Understanding Capital Taxation in Ramsey Models*

Stefania Albanesi
Columbia University, NBER and CEPR

Roc Armenter
Federal Reserve Bank of New York

December 12, 2007

Abstract

Most Ramsey models prescribe that capital taxes should be zero in the long run. We propose a new argument for this result that relies on the government's ability to reallocate distortions over time. The argument translates into the following principle: If all distortions can be front-loaded, any distortion on the intertemporal margin will be purely temporary. This front-loading principle is very general and delivers predictions on distortions, not merely taxes. We show that it can be applied to a very large class of Ramsey models and that it can be used to identify a class of Pareto-improving reforms of suboptimal policies.

*PRELIMINARY AND INCOMPLETE. The views expressed here do not necessarily reflect the views of the Federal Reserve Bank of New York or the Federal Reserve System. Contact: Stefania.Albanesi@columbia.edu and Roc.Armenter@ny.frb.org.

1 Introduction

The optimal setting of fiscal policies is a fundamental question in macroeconomics. The Ramsey approach is perhaps the most popular framework to explore this problem. This paradigm applies the tools of public finance in a dynamic general equilibrium setting. A benevolent government sets fiscal policy once and for all at an initial date. Fiscal policy consists of a combination of distortionary taxes on factor income or commodities. Lump sum taxes are ruled out.

A key lesson from public finance is that distortions should be optimally spread across all margins. Yet, the vast majority of Ramsey models prescribe that capital taxes should be zero in the long run so there are no permanent distortions in the intertemporal margin, even as distortions persist along other margins. Chamley (1986) and Judd (1985) first demonstrated this result in a deterministic setting. The absence of permanent intertemporal distortions in the optimum is very robust. Atkeson, Chari and Kehoe (1999) show that it holds in a broad class of deterministic economies. Zhu (1992) confirms the finding for representative agent economies with aggregate risk and Farhi (2006) extends the result to incomplete markets.[1] Aiyagari (1995) derives the result in an economy with idiosyncratic risk with incomplete insurance and borrowing constraints.[2] Given its generality, the absence of permanent intertemporal distortions has defined the conventional wisdom on the properties of optimal policies.[3]

The main goal of this paper is to deepen our understanding of intertemporal distortions in Ramsey taxation models. To this end, we propose a new argument for the Chamley-Judd result that relies on the government's ability to reallocate distortions over time. The reasoning can be heuristically described as follows. Consider a candidate optimal policy with a permanent *inter*temporal distortion corresponding to a positive capital tax in steady state. The wedge between the intertemporal rate of substitution and the return to capital implies that a reallocation of consumption from current to later periods would have first order welfare gains. Such a reallocation can be achieved by changing the timing of *intra*temporal distortions. For example, the government could raise labor taxes today to finance a tax cut in future periods. In a steady state, the direct welfare effects of the labor tax changes offset each other. Hence, the net welfare effect is the gain from the intertemporal reallocation of consumption. This contradicts the optimality of the candidate policy. A similar argument can be constructed if the candidate optimal policy features a capital subsidy.

This argument translates into the following principle: If it is possible for the government to front-load all distortions, there will be no permanent intertemporal wedge. This reasoning departs from the most common interpretation of the Chamley-Judd result based on uniform commodity taxation. Since a permanent capital tax is equivalent to an ever-increasing consumption tax, the uniform commodity taxation principle suggests that it should be zero.

[1] With aggregate risk, the optimal intertemporal wedge is or fluctuates around zero in the long run.

[2] In Aiyagari (1995), aggregate capital converges to the first best and individual intertetemporal wedges fluctuate around zero.

[3] The result fails in models with incomplete factor taxation, as shown in Correia (1996) and Jones, Manuelli, and Rossi (1997). Lansing (1999) presents an economy with a balanced budget constraint on the government and heterogeneous agents in which capital tazes are not zero in the long run under logarithmic utility.

