reinforcing) for consumption during  $[S, T-1]$ , but for capital taxes this is true only if  $T-S$  is sufficiently large.*

Positive distribution effects and  $S_{\alpha\sigma} < 0$  are consistent with some evidence. First, the result that positive distribution effects cause total tax effects on aggregate quantities to be smaller than in representative agent models agrees with recent evidence that the aggregate effects of tax policies are small (Slemrod, 1990, and Glick and Hutchinson, 1990) or insignificant (McGrattan, 1994). Second,  $S_{\alpha\sigma} < 0$  is consistent with some microeconomic evidence. In particular, the fact that consumption varies with wealth, combined with the observation that the dispersion of consumption is less than that of wealth (from a comparison of Cutler and Katz, 1992, and Wolff and Marley, 1989) or that the marginal propensity to consume varies inversely with wealth (Menchik and David, 1983, and Diamond and Hausman, 1984), implies  $S_{\alpha\sigma} < 0$  by (12).<sup>16</sup> Also, the observation that leisure does not vary much across wealth classes or that the dispersion of leisure is less than or equal to that of wealth, implies by (13) that  $S_{\alpha a} < 0$  or  $S_{\alpha z_a} < 0$ , which together with the preceding evidence supports the assumption that  $V_{\alpha} > 0$ . Thus, the evidence suggests that the rich are leisure lovers, have lower propensities to consume, and are more skilled at market and home production.<sup>17</sup>

---

<sup>16</sup> The latter two studies argue that savings propensities and wealth are positively correlated, which generally implies that consumption propensities and wealth are negatively correlated. The present model abstracts from heterogeneous savings propensities but allows heterogeneous consumption propensities.

<sup>17</sup> To see this most easily, divide the population into two groups where  $\sigma^R > \sigma^P$ . The micro evidence requires  $c^R/c^P < \sigma^R/\sigma^P$  and  $l^R/l^P \leq \sigma^R/\sigma^P$ . Thus, the strict inequality implies  $\alpha^R < \alpha^P$  and so  $S_{\alpha\sigma} < 0$ . Together with  $\alpha^R < \alpha^P$ , the weak inequality implies  $a^R > a^P$  or  $S_{\alpha a} < 0$ , which by definition means  $V_{\alpha} > 0$ . Because  $S_{\alpha\sigma} < 0$ , the rich cannot have low home productivity. Then, because  $V_{\alpha} > 0$ , the rich cannot also have low market skills.

#### IV. Wealth Distribution Effects of Factor Taxation

The effects of factor taxes on the first two moments of the wealth distribution are computed by differentiating equations (22), (24), and (26) and then substituting for  $\hat{\varepsilon}$  from (34). Thus,

$$(35) \quad \hat{E}_z = \frac{X_0}{1-\chi} \frac{dt_r}{1-t_r} - E_{(1-\alpha)\sigma} \hat{\varepsilon}$$

$$(36) \quad \frac{\hat{S}_{\sigma\sigma}}{2} = -P_0 l \Psi_w^{S_{\sigma\sigma}} \frac{dt_w}{1-t_w} - \frac{X_0}{1-\chi} \Psi_r^{S_{\sigma\sigma}} \frac{dt_r}{1-t_r}$$

where  $\Psi_w^{S_{\sigma\sigma}} = V_\sigma + \left( \frac{1+\lambda^{-1}}{1+\varepsilon} - V_\sigma \right) \Psi_w^\varepsilon$  and  $\Psi_r^{S_{\sigma\sigma}} = 1 - V_\sigma + h\varphi_0 V_\sigma + \left( \frac{1+\lambda^{-1}}{1+\varepsilon} - V_\sigma \right) \Psi_r^\varepsilon$ . Under the assumption of two groups with  $\sigma^R > 1 > \sigma^P$  where wealth shares can be described as a function of  $S_{\sigma\sigma}$ , one can relate changes in the variance to changes in cumulative wealth shares and to changes in the entropy measure of wealth inequality:

$$(37) \quad \frac{\hat{S}_{\sigma\sigma}}{2} = \left( 1 - \frac{1}{\sigma^R} \right)^{-1} \hat{\sigma}^R = - \left( \frac{1}{\sigma^P} - 1 \right)^{-1} \hat{\sigma}^P = \left( 1 - \frac{1}{\sigma^R} \right)^{-1} \left( \frac{1}{\sigma^P} - 1 \right)^{-1} \sum_i n^i \hat{\sigma}^i$$

This formulation leads to simple interpretations and, perhaps more importantly, will be used in deriving and ultimately quantifying the marginal welfare costs of factor taxes.

Factor tax effects on average wealth separate into a direct effect and a distribution effect which are, respectively, the first and second term in (35). The direct effect occurs even without heterogeneity and is zero for labor taxes and positive for capital taxes (more so if  $T-S$  rises or  $S$  falls). Under strong heterogeneity, a positive distribution effect reduces  $E_z$  because redistribution toward consumption lovers or agents with low home productivity tends to reduce leisure and wages and human wealth. Thus, while labor tax effects on average wealth depend solely on the sign and magnitude of  $S_{\sigma\sigma}$ , capital taxes increase average wealth unless  $S_{\sigma\sigma}$  is sufficiently negative.

Similarly, factor tax effects on  $S_{\sigma\sigma}$  depend on how much heterogeneity is allowed in (36). When there is only weak heterogeneity so that  $\Psi_i^\varepsilon = 0$ , factor taxes reduce wealth inequality. However, there are no inequality effects when there is no human wealth variation (or  $V_\sigma = 0$ ) for labor taxes or else when there is no endowment variation and capital taxes are effectively lump-sum (or

$V_\sigma = 1$  and  $\varphi_0 = 0$ ). With strong heterogeneity or  $S_{\alpha\sigma} \neq 0$ , one must consider how the  $V_x$  vary for  $x = \alpha, \sigma$ . A reasonable assumption is that  $V_\alpha$  rises with  $V_\sigma$  over the unit interval with endpoints of  $V_\sigma = 0 = V_\alpha$  and  $V_\sigma = 1 = V_\alpha$ . Thus, factor taxes reduce wealth inequality for  $V_\sigma, V_\alpha \in [0, 1]$ , except if  $V_\sigma = 0 = V_\alpha$  for labor taxes. In other words,  $\Psi_i^{S_{\alpha\sigma}} > 0$  which can be verified by looking at the limits of  $V_x$  (because all other possibilities are linear combinations of them) and letting  $S_{\alpha\sigma}$  vary. As  $S_{\alpha\sigma}$  rises from its lower limit to its upper limit, inequality effects shrink but remain negative.

For intuition, consider first the case of weak heterogeneity. Labor taxes have no average wealth effects because budget-balancing lump-sum transfers perfectly offset average human wealth changes. However, labor taxes have a wealth-equalizing differential effect by reducing human wealth that hurts productive agents proportionately more. By contrast, capital taxes raise average wealth because lump-sum tax rebates more than offset human wealth losses. Wealth inequality falls, because the poor -- for whom equal-sized rebates mean a disproportionately larger gain -- benefit more. Thus, both the average wealth effect and the differential wealth effect of capital taxes reduce wealth share inequality. Under strong heterogeneity, distribution effects on human wealth must be added to the previous effects. A positive distribution effect lowers leisure, wages, and human wealth, leading to lower average wealth and an inequality-increasing differential effect. Thus, a positive distribution effect lowers average wealth for labor taxes, but for capital taxes it raises average wealth by less than under weak heterogeneity. Also, wealth equalization is weakened. By contrast, when  $S_{\alpha\sigma} > 0$  positive average wealth and wealth equality effects are reinforced.

The discussion of this section is summarized by:

**Theorem 3.** *(Wealth Distribution Effects) Assume  $V_x \in [0, 1]$ ,  $x = \alpha, \sigma$  with  $V_x \neq 0$  for labor taxes and  $\varphi_0 > 0$  for capital taxes. Then, under weak heterogeneity, average wealth is unchanged for labor taxes and rises with capital taxes while wealth inequality falls for labor and capital taxes. Under strong heterogeneity, positive (negative) distribution effects cause average wealth to fall (rise) for labor taxes, strengthen (weaken) the positive average wealth effects for capital taxes, and*

*strengthen (weaken) the positive wealth equality effects for labor and capital taxes.*

The result that factor taxes lower wealth inequality is robust to the full range of distribution effects. However, the effects of factor taxes on aggregate quantities and wealth are much more sensitive to the size and sign of the distribution effects. Thus, an inequality-output tradeoff exists for  $[S, T-1]$  under weak heterogeneity for capital taxes, something that is not guaranteed otherwise. Also, the results on the wealth distribution effects of factor taxes are partially consistent with the data. Wolff and Marley (1989) show that aggregate wealth rose in the United States from the late 1940s to the middle 1980s, whereas wealth inequality tended to fall until the middle 1970s and then rose. During the same time, aggregate marginal labor tax rates trended up according to McGrattan (1994) while capital tax rates trended down. Thus, for  $S_{\alpha\sigma} < 0$ , labor taxes could explain a rise in aggregate wealth and wealth equality over several decades, particularly if -- as suggested by the simulations below -- labor taxes have greater wealth distribution effects than capital taxes.

## **V. Marginal Welfare Costs of Factor Taxation under Heterogeneity**

The marginal welfare cost of a tax is defined as the social welfare change of a distortionary tax per unit of lump-sum transfers financed by revenue changes. To find the amount of lump-sum transfers necessary to balance the budget after a change in factor tax rates, or  $dR_0$ , one differentiates excess distortionary revenues in (23). The change in social welfare,  $dU$ , is found by differentiating (27). Inserting the equilibrium responses for consumption and leisure from equations (29) and (32) yields the  $\sigma^i$ -constant or compensated welfare change, which can be further divided into an efficiency effect and a distribution effect.<sup>18</sup> Equity effects on social welfare are computed by relating entropy changes to changes in wealth inequality using (37). Welfare changes are transformed into wealth equivalents -- so that they are expressed in the same units as revenue changes -- by dividing  $dU$  by the steady-state marginal utility of consumption of the representative agent,  $u_c^{RA} = E_c / (1-\chi)(1-\gamma)$ .

---

<sup>18</sup> As noted by Guesnerie (1995) and Boadway (1994), strict separation between efficiency and equity, or in this case distribution effects, is unlikely in applied welfare analysis in a second-best world.

As noted by Judd (1987a) for a continuous time analog, the impact on social welfare of a tax change is equivalent to aggregate consumption rising by a constant equal to  $(1-\rho)dU/(\rho u_c^{RA})$ .

