

**Optimal Long-Run Fiscal Policy:  
Constraints, Preferences and the Resolution of Uncertainty**

Alan J. Auerbach  
University of California, Berkeley and NBER

Kevin Hassett  
American Enterprise Institute

August 2002

We are grateful to Ethan Auerbach and Philip Oreopoulos for excellent research assistance, and to the American Enterprise Institute and the Robert D. Burch Center for Tax Policy and Public Finance for financial support.

## **Abstract**

We construct a computational dynamic stochastic overlapping generations general equilibrium model with uncertain lifetimes and explore the impact of policy stickiness (specifically, a major reform will preclude future reforms for a generation) on optimal long-run fiscal policy. Under such circumstances, entitlement reforms exhaust a valuable option to move in the future. We explore the conditions under which the gain to waiting is large enough to induce optimizing policymakers to delay reforming a suboptimal system. We also allow for the uncertainty to have ARCH characteristics and explore the impact of time-varying uncertainty on the optimality of delayed policy action.

JEL Nos. E62, H62

Alan J. Auerbach  
Department of Economics  
University of California  
Berkeley, CA 94720-3880  
(510) 643-0711  
auerbach@econ.berkeley.edu

Kevin A. Hassett  
American Enterprise Institute  
1150 Seventeenth St. N.W.  
Washington, DC 20036  
(202) 862-7157  
khassett@aei.org

## 1. Introduction

A growing share of government expenditures, in the United States and in other developed countries, is devoted to transfer programs benefiting the elderly. These programs involve substantial and very long-term commitments to members of the current labor force. Combined with aging populations, this spending pattern has led to serious questions about the viability of current fiscal policy, both in the United States and in other developed countries (e.g., Auerbach *et al* 1999). But how and when to deal with these apparent fiscal imbalances is very difficult to ascertain, even if one's social objectives are clearly specified. With considerable uncertainty about future mortality, productivity and other factors affecting fiscal balance, the plausible range of outcomes under current policy is enormous (e.g., Lee and Tuljapurkar 2001). The political difficulty of changing such well-established programs imparts a further brake to the policy reform process. Policy-makers shy away from touching the "third rail" of politics, as the U.S. social security system has been called in graphic commemoration of its popularity.

This paper investigates the nature of optimal fiscal policy, in an environment endowed with many of the important characteristics of the current fiscal climate. Using an infinite horizon, overlapping-generations model, we estimate the optimal levels of taxes, transfers and debt when there is population uncertainty and restrictions on the government's ability to change policy. We also consider how optimal policy is influenced by household attitudes toward risk, and how the prospect of resolving uncertainty may affect actions today. This last point, in particular, is of interest, because it addresses the point one frequently hears in the policy debate, that with such enormous uncertainty about future circumstances, it is better to wait until more information is available. Clearly, the desirability of waiting depends on whether the future brings a resolution of uncertainty or simply new shocks to replace the old.

This paper builds on our previous work on the subject (Auerbach and Hassett 2001). In that paper, we provided a theoretical analysis of optimal fiscal policy in an overlapping-generations model with several sources of uncertainty, and then provided simulations of the model, with and without restrictions on government policy, to estimate the effects of imposing these constraints. The current paper extends our previous analysis in a number of important ways. Most fundamentally, using an alternative simulation technique, we are now able to derive optimal government policies in an infinite horizon setting, rather than in the short-horizon case to which the previous paper's simulations were limited. This change, in turn, means that we are now able to examine the stochastic steady states under different policy rules, to trace out not only the impact effects of government policies, but also the long-run effects. In addition, we consider variations in preferences and in the environment of uncertainty. We estimate the sensitivity of optimal policy to the degree of household risk aversion, and replace the assumption of stationary shocks to life expectancy with an ARCH model in which the current environment can either be more or less uncertain than the environment expected for the future.

The remainder of the paper is organized as follows. The next section sets up the basic model, and reviews the results of the previous literature. Section 3 describes the simulation techniques we use to solve for the government's value function and optimal policy in different environments, based on the use of neural networks. Section 4 presents our basic results for optimal policy and considers the impact of political constraints, while Section 5 considers the impact of varying household risk aversion, and Section 6 the stochastic process for life expectancy. Section 7 offers some concluding comments and suggestions for future research.

## 2. Government Policy in an Overlapping Generations Model

### *A. The Basic Setup*

Throughout the paper, we analyze policy design using a two-period overlapping-generations model with no bequests, heterogeneity within each generation, or capital market imperfections. There is no growth in population per generation or productivity, and one source of uncertainty in the economy, individual life span.<sup>1</sup> Individuals supply a unit of fixed labor when young (normalized to 1 unit per generation) and are retired when old, and make one decision based on preferences, how much to save for retirement consumption and how much to spend on consumption when young. A generation's life span is unknown when it makes its saving decision, but is revealed before it consumes in old age, so that there are no accidental bequests. The government can increase welfare because the market does not offer an asset that insures the current young against the possibility that they will have an unexpectedly long lifespan. The old, however, know their actual lifespan, so they do not have to purchase an annuity to insure against long lifetimes. Since precautionary saving by the young is a function of uncertainty concerning lifetime, our framework is related to that of Yaari (1965) and Sheshinski and Weiss (1981).

We assume that production is Cobb-Douglas, obeying the expression:

$$(1) \quad Y_t = K_t^\alpha$$

---

<sup>1</sup> Productivity shocks are also important in determining optimal fiscal policy, and their inclusion would be a useful extension of the current approach. We have not included them here because of the considerable complexity they would add to the simulation problem. The complication arises because, with productivity shocks, private assets are risky and it is necessary to determine portfolio equilibrium and the premium on risky assets relative to safe assets. With only life-span uncertainty, all assets are perfect substitutes, so it is unnecessary to solve for the risk-premium or portfolio decisions.

From (1), the competitive wage rate and interest rate at  $t$  are:

$$(2a) \quad w_t = (1 - \alpha)K_t^\alpha \quad (2b) \quad r_t = \alpha K_t^{\alpha-1}$$

We initially also assume that preferences are Cobb-Douglas,

$$(3) \quad U_t = \ln C_{1t} + E_t \beta_{t+1} \ln \left( \frac{C_{2t+1}}{\beta_{t+1}} \right)$$

where  $U_t$  is the expected utility of the generation born in period  $t$ ,  $C_{1t}$  ( $C_{2t}$ ) is the consumption of the younger (older) generation in period  $t$ , and  $\beta_t$  is the length of second-period life for the older generation in period  $t$ . The form of second period utility is based on the notion that the period is really divided into  $\beta$  sub-periods, each with equal weight and consumption.

Government policy at date  $t$  is specified by the level of safe government debt at that date,  $B_t$ , and taxes on the young and old generations,  $T_{1t}$  and  $T_{2t}$ . Together, these three variables determine the transition equation for government debt:

$$(4) \quad B_{t+1} = (1 + r_t)B_t - T_{1t} - T_{2t}$$

There are various possible interpretations of the terms  $T_{1t}$  and  $T_{2t}$ . One may think of them as representing the tax and benefit components, respectively, of a public pension scheme.

However, for convenience, we assume that the second-period tax,  $T_{2t}$ , is imposed as a proportional tax on second-period capital income. Although this appears to impose a distortion, the government's optimal policy will turn out to be one that eschews distortionary taxation.<sup>2</sup>

---

<sup>2</sup> That is, in the model in which policy is unconstrained, the government is able to achieve the same allocation as it would by directly choosing the levels of consumption for the old and the young. See Auerbach and Hassett (2001) for further discussion.

Utility maximization by the young generation at date  $t$  yields the expression for first-period consumption:

$$(5) \quad C_{1t} = \frac{w_t - T_{1t}}{1 + E_t \beta_{t+1}}$$

The expression for second-period consumption is:

$$(6) \quad C_{2t} = K_t(1 + r_t) + B_t(1 + r_t) - T_{2t} = K_t(1 + r_t) + B_{t+1} + T_{1t},$$

where the last substitution follows from (4). With use made of (1), (2), (4), (5) and (6), the capital transition equation is:

$$(7) \quad K_{t+1} = K_t^\alpha (1 - \alpha) \left( 1 - \frac{1}{1 + E_t \beta_{t+1}} \right) - T_{1t} \left( 1 - \frac{1}{1 + E_t \beta_{t+1}} \right) - B_{t+1}$$

Finally, we assume that life span evolves according to the first-order autoregressive process,

$$(8) \quad \beta_{t+1} = 1 + \rho(\beta_t - 1) + \varepsilon_t,$$

where, initially, the random term  $\varepsilon$  is assumed to be i.i.d. over time. In the simulations presented below, we assume that  $\varepsilon$  has a uniform distribution.

The preceding equations provide a complete solution for the economy's evolution at each date  $t$ .

## B. Optimal Policy without Constraints

We are now in a position to maximize social welfare through the choice of  $\{T_{1t}, T_{2t}, B_{t+1}\}$  at date  $t$ , subject to expectations at date  $t$ . We assume an additively separable social welfare function with weight  $\omega_t$  assigned to generation  $t$ . Our objective, therefore, is to maximize<sup>3</sup>

$$(9) \quad W_t = \omega_{t-1} \beta_t \ln C_{2t} + \omega_t \ln C_{1t} + E_t W_{t+1}$$

subject to three state variables, the values debt,  $B_t$ , capital,  $K_t$ , and life expectancy,  $\beta_t$ , at date  $t$ .

We further assume that the government's pure rate of time preference,  $\frac{\omega_{t-1}}{\omega_t} - 1$ , is constant over time, and denote this pure discount rate as  $r^*$ . Thus, we may rewrite (9) as:

$$(9') \quad V_t = \beta_t \ln C_{2t} + \frac{1}{1+r^*} \ln C_{1t} + \frac{1}{1+r^*} E_t V_{t+1}$$

where  $V_t = W_t / \omega_{t-1}$  is the government's objective at date  $t$ , normalized by the discount factor it applies to that date's old generation.