However, this logic is limited. While uniform commodity taxation requires weak separability between consumption and leisure[4], the absence of intertemporal distortions in the steady state does not hinge on the properties of preferences. Moreover, economies with private information, perhaps the best known example of permanent intertemporal distortions, satisfy uniform commodity taxation, as shown in Golosov, Kocherlakota and Tsyvinski (2003).

The notion that the absence of permanent intertemporal distortions depends on the ability to front-load distortions is very general. For one, the argument can be formulated for a virtually all Ramsey models. The presence of aggregate uncertainty, incomplete markets or additional frictions does not alter the structure of the proof. Moreover, the argument is more general since it delivers predictions on distortions, not merely taxes. There are several examples of economies where the capital tax is different from zero in the steady state but there are no permanent intertemporal distortions. One well known case is Aiyagari (1995). Under the optimal policy, aggregate capital attains the first be level while the individual intertemporal wedge fluctuates around zero. To implement this allocation, a positive steady state capital tax is necessary. Thus, the borrowing constrained agents face a negative intertemporal wedge, whereas the unconstrained agents face a positive wedge. Another notable example is Erosa and Gervais (2000). They analyze an overlapping generations economy in which capital taxes or subsidies are typically needed to implement the optimal allocation, but there are no permanent intertemporal distortions. A capital income taxes may be required to ensure that the social discount factor is equated to the marginal return to aggregate capital when the government's objective function does not weigh future generations enough. In addition, given that consumption growth varies over the lifecycle, agents' intertemporal wedge could be positive or negative depending on age. In these economies, the capital taxes serve to correct intrinsic distortions on the intertemporal margin which arise due to frictions that are not typically present in the standard Ramsey model.

Importantly, the front-loading principle also applies to a large variety of optimal policy problems outside the Ramsey paradigm. Albanesi and Armenter (2007) formally analyze a very general class of public finance problems that encompasses the Ramsey model, as well as economies with limited commitment, private information and political economy frictions. For this class of economies, a sufficient condition rules out permanent intertemporal distortions at the optimum: If there exists an admissible allocation that converges to the first best steady state, then all intertemporal distortions are temporary in the second best. This result bears a clear connection with the front-loading principle since convergence to the first best steady state requires all distortions to be front-loaded.

The implications of the front-loading principle go beyond the predictions for optimal policies. Ultimately, this reasoning implies that whenever private savings are distorted, the government should be reducing debt and lower taxes in future periods. This constitutes an important lesson for policy design and suggests a novel class of Pareto-improving reforms that do not rely on adjustment to the path of capital taxes.

To illustrate this point, we present a series of policy reforms in a two-class economy similar to Judd (1985). The status-quo policy is calibrated to match stylized facts about the U.S. fiscal

[4] See Atkinson and Stiglitz (1972).

policy — a relatively high capital tax rate among them. Our benchmark is a Pareto-improving reform that temporarily raises labor taxes and uses the proceeds to finance a cut in future labor taxes. The policy eventually returns to the initial status quo so all welfare changes are due to the temporary reallocation of labor taxes across periods. This reform amounts to a front-loading of intratemporal distortions and delivers welfare gains for a variety of parametrizations and specifications of the reform.

The paper is organized as follows. Section 2 describes our argument for the absence of intertemporal distortions in a benchmark Ramsey model. We argue for the generality of the argument in Section 3 and include a brief discussion of Ramsey models with uncertainty, incomplete markets, and overlapping generations. Section 4 discusses optimal policy in some environments where it is not possible to front-load all distortions and thus our argument does not apply. Section 5 describes a series of policy reforms which suggest that the front-loading principle can generally be applied even beyond optimal policy analysis. Section 6 concludes.