The resulting expressions for  $dR_0$  and  $dU/u_c^{RA}$  are combined with (34) and (36) to yield formulas for the marginal welfare cost,  $(dU/u_c^{RA})/dR_0$ . Then, following Judd (1987a), the formulas are transformed into  $MWC \equiv [(1-\rho)/\rho](dU/u_c^{RA})/dR_0$ , which is the per period (or flow) wealth equivalent of the change in welfare from using a factor tax rather than a lump-sum tax:

$$(38) \quad \frac{\rho}{(1-\rho)\lambda\varepsilon} MWC(t_w) = -U^w + \left(1 - \frac{1}{\sigma^R}\right) \left(\frac{1}{\sigma^P} - 1\right) \frac{l\Psi_w^{S_{\alpha\sigma}}}{E_\alpha} - (1+\lambda^{-1})U^\varepsilon\Psi_w^\varepsilon$$

$$(39) \quad \frac{\rho(\lambda\varepsilon - 1)}{(1-\rho)\lambda\varepsilon} MWC(t_r) = -\varphi_0 U^r + \left(1 - \frac{1}{\sigma^R}\right) \left(\frac{1}{\sigma^P} - 1\right) \frac{\Psi_r^{S_{\alpha\sigma}}}{E_\alpha} - (1+\lambda^{-1})U^\varepsilon\Psi_r^\varepsilon$$

where  $U^w \equiv \frac{h}{\varepsilon^{RA}} \left( \frac{\varepsilon^{RA}}{\varepsilon} \frac{\lambda\varepsilon}{\lambda^{NG}\varepsilon^{NG}} - 1 \right)$ ,  $U^r \equiv \frac{h}{\varepsilon^{RA}} \left( \frac{\varepsilon^{RA}}{\varepsilon} \frac{\lambda\varepsilon}{\lambda^{NG}\varepsilon^{NG}} \left( 1 + \frac{\chi^{NG} - \chi}{\chi(1-h)(1-\theta)} \right) - 1 \right)$  and

$U^\varepsilon \equiv \left( \frac{1}{\lambda^{NG}\varepsilon^{NG}} - \frac{1}{\lambda\varepsilon} \right) + \frac{S_{\alpha\sigma}}{E_\alpha} \left( 1 - \frac{1}{\lambda\varepsilon} \right)$ . Variables with superscript *RA* have been evaluated under

the weak and no-heterogeneity assumption of  $S_{\alpha\sigma} = 0$ . Alternatively, a superscript *NG* denotes a no-government regime where  $\gamma = t_w = t_r = 0$  in the initial steady state. The distinctions are useful for highlighting the independent effects of the taste distribution and of preexisting distortions on welfare costs, where preexisting distortions are assumed to satisfy  $\chi^{NG} \geq \chi$  and  $\lambda\varepsilon \geq \lambda^{NG}\varepsilon^{NG}$ .<sup>19</sup>

The interpretation of the formulas is that a negative term represents a cost to society while positive terms represent a gain. The *MWC* formulas are a sum of three parts. In the order in which they appear in (38) and (39), there is a representative agent or efficiency effect, which is also known as the marginal efficiency cost (or *MEC*), an inequality or equity effect, and finally a distribution effect. These terms are best understood by comparing the cases of no, weak, and strong heterogeneity, where all cases are evaluated without a government ( $\lambda\varepsilon = \lambda^{NG}\varepsilon^{NG}$ ) and with a government

<sup>19</sup> Note that  $\lambda\varepsilon$  is the ratio of the consumption share of net output to the after-tax wage share, which rises with preexisting tax rates. While  $\lambda\varepsilon > \lambda^{NG}\varepsilon^{NG}$  is a joint condition on initial tax rates and  $\gamma$ , the emphasis of this paper is on factor tax distortions, not the offset to distortions from  $\gamma$ . Also,  $\lambda^{NG}\varepsilon^{NG} > 1$  because  $\rho < 1$  and  $\theta < 1$ .

( $\lambda \varepsilon > \lambda^{NG} \varepsilon^{NG}$ ). The comparison also requires that the initial steady-state  $h$  and  $\chi$  are fixed to disentangle the effect of initial distortions and the initial taste distribution.

First, without any heterogeneity, or  $\varepsilon = \varepsilon^{RA}$  and  $\Psi_j^i = 0$ , factor taxes have a negative efficiency effect that rises with preexisting factor tax distortions, whereby taxes have no social welfare effect when there is no government ( $U^i = 0$ ). Second, under weak heterogeneity,  $\varepsilon = \varepsilon^{RA}$  and  $\Psi_j^\varepsilon = 0$  in (36), but ability and endowment variation imply  $\Psi_j^{S_{\alpha\sigma}} > 0$  under the assumptions of Theorem 3. In this case, factor taxes also promote wealth equality, which increases social welfare unless  $V_\sigma = 0$  for labor taxes. In other words, factor taxes cause positive equity effects on welfare that offset negative efficiency effects. Thus, the *MWC* is lower under weak heterogeneity than under no heterogeneity. However, because equity effects do not depend on initial distortions given  $h \varphi_0$ , efficiency costs increase in importance relative to equity gains as preexisting distortions rise.<sup>20</sup>

Third, under strong heterogeneity the distribution effect on welfare comes into play and the efficiency and equity effects are also altered because they depend on the taste distribution. The efficiency effect depends on the taste distribution through  $E_\alpha$ , which determines how strong consumption effects on welfare are relative to leisure. When  $E_\alpha$  is sufficiently large, factor taxes cause an efficiency loss because discounted consumption reductions outweigh discounted leisure increases, but an efficiency gain is possible with a sufficiently small  $E_\alpha$ . When  $h$  and  $\varepsilon$  are fixed,  $E_\alpha$  and  $S_{\alpha\sigma}$  vary according to  $S_{\alpha\sigma} = \varepsilon(1+\varepsilon)^{-1} - E_\alpha$  so that  $E_\alpha$  will be small (large) as  $S_{\alpha\sigma} > (<) 0$ . Thus, the efficiency coefficients  $U^w$  and  $U^r$  are smaller (larger) as  $S_{\alpha\sigma} > (<) 0$  and  $\varepsilon$  rises (falls) relative to  $\varepsilon^{RA}$ , thereby offsetting (enhancing) the efficiency cost under weak heterogeneity. Without a government, the efficiency effect on welfare is negative when  $S_{\alpha\sigma} < 0$  and positive when  $S_{\alpha\sigma} > 0$ . Also, higher initial distortions aggravate efficiency losses when  $S_{\alpha\sigma} < 0$  while for  $S_{\alpha\sigma} > 0$  efficiency effects are offset. As preexisting distortions increase, the response of consumption rises relative to that of

---

<sup>20</sup> Also, the coefficients on the left-hand side of (38) and (39) imply that initial tax distortions raise  $MWC(t_w)$  but weaken  $MWC(t_r)$ . The reason is that for labor taxes a higher initial level of discounted excess distortionary revenues,  $R_0$ , means that tax increases have a smaller revenue yield,  $dR_0$ . By contrast, initial distortions imply a higher revenue yield for capital taxes because the discount factor becomes more responsive to capital taxes when initial distortions are high, and this positive effect on  $dR_0$  overshadows the previous negative effect on  $dR_0$ .

leisure so that a negative efficiency effect is possible even when  $S_{\alpha\sigma}$  is positive but not too large.<sup>21</sup>

In contrast to the efficiency effect, the equity effect depends on the taste distribution via distribution effects on wages and human wealth. Theorem 2 shows that the distribution effect weakens (strengthens) the inequality effects when  $S_{\alpha\sigma} > (<) 0$ . Examination of  $\Psi_j^{S_{\alpha\sigma}}$  in (36) reveals that initial distortions weaken the equity effect when  $S_{\alpha\sigma} < 0$ . However, when  $S_{\alpha\sigma} > 0$  distortions have an indeterminate effect, either strengthening or weakening the equity gains. Thus, when  $S_{\alpha\sigma} < 0$ , initial distortions widen the difference between efficiency losses and equity gains. By contrast, when  $S_{\alpha\sigma} > 0$ , initial distortions offset efficiency gains, causing them to turn negative and have indeterminate effects on equity gains so that the net effect on welfare is an open question.<sup>22</sup>

Finally, when the distribution effect on welfare is considered, welfare losses under strong heterogeneity may, under some circumstances, be larger than under weak heterogeneity. It is clear from the coefficient  $U^\varepsilon$  that without preexisting distortions, positive and negative distribution effects reduce social welfare, but as preexisting distortions increase it becomes more likely that a positive distribution effect increases welfare. This surprising result can easily be explained. Because the distribution effect moves aggregate consumption and leisure in opposite directions, the net welfare effect depends on whether consumption or leisure dominates. The coefficient  $U^\varepsilon$  reveals that initial tastes and preexisting distortions determine the outcome. Consider first the case of no preexisting distortions. When  $S_{\alpha\sigma} < 0$ , agents see equi-proportional  $\sigma^i$ -compensated reductions in consumption and increases in leisure, or  $\hat{c}_s^i - \hat{\sigma}^i < 0 < \hat{l}_s^i - \hat{\sigma}^i$ , where wealth redistribution from rich ( $R$ ) to poor ( $P$ ) implies that uncompensated responses satisfy  $\hat{c}_s^P > 0 > \hat{c}_s^R$  and  $\hat{l}_s^P > 0 > \hat{l}_s^R$ . At the same time, when  $S_{\alpha\sigma} < 0$ ,  $E_\alpha$  is large, meaning that compensated consumption declines dominate leisure effects on the *MWC*. By contrast, when  $S_{\alpha\sigma} > 0$ , the above inequalities are reversed and  $E_\alpha$  is small, so that compensated leisure reductions prevail. Thus, positive and negative distribution effects aggravate the *MWC* with no initial tax distortions. Initial distortions increase the uncompen-

<sup>21</sup> As can be seen by comparing (29) and (32), the dynamic effect on consumption rises relative to leisure as  $l/h = \lambda$  or  $(\lambda\varepsilon/\lambda^{NG}\varepsilon^{NG})(\varepsilon^{RA}/\varepsilon)$  rises or as initial distortions increase.

<sup>22</sup> Note that the *MEC* of capital taxation tends to increase with  $\varphi_0$  or when  $S$  is larger and  $T-S$  smaller. The equity effect,  $\Psi_r^{S_{\alpha\sigma}}$ , rises with  $\varphi_0$  when  $S_{\alpha\sigma} = 0$ , but when  $S_{\alpha\sigma} < 0$ , the effect is weaker.

sated response of consumption relative to leisure, weakening the  $\sigma^i$ -compensated response of consumption and increasing the likelihood that the compensated leisure effect is dominant. Thus, with sufficiently large distortions so that the leisure effect does in fact dominate, *MWCs* will be aggravated by a negative distribution effect and weakened by a positive distribution effect.

This discussion on the marginal welfare costs of factor taxes can be summarized by

**Theorem 4.** (*Marginal Welfare Costs*) Assume initial distortions satisfy  $\chi^{NG} \geq \chi$  and  $\lambda \epsilon \geq \lambda^{NG} \epsilon^{NG}$ , fix the initial steady-state  $h$  and  $\chi$ , and divide the population into two groups according to their wealth. Then under the conditions of Theorem 3, factor taxes reduce social welfare under no heterogeneity through a negative efficiency effect that increases with preexisting tax distortions. Under weak heterogeneity, factor taxes have a positive equity effect that is independent of preexisting distortions and offsets the efficiency effect on welfare. Under strong heterogeneity, efficiency and equity effects are strengthened (offset) when  $S_{\alpha\sigma} < (>) 0$ , with efficiency gains a possibility when  $S_{\alpha\sigma} > 0$ . Distortions widen the difference between efficiency losses and equity gains when  $S_{\alpha\sigma} < 0$ . Also, a negative distribution effect on welfare must be added to efficiency and equity effects, but when preexisting distortions are sufficiently large and  $S_{\alpha\sigma} \ll 0$ , a positive distribution effect is possible. Increases of  $|S_{\alpha\sigma}|$  offset (enhance) the distribution effect when  $S_{\alpha\sigma} < (>) 0$ .

To put this section into perspective, it is clear that the *MWC* is lower under weak heterogeneity than under no heterogeneity, with the difference narrowing with initial distortions. How the strong heterogeneity scenario compares depends on the distribution of tastes and on initial distortions. A positive *MWC* is possible when  $S_{\alpha\sigma} > 0$  and initial distortions are small, something that seems counterintuitive and further strengthens the arguments against this specification of individual characteristics. Alternatively, when  $S_{\alpha\sigma} < 0$ , the difference between the *MWC* under strong and weak heterogeneity depends on whether the *MEC* rises more or less than the equity gain as  $S_{\alpha\sigma}$  falls. Also, the distribution effect aggravates the *MWC* when distortions are sufficiently small (and weak-

ens it otherwise). However, as  $S_{\alpha\sigma}$  becomes more negative, the effect of distortions on the distribution effect are offset. The problem is that, while the evidence favors  $S_{\alpha\sigma} < 0$  over  $S_{\alpha\sigma} \geq 0$ , no evidence exists on whether  $S_{\alpha\sigma}$  is small or large. Because it is unclear a priori which of these effects on the *MWC* dominates, a quantitative assessment of the *MWC* formulas is considered next.

