Department of Economics
School of Social Sciences

Analyzing Economic Policy Using High Order Perturbations

Michael Ben-Gad[1]
City University

Department of Economics
Discussion Paper Series
No. 08/07

[1] Department of Economics, City University, Northampton Square, London, EC1V 0HB, UK. Email:*mbengad@city.ac.uk*

Analyzing Economic Policy Using High Order Perturbations

Michael Ben-Gad
Department of Economics
City University
Northampton Square
London EC1V 0HB, UK*

September 24, 2008

Abstract

In this chapter I demonstrate the use of high order general perturbations to analyze policy changes in dynamic economic models. The inclusion of high moments in approximating the behavior of dynamic models is particularly necessary for welfare analysis. I apply the method of general perturbations to the analysis of permanent changes to a flat rate tax on the return to capital in the context of the standard Ramsey optimal growth model. Reliance on simple linearizations or quadratic approximations are adequate for generating impulse responses for the variables of interest or the welfare analysis of small policy changes. However when considering the welfare implications of sizable policy changes, the failure to include higher moments can lead not only to quantitatively serious inaccuracies, but even to spurious welfare reversals.

*This research was funded by the Israel Science Foundation, Grant No. 49/06. e-mail: *mbengad@city.ac.uk*.

normative analyses of policy changes, and particularly if the proposed policy changes are large, failure to consider higher order moments can generate results that are often quantitatively, and occasionally qualitatively misleading.

In this article I consider the welfare effects of large, but temporary changes to a flat rate tax on income derived from capital in the context of the standard continuous-time Ramsey optimal growth model. All revenue from the tax is returned to taxpayers as transfer payments, so changes in the tax rate only affect household welfare indirectly, through the excess burden or deadweight loss the distortionary tax generates. To analyze the changes to policy I employ the method of general perturbations first introduced into economics by Kenneth Judd in the 1980's (Judd 1982, 1985, 1987). In those original papers first order perturbations were employed to analyze relatively small changes in fiscal policy. Here because the changes I analyze are quite large, I demonstrate the use of high order perturbations (as outlined in Judd (1999)) and how the inclusion of successively higher moments can change our assessment of the welfare implications of sizable changes in policy.

There are three reasons why the method of general perturbations is particularly useful for this type of analysis. First, it yields explicit continuous time formulae that describe the dynamic behavior of the economic variables rather than policy functions defined by grids or collections of points. This is particularly advantageous for analyzing welfare in the context of continuous time models, as such analysis typically requires integration of utility functions that themselves are dependent on the time path of consumption. Second, the procedure permits high order approximations that are simple to implement and are analogous to Taylor expansions. Finally, the perturbations procedure is useful for analyzing complicated dynamic changes to policy, beyond the simple temporary changes I analyze here (Ben-Gad (2004, 2006, 2008)).

The policy change I consider is a temporary change in a flat rate tax paid on the returns generated by physical capital and used to fund transfers to the representative household. Starting at a baseline rate of 35%, I consider the quantitative effect on the welfare of the household of permanently shifting the tax rate between zero to 99% over the course five, ten, fifteen, and twenty years. I demonstrate that for tax cuts that are sufficiently large, failure to include high order moments in the approximation can generate the type of spurious welfare reversals documented by Tesar (1995) and Kim and Kim (2003) when they compare economies with complete and incomplete markets. Furthermore, even though qualitatively, a first order approximation may yield the same predicted welfare outcomes as a third order approximation, quantitatively the results may differ substantially. In the model I present, an analysis based on only the first moment of the perturbations method (essentially a linearization) can generate significant overestimates of the welfare gains from lowering the tax on income from capital and even larger underestimates of the welfare losses generated by raising the tax. This quantitative discrepancy is important, as rarely is such a policy considered in isolation, but only in tandem with considerations involving income distribution and tax incidence. Hence including the nonlinearities first order approximations ignore is essential for producing a realistic picture of the trade-offs such policy changes imply.

It is important to bear in mind that the optimal growth model considered here is generally perceived to be close to linear, which indeed it is. Hence if the changes considered are modest and only the impulse responses, and not their welfare implications are of interest, first order approximations may indeed be sufficient. Nonetheless, if the higher moments in this type of model can prove to be quantitatively important for welfare analysis, then they argue for great caution when analyzing models characterized by far less linear dynamics. Furthermore, the simple temporary changes considered here are not particularly dynamic themselves; they generate one-time jumps in consumption followed by a monotonic path in the opposite direction until the policy ends, and then gradual convergence back to the original steady state. More complicated policy changes, such as gradual, non-monotonic, or delayed changes to fiscal policy, even if more modest in scope than those considered here, provide plenty of opportunities for serious welfare miscalculations unless the higher moments are included in any approximation the time path the economy will follow after the policy is announced.

2 The Ramsey Optimal Growth Model with Capital Taxation

Consider a representative agent whose income is generated by wages $w(t)$ from fixed labor supply l, the return $r(t)$ net of the flat rate tax $\tau(t)$ on capital holdings $k(t)$ as well as a transfer payment $v(t)$. The agent maximizes the present value of utility $U : \mathbb{R}_{++} \rightarrow \mathbb{R}$ discounted at the rate ρ, generated by a continuous stream of consumption $c(t)$:

$$\max_{c} \int_{0}^{\infty} e^{-\rho t} U(c(t))\, dt \tag{1}$$

subject to the continuous time budget constraint:

$$\dot{k}(t) = w(t)l + (1 - \tau(t))\, r(t)k(t) - c(t) + v(t) \quad \forall t. \tag{2}$$

I assume the instantaneous utility function is of the constant intertemporal elasticity of substitution (constant relative risk aversion) form, $U(c) = \frac{c^{1-\sigma}}{1-\sigma}$, where $\sigma \geq 0$, is the Arrow Pratt measure of relative risk aversion (the inverse of the intertemporal elasticity of substitution).

The present value Hamiltonian of the optimization problem:

$$H\left(c\left(t\right),k\left(t\right),\lambda\left(t\right)\right)=e^{-\rho t}\frac{c\left(t\right)^{1-\sigma}}{1-\sigma}+\lambda(t)\left(w(t)l+\left(1-\tau\left(t\right)\right)r(t)k(t)-c(t)+v\left(t\right)\right) \quad (3)$$

yields the necessary first order conditions:

$$\frac{\partial H}{\partial c}:\frac{e^{\rho t}}{c(t)^{\sigma}}=\lambda(t) \quad (4)$$

$$\frac{\partial H}{\partial k}:\left(1-\tau\left(t\right)\right)\lambda(t)r(t)=-\dot{\lambda}(t) \quad (5)$$

and the transversality condition:

$$\lim_{t\rightarrow\infty}\lambda(t)k(t)=0. \quad (6)$$

Differentiating (4) with respect to t and substituting into (4) yields the law of motion for consumption:

$$\dot{c}(t)=\frac{1}{\sigma}\left[\left(1-\tau\left(t\right)\right)r(t)-\rho\right]c(t) \quad (7)$$

In this economy firms combine capital and labor to produce a single good that is both consumed employed as capital, maximizing:

$$\max_{k,l}\left\{F\left[k(t),l\right]-w(t)l-\left(r(t)-\delta\right)k(t)\right\} \quad (8)$$

where $F:\mathbb{R}_{+}^{2}\rightarrow\mathbb{R}_{+}$ is the production function and δ is the rate of capital depreciation. From the optimization problem (8), in equilibrium the returns to the factors of production are their marginal products:

$$F_{l}\left[k(t),l\right]=w(t), \quad (9)$$

$$F_{k}\left[k(t),l\right]=r(t)-\delta. \quad (10)$$

Assuming the production function takes the Cobb-Douglas form $F\left[k,l\right]=k(t)^{\alpha}l^{1-\alpha}$ and setting l=1, the laws of motion governing the dynamic behavior of the economy are the law of motion for consumption:

$$\dot{c}(t)=\frac{1}{\sigma}\left[\left(1-\tau\left(t\right)\right)\alpha k(t)^{\alpha-1}-\rho\right]c(t) \quad (11)$$

and from the market clearing condition, the law of motion for the capital stock:

$$\dot{k}(t)=k(t)^{\alpha}-c(t)-\delta k\left(t\right). \quad (12)$$

In Figure 1 I plot the loci that characterize the relationship between consumption and capital when in (11) $\dot{c}=0$, and when in (12) $\dot{k}=0$. In each panel σ =0.5, 1.5, or 2.5, the initial rate of taxation for income from capital is $\tau=0.35$, and the other parameter values are $\alpha=0.4$, $\delta=0.1$,

$\rho = 0.04$. The intersection between the two loci corresponds to steady state consumption and capital and the arrows in Figure 1 represent the vector field corresponding to the system (11) and (12). Notice neither the steady state values of consumption or capital, or indeed either of the loci are themselves functions of the curvature parameter σ, but its value does subtly influence the dynamics of the system and the shape of the saddle path along which the economy must converge towards steady state, and ultimately has a profound quantitative impact on the welfare effects of any change in policy.

If the government chooses to permanently double the tax rate on income from capital from 0.35 to 0.7, in order to fund more generous transfers, the locus corresponding to $\dot{k} = 0$ remains as it was, but the locus corresponding to $\dot{c} = 0$ in Figure 2 shifts to the right.[1] The vector field in Figure 2 correspond to the economy after the change in policy has been announced and indicates that consumption initially rises with the announced change in policy, before declining along with the capital stock until the system converges to its new steady state equilibrium. Similarly, a decision to eliminate transfers and the capital tax that funds them produces a qualitatively symmetric response in Figure 3—consumption initially declines but then both consumption and capital increase along a saddle path.

If the policy considered is temporary, convergence to the long-run steady state is no longer monotonic. Instead consumption experiences an immediate jump and then both consumption and capital gradually move back and forth along a vector field that is itself continuously changing. There are a number of competing methods for evaluating the evolution of $c(t)$ and $k(t)$ following such changes in policy or other exogenous shocks. It is important to emphasize that all methods, including numerical shooting are approximations. In the next section we consider one method that is both versatile and particularly suited to welfare analysis—the method of general perturbations. After that I demonstrate how it can be used to calculate the behavior of the system (11) and (12) to approximate the equilibrium values of $c(t)$ and $k(t)$.

3 The Method of General Perturbations

Consider a general dynamic system:

$$\dot{\mathbf{x}}(t) = \mathbf{\Gamma}\left[\mathbf{x}(t), \theta(t)\right], \tag{13}$$

where $\mathbf{x}(t) \in \mathbb{R}^n \times \mathbb{R}$, m is the number of control variables, $n-m$ is the number of state variables, $\theta(t) \in \mathbb{R}^p \times \mathbb{R}$ is a vector of policy variables. Assume that the system (13) is saddle path stable.

[1] The locus corresponding to $\dot{k} = 0$ would shift downward if the tax was used to finance government expenditure rather than a transfer payment.

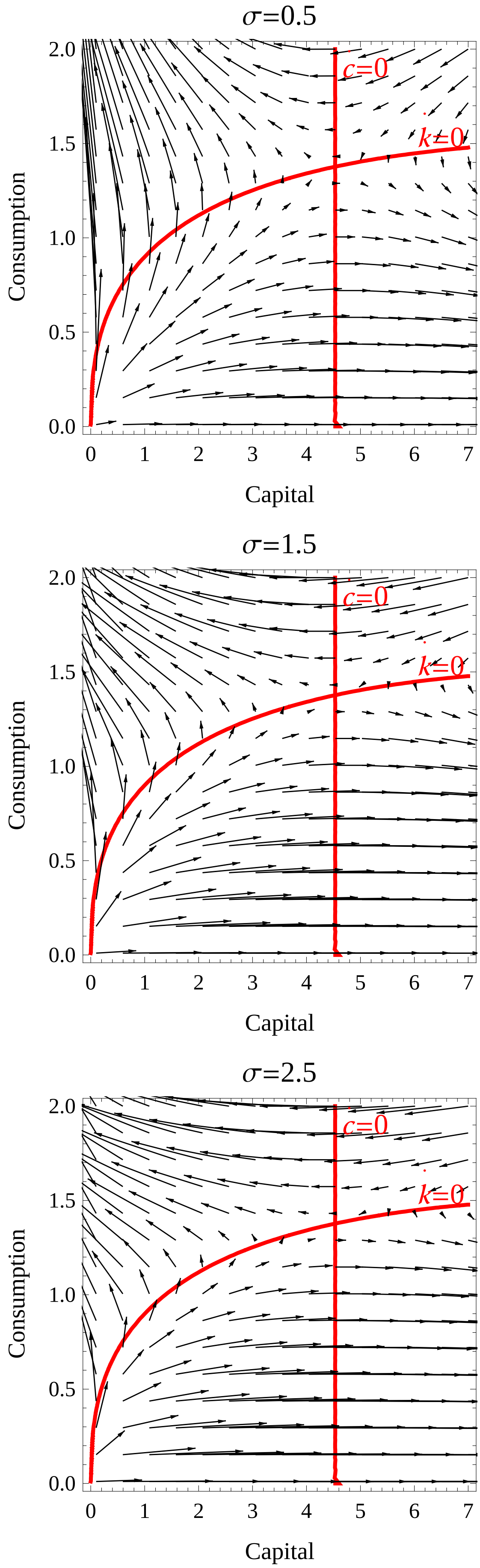

Figure 1: Vector field for the baseline model with $\tau = 0.35$.

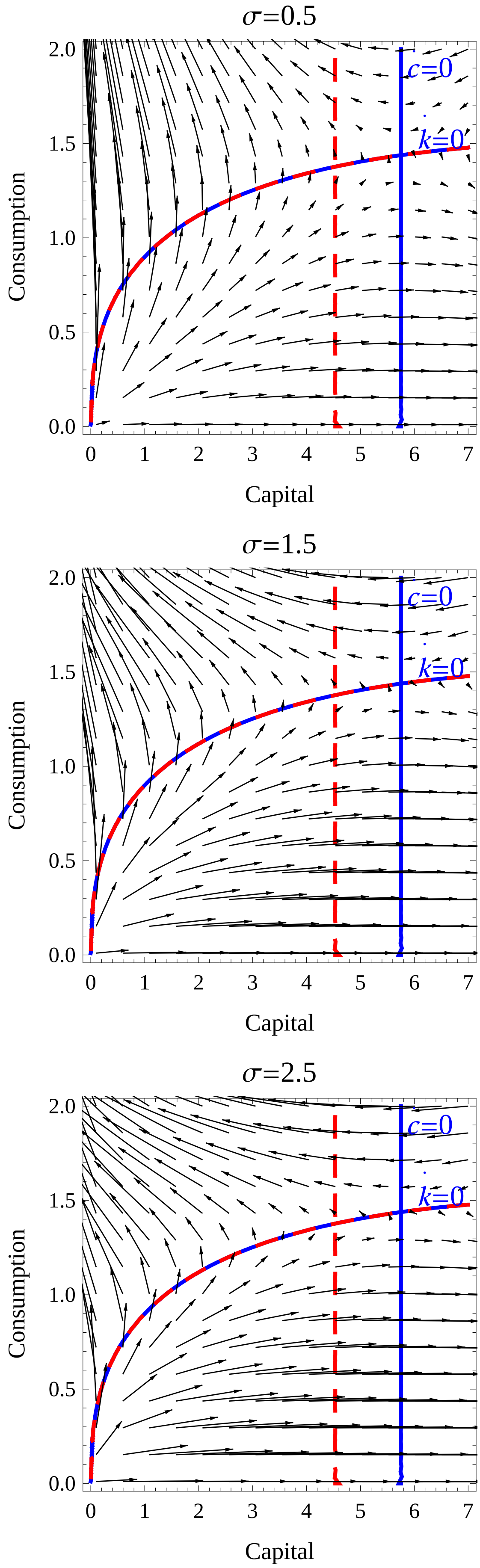

Figure 2: Vector field following abolition of the tax on capital income.

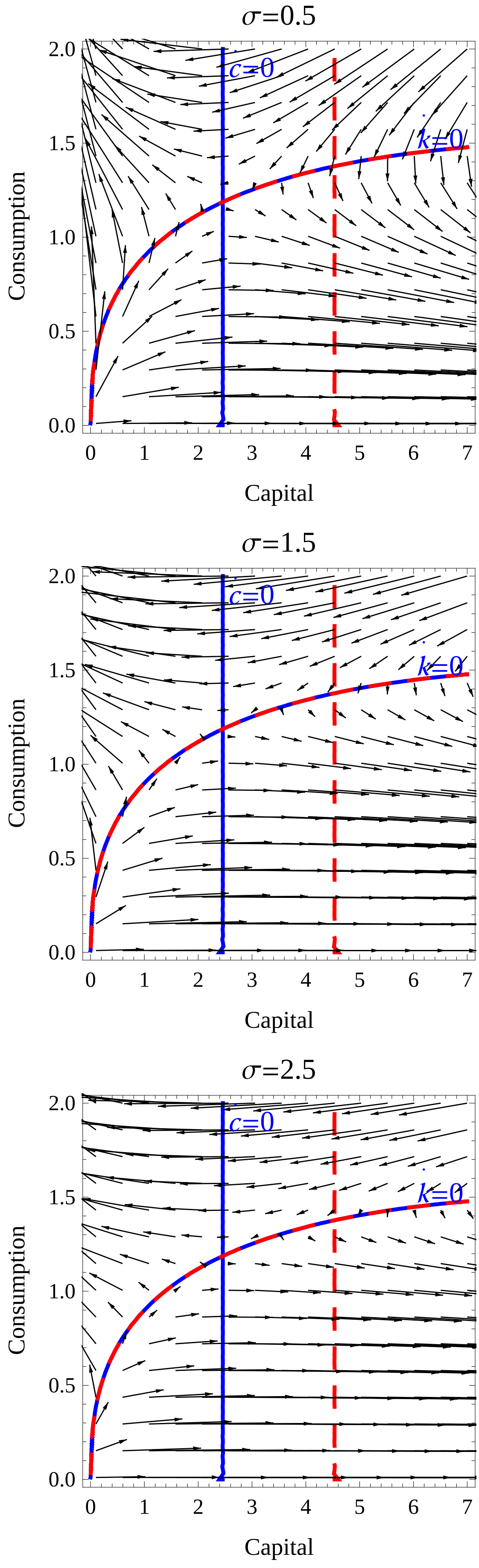

Figure 3: Vector field following the doubling of the tax on capital income.

