The Optimal Taxation of Asset Income when Government Consumption is Endogenous: Theory, Estimation and Welfare

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

December 2016

Abstract

This paper derives the Ramsey optimal policy for taxing asset income in a model where government expenditure is a function of net output or the inputs that produce it. Extending Judd (1999), I demonstrate that the canonical result that the optimal tax on capital income is zero in the medium to long term is a special case of a more general model. Employing a vector error correction model to estimate the relationship between government consumption and net output or the factor inputs that generate it for the United States between 1948Q1 to 2015Q4, I demonstrate that this special case is empirically implausible, and show how a cointegrating vector can be used to determine the optimal tax schedule. I simulate a version of the model using the empirical estimates to measure the welfare implications of changing the tax rate on asset income, and contrast these results with those generated in a version of the model where government consumption is purely exogenous. The shifting pattern of welfare measurements confirms the theoretical results. I calculate that the prevailing effective tax rate on net asset income in the US between 1970 and 2014 averaged 0.449. Hence abolishing the tax completely does generate welfare improvements, though only by the equivalent of between 1.103 and 1.616% percent permanent increase in consumption—well under half the implied welfare benefit when the endogeneity of the government consumption is ignored. The maximum welfare improvement from shifting part of the burden of tax from capital to labor is the equivalent of a permanent increase in consumption of between only 1.491 and 1.858%, and is attained when the tax rate on asset income is lowered to between 0.148 and 0.186. Allowing the tax rate to vary over time raises the maximum welfare benefit to 1.865%. All the results are very robust to a wide range of elasticities of labor supply.

JEL classification: E62, H21, H50

Keywords: Fiscal Policy; Optimal Taxation; Vector Error Correction Model

* *mbengad@city.ac.uk*. I would like to thank Giovanni Melina, Joseph Pearlman and Gustavo Sanchez as well as participants in presentations at City, University of London, The Annual ESRC/HMRC/HMT Tax Administration Research Centre Workshop 2014 at the University of Exeter, The Royal Economic Society Meetings, 2014, at the University of Manchester and The Society for Computational Economics Annual Conference, 2014 at BI Norwegian Business School for helpful comments and advice.

1 Introduction

If governments must fund their activities by taxing income, on which sources of income should the burden fall? In this paper I consider a general optimal growth model, one in which there is a direct link between either aggregate net output or the factor inputs that produce it, and the share of output allocated to government consumption. In such an economy the canonical results of zero taxation of capital income no longer hold. I demonstrate that for an empirically plausible specification of the link between government consumption and either net output or factor inputs, there is a simple relationship that can be employed to determine the optimal rate of tax on asset income and estimate a range of appropriate rates for the United States. Finally, I measure the welfare implications of shifting the burden of taxation between asset income and labor earnings. I demonstrate that the optimal tax rate on asset income is indeed positive, but that given the prevailing rates of taxation in the United States, the maximal welfare benefit that can be obtained from adopting an optimal policy is much smaller than what usually emerges when government consumption expenditure is assumed to be exogenously determined.

Atkinson and Stiglitz (1976) demonstrated how the tax on interest income depends on the complementarity or substitutability between consumption and leisure in a representative agent's instantaneous utility function. Additive separability between the two, implies the optimal tax rate on interest is zero. Chamley (1981) was the first to calculate the excess burden associated with the taxation of income from endogenously determined capital in a complete general equilibrium setting. Judd (1985) and Chamley (1986) extended this work and demonstrated that in standard optimal growth models, models where capital and consumption converge to a steady state, the optimal long-run policy sets the tax rate on income from capital to zero. Judd (1999) showed that for a wider class of dynamic models, particularly models that do not necessarily converge to a single steady state or balanced growth path, optimal policy still entails setting the tax rate on capital income to at least an average of zero over time.

In Chamley (1981), (1986), and Judd (1985), taxes are imposed to finance a fixed amount of government expenditure. By contrast, in Judd's (1999) more general formulation, government expenditure is a public good that enters the utility function of the representative agent. In none of this work is government expenditure directly related to economic output or its production. In Section 2, I adopt Judd's (1999) approach to determining optimal fiscal policy in models that may not necessarily possess a single steady state or balanced growth path, but I distinguish between public spending on transfer payments and government consumption. The latter is first, a general function of factor inputs, and then more specifically a function of the economic output the factors generate. Zero taxation of asset income does not emerge here as an optimal policy except as a special case. Instead, if we assume that government consumption and domestic output, net of depreciation, are related to each other in a particular way—one that can be easily estimated as a cointegrating vector—a simple formula for the optimal tax rate on asset income emerges; a formula independent of government spending on transfer payments.

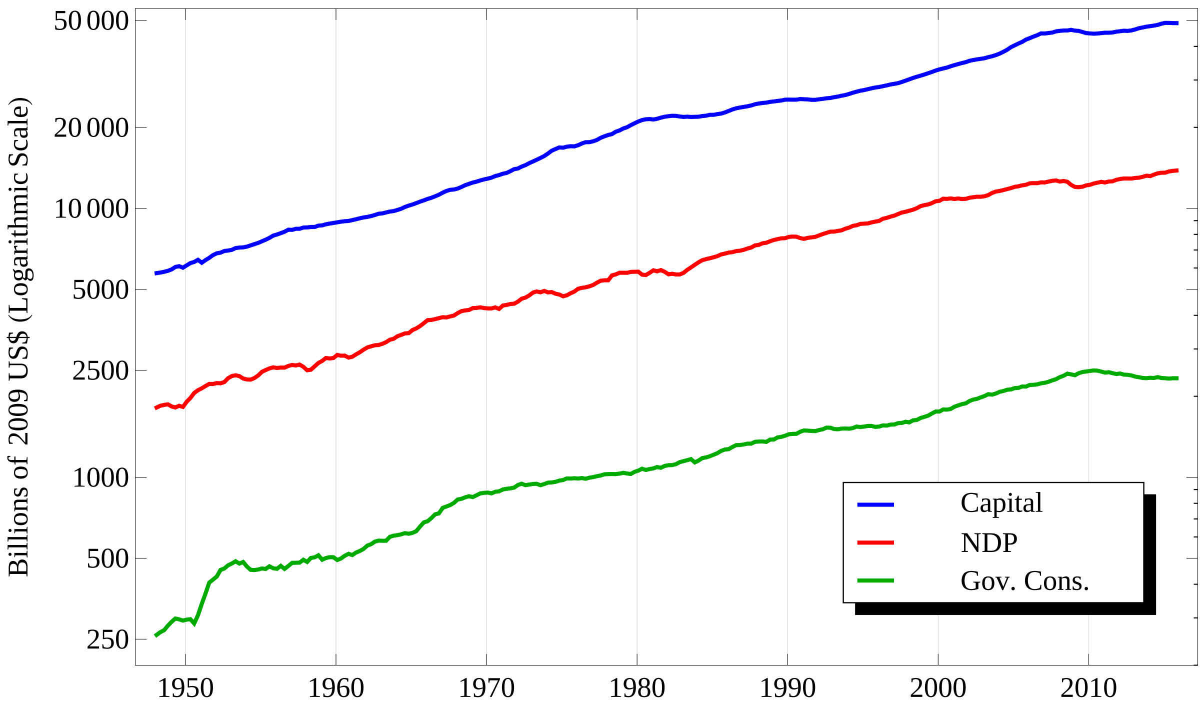

Figure 1: Capital, net domestic product and government consumption, quarterly in the United States, deflated using the NDP deflator, 1948Q1 to 2015Q4, natural logarithmic scale. *Data Source*: Federal Reserve Economic Data, Federal Reserve Bank of St. Louis; U.S. Department of Commerce: Bureau of Economic Analysis seasonally adjusted variables and Fernald "A Quarterly, Utilization-Adjusted Series on TFP." Federal Reserve Bank of San Francisco: Capital, Annual Net Assets [K1TTOTL1ES000] interpolated using quarterly changes in private capital from Fernald; Net Domestic Product [A362RC1Q027SBEA]; Government Consumption Expenditures [A955RC1Q027SBEA]; Deflator [A362RG3Q086SBEA]. http://research.stlouisfed.org/fred2/ and http://www.frbsf.org/economic-research/indicators-data/total-factor-productivity-tfp/.

Consider the behaviour of government consumption expenditure, capital and net domestic product in the United States from 1948 onward in Figure 1. Throughout this work I use nominal data deflated by the net domestic product deflator—the focus here is on the financing of government consumption expenditure, so real volume measures of government outputs would generate a distorted picture of how much net output is devoted to government consumption or what portion of the capital stock is devoted to its production. Whereas the rise in the amount spent on transfer payments has caused total government expenditure to grow at a faster pace than the economy as a whole, the portion of net domestic output devoted to direct government spending on goods and services closely tracks overall net domestic product—an impression reinforced if we consider in Figure 2, either the evolution of the share of government consumption in net domestic product, or the ratio of the trend components of each series.

As I demonstrate in the first part of Section 3, both are integrated series and there exists a cointegrating relationship between them that can be captured by estimating a vector error correction model.

In contrast to Chamley (1981), (1986), and Judd (1985), (1999), in the model in Section 2 more capital accumulation leads not only to higher output but to higher government consumption as well. Hence in the absence of a tax on the income it generates, the amount of capital the economy will accumulate will be innefficiently high. As it this underlying relationship between government consumption and capital that determines the optimal rate of taxation, in the second part of Section 3, I estimate the cointegrating relationship between government consumption and factor inputs. Forecasts generated by these very same vector error correction models can then be used to provide estimates of the long-run optimal tax rate on asset income.

Finally in Section 4, I incorporate the estimates from Section 3 into the calibration of an optimal growth model with elastic labor supply and measure the welfare implications of shifting the tax burden from income derived from assets to labor earnings. As has been demonstrated in previous studies by Coleman (2000), Domeij and Heathcote (2004), Eerola and Määttänen (2013), İmrohoroğlu (1998), Laitner (1995) and Lucas (1990), the prevailing rate of tax on asset income is sufficiently high in the United States that in the context of a representative agent framework, eliminating it completely and shifting the burden to labor income has the potential to generate a substantial positive welfare benefit. Qualitatively this effect is retained, but if government consumption flows are directly related to economic activity (production or capital accumulation) the magnitude of the benefit will be significantly smaller. Indeed, rather than eliminating the tax completely, a more modest shift, one that lowers the effective tax rate on net asset income from its recent long-run average of 0.449 to between 0.148 and 0.186 (depending on the model specification chosen), generates the greatest (though still relatively modest) improvement in welfare, equivalent to a permanent increase in consumption of between 1.491% and 1.858%. These numbers can be improved upon, though only to a very small extent, if the tax rate is permitted to shift slightly over time in the vicinity of this range.

As far back as Adolph Wagner (1883) and Henry Carter Adams (1898), Economists, have postulated a close relationship between the amount of government expenditure and the overall size of the economy. Indeed, a sizable empirical literature has developed to examine and explain this relationship, starting with the seminal work by Peacock and Wiseman (1961) for the United Kingdom. Yet rarely is this feature incorporated into models studying optimal fiscal policy. Taken together, the theoretical and empirical results of this work suggest that failure to consider the relationship between the share of net output devoted to government activity that taxes help finance and the overall size and productive capacity of the economy in general has the potential to skew our conclusions regarding the best allocation of the tax burden across the different input factors.

In 1990, Robert E. Lucas wrote:

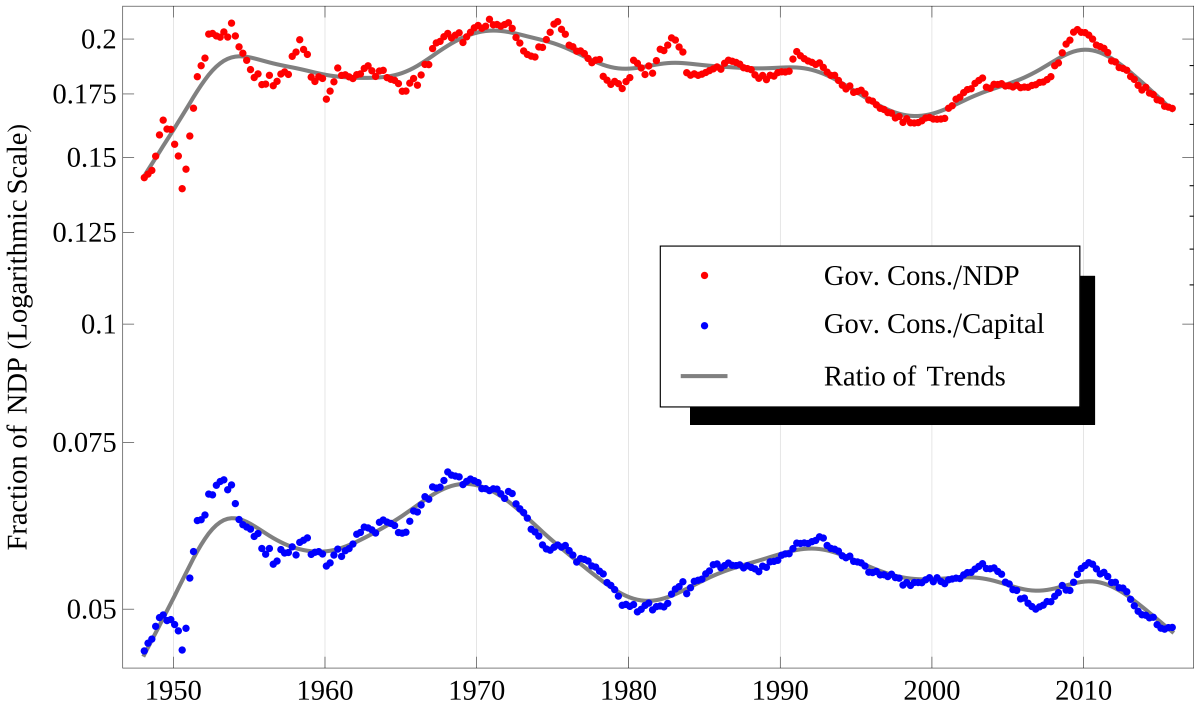

Figure 2: The ratios of government consumption expenditure to net domestic product or capital stock in the United States, combined with the ratios of each of their trend components from Hodrick Prescott Filters (the value of the penalty parameter is set to λ=1600), 1948Q1 to 2015Q4.