## VI. Computational Experiments

Numerical examples are provided to shed light on how the *MWC* and its critical components (changes in leisure, consumption, and inequality) react to variations of  $S_{\alpha\sigma}$ ,  $V_\alpha$ , and  $V_\sigma$ . For simplicity, this section looks at the effects of permanent factor tax increases with  $S \geq 1$  and  $T \rightarrow \infty$ . Figures 1 and 2 graph the results of an initial experiment where  $S = 3$  for a given set of parameters and predetermined variables that are discussed below.<sup>23</sup> In particular, Figure 1 looks at the transition path of consumption and leisure following a 10 percent change in labor or capital taxes, and Figure 2 looks at the inequality effects (as measured by the percentage change of the wealth share of the poorest 90th percentile of the population) and the *MWC*. Table 1 summarizes this experiment and compares it with alternative experiments that vary  $S$ , initial tax rates, and other critical parameters.

### *Parameter Choices*

The parameters for the initial experiment are  $\rho = 0.99$ ,  $\theta = 0.4$ ,  $t_w = 0.25$ ,  $t_r = 0.5$ , and  $\gamma = 0.2$ , and predetermined variables are  $h = 0.2$ ,  $n^R = 0.1$ , and  $n^R \sigma^R = 0.7$ . Since most variables in the model are described as shares of output, the initial steady-state output is normalized to  $y_0 = 1$ . The parameters are close to point estimates by McGrattan (1994).<sup>24</sup> First, the production sector is defined by the capital share parameter, which has shown a wide dispersion in the literature. Numb-

---

<sup>23</sup> The timing of policies is a key issue in dynamic models. Although many alternatives exist, having  $S = 3$  and  $T \rightarrow \infty$  illustrates the idea that most taxes are anticipated because legislation takes time and are rarely legislated to be temporary. The choice also allows an evaluation of anticipation effects and perhaps also a fairer appraisal of the merits of wage and capital taxes given that unanticipated capital taxes are partly lump-sum.

<sup>24</sup> Specifically, the numbers have been rounded to simplify comparison with other research. The difference between the rounded numbers and McGrattan's point estimates is statistically insignificant. Also, the simulations suggest that the rounding differences are quantitatively insignificant.

ers range from 0.25 as in Judd (all) and Auerbach and Kotlikoff (1987) to 0.4 in Cooley and Prescott (1995) and 0.43 in King et al. (1988a). Specific choices depend on how income is assigned to capital and labor (Christiano, 1988), which depends partly on how government capital and household durables are counted (Cooley and Prescott, 1995). Next, the government sector is defined by spending and tax parameters. The spending share is chosen to approximate the postwar average in the United States and give a reasonable value for initial steady-state government debt. There is considerably more debate about the appropriate tax rates. Hansson (1985) and Judd (1987a) summarize a variety of studies and argue that estimates for aggregate marginal labor tax rates have ranged from 0.2 to 0.4 and estimates for capital tax rates from 0.3 to 0.5. McGrattan computes aggregate marginal factor tax rates for 1947 through 1988 and values close to the sample averages have been adopted.

Finally, the household sector is described by a few summary statistics of the joint distribution of agent characteristics and aggregate labor. First, a value of  $\rho = 0.99$  for the utility discount factor is often assumed in the literature (as in Judd, 1987a,b). The value defines the length of a period as a quarter of a year and implies an average annual after-tax real rate of return of 4.1 percent. However, choices can range from 0.987 (Cooley and Prescott, 1995) to McGrattan's estimate of 0.993. In the aggregate, households are also assumed to spend 20 percent of their substitutable time working in the original steady state. Calibrated examples of dynamic models consider a wide range of assumptions for  $h$ , such as 0.33 in Cooley and Prescott (1995) or 0.4 in Auerbach and Kotlikoff (1987). As Eichenbaum, Hansen, and Singleton (1988) explain, such choices largely depend on how the time endowment is computed and whether only employed households or all households are counted.<sup>25</sup> Once  $h$  is determined,  $\varepsilon$  is pegged and  $\varepsilon(1+\varepsilon)^{-1} = S_{\alpha\sigma} + E_{\alpha}$  determines how  $E_{\alpha}$  and  $S_{\alpha\sigma}$  vary. However, there is little evidence on tastes for consumption relative to leisure. Under the implicit assumption that  $S_{\alpha\sigma} = 0$ , Auerbach and Kotlikoff assume  $E_{\alpha} = 0.4$  to tie down aggregate labor and McGrattan estimates  $E_{\alpha} = 0.25$ . However,  $E_{\alpha}$  does not tie down  $h$  unless  $S_{\alpha\sigma}$  is identified and, as

---

<sup>25</sup> Strictly speaking,  $h$  is an efficiency-weighted measure of aggregate labor, which differs from aggregate labor or employment, depending on the distribution of skills (for work in this area see Kydland, 1984, or Kydland and Prescott, 1993). However, a numerical example shows that the *MWC* are relatively insensitive to variations in  $h$ .

argued earlier, no evidence exists to support any particular identification beyond  $S_{\alpha\sigma} < 0$ . To be safe, the simulations consider the whole range of  $S_{\alpha\sigma}$  for a given  $h$ , but when using the range is too unwieldy,  $S_{\alpha\sigma} \in \{0.15, 0, -0.15\}$  is assumed.<sup>26</sup> It was also argued earlier that the evidence seems consistent with  $V_x \in (0, 1]$  for  $x = \alpha, \sigma$  and  $S_{\alpha\sigma} < 0$ . However, the problem is that there is no evidence on the exact value of  $V_x$ .<sup>27</sup> Thus, the polar cases of no-ability variation and no-endowment variation are examined to see the range of possibilities. Finally, following Wolff (1994) and Wolff and Marley (1989), the top wealth decile is assumed to hold 70 percent of aggregate wealth, which is an average of marketable wealth holdings in the United States from the early 1960s to the late 1980s.

### *Simulation Results*

Figure 1 shows the transition effects of aggregate leisure and consumption (measured as percent deviations from the initial steady state) when factor taxes rise by 10 percent. Transition effects are compared for  $S_{\alpha\sigma} \in \{-0.15, 0, 0.15\}$ . It is apparent that when  $S_{\alpha\sigma} < 0$  the distribution effect for consumption and leisure is much likelier to dominate the representative agent effect for labor taxes than for capital taxes. The reason is that capital taxes have much larger representative agent effects on capital than labor taxes. In fact, the representative agent effect of capital taxes is so large that the distribution effect will, in most cases, be dominated, even when  $S_{\alpha\sigma}$  approaches its lower limit. By contrast, the distribution effect of labor taxes is completely offsetting when  $S_{\alpha\sigma}$  falls slightly below -0.15. In addition, the distribution effect is comparably large in the anticipation phase for labor taxes and also for capital taxes very early in the anticipation phase of consumption. Thus, because of the distribution effect when  $S_{\alpha\sigma} < 0$ , labor taxes can easily cause non-standard transition dynamics, something that is unlikely for capital taxes except very early in the anticipation phase.

Figure 2 graphs the wealth inequality effects and the *MWC* for the whole range of  $S_{\alpha\sigma}$  that

---

<sup>26</sup> The range of  $S_{\alpha\sigma}$  is  $(-(1+\varepsilon)^{-1}, \varepsilon(1+\varepsilon)^{-1})$ , which by  $\lambda\varepsilon$  shifts to the right as  $h$  or initial taxes rise. The initial parameter assumptions imply  $S_{\alpha\sigma} \in (-1.33, 0.251)$ , so that  $S_{\alpha\sigma} = -0.15$  is very conservative, contrary to  $S_{\alpha\sigma} = 0.15$ . Thus, the interesting case of  $S_{\alpha\sigma} = -0.15$  represents a fairly small deviation from the representative agent case.

<sup>27</sup> While there has been considerable debate about the endowment or bequest share in aggregate wealth (see Kessler and Masson, 1988), there has been no discussion of how endowments and abilities covary across the population and how these components of wealth covary with tastes. Both issues lie at the heart of determining  $V_x$ .

corresponds with the initial experiment. Specifically, the upper panels examine inequality effects and the lower panels of Figure 2 compare the marginal efficiency cost of factor taxes with the *MWC* for  $V_x = 0$  and  $V_x = 1$ . As a point of reference, having  $S_{\alpha\sigma} = 0$  isolates the wealth inequality effect under weak heterogeneity and the equity effect, which is the gap between the *MEC* and the *MWC*. For labor taxes, the equity effect is zero when  $V_x = 0$  but soon becomes much larger than the efficiency loss as  $V_x$  rises or as the differential human wealth effect becomes stronger and with it the wealth inequality effect. While the equity effect is not as pronounced for capital taxes, it still provides a non-trivial offset to the *MEC* and the *MWC*. However, the difference between the *MWC* evaluated at  $V_x = 0$  and at  $V_x = 1$  plays a minor role, because when differential human wealth effects become weaker, they are balanced by stronger average wealth effects. As  $S_{\alpha\sigma}$  falls from zero, the net effect of taste heterogeneity is to lower the *MWC* for labor taxes, but it still stays positive, while for capital taxes the net effect aggravates the negative *MWC*. Thus, on net tastes tend to play a relatively minor role for labor taxes, but for capital taxes they are more important, causing the *MWC* under strong heterogeneity to approach the *MEC*.<sup>28</sup> The observation that tastes may be more important for  $MEC(t_r)$  than for  $MEC(t_w)$  as  $S_{\alpha\sigma}$  becomes more negative tends to be valid for all experiments considered in this section.

Table 1 presents comparisons of the initial experiment with alternative cases to explore the sensitivity of the model. Rows of the table show the long-run response of leisure and consumption, inequality effects, and the *MWC* under no and strong heterogeneity, where the range of possible inequality effects and the *MWC* is indicated by  $V_x = 0$  and  $V_x = 1$ . Columns of the table distinguish the results according to  $S_{\alpha\sigma} \in \{-0.15, 0, 0.15\}$ . For example, when  $S_{\alpha\sigma} = 0$  the *MEC* is -0.008 for labor taxes and -0.342 for capital taxes in case 1. The numbers also show how strong the equity gain for labor taxes is compared with efficiency costs and distribution effects. When  $S_{\alpha\sigma} = 0$ , the equity

---

<sup>28</sup> Recall that strong heterogeneity has three separate effects on the *MWC*. As  $S_{\alpha\sigma} < 0$  grows more negative, the *MEC* and the equity gain both rise. Thus, the gap between the *MEC* and *MWC* curves widen because the rise of the equity gain dominates as  $S_{\alpha\sigma}$  falls. The third effect on the *MWC* arises from the distribution effect and is usually very small compared to the *MEC* and the equity gain. For the cases considered in this section, initial distortions are sufficiently high that the distribution effect offsets the *MEC* and adds to the equity gain when  $S_{\alpha\sigma} < 0$ .

effect raises the *MWC* for labor taxes to 0.066 when  $V_x = 1$  and diminishes the *MWC* for capital taxes by at least two-thirds. However, when  $S_{\alpha\sigma} = -0.15$ , the *MWC* for capital taxes under strong heterogeneity moves close to the *MEC* despite positive distribution effects that cause underlying changes in consumption, leisure, and inequality to be smaller than under weak heterogeneity.