Assume the vector $\theta(t)$ is a function of a set of dynamic perturbations $\pi(t) \in \mathbb{R}^p \times\mathbb{R}$, the magnitude of its effect on the system governed by a scalar ϵ so that $\theta(t) = \theta + \epsilon\pi(t)$. I rewrite the system as:

$$\dot{\mathbf{x}}(t;\epsilon) = \mathbf{\Gamma}\left[\mathbf{x}(t;\epsilon), \theta + \epsilon\pi(t)\right], \tag{14}$$

and differentiate with respect to ϵ :

$$\dot{\mathbf{x}}_\epsilon(t;\epsilon) = \mathbf{\Gamma}_\mathbf{x}\mathbf{x}_\epsilon(t;\epsilon) + \mathbf{\Gamma}_\theta\pi(t), \tag{15}$$

where $\mathbf{\Gamma}_\mathbf{x}$ and $\mathbf{\Gamma}_\theta$ are Jacobian matrices evaluated at steady state values (corresponding to $\epsilon = 0$). A first order approximation of the solution to (13) is:

$$\mathbf{x}(t) \approx \mathbf{x}(0;0) + \epsilon\mathbf{x}_\epsilon(t;\epsilon). \tag{16}$$

To solve the system (15), apply Laplace transforms to both sides:

$$\mathcal{L}_s\left[\dot{\mathbf{x}}_\epsilon\right] = \mathbf{\Gamma}_\mathbf{x}\mathcal{L}_s\left[\mathbf{x}_\epsilon\right] + \mathbf{\Gamma}_\theta\mathcal{L}_s\left[\pi\right] \tag{17}$$

where the Laplace transform of a function $f(t)$ is $\mathcal{L}_s[f] = \int_0^\infty f(t)\, e^{-st}dt$ and s is an arbitrary positive scalar. Applying the relationship: $\mathcal{L}_s\left[\dot{f}\right] = s\mathcal{L}_s[f] - f(0)$ to (17) and solving for $\mathcal{L}_s[\mathbf{x}_\epsilon]$ yields:

$$\mathcal{L}_s[\mathbf{x}_\epsilon] = \left[s\mathbf{I} - \mathbf{\Gamma}_\mathbf{x}\right]^{-1}\left(\mathbf{x}_\epsilon(0) + \mathbf{\Gamma}_\theta\mathcal{L}_s[\pi]\right), \tag{18}$$

which can be interpreted as providing a convenient relationship between the time discounted value of the variables of the model $\mathcal{L}_s[\mathbf{x}_\epsilon]$, the discounted value of the shocks $\mathcal{L}_s[\pi]$, and the initial change in the variables at the very moment a policy change becomes known $\mathbf{x}_\epsilon(0)$. Since $\mathcal{L}_s[\mathbf{x}_\epsilon]$ must be bounded for any positive value of s it must be bounded for any of the positive eigenvalues of $\mathbf{\Gamma}_\mathbf{x}$: μ_i, $i \in \{1, 2, ..., m\}$.[2] The determinants $|\mu_i\mathbf{I} - \mathbf{\Gamma}_\mathbf{x}|$ for each $i \in \{1, 2, ..., m\}$ equal zero by definition. Therefore the only way for the system to be bounded when s=μ_i, $i \in \{1, 2, ..., m\}$, is for the numerator of $\left[\mu_i\mathbf{I} - \mathbf{\Gamma}_\mathbf{x}\right]^{-1}$, the adjoint matrix of $\mu_i\mathbf{I} - \mathbf{\Gamma}_\mathbf{x}$ multiplied by the vector $\left(\mathbf{x}_\epsilon(0) + \mathbf{\Gamma}_\theta\mathcal{L}_s[\pi]\right)$ to be equal to zero. Hence the value of the non-zero elements of $\mathbf{x}_\epsilon(0)$, the first order approximation of the changes in the control variables that occur the moment the new policy is announced is the solution to:

$$adj\left[\mu_i\mathbf{I} - \mathbf{\Gamma}_\mathbf{x}\right]\left(\mathbf{x}_\epsilon(0) + \mathbf{\Gamma}_\theta\mathcal{L}_{\mu_i}[\pi]\right) = 0, i \in \{1, 2, ..., m\}. \tag{19}$$

Taking the inverse Laplace transform of (18) yields the first moment of the approximation:

$$\mathbf{x}_\epsilon(t;\epsilon) = e^{\mathbf{\Gamma}_\mathbf{x}t}\mathbf{x}_\epsilon(0) + \int_0^t e^{\mathbf{\Gamma}_\mathbf{x}(t-r)}\mathbf{\Gamma}_\theta\pi(r)dr. \tag{20}$$

[2]The number of positive eigenvalues must equal the number of controls if the system is saddle path stable.

A second order approximation is obtained by differentiating (15) with respect to ϵ :

$$\dot{\mathbf{x}}_{\epsilon\epsilon}(t;\epsilon) = \mathbf{\Gamma_x x}_{\epsilon\epsilon}(t;\epsilon) + \omega(t), \tag{21}$$

where $\omega(t)$ is a quadratic function of $\pi(t)$ and $\mathbf{x}_\epsilon(t;\epsilon)$ and the tensors $\mathbf{\Gamma_{xx}}$, $\mathbf{\Gamma_{x\theta}}$, $\mathbf{\Gamma_{\theta\theta}}$ yielding

$$\mathbf{x}(t;\epsilon) \approx \mathbf{x}(0;0) + \epsilon\mathbf{x}_\epsilon(t;\epsilon) + \frac{1}{2}\epsilon^2\mathbf{x}_{\epsilon\epsilon}(t;\epsilon). \tag{22}$$

The process can then be repeated to produce ever-closer approximations; an approximation of degree Z is:

$$\mathbf{x}(t;\epsilon) = \mathbf{x}(0;0) + \sum_{i=1}^{Z} \frac{\epsilon^i}{i!}\frac{\partial^i}{\partial\epsilon^i}\mathbf{x}(t;\epsilon) + \mathcal{O}_Z. \tag{23}$$

As an example, suppose $n = 2$ and $m = p = 1$ so that there is one control variable x_1, one state variable x_2, and only one policy variable changes. Assume the policy change is immediately implemented after it is announced, lasts for only T periods and is constant throughout, so $\pi(t) = \mathcal{U}(t-T)$ is the indicator function $\mathcal{U}(x) = 0$ if $x \leq 0$, and $\mathcal{U}(x) = 1$ if $x > 0$. Then (20) is:

$$\begin{aligned} x_{1,\epsilon}(t) &= \frac{\gamma_{x_{12}}\gamma_{\theta_2} - \gamma_{x_{22}}\gamma_{\theta_1}}{\mu_2\mu_1} + \frac{e^{t\mu_2 - T\mu_1}\left(1-e^{T\mu_1}\right)\left(\gamma_{x_{12}}\gamma_{\theta_2} + \gamma_{\theta_1}\left(\mu_1 - \gamma_{x_{22}}\right)\right)\left(\mu_2-\gamma_{x_{22}}\right)}{\left(\gamma_{x_{22}}-\mu_1\right)\left(\mu_1-\mu_2\right)\mu_1} \\ &+ \frac{1}{\mu_1\mu_2\left(\mu_2-\mu_1\right)}\left\{e^{t\mu_2}\left(\gamma_{x_{12}}\left(\gamma_{\theta_2}\mu_1 - \gamma_{x_{21}}\gamma_{\theta_1}\right) + \gamma_{x_{22}}\gamma_{\theta_1}\left(\mu_2-\gamma_{x_{22}}\right)\right)\right. \\ &+ \Big[\left(\gamma_{x_{12}}\left(\gamma_{x_{21}}\gamma_{\theta_1} - \gamma_{\theta_2}\mu_1\right) + \gamma_{x_{22}}\gamma_{\theta_1}\left(\gamma_{x_{22}}-\mu_2\right)\right)e^{\mu_2(t-T)} + \left(\gamma_{x_{22}}\gamma_{\theta_1} - \gamma_{x_{12}}\gamma_{\theta_2}\right)\left(\mu_2-\mu_1\right) \\ &\left. + e^{\mu_1(t-T)}\left(\gamma_{x_{22}}\gamma_{\theta_1}\left(\gamma_{x_{11}}-\mu_2\right) + \gamma_{x_{12}}\left(\gamma_{\theta_2}\mu_2 - \gamma_{x_{21}}\gamma_{\theta_1}\right)\right)\Big]\mathcal{U}(t-T)\right\}, \end{aligned} \tag{24}$$

$$\begin{aligned} x_{2,\epsilon}(t) &= \frac{\gamma_{x_{21}}\gamma_{\theta_1} - \gamma_{x_{11}}\gamma_{\theta_2}}{\mu_1\mu_2} + \frac{e^{t\mu_2 - T\mu_1}\left(1-e^{T\mu_1}\right)\gamma_{x_{21}}\left(\gamma_{x_{12}}\gamma_{\theta_2} + \gamma_{\theta_1}\left(\mu_1-\gamma_{x_{22}}\right)\right)}{\left(\gamma_{x_{22}}-\mu_1\right)\left(\mu_1-\mu_2\right)\mu_1} \\ &- \frac{1}{\mu_1\mu_2\left(\mu_1-\mu_2\right)}\left\{e^{t\mu_2}\left(\gamma_{x_{21}}\gamma_{\theta_1}\left(\gamma_{x_{11}}+\gamma_{x_{22}}-\mu_2\right) + \gamma_{\theta_2}\left(\mu_2\gamma_{x_{11}} - \gamma_{x_{12}}\gamma_{x_{21}} - \gamma_{x_{11}}^2\right)\right)\right. \\ &- \Big[e^{(t-T)\mu_1}\left(\gamma_{\theta_2}\left(\gamma_{x_{11}}-\mu_1\right) - \gamma_{x_{21}}\gamma_{\theta_1}\right)\mu_2 + \left(\gamma_{x_{21}}\gamma_{\theta_1} - \gamma_{x_{11}}\gamma_{\theta_2}\right)\left(\mu_2-\mu_1\right) \\ &\left. + e^{(t-T)\mu_2}\left(\gamma_{x_{21}}\gamma_{\theta_1}\left(\gamma_{x_{11}}+\gamma_{x_{22}}-\mu_2\right) + \gamma_{\theta_2}\left(\mu_2\gamma_{x_{11}} - \gamma_{x_{12}}\gamma_{x_{21}} - \gamma_{x_{11}}^2\right)\right)\Big]\mathcal{U}(t-T)\right\}. \end{aligned} \tag{25}$$

If on the other hand the change in policy is permanent as in Figures 1 to 3, then $\pi = 1$, and (20) reduces to:

$$\begin{aligned} x_{1,\epsilon}(t) &= \frac{\gamma_{\mathbf{x}_{12}}\gamma_{\theta_2} - \gamma_{\mathbf{x}_{22}}\gamma_{\theta_1}}{\mu_1\mu_2} \\ &+ \frac{\left(\gamma_{\mathbf{x}_{11}}\gamma_{\theta_2} - \gamma_{\mathbf{x}_{21}}\gamma_{\theta_1}\right)\left(\left(\left(\gamma_{\mathbf{x}_{11}}-\mu_1\right)\left(\gamma_{\mathbf{x}_{22}}-\mu_2\right)\gamma_{\mathbf{x}_{22}}e^{\mu_2 t} - \left(\gamma_{\mathbf{x}_{22}}+\mu_1-\mu_2\right)\gamma_{\mathbf{x}_{12}}\gamma_{\mathbf{x}_{21}}\right)\right)}{\left(\mu_1-\mu_2\right)\gamma_{\mathbf{x}_{12}}{\mu_1}^2\mu_2}, \end{aligned} \tag{26}$$

$$x_{2,\epsilon}(t) = \frac{\gamma_{\mathbf{x}_{21}}\gamma_{\theta_1} - \gamma_{\mathbf{x}_{11}}\gamma_{\theta_2}}{\mu_1\mu_2}\left[1 - e^{\mu_2 t}\right]. \tag{27}$$

The values of $x_{1,\epsilon}(t)$ and $x_{2,\epsilon}(t)$ from (24) and (25), or (26) and (27), become inputs in the shock processes used to calculate the second moments $x_{1,\epsilon\epsilon}(t)$ and $x_{2,\epsilon\epsilon}(t)$, which in turn are used to calculate third moments. As already mentioned, this process can be repeated indefinitely, though each successive moment demands ever more computer resources to calculate.

4 Applying the Method of General Perturbations to a Permanent Change in the Rate of Capital Taxation

To implement the method of perturbations to the question of a how change in the tax rate on capital income affect the economy, substitute for the tax rate on capital income in (11) and (12) $\tau(t) = \tau + \epsilon\pi(t)$ where $\pi(t)$ is any bounded dynamic path and ϵ is a small number that regulates its magnitude:

$$\begin{bmatrix} \mathcal{L}_s[c_\epsilon] \\ \mathcal{L}_s[k_\epsilon] \end{bmatrix} =$$

$$\begin{bmatrix} s & -\frac{(1-\alpha)(\rho+(1-\tau)\delta)}{\alpha\sigma}\left((1-\alpha)\delta + \frac{\rho}{1-\tau}\right) \\ -1 & s - \frac{\rho}{1-\tau} \end{bmatrix}^{-1} \begin{bmatrix} c_\epsilon(0) + \frac{\rho}{\sigma(1-\tau)}\left(\left(\frac{\alpha(1-\tau)}{\rho+(1-\tau)\delta}\right)^{\frac{1}{1-\alpha}} + \left(\frac{\alpha(1-\tau)}{\rho+(1-\tau)\delta}\right)^{\frac{\alpha}{1-\alpha}}\right)\mathcal{L}_s[\pi] \\ 0 \end{bmatrix}. \tag{28}$$

The initial change in the control variable consumption is:

$$c_\epsilon(0) = \frac{\alpha^{\frac{\alpha}{1-\alpha}}\rho(\rho + (1-\alpha)\delta(1-\tau))(1-\tau)^{\frac{1}{1-\alpha}-2}}{\sigma(\rho+\delta(1-\tau))^{\frac{1}{1-\alpha}}}\mathcal{L}_{\mu_2}[\pi] \tag{29}$$

where $\mu_1 = \frac{1}{2}\left(\frac{\rho}{1-\tau} + \sqrt{\frac{\rho^2}{(1-\tau)^2} + \frac{4(1-\alpha)(\rho+(1-\tau)\delta)}{\alpha\sigma}\left((1-\alpha)\delta + \frac{\rho}{1-\tau}\right)}\right)$. The first order effects of an immediate shock to taxation that lasts until period T is:

$$\begin{aligned} c_\epsilon(t) \quad = \quad & -\frac{\alpha^{\frac{\alpha}{1-\alpha}}\rho(\rho+(1-\alpha)\delta(1-\tau))(1-\tau)^{\frac{1}{1-\alpha}}}{\sigma\mu_1(\rho-(1-\tau)\mu_1)(\rho-2(1-\tau)\mu_1)((1-\tau)\delta+\rho)^{\frac{1}{1-\alpha}}} \\ & \times\left[\frac{\left(e^{(t-T)\mu_1}(\rho-(1-\tau)\mu_1)^2 - \rho(\rho-2(1-\tau)\mu_1)\right)(1-\mathcal{U}(t-T))}{(1-\tau)^2}\right. \\ & \left. e^{t\mu_2}\mu_1\left(e^{-\mu_2 T}\mu_1\mathcal{U}(t-T) - \frac{e^{-\mu_1 T}\left(\left(1-e^{T\mu_1}\right)\rho - \left(1-2e^{T\mu_1}\right)(1-\tau)\mu_1\right)}{1-\tau}\right)\right], \end{aligned} \tag{30}$$

$$k_{\epsilon}(t) = -\frac{\alpha^{\frac{\alpha}{1-\alpha}}\rho(\rho+(1-\alpha)\delta(1-\tau))(1-\tau)^{\frac{1}{1-\alpha}}}{\sigma\mu_1\left(\rho-(1-\tau)\mu_1\right)\left(\rho-2(1-\tau)\mu_1\right)\left((1-\tau)\delta+\rho\right)^{\frac{1}{1-\alpha}}} \tag{31}$$

$$\left[\left(\left(e^{t\mu_2}-e^{t\mu_1}\right)\left(\mu_1-\frac{\rho}{1-\tau}\right)e^{-T\mu_1}-\frac{\left(1-e^{t\mu_2}\right)\left(\rho-2(1-\tau)\mu_1\right)}{1-\tau}\right)\right.$$

$$\left.+\frac{\left(\left(1-e^{(t-T)\mu_1}\right)\rho-\left(2-e^{(t-T)\mu_1}-e^{(t-T)\mu_2}\right)(1-\tau)\mu_1\right)\mathcal{U}(t-T)}{1-\tau}\right],$$

where $\mu_2 = \frac{1}{2}\left(\frac{\rho}{1-\tau}-\sqrt{\frac{\rho^2}{(1-\tau)^2}+\frac{4(1-\alpha)(\rho+(1-\tau)\delta)}{\alpha\sigma}\left((1-\alpha)\delta+\frac{\rho}{1-\tau}\right)}\right)$.