> When I left graduate school, in 1963, I believed that the single most desirable change in the U.S. tax structure would be the taxation of capital gains as ordinary income. I now believe that neither capital gains nor any of the income from capital should be taxed at all.

Yet the continued development of dynamic general equilibrium models that endogenise the supply of capital has not settled the argument regarding the efficacy of taxing the income it generates. Aiyagari (1995), Correia (1996), Reis (2011) and others, all find that under conditions of uninsurable idiosyncratic risk, asymmetric information, or the inability of governments to tax some factor inputs, Ramsey optimal policies will include some taxation of capital income. This work implies that even in the absence of uncertainty, incomplete markets, or asymmetric information, imposing some of the burden of funding government expenditure on capital income can be an economically efficient policy that maximises the welfare of a representative agent, provided there is a functional relationship between government consumption, the capital stock and the overall size of the economy. Indeed, rather than setting the tax rate on asset income to zero, a welfare optimising policy for the United States would imply the near equalising of net rates of taxation across different sources of income. The only difference between the tax rates imposed on asset income and earnings will stem from the burden of debt service, which should fall solely on the latter.

2 Theory: Ramsey Optimal Policy

2.1 The Representative Household's Problem

We begin by reformulating Judd's (1999) optimal taxation argument in discrete time and also alter his model to make government consumption a function of either factor inputs or the net output they together produce. Assume an economy in which all participants are members of households that share the instantaneous utility function $u : \mathbb{R}_+^2 \to \mathbb{R}$, which maps preferences over consumption and labor, and a discount factor $\beta \in (0, 1)$. Utility is strictly increasing in consumption, and strictly decreasing in labor. I normalize the size of the initial population to N_0=1, and a representative household chooses its consumption c_t and labor input l_t to maximise its infinite horizon discounted utility:

$$\max_{c,l} \sum_{t=0}^{\infty} \beta^t N_t u \left[c_t, l_t\right] \tag{P.1}$$

subject to

$$N_{t+1} a_{t+1} = N_t \left(\bar{w}_t l_t + (1 + \bar{r}_t) a_t - \bar{p}_t c_t + h_t\right) \tag{1}$$

where a_t represents assets (both bonds and capital); h_t represents net government transfer payments; $\bar{w}_t$, $\bar{r}_t$ and $\bar{p}_t$ represent the time t after tax wage rate, after-tax rate of return on asset holding and after-tax price of consumption; the size of the population is N_t, and the net rate of population growth between time t and $t + 1$ is $N_{t+1}/N_t - 1$.

Differentiating the optimization problem P.1 with respect to c_t, c_{t+1}, l_t and a_{t+1} yields the first order conditions:

$$u_c \left[c_t, l_t\right] - \lambda_t \bar{p}_t = 0, \tag{2}$$

$$u_l \left[c_t, l_t\right] + \lambda_t \bar{w}_t = 0, \tag{3}$$

$$-\lambda_t + \beta \lambda_{t+1} \left(1 + \bar{r}_{t+1}\right) = 0, \tag{4}$$

where $u_c \left[c_t, l_t\right] > 0$ and $u_l \left[c_t, l_t\right] < 0$ are the marginal utilities of consumption and labor, and λ_t is a current value costate variable that expresses the marginal utility derived by the representative household from a positive increment to asset wealth.[1]

2.2 The Social Planner's Problem

Output in this economy is produced by combining aggregate capital K_t and aggregate effective labor $z_t L_t$, which is the aggregate labor input itself $L_t = N_t l_t$ multiplied by labor augmenting technology z_t. I denote the production function as $F : \mathbb{R}_+^2 \to \mathbb{R}_+$. Capital depreciates at the

[1]For the special case where $\forall l$, $u_l \left[c, l\right] = 0$, the first order conditions reduce to (2) and (4) only.

constant rate $\delta \geq 0$, and so net domestic product, defined as output net of capital depreciation is $Y_t \equiv F[K_t, z_t L_t] - \delta K_t$. We assume competitive firms maximise profits, so that pre-tax factor returns r_t and w_t equal their marginal products. I also assume the production technology F is homogenous of degree 1, so that in equilibrium:

$$r_t = F_1[K_t, z_t L_t] - \delta, \tag{5}$$

$$w_t = z_t F_2[K_t, z_t L_t]. \tag{6}$$

The government raises revenue in each period t by selling one period bonds B_{t+1}, collecting a tax on labor earnings $\tau_t^l = 1 - \frac{\bar{w}_t}{w_t}$, collecting a tax $\tau_t^a = 1 - \frac{\bar{r}_t}{r_t}$ levied on income from the returns generated by either of the two assets, physical capital or the bonds themselves, and collecting an *ad valorem* tax $\tau_t^c = \frac{\bar{p}_t}{p_t} - 1$ on consumption c_t. In this real economy we normalise the pre-tax price of the consumption good p_t to one. Together all these revenues finance government consumption, which here is confined to that portion of government activity represented on the expenditure side of the national accounts which is not designated as investment, finance an exogenous stream of transfer payments, or redeem the interest and principal of all the outstanding debt incurred in the period prior. In contrast to most of the optimal tax literature, where government consumption is fixed, or Judd (1999), where it enters the utility function of both the representative agent and the social planner, here I assume that like output, it is a function of both aggregate inputs, $G : \mathbb{R}_+^2 \to \mathbb{R}_+$. The government's budget constraint is:

$$B_{t+1} = G[K_t, z_t L_t] + H_t - \tau_t^l w_t L_t - \tau_t^a r_t \left(K_t + B_t\right) + \left(1 + r_t\right) B_t - \tau_t^c C_t, \tag{7}$$

where B_t is the aggregate stock of government bonds at the beginning of time t, $C_t = N_t c_t$ represents aggregate consumption flows during this period, and $H_t = N_t h_t$ represents aggregate government transfer payments.

Now consider the Ramsey problem of a policy maker who chooses per-capita consumption c_t, leisure l_t, the after-tax price of consumer goods $\bar{p}_t$, and after-tax factor returns $\bar{r}_t$ and $\bar{w}_t$ which maximise the representative households' discounted utility:

$$\max_{c,l,\bar{r},\bar{w},\bar{p}} \sum_{t=0}^{\infty} \beta^t N_t u\left[c_t, l_t\right] \tag{P.2}$$

subject to the incentive compatibility constraints (2) to (4), the feasibility condition:

$$K_{t+1} = F\left[K_t, z_t L_t\right] - G\left[K_t, z_t L_t\right] + \left(1 - \delta\right) K_t - C_t, \tag{8}$$

and assuming the aggregate production function F is homogenous of degree one, the government's budget constraint (7), which can be reformulated as:

$$B_{t+1} = \bar{r}_t \left(K_t + B_t\right) + \bar{w}_t L_t + \delta K_t - F\left[K_t, z_t L_t\right] + G\left[K_t, z_t L_t\right] - \left(\bar{p}_t - 1\right) C_t + B_t + H_t. \tag{9}$$

Budget constraints (8) and (9), when combined imply (1) after it is aggregated. I assume $\lim_{t->\infty} |b_t| < \infty$ and also:

$$\bar{r}_t \geq 0, \tag{10}$$

$$\bar{w}_t \geq 0, \tag{11}$$

$$\bar{p}_t \geq 0. \tag{12}$$

Differentiating the social planner's optimization problem P.2 with respect to $\bar{p}_t$, $\bar{w}_t$, $\bar{r}_t$, λ_t, and the per-capita values c_t, l_t, k_{t+1}, and b_{t+1} yields the first order conditions:

$$u_c\left[c_t, l_t\right] - \phi_t^k + \phi_t^c u_{cc}\left[c_t, l_t\right] + \phi_t^l u_{cl}\left[c_t, l_t\right] - \mu_t\left(\bar{p}_t - 1\right) = 0, \tag{13}$$

$$\begin{aligned} u_l\left[c_t, l_t\right] + z_t\phi_t^k\left(F_2\left[K_t, z_tL_t\right] - G_2\left[K_t, z_tL_t\right]\right) + \mu_t\left(\bar{w}_t - z_tF_2\left[K_t, z_tL_t\right] + z_tG_2\left[K_t, z_tL_t\right]\right) \\ +\phi_t^c u_{cl}\left[c_t, l_t\right] + \phi_t^l u_{ll}\left[c_t, l_t\right] = 0, \end{aligned} \tag{14}$$

$$-\mu_t\frac{c_t}{p_t} - \phi_t^c\lambda_t + \nu_t^p = 0, \tag{15}$$

$$\mu_t l_t + \lambda_t\phi_t^l + \nu_t^w = 0, \tag{16}$$

$$\phi_{t-1}^{\lambda}N_{t-1}\lambda_t + \mu_t\left(K_t + B_t\right) + N_t\nu_t^r = 0, \tag{17}$$

$$N_{t-1}\phi_{t-1}^{\lambda}\left(1 + \bar{r}_t\right) - N_t\left(\beta\phi_t^{\lambda} + \phi_t^c\bar{p}_t - \phi_t^l\bar{w}_t\right) = 0, \tag{18}$$

$$\begin{aligned} -\phi_t^k + \beta\phi_{t+1}^k\left(F_1\left[K_{t+1}, z_{t+1}L_{t+1}\right] - G_1\left[K_{t+1}, z_{t+1}L_{t+1}\right] + 1 - \delta\right) \\ +\beta\mu_{t+1}\left(\bar{r}_t + \delta - F_1\left[K_{t+1}, z_{t+1}L_{t+1}\right] + G_1\left[K_{t+1}, z_{t+1}L_{t+1}\right]\right) = 0, \end{aligned} \tag{19}$$

$$-\mu_t + \beta\mu_{t+1}\left(1 + \bar{r}_{t+1}\right) = 0. \tag{20}$$

where ϕ_t^c, ϕ_t^l, ϕ_t^{λ}, μ_t, ϕ_t^k, ν_t^r, ν_t^w, and ν_t^p are the costates variables associated with (2), (3), (4), (8), (9), (10), (11) and (12). Straub and Werning (2015) demonstrate that under certain conditions, a social planner will prefer policies such as setting $\bar{r}_t = 0$, $\forall t$, that have the effect of driving both the stock of capital and consumption to zero in the long-run, rather than interior solutions. In what follows, I restrict my attention to interior solutions to (13) to (20) and assume that $\nu_t^r = \nu_t^p = \nu_t^w = 0$.

In the absence of any tax distortions, the marginal value at time t of an increment of capital for the representative household, $\lambda_t > 0$ is equal to its marginal value for the social planner, $\phi_t^k > 0$. Similarly, in a model in which taxes are not distortionary and do not generate excess burdens, Ricardian equivalence prevails, and the neutrality of public debt held in household asset portfolios implies that $\mu_t = 0$. Otherwise, as is the case here, servicing any increase in the public debt burden entails deadweight losses so that $\mu_t < 0$. Following Judd (1999), I define $\Lambda_t \equiv \frac{1}{\bar{p}_t}\frac{\phi_t^k - \mu_t}{\lambda_t}$, which is a measure of the social value of an increment to physical capital when the value of private assets (comprising both capital and public debt) is held constant, and the reciprocal of $\bar{p}_t$ corrects for the distorting effect of ad-valoreum taxes paid on private consumption. Again, in a model without distortionary taxation, $\bar{p}_t = 1$, $\phi_t^k = \lambda_t$ and $\mu_t = 0$ so therefore $\Lambda_t = 1$.

Judd (1999) assumes that government expenditure enters the utility function as a way to ensure that the value of $\Lambda_t > 0$. It is, however, possible to achieve the same result by placing a few restrictions on preferences and on the production and government consumption functions.

Lemma 1. *A sufficient condition that ensures that* $\Lambda_t > 0$ *for all* $c_t > 0$, $k_t > 0$ *and* $l_t > 0$, *is* $u_l[c_t, l_t] + lu_{ll}[c_t, l_t] + cu_{cl}[c_t, l_t] \leq 0$ *and* $z_t(F_2[K_t, z_tL_t] - G_2[K_t, z_tL_t]) > 0$.