The first set of alternative experiments varies the initial tax distortions, with case 2 assuming  $t_w = 0.3$  and  $t_r = 0.3$  and case 3 assuming  $t_w = 0.3$  and  $t_r = 0.6$ . These cases confirm that the inequality effect of labor taxes tends to be much stronger than for capital taxes. As the capital tax rate rises, the consumption effect becomes more pronounced as compared with leisure so that the *MEC* rises; yet there is no commensurate reduction in inequality. One reason that the capital tax has comparatively large efficiency effects and small inequality effects is because it is so high. When the capital tax rate is the same size as the labor tax rate, a positive *MWC* is possible despite inequality effects that appear small. Another set of experiments -- case 4 where  $S = 1$  and case 5 where  $S = 5$  -- shows that the much discussed timing issue remains important when considering heterogeneity, although mainly for capital taxes. While changes in  $S$  tend to have modest effects for labor taxes, for capital taxes when  $S$  rises the rapid rise of the *MEC* overwhelms everything else and the *MWC* becomes more negative. When  $S = 1$ , the capital tax is partially a lump-sum tax. In this case, capital taxes have a small efficiency effect and a sizeable inequality effect when wealth variation is mainly due to endowment variation. The final two cases presented in the table vary parameters that have shown wide dispersion in the literature. Specifically, it is shown that varying  $\theta$ , which in case 6 falls to 0.25, is quantitatively important but that doubling  $h$  in case 7 does not appreciably affect the results. Increasing the capital share leaves long-run quantity and inequality effects relatively unchanged for labor taxes and subdues them for capital taxes. However, because the adjustment time toward a lower steady-state is also reduced, the *MEC* and *MWC* for labor and capital taxes tend to rise. Finally, reasonable variations in  $\rho$  and  $n^R \sigma^R$  create no significant qualitative differences.

Surprisingly, the *MECs* in the initial experiment are very close to those in Judd (1987a), despite differences in timing assumptions and parameters. Judd calculates that the *MEC* ranges from

-0.04 to -0.02 for labor taxes and from -0.38 to -0.15 for capital taxes for experiments that come nearest to the ones discussed here. These ranges can easily be replicated. Assume, as Judd does, that  $S = 1$  and  $\theta = 0.25$ , let  $t_w = t_r = 0.3$  in one case and  $t_w = 0.4$  and  $t_r = 0.5$  in the other, constrain the initial government debt to be zero, and let  $h \in [0.2, 0.4]$ . Under these circumstances,  $MEC(t_w) \in (-0.02, -0.01)$  and  $MEC(t_r) \in (-0.36, -0.15)$ . These findings echo Judd's and Auerbach and Kotlikoff's conclusion that capital taxes impose a significantly higher efficiency cost on society than labor taxes. By comparison, case 4 and results by McGrattan (1994), where the tax parameters mimic Judd but all other parameters are close to case 4, suggest that the disparity between  $MEC(t_r)$  and  $MEC(t_w)$  may not be quite so drastic. One reason for the disparity is that Judd and Auerbach and Kotlikoff constrain the initial government debt or deficit to be zero when initial tax rates are raised while the other cases constrain  $\gamma$  to remain unchanged. When  $\gamma$  is adjusted upward to balance an increase in initial tax rates, initial excess distortionary revenues fall, which tends to raise the revenue yield of the labor tax and thus lowers the  $MWC(t_w)$ . However, for capital taxes this outcome is overshadowed by an enhanced responsiveness of the discount factor to tax changes that tends to lower the revenue yield and raise the  $MWC(t_r)$ . Case 4 and McGrattan allow a widening of a preexisting deficit to occur instead of adjusting  $\gamma$ , which compresses the disparity between  $MEC(t_r)$  and  $MEC(t_w)$ . In fact, through this same process, the  $MEC(t_w)$  may even exceed the  $MEC(t_r)$  for an initial surplus when  $S = 1$ . However, the  $MEC(t_r)$  quickly eclipses the  $MEC(t_w)$  as  $S$  rises, illustrating not only the special nature of assuming  $S = 1$  and constraining initial debt but also reaffirming how important timing assumptions are for capital taxes. In fact, when  $S = 1$ , capital taxes always cause large equity gains that dominate small efficiency losses for  $V_x = 0$ , but when  $S > 1$ , capital taxes lead to large efficiency losses that dominate equity gains.

While the welfare costs of capital taxes tend to be sensitive to parameter variation, the numerical examples yield a few robust conclusions when  $S_{\alpha\sigma} < 0$  (which seems reasonable in light of the previous arguments against  $S_{\alpha\sigma} \geq 0$ ) and  $V_x > 0$ . Labor taxes have a distribution effect that may dominate the representative agent effect along the adjustment path even when  $S_{\alpha\sigma}$  is small, but for cap-

ital taxes representative agent effects always rule. Thus, labor taxes can easily have small negative or even positive effects on leisure and consumption. Also, labor taxes tend to be more effective at reducing wealth inequality than capital taxes, except when abilities contribute little to wealth variation and taxes are unanticipated. However, inequality reductions are small: following a 10 percent rise in tax rates, the wealth share of the poorest 90th percentile rises from 30 percent by at least 1.5 percentage points for labor taxes and by no more than 0.5 percentage points for capital taxes. Still, the equity gains from inequality reductions are not trivial when compared with the efficiency losses of labor and capital taxes. In fact, the equity gain for labor taxes usually dominates the marginal efficiency cost under weak heterogeneity. This dominance is unlikely for anticipated capital taxes.

Under strong heterogeneity, equity gains still dominate for labor taxes and cause a positive marginal welfare cost. By contrast, efficiency costs and distribution effects dominate equity gains for capital taxes. In this case, as  $S_{\alpha\sigma}$  becomes more negative, the marginal welfare cost under strong heterogeneity approaches the marginal efficiency cost, which falls in the range computed by Judd (1987a).

## VII. Conclusion

For some economic questions, such as the welfare cost of taxation, it is essential that heterogeneity approximate the real world in order to capture potential distribution effects. However, once heterogeneity is taken seriously in a dynamic model, one must also analyze the consequences for the transitional dynamics because these enter into the welfare calculation. To illustrate the dangers of ignoring heterogeneity, a simple deterministic dynamic model was developed where agents differ in their tastes, abilities, and endowments, as in traditional public finance models. The model is tractable and yields explicit formulas for dynamics, wealth distribution effects, and marginal welfare costs. It is shown that although weak heterogeneity (abilities and endowments) does not affect dynamics, welfare costs are lower than without heterogeneity because factor taxes reduce wealth inequality. In fact, numerical examples show that the equity gain for labor taxes is likely to dominate the efficiency loss, while for capital taxes the equity gain can provide a considerable offset to the effi-

ency loss but most likely is not dominant. Strong heterogeneity (abilities, endowments, and tastes) creates feedback between the wealth distribution and the path of aggregate variables, and this distribution effect can complicate adjustment dynamics and usually aggravates welfare losses. Numerical examples reveal that the distribution effects can easily dominate the adjustment dynamics for labor taxes, but not for capital taxes. Also, equity gains tend to dominate efficiency losses and distribution effects when computing the marginal welfare cost of labor taxes and vice versa for capital taxes.

While the functional form assumptions might appear restrictive, it is not too difficult to show that the heterogeneity considered here still matters for dynamics and welfare costs with more general functional forms, as long as utility is intratemporally and intertemporally separable. The model also generalizes quite naturally to other forms of heterogeneity, such as time preference and age effects. Preliminary work suggests that embedding the model in Blanchard's (1985) overlapping generations framework and assuming heterogeneous survival rates matters for dynamics only if intratemporal tastes differ. However, if annuity markets are imperfect in the sense of Abel (1989), a new intertemporal distribution effect arises via the aggregate Euler equation, which is complementary to evidence by Attanasio and Weber (1993) that age-distribution effects matter. Naturally, future work might also consider stochastic environments along the lines of Long and Plosser (1983), or else such work may explore how heterogeneity alters the outcomes of other policies (see Betsi, 1993).

### Appendix

**The Initial Steady State.** A few important initial steady-state relationships are presented at the outset. First, normalize output to  $y_0 \equiv 1$ . Evaluating (19) through (21) in steady state implies

$$(A1) \quad k_0 = \chi(1-\gamma), \text{ where } \chi = (1-\gamma)^{-1} \theta \rho (1-t_r),$$

$$(A2) \quad h = (1+\lambda)^{-1} = \lambda^{-1} l, \text{ where } \lambda \equiv \frac{(1-\chi)(1-\gamma)}{(1-t_w)(1-\theta)\varepsilon}$$

Given  $\rho > \theta$ , then  $\rho(1-t_r) > 1-\gamma$  implies  $\rho > \chi > \theta$ . Also,  $(1-\chi)(1-\gamma) \geq (1-t_w)(1-\theta)$  implies  $\lambda \varepsilon \geq 1$ .

Evaluating discounted output in (25) yields  $\pi_s = \rho^s$  and  $P \equiv \sum_{s \geq 1} \pi_s = \rho/(1-\rho)$ . Thus, individual wealth in (6), aggregate wealth in (24), and the government's budget constraint can be written as

$$(A3) \quad z^i = k_0^i + b_0^i + P(1-t_w)(1-\theta)(1+\lambda)a^i + Pt$$

$$(A4) \quad E_z = k_0 + P(1-t_w)(1-\theta)\lambda(1+\varepsilon) - P(1-\rho)\rho^{-1}\chi(1-\gamma) = P(1-t_w)(1-\theta)\lambda(1+\varepsilon)$$

$$(A5) \quad b_0 = Pt + P(\tau - \gamma) = Pt + (P(1-t_w)(1-\theta)(\lambda\varepsilon - 1) - k_0)$$

where  $1 + \varepsilon = 1/E_{(1-\alpha)\sigma}$ . (A4) can be used to derive:

$$(A6) \quad E_z S_{x\sigma} = S_{xz} = S_{x(k+b)} + S_{xz_a} = S_{x(k+b)} + V_x(E_z S_{x\sigma}), \quad x = \alpha, \sigma$$

where using (A5) a crucial term is  $V_x S_{x\sigma} \equiv \frac{S_{xz_a}}{E_z} \equiv P(1-t_w)(1-\theta)(1+\lambda)\frac{S_{xa}}{E_z} = \frac{E_{(1-\alpha)\sigma}}{l} S_{xa} = \frac{1+\lambda^{-1}}{1+\varepsilon} S_{xa}$ .

**Derivation of (28)-(33)** Given  $I_s = 1 \forall s \in [S, T-1]$  and  $I_s = 0 \forall s \notin [S, T-1]$ , it follows that

$$(A7) \quad X_s \equiv (1-\chi) \sum_{v \geq 1} \chi^{v-1} I_{s-v} = \begin{cases} \chi^{S-1-s}(1-\chi^{T-S}) & s \leq S-1 \\ 1-\chi^{T-1-s} & S \leq s \leq T-1 \\ 0 & s \geq T \\ 0 & s \leq S-1 \end{cases}$$

$$(A8) \quad K_s^I \equiv \sum_{v=0}^{s-1} \theta^v I_{s-v} = \begin{cases} \frac{1-\theta^{s+1-S}}{1-\theta} & S \leq s \leq T-1 \\ \theta^{s+1-T} K_{T-1}^I & s \geq T \end{cases}$$

$$(A9) \quad K_s^\chi \equiv \sum_{v=0}^{s-1} \theta^v X_{s-v} = \begin{cases} \frac{1-(\theta\chi)^s}{1-\theta\chi} X_s & s \leq S-1 \\ \frac{1-(\theta\chi)^s}{1-\theta\chi} (X_s - 1) + K_s^I + \left(1 - (1-\theta)K_s^I\right) \frac{1-(\theta\chi)^{S-1}}{1-\theta\chi} & S \leq s \leq T-1 \\ \theta^{s+1-T} K_{T-1}^\chi & s \geq T \end{cases}$$

Thus, in the long run or as  $s \rightarrow \infty$ , a permanent shock (or  $T-S \rightarrow \infty$ ) implies  $I_\infty = X_\infty = 1$  and

$K_\infty^\chi = K_\infty^I = 1/(1-\theta)$ , while a temporary shock (or  $T-S < \infty$ ) implies  $I_\infty = X_\infty = K_\infty^\chi = K_\infty^I = 0$ .