If the change in the rate of taxation is permanent, $\pi(t)=1$, and the first order perturbations that correspond to (26) and (27) are:

$$c_{\epsilon}(t) = \frac{\alpha^{\frac{1}{1-\alpha}}\rho\left((1-\tau)\mu_1 e^{\mu_2 t}-\rho\right)}{(1-\alpha)(1-\tau)^{\frac{1-2\alpha}{1-\alpha}}(\delta+\rho-\tau\delta)^{\frac{2-\alpha}{1-\alpha}}}, \tag{32}$$

$$k_{\epsilon}(t) = \frac{\alpha^{\frac{1}{1-\alpha}}\rho\left(e^{\mu_2 t}-1\right)(1-\tau)^{\frac{\alpha}{1-\alpha}}}{(1-\alpha)(\delta+\rho-\tau\delta)^{\frac{2-\alpha}{1-\alpha}}}, \tag{33}$$

As stated above, the process can be repeated to attain higher order approximations. The second order shock process in (21) are functions of the values of $c_{\epsilon}(t)$ and $k_{\epsilon}(t)$. Those that correspond to temporary perturbations are long and calculated expressions but those that correspond to permanent changes are:

$$\omega_1(t) = \frac{2}{\sigma}\left\{\delta c_{\epsilon}(t)-\alpha\left(\frac{\rho+(1-\tau)\delta}{\alpha(1-\tau)}\right)^{\frac{2-\alpha}{1-\alpha}}\left[\left(\left(\frac{\alpha(1-\tau)}{\rho+(1-\tau)\delta}\right)^{\frac{1}{1-\alpha}}+(1-\alpha)(1-\tau)k_{\epsilon}(t)\right)c_{\epsilon}(t)\right.\right. \tag{34}$$

$$\left.\left.+(1-\alpha)\left(\frac{\rho+(1-\alpha)(1-\tau)\delta}{\alpha(1-\tau)}\right)\left(\left(\frac{\alpha(1-\tau)}{\rho+(1-\tau)\delta}\right)^{\frac{1}{1-\alpha}}+(1-\frac{\alpha}{2})(1-\tau)k_{\epsilon}(t)\right)k_{\epsilon}(t)\right]\right\},$$

$$\omega_2(t) = -\alpha(1-\alpha)\left(\frac{\rho+(1-\tau)\delta}{\alpha(1-\tau)}\right)^{\frac{2-\alpha}{1-\alpha}}k_{\epsilon}(t). \tag{35}$$

In Figure 4, I set the value of T=5, in Figure 5, T=10, in Figure 6, T=15, and in Figure 7, T=20. I then plot the values of the first three moments of the approximation of consumption, $c_{\epsilon}(t)$, $c_{\epsilon\epsilon}(t)$, and $c_{\epsilon\epsilon\epsilon}(t)$, for three different values of the parameter that governs the curvature of the utility function, $\sigma=0.5$, 1.5, 2.5. Note three important results. First, contrary to what we might expect, in each figure the lower the curvature of the utility function (the lower the value of σ), the larger the third order moments vary relative to the first moment. Second, the higher the value of T, the greater the two higher order moments vary relative to the first, and the greater the inaccuracy their exclusion generates. Hence the third order degree of non-linearity is greatest, the longer the policy lasts and the less averse are agents in the economy to substituting consumption between different periods of time. Finally, the second moment varies most relative

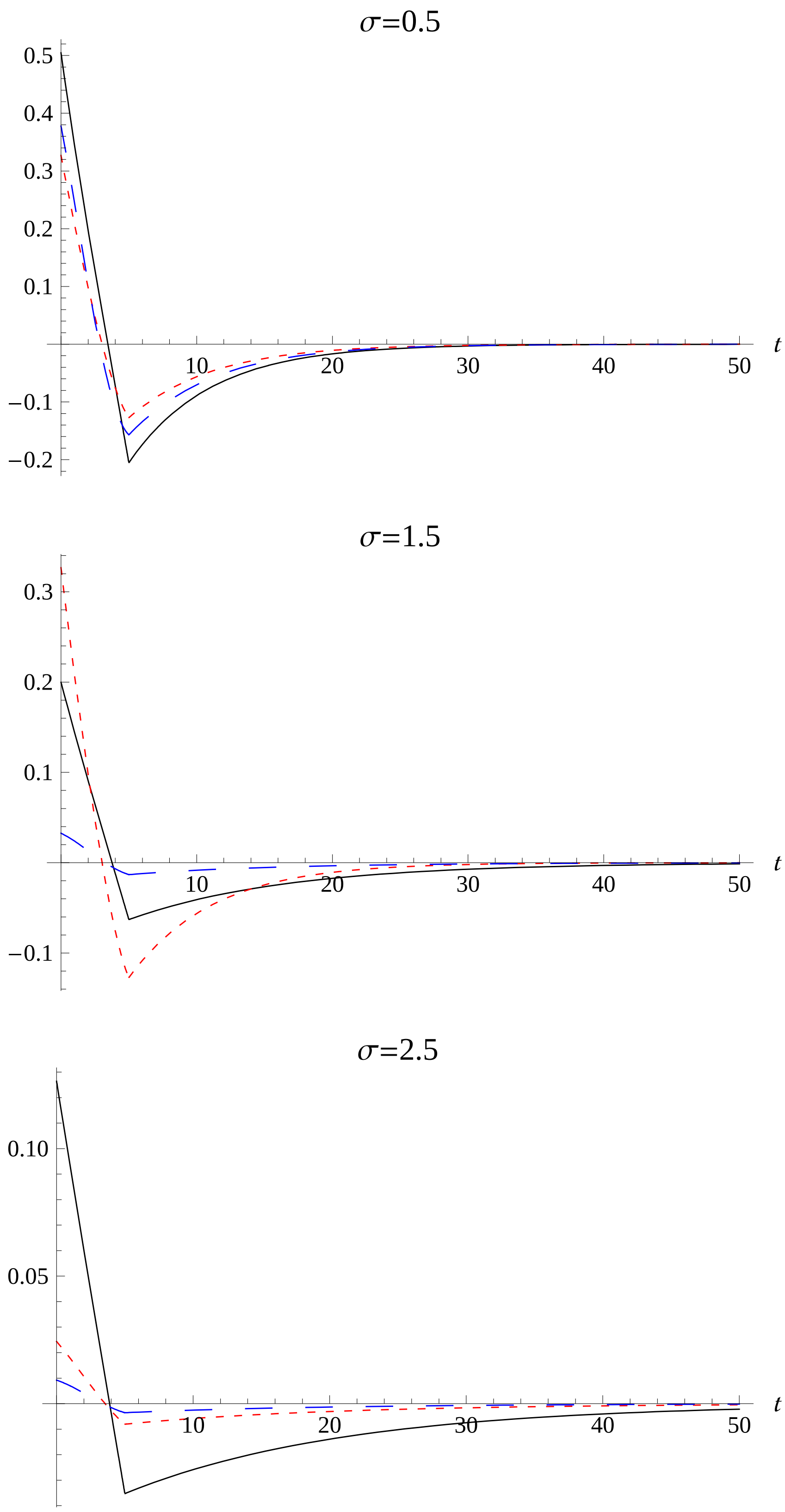

Figure 4: The first three moments of the approximation of consumption, $c_\epsilon(t)$ (solid black curve), $c_{\epsilon\epsilon}(t)$ (small red dashes), and $c_{\epsilon\epsilon\epsilon}(t)$ (large blue dashes), for $T = 5$ and σ =0.5, 1.5, 2.5.

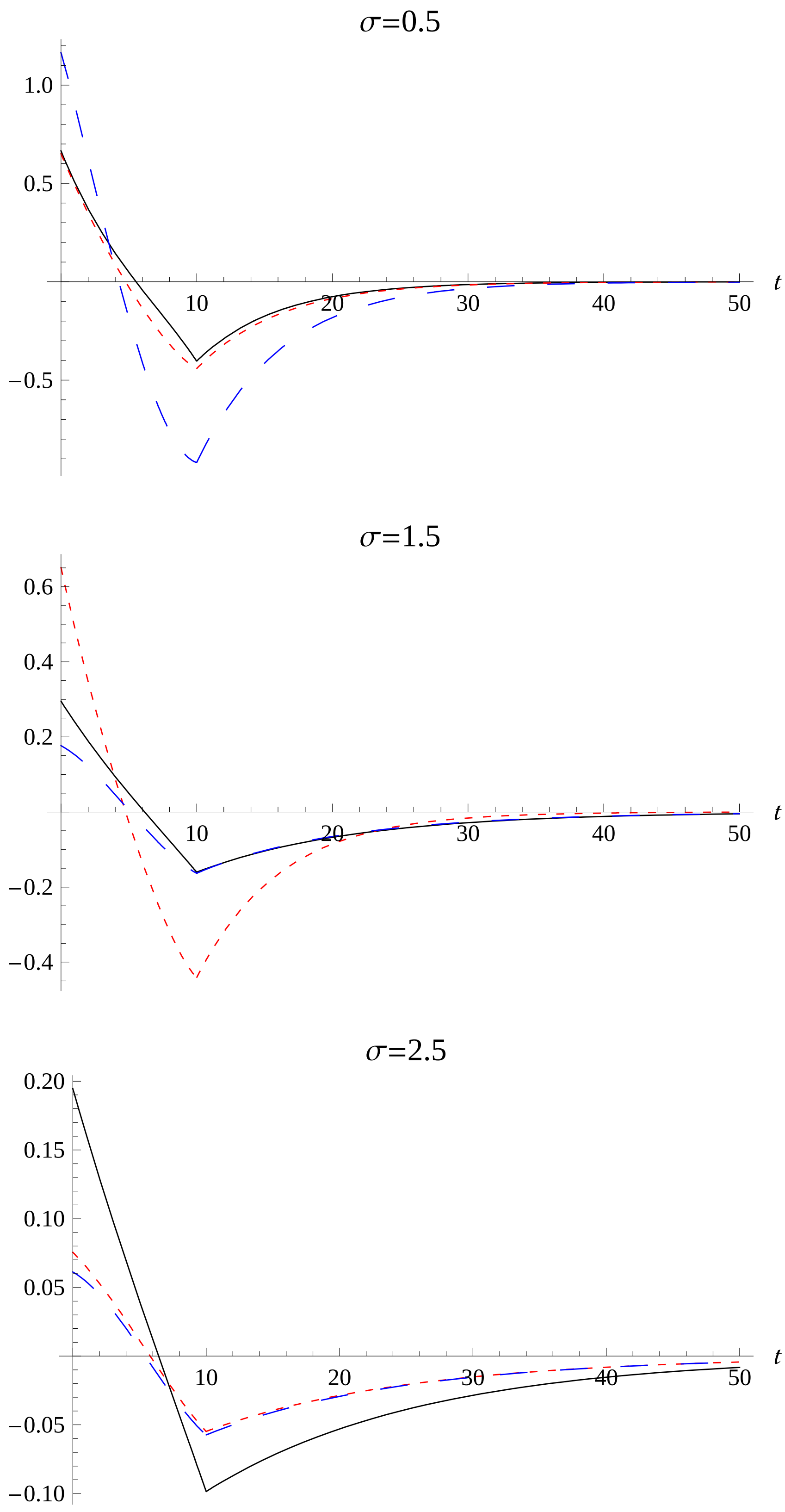

Figure 5: The first three moments of the approximation of consumption, $c_\epsilon(t)$ (solid black curve), $c_{\epsilon\epsilon}(t)$ (small red dashes), and $c_{\epsilon\epsilon\epsilon}(t)$ (large blue dashes), for $T = 10$ and σ =0.5, 1.5, 2.5.

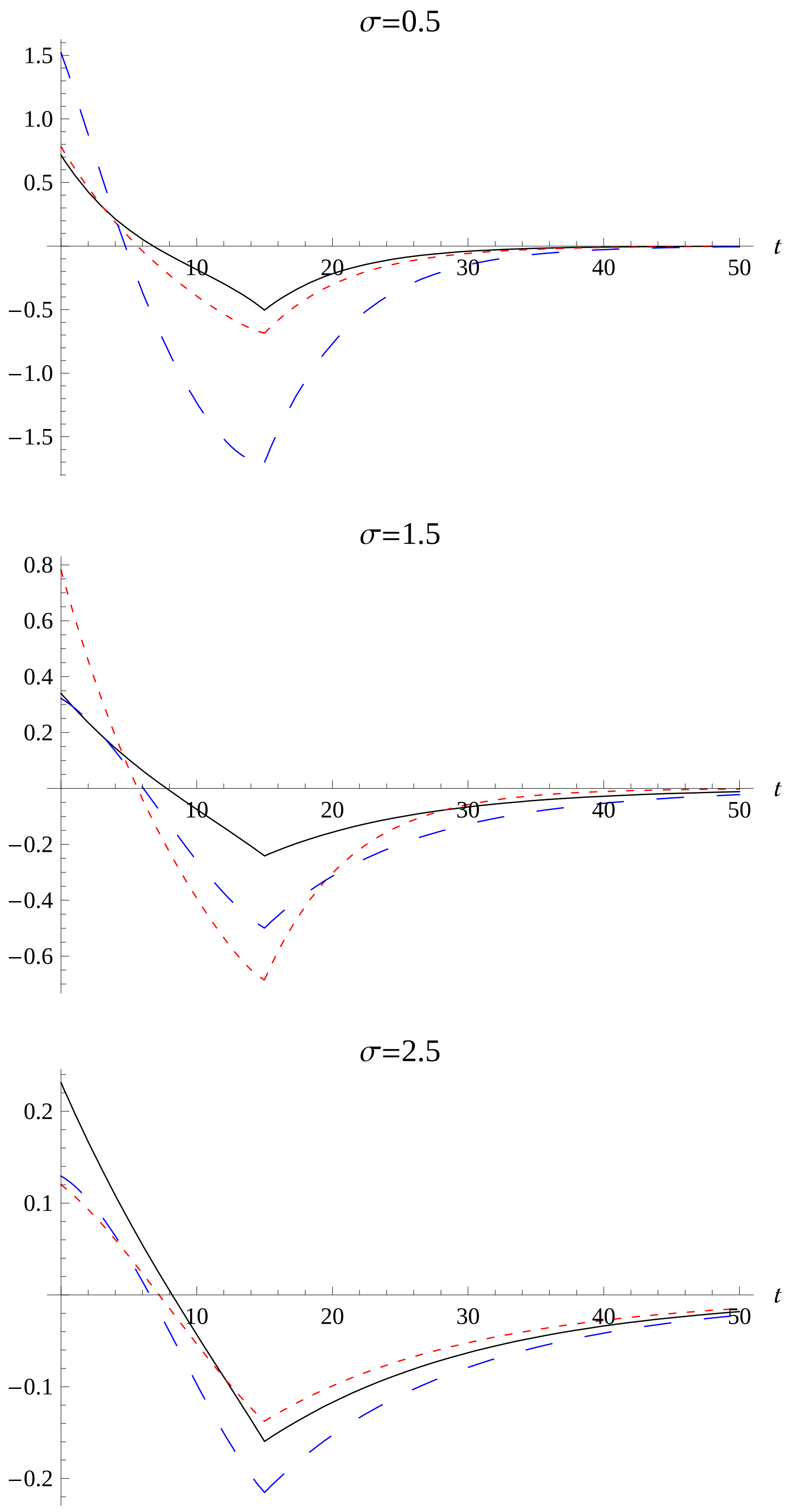

Figure 6: The first three moments of the approximation of consumption, $c_\epsilon(t)$ (solid black curve), $c_{\epsilon\epsilon}(t)$ (small red dashes), and $c_{\epsilon\epsilon\epsilon}(t)$ (large blue dashes), for $T = 15$ and σ =0.5, 1.5, 2.5.