In our earlier paper, we discussed the characteristics of the optimal solution to this problem and derived some results based on a linearization of the model around its stochastic steady state. While we will not try to summarize all these results, a couple of them are worth noting.

First, the government's optimal policy calls for the marginal utility of first-period consumption to follow a random walk, with drift of  $(1+r^*)/(1+r_{t+1}) \approx r^* - r_{t+1}$ ,

---

<sup>3</sup> There is an additional term in the objective function,  $-\omega_{t-1} \ln \beta_t$ , but this doesn't vary with government policy.

$$(10) \quad E_t \left[ \frac{1}{C_{t+1}} \right] = \left( \frac{1+r^*}{1+r_{t+1}} \right) \frac{1}{C_{1t}}$$

Note that expected growth of consumption itself depends on the level of uncertainty, because of the convexity of the function  $U_1(C_1) = 1/C_1$ . A mean-preserving spread in next period's consumption will increase expected marginal utility, because the decrease in marginal utility associated with positive deviations of consumption from its mean will be more than offset by the increase in marginal utility associated with negative consumption deviations. The uncertainty of future consumption increases its value and makes it optimal for the household to increase mean expected consumption and hence the consumption growth rate. As first noted by Leland (1968), this increase in precautionary saving in response to uncertainty is implied by the convexity of the marginal utility of consumption or, equivalently, a positive third derivative of the utility function. Any utility function with the desirable property of constant or decreasing absolute risk aversion—including the constant relative risk aversion class considered below, of which logarithmic utility is a special case—satisfies this property. Indeed, the greater the fluctuations in consumption, the greater this precautionary saving should be.

Second, a positive shock to life span should reduce current capital accumulation. With more consumption needed in the near term, less capital should be accumulated. This result may seem counterintuitive, if one thinks of longer life span as making individuals better off and more able to make transfers to other generations. But the key here is the impact on the marginal value of consumption; with longer life span, a generation must spread its resources over a longer period. Also, the more durable this positive shock, the smaller the reduction in capital accumulation – and the greater the decline in consumption – that should occur. This is because future generations are expected to be less able to provide resources for current ones.

### C. Limits on Policy Changes

Thus far, we have considered government policy when there is sufficient instrument flexibility for the government to control consumption directly. In a more realistic setting, this is unlikely to be the case.

One complication that may arise is that it may not be possible to change government's instruments in every period. This may reflect political difficulties, or implicitly the large fixed costs associated with major policy changes. To be concrete, let us suppose that the tax rates  $T_1$  and  $T_2$  cannot be changed in successive periods. Given that each period in an overlapping generations model corresponds to roughly 30 years, this restriction corresponds to the notion that major changes in, say, the Social Security or Medicare system may be possible only once every few decades. In the U.S. currently, for example, it has been almost 20 years since the Greenspan commission's recommendations lead to an overhaul of the Social Security system, despite the fact that the long run balance of the system has been a clear issue for at least a decade.

With this restriction on frequent changes, the government's problem now can be viewed as depending on an indicator variable  $d_t$ , which equals 0 if  $T_{1t-1} = T_{1t-2}$  and  $T_{2t-1} = T_{2t-2}$  and 1 otherwise. Letting  $V_t^0$  ( $V_t^1$ ) be the government's normalized objective function in period  $t$  if  $d_t = 0$  (1), we may express these functions as:

$$(11a) \quad V_t^0 = \max \left[ \beta_t \ln C_{2t} + \frac{1}{1+r^*} \ln C_{1t} + \frac{1}{1+r^*} E_t(V_{t+1}^1), \quad V_t^1 \right]$$

$$(11b) \quad V_t^1 = \beta_t \ln \bar{C}_{2t} + \frac{1}{1+r^*} \ln \bar{C}_{1t} + \frac{1}{1+r^*} E_t(V_{t+1}^0)$$

where  $\bar{C}_i$  is the value of  $C_{it}$  chosen by the household of age  $i$  when  $\{T_{1t}, T_{2t}\}$  equals  $\{T_{1t-1}, T_{2t-1}\}$ .

The government's optimization problem at date  $t$  thus consists of maximizing  $V_t^0$  if it did not change taxes in the previous period, subject to five state variables, the original set  $\{B_t, K_t, \beta_t\}$  plus lagged taxes,  $\{T_{1t-1}, T_{2t-1}\}$ .

Despite the apparent simplicity of this objective, the presence of two value functions and five state variables makes even numerical solution challenging. Using the approach of backward induction, starting from an assumed terminal period, we were able in our previous paper to arrive at solutions only for a four-period horizon, because of the growing dimensionality of the optimization problem.

### 3. Simulation Methodology

In this paper, we take a different approach to solving for the government's optimal policy. Consider first the case in which the government is unconstrained in its choice of  $T_{1t}$  and  $T_{2t}$ , regardless of its policy decisions the previous period. By the stationarity of the infinite-horizon problem, the value function will depend only on the state variables at date  $t$ , and not the date itself. That is, we can express Bellman's equation corresponding to the optimization problem as:

$$(12) \quad V(K_t, B_t, \beta_t) = \max_{T_{1t}, T_{2t}} \left[ \beta_t \ln C_{2t} + \frac{1}{1+r^*} \ln C_{1t} + \frac{1}{1+r^*} E_t V(K_{t+1}, B_{t+1}, \beta_{t+1}) \right]$$

where we use expressions (2) and (4)-(8) to express the variables on the right-hand side of (12) in terms of the state variables at date  $t$  and the instruments  $T_{1t}$  and  $T_{2t}$ .

To avoid the “curse of dimensionality,” we attempt to solve (12) directly for the stationary value function, rather than by converging to this function through backward induction. That is, we start with an initial guess of the function  $V(\cdot)$ , say  $\tilde{V}(\cdot)$ , defined over a grid of values for the vector of state variables, and calculate a new guess of the function, say  $\bar{V}(\cdot)$ , using equation (12),

$$(13) \quad \bar{V}(K_t, B_t, \beta_t) = \max_{T_{1t}, T_{2t}} \left[ \beta_t \ln C_{2t} + \frac{1}{1+r^*} \ln C_{1t} + \frac{1}{1+r^*} E_t \tilde{V}(K_{t+1}, B_{t+1}, \beta_{t+1}) \right],$$

and iterate on this expression until the function converges, i.e.,  $\bar{V}(\cdot) \equiv \tilde{V}(\cdot)$ . When this occurs, we have solved for the value function  $V(\cdot)$ . Because we iterate over the same value function over and over, this problem has the same dimensionality as that of the two-period finite horizon case, making this aspect of the problem much simpler than under the backward-induction approach. But the trade-off is that we must start with a guess of the value function, and must use a method of approximation that is flexible enough to represent a potentially very complicated function of unknown shape.

The problem of approximating an unknown function is even more challenging for the case in which government is constrained not to move every period, for then the value function  $V^0$  is an upper envelope of two functions, based on this period’s  $V^1$  and next period’s  $V^0$ . Bellman’s equations corresponding to (11a) and (11b) are:

$$(14a) \quad V^0(X_t) = \max_{T_{1t}, T_{2t}} \left[ \beta_t \ln C_{2t} + \frac{1}{1+r^*} \ln C_{1t} + \frac{1}{1+r^*} E_t (V_{t+1}^1(X_{t+1})), \quad V^1(X_t) \right]$$

$$(14b) \quad V^1(X_t) = \beta_t \ln \bar{C}_{2t} + \frac{1}{1+r^*} \ln \bar{C}_{1t} + \frac{1}{1+r^*} E_t (V^0(X_{t+1}))$$

where  $X_t = (K_t, B_t, \beta_t, T_{1t-1}, T_{2t-1})$ . Substituting the expressions for  $V^1(X_t)$  and  $V^1(X_{t+1})$  from (14b) into that for  $V^0(X_t)$  in (14a) and applying the law of iterated expectations yields an expression in the single value function  $V^0(\cdot)$  at different dates,

$$(15) \quad V^0(X_t) = \max_{T_{1t}, T_{2t}} \left[ \begin{array}{l} \beta_t \ln C_{2t} + \frac{1}{1+r^*} \ln C_{1t} + \\ \frac{1}{1+r^*} E_t \left( \beta_{t+1} \ln \bar{C}_{2t+1} + \frac{1}{1+r^*} \ln \bar{C}_{1t+1} + \frac{1}{1+r^*} V^0(X_{t+2}) \right), \\ \beta_t \ln \bar{C}_{2t} + \frac{1}{1+r^*} \ln \bar{C}_{1t} + \frac{1}{1+r^*} E_t (V^0(X_{t+1})) \end{array} \right]$$

As in the simpler, unconstrained case, we may iterate on the value function in (15) by replacing each of the value functions on the right-hand side of the equation with an initial guess  $\tilde{V}^0(\cdot)$  and solving for a new function  $\bar{V}^0(\cdot)$  on the left-hand side. Once convergence is reached, we can immediately solve for the function  $V^1(\cdot)$  by plugging the solved equation for  $V^0(\cdot)$  into the right-hand side of (14b).

Following the suggestion in Judd (1998), we use neural networks to approximate the value function for each iteration of the process just described. A neural network approximates functions using a succession of linear and nonlinear transformations of the vector of state variables, the parameters of these transformations being updated with each iteration in order to approximate the newly obtained estimate of the value function,  $\bar{V}(\cdot)$  or  $\bar{V}^0(\cdot)$ . The Appendix provides further details of the simulation methodology.<sup>4</sup>

---

<sup>4</sup> We experimented with alternative solution algorithms including the use of Chebyshev polynomials, but found that the neural network approach was more successful in approximating the value function for the more complicated problem with government policy constraints.