To derive equation (28), totally differentiate (19) and insert (A7). Doing so yields

$$(A10) \quad \hat{\lambda}_s = -(1-\gamma)\frac{dt_{rs+1}}{1-t_r} + \gamma\hat{\lambda}_{s+1} = -(1-\gamma)\sum_{v \geq 1} \gamma^{v-1} I_{s+v} \frac{dt_r}{1-t_r} = -X_s \frac{dt_r}{1-t_r}$$

With this calculation, equation (29) follows directly from (20) differentiated and (A10):

$$(A11) \quad \hat{h}_s = -(1-h)\hat{\lambda}_s, \quad \hat{\lambda}_s = -\hat{\varepsilon} + I_s \frac{dt_w}{1-t_w} - \frac{\chi}{1-\chi} \hat{\lambda}_s = -\hat{\varepsilon} + I_s \frac{dt_w}{1-t_w} + \frac{\chi X_s}{1-\chi} \frac{dt_r}{1-t_r}$$

Then, to derive (30), totally differentiate (21), substitute for  $\hat{h}_s$  from (A12), collect  $\hat{\lambda}_s$  coefficients, simplify using (A11), and define  $1-\varphi \equiv (1-h)(1-\theta)$ :

$$(A12) \quad \hat{k}_s = \sum_{v=0}^{s-1} \theta^v \left[ (1-\varphi)\hat{\varepsilon} + (1-\varphi)I_{s-v} \frac{dt_w}{1-t_w} + \frac{1-\varphi\chi}{1-\chi} \hat{\lambda}_{s-v} \right] = (1-h)(1-\theta^s)\hat{\varepsilon} - (1-\varphi)K_s^I \frac{dt_w}{1-t_w} - \frac{1-\varphi\chi}{1-\chi} K_s^\chi \frac{dt_r}{1-t_r}$$

Thus, output and consumption effects in equations (31) and (32) are

$$(A13) \quad \hat{y}_s = -\hat{\lambda}_s + \hat{k}_s = (1-h)(1-\theta^s)\hat{\varepsilon} - (1-\varphi)K_s^I \frac{dt_w}{1-t_w} - \left( \frac{1-\varphi\chi}{1-\chi} K_s^\chi - X_s \right) \frac{dt_r}{1-t_r}$$

$$(A14) \quad \hat{c}_s = -(1-\gamma)^{-1}\hat{\lambda}_s + \hat{k}_s = (1-h)(1-\theta^s)\hat{\varepsilon} - (1-\varphi)K_s^I \frac{dt_w}{1-t_w} - \left( \frac{(1-\varphi\chi)K_s^\chi - X_s}{1-\chi} \right) \frac{dt_r}{1-t_r}$$

Equation (33) uses (A13) and (A11) and links to the capital-labor ratio via  $-(\theta/(1-\theta))\hat{r}_s = \hat{w}_s = \theta(\hat{k}_{s-1} - \hat{h}_s)$ :

$$(A15) \quad \hat{w}_s = \hat{y}_s - \hat{h}_s = -(1-h)\theta^s \hat{\varepsilon} + \left( (1-\varphi)(I_s - K_s^I) + \theta(1-h)I_s \right) \frac{dt_w}{1-t_w} - \left( \frac{1-\varphi\chi}{1-\chi} K_s^\chi - \frac{1-\chi h}{1-\chi} X_s \right) \frac{dt_r}{1-t_r}$$

Finally, total differentiation of equation (28) together with (A10) and (A11) produces

$$(A16) \quad \hat{\pi}_s = -\hat{\lambda}_s + \sum_{u=1}^s \left( \hat{\lambda}_u + I_u \frac{dt_r}{1-t_r} \right) = -\hat{\lambda}_s + \sum_{u=1}^s \left( \hat{\lambda}_u - \frac{1}{1-\chi} (\hat{\lambda}_u - \chi \hat{\lambda}_{u-1}) \right) = \frac{\chi \hat{\lambda}_s - \hat{\lambda}_0}{1-\chi} = - \left( \frac{\chi X_s - X_0}{1-\chi} \right) \frac{dt_r}{1-t_r}$$

$$(A17) \quad \hat{\pi}_s - I_s \frac{dt_w}{1-t_w} + \hat{\lambda}_s = -\hat{\varepsilon} + \frac{X_0}{1-\chi} \frac{dt_r}{1-t_r}, \quad \hat{\pi}_s - I_s \frac{dt_w}{1-t_w} + \frac{\lambda}{1+\lambda} \hat{\lambda}_s = -\hat{\varepsilon} + \frac{X_0}{1-\chi} \frac{dt_r}{1-t_r} - \frac{1}{1+\lambda} \hat{\lambda}_s$$

### Derivation of (34)-(36).

This section finds solutions for two similar equations:

$$(A18) \quad (E_{\alpha\sigma} E_{(1-\alpha)\sigma}) \hat{\varepsilon} = dS_{\alpha\sigma} = E_z^{-1} \sum_i n^i (\alpha^i - E_\alpha) (dz_1^i - dE_z) - S_{\alpha\sigma} \hat{E}_z$$

$$(A19) \quad \frac{1}{2} dS_{\sigma\sigma} = E_z^{-1} \sum_i n^i (\sigma^i - 1) (dz^i - dE_z) - S_{\sigma\sigma} \hat{E}_z$$

Straightforward differentiation of average wealth and inserting (A17) implies simply

$$(A20) \quad \hat{E}_z = \frac{(1-t_w)(1-\theta)\lambda(1+\varepsilon)}{E_z} \sum_{s \geq 1} \rho^s \left( -\hat{\varepsilon} + \frac{X_0}{1-\chi} \frac{dt_r}{1-t_r} + \frac{\varepsilon}{1+\varepsilon} \hat{\varepsilon} \right) = -E_{(1-\alpha)\sigma} \hat{\varepsilon} + \frac{X_0}{1-\chi} \frac{dt_r}{1-t_r}$$

The differential effect is derived by differentiating  $S_{x\sigma}$  in (22), treating  $x$  as fixed, and inserting

(A17). These steps yield the same answer as computing  $dz_1^i - dE_z$  and then aggregating. Thus,

$$\sum_i n^i (x^i - E_x)(dz^i - dE_z) = S_{xa}(1-t_w)(1-\theta)(1+\lambda) \sum_{s \geq 1} \rho^s \left( -\hat{\varepsilon} + \frac{X_0}{1-\chi} \frac{dt_r}{1-t_r} - \frac{1}{1+\lambda} \hat{\lambda}_s \right)$$

Substituting equations (A11), (A4) and (A16) into this expression yields

$$\sum_i n^i (x^i - E_x)(dz^i - dE_z) = -S_{xa} E_z E_{(1-\alpha)\sigma} \left( \hat{\varepsilon} + \frac{P_0}{\lambda} \frac{dt_w}{1-t_w} - \frac{X_0}{1-\chi} \frac{dt_r}{1-t_r} \right) - \frac{S_{xa} E_z E_{(1-\alpha)\sigma}}{\lambda} \Psi \frac{dt_r}{1-t_r}$$

where

$$P_0 \equiv P^{-1} \sum_{s \geq 1} \rho^s I_s = \rho^{S-1} (1 - \rho^{T-S}), \quad \Psi \equiv \frac{\chi P^X}{(1-\chi)P} - \frac{X_0}{1-\chi} \quad \text{and}$$

$$P^X \equiv \sum_{s \geq 1} \rho^s X_s = (1 - \chi^{T-S}) \rho^{S-1} \sum_{s=1}^{S-1} \left( \frac{\chi}{\rho} \right)^{S-1-s} + \sum_{s=S}^{T-1} \rho^s - \rho^{T-1} \sum_{s=S}^{T-1} \left( \frac{\chi}{\rho} \right)^{T-1-s} = \frac{1-\chi}{\rho-\chi} P_0 P - \frac{\rho}{\rho-\chi} X_0$$

Simplifying using (A6) results in the differential effect for  $x = \alpha, \sigma$  :

$$(A21) \quad \sum_i (x^i - E_x) \frac{dz^i - dE_z}{E_z} = -S_{xa} E_{(1-\alpha)\sigma} \left( \hat{\varepsilon} + \frac{P_0}{\lambda} \frac{dt_w}{1-t_w} - \frac{X_0}{1-\chi} \frac{dt_r}{1-t_r} \right) - \left( S_{x\sigma} - \frac{S_{x(k+b)}}{E_z} - E_{(1-\alpha)\sigma} S_{xa} \right) \Psi \frac{dt_r}{1-t_r}$$

Finally, inserting (A20) and (A21) into (A18) and (A19) and collecting  $\hat{\varepsilon}$ -terms yields

$$(A22) \quad E_{(1-\alpha)\sigma} (E_\alpha + S_{\alpha\alpha}) \hat{\varepsilon} = -E_{(1-\alpha)\sigma} S_{\alpha\alpha} \frac{P_0}{\lambda} \frac{dt_w}{1-t_w} - \frac{S_{\alpha(k+b)} X_0}{E_z (1-\chi)} \frac{dt_r}{1-t_r} - \left( S_{\alpha\sigma} - \frac{S_{\alpha(k+b)}}{E_z} - E_{(1-\alpha)\sigma} S_{\alpha\alpha} \right) \Psi_0 \frac{dt_r}{1-t_r}$$

$$(A23) \quad \frac{dS_{\sigma\sigma}}{2} = E_{(1-\alpha)\sigma} (S_{\sigma\sigma} - S_{\sigma\alpha}) \hat{\varepsilon} - E_{(1-\alpha)\sigma} S_{\sigma\alpha} \frac{P_0}{\lambda} \frac{dt_w}{1-t_w} - \frac{S_{\sigma(k+b)} X_0}{E_z (1-\chi)} \frac{dt_r}{1-t_r} - \left( S_{\sigma\sigma} - \frac{S_{\sigma(k+b)}}{E_z} - E_{(1-\alpha)\sigma} S_{\sigma\alpha} \right) \Psi_0 \frac{dt_r}{1-t_r}$$

$$\text{where } \Psi_0 \equiv \Psi + \frac{X_0}{1-\chi} \equiv \frac{\chi}{1-\chi} \frac{P^X}{P} = \frac{\chi(1-\rho)}{\rho-\chi} \left( \frac{\rho^{S-1}(1-\rho^{T-S})}{1-\rho} - \frac{\chi^{S-1}(1-\chi^{T-S})}{1-\chi} \right) = \frac{\chi(1-\rho)}{\rho-\chi} \left( \frac{P_0}{1-\rho} - \frac{X_0}{1-\chi} \right).$$

Equations (A20), (A22), and (A23) are simplified using (A2) and (A6) and defining  $x = \alpha, \sigma$ :

$$\Psi_w^x \equiv l \frac{S_{xz_a}}{E_z} P_0 = S_{x\sigma} P_0 \frac{V_x}{1+\lambda^{-1}}, \quad \Psi_r^x \equiv \left( S_{x\sigma} - \frac{S_{xz_a}}{E_z} \right) \frac{X_0}{1-\chi} + h \frac{S_{xz_a}}{E_z} \Psi_0 = S_{x\sigma} \frac{X_0}{1-\chi} (1 - V_x + h \phi_0 V_x)$$

$$\Psi_\varepsilon^\alpha \equiv E_{(1-\alpha)\sigma} E_\alpha + l \frac{S_{\alpha z_a}}{E_z} = S_{x\sigma} l \left( \frac{1+\lambda^{-1}}{1+\varepsilon} \frac{E_\alpha}{S_{\alpha\sigma}} + V_\alpha \right), \quad \Psi_\varepsilon^\sigma \equiv E_{(1-\alpha)\sigma} S_{\sigma\sigma} - l \frac{S_{\sigma z_a}}{E_z} = S_{\sigma\sigma} l \left( \frac{1+\lambda^{-1}}{1+\varepsilon} - V_\sigma \right)$$

$$\text{where } \phi_0 \equiv \frac{\Psi_0(1-\chi)}{X_0} = \frac{\chi(1-\rho)}{\rho-\chi} \left( \frac{P_0}{X_0} \frac{1-\chi}{1-\rho} - 1 \right) > 0 \text{ for } T \rightarrow \infty \text{ or } S = T-1 > 1 \text{ and } \phi_0 = 0 \text{ for } S = T-1 = 1.$$