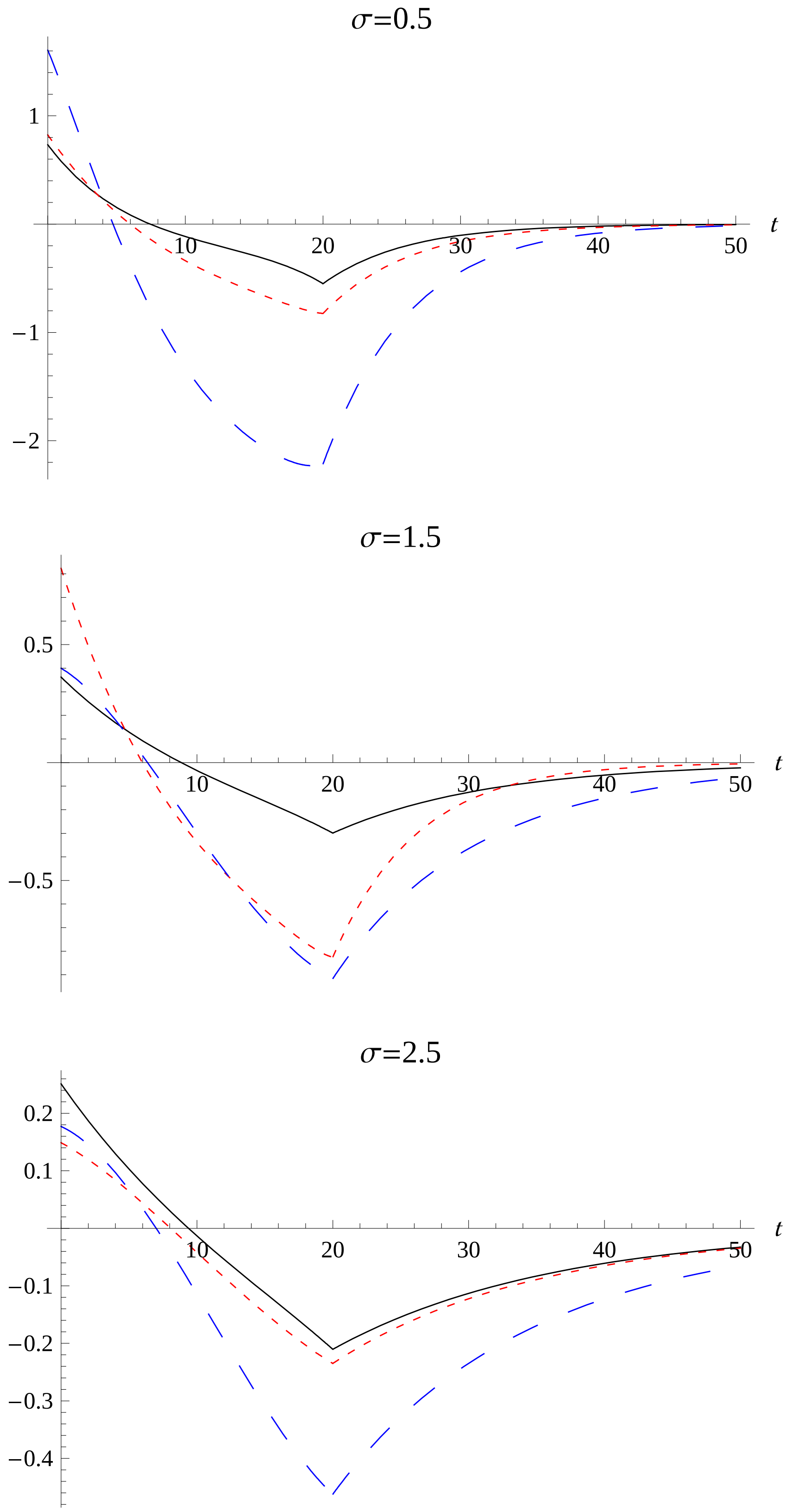

Figure 7: The first three moments of the approximation of consumption, $c_\epsilon(t)$ (solid black curve), $c_{\epsilon\epsilon}(t)$ (small red dashes), and $c_{\epsilon\epsilon\epsilon}(t)$ (large blue dashes), for $T = 20$ and σ =0.5, 1.5, 2.5.

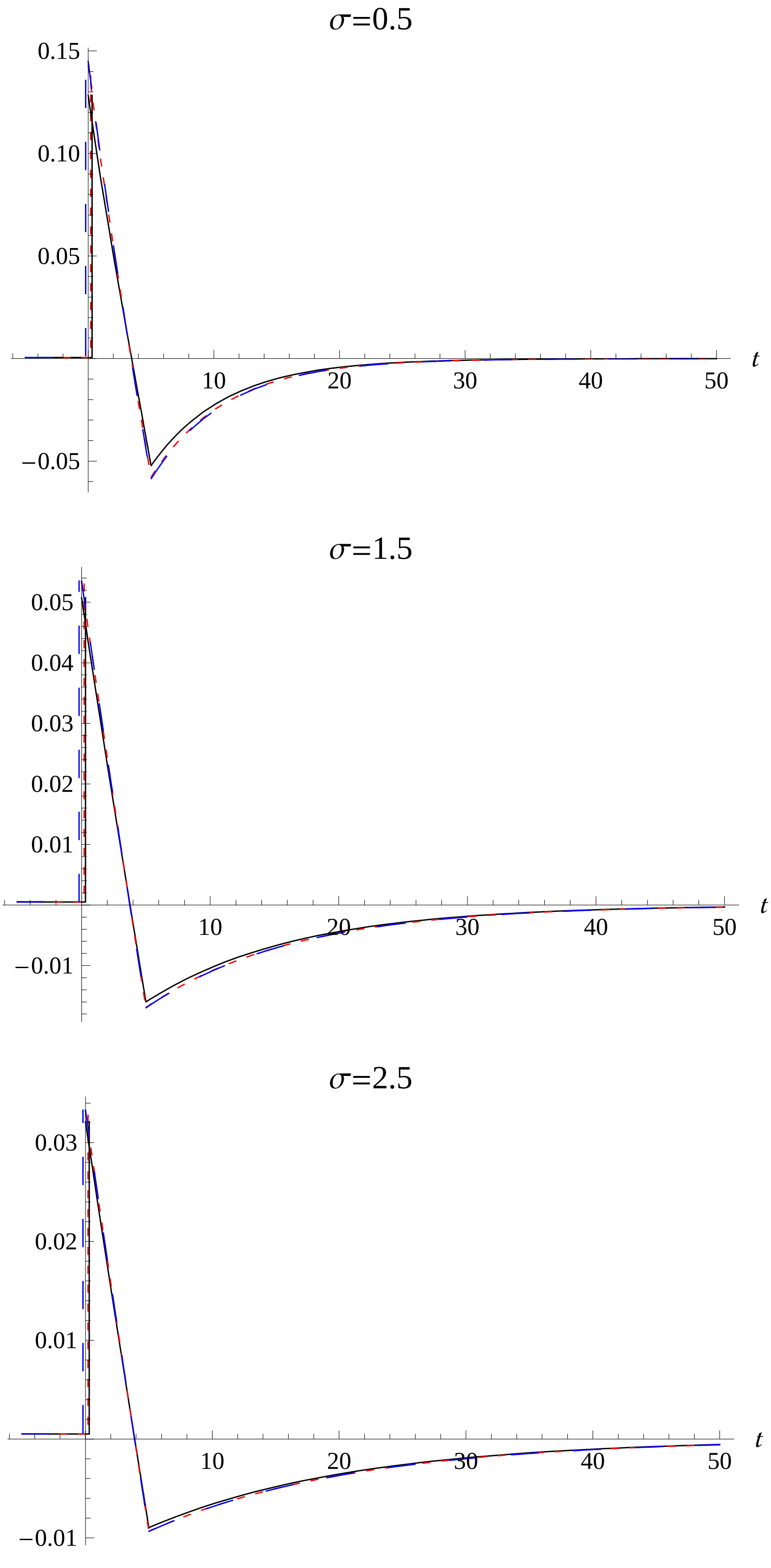

Figure 8: The behavior of consumption following the doubling of the tax on income from capital from 0.35 to 0.7 for $T = 5$ years. The solid black curve represents a first order approximation, the curve with small red dashes, a second order approximation, the large blue dashes, a third order approximation.

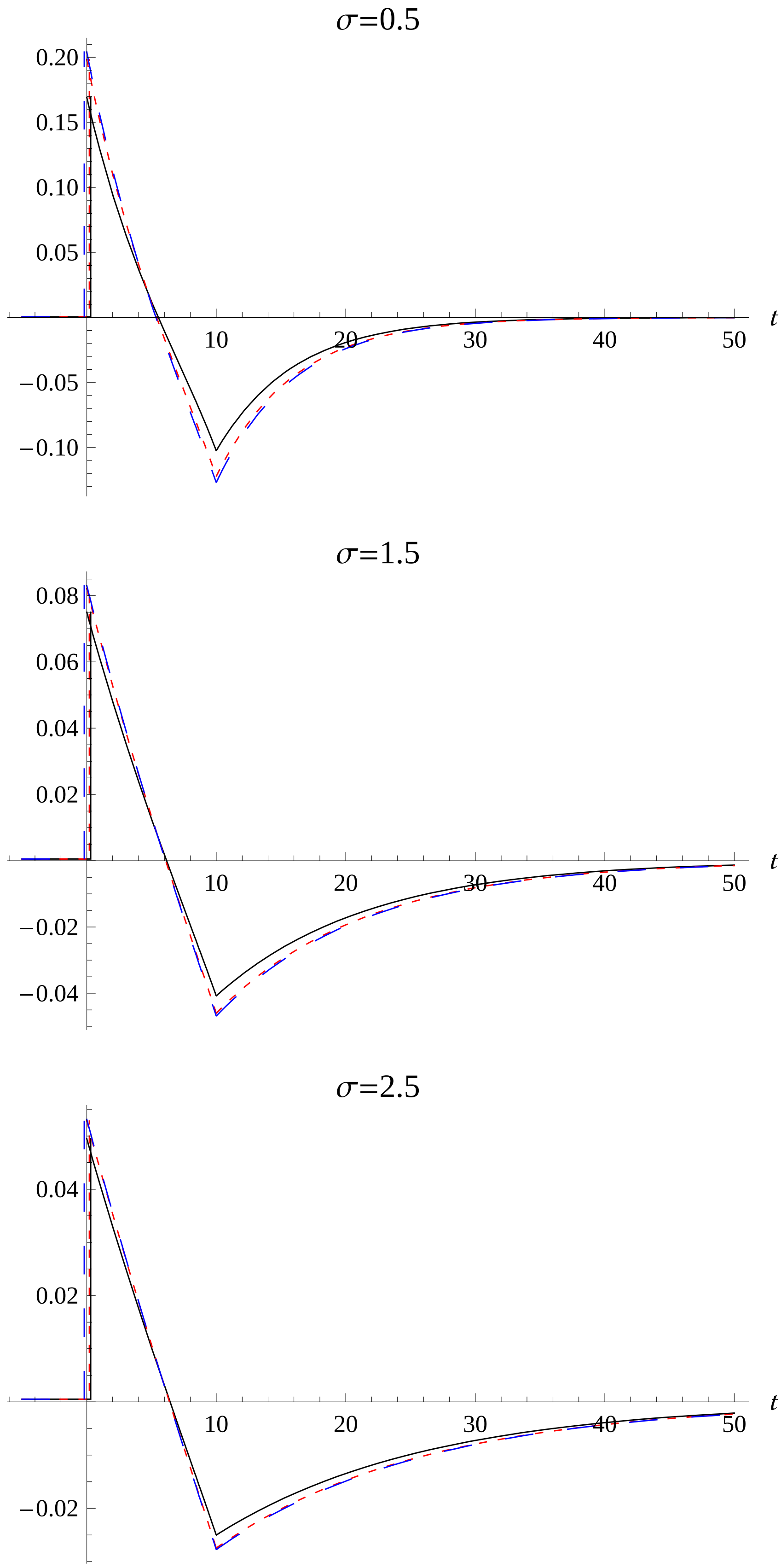

Figure 9: The behavior of consumption following the doubling of the tax on income from capital from 0.35 to 0.7 for $T = 10$ years. The solid black curve represents a first order approximation, the curve with small red dashes, a second order approximation, the large blue dashes, a third order approximation.

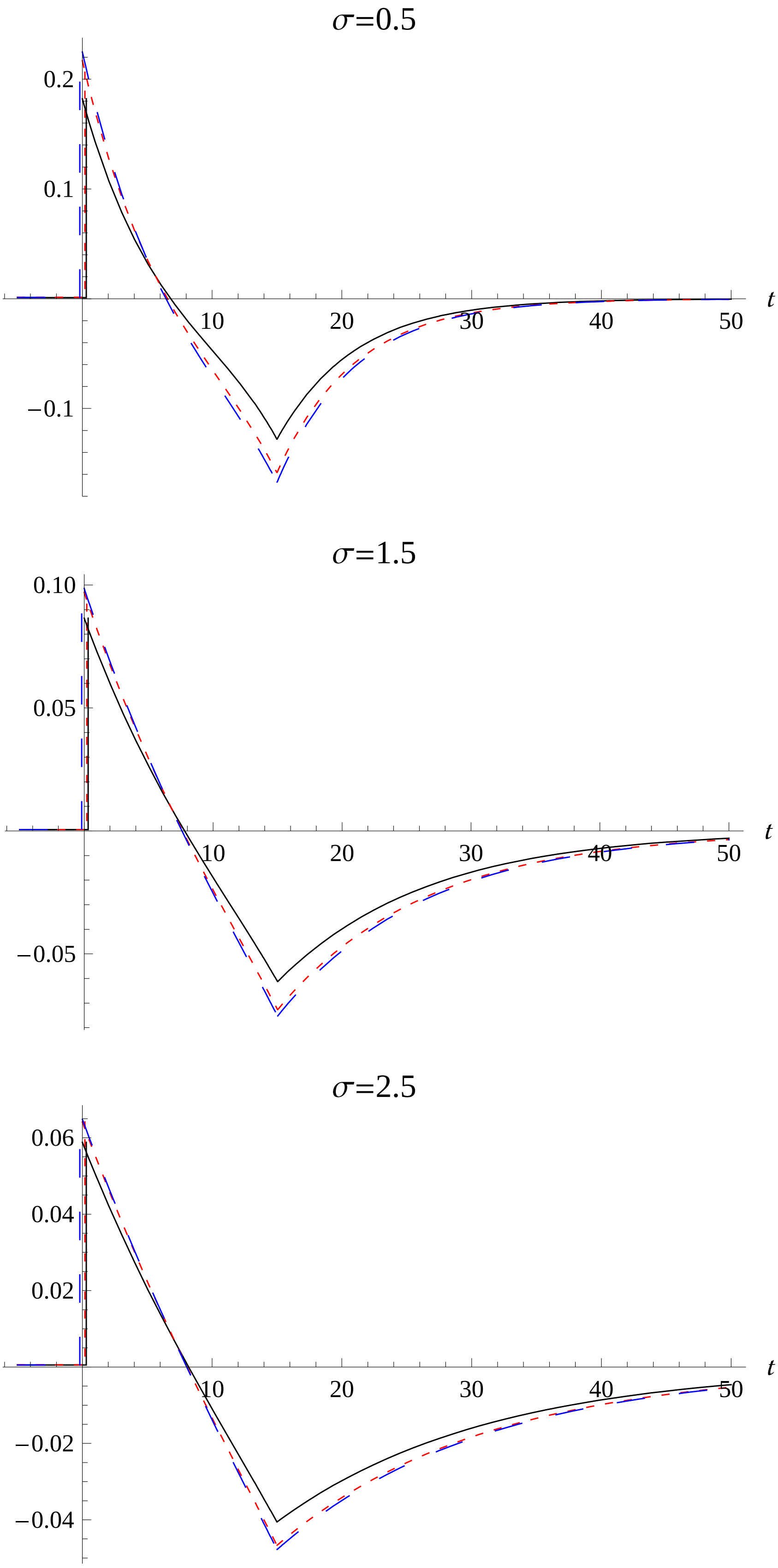

Figure 10: The behavior of consumption following the doubling of the tax on income from capital from 0.35 to 0.7 for $T = 15$ years. The solid black curve represents a first order approximation, the curve with small red dashes, a second order approximation, the large blue dashes, a third order approximation.

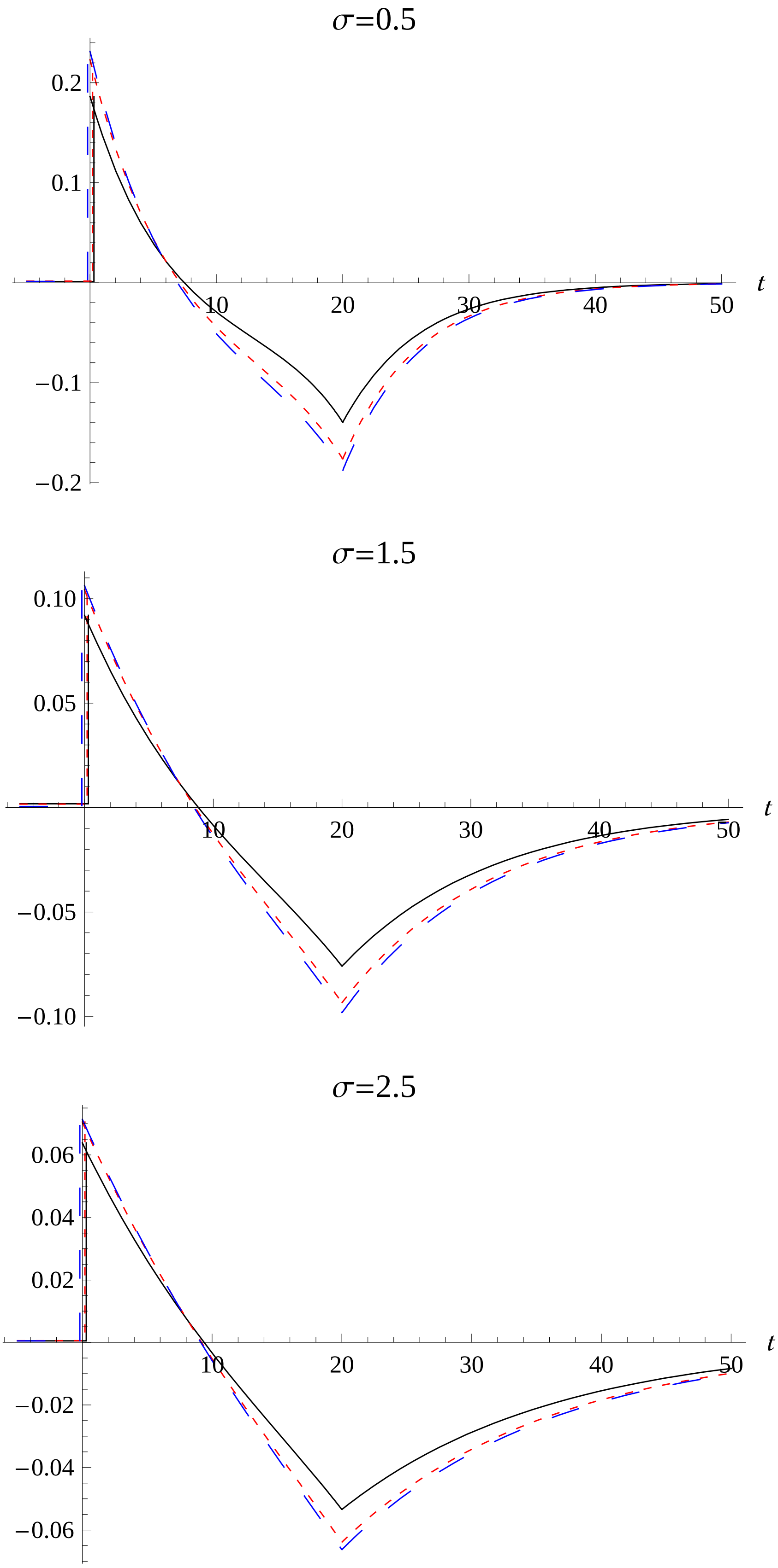

Figure 11: The behavior of consumption following the doubling of the tax on income from capital from 0.35 to 0.7 for $T = 20$ years. The solid black curve represents a first order approximation, the curve with small red dashes, a second order approximation, the large blue dashes, a third order approximation.

to the first if $\sigma = 1.5$. For both higher and lower values of σ, the quadratic component of non-linearity in the response of consumption is smaller.