Proof. Solving (14) for ϕ_t^k, subtracting μ_t, substituting the interior solutions for (15) and (16) for ϕ_t^c and ϕ_t^l (with $\nu_t^p = \nu_t^w = 0$), and replacing $\bar{w}_t$ using (3) yields:

$$\phi_t^k - \mu_t = \frac{-u_c[c_t, l_t]\, u_l[c_t, l_t] + \mu_t(u_l[c_t, l_t] + l_t u_{ll}[c_t, l_t] + c_t u_{cl}[c_t, l_t])}{u_c[c_t, l_t]\, z_t(F_2[K_t, z_tL_t] - G_2[K_t, z_tL_t])}. \tag{21}$$

From the assumptions that $u_c[c_t, l_t] > 0$ and $u_l[c_t, l_t] < 0$, $\phi_t^k - \mu_t > 0$ if $u_l[c_t, l_t] + lu_{ll}[c_t, l_t] + cu_{cl}[c_t, l_t] \leq 0$ and $z_t(F_2[K_t, z_tL_t] - G_2[K_t, z_tL_t]) > 0$. Finally from (2) $\lambda_t > 0$ and hence $\Lambda_t > 0$. □

After substituting (4), (19) and (20), the value of Λ_t evolves according to:

$$\frac{\Lambda_{t+1}}{\Lambda_t} = \frac{\bar{p}_t}{\bar{p}_{t+1}} \frac{1 + \bar{r}_{t+1}}{1 + F_1[K_{t+1}, z_{t+1}L_{t+1}] - G_1[K_{t+1}, z_{t+1}L_{t+1}] - \delta}. \tag{22}$$

The numerator in the right-hand side of (22), $1 + \bar{r}_{t+1}$ when multiplied by the price ratio $\frac{\bar{p}_t}{\bar{p}_{t+1}}$, is the cost agents in this economy face when they shift a unit of consumption between period t and $t+1$. The denominator reflects the cost of this shift in terms of production, which here includes the portion of extra output lost to additional government consumption. Iterating (22) from period t backwards:

$$\frac{\Lambda_t}{\Lambda_0} = \frac{\bar{p}_0}{\bar{p}_t} \prod_{i=1}^{t} \frac{1 + \bar{r}_i}{1 + F_1[K_i, z_iL_i] - G_1[K_i, z_iL_i] - \delta}. \tag{23}$$

Assume that along a balanced growth path the ad valoreum tax rate on consumption is constant so that $\bar{p}_t = \bar{p}_0$. Comparing the growth rates for the costate variables μ_t and λ_t in (4) and (20), we know the ratio μ_t/λ_t is always constant. If this economy converges to a steady state or a balanced growth path, once convergence is complete, the ratio ϕ_t^k/λ_t will have converged to a constant value as well. Hence for an economy that has converged, a solution to the social planner's problem implies the value Λ_t is a constant and (23) implies that the sequence $\{1 + \bar{r}_i\}_{i=1}^t$ must be set to ensure the ratio Λ_t/Λ_0 is equal to one.

Theorem 1. *Suppose an economy converges to a steady state or balanced growth path, then assuming an interior solution for (13) to (20), the long-run optimal policy is to set* $\bar{r}$ *equal to* $\lim_{t\to\infty} F_1[K_t, z_tL_t] - G_1[K_t, z_tL_t] - \delta$. *The social planner accomplishes this by setting the long-run tax rate* τ^a *equal to* $\lim_{t\to\infty} G_1[K_t, z_tl_t] / (F_1[K_t, z_tl_t] - \delta)$.

Proof. Follows from $\Lambda_t/\Lambda_0 = 1$ in (23). □

Indeed, notice how the same trajectory of Λ_t/Λ_0 can be generated by choosing either the sequence $\{1 + \bar{r}_i\}_{i=1}^t$ or the ratio $\frac{\bar{p}_0}{\bar{p}_t}$. The social planner has more policy instruments than necessary to achieve an optimal solution to P.2.

Corollary 1. *The availability of a consumption tax is not necessary to ensure a Ramsey second best allocation associated with the optimization problem P.2.*

Proof. From (4) and (20) we know that $\frac{\mu_{t+1}}{\lambda_{t+1}} = \frac{\mu_t}{\lambda_t}$ $\forall t$. Combining (17) with (18) and inserting the values of K_{t+1} and B_{t+1} from (8) and (9) yields:

$$\frac{\mu_t}{\lambda_t}\left(\bar{w}_t l_t - c_t\right) - \phi_t^c + \phi_t^l \bar{w}_t = 0.$$

Combining this with (16) while assuming an interior solution so that $\nu_t^w = 0$ yields:

$$\phi_t^c = -\frac{\mu_t}{\lambda_t} c_t,$$

which replicates (15) as long as $\nu_t^p = 0$. Since (15) is implied by the other first order conditions, for any interior solution the availability to the social planner of a consumption tax does not alter the result in Theorem 1. □

In what follows, we assume that $\bar{p}_t$ is always constant. A number of special cases emerge from Theorem 1, depending on how the function G is specified. For example, if $G_1\left[K_t, z_t L_t\right] = 0$ so that government consumption is not a function of the capital stock, we recover the canonical Chamley-Judd result of zero taxation on asset income as the long-run optimising policy. This is the case even if $G_2\left[K_t, z_t L_t\right] \neq 0$, and government consumption is still a function of the amount of effective labor employed in production. Alternatively if $G_1\left[K_t, z_t L_t\right] \neq 0$, then endogenous government expenditure creates a wedge between the net marginal product of capital $F_1\left[K, zL\right] - \delta$, and corresponding interest rate r, that confronts individuals in this economy, and the full marginal product of capital $F_1\left[K, zL\right] - G_1\left[K, zL\right] - \delta$, as it is perceived by the social planner. The tax rate imposed on assets τ^a serves to compensate for this disparity.

For example, if $G_1\left[K_t, z_t L_t\right] < 0$, then any policy that encourages capital accumulation depresses the amount of output diverted to government consumption, and the optimal policy is to subsidise capital income by setting $\bar{r}$ to be less than r and $\tau^a < 0$. If $G\left[K_t, z_t L_t\right] = gK_t^{-\beta}\left(z_t L_t\right)^{1+\beta}$, then even in a model with exponential steady state growth, government expenditure as a share of GDP still converges to a strictly positive amount, and yet the optimal tax is still negative. Finally, if $G_1\left[K_t, z_t L_t\right] > 0$ then the optimal tax rate is positive.

In an economy in which government expenditure is exogenously determined, the long-run supply curve for capital is infinitely elastic at a given interest rate. This is why the distortions associated with policies that lower the after-tax rate of return dominate those that directly affect labor supply. By contrast, for the type of economies specified in Theorem 1, a change in the tax rate on asset income alters not just the amount of capital available to produce the consumption good, but indirectly affects the overall amount of government consumption, which here does not have the usual lump-sum quality. Instead, government consumption is itself a type of distortion that asset taxation serves to mitigate. This remains the case even if the economic activity from which it is derived necessitates the government's consumption.

Consider the case of the power function $G[K_t, z_t L_t] = g(F[K_t, z_t L_t] - \delta K_t)^\gamma$, and $g > 0$ and $\gamma > 0$. Suppose the values of z_t and N_t converge to constants, and the economy converges to a stationary steady state. Government consumption converges to a positive share of net output and the optimal long-run tax rate on asset income is $\gamma g(F[K_t, z_t L_t] - \delta K_t)^{\gamma-1}$.[2] By contrast, if we assume z_t and/or N_t are growing, we must constrain the value of γ to be less than or equal to one, to ensure that government consumption does not ultimately exceed net output. If there exists a balanced growth path and $\gamma = 1$, government consumption converges to a positive share of output g, and the long-run optimal policy will be one where the tax rate is positive so that $\tau^a = g$. If, however, $\gamma < 1$, and the aggregate economy is growing, then $\lim_{t\to\infty} g(F[K_t, z_t L_t] - \delta K_t)^{\gamma-1} = 0$, which means we recover the Chamley-Judd optimal long-run policy of setting $\bar{r} = F_1[K, zL] - \delta$ and $\tau^a = 0$ in the limit—though as I will demonstrate below, until the economy converges, the optimal policy may be very different.

Theorem 1 applies to economies that have converged to a balanced growth path. What then should be the policy if the economy does not converge to a balanced growth path but is characterized by cycles, or, alternatively, convergence is achieved only over a very long time horizon. Can we say anything about optimal policy in the interim?

First, the results in Theorem 1 and the after tax rate of return as $t \to \infty$ do not depend on the distorting properties of the other taxes, on the evolution of lump-sum transfers or whether they are positive or negative. Indeed it remains valid even if in the initial period the social planner is able to set $\bar{r}_0 = 0$ and confiscate all asset income in the initial period. In that case an optimising social planner will deploy the additional revenue from what amounts to an initial lump sum tax towards reducing the distortionary impact of wage taxation. To see this, assume the utility function is not a function of labor hours supplied in the market, but depends on consumption alone, and that H_t is a policy variable. Labor taxation and transfers are now completely interchangable and for these two policy instruments Ricardian equivalence prevail—the shadow price of government debt μ_t is equal to zero. Yet the reasoning behind Theorem 1 remains intact. The social planner will set the value of $\bar{r} = \lim_{t\to\infty} F_1[K_t, z_t L_t] - G_1[K_t, z_t L_t] - \delta$ to ensure that $\phi_t^k = \lambda_t$. More often, when analysing optimal factor taxation, we exclude the option of resorting to lump-sum taxation as an alternative source of revenue. Suppose, in what follows, we constrain H_t to be equal to zero.

The challenge here is that regardless of what the long-run optimum policy is, during the initial period, the social planner might want to set the value $\bar{r}_0$ very low to exploit the time-zero inelasticity of capital supply and replicate the now missing option of imposing a lump sum tax. It is this reasoning that gives rise to the "bang-bang" pattern of optimal taxation described by Chamley (1986). This is why, even though the more a sequence of tax rates on asset income causes the value of Λ_t to deviate from one, the more it distorts the economy and generates welfare losses, it is not possible to pin down the initial value of Λ_0 or assume it equals one. Yet

[2] Ben-Gad (2003) analyses long-run optimal fiscal policy along a balanced growth path in the context of a two-sector endogenous growth model for the case of $g > 0$ and $\gamma = 1$.

if we assume that an optimal programme will seek to minimise distortions beyond an initial period of high taxation, subsequent values of Λ_t must be bounded below and above over time: $\Lambda_\infty < \Lambda < \Lambda^\infty$. Setting the bounds and rewriting the inequalities in logarithms yields:

$$\ln\left(\frac{\Lambda_0}{\Lambda^\infty}\right) \leq \sum_{i=1}^{t} \ln\left(\frac{1+F_1\left[K_i, z_i L_i\right]-G_1\left[K_i, z_i L_i\right]-\delta}{1+\bar{r}_i}\right) \leq \ln\left(\frac{\Lambda_0}{\Lambda_\infty}\right), \tag{24}$$

which implies that as the value of Λ_t in (23) evolves over a sufficiently long period of time, it must on average be equal to one and the average distortion measured as deviations from $\Lambda_t = 1$ approaches zero in the limit. Extending Judd (1999), this implies that for all $t_1 \geq 0$, any long-run constant value of $\bar{r}$ must satisfy:

$$\lim_{t_2\to\infty} \frac{1}{t_2} \sum_{i=t_1}^{t_1+t_2} \ln\left(\frac{1+F_1\left[K_i, z_i L_i\right]-G_1\left[K_i, z_i L_i\right]-\delta}{1+\bar{r}}\right) = 0, \tag{25}$$

which in turn implies that if the social planner must choose a particular constant tax rate, then:

Theorem 2. *Assume there exists an interior solution for (13) to (20), for any $t_1 \geq 0$, if the value of $\bar{r}$ is fixed, then the long-run optimal policy is to set it to satisfy (25).*

Theorem 2 generalises Theorem 1 to economies that may not converge to balanced growth paths or steady states. For example, if the dynamic behaviour of the economy is characterized by permanent cycles, and $\bar{r}$ is to be fixed to any value, it will be optimally so, if on average it equals $F_1\left[K_t, z_t L_t\right] - G_1\left[K_t, z_t L_t\right] - \delta$ and the long-run tax rate τ^a on asset income is set to the average value of $G_1\left[K_t, z_t L_t\right] / \left(F_1\left[K_t, z_t L_t\right] - \delta\right)$. Yet this leaves the question of what is the best policy if the policy maker is not necessarily constrained to choose fixed values for $\bar{r}_t$ and τ_t^a?

To avoid the issue of time inconsistency, assume a policy maker commits to an infinite sequence of $\bar{r}_t$ that need not be constant. An infinite number of different sequences satisfy the boundary conditions in (24) and hence also satisfy

$$\lim_{t_2\to\infty} \frac{1}{t_2} \sum_{i=t_1}^{t_1+t_2} \ln\left(\frac{1+F_1\left[K_i, z_i L_i\right]-G_1\left[K_i, z_i L_i\right]-\delta}{1+\bar{r}_i}\right) = 0. \tag{26}$$

Yet only by committing to a policy of setting $\bar{r}_i$ equal to $F_1\left[K_i, z_i L_i\right] - G_1\left[K_i, z_i L_i\right] - \delta$ and tax rates τ_i^a equal to $G_1\left[K_i, z_i L_i\right] / \left(F_1\left[K_i, z_i L_i\right] - \delta\right)$ in each period does a policy maker both satisfy (25) and minimise deviations from $\Lambda_t = 1$.