## 4. Basic Results

Table 1 presents the simulation results for optimal policy under both assumptions about government policy constraints, and for a variety of initial values of life expectancy,  $\beta$ , and the persistence of shocks, as represented by the autoregressive term,  $\rho$ . For all simulations in this table, the random term  $\varepsilon$  is assumed to be drawn from a distribution that is uniform over the interval  $[-0.1, 0.1]$ , and the government's pure rate of time preference,  $r^*$ , equals 2.<sup>5</sup> The initial stock of capital is set equal to its steady state value in the absence of shocks, and the initial stock of debt is set to zero.<sup>6</sup>

Let us consider first the results for the unconstrained model. In this case, it can be shown (see Auerbach and Hassett 2001) that the government's decision simplifies to a two-step problem: it allocates consumption to each generation in a given period to equal  $C_2/C_1 = (1+r^*)\beta$  (which equals  $3\beta$  for our assumption that  $r^* = 2^7$ ), and then chooses the capital stock to spread resources over time, consistent with its Euler equation relating the marginal utilities of successive young generations.

The unconstrained results are also consistent with the prediction regarding the impact of  $\rho$  on the response of consumption to a shock to current life span,  $\beta$ . Recall that as the shock to  $\beta$  is expected to be more permanent ( $\rho$  is large), we can less afford an increase in the current consumption of the elderly. Thus, consumption of the elderly should increase less, and saving

---

<sup>5</sup> Remember that each period in this life-cycle model represents a generation of perhaps 30 years, so a discount factor of 2 corresponds to a compounded annual discount factor of around 3.7 percent.

<sup>6</sup> The simulations in this table are based on the same parameterizations as those in Table 1 of Auerbach and Hassett (2001). Although the patterns of the results are reasonably similar when  $\rho=0.1$ , they are less so when  $\rho=0.9$ , suggesting that the approximation error involved in using a finite horizon is more significant when shocks are persistent.

<sup>7</sup> That this relationship does not hold exactly in each simulation reflects rounding and slight interpolation error.

should be higher, when  $\rho$  is high. Indeed, this is quite evident when we compare the changes in  $C_2$  as  $\beta$  increases for  $\rho = .1$  and  $\rho = .9$ .

Another result worth noting is that, as the level of uncertainty increases, so does optimal saving. For a starting value of  $\beta=1$ , there is no expected trend in life span. The larger  $\rho$  is, however, the more persistence is expected for positive or negative life-span shocks, and hence these shocks have more important effects. The increased risk at  $\beta = 1$  associated with an increase in  $\rho$  from 0.1 to 0.9 raises the optimal level of  $K_{t+1}$ . Indeed, for  $\rho = 0.9$ , the optimal capital stock is higher in all cases than in the deterministic steady-state value. Thus, with sufficient uncertainty, it is optimal to accumulate additional capital, even if current needs are unusually high (i.e.,  $\beta > 1$ ).

The simulations for the constrained model show the optimal policy choices, if the government is able to act and if it is optimal to do so. Whether moving is optimal, in turn, will depend on the lagged values of  $T_1$  and  $T_2$ .

As one would expect, anticipating a lack of flexibility leads the government to sacrifice current consumption to provide for the future. In all cases, both  $C_1$  and  $C_2$  are lower than in the corresponding unconstrained specification. In percentage terms, the reduction in  $C_1$  is larger than that in  $C_2$ . Intuitively, this reflects the fact that the young may be expected to “recover” part of the government’s extra precautionary saving when they are old. Also, the higher is current life span,  $\beta$ , the greater is amount of added precautionary saving. The intuition for this is that, when the current value of  $\beta$  is high, future generations are, in general, expected to be worse off – they are also expected to have a high  $\beta$ , and hence less consumption per unit time for a given level of resources. With their higher associated marginal utility of consumption, the welfare cost of having the “wrong” fiscal policy in place when the government cannot move will be greater.

This leads the government to exact a greater sacrifice from current generations, particularly the young.

In Figures 1 and 2, we present estimates of the range of inaction in the policy-constrained version of the model, for the cases of  $\beta = 1$  and  $\rho = 0.9$  and  $0.1$ , respectively. In each figure, the star represents the optimal value of the pair  $\{T_1, T_2\}$ , corresponding to the values in bold in Table 1. These values of the lagged tax rates would yield the highest value of the current value function. As the lagged tax rates move away from this point, it is more and more costly for the government to maintain its option to move next period. When the boundary of the inaction range is passed, it becomes optimal for the government to change policy in the current period, extinguishing the option to move in the next period.

There are a number of interesting patterns in these two figures. First, the area of the inaction range is substantially bigger for  $\rho = .9$  than for  $\rho = .1$ . This difference suggests that there is much greater value in waiting to adjust policy if next period's information is more "permanent." Another interesting pattern in these results is that the optimal values of  $T_1$  and  $T_2$  each lie much closer to the upper boundary of their respective inaction ranges than to the lower boundary. This is especially true for  $\rho = 0.1$ . We may infer from this that suboptimal current tax rates are more costly if they are too high than if they are too low.

The intuition for this finding is as follows, and depicted in Figure 3. Mistakes in either direction have a roughly symmetric impact on the welfare of future generations; because these deviations in resources from the optimal policy are spread over a large number of generations, the deviation for any one generation is very small and hence valued at an approximately constant marginal utility for movements in either direction. This is depicted by the dashed horizontal line in Figure 3, characterizing the valuation by future generations of movements away from the

optimal distribution of resources at point  $a^*$ . For current generations, however, the resource fluctuations from optimal policy are large, so that the marginal utility applicable to added resources will be lower than the marginal utility applicable to resources taken away, as depicted by the downward sloping line in Figure 3. There is a net social cost of moving in either direction from point  $a^*$ , but, for equal movements in either direction (to point  $a_l$  or point  $a_h$ ) the cost of being too generous to current generations—the area  $H$  in the figure—is smaller than the cost of being too harsh—point  $L$  in the figure—as long as the marginal valuation curve of current generations is convex. This convexity condition on marginal utility, i.e., that the utility function's third derivative is positive, is precisely the same condition discussed above that causes uncertainty to increase precautionary saving. Hence, a factor that tends to increase saving in the unconstrained case and, presumably, in the constrained case as well, also works in the opposite direction in the constrained case, causing the government to be more likely to cut taxes than to raise them when current policy is not optimal.

Thus, when taking action, the government is likely to engage in more precautionary saving in the constrained model, but the timing of its actions tend to favor less saving. Taking into account both periods of action and periods of inaction, taxes could be higher or lower in the constrained world than in the unconstrained world. We address this question below, but consider first the impact of the results presented thus far of the government's discount rate.

Table 2, and Figures 4 and 5, provide simulations and inaction ranges comparable to those in Table 1 and Figures 1 and 2, but for a government discount rate of  $r^* = 1$  instead of  $r^* = 2$ . Although the patterns are generally similar to those already discussed, there are some interesting and intuitive differences. Note, for example, what happens to capital accumulation in the constrained case for  $\rho = 0.9$ , as  $\beta$  rises. Previously, we found that capital accumulation

should fall, because the current elderly need more resources. The high value of  $\rho$  mitigates this, because the elderly in the near future will also need more resources. Here, though, the government actually chooses to accumulate *more* when  $\beta$  is high. The reason is that  $\beta$  could go even higher, and a government constrained not to move in the future, with a low discount rate, perceives the need to provide for this eventuality.

Another effect of the lower discount rate is the expansion of the government's inaction range. This is most evident when the range is large to begin with, for the case of  $\rho = 0.9$  (Figure 4 versus Figure 1). Imbued with a greater concern for future generations, the government values the option to move in the future more highly.

We now return to the question raised earlier, whether the imposition of policy constraints reduces saving and capital accumulation in the long run. As discussed, there are offsetting factors, as government engages in more precautionary saving when it changes policy, but is more likely to change policy when saving is too high, rather than too low. Figures 6 and 7 provide comparisons of steady-state capital stock distributions for constrained and unconstrained policy action, for the case of persistent life-span shocks ( $\rho = 0.9$ ). Figure 6 is for the case of a high government discount rate ( $r^* = 2$ ), corresponding to Figure 1 and the left side of Table 1. Figure 7 is for the case of a low discount rate ( $r^* = 1$ ), corresponding to Figure 4 and the left side of Table 2. In each of the new figures, the distribution of the capital stock is tighter when there are no policy constraints, as one would expect. When  $r^* = 1$ , the capital stock distribution has a higher mean as well, reflecting the precautionary saving motive. But, when  $r^* = 2$ , the asymmetry of the inaction range dominates, and the capital stock is actually lower, on average, with constraints on government action.

## 5. The Impact of Risk Aversion

Thus far, we have considered variations in initial conditions, government constraints and the environment of uncertainty, but not in household preferences and, by extension, the risk aversion implicit in government policy making. Clearly, risk aversion is an important element of the problem, so it is useful to see how variations in risk aversion affect optimal government policy. In place of the Cobb-Douglas preferences assumed thus far, we now consider the general constant elasticity of substitution (CES) preferences,

$$(16) \quad U_t = \frac{1}{1-\gamma} C_{1t}^{1-\gamma} + \frac{1}{1-\gamma} E_t \left[ \beta_{t+1} \left( \frac{C_{2t+1}}{\beta_{t+1}} \right)^{1-\gamma} \right]$$

with coefficient of risk aversion  $\gamma$ . When  $\gamma = 1$ , preferences are Cobb-Douglas.

For general CES preferences, the first-order condition for first-period consumption becomes, in place of (5),

$$(5') \quad C_{1t} = \frac{w_t - T_{1t}}{1 + x_{t+1}}$$

where

$$(17) \quad x_t = \left\{ E_t \left[ \beta_{t+1}^\gamma (1 + r_{t+1}^n)^{1-\gamma} \right] \right\}^{1/\gamma}$$

and  $r_{t+1}^n$  is the after-tax rate of return in period  $t+1$ . When  $\gamma=1$ ,  $x_{t+1}$  reduces to  $E_t \beta_{t+1}$  and (5') reduces to (5). Otherwise, the distribution of  $r_{t+1}^n$  must be known to calculate  $C_{1t}$ . As this complication would make the simulation exercise considerably more complicated, we approximate  $r_{t+1}^n$  with  $r^*$  in (17). The approximation error involved in doing so is likely to be

small, given the fairly tight distribution for the capital stock in the simulations we are considering, with  $r^*=2$  and  $\rho=0.9$  and government unconstrained (see Figure 5).