I present the impulse responses of consumption itself following the doubling of the tax in Figures 8-11 or elimination of the tax in Figures 12-15. Doubling the rate of tax lowers the net rate of return, and the representative agent responds by dissaving, immediately raising consumption. From this high point consumption gradually declines and savings increase in response to the subsequent rise in the rate of return to capital. The rate of return to capital which is merely its marginal product increases because the capital stock deteriorates as long as the total amount saved is insufficient to compensate for depreciation. Approximately halfway between the time the policy is implemented and the old policy is restored, consumption drops below its long-run value as agents anticipate the further rise in net returns that will follow the expiration of the policy. Finally, once the old tax rate is restored, consumption increases once again, this time gradually, until it returns to its initial level. Of course the longer the period over which the period is implemented, the greater the initial instantaneous increase in consumption. Similarly the less curvature in the utility function the greater the initial response and subsequent adjustments.

The linear approximation of consumption following the doubling the tax rate for five years in Figure 8 appears to roughly coincide with the impulse responses generated with the additional higher moments—particularly if the value of σ is high. However as the tax increase is extended over a longer period Figures 9 through 11 reveal that the impulse response generated by only the first moment underestimates the initial increase in consumption as well as its subsequent gradual decline. The result is that by omitting higher moments in the approximation, we risk underestimating the volatility in the path of consumption such an increase in taxes generates.

Now consider the temporary elimination of the tax in Figures 12-15. The immediate increase in capital's net rate of return, provides a strong incentive for higher savings and lower consumption—consumption drops immediately with the announced change in policy. Over time the higher savings generates faster accumulation of capital and a decline in the marginal product of capital and therefore also its rate of return. Consumption gradually recovers until eventually as agents in the economy anticipate the resumption of the tax, consumption rises above its initial level and continues to rise until the tax is reimposed. From that point consumption declines gradually until eventually it returns to its initial level as the economy converges.

As in Figure 8, the difference between the different impulse responses in each graph in Figure 12, where the tax cut lasts only five years, is not very large. However as the time period lengthens in Figures 13-15, the differences between the approximated paths of consumption become more pronounced across the different degrees of approximation. Unlike before, the

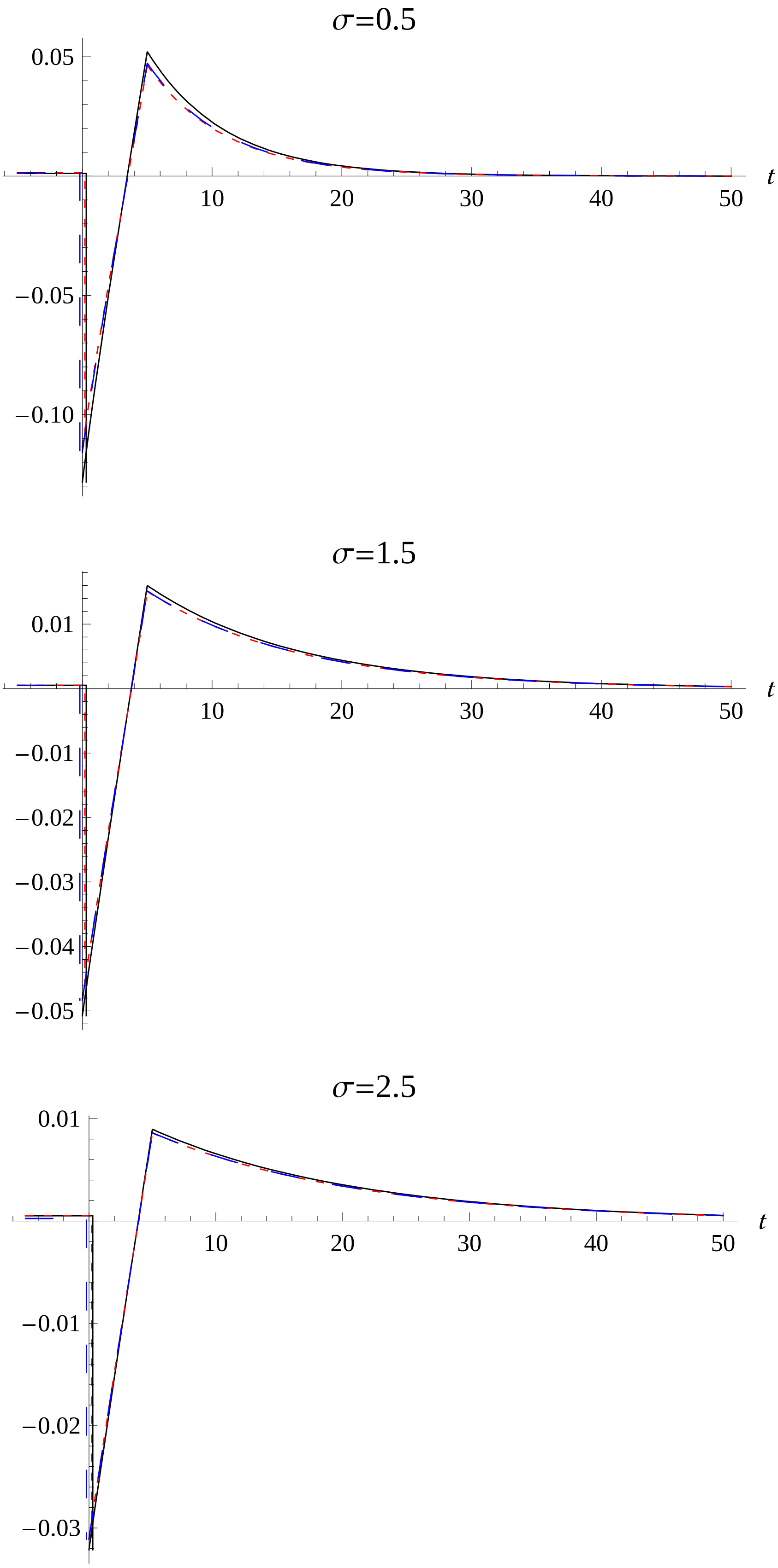

Figure 12: The behavior of consumption following the reduction of the tax on income from capital from 0.35 to zero for $T = 5$ years. The solid black curve represents a first order approximation, the curve with small red dashes, a second order approximation, the large blue dashes, a third order approximation.

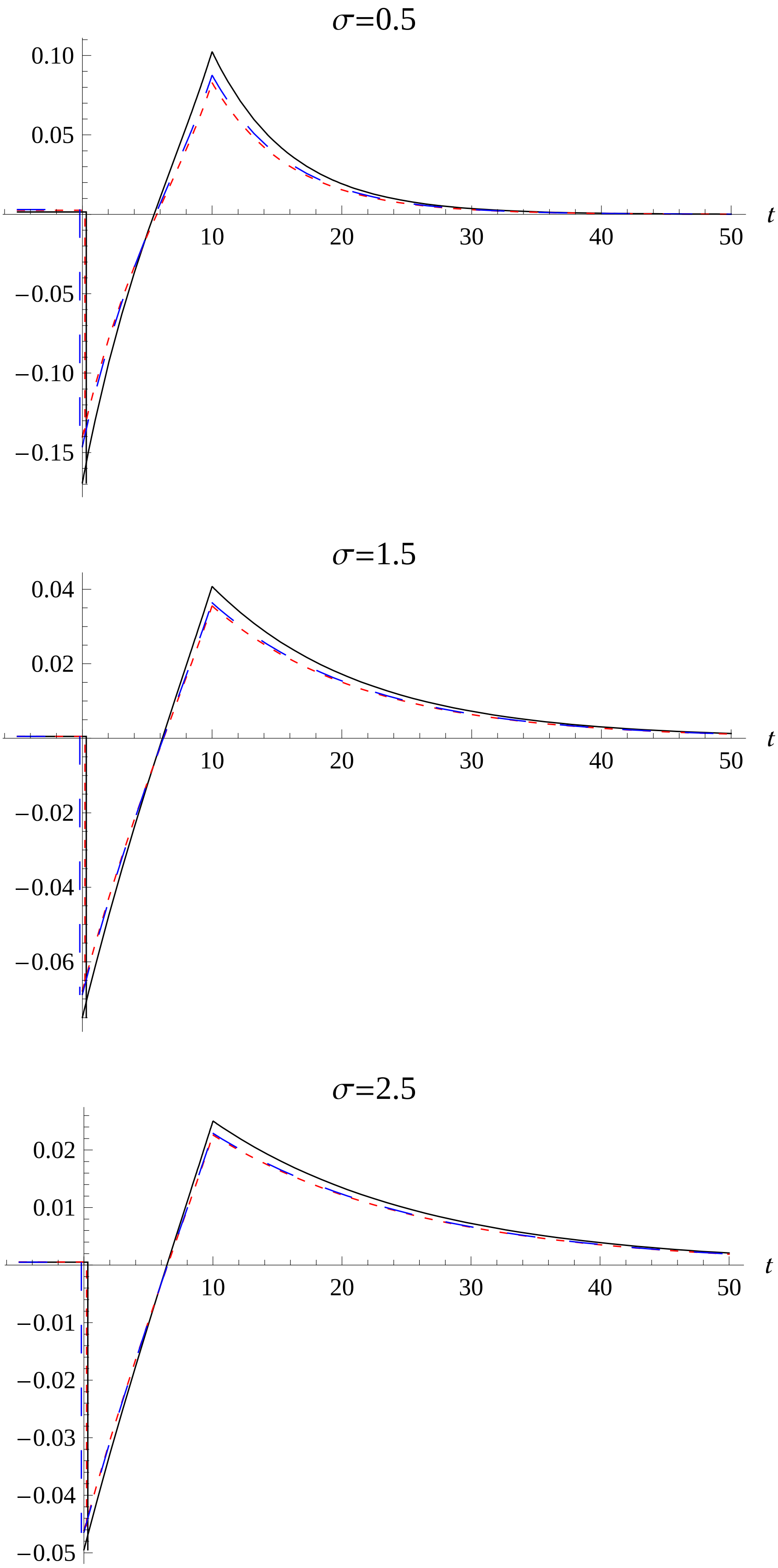

Figure 13: The behavior of consumption following the reduction of the tax on income from capital from 0.35 to zero for $T = 10$ years. The solid black curve represents a first order approximation, the curve with small red dashes, a second order approximation, the large blue dashes, a third order approximation.

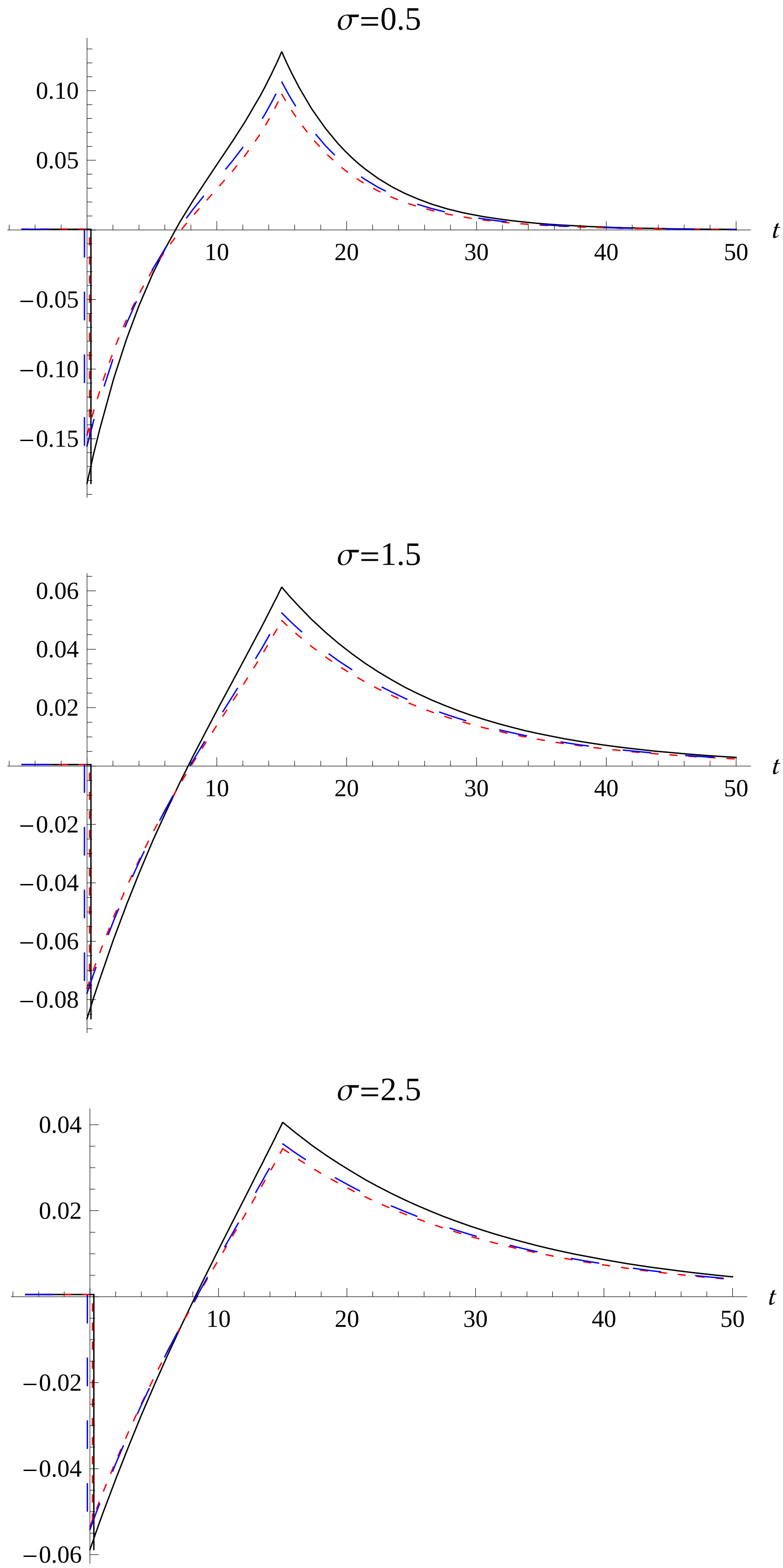

Figure 14: The behavior of consumption following the reduction of the tax on income from capital from 0.35 to zero for $T = 15$ years. The solid black curve represents a first order approximation, the curve with small red dashes, a second order approximation, the large blue dashes, a third order approximation.

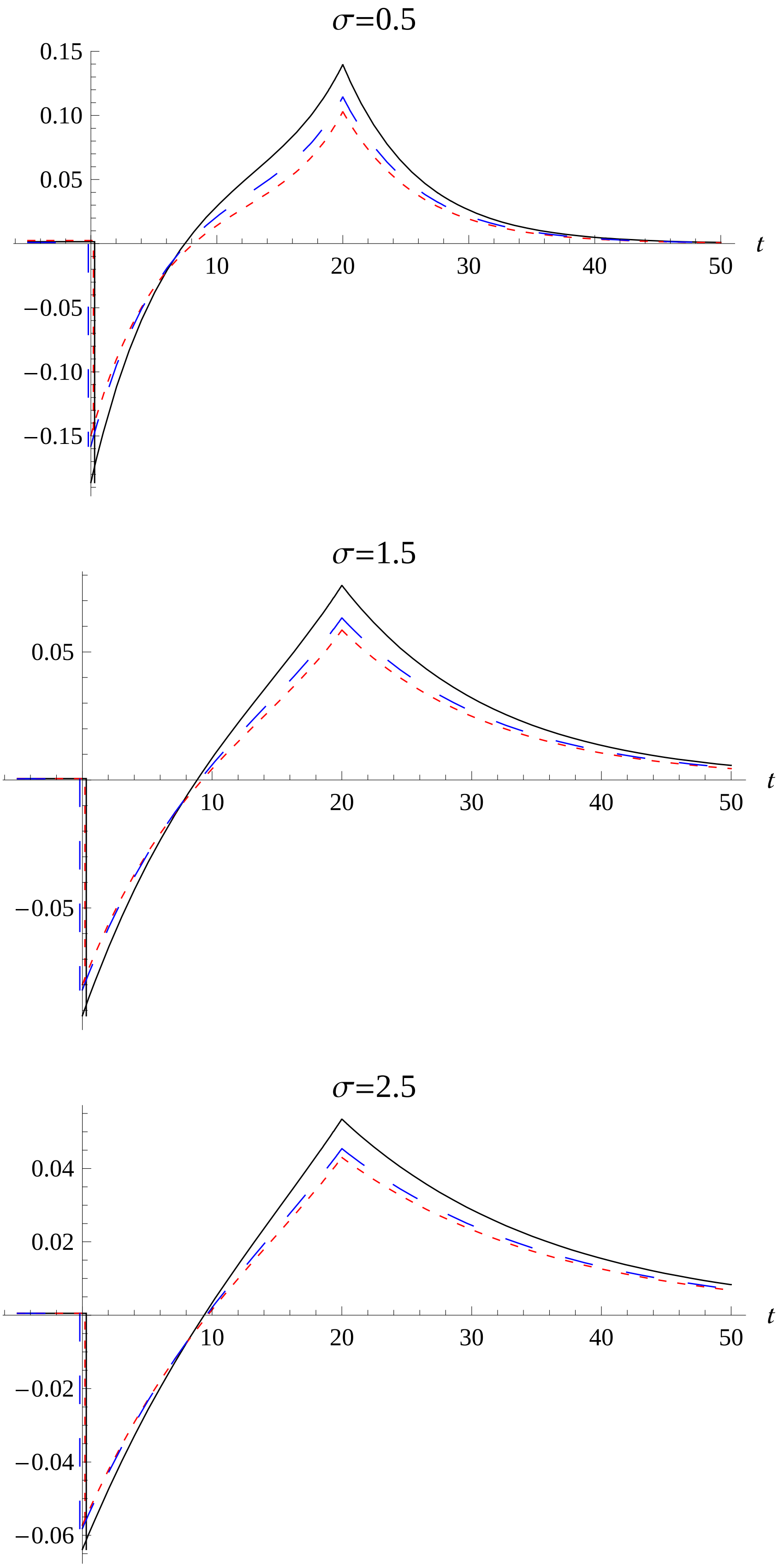

Figure 15: The behavior of consumption following the reduction of the tax on income from capital from 0.35 to zero for $T = 20$ years. The solid black curve represents a first order approximation, the curve with small red dashes, a second order approximation, the large blue dashes, a third order approximation.

linear approximation now overstates the volatility of consumption that the tax cut generates. The second order approximation yields both a lower initial drop in consumption, and a lower increase in consumption at the point where the policy expires. The third moment, rather than reinforcing the effect of the second moment, as was the case for the tax increase, now slightly moderates its effect.