2 Front-loading and Capital Taxation

We start by describing the front-loading argument for the benchmark deterministic Ramsey model. A fiscal policy is a sequence of linear tax rates on labor and capital $\left\{\tau_t^n, \tau_t^k\right\}_{t=0}^{\infty}$ as well as debt holdings $\{b_t\}_{t=1}^{\infty}$. Tax revenues finance an exogenous stream of government consumption $\{g_t\}_{t=0}^{\infty}$ which, for simplicity, we set to be constant $g_t = g$. The absence of lump sum taxes is the only source of frictions in this economy.A competitive equilibrium requires firms and households to make optimal decisions given prices and policy, the government budget constraint to be satisfied, and all markets to clear.[5] Consumption, investment, and labor decisions $\{c_t, k_{t+1}, n_t\}_{t=0}^{\infty}$ must conform to a competitive equilibrium given taxes and prices.

A key feature of the competitive equilibrium is that the government's ability to borrow and lend implies that the sequence of flow budget constraints can be collapse into one present value constraint at $t = 0$:

$$\sum_{t=0}^{\infty} q_0^t c_t \leq b_0 + \left(\left(1 - \tau_0^k\right) r_0 + 1 - \delta\right) k_0 + \sum_{t=0}^{\infty} q_0^t \left(1 - \tau_t^n\right) w_t n_t. \tag{1}$$

Here, q_0^t is the discount price of a bond that pays one unit of consumption at date t, w_t and r_t are the rental rates of labor and capital respectively, and δ the depreciation rate.

A Ramsey equilibrium is simply the competitive equilibrium that maximizes the households' welfare from the standpoint of date $t = 0$. Optimal policies implement the Ramsey equilibrium allocation.

Consider a *candidate optimal policy* with positive capital taxation in the steady state:

$$\begin{aligned} \lim_{t\to\infty} \tau_t^k &= \tau^k > 0, \\ \lim_{t\to\infty} \tau_t^n &= \tau^n \geq 0. \end{aligned}$$

[5] We omit a formal description of the environment and competitive equilibrium conditions. See Chari and Kehoe (1999) for a detailed discussion.

The candidate policy must satisfy the government budget constraint and conform to a competitive equilibrium. Since our argument is formulated at steady state, we can abstract from the properties of this policy on the transition path, provided the allocation converges. We will show that such a policy cannot be optimal by constructing an alternative policy $\left\{\tilde{\tau}^n_t, \tilde{\tau}^k_t\right\}_{t=0}^{\infty}$ which delivers strictly higher welfare.

First, consider the household's equilibrium conditions in steady state under the candidate policy. The Euler equation is:

$$u^c_{ss} = \beta u^c_{ss} \left\{ \left(1 - \tau^k\right) r_{ss} + 1 - \delta \right\},$$

where we use subscript ss to denote the steady state variables. Equating the rental rate to the marginal product of capital, the positive capital tax implies:

$$\beta^{-1} < F^k_{ss} + 1 - \delta. \tag{2}$$

This defines the *intertemporal wedge*. As long as there is a positive capital tax, there will be intertemporal distortions.

Similarly, the household equates the after-tax wage rate to the marginal rate of substitution,

$$-\frac{u^n_{ss}}{u^c_{ss}} = \left(1 - \tau^n\right) w_{ss}.$$

The wage rate is equal to the marginal product of labor and the marginal rate of substitution is thus below the marginal rate of transformation

$$-\frac{u^n_{ss}}{u^c_{ss}} \leq F^n_{ss}. \tag{3}$$

This defines the *intratemporal wedge*. As long as there is a positive labor tax, there will be intratemporal distortions.[6]

The intertemporal wedge (2) implies that, ceteris paribus, there are first-order welfare gains from reallocating consumption from date t to $t+1$. Now, this is exactly what a tax reform can do by changing *the timing of intratemporal distortions*. To this end the government can raise taxes and run a primary surplus at date t, and then use the savings to finance a tax cut at date $t+1$. By front-loading the distortions, the government effectively transfers resources from t to $t+1$. The welfare impact of the labor supply distortions at dates t and $t+1$ are symmetric and cancel each other when evaluated at the steady state. As a result the only first-order welfare effect is the gain from reallocating consumption intertemporally.