Theorem 3. *For a sufficiently large value of $t_1 \geq 0$, an optimal tax policy for asset income is such that the values of $\bar{r}_i$ are set equal to $F_1\left[K_i, z_i L_i\right] - G_1\left[K_i, z_i L_i\right] - \delta$.*

Proof. By satisfying (26) such a policy simultaneously minimizes deviations from $\Lambda_t = 1$ over time and satisfies the boundary conditions in (24). It generalizes Theorem 6 in Judd (1999) as well as Theorem 2 above to the case where the policy maker is not constrained to set tax rates to one constant value. □

Theorem 3 can be implemented by setting the sequence of tax rates τ_i^a equal to $G_1[K_i, z_iL_i] / (F_1[K_i, z_iL_i] - \delta)$. There is however an obvious limitation to the practical applicability of the Theorem—the difficulty in determining the appropriate size of the initial t_1 periods during which the social planner may choose to set the tax rate on asset income very high to exploit the short-term inelasticity in the supply of capital.[3] Yet regardless of the length of t_1, we can utilize the intuition that underlies Theorem 3 to generate a useful conjecture about how different tax policies are likely to compare. Once again we focus on the power function.

Conjecture 1. *Suppose government consumption is a power function of net domestic product,* $G[K_t, z_tL_t] = g(F[K_t, z_tL_t] - \delta K_t)^\gamma$. *Then a policy of setting the sequence of tax rates* τ_i^a *equal to* $\gamma G[K_t, z_tL_t] / (F[K_t, z_tL_t] - \delta K_t)$, *which equals* $\gamma g(F[K_t, z_tL_t] - \delta K_t)^{\gamma-1}$ *for all periods* $t \geq 0$, *weakly dominates a policy of fixing* τ^a *to any fixed value. Furthermore, if* $\gamma = 1$, *then the policy of fixing* $\tau^a = g$ *strictly dominates the policy of fixing* τ^a *to any value* $\tau^a \neq g$.

Setting aside the possibility of employing a "bang-bang" optimal control policy through period t_1, Conjecture 1 predicts how the welfare effects of different tax policies, if implemented immediately, are likely to compare, and not merely over the long-run, if the relationship between government consumption and net output takes a very particular form, the power function.

More generally, one possible interpretation of the function $G[K_t, z_tL_t]$, one that generalizes beyond the context of the strictly theoretical models considered in this section, is that it expresses a long-run equilibrating relationship between government consumption and either factor inputs or the economic output, net of depreciation, they generate. Provided government consumption and net output are integrated $I(1)$ processes, perhaps because they share a trend driven by labor augmenting technology and/or population growth as in the model above, the specific case of the power function is easily estimated, as it corresponds in its logarithmic form to the cointegrating relationship in Johansen's Vector Error Correction Model (VECM). If capital and effective labor are I(1) as well, it is also possible to replace net output with factor inputs and establish the cointegrating relationship between capital and government expenditure that is fundamental to the theory. In the next section I use the VECM to estimate both types of relationship, and then in Section 4 I incorporate these estimates into the calibration of a model designed to numerically evaluate the main implications of Conjecture 1.

3 Estimation: Error Correction Models for Government Consumption

3.1 Government Consumption and Net Output

We start by examining the properties of government consumption G_t, and net domestic product in the United States Y_t, using quarterly data from the first quarter of 1948 to the fourth quarter

[3]Chamley (1986) provides one method for approximating t_1.

of 2014. In Table 1, neither the augmented Dickey-Fuller, the Dickey–Fuller test with GLS detrending, the Elliott-Rothenberg-Stock point-optimal test, Phillips-Perron test or Ng and Perron's MZ_α, MZ_t, MS_B and MP_t tests cannot reject the null hypothesis of a unit root at the 1% critical level when applied to the levels of each series, but all reject the existence of unit roots at the 1% level when applied to the series' first differences. Perron and Vogelsang's or Zivot and Andrew's Dickey-Fuller tests which each accounts for breaks in both the trend or constant (in levels) or the constant (in differences) generate the same pattern, and fail to find evidence of spurious unit root behavior. Similarly, the KPSS test rejects the null hypothesis of stationarity in the levels of each of the two series at the 1% level, but cannot reject the null hypothesis when applied to differences, as long as the test includes a deterministic time trend (which in each case is statistically significant).

For the ratio of the two series (log differences) in the last two columns, the augmented Dickey-Fuller and Phillips-Perron test each rejects the existence of a unit root at the 1% critical level. Accounting for break points weakens but does not contradict these results and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) cannot reject the null hypothesis of stationarity of the ratio at the 10% critical value (the deterministic time trend in KPSS_t is not statistically significant).[4] However the the Dickey–Fuller test with GLS detrending, the Elliott-Rothenberg-Stock point-optimal test, Ng and Perron's MZ_α, MZ_t, and MP_t cannot reject the null hypothesis of a unit root even at the 10% critical level. Together, these contradictory results could imply that if the two series are cointegrated and characterized by the exponential relationship in Section 2 the value of γ might fall somewhere in the vicinity of one.

To test for cointegration, I begin by estimating an unrestricted VAR for the two time series. The optimal lag length p for the estimated VAR indicated by the Aikake's information criterion, Akaike's final prediction error (FPE) and the likelihood ratio (LR) is $q = 4$, but Schwarz's Bayesian information criterion (SBIC) and the Hannan and Quinn information criterion (HQIC) indicate, as is often the case, a more parsimonious optimal lag length, $q = 2$, which is what I choose to use. If government consumption and net domestic product do indeed share a common stochastic trend, the largest eigenvalue of the system must be equal to one, and to guarantee stability all the others must be (in modulus) less than one. The largest eigenvalue here is .997, the next highest is .911, and the two complex eigenvalues $0.341 + .030i$ the fall well within the unit circle.

To determine the cointegrating vector itself, I estimate the Vector Error Correction Model

[4] A Quandt-Andrews breakpoint test performed on an AR(1) estimation of the log ratio cannot reject the null hypothesis of no breakpoints at a confidence level of 5% when trimming 5, 10 or 15% of the data, though a Bai Perron breakpoint test with 5% trimming suggests one or two possible breaks in the early 1950's coinciding with large changes in defense spending associated with the beginning and end of the Korean War.

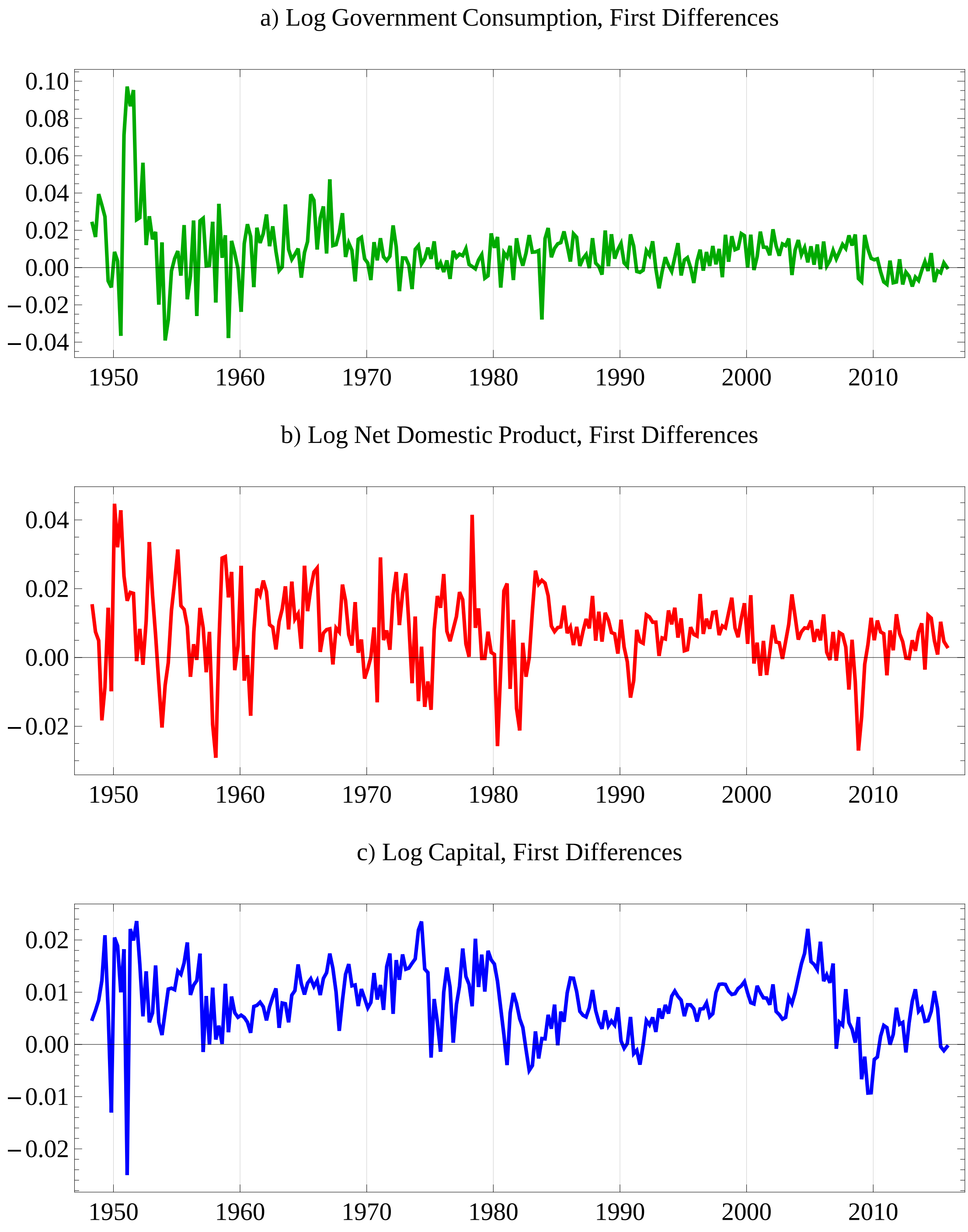

Figure 3: Log differences of government consumption, net domestic product and log capital, quarterly in the United States, seasonally adjusted annual rate deflated by the NDP deflator, 1948Q1 to 2015Q4, natural logarithmic scale. *Data Source*: See caption in Figure 1.

Method	Gov. Consumption		Net Domestic Product		Gov. Cons./NDP
	Level	First Diff.	Level	First Diff.	Level
ADF	-2.678	-5.895***	-1.678	-11.028***	-4.328***
DF-GLS	-0.736	-3.540***	-0.970	-6.543***	-1.613
P_T-GLS	46.307	0.948***	26.203	0.319***	17.658
PP	-2.916	-12.085***	-1.509	-11.066***	-3.849**
MZ_α	-2.916	-22.231***	-3.232	-66.434***	-6.096
MZ_t	-0.828	-3.296***	-1.008	-5.749***	-1.658
MS_B	0.377	0.148***	0.312	0.087***	0.272
MP_t	31.315	1.236***	23.092	0.402***	14.900
DF-PV	-4.352	-7.099***	-3.595	-11.392***	-4.811
DF-ZA	-4.369	-8.207***	-3.743	-8.576***	-5.508**
KPSS_{nt}	2.137†††	0.564††	2.181†††	0.387†	0.214
KPSS_t	0.356†††	0.071	0.391†††	0.027	0.189††

Table 1: Nominal data deflated by NDP deflator, in natural logarithms. ADF is the augmented Dickey–Fuller test. DF-GLS is the Dickey–Fuller test with GLS detrending. P_T-GLS is the Elliott-Rothenberg-Stock point-optimal test statistic. PP is the Phillips-Perron test. MZ_a, MZ_t, MS_B, and MP_t are the modified tests in Ng and Perron (2001). DF-PV and DF-ZA are the Perron and Vogelsang and Zivot and Andrews unit root tests with intercept and trend breaks. All these tests are conducted with a constant term and trend for levels and ratios, and a constant term only for first differences. Finally, KPSS_{nt} is the Kwiatkowski-Phillips-Schmidt-Shin test without deterministic time trend while KPSS_t includes the trend. * Reject the null hypothesis of a unit root at the 10% confidence level. ** Reject the null hypothesis of a unit root at the 5% confidence level. *** Reject the null hypothesis of a unit root at the 1% confidence level. † Reject the null hypothesis of stationarity at the 10% confidence level. †† Reject the null hypothesis of stationarity at the 5% confidence level, ††† Reject the null hypothesis of stationarity at the 1% confidence level.

(VECM):

$$\begin{aligned}\begin{pmatrix} \Delta \ln G_t \\ \Delta \ln Y_t \end{pmatrix} &= \begin{pmatrix} \theta_G \\ \theta_Y \end{pmatrix} (\ln G_{t-1} - \gamma \ln Y_{t-1} - \ln g) \\ &+ \sum_{i=1}^{q} \begin{pmatrix} \zeta_{GG}(i) & \zeta_{GY}(i) \\ \zeta_{YG}(i) & \zeta_{YY}(i) \end{pmatrix} \begin{pmatrix} \Delta \ln G_{t-i} \\ \Delta \ln Y_{t-i} \end{pmatrix} + \begin{pmatrix} \eta_G \\ \eta_Y \end{pmatrix} + \begin{pmatrix} \varepsilon_{G,t} \\ \varepsilon_{Y,t} \end{pmatrix},\end{aligned} \tag{27}$$

where θ_G and θ_Y represent the spead of adjustment, the vector $(1, \gamma)'$ represents the normalised cointegrating vector and g is a constant. If $\gamma = 1$, then the long-run relationship between government consumption and net domestic product is a fixed proportion, represented by the constant term g. Together γ and g correspond to the parameters in the power function that expresses the long-run relationship between government consumption and net output in Section 2.