5 Welfare Calculations

In practice, the lower order approximations in Figures 8 through 15 are often sufficiently accurate for most positive economic analyses. Not so for normative analyses—as I will demonstrate here, the subtle nonlinearities that can be safely ignored in other contexts have profound implications when analyzing welfare effects of large policy changes. Furthermore, the more direct numerically driven methods for approximation such as numerical shooting do not produce explicit formulae, but rather specific numerical values. When evaluating welfare in the context of a continuous time model, the use of numerical shooting to analyze the behavior of the model necessitates the use of interpolation between each point as the utility function of the time path of consumption is integrated. Thus the accuracy of the welfare calculations associated with each change in policy is inversely related to the size of the gaps between the numerical values. This creates a second source of inaccuracy. By contrast, perturbations methods produce explicit formulae describing the time path of consumption whose explicit integral can often be expressed algebraically (at least the first moment as in (30) and (31), or (32) and (33)). In this example all the moments that approximate the time path of consumption can be characterized as two simple sums of exponential functions, each multiplied by an indicator function.

To calculate the change in welfare generated by the changes in fiscal policy I calculate the compensating differential q, the fractional change in the value of initial steady state consumption $\bar{c}$ necessary to equal the utility generated by the time path of consumption $c(t;\epsilon)$, following the change in policy:

$$\int_0^\infty e^{-\rho t}\frac{\left((1+q)\,\bar{c}\right)^{1-\sigma}}{1-\sigma}dt = \int_0^\infty e^{-\rho t}\frac{c(t;\epsilon)^{1-\sigma}}{1-\sigma}dt.$$

Solving for q yields:

$$q = \left[\rho\int_0^\infty e^{-\rho t}\left(\frac{c(t;\epsilon)}{\bar{c}}\right)^{1-\sigma}dt\right]^{\frac{1}{1-\sigma}} - 1.$$

Between Tables 1 to 4, and Figures 16 to 20, I vary the time period T of the tax changes between five, ten, fifteen, and twenty years. In each I present the values of the compensating differential q as the temporary tax rate on capital income varies between zero to 0.9 for the values of $\sigma = 0.5$, 1.5, 2.5, each corresponding to a time path of consumption calculated as first,

Tax		$\sigma = 0.5$			$\sigma = 1.5$			$\sigma = 2.5$	
Rate	First	Second	Third	First	Second	Third	First	Second	Third
0	0.0012	0.0010	0.0010	0.0008	0.0007	0.0007	0.0006	0.0006	0.0006
0.1	0.0009	0.0008	0.0008	0.0006	0.0005	0.0005	0.0005	0.0004	0.0004
0.2	0.0006	0.0005	0.0005	0.0004	0.0003	0.0004	0.0003	0.0003	0.0003
0.3	0.0002	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
0.4	-0.0002	-0.0002	-0.0002	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
0.5	-0.0007	-0.0007	-0.0007	-0.0004	-0.0004	-0.0004	-0.0003	-0.0003	-0.0003
0.6	-0.0012	-0.0014	-0.0014	-0.0007	-0.0008	-0.0008	-0.0005	-0.0006	-0.0006
0.7	-0.0017	-0.0021	-0.0022	-0.0010	-0.0011	-0.0011	-0.0008	-0.0008	-0.0008
0.8	-0.0023	-0.0029	-0.0031	-0.0014	-0.0015	-0.0016	-0.001	-0.0011	-0.0011
0.9	-0.0029	-0.0039	-0.0042	-0.0017	-0.002	-0.002	-0.0013	-0.0014	-0.0015

Table 1: Welfare effects q generated by changing the rate of taxation on income from capital from the baseline rate of 0.35 for five years.

second or third order approximations. Once again $\alpha = 0.4$, $\delta = 0.1$, $\rho = 0.04$ and the initial rate of taxation on income from capital is $\tau = 0.35$. For small changes all the approximations yield welfare effects that are symmetric, with little difference between the different approximations. A five year rise in the rate of taxation from 0.35 to 0.4 generates a loss in welfare equivalent to a permanent drop in consumption of between 0.06% if $\sigma = 0.5$, to 0.03% if $\sigma = 2.5$. The same rise in the rate of taxation extended to twenty years costs agents in the economy the equivalent of a permanent 0.14% in consumption if $\sigma = 0.5$, and 0.09% if $\sigma = 2.5$.[3] Lowering the tax rate to 0.3 for the same number of years generates a welfare increases of nearly identical magnitudes. Again it is important to emphasize that these welfare effects are generated by changes in the excess burden the tax on capital income generates, rather than the direct effects of taxation on net income, as all proceeds from the tax are returned to the same representative agent as direct transfer payments.

The extent to which the different degrees of approximation diverge depends somewhat on

[3]The magnitudes of both the modest welfare gains generated by a given tax cut and the relatively larger welfare losses generated by a similar-sized tax increase are inversely related to the value of σ. Hence the higher the intertemporal elasticity of substitution the more sensitive the economy to changes in the tax rate on capital income. Although this parameter plays no role in determining the steady state values of consumption and capital, it does determine the speed of convergence. The more willing the representative agent is to substitute consumption between periods in response to changes in the net rate of return, the faster the agent accumulates or disaccumulates capital in Figures 8 through 15.

Tax		$\sigma = 0.5$			$\sigma = 1.5$			$\sigma = 2.5$	
Rate	First	Second	Third	First	Second	Third	First	Second	Third
0	0.0034	0.0018	0.0023	0.0022	0.0016	0.0018	0.0018	0.0015	0.0015
0.1	0.0026	0.0017	0.0019	0.0017	0.0014	0.0014	0.0014	0.0012	0.0012
0.2	0.0016	0.0013	0.0014	0.0011	0.0010	0.0010	0.0009	0.0008	0.0008
0.3	0.0006	0.0005	0.0005	0.0004	0.0004	0.0004	0.0003	0.0003	0.0003
0.4	-0.0006	-0.0006	-0.0006	-0.0004	-0.0004	-0.0004	-0.0003	-0.0003	-0.0003
0.5	-0.0019	-0.0022	-0.0023	-0.0013	-0.0014	-0.0014	-0.0010	-0.0011	-0.0011
0.6	-0.0033	-0.0043	-0.0046	-0.0022	-0.0026	-0.0027	-0.0017	-0.0020	-0.0020
0.7	-0.0048	-0.0069	-0.0076	-0.0032	-0.004	-0.0042	-0.0026	-0.0030	-0.0031
0.8	-0.0064	-0.0100	-0.0116	-0.0043	-0.0057	-0.0061	-0.0034	-0.0042	-0.0043
0.9	-0.0081	-0.0138	-0.0168	-0.0055	-0.0076	-0.0083	-0.0043	-0.0056	-0.0058

Table 2: Welfare effects q generated by changing the rate of taxation on income from capital from the baseline rate of 0.35 for ten years.

the curvature of the utility function but primarily on how long the tax rate changes. In Table 1, the value of the compensating differentials does not vary substantially between the different columns following a reduction in taxation for a period of five years. This is true across the first row, which corresponds to the welfare effects of completely eliminating the tax during this period. The same is true for doubling the tax rate. Only if the tax rate rises to 90% do we see significant evidence of divergence and then only if the curvature parameter σ is set to 0.5. Even then as is evident in Figure 16, the second order approximation is sufficient to capture nearly all the non-linearity that might affect welfare.

Double the time over which the tax rate changes to ten years and more significant differences emerge. For $\sigma = 0.5$, a decade long elimination of the tax on capital raises welfare by the equivalent of a 0.34% rise in consumption according to the approximation generated with a first order perturbation, but only 0.18% according to the calculation obtained with a second order perturbation, and 0.23% when the third order perturbation is included. Raise the rate of taxation to 60% or higher and we see similar divergence. Hence obtaining an accurate estimate of the welfare effects of large changes in taxes for a period of ten years or longer necessitates the inclusion of at very least second, if not third order perturbations.

This last conclusion is further reinforced by the results in Tables 3 and 4. In each instance and for every value of σ, the first order approximations significantly overestimate the welfare

Tax Rate	$\sigma = 0.5$ First	$\sigma = 0.5$ Second	$\sigma = 0.5$ Third	$\sigma = 1.5$ First	$\sigma = 1.5$ Second	$\sigma = 1.5$ Third	$\sigma = 2.5$ First	$\sigma = 2.5$ Second	$\sigma = 2.5$ Third
0	0.0056	0.0020	0.0036	0.0038	0.0023	0.0028	0.0031	0.0021	0.0024
0.1	0.0042	0.0023	0.0029	0.0029	0.0021	0.0023	0.0024	0.0019	0.0019
0.2	0.0027	0.0020	0.0021	0.0019	0.0016	0.0016	0.0015	0.0013	0.0013
0.3	0.0009	0.0009	0.0009	0.0007	0.0006	0.0006	0.0005	0.0005	0.0005
0.4	-0.0010	-0.0011	-0.0011	-0.0007	-0.0007	-0.0007	-0.0006	-0.0006	-0.0006
0.5	-0.0031	-0.0038	-0.0040	-0.0022	-0.0026	-0.0026	-0.0018	-0.0020	-0.0020
0.6	-0.0053	-0.0075	-0.0082	-0.0039	-0.0049	-0.0051	-0.0032	-0.0038	-0.0039
0.7	-0.0077	-0.0123	-0.0142	-0.0057	-0.0078	-0.0084	-0.0046	-0.0060	-0.0063
0.8	-0.0103	-0.0181	-0.0223	-0.0076	-0.0113	-0.0126	-0.0062	-0.0085	-0.0092
0.9	-0.0130	-0.0251	-0.0330	-0.0097	-0.0154	-0.0180	-0.0079	-0.0115	-0.0127

Table 3: Welfare effects q generated by changing the rate of taxation on income from capital from the baseline rate of 0.35 for fifteen years.

gains from the temporary abolition of the tax on capital. In the most extreme case with the value of σ set to 0.5, and the tax eliminated for twenty years, the first order approximation predicts a welfare gain equivalent to 0.78% permanent increase in consumption, whereas the third order approximation predicts a far more modest 0.48%.

Based on the third order approximation, doubling the tax rate on capital from 0.35 to 0.7 for twenty years, yields a welfare loss equivalent to a permanent decrease in consumption that ranges between just over two percent if the intertemporal elasticity of substitution is high ($\sigma = 0.5$), and just under one percent if the intertemporal elasticity of substitution is low ($\sigma = 2.5$). For the U.S., U.K., or Japanese economies this is the equivalent of between four quarters (in the first instance) and two quarters (in the second instance) of average per-capita output growth. By contrast should policy makers rely on only first order approximations they would conclude that the welfare losses from eliminating the tax are perhaps only half as big as they are likely to be.

The difference between the compensating differentials estimated using first and second order perturbations versus including third order perturbations as well, is far smaller than the difference between the compensating differentials estimated using only the first order perturbation and the first and second order perturbations. This indicates that with the addition of each successive moment, there is convergence to increasingly more accurate approximations of welfare changes.

Tax	$\sigma = 0.5$			$\sigma = 1.5$			$\sigma = 2.5$		
Rate	First	Second	Third	First	Second	Third	First	Second	Third
0	0.0078	0.0019	0.0048	0.0054	0.0025	0.0036	0.0044	0.0025	0.0031
0.1	0.0058	0.0027	0.0038	0.0041	0.0026	0.0030	0.0034	0.0024	0.0026
0.2	0.0036	0.0025	0.0027	0.0026	0.0021	0.0022	0.0022	0.0018	0.0018
0.3	0.0013	0.0011	0.0011	0.0009	0.0009	0.0009	0.0008	0.0007	0.0007
0.4	-0.0013	-0.0014	-0.0014	-0.0010	-0.0010	-0.0011	-0.0008	-0.0009	-0.0009
0.5	-0.0040	-0.0053	-0.0055	-0.0031	-0.0037	-0.0038	-0.0026	-0.0030	-0.0031
0.6	-0.0070	-0.0104	-0.0116	-0.0054	-0.0073	-0.0077	-0.0045	-0.0058	-0.0060
0.7	-0.0101	-0.0171	-0.0203	-0.0079	-0.0117	-0.0130	-0.0067	-0.0092	-0.0099
0.8	-0.0134	-0.0253	-0.0324	-0.0106	-0.0172	-0.0201	-0.0090	-0.0133	-0.0149
0.9	-0.0169	-0.0353	-0.0486	-0.0135	-0.0237	-0.0293	-0.0114	-0.0183	-0.0212

Table 4: Welfare effects q generated by changing the rate of taxation on income from capital from the baseline rate of 0.35 for twenty years.

Nonetheless in Table 3, for $\sigma = 0.5$, the welfare gain from reducing the tax rate from the baseline rate of 0.35 to 0.1 is slightly larger than the welfare gain generated by reducing it to zero. This phenomenon emerges in Table 4 as well for both $\sigma = 0.5$ and $\sigma = 1.5$, and in each case implies that reducing the magnitude of a distortionary tax is preferable to eliminating it. Such welfare reversals are clearly spurious and underline the need to consider third or even higher approximations in these instances.

Another issue at least partially obscured if higher order moments are ignored, is the asymmetry between the welfare gains generated by tax cuts, and the losses generated by tax increases. This can be seen in the way the curves in Figures 16 to 20 become progressively less linear as we pass from simulations based on first order perturbations only to simulations that include second and then third order perturbations as well. This nonlinearity is precisely analogous to the Harberger triangle—the property of the excess burden increasing at an approximately quadratic rate in the magnitude of the distortionary tax.

6 Conclusion

The results derived in this article demonstrate not only the method for analyzing high order approximations of dynamic non-linear models, but the pitfalls of failing to account for high order nonlinearities when considering the welfare effects of policy changes. It is important to

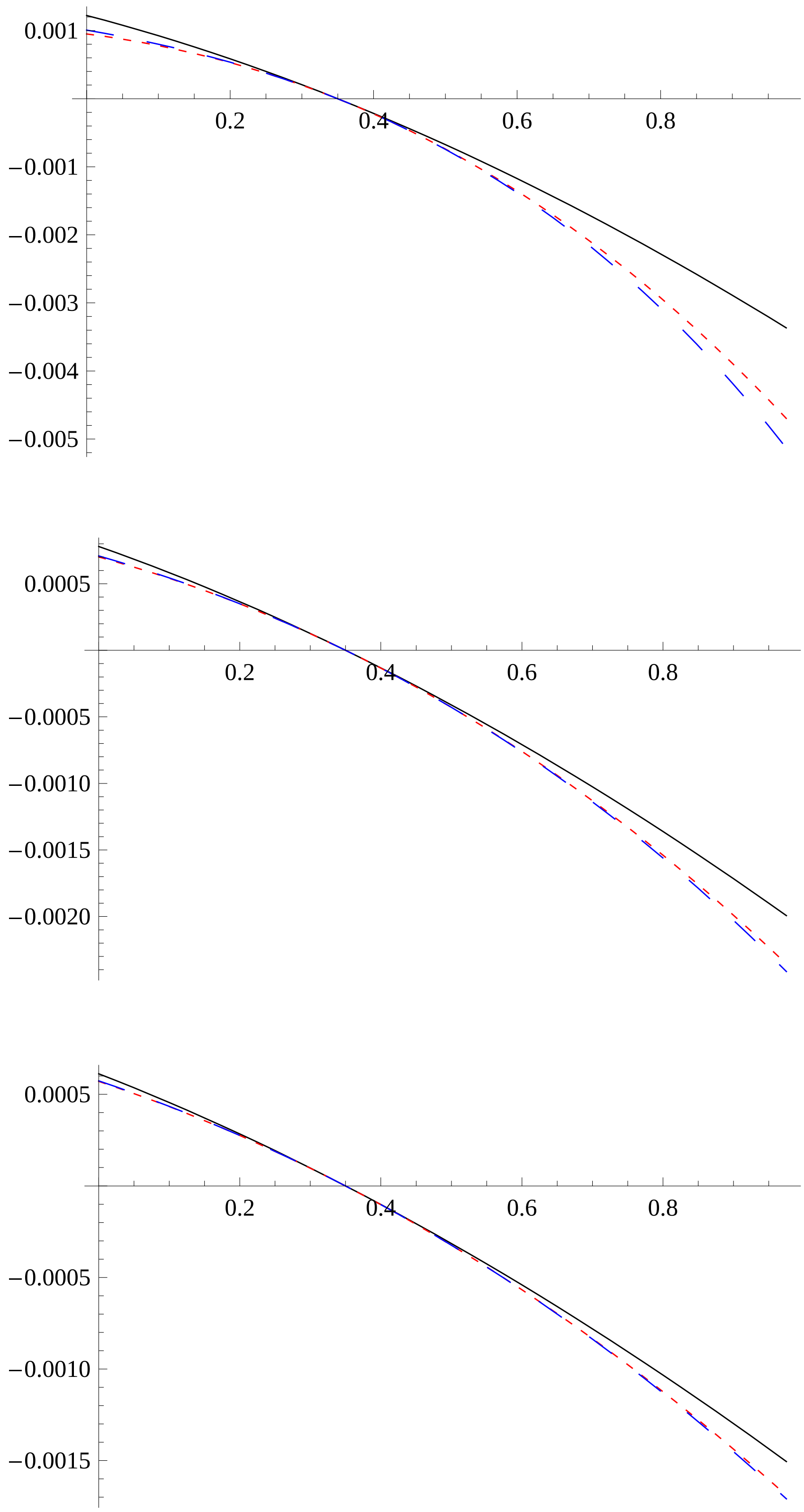

Figure 16: The compensating differential q (on the vertical axis) representing welfare effects of changing the tax rate (on the horizontal axis) from its baseline rate of 0.35 for five years Using the first order (solid black curve), second order (small red dashes) and third order (large blue dashes) perturbation methods.