We now formally lay out the proposed tax reform. Let allocations $\left\{\tilde{c}_t, \tilde{k}_{t+1}, \tilde{n}_t\right\}_{t=0}^{\infty}$ be the competitive equilibrium induced by the reformed policy $\left\{\tilde{\tau}^n_t, \tilde{\tau}^k_t\right\}_{t=0}^{\infty}$. The tax reform involves only two dates, t and $t+1$; allocations and policies at all other periods are identical than under the candidate policy, $\tilde{c}_j = c_j$, $\tilde{n}_j = n_j$ and so on for all dates $j \neq t, t+1$. Date t is assumed to be late enough such that the economy under the candidate policy is in the steady state for date

[6] In this simple Ramsey model, as in most of them, all wedges stem from distortionary taxation. We will later discuss economies with intrinsic frictions, such as incomplete markets.

t and later, $c_j = c_{ss}$, $n_j = n_{ss}$, and so on for all $j \geq t$. The policy reform should satisfy not only the government budget constraint but all equilibrium conditions, including those leading to date t.

We start by raising intratemporal distortions at date t. As we perturb the intratemporal margin, we adjust the tax rates and debt holdings so that all equilibrium conditions are restored without any change at later dates. We illustrate this adjustment in the space of allocations. To this end, we substitute the equilibrium conditions for bond holdings and labor:

$$
\begin{aligned}
q_0^t &= \beta^t \frac{u_t^c}{u_0^c}, \\
-\frac{u_t^n}{u_t^c} &= \left(1 - \tau_t^n\right) w_t,
\end{aligned}
$$

into constraint (1) to obtain:

$$
\left(u_0^c\right)^{-1} \sum_{t=0}^{\infty} \beta^t \left(u_t^c c_t + u_t^n n_t\right) = b_0 + \left(\left(1 - \tau_0^k\right) r_0 + 1 - \delta\right) k_0. \tag{4}
$$

By Walras' law, (4) ensures that both the household and the government's budget constraint are satisfied. Indeed, it is sufficient that condition (4) is satisfied for allocations to constitute a competitive equilibrium.[7]

Let the reduction in consumption at time t be given by:

$$
dc_t = -\varepsilon,
$$

for $\varepsilon > 0$ arbitrarily small. The corresponding adjustment in labor supply must satisfy:

$$
d\left[u_t^c c_t + u_t^n n_t\right] = 0,
$$

in order to restore the competitive equilibrium conditions without resorting to changes at later dates. Hence, the competitive equilibrium conditions dictate:

$$
dn_t = \alpha dc_t \tag{5}
$$

for some α solving $d\left[u_t^c c_t + u_t^n n_t\right] = 0$. It follows that:

$$
dn_t = -\alpha\varepsilon.
$$

This perturbation increases distortions and time t by reducing consumption. Labor supply can contract or expand as a consequence of this adjustment. No restrictions are needed on the sign of α.

The next step is to derive the corresponding variation in investment. The resource constraint at date t is:

$$
c_t + k_{t+1} + g \leq F\left(k_t, n_t\right) + \left(1 - \delta\right) k_t.
$$

[7] See Chari and Kehoe (1999) for a formal proof. The reader familiar with Ramsey problems will recognize in the discussion the logic of the implementability constraint and the primal approach. We return to them in the next Section.