Define the 2×2 matrix $\Phi \equiv \begin{pmatrix} \theta_G \\ \theta_Y \end{pmatrix} \begin{pmatrix} 1 & \gamma \end{pmatrix}$. From Johansen's trace and maximum eigenvalue test in Table 2 we can reject the null of no cointegrating vector at nearly any reasonable

Rank	Eigen.	Trace	p-Value	Max-Eigen.	p-Value
0	0.092	29.402	<0.001	25.986	0.001
1	0.013	3.416	0.065	3.416	0.065

Table 2: Johansen's trace and maximum eigenvalue tests along with information criteria for the rank of the matrix $\alpha\beta$ for government consumption and net domestic product.

confidence level but do not reject the hypothesis of one cointegrating vector (rather than full rank). Together this implies that $(1,\gamma)'$ represents a valid cointegrating vector that, together with the estimated value of g, represents the long-run relationship between government consumption and net domestic product.[5]

Given the contradictory results the different tests for the stationarity of the ratio between government consumption and net output yield, it is not surprising that the estimate of $\gamma = 0.927$ in the unrestricted vector error correction model (27) in column (1) of Table 3 is close to, yet still more than two and a half standard deviations from one. Hence we can reject the restriction of γ to one in column (2) with a p-value of 0.008. The estimate of θ_G is statistically significant but the estimate of θ_Y is not, implying that net output is weakly exogenous.That is consistent with the theoretical model in Section 2 which implies that government consumption is directly determined by the values of the parameters g and γ and the level of economic activity. Hence we cannot reject the restriction that the value of θ_Y is zero in column (3), in which case the value of γ rises slightly to 0.935. However we can reject, at least a the 5% level, the imposition of both constraints in column (4).

So what do the results in Table 3 tell us about optimal tax policy? First, if we do impose the constraint $\gamma = 1$, the optimal long-run tax on asset income is simply equal to the value of g or 0.184 which is also the long-run average share of government consumption in net output between 1948 and 2015. By contrast, if $\gamma < 1$, as our estimates imply, and the size of the economy continues to grow in the future, either because of per-capita output growth or the increasing size of the work force, the share devoted to government consumption declines as its growth fails to completely keep pace with that of net output. As mentioned in Section 2, such scale effects mean the limiting optimal tax rate as $t \to \infty$ coincides with the Chamley-Judd rate of zero.

Yet immediately setting the rate of tax to zero would not minimise distortions as described in Theorem 3. Even if we believe the estimated power relationship is stable over long periods of time, decades or even centuries may pass before the value of $\gamma g \left(F\left[K_t, z_t L_t\right] - \delta K_t\right)^{\gamma-1}$ decays by any significant amount.

[5]Wickens (1996) demonstrates that the particular factorisation of Φ chosen by VECM is not necessarily economically meaningful and the estimated cointegrating vector is an unknown linear transform of the underlying long-run structural relationship. However this critique is not pertinent to the estimation of a single structural equation such as the power function relationship estimated here.

	(1)		(2)		(3)	(4)
	$\Delta \ln G_t$	$\Delta \ln Y_t$	$\Delta \ln G_t$	$\Delta \ln Y_t$	$\Delta \ln G_t$	$\Delta \ln G_t$
Cointegrating eq.						
γ	0.927^{***} (0.025)		1		0.935^{***} (0.025)	1
g	0.248		0.184		0.240	0.184
Error correction						
θ_G, θ_Y	-0.059^{***} (0.012)	-0.007 (0.008)	-0.055^{***} (0.013)	-0.003 (0.009)	-0.060^{***} (0.012)	-0.056^{***} (0.013)
$\Delta \ln G_{t-1}$	0.227^{***} (0.058)	-0.066 (0.043)	0.251^{***} (0.059)	-0.061 (0.042)	0.228^{***} (0.058)	0.251^{***} (0.059)
$\Delta \ln G_{t-2}$	0.173^{***} (0.058)	0.035 (0.042)	0.201^{***} (0.058)	0.038 (0.042)	0.176^{***} (0.057)	0.201^{***} (0.058)
$\Delta \ln Y_{t-1}$	0.067 (0.085)	0.337^{***} (0.062)	0.096 (0.085)	0.348^{***} (0.062)	0.068 (0.085)	0.096 (0.085)
$\Delta \ln Y_{t-2}$	0.001 (0.086)	0.101 (0.063)	0.018 (0.087)	0.115^{*} (0.063)	-0.001 (0.086)	0.018 (0.087)
η_G, η_Y	0.004^{***} (0.001)	0.004^{***} (0.001)	0.003^{***} (0.001)	0.004^{***} (0.001)	0.004^{***} (0.001)	0.003^{***} (0.001)
R^2	0.250	0.161	0.230	0.161	0.250	0.230
$\bar{R}^2$	0.236	0.146	0.215	0.144	0.236	0.215
F-statistic	17.542	10.128	15.683	9.984	17.548	15.683
Constraints	-		$\gamma = 1$		$\theta_Y = 0$	$\gamma = 1$, $\theta_Y = 0$
Likelihood Value	1624.087		1620.537		1623.777	1620.491
LR Statistic	-		7.101		0.621	7.19
Deg. of Freed.	-		1		1	2
p-Value	-		0.008		0.431	0.027

Table 3: Estimated Vector Error Correction Model for US data on government consumption and net domestic product, 1948Q1 to 2015Q4. Standard errors in parentheses.

Method	Capital		Effective Labor		Gov. Cons./Capital
	Level	First Diff.	Level	First Diff.	Level
ADF	-1.233	-4.659***	-1.674	-13.092***	-3.493**
DF-GLS	-1.022	-4.638***	-1.273	-13.109***	-1.213
P_T-GLS	22.932	0.856***	21.457	0.236***	29.040
PP	-0.911	-11.361***	-1.550	-13.149***	-3.662**
MZ_α	-4.879	-33.700***	-4.176	-128.632***	-3.687
MZ_t	-1.351	-4.058***	-1.254	-7.993***	-1.245
MS_B	0.277	0.120***	0.300	0.062***	0.338
MP_t	17.609	0.870***	19.999	0.236***	23.084
DF-PV	-2.581	-5.541***	-3.325	-13.536***	-4.956*
DF-ZA	-3.077	-5.443***	-3.760	-7.386***	-4.965*
$KPSS_{nt}$	2.173†††	0.384†	2.151†††	0.248	0.923†††
$KPSS_t$	0.357†††	0.072	0.260†††	0.046	0.152††

Table 4: Nominal data deflated by NDP deflator, in natural logarithms. For description of the tests, see caption in Table 1

3.2 Government Consumption and Factor Inputs

It is the underlying connection between government consumption and capital that generates the policy prescriptions in Theorems 1 through 3. Is there in fact a cointegrating relationship between government consumption and the two aggregate factor inputs, and in particular to capital? First, to generate data for the total capital stock K_t at a quarterly frequency in Figures 1 and 2, I using the quarterly rates of change in private capital estimated by Fernald (2014) to interpolate quarterly estimates of total fixed assets between the available end of year figures. Then using nominal gross domestic product, and assuming a constant returns to scale Cobb-Douglas production function, I employ simple growth accounting to generate a series that corresponds to effective labor, $z_t L_t$. When applied to each factor input, the unit roots tests in Table 4 reveal a pattern similar to those in Table 1—alone each series is $I(1)$ but tests on the ratio of government consumption to capital yield results that are as contradictory as those for the ratio of government consumption to net output.

When applied to estimates of unrestricted vector autoregressions with the logarithm of government expenditure $\ln G_t$, the logarithm of aggregate capital $\ln K_t$, and the logarithm of effective labor $\ln(z_t L_t)$, SBIC once again suggests the optimal number of lags is equal to 2. The eigenvalues of the system are stable, the largest equal to 0.998, and the next highest 0.971. Both Johansen's trace and maximum eigenvalue tests for cointegration in Table 5 support the existence of one and only one cointegrating vector.

Rank	Eigen.	Trace	p-Value	Max-Eigen.	p-Value
0	0.078	32.846	0.022	22.029	0.037
1	0.032	10.817	0.223	8.710	0.311
2	0.008	2.106	0.147	2.106	0.147

Table 5: Johansen's trace and maximum eigenvalue tests along with information criteria for the rank of the matrix $\alpha\beta$ for government consumption, aggregate capital and effective labor.

The counterpart here to (27) is:

$$\begin{pmatrix} \Delta \ln G_t \\ \Delta \ln K_t \\ \Delta \ln (z_t L_t) \end{pmatrix} = \begin{pmatrix} \theta_G \\ \theta_K \\ \theta_{zL} \end{pmatrix} (\ln G_{t-1} - \gamma\alpha \ln K_{t-1} - \gamma (1-\alpha) \ln (z_{t-1} L_{t-1}) - \gamma \ln \Phi - \ln g) \tag{28}$$
$$+ \sum_{i=1}^{q} \begin{pmatrix} \zeta_{GG}(i) & \zeta_{GK}(i) & \zeta_{GzL}(i) \\ \zeta_{KG}(i) & \zeta_{KK}(i) & \zeta_{KzL}(i) \\ \zeta_{zLG}(i) & \zeta_{zLK}(i) & \zeta_{zLzL}(i) \end{pmatrix} \begin{pmatrix} \Delta \ln G_{t-i} \\ \Delta \ln Y_{t-i} \\ \Delta \ln (z_{t-i} L_{t-i}) \end{pmatrix} + \begin{pmatrix} \eta_G \\ \eta_Y \\ \eta_{zL} \end{pmatrix} + \begin{pmatrix} \varepsilon_{G,t} \\ \varepsilon_{Y,t} \\ \varepsilon_{zL,t} \end{pmatrix},$$

where Φ relates to the rate of depreciation, or more precisely represents the ratio between net domestic product Y_t and gross domestic product $\tilde{Y}_t = Y_t + \delta K_t$. I treat this as a constant and set its value to its long-run average between 1948 to 2015, Φ=0.889.[6] Given that the series for effective labor was constructed using a constant returns to scale production technology, the value of the estimated coefficients on $\ln K_{t-1}$ and $\ln (z_{t-1} L_{t-1})$ in (28) must by construction equal γ, which allows me to isolate the value of g from the estimate of the constant in the cointegrating vector.

The unrestricted estimates of the model in the columns labeled (1) in Table 6 imply values of $\alpha = 0.441$, $g = 0.252$ and $\gamma = 0.887$—the latter two estimates are not significantly different from the corresponding estimates in Table 3.[7] The constraint that the sum of the values of the two coefficients in (28) are equal to one can be rejected at the 1% level. However, once again the estimates this time of θ_K and θ_{zL} suggest that both aggregate capital and effective labor are weakly exogenous, and this is confirmed in the column labeled (2) where the p-value associated with the constraint is 0.607. In this specification the estimates imply: $\alpha = 0.476$, $g = 0.230$ and $\gamma = 0.897$. The column labeled (3) includes estimates where the constraints $\gamma = 1$ as well as $\theta_K = 0$ and $\theta_{zL} = 0$. Here the p-value associated with the constraint is only 0.011, so we can once again reject this additional restriction. Again the estimates in both Tables 3 and 6 confirm what the tests for stationarity of the ratios in Tables 1 and 4 suggest—that the value of γ is close to, yet statistically speaking, significantly less than one.

[6]Estimating Φ as the constant term in the cointegrating vector for Y_{t-1} and $\tilde{Y}_{t-1}$ yields nearly identical results.

[7]The restriction that γ=0.927 cannot be rejected at a probability of 0.232.