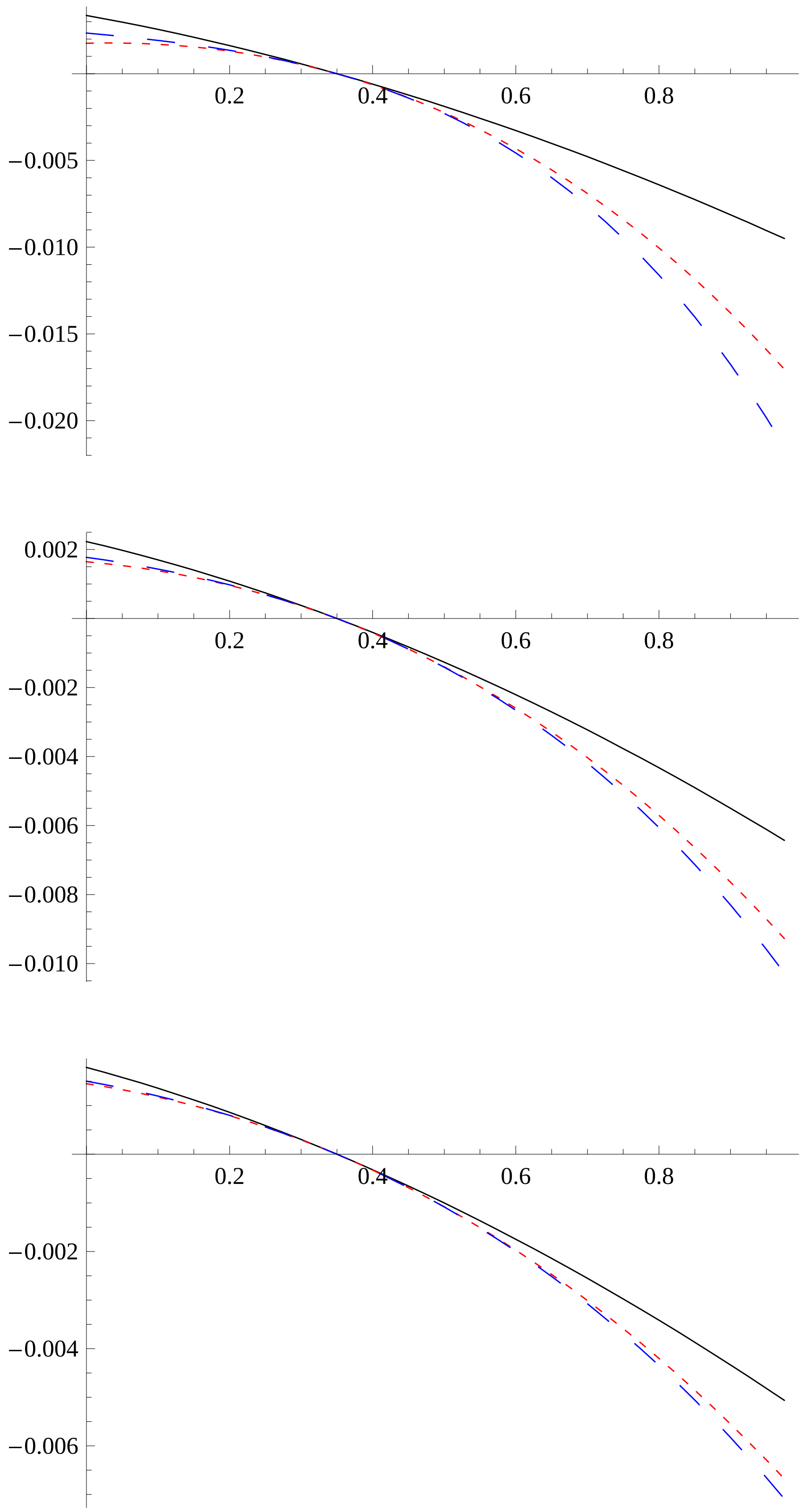

Figure 17: The compensating differential q (on the vertical axis) representing welfare effects of changing the tax rate (on the horizontal axis) from its baseline rate of 0.35 for ten years Using the first order (solid black curve), second order (small red dashes) and third order (large blue dashes) perturbation methods.

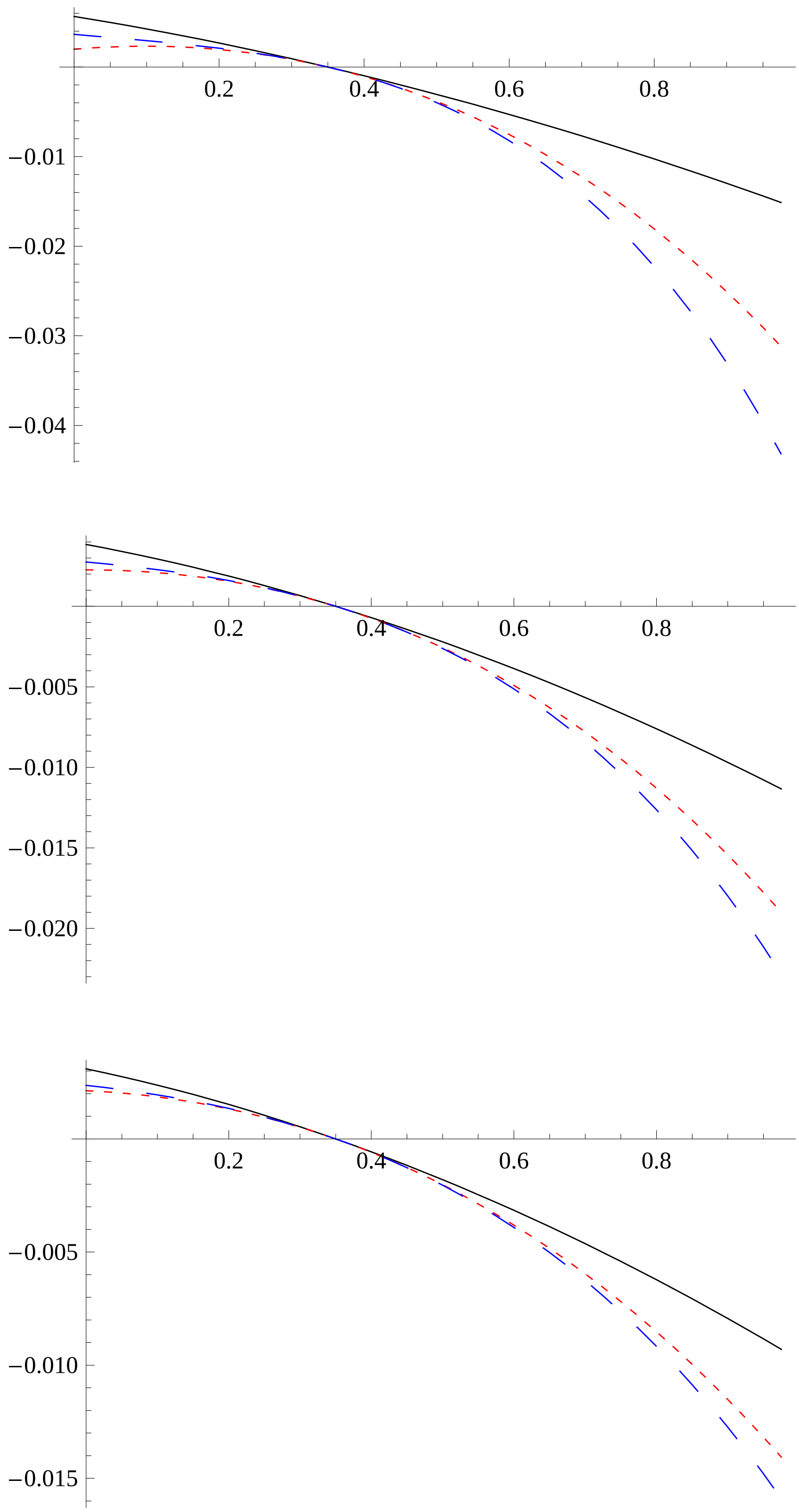

Figure 18: The compensating differential q (on the vertical axis) representing welfare effects of changing the tax rate (on the horizontal axis) from its baseline rate of 0.35 for fifteen years Using the first order (solid black curve), second order (small red dashes) and third order (large blue dashes) perturbation methods.

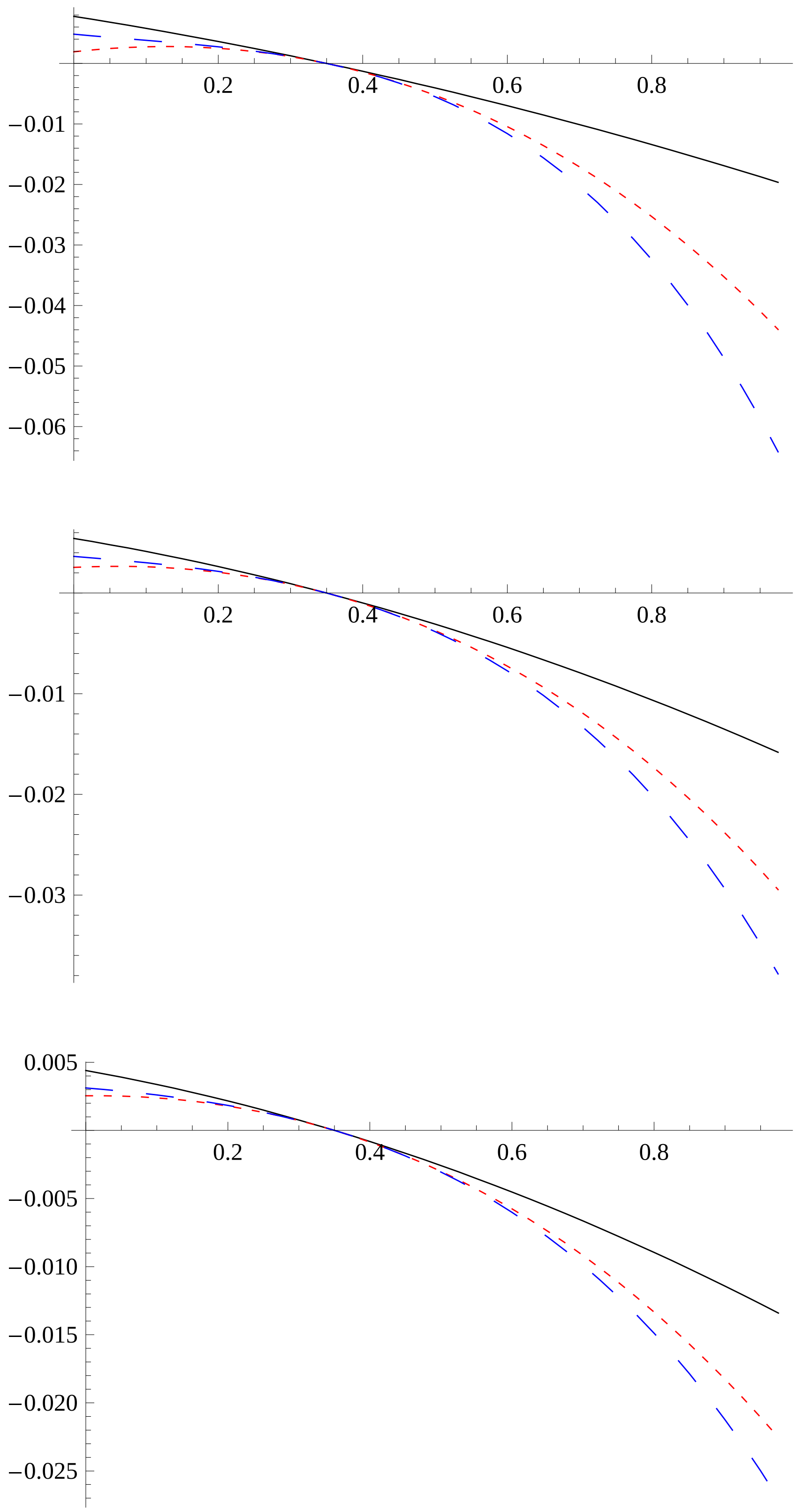

Figure 19: The compensating differential q (on the vertical axis) representing welfare effects of changing the tax rate (on the horizontal axis) from its baseline rate of 0.35 for twenty years Using the first order (solid black curve), second order (small red dashes) and third order (large blue dashes) perturbation methods.

emphasize that the model as well as the policy changes considered here are each about the simplest possible.

A more complicated change in the tax rate, for example one that whose duration is unknown, or one whose implementation is preceded by a long delay will produce far more complicated dynamics, necessitating high order approximations even if the magnitudes of the policy changes are far more modest. Similarly, this model has only a single sector, a single representative agent and a single simple distortion. In any richer model, one with heterogenous agents, multiple sectors, external effects, more activities taxed, or graduated tax rates, linear or even quadratic approximations will yield misleading welfare predictions, possibly even spurious welfare reversals, though the proposed changes in policy might well be relatively simple in nature or far more modest in scale than the changes in the tax on capital income considered here.

References

[1] Ben-Gad, Michael, "The impact of immigrant dynasties on wage inequality," *Research in Labor Economics* 24, 2006, 77-134.

[2] Ben-Gad, Michael "Capital-skill complementarities and the immigration surplus," *Review of Economic Dynamics* 11, 2008, 335-365.

[3] Ben-Gad, Michael, "The economic effects of immigration—a dynamic analysis," *Journal of Economic Dynamics and Control* 28, 2004, 1825-1845.

[4] Judd, Kenneth L., "An alternative to steady-state comparisons in perfect foresight models," *Economics Letters* 10, 1982, 55-59.

[5] Judd, Kenneth L., "The welfare cost of factor taxation in a perfect-foresight model," *Journal of Public Economics* 28, 1985, 59-83.

[6] Judd, Kenneth L., "Redistributive taxation in a simple perfect foresight model," *Journal of Political Econ*omy 95, 1987, 675-709.

[7] Judd, Kenneth L., *Numerical Methods in Economics*, MIT Press, 1998.

[8] Kim, Jinill, and Sunghyun Henry Kim, "Spurious welfare reversals in international business cycle models," *Journal of International Economics* 60, 2003, 471-500.

[9] Tesar, Linda, "Evaluating the gains from international risksharing," *Carnegie-Rochester Conference Series on Public Policy*, 42, 95-143.

7 Appendix

The moments of the approximation for consumption for the Ramsey optimal growth model with $\alpha = 0.4$, $\delta = 0.1$, $\rho = 0.04$ and $\tau = 0.35$. Note that these are all the sum of exponential functions. The third order approximation of the time path of consumption is the sum of these moments weighted by ϵ, $\epsilon^2/2$, and $\epsilon^3/6$.