Inserting these definitions into (A22) yields (34), where the coefficients are  $P_0 \Psi_w^\varepsilon \equiv \Psi_w^\alpha / \Psi_\varepsilon^\alpha$  and

$$(1+\lambda^{-1}) \frac{X_0}{1-\chi} \Psi_r^\varepsilon \equiv \frac{\Psi_r^\alpha}{\Psi_\varepsilon^\alpha}. \text{ To evaluate } \Psi_w^\varepsilon \text{ and } \Psi_r^\varepsilon \text{ for } V_x \in [0, 1], \text{ fix } \lambda \text{ or } \frac{\varepsilon}{1+\varepsilon} = E_\alpha + S_{\alpha\sigma}. \text{ For } E_\alpha \in (0, 1) \text{ it}$$

follows that  $S_{\alpha\sigma} \in \left( \frac{\varepsilon}{1+\varepsilon}, \frac{-1}{1+\varepsilon} \right)$  and  $\frac{E_\alpha}{S_{\alpha\sigma}} \in (0, -(1+\varepsilon))$ . As  $S_{\alpha\sigma}$  rises from its lower limit,  $\frac{1+\lambda^{-1}}{1+\varepsilon} \frac{E_\alpha}{S_{\alpha\sigma}} + V_\alpha$

rises from a lower limit of  $V_\alpha - (1+\lambda^{-1})$  to an upper limit of  $V_\alpha$ . Both  $\Psi_w^\varepsilon$  and  $\Psi_r^\varepsilon$  rise with  $S_{\alpha\sigma}$  from a lower negative limit to zero (when  $S_{\alpha\sigma} = 0$ ) and then to an upper positive limit. Thus,  $\Psi_w^\varepsilon|_{V_\alpha=0} = 0$  and  $\text{sign}\{\Psi_w^\varepsilon|_{V_\alpha \in (0,1)}\} = \text{sign}\{\Psi_r^\varepsilon|_{V_\alpha \in [0,1]}\} = \text{sign}\{S_{\alpha\sigma}\}$ . Also,  $|\Psi_w^\varepsilon|$  and  $|\Psi_r^\varepsilon|$  rise with  $|S_{\alpha\sigma}|$ .

Then inserting (34) and the above definitions into (A20) and (A23) yields (35) and (36) via:

$$(A24) \quad \hat{E}_z = -\frac{\hat{\varepsilon}}{1+\varepsilon} + \frac{X_0}{1-\chi} \frac{dt_r}{1-t_r} \equiv \frac{P_0 \Psi_w^\varepsilon}{1+\varepsilon} \frac{dt_w}{1-t_w} + \frac{X_0}{1-\chi} \left( 1 + \frac{1+\lambda^{-1}}{1+\varepsilon} \Psi_r^\varepsilon \right) \frac{dt_r}{1-t_r}$$

$$(A25) \quad \frac{dS_{\alpha\sigma}}{2} = \Psi_\varepsilon^\sigma \hat{\varepsilon} - \Psi_w^\sigma \frac{dt_w}{1-t_w} - \Psi_r^\sigma \frac{dt_r}{1-t_r} \equiv -P_0 l S_{\alpha\sigma} \Psi_w^{S_{\alpha\sigma}} \frac{dt_w}{1-t_w} - \frac{X_0 S_{\alpha\sigma}}{1-\chi} \Psi_r^{S_{\alpha\sigma}} \frac{dt_r}{1-t_r}$$

Labor taxes only act through  $\hat{\varepsilon}$  in (A24) wealth; thus, the analysis is as above. Capital taxes have a positive direct effect and a distributional effect that rises with  $S_{\alpha\sigma}$  through  $\Psi_r^\varepsilon$ . In fact,  $\Psi_r^\varepsilon = 0$  for  $S_{\alpha\sigma} = 0$

and  $1 + \frac{1+\lambda^{-1}}{1+\varepsilon} \Psi_r^\varepsilon|_{V_\alpha=0} \in \left( \frac{\varepsilon}{1+\varepsilon}, +\infty \right)$  and  $1 + \frac{1+\lambda^{-1}}{1+\varepsilon} \Psi_r^\varepsilon|_{V_\alpha=1} \in \left( 1 - \frac{\varphi_0}{1+\varepsilon}, 1 + \frac{\lambda^{-1}\varphi_0}{1+\varepsilon} \right)$ . The lower bound for

$V_\alpha = 1$  is positive if  $1+\varepsilon > \varphi_0$  and less than the lower bound for  $V_\alpha = 0$  if  $\varphi_0 > 1$ . Since  $\varphi_0$  can be large (if  $S$  is large and  $T-S$  small), the lower bound for  $V_\alpha = 1$  may be negative.

Bounds on the coefficients in (A25) are  $\Psi_w^{S_{\alpha\sigma}}|_{V_\alpha=V_\sigma=0} = 0$  and  $\Psi_w^{S_{\alpha\sigma}}|_{V_\alpha=V_\sigma=1} \in \left( \varepsilon \lambda \frac{1+\lambda^{-1}}{1+\varepsilon}, \frac{1+\lambda^{-1}}{1+\varepsilon} \right)$ ,

$\Psi_r^{S_{\alpha\sigma}}|_{V_\alpha=V_\sigma=0} \in \left( \frac{\varepsilon}{1+\varepsilon}, +\infty \right)$ , and  $\Psi_r^{S_{\alpha\sigma}}|_{V_\alpha=V_\sigma=1} \in \left( \frac{\varphi_0 \varepsilon}{1+\varepsilon}, \frac{\varphi_0 \lambda^{-1}}{1+\varepsilon} \right)$ , where  $\lambda \varepsilon \geq 1$  in (A2). Also, the

$\Psi_x^{S_{\alpha\sigma}}|_{V_\alpha, V_\sigma \in [0,1]}$  depend positively on  $S_{\alpha\sigma}$  through  $\Psi_x^\varepsilon$  with  $S_{\alpha\sigma} = 0$  implying  $\Psi_w^{S_{\alpha\sigma}} = V_\sigma$  and

$$\Psi_r^{S_{\alpha\sigma}} = 1 - V_\sigma + h\varphi_0 V_\sigma.$$

**Derivation of (38)-(39).** The first step toward solving for the welfare costs of factor taxation is differentiating discounted excess distortionary tax revenues in (23):

$$dR_0 = \sum_s \rho^s \left( \left( (1-\chi)(1-\gamma) - (1-\theta)(1-t_w) \right) \hat{\pi}_s + \left( -(1-\gamma)\chi \hat{\lambda}_s + (1-\theta)(1-t_w) \frac{dt_w}{1-t_w} \right) \right)$$

which is evaluated using (A16), (A10), and the relationship  $(1-\chi)(1-\gamma) - (1-\theta)(1-t_w) = (1-\theta)(1-t_w)(\lambda\varepsilon - 1)$ .

Simplifying the resulting expression yields the first equation required to calculate marginal welfare

costs:

$$(A26) \quad \frac{1-\rho}{\rho} dR_0 = (1-\theta)(1-t_w) \left( P_0 \frac{dt_w}{1-t_w} + (\lambda\varepsilon-1) \frac{X_0}{1-\chi} \frac{dt_r}{1-t_r} \right)$$

The second step is taking the total differential of welfare in (27) which results in

$$dU = P \sum_i n^i \hat{\delta}^i + E_\alpha \sum_s \rho^s \hat{c}_s^{RA} + E_{1-\alpha} \sum_s \rho^s h \hat{\lambda}_s^{RA} + P \left( \frac{\lambda\varepsilon-1}{(1+\lambda)(1+\varepsilon)} - \frac{E_\alpha}{1+\lambda^{-1}} \frac{\theta(1-\rho)}{1-\rho\theta} \right) \hat{\varepsilon}$$

The crucial step is evaluating the discounted sum of the consumption and leisure effects, which requires finding the discounted sum of (A11) and (A14) using (A8) and (A9). It can be shown that (A8) and (A9)

imply  $\sum_s \rho^s K_s^I = \frac{P_0 P}{1-\rho\theta}$  and  $\sum_s \rho^s K_s^X = \frac{P^X}{1-\rho\theta}$ . Consequently,

$$E_\alpha \sum_s \rho^s \hat{c}_s^{RA} + E_{1-\alpha} \sum_s \rho^s h \hat{\lambda}_s^{RA} = P_0 P \left( -E_\alpha \frac{1-\phi}{1-\rho\theta} + E_{1-\alpha} h \right) \frac{dt_w}{1-t_w} + \frac{P^X}{1-\chi} \left( -E_\alpha \frac{\rho\theta - \phi\chi}{1-\rho\theta} + E_\alpha h \chi \right) \frac{dt_r}{1-t_r}$$

Inserting  $1-\phi \equiv (1-h)(1-\theta)$  and  $\chi P^X = \phi_0 X_0 P$  and then rearranging the coefficients yields the second equation required to calculate marginal welfare costs:

$$(A27) \quad \left( \frac{1-\rho}{\rho} \frac{1}{(1-\chi)(1-\gamma)} \right) \frac{dU}{u_c^{RA}} = -U^w P_0 \frac{dt_w}{1-t_w} - U^r \frac{\chi X_0 \phi_0}{1-\chi} \frac{dt_r}{1-t_r} + \left( 1 - \frac{1}{\sigma^R} \right) \left( \frac{1}{\sigma^P} - 1 \right) \frac{\hat{S}_{\sigma\sigma}}{2E_\alpha} + U^\varepsilon \hat{\varepsilon}$$

where  $U^w$ ,  $U^r$ , and  $U^\varepsilon$  are defined in the main text. Combining (A26) and (A27) yields (38) and (39).