		(1)		(2)	(3)
	$\Delta \ln G_t$	$\Delta \ln K_t$	$\Delta \ln (z_t L_t)$	$\Delta \ln G_t$	$\Delta \ln G_t$
Cointegrating eq.					
$\alpha\gamma$		0.391^{***}		0.427^{***}	0.231
		(0.147)		(0.143)	(0.192)
$\gamma(1-\alpha)$		0.496^{***}		0.470^{***}	0.769^{***}
		(0.159)		(0.154)	(0.192)
g		0.252		0.230	0.217
Error correction					
$\theta_G, \theta_K, \theta_{zL}$	-0.051^{***}	-0.003	-0.012	-0.052^{***}	-0.038^{***}
	(0.011)	(0.004)	(0.014)	(0.013)	(0.012)
$\Delta \ln G_{t-1}$	0.253^{***}	-0.020	0.201^{***}	0.256^{***}	0.283^{***}
	(0.060)	(0.023)	(0.079)	(0.060)	(0.061)
$\Delta \ln G_{t-2}$	0.176^{***}	0.084^{***}	0.069	0.180^{***}	0.199^{***}
	(0.060)	(0.023)	(0.079)	(0.060)	(0.062)
$\Delta \ln K_{t-1}$	-0.238	0.345^{***}	-0.132	-0.231	-0.210
	(0.154)	(0.060)	(0.203)	(0.154)	(0.158)
$\Delta \ln K_{t-2}$	0.126	0.167^{***}	-0.173	0.133	0.172
	(0.154)	(0.059)	(0.202)	(0.153)	(0.157)
$\Delta \ln (z_{t-1} L_{t-1})$	0.030	0.045	0.139^{**}	0.031	0.031
	(0.047)	(0.018)	(0.062)	(0.047)	(0.048)
$\Delta \ln (z_{t-2} L_{t-2})$	-0.047	-0.026	0.067	-0.046	-0.048
	(0.048)	(0.018)	(0.063)	(0.048)	(0.049)
$\eta_G, \eta_K, \eta_{zL}$	0.006^{***}	0.003^{***}	0.006^{***}	0.005^{***}	0.005^{***}
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
R^2	0.234	0.272	0.098	0.233	0.199
$\bar{R}^2$	0.213	0.253	0.074	0.213	0.178
F-statistic	11.369	13.955	4.073	11.377	10.509
Constraints		-		$\theta_Y = 0$	$\gamma = 1$, $\theta_Y = 0$
Likelihood Value		2489.455		2488.957	2483.863
LR Statistic		-		0.998	11.185
Deg. of Freed.		-		2	3
p-Value		-		0.607	0.011

Table 6: Estimated Vector Error Correction Model for US data on government consumption, aggregate capital stock and effective labor, 1948Q1 to 2015Q4. Standard errors in parentheses.

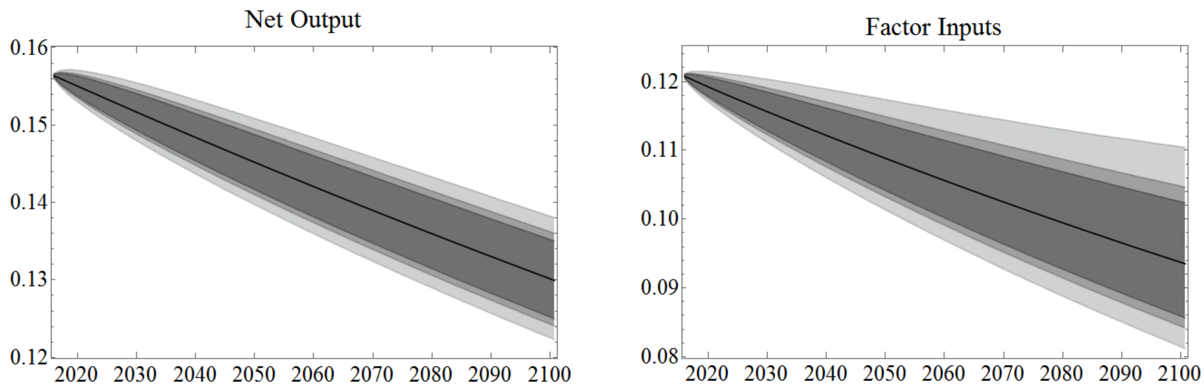

Figure 4: Forecasts for the value of $\gamma G_t/Y_t$ from 2016Q1 to 2100Q4 derived from the vector error correction model with net output in (27) and estimated in column (3) of Table 3, and the model with factor outputs in (28) estimated in column (2) of Table 6. Grey shading indicates confidence bands of 90%, 95% and 99%.

3.3 Forecasting the Optimal Time Path for the Tax Rate

How then should a government that wishes to set policy well in advance, choose to set tax rates? One option is to set them to either the long run average ratio of government consumption to net output, or the estimates of g in columns (2) and (4) in Table 3 or column (3) in Table 6. That means setting the tax rate to a constant value of between 0.184 and 0.217, depending on which estimate is chosen. Alternatively, as the likelihood ratio tests consistently imply we can reject the constraint that sets $\gamma = 1$ with a high degree of confidence, it might be more appropriate for the government to choose a path of taxes that matches the future values of $\gamma g \left(F\left[K_t, z_t L_t\right] - \delta K_t\right)^{\gamma-1}$ that the vector error correction models forecast.[8]

It is the behavior of factor inputs, and the net output they generate, that determines government expenditure in the theoretical model in Section 2. Hence the versions of the vector error correction models that most closely match that theoretical model are those where net output in Table 3, or inputs in Table 6, are weakly exogenous. As in fact, we cannot reject the hypothesis that constrains θ_Y, θ_K, or θ_{zL} to equal zero but can reject constraining γ to equal 1, I will use the estimates in columns (3) in Table 3 and (2) in Table 6 to generate long-run forecasts of $\gamma g \left(F\left[K_t, z_t L_t\right] - \delta K_t\right)^{\gamma-1}$ in Figure 4 that begin in the first quarter of 2016 and continue to the last quarter of 2100. These coincide with the optimal tax rates on asset income implied by both Theorem 3 and Conjecture 1 in Section 2.

When the estimates from Table 3 are used, the forecasts of $\gamma g \left(F\left[K_t, z_t L_t\right] - \delta K_t\right)^{\gamma-1}$ corresponding to column (3) imply an optimal policy that entails slowly lowering the tax rate on

[8]Note that in the theoretical model this term exactly equals the ratio of government expenditure to net output, but forecasts of $\gamma G\left[K_t, z_t L_t\right] / \left(F\left[K_t, z_t L_t\right] - \delta K_t\right)$ can diverge somewhat from $\gamma g \left(F\left[K_t, z_t L_t\right] - \delta K_t\right)^{\gamma-1}$. The former contains the independent behaviour of government consumption not explained by the cointegrating relationship. That part of its behavior is not relevant for determining the optimal tax rate on asset income.

asset income from 0.156 at the beginning of 2016, to 0.130 the end of 2100. If we rely on the estimates in column (2) of Table 6, the optimal tax rate is somewhat lower, dropping from an initial value of 0.121 and reaching 0.094 at the end of the century. In each case, the values associated with the fixed tax rates 0.184 and 0.217 mentioned above are not only significantly higher, but fall well outside even the 99% confidence intervals.

Having established in this section a strong empirical case that government consumption evolves as a function of net output, or indeed the aggregate factor inputs that generate it, and having already demonstrated in the previous section how such an assumption alters the nature of optimizing fiscal policy, I consider in the next section the quantitative welfare implications of shifting the burden of tax between income generated from asset holdings and labor earnings, using these estimates. Is Conjecture 1 valid for these sets of parameter choices? Given that it is far easier to implement a fixed rate of tax on asset income, I can evaluate how much of the maximum potential welfare gain is sacrificed if that is the policy adopted. Furthermore, I can not only measure the welfare benefit of implementing these different policies, but also evaluate to what degree, in terms of welfare, the small differences between the different estimated versions in (27) and (28) matter when used to inform tax policy.

4 Welfare Analysis

The purpose of this section is to numerically assess the theories and conjecture in Section 2, so as to quantify the magnitude of the welfare effects they imply for the US economy while incorporating the estimates from Section 3 regarding the relationship between government consumption and net output, and to juxtapose these results with those predicted by the canonical Chamley-Judd formulation, where government consumption is assumed to be growing at a fixed exogenous rate. To proceed, I assume a functional form for the utility function

$$u\left[c_t, l_t\right] = \ln c_t - \frac{l_t^{1+\frac{1}{v}}}{1+\frac{1}{v}} \tag{29}$$

where v corresponds to the Frisch elasticity of labor supply. I also assume that the aggregate production function takes the Cobb-Douglas form:

$$F\left[K_t, z_t L_t\right] = K_t^{\alpha}\left(z_t L_t\right)^{1-\alpha}. \tag{30}$$

To compare the welfare implications of shifts in fiscal policy, I calculate compensating differentials, measured in terms of a permanent increase in consumption. More formally, define $\{\breve{\tau}_t^a\}_{t=1}^T$ as the sequence of tax rates on asset income associated with the new fiscal policy we wish to evaluate. The value of T may be finite or infinite, depending on whether the policy is assumed to be temporary or permanent. Such a policy generates flows of consumption and labor $\{c_t, l_t\}_{t=0}^{\infty}$, which can then be compared to $\left\{\breve{c}_t, \breve{l}_t\right\}_{t=0}^{\infty}$, the agents' counterfactual flows

Parameters and Initial Tax Rates	
α	0.320
τ^c	0.050
τ^a	0.449
Discount Rate and Depreciation (Quarterly)	
β	0.9953
δ	0.0098
Exogenous Growth Rates (Quarterly)	
N_t/N_{t-1}	0.003
z_t/z_{t-1}	0.005
Initial Asset Holdings (as Ratio of Annual Gross Output)	
$B_0/F\left[K_0, z_0L_0\right]$	0.921
$K_0/F\left[K_0, z_0L_0\right]$	2.908

Table 7: The calibrated parameters of the model. The effective tax rate on asset income τ^a is the rate prior to policy changes and is calibrated using the procedure in Trabandt and Uhlig (2011).

of consumption and labor given the initial tax policy. The compensating differential is

$$\pi\left(\breve{\tau}_t^a|_{t=0}^T\right) = e^{-\sum_{t=0}^{\infty}\beta^t N_t} e^{\sum_{t=0}^{\infty}\beta^t N_t\left[\ln\frac{c_t}{\breve{c}_t}+\frac{v}{1+v}\left(\breve{l}_t^{1+\frac{1}{v}}-l_t^{1+\frac{1}{v}}\right)\right]} - 1. \tag{31}$$

The welfare implications of different policies can now be evaluated by feeding their associated impulse responses into (31), all the while assuming that the tax on labor earnings adjusts endogenously to keep the debt burden fixed over time.

4.1 Calibration

Microeconometric estimates of the Frisch elasticity labor supply vary from nearly zero to 0.5 for men and slightly higher for women. Surveying the recent literature, Reichling and Whalen (2012) conclude that a value of 0.4 provides the best central estimate, and this is the one employed by the Congressional Budget Office. By contrast, most dynamic stochastic general equilibrium models assume elasticities of one or two. Here I consider the two most extreme values analysed by Keane and Rogerson (2012): v=0.1 and v=2. As will be demonstrated below, the welfare effects generated by shifting the tax burden between capital and labor are remarkably robust to these two very different assumptions about the magnitude of this parameter, obviating the need to consider the implications of intermediate values.

I use the estimates in Tables 3 and 6 of the cointegrating vectors in (27) and (28) to set the values of g and γ. Given the potential scale effects in this economy when $0 < \gamma < 1$, it seems most appropriate to use long-run averages for the underlying growth rates for population and labor augmenting technology. Hence for both N_t/N_{t-1} and z_t/z_{t-1}, the growth rates match the average growth rates from 1948 to 2015 (quarterly), where I assume that for an economy close to the balanced growth path the latter can be approximated by the per-capita growth rate of gross domestic product. Similarly, the values of the factor share of capital in output α, depreciation δ and the initial capital-output ratio $K_0/F\left[K_0, z_0 L_0\right]$ are taken from the mean value of the data from this period used in Section 3. Tax rates are calculated using the available data from 1970 to 2014, following the procedure employed by Trabandt and Uhlig (2011) and using both the OECD and the European Union's AMECO databases. I choose the value of the subjective quarterly discount factor β that is consistent with the evolution of the Euler equation for consumption, given the long-run rates of growth, factor shares, the tax rate on assets and the capital output ratio.

Throughout, I constrain all shifts in fiscal policy to be fully financed—changes in the tax rate on asset income are fully compensated by offsetting changes to the tax rate on labor earnings that adjusts to maintain a constant ratio of public debt to output—both on and off the balanced growth path—all the while the *ad valorem* tax rate on consumption remains constant.[9] Instead of averaging over the preceding years, I choose for this ratio the sum of publicly held US federal, state and local debt relative to gross output at the end of 2015Q4.[10]

4.2 Implications of the constrained estimates: $\gamma = 0$ versus $\gamma = 1$

Before considering the implications of the theoretical and empirical results from Sections 2 and 3, it is worthwhile to first examine how welfare would be affected by changes to US tax policy if, as was assumed by Chamley (1981), (1986), and Judd (1985), (1999), government consumption is exogenously determined. In our model, this is analogous to assuming that the value of γ in Sections 2 and 3 is equal to zero. Here, the immediate impact of any shift in the tax burden away from income derived from assets and towards labor earnings is to immediately raise the after tax rate of return on capital. This induces both an immediate drop in consumption and an increase in the amount of labor agents supply. The combined effect is to initiate a long period of capital accumulation during which de-trended per-capita consumption first recovers and ultimately exceeds its initial value, while the amount of labor gradually declines in response to both the higher rate at which it is being taxed, and an income effect associated with a diminution of the excess burden from capital income tax.

Inserting the impulse responses associated with different lower rates of taxation of asset income into (31), the welfare effects, expressed as compensating differentials in terms of permanent increases to consumption, increase as the burden of tax is shifted from assets and is maximised (or nearly so) for the two curves labeled $\gamma = 0$, $\nu = 0.1$ and $\gamma = 0$, $\nu = 2$ in Figure 5 and Table 8, in a manner consistent with the Chamley-Judd results, when the tax on asset

[9] All impulse responses are calculated using a shooting algorithm. Programs available from the author by request.