7.1 $\sigma = 0.5$

$\boldsymbol{T = 5}$ **:**

$$c_\epsilon(t) = \left[-0.2724 + 0.8768e^{-0.1673t} - 0.09962e^{0.2288t}\right] U(5-t) - 0.474e^{-0.1673t} U(t-5)$$

$$\begin{aligned} c_{\epsilon\epsilon}(t) = &\left[e^{0.2288t}(-0.007557t - 0.005527) + 0.07867e^{-0.3346t} - 0.09758e^{0.06154t}\right. \\ &\left.+0.02444e^{0.4577t} + e^{-0.1673t}(0.09951t + 1.158) - 0.8316\right] U(5-t) \\ &+ \left[0.023e^{-0.3346t} - 0.3037e^{-0.1673t}\right] U(t-5) \end{aligned}$$

$$\begin{aligned} c_{\epsilon\epsilon\epsilon}(t) = e^{-1.11t} &\left[e^{1.171t}(-0.02772t - 0.7098) - 0.02657e^{0.6077t} + 0.06589e^{1.004t} - 3.412e^{1.11t}\right. \\ &+ 0.0555e^{1.4t} - 0.01088e^{1.796t} + e^{1.567t}(0.005562t + 0.06804) + e^{0.775t}(0.02678t - 0.6782) \\ &\left.-0.000430e^{1.338t}(t - 9.16)(t + 196.5) + e^{0.9423t}\left(0.008469t^2 + 0.2601t + 4.252\right)\right] U(5-t) \\ &+ \left[0.004199e^{-0.5019t} + 0.04420e^{-0.3346t} - 0.3823e^{-0.1673t}\right] U(t-5) \end{aligned}$$

$\boldsymbol{T = 10}$ **:**

$$c_\epsilon(t) = \left[-0.2724 + 0.9697e^{-0.1673t} - 0.03173e^{0.2288t}\right] U(10-t) - 2.148e^{-0.1673t} U(t-10)$$

$$\begin{aligned} c_{\epsilon\epsilon}(t) = &\left[e^{0.2288t}(-0.002407t - 0.002652) + 0.09622e^{-0.3346t} - 0.03437e^{0.06154t}\right. \\ &\left.+0.002479e^{0.4577t} + e^{-0.1673t}(0.11t + 1.421) - 0.8316\right] U(10-t) \\ &+ \left[0.4723e^{-0.3346t} - 2.439e^{-0.1673t}\right] U(t-10) \end{aligned}$$

$$\begin{aligned} c_{\epsilon\epsilon\epsilon}(t) = e^{-1.11t} &\left[e^{1.171t}(-0.009762t - 0.2589) - 0.03594e^{0.6077t} + 0.02566e^{1.004t} - 3.412e^{1.11t}\right. \\ &+ 0.006226e^{1.4t} - 0.0003514e^{1.796t} + e^{1.567t}(0.0005642t + 0.00711) + e^{0.775t}(0.03276t - 0.7878) \\ &\left.-0.0001369e^{1.338t}(t - 11.37)(t + 199.4) + e^{0.9423t}\left(0.009366t^2 + 0.3114t + 5.31\right)\right] U(10-t) \\ &+ \left[0.3909e^{-0.5019t} + 1.609e^{-0.3346t} - 5.212e^{-0.1673t}\right] U(t-10) \end{aligned}$$

$\boldsymbol{T = 15}$ **:**

$$c_\epsilon(t) = \left[-0.2724 + 0.9993e^{-0.1673t} - 0.01010e^{0.2288t}\right] U(15-t) - 6.198e^{-0.1673t} U(t-15)$$

$$\begin{aligned} c_{\epsilon\epsilon}(t) = &\left[e^{0.2288t}(0.0008922 - 0.0007665t) + 0.1022e^{-0.3346t} - 0.01128e^{0.06154t}\right. \\ &\left.+0.0002514e^{0.4577t} + e^{-0.1673t}(0.1134t + 1.521) - 0.8316\right] U(15-t) \\ &+ \left(3.931e^{-0.3346t} - 8.741e^{-0.1673t}\right) U(t-15) \end{aligned}$$

$$\begin{aligned} c_{\epsilon\epsilon\epsilon}(t) = e^{-1.11t} &\left[e^{1.171t}(-0.003204t - 0.08302) - 0.03933e^{0.6077t} + 0.00868e^{1.004t} - 3.412e^{1.11t}\right. \\ &+ 0.0006508e^{1.4t} - 0.00001135e^{1.796t} + e^{1.567t}(0.00005722t + 0.0005915) + e^{0.775t}(0.03479t - 0.8191) \\ &\left.-0.0000436e^{1.338t}(t - 14.62)(t + 198.1) + e^{0.9423t}\left(0.009652t^2 + 0.3306t + 5.738\right)\right] U(15-t) \\ &+ \left[9.386e^{-0.5019t} + 16.63e^{-0.3346t} - 22.34e^{-0.1673t}\right] U(t-15) \end{aligned}$$

$\boldsymbol{T = 20}$ **:**

$$c_\epsilon(t) = \left[-0.2724 + 1.009e^{-0.1673t} - 0.003218e^{0.2288t}\right] U(20-t) - 15.61e^{-0.1673t} U(t-20)$$

$$c_{\epsilon\epsilon}(t) = \big[e^{0.2288t}(0.001202 - 0.0002441t) + 0.1041e^{-0.3346t} - 0.003626e^{0.06154t}$$
$$+0.00002550e^{0.4577t} + e^{-0.1673t}(0.1145t + 1.553) - 0.8316\big] U(20-t)$$
$$+ \left[24.92e^{-0.3346t} - 24.31e^{-0.1673t}\right] U(t-20)$$

$$c_{\epsilon\epsilon\epsilon}(t) = e^{-1.11t}\big[e^{1.171t}(-0.001030t - 0.02524) - 0.04045e^{0.6077t} + 0.002817e^{1.004t}$$
$$- 3.412e^{1.11t} + 0.00006663e^{1.4t} - \left(3.666\times 10^{-7}\right) e^{1.796t}$$
$$+ e^{1.567t}\left(\left(5.804\times 10^{-6}\right) t + 0.00003818\right) + e^{0.775t}(0.03545t - 0.8291)$$
$$-0.00001389e^{1.338t}(t-18.72)(t+194.7) + e^{0.9423t}\left(0.009743t^2 + 0.3368t + 5.86\right)\big] U(20-t)$$
$$+ \left[149.8e^{-0.5019t} + 116.5e^{-0.3346t} - 67.22e^{-0.1673t}\right] U(t-20)$$

7.2 $\sigma = 1.5$

$T = 5$:

$$c_\epsilon(t) = \left[-0.2724 + 0.5393e^{-0.08631t} - 0.06725e^{0.1479t}\right] U(5-t) - 0.09696e^{-0.08631t} U(t-5)$$

$$c_{\epsilon\epsilon}(t) = \big[e^{0.1479t}(0.1047 - 0.006769t) - 0.06753e^{-0.1726t} - 0.1216e^{0.06154t}$$
$$+0.01830e^{0.2957t} + e^{-0.08631t}(0.04784t + 0.9560) - 0.8316\big] U(5-t)$$
$$+ \left[-0.002183e^{-0.1726t} - 0.03029e^{-0.08631t}\right] U(t-5)$$

$$c_{\epsilon\epsilon\epsilon}(t) = e^{-0.5427t}\big[e^{0.6042t}(-0.03453t - 0.6726) + 0.04753e^{0.2837t} - 0.05374e^{0.5179t} - 3.412e^{0.5427t}$$
$$+ 0.08575e^{0.752t} - 0.009797e^{0.9862t} + e^{0.37t}(-0.01797t - 0.7028) + e^{0.8384t}(0.005525t - 0.02099)$$
$$-0.0005110e^{0.6905t}(t-20.63)(t+89.50) + e^{0.4563t}\left(0.003182t^2 + 0.1896t + 3.828\right)\big] U(5-t)$$
$$+ \left[-0.0002762e^{-0.2589t} - 0.002046e^{-0.1726t} - 0.01888e^{-0.08631t}\right] U(t-5)$$

$T = 10$:

$$c_\epsilon(t) = \left[-0.2724 + 0.5995e^{-0.08631t} - 0.03211e^{0.1479t}\right] U(10-t) - 0.3802e^{-0.08631t} U(t-10)$$

$$c_{\epsilon\epsilon}(t) = \big[e^{0.1479t}(0.05210 - 0.003232t) - 0.08344e^{-0.1726t} - 0.06452e^{0.06154t}$$
$$+0.004171e^{0.2957t} + e^{-0.08631t}(0.05318t + 1.083) - 0.8316\big] U(10-t)$$
$$+ \left[-0.03355e^{-0.1726t} - 0.2643e^{-0.08631t}\right] U(t-10)$$

$$c_{\epsilon\epsilon\epsilon}(t) = e^{-0.5427t}\big[e^{0.37t}(-0.02220t - 0.8771) + 0.06528e^{0.2837t} - 0.03171e^{0.5179t}$$
$$- 3.412e^{0.5427t} + 0.02173e^{0.752t} - 0.001066e^{0.9862t}$$
$$+ e^{0.6042t}(-0.01833t - 0.3540) + e^{0.8384t}(0.00126t - 0.005611)$$
$$-0.0002440e^{0.6905t}(t-22.92)(t+90.47) + e^{0.4563t}\left(0.003537t^2 + 0.2135t + 4.266\right)\big] U(10-t)$$
$$+ \left[-0.01664e^{-0.2589t} - 0.06996e^{-0.1726t} - 0.3544e^{-0.08631t}\right] U(t-10)$$

$T = 15$:

$$c_\epsilon(t) = \left[-0.2724 + 0.6283e^{-0.08631t} - 0.01533e^{0.1479t}\right] U(15-t) - 0.8801e^{-0.08631t}U(t-15)$$

$$c_{\epsilon\epsilon}(t) = \left[e^{0.1479t}(0.02703 - 0.001543t) - 0.09163e^{-0.1726t} - 0.03228e^{0.06154t} \right.$$
$$\left. +0.0009510e^{0.2957t} + e^{-0.08631t}(0.05573t + 1.161) - 0.8316\right] U(15-t)$$
$$+ \left[-0.1798e^{-0.1726t} - 0.8874e^{-0.08631t}\right] U(t-15)$$

$$c_{\epsilon\epsilon\epsilon}(t) = e^{-0.5427t} \left[e^{0.37t}(-0.02438t - 0.9745) + 0.07513e^{0.2837t} - 0.01663e^{0.5179t} - 3.412e^{0.5427t} \right.$$
$$+ 0.005191e^{0.752t} - 0.0001161e^{0.9862t} + e^{0.6042t}(-0.009169t - 0.1723) + e^{0.8384t}(0.0002872t - 0.001680)$$
$$\left. -0.0001165e^{0.6905t}(t - 25.16)(t + 89.92) + e^{0.4563t}\left(0.003707t^2 + 0.2272t + 4.556\right)\right] U(15-t)$$
$$+ \left[-0.2065e^{-0.2589t} - 0.5439e^{-0.1726t} - 1.659e^{-0.08631t}\right] U(t-15)$$

$\boldsymbol{T = 20:}$

$$c_\epsilon(t) = \left[-0.2724 + 0.642e^{-0.08631t} - 0.007320e^{0.1479t}\right] U(20-t) - 1.68e^{-0.08631t}U(t-20)$$

$$c_{\epsilon\epsilon}(t) = \left[e^{0.1479t}(0.01469 - 0.0007368t) - 0.09568e^{-0.1726t} - 0.01575e^{0.06154t} \right.$$
$$\left. +0.0002168e^{0.2957t} + e^{-0.08631t}(0.05694t + 1.202) - 0.8316\right] U(20-t)$$
$$+ \left(-0.6555e^{-0.1726t} - 2.090e^{-0.08631t}\right) U(t-20)$$

$$c_{\epsilon\epsilon\epsilon}(t) = e^{-0.5427t} \left[e^{0.37t}(-0.02546t - 1.024) + 0.08016e^{0.2837t} - 0.008289e^{0.5179t} \right.$$
$$- 3.412e^{0.5427t} + 0.001209e^{0.752t} - 0.00001264e^{0.9862t}$$
$$+ e^{0.6042t}(-0.004474t - 0.07886) + e^{0.8384t}(0.00006547t - 0.0005414)$$
$$\left. -0.00005562e^{0.6905t}(t - 27.90)(t + 87.82) + e^{0.4563t}\left(0.003788t^2 + 0.2341t + 4.706\right)\right] U(20-t)$$
$$+ \left[-1.438e^{-0.2589t} - 2.446e^{-0.1726t} - 4.669e^{-0.08631t}\right] U(t-20)$$

7.3 $\sigma = 2.5$

$\boldsymbol{T = 5:}$

$$c_\epsilon(t) = \left[-0.2724 + 0.4483e^{-0.06199t} - 0.04944e^{0.1235t}\right] U(5-t) - 0.04810e^{-0.06199t}U(t-5)$$

$$c_{\epsilon\epsilon}(t) = \left[e^{0.1235t}(0.1314 - 0.005526t) - 0.07117e^{-0.1240t} - 0.1238e^{0.06154t} \right.$$
$$\left. +0.01322e^{0.2471t} + e^{-0.06199t}(0.03478t + 0.9063) - 0.8316\right] U(5-t)$$
$$+ \left[-0.0008194e^{-0.124t} - 0.01037e^{-0.06199t}\right] U(t-5)$$

$$c_{\epsilon\epsilon\epsilon}(t) = e^{-0.3724t} \left[e^{0.4339t}(-0.03517t - 0.5268) + 0.04610e^{0.1864t} - 0.1415e^{0.3719t} - 3.412e^{0.3724t} \right.$$
$$+ 0.08688e^{0.5574t} - 0.007174e^{0.7429t} + e^{0.2484t}(-0.01656t - 0.5880) + e^{0.6194t}(0.004432t - 0.05248)$$
$$\left. -0.0004632e^{0.4959t}(t - 29.35)(t + 63.46) + e^{0.3104t}\left(0.002024t^2 + 0.1543t + 3.742\right)\right] U(5-t)$$
$$+ \left[-0.00005695e^{-0.186t} - 0.0005299e^{-0.124t} - 0.004432e^{-0.06199t}\right] U(t-5)$$

$\boldsymbol{T = 10:}$

$$c_\epsilon(t) = \left[-0.2724 + 0.4937e^{-0.06199t} - 0.02666e^{0.1235t}\right] U(10-t) - 0.1831e^{-0.06199t} U(t-10)$$

$$\begin{aligned} c_{\epsilon\epsilon}(t) = &\left[e^{0.1235t}(0.07474 - 0.002980t) - 0.08631e^{-0.124t} - 0.07353e^{0.06154t}\right. \\ &\left. + 0.003843e^{0.2471t} + e^{-0.06199t}(0.03830t + 0.9883) - 0.8316\right] U(10-t) \\ &+ \left[-0.01187e^{-0.124t} - 0.09547e^{-0.06199t}\right] U(t-10) \end{aligned}$$

$$\begin{aligned} c_{\epsilon\epsilon\epsilon}(t) = e^{-0.3724t} &\left[e^{0.4339t}(-0.02088t - 0.2946) + 0.06157e^{0.1864t} - 0.09256e^{0.3719t} - 3.412e^{0.3724t}\right. \\ &+ 0.02782e^{0.5574t} - 0.001125e^{0.7429t} + e^{0.2484t}(-0.02009t - 0.7080) + e^{0.6194t}(0.001289t - 0.01694) \\ &\left. - 0.0002498e^{0.4959t}(t - 31.79)(t + 63.29) + e^{0.3104t}\left(0.002229t^2 + 0.1688t + 3.995\right)\right] U(10-t) \\ &+ \left[-0.003139e^{-0.186t} - 0.01857e^{-0.124t} - 0.09568e^{-0.06199t}\right] U(t-10) \end{aligned}$$

T = 15 :

$$c_\epsilon(t) = \left[-0.2724 + 0.5182e^{-0.06199t} - 0.01438e^{0.1235t}\right] U(15-t) - 0.4045e^{-0.06199t} U(t-15)$$

$$\begin{aligned} c_{\epsilon\epsilon}(t) = &\left[e^{0.1235t}(0.04259 - 0.001607t) - 0.09508e^{-0.124t} - 0.04161e^{0.06154t}\right. \\ &\left. + 0.001117e^{0.2471t} + e^{-0.06199t}(0.04020t + 1.045) - 0.8316\right] U(15-t) \\ &+ \left[-0.05793e^{-0.124t} - 0.3257e^{-0.06199t}\right] U(t-15) \end{aligned}$$

$$\begin{aligned} c_{\epsilon\epsilon\epsilon}(t) = e^{-0.3724t} &\left[e^{0.2484t}(-0.02213t - 0.7839) + 0.07119e^{0.1864t} - 0.05498e^{0.3719t} - 3.412e^{0.3724t}\right. \\ &+ 0.00849e^{0.5574t} - 0.0001764e^{0.7429t} + e^{0.4339t}(-0.01182t - 0.1577) + e^{0.6194t}(0.0003747t - 0.005459) \\ &\left. - 0.0001347e^{0.4959t}(t - 33.87)(t + 62.52) + e^{0.3104t}\left(0.002339t^2 + 0.1780t + 4.179\right)\right] U(15-t) \\ &+ \left[-0.03386e^{-0.186t} - 0.1399e^{-0.124t} - 0.4849e^{-0.06199t}\right] U(t-15) \end{aligned}$$

T = 20 :

$$c_\epsilon(t) = \left[-0.2724 + 0.5314e^{-0.06199t} - 0.007752e^{0.1235t}\right] U(20-t) - 0.7265e^{-0.06199t} U(t-20)$$

$$\begin{aligned} c_{\epsilon\epsilon}(t) = &\left[e^{0.1235t}(0.02473 - 0.0008664t) - 0.09999e^{-0.124t} - 0.02301e^{0.06154t}\right. \\ &\left. + 0.0003249e^{0.2471t} + e^{-0.06199t}(0.04123t + 1.078) - 0.8316\right] U(20-t) \\ &+ \left[-0.1869e^{-0.124t} - 0.7576e\right] U(t-20) \end{aligned}$$

$$\begin{aligned} c_{\epsilon\epsilon\epsilon}(t) = e^{-0.3724t} &\left[e^{0.2484t}(-0.02327t - 0.8284) + 0.07677e^{0.1864t} - 0.03118e^{0.3719t}\right. \\ &- 3.412e^{0.3724t} + 0.002531e^{0.5574t} - 0.00002765e^{0.7429t} \\ &+ e^{0.4339t}(-0.006536t - 0.0798) + e^{0.6194t}(0.0001089t - 0.001809) \\ &\left. - 0.00007263e^{0.4959t}(t - 36.13)(t + 60.71) + e^{0.3104t}\left(0.002399t^2 + 0.1833t + 4.292\right)\right] U(20-t) \\ &+ \left[-0.1962e^{-0.186t} - 0.5847e^{-0.124t} - 1.412e^{-0.06199t}\right] U(t-20) \end{aligned}$$