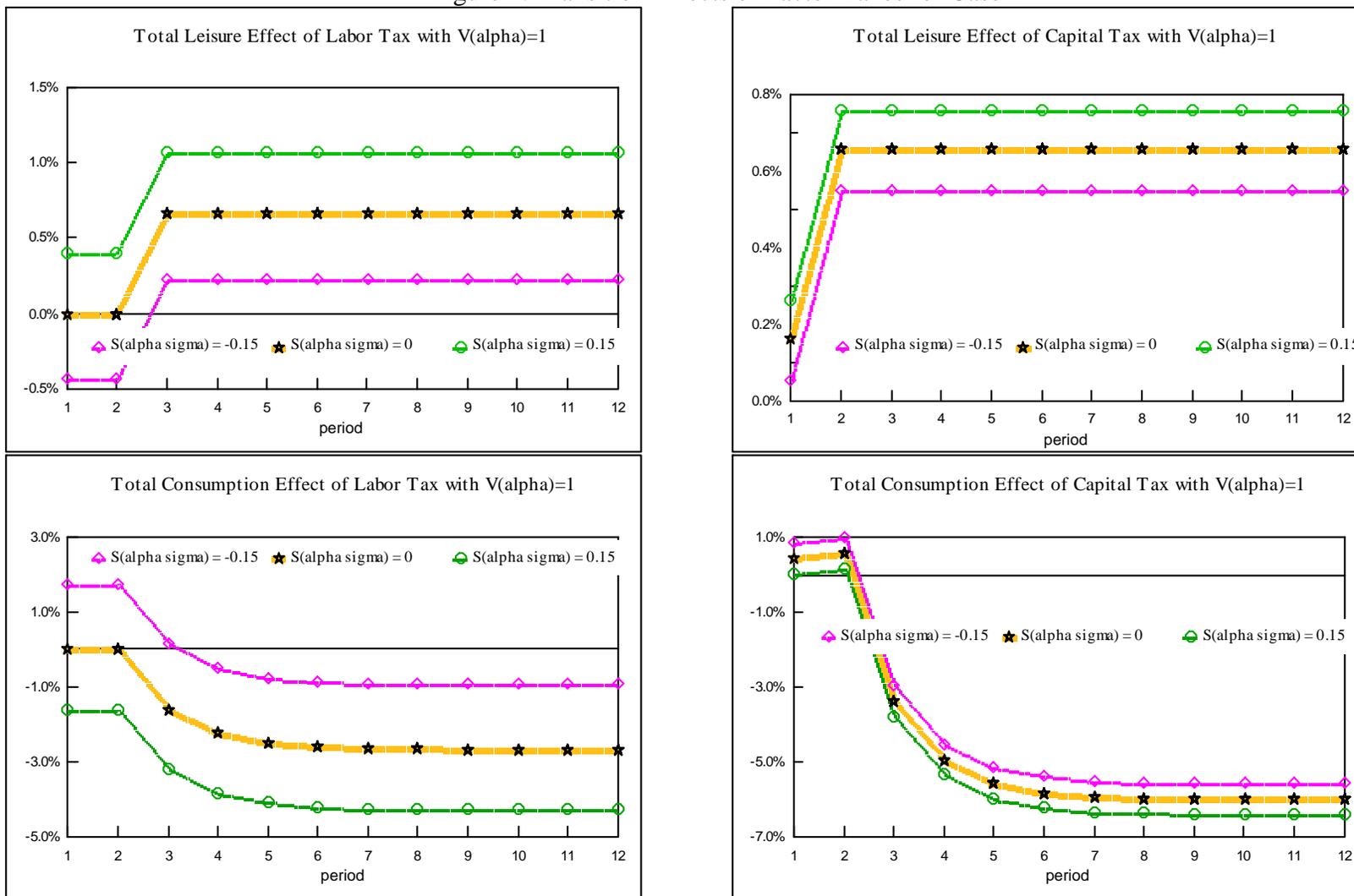
**References**

- Abel, Andrew B. (1989), "Birth, Death and Taxes," *Journal of Public Economics*, 39, 1-15.
- Aiyagari, S. Rao (1994), "Uninsured Idiosyncratic Risk and Aggregate Savings," *Quarterly Journal of Economics*, 109, 659-684.
- Atkeson, A., and R. E. Lucas (1992), "On Efficient Distribution With Private Information," *Review of Economic Studies*, 59: 427-453.
- Attanasio, O.P., and G. Weber (1993), "Consumption Growth, the Interest Rate and Aggregation," *Review of Economic Studies*, 60(3), 631-649.
- Auerbach, A. and L. Kotlikoff (1987), *Dynamic Fiscal Policy*, Cambridge: Cambridge University Press.
- Becsi, Zsolt (1993), "Dynamic Fiscal Policy with Heterogeneous Consumer/ Workers," mimeo.
- Bencivenga, Valerie R. (1992), "An Econometric Study of Hours and Output Variation with Preference Shocks," *International Economic Review*, 33(2), 449-471.
- Blanchard, Olivier J. (1985), "Debt, Deficits and Finite Horizons," *Journal of Political Economy*, 93(2), 223-247.
- Blanchard, O. J., and S. Fischer (1989). *Lectures on Macroeconomics*. Cambridge: MIT Press.
- Blundell, R., P. Pashardes, and G. Weber (1993), "What Do We Learn About Consumer Demand Patterns from Micro Data?," *American Economic Review*, 83(3), 570-97.
- Boadway, Robin W. (1994), "The Role of Second-Best Theory in Public Economics," Queen's University, Institute for Economic Research Discussion Paper 910, October.
- Bourguignon, Francois (1979), "Decomposable Income Inequality Measures," *Econometrica*, 47, 901-20.
- Chamley, Christophe (1986), "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54(3), May 1986, 607-22.
- Christiano, Lawrence J. (1988), "Why Does Inventory Investment Fluctuate So Much?," *Journal of Monetary Economics*, 21(2/3), 247-80.
- Cogley, T., and J. M. Nason (1995) "Output Dynamics in Real Business Cycle Models," *American Economic Review*, 85(3), 492-511.
- Cooley, T.F., and E.C. Prescott, (1995), "Economic Growth and Business Cycles," T.F. Cooley, ed., *Frontiers of Business Cycle Research*, Princeton: Princeton University Press.
- Cowell, Frank A. (1980), "On the Structure of Additive Inequality measures," *Review of Economic Studies*, 47, 521-531.

- Cowell, Frank A. (1995), *Measuring Inequality*, Second Edition, New York: Prentice Hall.
- Cutler, D. M., and L. F. Katz (1992), "Rising Inequality? Changes in the Distribution of Income and Consumption in the 1980s," *American Economic Review*, 82(2), 546-551.
- Diamond, Peter A. (1975), "A Many-Person Ramsey Tax Rule," *Journal of Public Economics*, 4, 335-42.
- Diamond, P.A., and J.A. Hausman (1984), "Individual Retirement and Savings Behavior," *Journal of Public Economics*, 23(1/2), 81-114.
- Eichenbaum, M., L. Hansen, and K. Singleton (1988), "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice Under Uncertainty," *Quarterly Journal of Economics*, 103, 51-78.
- Glick, R., and M. Hutchinson (1990), "New Results in Support of the Fiscal Policy Ineffectiveness Proposition," *Journal of Money, Credit, and Banking*, 22(3), 288-304
- Greenwood, J., R. Rogerson and R. Wright (1995), "Household Production in Real Business Cycle Theory," T. Cooley, ed., *Frontiers of Business Cycle Research*, Princeton: Princeton University Press.
- Guesnerie, R. (1995), "The Genealogy of Modern Theoretical Public Economics: From First Best to Second Best," *European Economic Review*, 39, 353-381.
- Hall, Robert E. (1971), "The Dynamic Effects of Fiscal Policy in an Economy with Foresight," *Review of Economic Studies*, 38(114), 229-244.
- Hansson, Ingemar (1985), "An Evaluation of the Evidence of the Impact of Taxation on Capital Formation," International Institute of Public Finance - 41st Congress: Madrid, Spain, August.
- Judd, Kenneth L. (1985), "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 28(1), 59-83.
- Judd, Kenneth L. (1987a), "Welfare Cost of Factor Taxation in a Perfect Foresight Model," *Journal of Political Economy*, 95(4), 675-709.
- Judd, Kenneth L. (1987b), "Debt and Distortionary Taxation in a Simple Perfect Foresight Model," *Journal of Monetary Economics*, 20(1), 51-72.
- Kessler, Denis, and André Masson (1988), "On Five Hot Issues on Wealth Distribution," *European Economic Review* 32, 644-653.
- King, R., C. Plosser and S. Rebelo. (1988a), "Production, Growth and Business Cycles: I. The Basic Neoclassical Model," *Journal of Monetary Economics*, 21(2/3), 195-232.
- King, R., C. Plosser and S. Rebelo. (1988b), "Production, Growth and Business Cycles: II. New Directions," *Journal of Monetary Economics*, 21(2/3), 309-341.

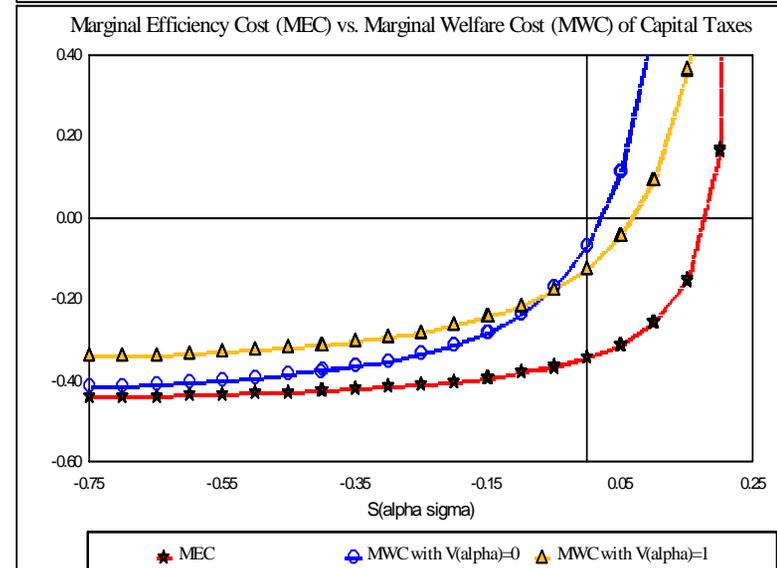
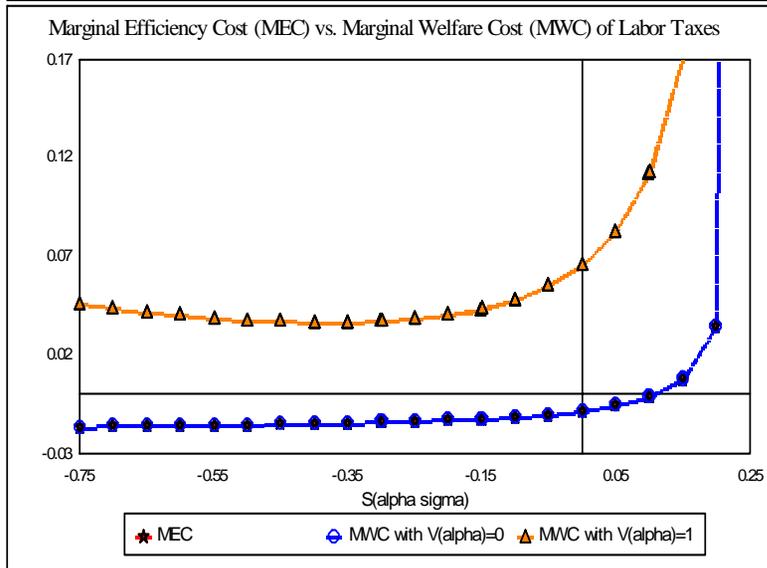
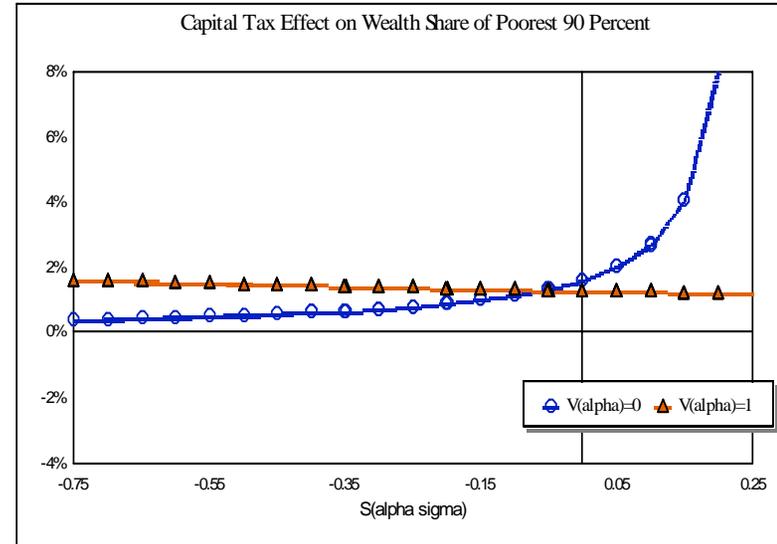
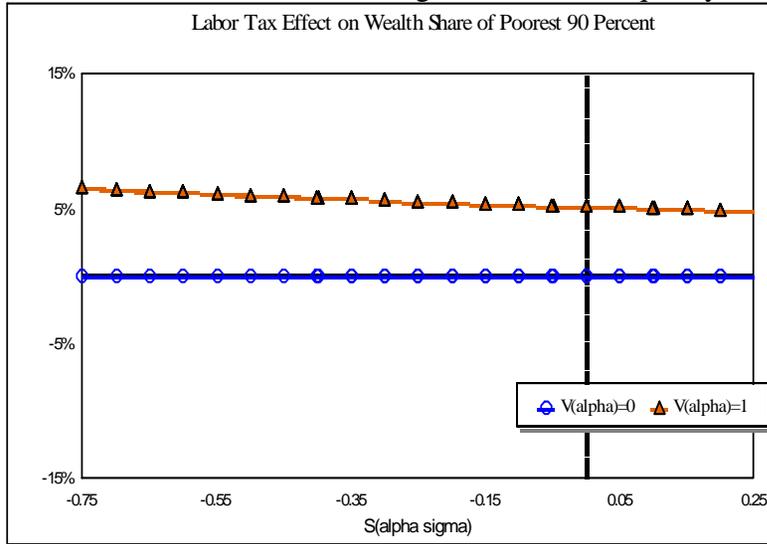
- Killingsworth, Mark R. (1983), *Labor Supply*, New York: Cambridge University Press.
- Kirman, Alan P. (1992), "Whom or What Does the Representative Individual Represent?," *Journal of Economic Perspectives*, 6(2), 117-136.
- Kydland, Finn E. (1984), "Labor Force Heterogeneity and the Business Cycle." *Carnegie-Rochester Conference Series on Public Policy*, 21(0), 173-208.
- Kydland, F. E., and E. C. Prescott (1993), "Cyclical Movements of the Labor Input and Its Implicit Real Wage," Federal Reserve Bank of Cleveland *Economic Review*, 29(2), 12-23.
- Laitner, John (1990), "Tax Changes and Phase Diagrams for an Overlapping Generations Model," *Journal of Political Economy*, 98(1), 193-220.
- Long, J., and C. Plosser (1983), "Real Business Cycles," *Journal of Political Economy*, 91(1), 39-69.
- McGrattan, Ellen R. (1994), "The Macroeconomic Effects of Distortionary Taxation," *Journal of Monetary Economics*, 33(3), 573-601.
- Menchik, P.L., and M. David (1983), "Income Distribution, Lifetime Savings, and Bequests," *American Economic Review*, 73(4), 672-690.
- Mirrlees, James A. (1971), "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38(114), 175-208.
- Rios-Rull, Jose-Victor (1995), "Models with Heterogeneous Agents," T.F. Cooley, ed., *Frontiers of Business Cycle Research*, Princeton: Princeton University Press.
- Shorrocks, Anthony F. (1980), "The Class of Additively Decomposable Inequality Measures," *Econometrica*, 48(3), 613-25.
- Slemrod, Joel (1990), "Do Taxes Matter? Lessons From the 1980s," J. Slemrod, ed., *Do Taxes Matter: The Impact of the Tax Reform Act of 1986*. Cambridge: MIT Press.
- Stoker, Thomas M. (1986), "Simple Tests of Distributional Effects on Macroeconomic Equations," *Journal of Political Economy*, 94(4), 763-794.
- Stoker, Thomas M. (1993), "Empirical Approaches to the Problem of Aggregation Over Individuals," *Journal of Economic Literature*, 31(4), 1827-74.
- Wolff, Edward N. (1994), "Trends in Household Wealth in the United States during the 1980s," *Review of Income and Wealth*, 40(2), 143-174.
- Wolff, E. N., and M. Marley (1989), "Long-Term Trends in U.S. Wealth Inequality: Methodological Issues and Results," R.E. Lipsey and H. S. Tice, eds., *The Measurement of Saving, Investment, and Wealth*, Chicago: University of Chicago Press.