Table 3 presents simulations for three different values of the degree of risk aversion,  $\gamma$ , 1.1, 3.0 and 5.0<sup>8</sup>, and two values for the interval of the uniform distribution. The interval used above,  $[-0.1, 0.1]$ , has length 0.2. To consider the effects of risk and risk aversion, we utilize two extreme values for the interval length, 0.1 and 1.1, corresponding to the uniform distributions  $[-.05, 0.05]$  and  $[-0.55, 0.55]$ .

As the table shows, the impact of risk on behavior is very small for a moderate degree of risk aversion. When  $\gamma=1.1$ , there is a slight decline in the optimal level of  $C_2$  caused by increased risk, and a decline in the optimal level of  $C_1$  so small that it is not perceptible to the third decimal place. Thus, although the point emphasized above about the pure impact of risk on precautionary saving is true, it is not particularly significant here, compared to the impact, say, of the government's inability to change policy. The importance of risk becomes more important as the degree of risk aversion grows. For the empirically large (but not implausible) risk aversion coefficient of 5, first- and second-period consumption levels are noticeably lower for the higher-risk simulation. This is consistent with the fact that, as risk-aversion grows, so (for the CRRA utility function) does the strength of the precautionary saving motive, as measured by the degree of "prudence" (Kimball 1990).<sup>9</sup>

---

<sup>8</sup> We use  $\gamma=1.1$  instead of 1.0 to simplify the programming, which would require additional statements to handle exact Cobb-Douglas preferences.

<sup>9</sup> Kimball defines "prudence" as minus the ratio of the third utility derivative to the second utility derivative and shows that the precautionary saving motive increases with this measure, which equals  $(1+\gamma)/C$  for risk-aversion coefficient  $\gamma$ . Thus, for a given value of consumption, prudence increases with risk aversion.

## 6. Why Wait? Changes in the Degree of Uncertainty

Up until this point, we have analyzed the impact on optimal policy of the level of uncertainty, but the distribution of lifetime uncertainty has been fixed for each simulation. This assumption ignores a second possible source of gain to policy delay: the possibility that the government can learn more about the underlying parameters over time. As many have documented, programs like Social Security and Medicare appear unsustainable and in need of substantial change. A commonly heard argument, though, is that for programs like these that are subject to enormous uncertainty, it is impossible and hence unwise to take any significant action immediately, when so little is known. While this argument may ring true, there is little in the analysis thus far to support it. Indeed, there are two factors working in precisely the opposite direction. First, as just discussed, risk aversion should lead to precautionary saving, and more risk should lead to more saving. Second, to the extent that government may face limits on its ability to act, any action it takes should involve even more stringent policy measures. It is true, as we have seen, that constraints on policy also leave room for ranges of inaction when circumstances are not far enough away from optimal, and that the inaction range expands with uncertainty. But is there more to the story than this?

One possibility involves the distinction between risk and uncertainty that goes back to Knight (1921) and has been considered anew in the recent literature (see, e.g., Schmeidler (1989)) The argument may be that the parameters of the distributions of the stochastic terms are unknown, but that the government is in the process of learning them. Without departing from the standard paradigm for analyzing optimal behavior under uncertainty, though, one may make sense of the argument for waiting if a resolution of the existing uncertainty may be expected. That is, it makes little sense to wait for today's uncertainty to be resolved, if more is on the way

tomorrow. But if today's uncertainty is greater than that anticipated for the future, waiting might make sense.

To some extent, an ARCH process that allows the parameters of the mortality distribution to evolve over time in a predictable way parametrically characterizes such a situation, but in an admittedly incomplete manner since it abstracts from the learning process and occurs within a well defined probability space. To make this intuition concrete, we adapt the model considered thus far, to allow changes in the degree of life-span uncertainty over time. We continue to assume that life span is stochastic, and described by the first-order autoregressive process in (8). But now, instead of assuming that the random terms  $\varepsilon_t$  are i.i.d., we assume that they are determined by an ARCH(1) process, in which the variance of the distribution itself varies over time around some long-run value.

Let  $v_t$  be a random draw from the uniform  $[-0.5, 0.5]$  distribution at date  $t$ , and let  $m$  be the term that multiplies this to determine the actual error term  $\varepsilon_t$  and the range of the corresponding uniform distribution. For example, for the range equal to 1.1 used in Table 3,  $m \equiv 1.1$ . Now, suppose that  $m$  is time varying, with  $\varepsilon_t = m_{t-1} v_t$ .<sup>10</sup> We specify that  $m$  evolves according to:

$$(18) \quad m_t^2 = (1 - \alpha)\bar{m}^2 + \alpha m_{t-1}^2 v_t^2 / \sigma_v^2$$

where  $\sigma_v^2$  is the (constant) variance of the innovation  $v$ . In this expression, the unconditional expectation of  $m_t^2$  is  $\bar{m}^2$ , and the expectation conditional on  $m_{t-1}^2$  is  $(1 - \alpha)\bar{m}^2 + \alpha m_{t-1}^2$ . Thus,  $m_t^2$  follows a first-order AR process. For  $\alpha = 0$ , the process reduces to the one previously

---

<sup>10</sup> We choose the timing of  $m$  in this way because it is known as of the end of period  $t-1$ , while  $v_t$  is not.

considered, with  $m \equiv \bar{m}$ . The larger  $\alpha$  is, the more persistence there is in the impact of shocks to the variance itself. Given that  $\sigma_v^2 = \int_{-.5}^5 v^2 dv = \frac{1}{3}(.5^3 + .5^3) = \frac{1}{12}$ , equation (18) becomes:

$$(19) \quad m_t^2 = (1 - \alpha)\bar{m}^2 + 12\alpha m_{t-1}^2 v_t^2$$

Using the ARCH specification, we can contrast two situations in which current conditions are the same. In the first, the variance is  $m_{t-1}$ , and expected to stay at this value. In another, the current variance is  $m_{t-1}$ , but  $m$  follows an ARCH(1) process and is expected to fall in the future, because  $m_{t-1} > \bar{m}$ . In this case,  $m_{t-1}$  is an additional state variable of the government's optimization problem.

Table 4 provides simulations that allow us to measure the impact of this distinction. The first three columns repeat the last three from Table 3, with  $r^* = 2$ ,  $\beta = 1$ ,  $\rho = 0.9$ , and  $m \equiv 1.1$ . The next three columns present simulations for the same current values of these parameters, but with  $m$  determined by an ARCH(1) process with  $\alpha = 0.2$ . In all simulations, the capital stock is increased by policy from its initial level, so the situation corresponds to one in which “fiscal discipline” is needed. The results indicate that the intuition provided above holds, but only for a sufficiently risk-averse population.

When  $\gamma = 5$ , the prospect of reduced future variance lowers the perceived need for precautionary saving. Consumption by both young and old is higher than under the constant-variance assumption, and capital accumulation is substantially lower. This effect, however, is substantially weaker for the moderate degree of risk aversion,  $\gamma = 3$ . For  $\gamma = 1.1$ , the effect is actually reversed. Although the model is sufficiently complicated that no simple intuition seems to apply, we conjecture that this reversal has something to do with a second aspect of the ARCH

process. While  $m$  is expected to fall in the long run, it also varies with actual shocks in the short run. In particular, if there is an unexpected increase in life span – which is bad news for fiscal policy – this will also induce an increase in  $m$ . Thus, negative fiscal shocks will be associated with increased uncertainty, and this positive relationship may lead to more precautionary saving, offsetting the effect of the long-run decline in  $m$  that is anticipated. In summary, unless the resolution of uncertainty is itself certain, its possibility does not necessarily dictate delay in remedial fiscal actions.

## 7. Conclusions

Using numerical simulations, we have considered the nature of optimal long-run fiscal policy in a variety of environments. Although it may be helpful to think of these results as relating specifically to optimal Social Security policy, the model is more general and abstracts from the specific institutions of Social Security.<sup>11</sup>

Our results indicate that limits on the ability of government to readily adjust policy exert a significant impact on the nature of optimal policy. These limits lead to more precautionary saving when policy changes, but also to an asymmetry in policy *responsiveness* that favors current generations. Thus, in the long run, policy constraints can contribute either to more saving or less, and make the distribution of the long-run capital stock much wider. We also find that the resolution of uncertainty may lead to circumstances where it is optimal to “wait and see,” but that powerful precautionary motives work in the opposite direction.

---

<sup>11</sup> For a good recent analysis of optimal Social Security policy with respect to risk-sharing in a model that incorporates many of Social Security’s institutional characteristics, see Bohn (2001). Bohn’s analysis relates most closely to that of our unconstrained model, which we explored more fully in Auerbach and Hassett (2001).

## Appendix: Neural Networks

Figure 8 provides a schematic representation of the neural network we use here, for the simple unconstrained problem in which there are three state variables, the capital stock,  $K$ , the level of debt,  $B$ , and life span,  $\beta$ .