[10] Federal Debt Held by the Public [FYGFDPUN] and Liability of State and Local Governments, Excluding Employee Retirement Funds; Credit Market Instruments [SLGSDODNS]. *Data Source*: http://research.stlouisfed.org/fred2/.

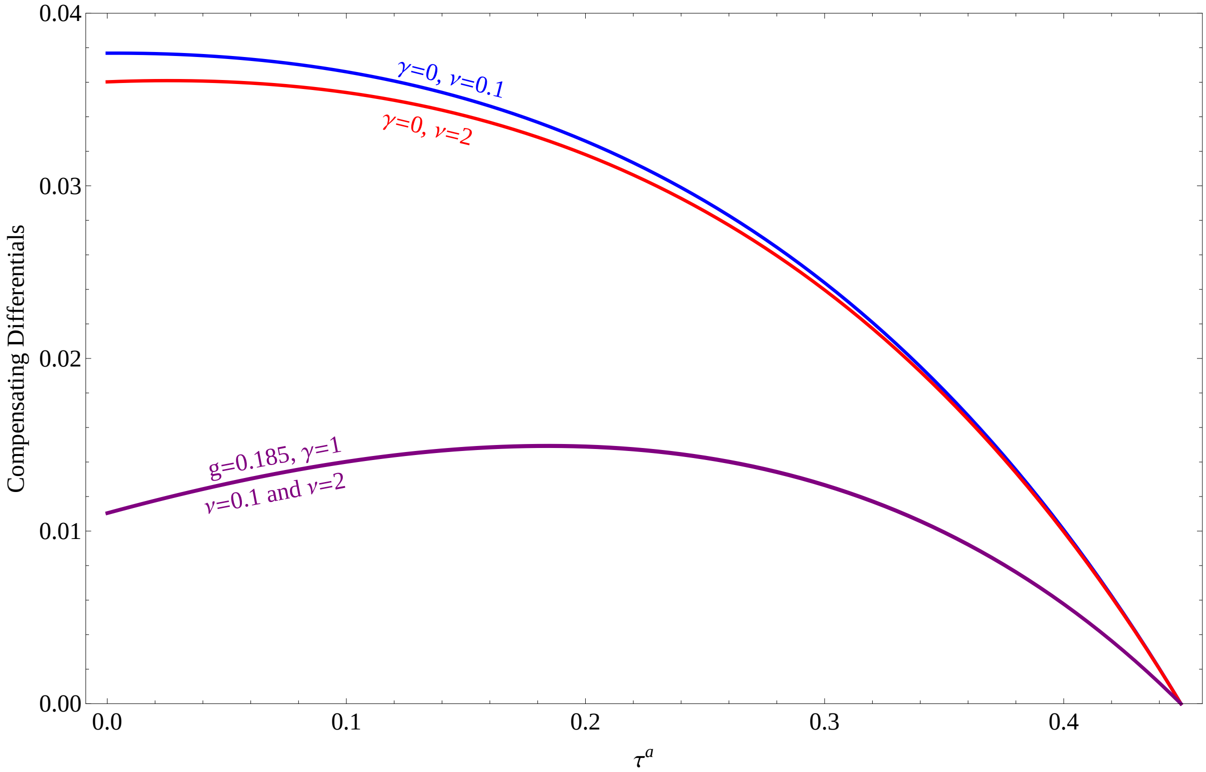

Figure 5: Welfare effects of lowering the tax rate on asset income from 0.449 in terms of permanent increases in consumption for γ=0 and γ=1.

income is completely eliminated ($\bar{\tau}^a = 0$).[11] Reflecting the logic of Harberger triangles, nearly half the maximum welfare benefit is achieved when the tax rate τ^a is lowered from its initial value of 0.449 to 0.35. Beyond that point marginal improvements in welfare decrease rapidly reaching the equivalent of a 3.768% permanent increase in consumption if $\nu = 0.1$, and 3.602% if $\nu = 2$ (denoted by $\pi(0)$ in Table 8). The (rather small) difference between the two welfare measures reflects the degree to which the higher taxes on labor are themselves distortionary owing to the elasticity of the labor supply. Overall, the impact of eliminating the tax on asset income is the equivalent in consumption terms of about six or seven quarters of per-capita output growth.

Consider how this measure contrasts with the results when we use the values of the estimation from columns (2) and (4) in Table 3, where γ is constrained to equal one, and the estimated

[11] In fact the maximum is attained at the small positive rates of taxation of 0.004 and 0.026 when $\nu = 0.1$ and $\nu = 2$, respectively (the first two rows of Table 8). The reason is that the tax on asset income is immediately reduced at the same time as it is announced. This has two contradictory effects on welfare: it reduces distortions because of the elasticity of capital supply in the long run, but, because capital supply is inelastic in the short run, it also bestows a costly lump-sum subsidy that the government must finance in the future. Only in the vicinity of the point where the tax is eliminated, and the welfare benefits of reducing the long-run distortions are nearly exhausted, does the this latter effect dominate, particularly when the elasticity of labour supply is relatively high. Even then it is too small to discern in Figure 5. Introducing a lag of two to three years between the time the policy is announced and its implementation, eliminates this slight non-monotonicity.

value of g corresponds to the long-run average share of government consumption in net domestic product, 0.184. First, decreases in the tax rate on asset income generate patterns of welfare changes, using the two different values of ν, that are so close together they are represented by one curve in Figure 5. The elasticity of labor supply is nearly irrelevant here. Second, welfare is maximised not at the zero tax rate, where the value of $\pi(0)$ in Table 8 indicates a welfare increase equivalent to a 1.103% to 1.105% increment to permanent consumption. Instead the maximum welfare benefit is attained at the value of $\bar{\tau}^a = 0.184$, validating the prediction made in Conjecture 1 that the optimal fixed rate of tax τ^a is equal to g. Indeed, although Theorem 1 refers to the long run, the results generated by an immediate and permanent change to the tax rate conform with their predictions, and those of Theorem 2 as well. Third, the maximum welfare gain, denoted by $\pi(\bar{\tau}^a)$ in Table 8, is equivalent to only a 1.491% to 1.495% permanent increase in consumption, less than half of what pertains when $\gamma = 0$ and the tax is eliminated. The implication is that, though from the perspective of a welfare maximising representative agent, the existing tax rate on asset income in the United States is presently set too high, the maximum benefit of reducing it is achieved when it is cut by slightly more than half, from 0.449 to 0.184, rather than by eliminating it completely. Furthermore, the maximal benefit to welfare that can be attained from any shift in the burden of taxation between asset income and labor earnings is much smaller than is commonly asserted in the literature, because there, government consumption is typically assumed not to respond to changes in the overall size of the economy.

4.3 Implications of the unconstrained estimates: $0 < \gamma < 1$, with constant rates of taxation

Suppose the value of γ is set between zero and one, in accordance with the estimates in Table 3 where the value of γ remains unconstrained. Unlike the cases where $\gamma = 0$ or $\gamma = 1$, here the share of government expenditure in net output, and hence also the optimal tax rate, decline over time, as seen in the first two rows of Figure 4. This also means that when considering the welfare effects implied by a change in tax policy, not only are the new paths of consumption and labor $\left\{\breve{c}_t, \breve{l}_t\right\}_{t=0}^{\infty}$ dynamic, but the paths associated with the initial policy $\{\bar{c}_t, \bar{l}_t\}_{t=0}^{\infty}$, are dynamic as well. Rather than considering the evolution of the economy as it adjusts from one balanced growth path to another, the point of comparision here is an economy that is, before the change in policy is initiated, already a very great distance from convergence to a balanced growth path.

So if at the moment when the policy changes, the economy has not converged to a balanced growth path, what starting point best matches the analysis above? More specifically, given the parameters, growth rates and initial tax rates in Table 7, along with particular estimated values of γ and g, is it possible to set the initial capital output ratio, the debt burden and the share of government expenditure independently? The answer is no.

As explained in Sections 2 and 3, if the aggregate economy continues to grow because of

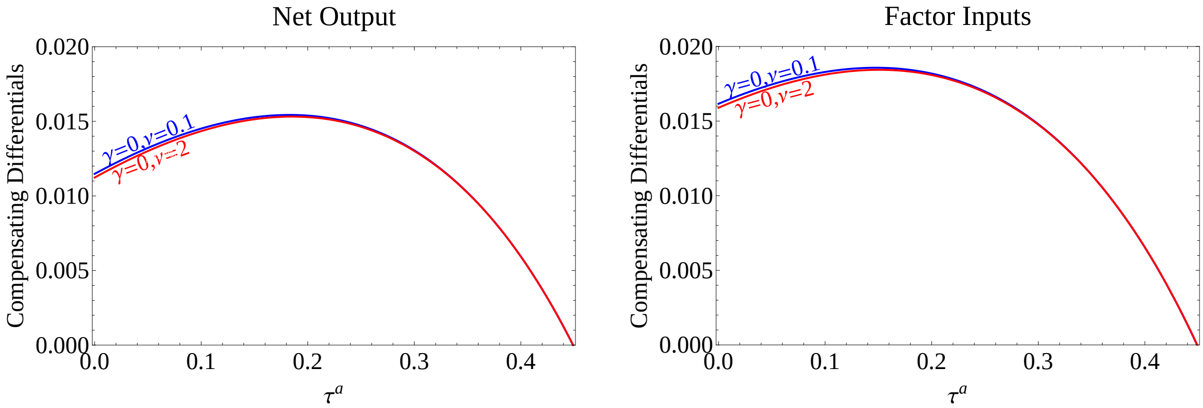

Figure 6: Welfare effects of lowering the tax rate on asset income from 0.449 in terms of permanent increases in consumption. The values of g and γ correspond to the estimates using net output in column (3) of Table 3, and using factor inputs in column (2) of Table 6.

population increase $N_t/N_{t-1} > 1$, technological improvement $z_t/z_{t-1} > 1$, or both, and the value of γ falls strictly between zero and one, the share of government consumption in net output shrinks and any change in policy affects an economy that has not yet reached its balanced growth path. Instead capital accumulates at a rate higher than what would be associated with natural population growth or the evolution of labour augmenting technology. Hence, to ensure impulse responses are monotonic, I assume that on impact the initial amount of per-capita physical capital amounts to 95% of its long-run value. I maintain consistency between the different versions of the model by keeping the same parameter values, hold fixed the ratio of public debt to annual output at 0.921 throughout, but allow the declining share of government expenditure to initially be a bit higher than 0.184. Furthermore, to avoid the unrealistic outcome that the share of government consumption in output ultimately approaches zero in the limit, I assume that government consumption evolves according to the power function relationship $G[K_t, z_t L_t] = g(F[K_t, z_t L_t] - \delta K_t)^\gamma$ from the initial period $t = 0$, when the new policy is announced and implemented, to the end of period $T-1$, when it reaches $g(F[K_{T-1}, z_{T-1}L_{T-1}] - \delta K_{T-1})^\gamma$. From period T onward, subsequent government consumption evolves by simply growing at the same exogenous rate as technology z_t and population N_t. I set the values of T=340 which means the period when government consumption is endogenous corresponds to the period between 2016Q1 till the end of the century, 2100Q4.

I proceed by first considering policies identical to those analysed for the case of $\gamma = 1$, where the tax rate on asset income changes to a new fixed value and then never changes again, but using the values of g and γ that correspond to the estimates from column (3) in Table 3 and column (2) in Table 6, where net output is weakly exogenous. Given that the estimated values of γ are relatively close in value to one, it is not surprising that the general pattern of compensating differentials in Figure 6 appears so similar to the curve that corresponds to $\gamma = 1$ in Figure 5. Once again the complete abolition of the tax on asset income is welfare improving, because the initial tax rate of 0.449 is so very high. Overall the benefits are not particularly large, and again hardly influenced by the value we choose for the elasticity of labor supply.

When I set $\gamma = 0.895$, $g = 0.240$, corresponding to estimation of (28) in column (2) of Table 6, the value of $\pi(0)$ in Table 8 is the equivalent of only a 1.616% permanent increase in consumption if $v = 0.1$, and drops to only 1.589% if the elasticity of labor supply is raised to $v = 2$. If the social planner is constrained to immediately change the tax rate to one fixed value, the highest possible welfare improvement, the equivalent of a 1.858% [1.844%] increase in consumption is achieved here if $v = 0.1$ [$v = 2$] and the tax rate on asset income is lowered from its initial value of 0.449 to $\bar{\tau}^a$ =0.148 [$\bar{\tau}^a$ =0.151]. When the model is calibrated using the estimation of (27) in column (3) of Table 3 the values of $\pi(0)$ drop to 1.148% [1.123%] if $v = 0.1$ [$v = 2$]. If $v = 0.1$ [$v = 2$], the planner attains the maximum welfare improvement equivalent to a permanent increase in consumption equivalent to 1.543% [1.532%] by fixing the tax rate on asset income at $\bar{\tau}^a$ =0.183 [$\bar{\tau}^a$ =0.186].

Overall, given that the estimations in Tables 3 and 6 reject constraints that fix the value of γ=1, but not the assumption of weak exogeneity we are left to conclude that the optimal fixed rate of taxation on asset income is somewhere between 0.148 and 0.186. A policy of lowering tax rates on asset income to within this range will produce the maximum welfare benefit equivalent to between a 1.532% and 1.858% increase in consumption, and no more. Given that the shapes of the parabolas in Figure 6 flatten out near their optima, any tax rate chosen within this region will capture nearly all the potential welfare improvement that is possible.