Figure 1: Transition Effects of Factor Taxes for Case 1\*



\*Note: Case 1 assumes  $S = 3$ ,  $\rho = 0.99$ ,  $\theta = 0.4$ ,  $t_w = 0.25$ ,  $t_r = 0.5$ ,  $\gamma = 0.2$ ,  $h = 0.2$ ,  $n^R = 0.1$ , and  $n^R \sigma^R = 0.7$ .  
 Also, the Total Effect of a tax change equals a Representative Agent Effect plus a Distribution Effect for select  $S_{\alpha\sigma}$ .

Figure 2: Wealth Equality and Welfare Effects of Factor Taxes for Case 1\*



\*Note: Case 1 assumes  $S = 3$ ,  $\rho = 0.99$ ,  $\theta = 0.4$ ,  $t_w = 0.25$ ,  $t_r = 0.5$ ,  $\gamma = 0.2$ ,  $h = 0.2$ ,  $n^R = 0.1$ , and  $n^R \sigma^R = 0.7$ .

Table 1: Comparison of Computational Experiments

	$V_x$	Labor Tax			Capital Tax			Experiments with operative parameters
		$S_{\sigma\sigma}$ -0.15	0	0.15	$S_{\sigma\sigma}$ -0.15	0	0.15	
Long-Run Leisure Effect ( $\hat{l}_\infty$ )	0	0.67%	0.67%	0.67%	0.58%	0.66%	0.98%	<b>Case 1</b> $S = 3$ $\theta = 0.4$ $\rho = 0.99$ $h = 0.2$ $t_w = 0.25$ $t_r = 0.5$ $n^R = 0.1$ $\sigma^R = 7$
	1	0.23%	0.67%	1.07%	0.55%	0.66%	0.76%	
Long-Run Consumption Effect ( $\hat{c}_\infty$ )	0	-2.67%	-2.67%	-2.67%	-5.68%	-6.01%	-7.30%	
	1	-0.93%	-2.67%	-4.27%	-5.57%	-6.01%	-6.41%	
Wealth Equality Effect ( $\hat{\sigma}^p$ )	0	0	0	0	1.02%	1.63%	4.06%	
	1	5.45%	5.23%	5.02%	1.36%	1.31%	1.25%	
Marginal Efficiency Cost (MEC)		-0.012	-0.008	0.008	-0.390	-0.342	-0.151	
Marginal Welfare Cost (MWC)	0	-0.012	-0.008	0.008	-0.278	-0.068	1.516	
	1	0.043	0.066	0.173	-0.240	-0.122	0.365	
Long-Run Leisure Effect ( $\hat{l}_\infty$ )	0	0.86%	0.86%	0.86%	0.37%	0.45%	0.81%	
	1	0.28%	0.86%	1.40%	0.38%	0.45%	0.53%	
Long-Run Consumption Effect ( $\hat{c}_\infty$ )	0	-3.43%	-3.43%	-3.43%	-2.08%	-2.40%	-3.82%	
	1	-1.13%	-3.43%	-5.59%	-2.09%	-2.40%	-2.69%	
Wealth Equality Effect ( $\hat{\sigma}^p$ )	0	0	0	0	0.96%	1.57%	4.28%	
	1	6.94%	6.72%	6.52%	0.93%	0.90%	0.88%	
Marginal Efficiency Cost (MEC)		-0.011	-0.007	0.011	-0.185	-0.137	0.076	
Marginal Welfare Cost (MWC)	0	-0.011	-0.007	0.011	-0.042	0.234	2.787	
	1	0.042	0.065	0.192	-0.047	0.076	0.631	
Long-Run Leisure Effect ( $\hat{l}_\infty$ )	0	0.86%	0.86%	0.86%	0.67%	0.74%	0.98%	<b>Case 3</b>         $t_w = 0.3$ $t_r = 0.3$
	1	0.32%	0.86%	1.33%	0.62%	0.74%	0.84%	
Long-Run Consumption Effect ( $\hat{c}_\infty$ )	0	-3.43%	-3.43%	-3.43%	-8.97%	-9.26%	-10.22%	
	1	-1.29%	-3.43%	-5.34%	-8.79%	-9.26%	-9.68%	
Wealth Equality Effect ( $\hat{\sigma}^p$ )	0	0	0	0	0.95%	1.47%	3.21%	
	1	7.13%	6.72%	6.36%	1.56%	1.47%	1.39%	
Marginal Efficiency Cost (MEC)		-0.015	-0.011	0.003	-0.508	-0.471	-0.345	
Marginal Welfare Cost (MWC)	0	-0.015	-0.011	0.003	-0.426	-0.289	0.491	
	1	0.047	0.066	0.147	-0.374	-0.289	0.017	
Long-Run Leisure Effect ( $\hat{l}_\infty$ )	0	0.67%	0.67%	0.67%	-0.67%	0.66%	5.94%	
	1	0.22%	0.67%	1.08%	0.55%	0.66%	0.76%	
Long-Run Consumption Effect ( $\hat{c}_\infty$ )	0	-2.67%	-2.67%	-2.67%	-0.70%	-6.01%	-27.16%	
	1	-0.89%	-2.67%	-4.30%	-5.57%	-6.01%	-6.41%	
Wealth Equality Effect ( $\hat{\sigma}^p$ )	0	0	0	0	16.63%	26.58%	66.20%	
	1	5.56%	5.33%	5.13%	1.37%	1.32%	1.26%	
Marginal Efficiency Cost (MEC)		-0.012	-0.008	0.008	-0.024	-0.021	-0.009	
Marginal Welfare Cost (MWC)	0	-0.012	-0.008	0.008	0.088	0.253	1.657	
	1	0.043	0.066	0.173	-0.015	-0.008	0.023	
Long-Run Leisure Effect ( $\hat{l}_\infty$ )	0	0.67%	0.67%	0.67%	0.65%	0.66%	0.68%	<b>Case 5</b> $S = 5$
	1	0.24%	0.67%	1.06%	0.55%	0.66%	0.76%	
Long-Run Consumption Effect ( $\hat{c}_\infty$ )	0	-2.67%	-2.67%	-2.67%	-5.99%	-6.01%	-6.09%	
	1	-0.96%	-2.67%	-4.24%	-5.58%	-6.01%	-6.40%	
Wealth Equality Effect ( $\hat{\sigma}^p$ )	0	0	0	0	0.06%	0.10%	0.25%	
	1	5.34%	5.12%	4.92%	1.33%	1.28%	1.23%	
Marginal Efficiency Cost (MEC)		-0.012	-0.008	0.008	-6.237	-5.470	-2.414	
Marginal Welfare Cost (MWC)	0	-0.012	-0.008	0.008	-6.125	-5.196	-0.747	
	1	0.043	0.066	0.173	-3.849	-1.955	5.846	
Long-Run Leisure Effect ( $\hat{l}_\infty$ )	0	0.67%	0.67%	0.67%	0.34%	0.37%	0.50%	
	1	0.21%	0.67%	1.10%	0.30%	0.37%	0.43%	
Long-Run Consumption Effect ( $\hat{c}_\infty$ )	0	-2.67%	-2.67%	-2.67%	-2.85%	-2.97%	-3.51%	
	1	-0.85%	-2.67%	-4.39%	-2.72%	-2.97%	-3.21%	
Wealth Equality Effect ( $\hat{\sigma}^p$ )	0	0	0	0	0.34%	0.57%	1.61%	
	1	5.37%	5.23%	5.09%	0.75%	0.73%	0.71%	
Marginal Efficiency Cost (MEC)		-0.025	-0.021	-0.002	-1.741	-1.610	-0.994	
Marginal Welfare Cost (MWC)	0	-0.025	-0.021	-0.002	-1.564	-1.165	2.550	
	1	0.042	0.051	0.173	-1.357	-1.039	0.560	
Long-Run Leisure Effect ( $\hat{l}_\infty$ )	0	1.33%	1.33%	1.33%	1.17%	1.32%	1.60%	<b>Case 7</b>         $h = 0.4$
	1	0.84%	1.33%	1.79%	0.99%	1.32%	1.62%	
Long-Run Consumption Effect ( $\hat{c}_\infty$ )	0	-2.00%	-2.00%	-2.00%	-5.13%	-5.35%	-5.78%	
	1	-1.26%	-2.00%	-2.68%	-4.86%	-5.35%	-5.80%	
Wealth Equality Effect ( $\hat{\sigma}^p$ )	0	0	0	0	1.24%	1.63%	2.39%	
	1	4.10%	3.92%	3.76%	2.73%	2.61%	2.50%	
Marginal Efficiency Cost (MEC)		-0.009	-0.006	-0.001	-0.350	-0.317	-0.253	
Marginal Welfare Cost (MWC)	0	-0.009	-0.006	-0.001	-0.262	-0.171	0.049	
	1	0.020	0.023	0.034	-0.155	-0.084	0.064	