The neural network approximates a function using a sequence of linear and nonlinear transformations of the state variables. In the first stage of each round, a pre-specified number, say  $m$ , of linear combinations of the state variables are constructed, using the weights  $w_{11} \dots w_{3m}$ .<sup>12</sup> Each linear combination corresponds to a “node” of the first “hidden” layer of the neural network, and is passed through a nonlinear transformation, denoted  $f$  in Figure 8. If there is just one hidden layer, then the resulting scalars,  $x_1 \dots x_m$ , are aggregated using linear weights  $u_i$  to form an approximation of the value function evaluated at the particular inputs,  $V(\beta, K, B)$ . If there is a second hidden layer, as depicted in Figure 8, then there is an additional step before this final aggregation, in which the initial process is repeated; if there are  $n$  nodes in the second hidden layer, we form  $n$  linear combinations of the outputs of the first hidden layer,  $x_1 \dots x_m$ , and pass these linear combinations through another nonlinear transformation, denoted  $g$  in Figure 8. The function depicted in the figure is:

$$(A1) \quad \hat{V}(\beta, K, B) = \sum_{i=1}^n u_i g \left( \sum_{j=1}^m v_{ji} f \left[ \sum_{k=1}^3 w_k z_{kj} \right] \right)$$

where  $\{z_1, z_2, z_3\} = \{\beta, K, B\}$ . The unknown parameters of this function, the vectors  $\mathbf{u}$ ,  $\mathbf{v}$ , and  $\mathbf{w}$ , are chosen to minimize the sum of squared residuals,  $(\hat{V}(\beta, K, B) - \bar{V}(\beta, K, B))^2$ , over the points in the state variable grid.

---

<sup>12</sup> The linear transformations also include a constant.

The flexibility of the neural network increases with the number of hidden layers and the number of nodes in each layer, but so does computational time. Also with too many nodes and/or layers than are necessary for a close approximation, we lose degrees of freedom and may hamper our ability to forecast the function for points not on the grid. Thus, we experiment with different specifications, adding layers or nodes until sufficient flexibility is achieved, and experimenting with different nonlinear transformations as well. For the unconstrained model and Cobb-Douglas preferences, we used one hidden layer and six nodes. For the unconstrained model with CES preferences, and for the ARCH model, we used one hidden layer and 10 nodes. For the most complex simulations, those based on the constrained model, we added a second hidden layer, with 16 nodes in the first layer and seven nodes in the second layer. For each specification, we used the hyperbolic tangent sigmoid transfer function

$$(A2) \quad \frac{2}{1 + e^{-2a}} - 1$$

for the first hidden layer. For the second hidden layer in the constrained simulations, we used the log sigmoid transfer function

$$(A3) \quad \frac{1}{1 + e^{-a}}$$

All simulations were carried out in MATLAB.

## References

- Auerbach, Alan J. and Kevin Hassett, 2001, "Uncertainty and the Design of Long-Run Fiscal Policy," in A. Auerbach and R. Lee, eds., *Demographic Change and Fiscal Policy*. Cambridge, UK: Cambridge University Press, 73-92.
- Auerbach, Alan J., Laurence J. Kotlikoff, and Willi Leibfritz, eds., 1999, *Generational Accounting Around the World*. Chicago: University of Chicago Press.
- Bohn, Henning, 2001, "Social Security and Demographic Uncertainty: The Risk Sharing Properties of Alternative Policies," in J. Campbell and M. Feldstein, eds., *Risk Aspects of Investment-Based Social Security Reform*. Chicago: University of Chicago Press, 203-241.
- Judd, Kenneth L., 1998, *Numerical Methods in Economics*. Cambridge: MIT Press.
- Kimball, Miles S., 1990, "Precautionary Saving in the Small and the Large," *Econometrica*, January, 53-73.
- Knight, Frank, 1921, *Uncertainty and Profit*. Boston, Houghton Mifflin Company.
- Lee, Ronald, and Shripad Tuljapurkar, 2001, "Population Forecasting for Fiscal Planning: Issues and Innovations," in A. Auerbach and R. Lee, eds., *Demographic Change and Fiscal Policy*. Cambridge: Cambridge University Press, 7-57.
- Leland, Hayne, 1968, "Saving and Uncertainty: the Precautionary Demand for Saving," *Quarterly Journal of Economics*, August, 465-73.
- Schmeidler, David, 1989 "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, May, 571-87.
- Sheshinski, Eytan, and Yoram Weiss, 1981, "Uncertainty and Optimal Social Security Systems," *Quarterly Journal of Economics*, May, 189-206.
- Yaari, Menahem E., 1965, "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," *Review of Economic Studies*, April, 137-150.

**Table 1. Optimal Policy Rules**  
( $r^* = 2$ )

<b>Unconstrained</b>						
$\rho$	0.9	0.9	0.9	0.1	0.1	0.1
$\beta$	0.7	1.0	1.3	0.7	1.0	1.3
$V$	-2.096	-2.534	-2.939	-2.203	-2.533	-2.833
$C_1$	0.157	0.124	0.100	0.154	0.124	0.104
$C_2$	0.332	0.368	0.393	0.322	0.373	0.403
$K_{t+1}$	0.065	0.064	0.062	0.079	0.059	0.048
$T_1$	0.099	0.124	0.144	0.069	0.124	0.160
$T_2$	-0.149	-0.184	-0.209	-0.138	-0.189	-0.219
$B_{t+1}$	0.050	0.060	0.065	0.070	0.065	0.060
<b>Constrained</b>						
$\rho$	0.9	0.9	0.9	0.1	0.1	0.1
$\beta$	0.7	1.0	1.3	0.7	1.0	1.3
$V$	-2.097	-2.549	-2.968	-2.216	-2.550	-2.858
$C_1$	0.139	0.108	0.084	0.132	0.105	0.087
$C_2$	0.325	0.357	0.382	0.319	0.363	0.389
$K_{t+1}$	0.090	0.090	0.089	0.104	0.087	0.079
$T_1$	0.130	<b>0.156</b>	0.181	0.111	<b>0.162</b>	0.194
$T_2$	-0.142	<b>-0.173</b>	-0.199	-0.135	<b>-0.180</b>	-0.205
$B_{t+1}$	0.011	0.018	0.018	0.024	0.018	0.011

Initial values:  $K_t = .06$ ;  $B_t = 0$

**Table 2. Optimal Policy Rules**  
( $r^* = 1$ )

<b>Unconstrained</b>						
$\rho$	0.9	0.9	0.9	0.1	0.1	0.1
$\beta$	0.7	1.0	1.3	0.7	1.0	1.3
$V$	-2.806	-3.294	-3.729	-2.987	-3.292	-3.555
$C_1$	0.260	0.212	0.176	0.249	0.212	0.179
$C_2$	0.363	0.418	0.458	0.347	0.418	0.474
$K_{t+1}$	0.165	0.158	0.153	0.192	0.158	0.135
$T_1$	0.023	0.049	0.074	-0.017	0.049	0.109
$T_2$	-0.048	-0.103	-0.143	-0.032	-0.103	-0.159
$B_{t+1}$	0.024	0.055	0.070	0.050	0.055	0.050

---

<b>Constrained</b>						
$\rho$	0.9	0.9	0.9	0.1	0.1	0.1
$\beta$	0.7	1.0	1.3	0.7	1.0	1.3
$V$	-2.821	-3.301	-3.758	-2.991	-3.300	-3.572
$C_1$	0.253	0.200	0.159	0.235	0.196	0.172
$C_2$	0.362	0.406	0.444	0.349	0.406	0.450
$K_{t+1}$	0.173	0.182	0.184	0.204	0.185	0.166
$T_1$	0.035	<b>0.073</b>	0.111	0.010	<b>0.080</b>	0.124
$T_2$	-0.047	<b>-0.091</b>	-0.129	-0.034	<b>-0.091</b>	-0.135
$B_{t+1}$	0.011	0.018	0.018	0.024	0.011	0.011

Initial values:  $K_t = .1575$ ;  $B_t = 0$

**Table 3. The Impact of Risk Aversion**(Unconstrained model;  $r^* = 2$ ,  $\rho = .9$ ,  $\beta = 1.0$ )

range	0.1	0.1	0.1	1.1	1.1	1.1
$\gamma$	1.1	3.0	5.0	1.1	3.0	5.0
$V$	-22.73	-15.04	-126.42	-22.74	-15.65	-136.95
$C_1$	0.132	0.200	0.218	0.132	0.197	0.211
$C_2$	0.356	0.286	0.271	0.353	0.286	0.266
$K_{t+1}$	0.067	0.069	0.066	0.070	0.072	0.078
$T_1$	0.120	0.075	0.063	0.120	0.079	0.073
$T_2$	-0.172	-0.102	-0.087	-0.169	-0.102	-0.082
$B_{t+1}$	0.052	0.027	0.024	0.050	0.023	0.009