4.4 Sequences with changing rates of tax

Suppose policy makers are no longer restricted to shifting between one fixed rate of tax on asset income and another as assumed in Theorem 2, but instead can choose tax rates that change over time, in accordance with Theorem 3 and Conjecture 1. Let $\{\vec{\tau}_t^a\}_{t=1}^{\infty}$ denote the sequence of tax rates that obtain when the model is simulated with the tax rate on asset income set to equal $\gamma g \left(F\left[K_t, z_t L_t\right] - \delta K_t\right)^{\gamma-1}$, and let $\pi\left(\vec{\tau}_t^a\right)$ denote the welfare effects in terms of compensating differentials, generated by adoption of this policy. Conjecture 1 indicates that the improvements to welfare associated with setting the tax rate equal to $\gamma g \left(F\left[K_t, z_t L_t\right] - \delta K_t\right)^{\gamma-1}$ should dominate the policy associated with the fixed tax rates $\bar{\tau}^a$. If $\gamma = 0$ or $\gamma = 1$, this distinction is not meaningful, and the optimal policy is still to set the tax rate in each period to $\tau_t^a = 0$ or $\tau_t^a = g$ respectively, so the focus here is on those instances where $0 < \gamma < 1$. Again, given that net output is growing, the value of $\gamma g \left(F\left[K_t, z_t L_t\right] - \delta K_t\right)^{\gamma-1}$ declines and so any sequence of optimal tax rates is declining over time as well. As before, to prevent the share of government consumption from reducing to zero in the limit, I assume that from period $T = 340$ onwards, its share of net output stabilises.

In every case in Table 8, the values of $\pi\left(\vec{\tau}_t^a\right)$ are greater than $\pi\left(\bar{\tau}^a\right)$, confirming the predictions of Conjecture 1. At the same time, the differences are not large—allowing the tax

γ	g	v	$100 \times \pi(0)$	$\bar{\tau}^a$	$100 \times \pi(\bar{\tau}^a)$	$100 \times \pi(\vec{\tau}_t^a)$	$100 \times \pi(\tilde{\tau}_t^a)$
0	-	0.1	3.768	0.004	3.769	-	-
		2.0	3.602	0.026	3.609		
0.894	0.230	0.1	1.616	0.148	1.858	1.865	1.822
		2.0	1.589	0.151	1.844	1.850	1.806
0.935	0.240	0.1	1.148	0.183	1.543	1.550	1.508
		2.0	1.123	0.186	1.532	1.536	1.493
1	0.185	0.1	1.103	0.184	1.491	-	-
		2.0	1.105	0.184	1.495		

Table 8: Optimising policies and corresponding welfare measures for different values of γ and g.

rate to vary yields only small increments to welfare beyond those already achieved by lowering them to the relevant fixed value $\bar{\tau}^a$. If indeed the share of government expenditure continues to gently decline before stabilising at the end of the century, the best this policy can achieve is a welfare improvement of 1.865% in terms of permanent consumption for the case where $g = 0.230$, $\gamma = 0.895$ and $v = 0.1$, compared to 1.858% if the tax rate on asset income is fixed at 0.148. Once more, though the prevailing burden on asset income is considerably higher than is optimal, the potential benefits of reducing it are still far more modest than is often assumed to be the case.

Finally, for completeness' sake, consider the welfare implications of introducing, not the sequence of tax rates on asset income denoted by $\{\tilde{\tau}_t^a\}_{t=1}^{\infty}$ that match the values of $\gamma g \left(F\left[K_t, z_t L_t\right] - \delta K_t\right)^{\gamma-1}$, as calculated within the model simulation, but rather the sequence of forecasts generated by the estimation of the models (27) and (28) in Figure 4. The compensating differentials, designated as $\pi(\tilde{\tau}_t^a)$ in Table 8 are uniformly lower than either $\pi(\tilde{\tau}_t^a)$ or $\pi(\bar{\tau}^a)$ for each parameterisation of the model, but the differences are not substantial. Either way, it seems that a policy of fixing the tax rate on asset income to the appropriate level $\bar{\tau}^a$ in Table 8 is one that both is easy to implement and secures much of any potential welfare benefit that can be attained by reforming fiscal policy.

5 Conclusion

Peacock and Wiseman's study *The Growth of Public Expenditure in the United Kingdom*, first published in 1961, was partly motivated by the desire to test a proposition first stated by Adolph Wagner in 1883:

> The "law of increasing expansion of public, and particularly state, activities" becomes for the fiscal economy the law of increasing expansion of fiscal requirements.

> Both the State's requirements grow and, often more so, those of local authorities.....That law is the result of empirical observation in progressive countries, at least in our Western European civilization; its explanation, justification and cause is the pressure for social progress and the resulting changes in the relative spheres of private and public economy, especially compulsory public economy. Financial stringency may hamper the expansion of state activities, causing their extent to be conditioned by revenue rather than the other way round, as is more usual. But in the long run the desire for development of a progressive people will always overcome these financial difficulties.[12]

Their work appeared to confirm Wagner's prediction that government expenditure would not only grow with the size of the economy, it would take an increasing share of output—they found that government consumption of goods and services rose from 6.6% to 22.7% of gross national product in the years between 1890 and 1955 (spending on transfers and subsidies rose from 2.3% to 13.9%). Similarly, in the US current government expenditure between 1948 and 2015 rose from 24.5% to 39.4% of net domestic product. But this is where the comparison ends. When we confine ourselves to the resources consumed by the government, consumption grows with the economy, but as a share of net output, it is either stable, or as implied by my estimates, declining very slowly. This is more consistent with the type of relationship described by an American contemporary of Wagner, Henry Carter Adams, writing in 1898:

> On the contrary, it seems reasonable to assume that with each increment in the social product the people will conceive it to be to their advantage to invest added sums in the machinery of government. From the point of view of investment, therefore, as well as from a consideration of the satisfaction to be secured from the activities of the State, may we conclude that the fiscal demands of government will increase along with, if not in proportion to, the general social income.[13]

Given the high effective rates of taxation on net asset income that currently prevail in the United States, any reduction, including all the way to zero, will yield some welfare benefit. Nonetheless, once the link between government consumption and net output is recognised, the optimality of setting the tax rate to zero and shifting the burden to labor earnings disappears. Instead a more modest shift, one that would see the rate of tax on asset income drop by slightly more than half, will yield the highest welfare improvement. Moreover, the potential welfare gain is less than half what we would expect if we ignore the linkage between government consumption and output. This result is highly robust to both the alternative specifications of

[12]Adolph Wagner, *Finanzwissenschaft*, Third edition, Leipzig 1883. Translated to English by Nancy Cooke in Richard A. Musgrave and Alan T. Peacock, ed. *Classics in the Theory of Public Finance*, New York: St Martin's Press,1967, pg. 8.

[13]Henry Carter Adams, *The Science of Finance: An Investigation of Public Expenditures and Public Revenues, New York: Henry Holt and Company*, 1898, pg. 39.

the cointegrating relationship between government consumption and either net output or the underlying factor inputs, as well as the widest plausible range of possible values for the elasticity of labor supply.

The policy implication of this paper, that an efficient fiscal policy is one that sets the tax on asset income to a positive rate only slightly below the rate of tax on labor earnings, is in qualitative terms nearly identical to that in Reis (2011). In her paper the taxing authority cannot distinguish between entrepreneurial labor income and returns to capital. Similarly, Correia (1996) demonstrates that if some productive factors cannot be taxed, some of the burden of financing government expenditure should fall on capital income. Banks and Diamond (2010) cite these considerations as underpinning their rejection of zero taxation of asset income in their recommendations published in the Mirlees Report (2010) which studied possible reforms to the United Kingdom's tax system. A parallel strand of the literature first developed by Aiyagari (1995) argues in favour of taxing the income derived from capital as a means of suppressing its overaccumulation because uninsurable idiosyncratic risk leads to precautionary saving. However, there is no reason to assume that a model that incorporates both mechanisms, the endogeneity of government consumption and the difficulty of distinguishing between or imposing a tax on some productive inputs, will generate yet higher optimal taxes on asset income that compound these two effects.

Similarly much of the literature on optimal taxation derived in a Mirleesian, rather than a Ramsey framework, generally implies positive asset income taxation as well. However the underlying mechanism that justifies taxes in a Mirleesian model of optimal taxation, the asymmetry in the availability of information available to agents and the tax authority as in Golosov *et al.* (2003), is unlikely to interact with the reasoning based on the endogeneity of government consumption here, in any manner that would indicate that two arguments should be taken together to justify rates of tax higher than those implied by each argument in isolation.[14]

References

[1] Adams, Henry Carter, *The Science of Finance: An Investigation of Public Expenditures and Public Revenues*, New York: Henry Holt and Company, 1898.

[2] Aiyagari, S Rao, "Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting," *Journal of Political Economy* 103:6 (1995), 1158-75.

[14] There is of course one important caveat. Where policy makers employ asset taxation for the purposes of redistribution, perhaps shifting resources to people with higher marginal utilities, as is in Conesa *et al.* (2009) and Fehr and Kindermann (2014), these effects would compound by lowering the potential efficiency losses that such policies might otherwise entail.

[3] Atkinson, Anthony B. and Joseph E. Stiglitz, "The design of tax structure: direct versus indirect taxation," *Journal of Public Economics* 6:(1-2) (1976), 55–75.

[4] Banks, James, and Peter Diamond. "The base for direct taxation, " *In Dimensions of Tax Design: The Mirrlees Review*, ed. Institute for Fiscal Studies, 548–648. Oxford University Press, 2010.

[5] Ben-Gad, Michael,. "Fiscal policy and indeterminacy in models of endogenous growth," *Journal of Economic Theory* 108:2 (2003), 322–344.

[6] Chamley, Christophe, "The welfare cost of capital income taxation in a growing economy," *Journal of Political Economy*, 89:3, (1981) 468–96.

[7] Chamley, Christophe, "Optimal taxation of capital income in general equilibrium with infinite lives," *Econometrica* 54:3 (1986), 607-622.

[8] Coleman II, Wilbur John, "Welfare and optimum dynamic taxation of consumption and income," *Journal of Public Economics*, 76:1, (2000), 1-39.

[9] Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger, "Taxing capital? Not a bad idea after all!," *American Economic Review*, 99:1, (2009), 25-48.

[10] Correia, Isabel H., "Should capital income be taxed in the steady state?," *Journal of Public Economics*, 60:1, (1996), 147-151.

[11] Domeij, David and Jonathan Heathcote, "On the distributional effects of reducing capital taxes," *International Economic Review*, 45:2, (2004), 523-553.

[12] Eerola, Essi and Niku Määttänen, "The optimal tax treatment of housing capital in the neoclassical growth model," *Journal of Public Economic Theory*, 15:6, (2013), 912-938.

[13] Fehr, Hans and Fabian Kindermann, "Taxing capital along the transition—Not a bad idea after all?," *Journal of Economic Dynamics and Control*, 51:1, (2015), 64-77.

[14] Fernald, John, A Quarterly, "Utilization-Adjusted Series on Total Factor Productivity," Federal Reserve Bank of San Francisco Working Paper (2014).

[15] Golosov, Mikhail, Narayana Kocherlakota and Aleh Tsyvinski, "Optimal indirect and capital taxation," *Review of Economic Studies*, 70:3, (2003), 569-587.

[16] İmrohoroğlu, Selahattin, "A quantitative analysis of capital income taxation," *International Economic Review*, 395:2, (1998), 307-328.

[17] Johansen, Søren, *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press, 1995.

[18] Judd, Kenneth L., "Redistributive taxation in a simple perfect foresight model," *Journal of Public Economics* 28:1 (1985), 59–83.

[19] Judd, Kenneth L., "Optimal taxation and spending in general competitive growth models," *Journal of Public Economics* 71:1 (1999), 1–26.

[20] Lucas, Robert E. Jr., "Supply-side economics: an analytical review," *Oxford Economic Papers* 42:2 (1990), 293-316.

[21] Musgrave, Richard A. and Alan T. Peacock, ed. *Classics in the Theory of Public Finance*, New York: St Martin's Press, 1967.

[22] Peacock, Alan T., and Jack Wiseman, *The Growth of Public Expenditure in the United Kingdom,* Princeton: Princeton University Press, 1961.

[23] Reichling, Felix and Charles Whalen, "Review of estimates of the Frisch elasticity of labor supply," Working Paper Series, Working Paper 2012-13, (2012), Congressional Budget Office, Washington, D.C.

[24] Reis, Catarina, "Entrepreneurial labor And capital taxation," *Macroeconomic Dynamics* 15:3 (2011), 326–335.

[25] Straub, Ludwig and Iván Werning, "Positive long run capital taxation: Chamley-Judd revisited," MIT working paper, February 2015.

[26] Trabandt, Mathias and Harald Uhlig, "The Laffer curve revisited," *Journal of Monetary Economics*, 58:4 (2011), 305–327.

[27] Wickens, Michael R., "Interpreting cointegrating vectors and common stochastic trends," *Journal of Econometrics*, 74:2 (1996), 255–271.