Initial values:  $K_t = .06$ ;  $B_t = 0$

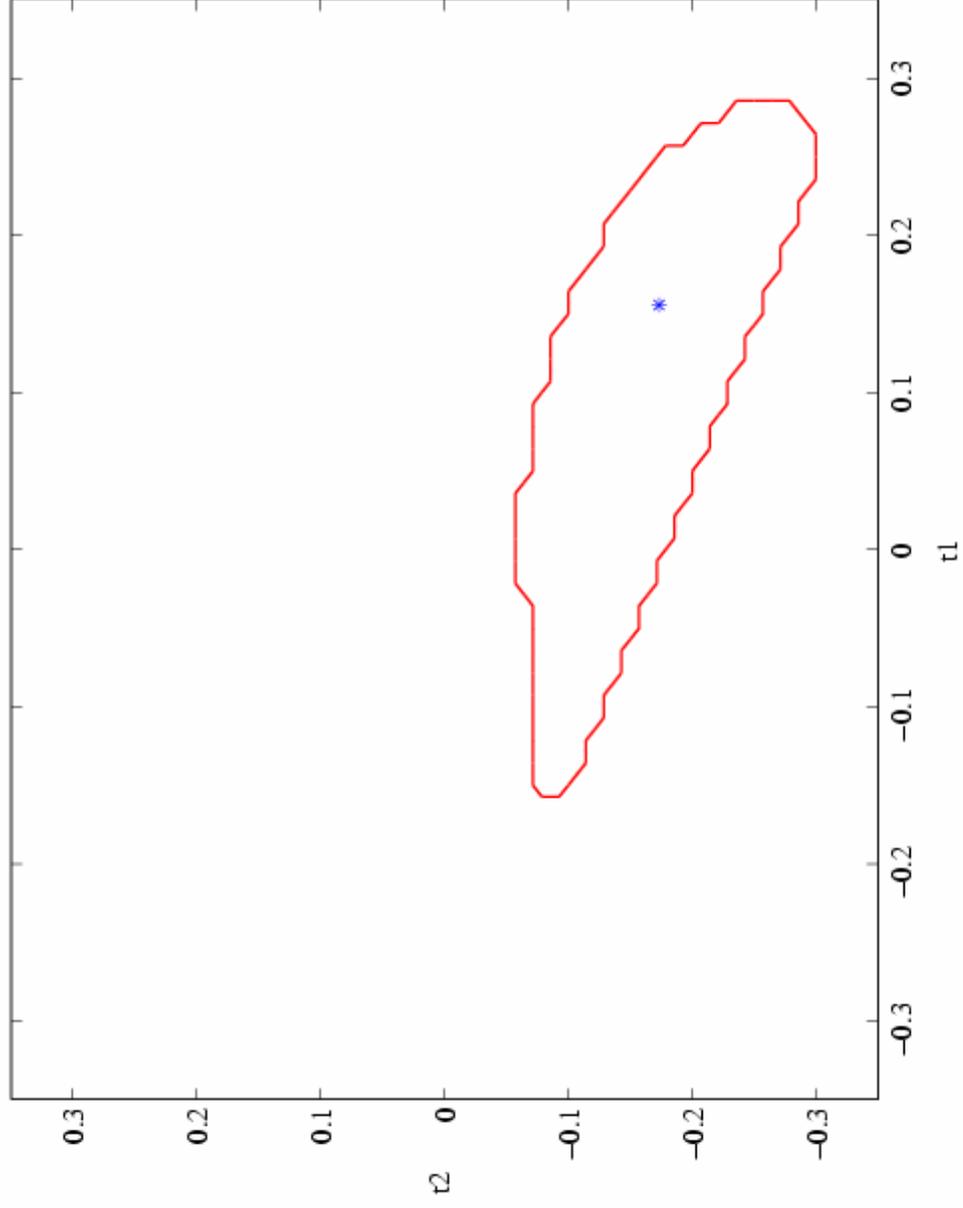
**Table 4. Changing Uncertainty: ARCH Model**  
(Unconstrained model;  $r^* = 2$ ,  $\rho = .9$ ,  $\beta = 1.0$ , Initial Range = 1.1)

range $\gamma$	Constant Range			ARCH(1)		
	1.1	3.0	5.0	1.1	3.0	5.0
$V$	-22.74	-15.65	-136.95	-22.73	-15.31	-123.62
$C_1$	0.132	0.197	0.211	0.131	0.200	0.223
$C_2$	0.353	0.286	0.266	0.351	0.286	0.278
$K_{t+1}$	0.070	0.072	0.078	0.074	0.069	0.054
$T_1$	0.120	0.079	0.073	0.123	0.071	0.046
$T_2$	-0.169	-0.102	-0.082	-0.167	-0.103	-0.095
$B_{t+1}$	0.050	0.023	0.009	0.044	0.032	0.048

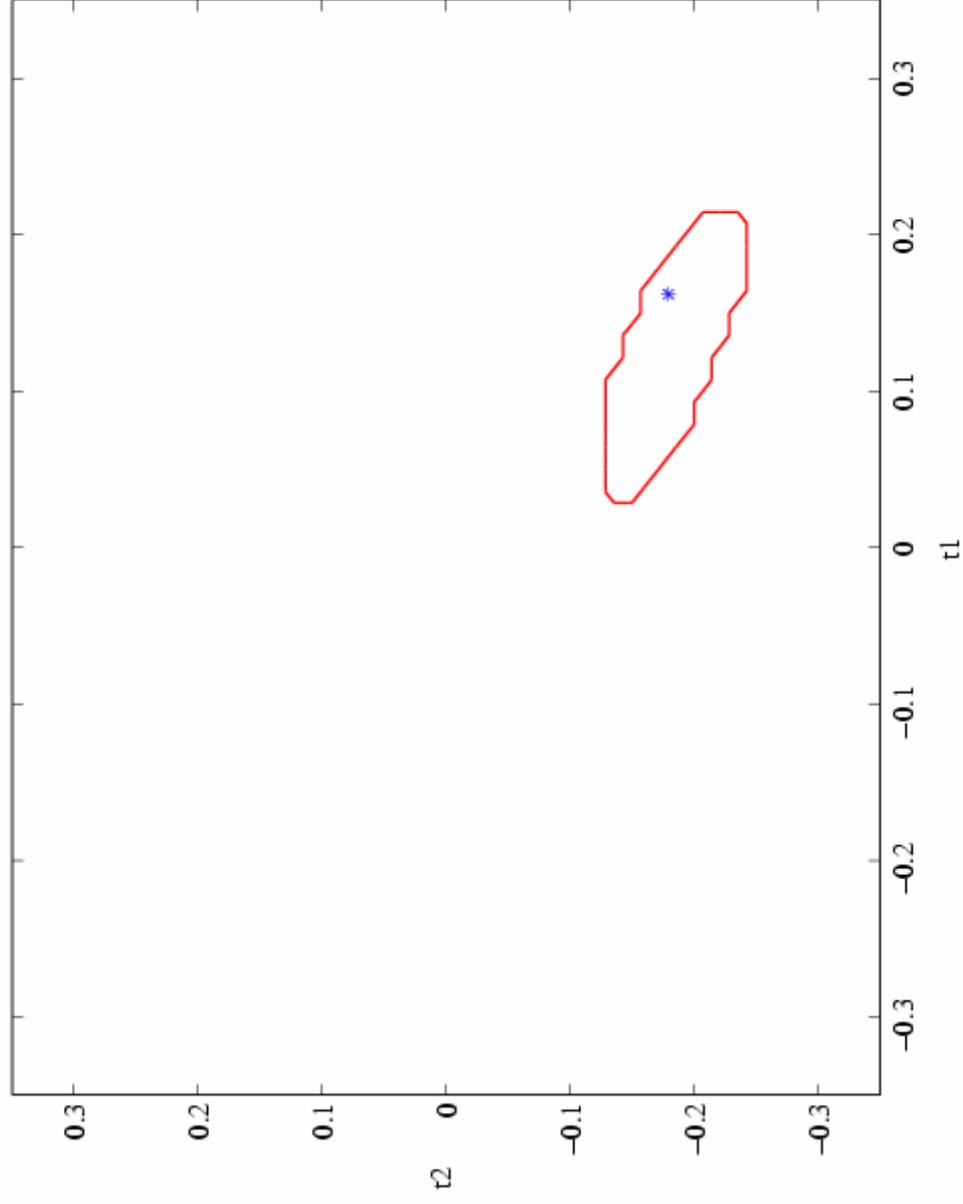
Initial values:  $K_t = .06$ ;  $B_t = 0$ ; ARCH parameters:  $\alpha = 0.2$ ,  $\bar{m} = 0.1$

**Figure 1. Inaction Range**

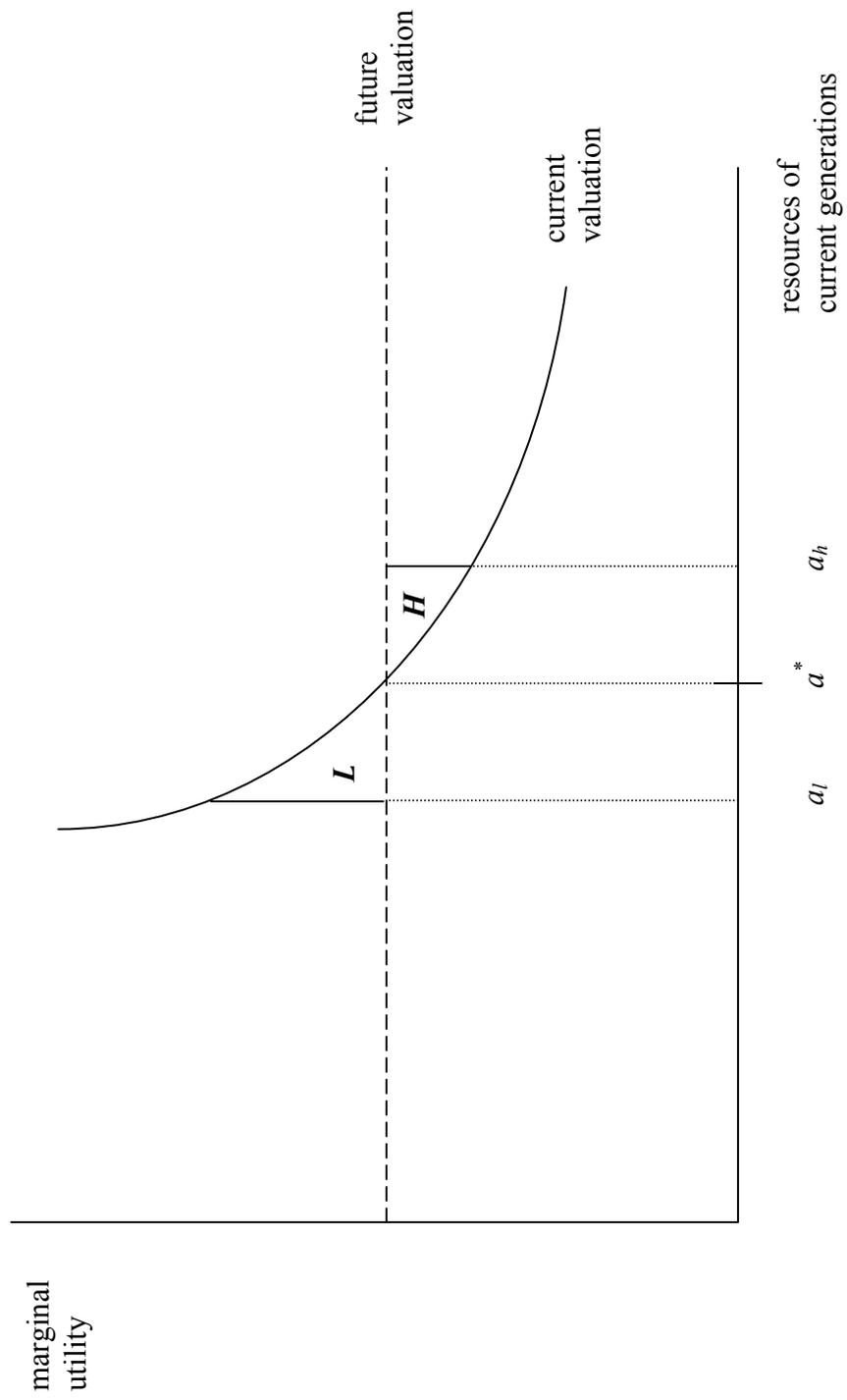
$(r^* = 2, \beta = 1.0, \rho = 0.9)$



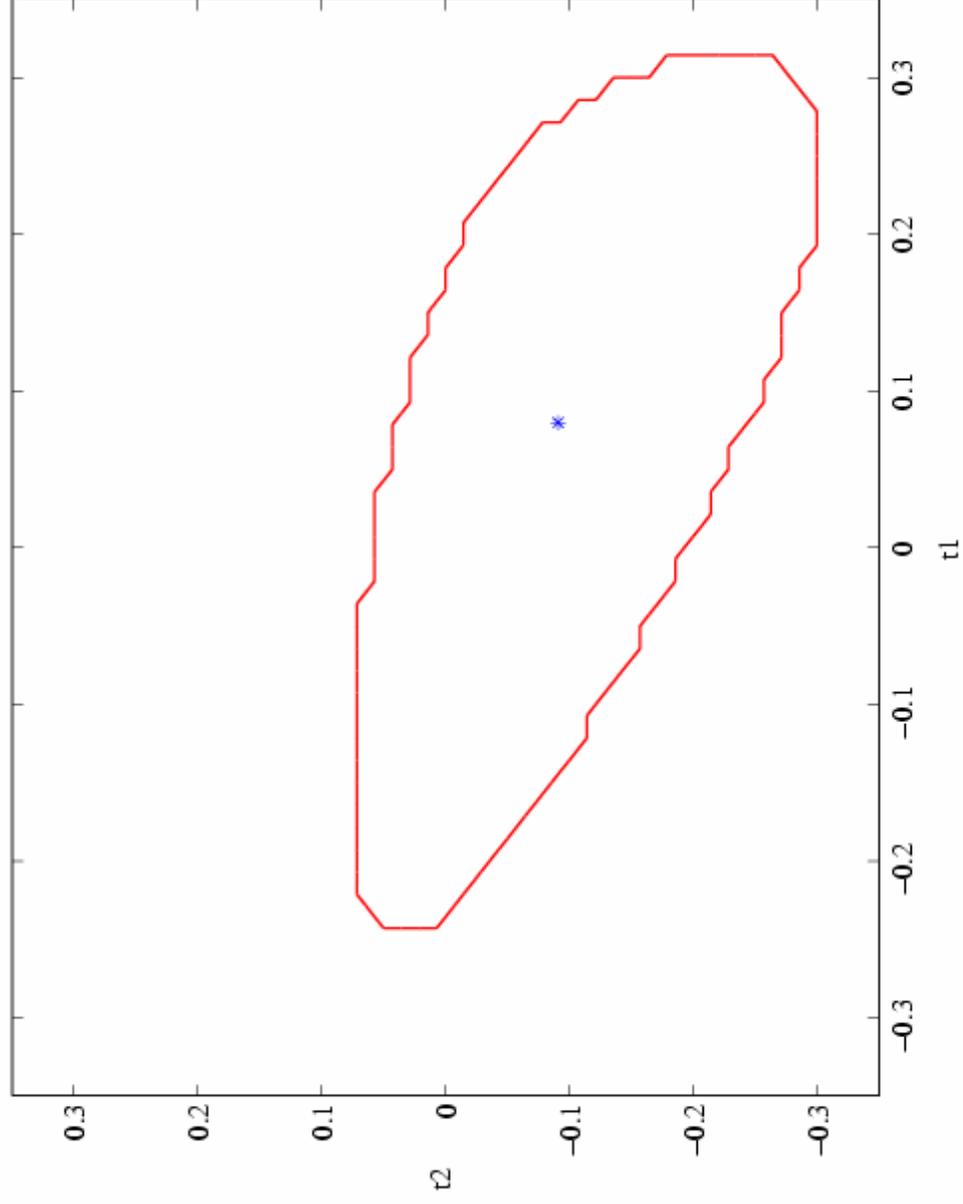
**Figure 2. Inaction Range**  
( $r^* = 2, \beta = 1.0, \rho = 0.1$ )



**Figure 3. The Asymmetry of Inaction Ranges**

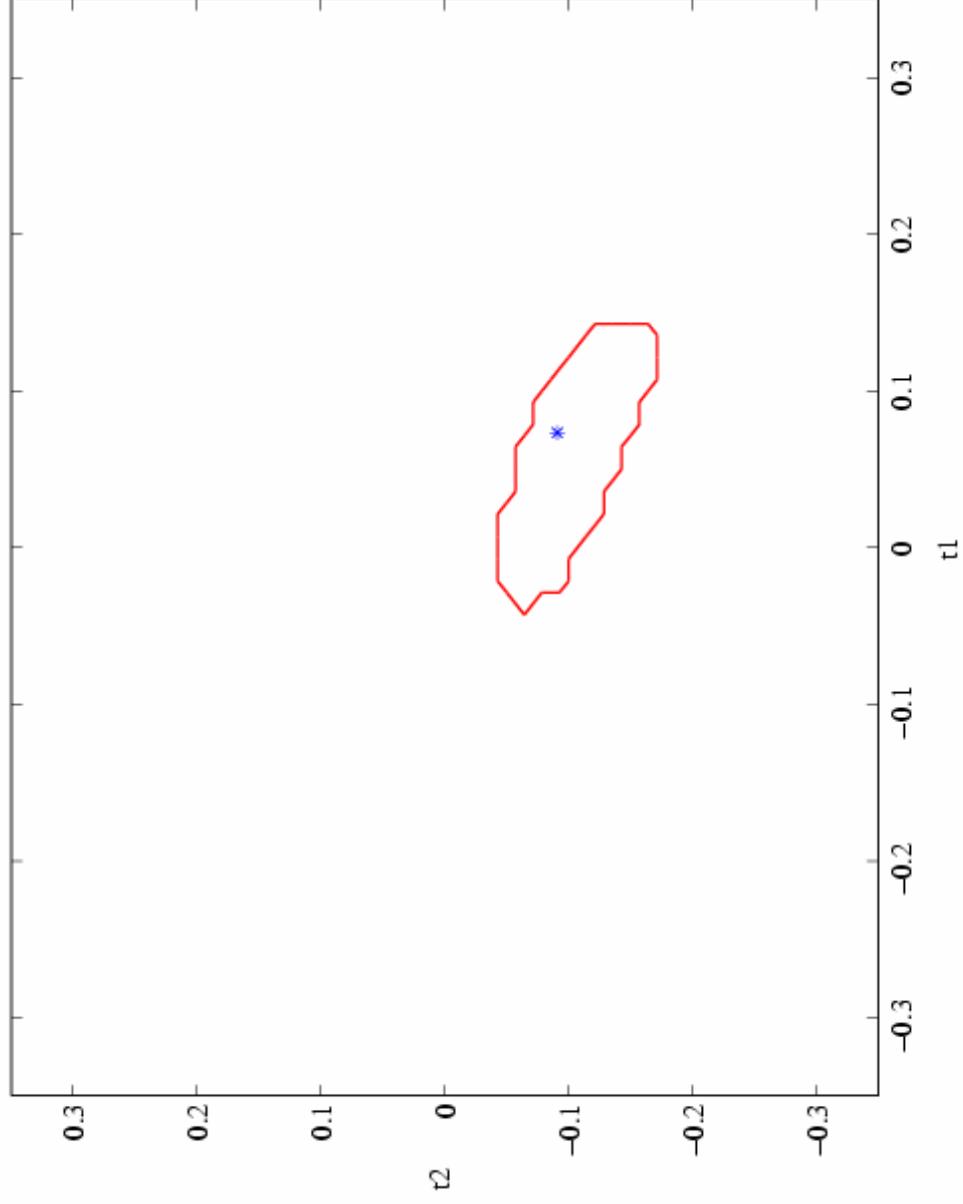


**Figure 4. Inaction Range**  
( $r^* = 1, \beta = 1.0, \rho = 0.9$ )



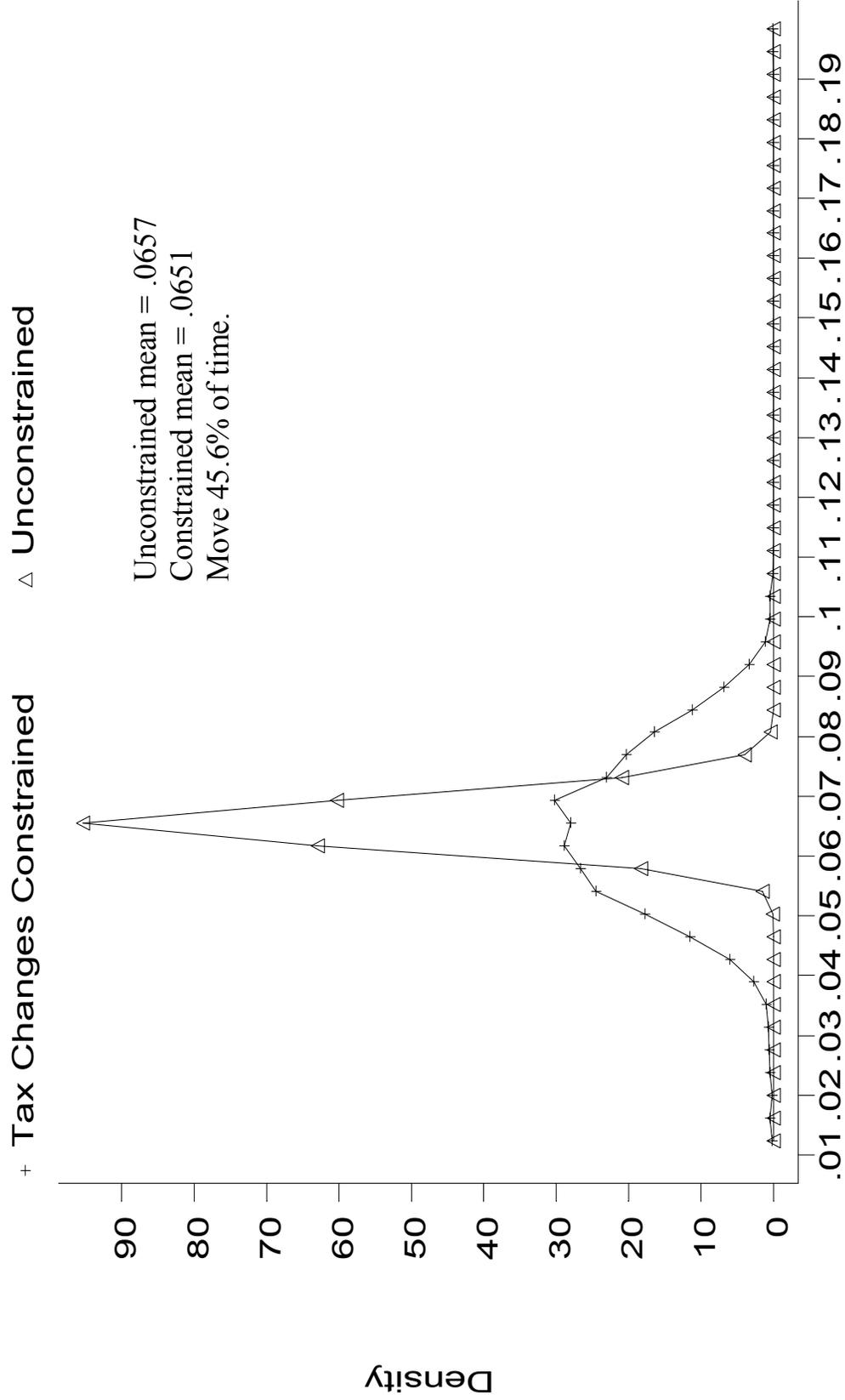
**Figure 5. Inaction Range**

$(r^* = 1, \beta = 1.0, \rho = 0.1)$



**Figure 6. Steady State Capital Stock Distribution**

$(r^* = 2, \rho = 0.9)$



**Figure 7. Steady State Capital Stock Distribution**

$(r^* = 1, \rho = 0.9)$

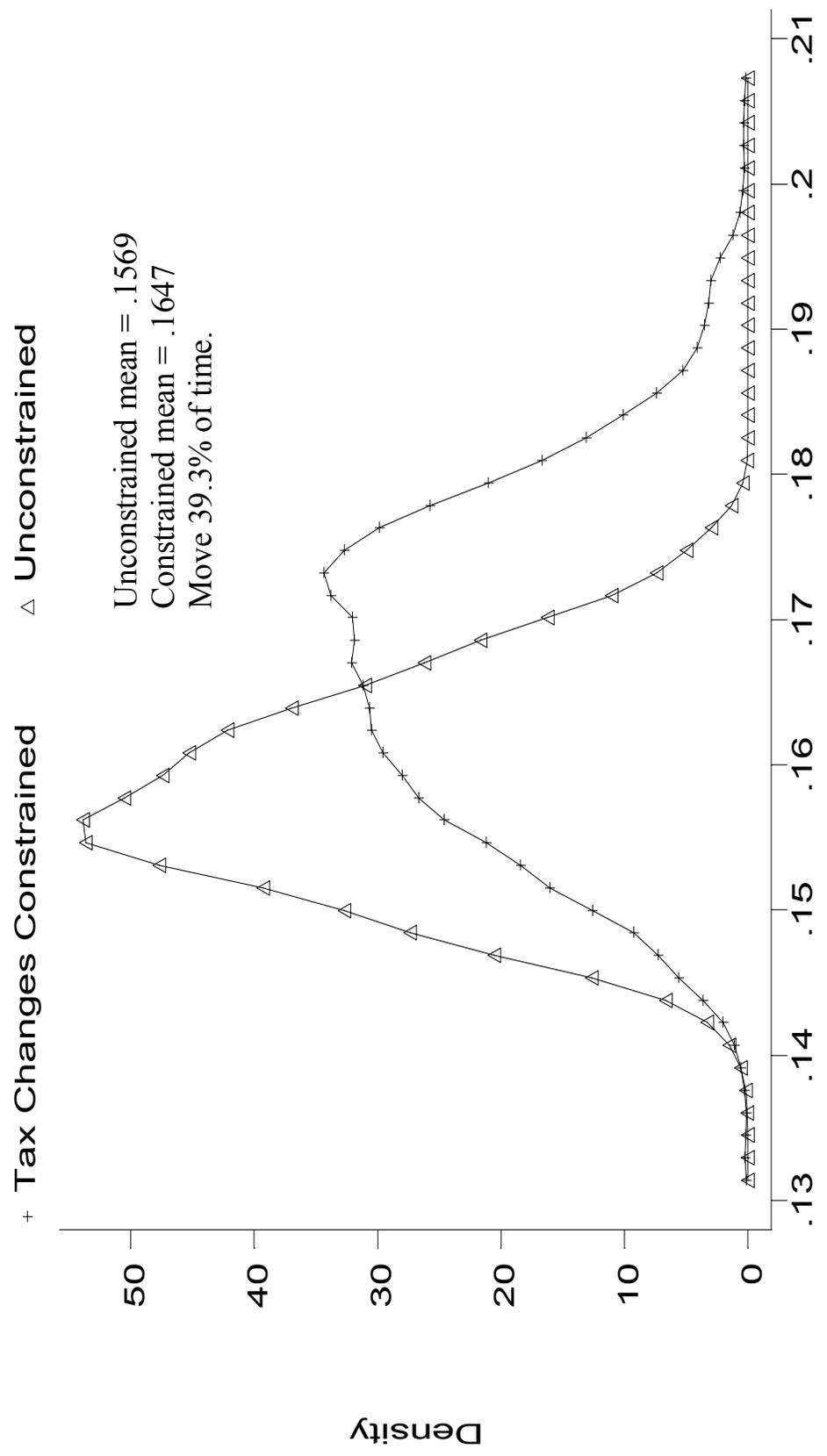
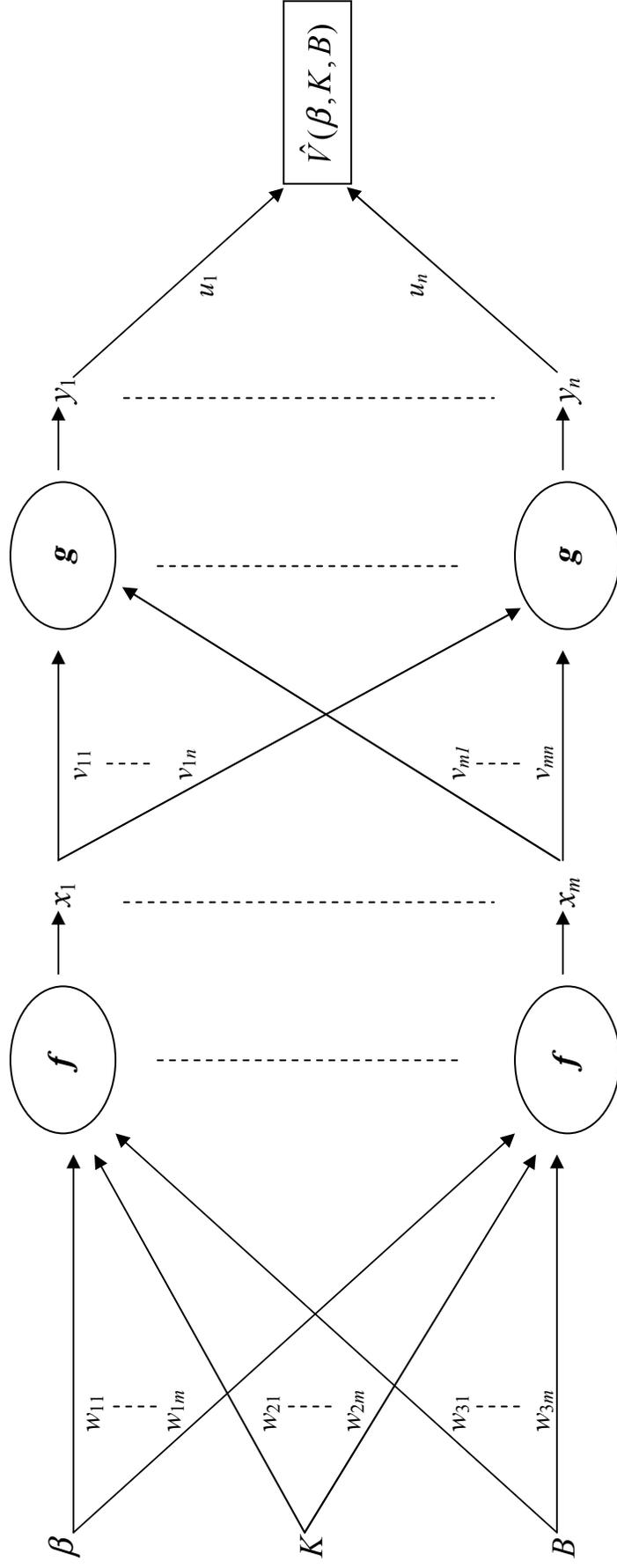


Figure 8. A Neural Network



Network illustrated has two hidden layers, with  $m$  and  $n$  nodes, respectively.