Then, the change in investment must satisfy :

$$dk_{t+1} = F^n_{ss} dn_t - dc_t,$$

which implies $dk_{t+1} = (1 - \alpha F^n_{ss})\,\varepsilon$. The drop in consumption will not map one to one into investment, i.e., $dk_{t+1} \neq -dc_t$, as long as the labor supply responds. Again, the sign of the change in investment is inconsequential, though we would expect it to rise.[8]

The next step is to use the additional resources to increase consumption and reduce intratemporal distortions at date $t+1$. As at date t, the variation in consumption must be accompanied by a corresponding adjustment in labor supply to restore the equilibrium conditions:

$$d\left[u^c_{t+1} c_{t+1} + u^n_{t+1} n_{t+1}\right] = 0.$$

Since we are evaluating a small steady state variation, as in (5), the relationship between the change in labor and consumption satisfies:

$$dn_{t+1} = \alpha dc_{t+1}.$$

We differentiate the resource constraint at date $t+1$ in order to pin down the change in consumption:

$$dc_{t+1} = F^n_{ss} dn_{t+1} + \left(F^k_{ss} + 1 - \delta\right) dk_{t+1}.$$

Using the derived change in investment, $dk_{t+1} = (1 - \alpha F^n_{ss})\,\varepsilon$, and the relationship between labor and consumption $dn_{t+1} = \alpha dc_{t+1}$, we have that

$$dc_{t+1} = \alpha F^n_{ss}\left(dc_{t+1} - \left(F^k_{ss} + 1 - \delta\right)\varepsilon\right) + \left(F^k_{ss} + 1 - \delta\right)\varepsilon.$$

It is then clear that:

$$dc_{t+1} = \left(F^k_{ss} + 1 - \delta\right)\varepsilon.$$

It is crucial that the government, by front-loading distortions, can postpone consumption and obtain the pre-tax return, $F^k_{ss} + 1 - \delta$, higher than the rate of return faced by householdsunder the candidate policy.

Finally, collecting all the welfare effects at date t and date $t+1$:

$$d\left[u\left(c_t, n_t\right) + \beta u\left(c_{t+1}, n_{t+1}\right)\right] = u^c_{ss}\left(dc_t + \beta dc_{t+1}\right) + u^n_{ss}\left(dn_t + \beta dn_{t+1}\right).$$

The relationship between the change in consumption and labor is constant at both dates,

$$dn_t + \beta dn_{t+1} = \alpha\left(dc_t + \beta dc_{t+1}\right),$$

and as a result the corresponding welfare effect is:

$$d\left[u\left(c_t, n_t\right) + \beta u\left(c_{t+1}, n_{t+1}\right)\right] = \left(u^c_{ss} + \alpha u^n_{ss}\right)\left(dc_t + \beta dc_{t+1}\right).$$

[8] The argument would break if some additional constraint imposed a cap in aggregate capital. Indeed the government may be prevented from saving or borrowing, but if it is still can manipulate aggregate capital the argument would go through. We return to this discussion in Section 3.

There are two cases. If $u_{ss}^c + \alpha u_{ss}^n \leq 0$, the first term is negative. It is then possible to improve welfare by raising distortions at date t without any need to decrease them at date $t+1$.[9] This counter-intuitive case may arise if the economy is highly distorted. Loosely speaking, marginal taxes are so high that their reduction gives rise to an increase in revenues. It is clear that such a policy holds little interest as optimum candidate.

The interesting case corresponds to $u_{ss}^c + \alpha u_{ss}^n > 0$. The welfare change is then proportional to the change in the consumption profile:

$$d\left[u\left(c_t, n_t\right) + \beta u\left(c_{t+1}, n_{t+1}\right)\right] \propto dc_t + \beta dc_{t+1}.$$

Substituting the derived consumption change at both dates:

$$dc_t + \beta dc_{t+1} = \left(\beta\left(F_{ss}^k + 1 - \delta\right) - 1\right)\varepsilon.$$

The presence of an intertemporal wedge (2) in itself implies that the proposed policy reform delivers strictly positive welfare gains.[10]

To summarize, we have considered a candidate policy with positive capital taxes in the long run. Starting from the steady state, we show how to construct a policy reform that delivers strictly higher welfare by front-loading distortions. Since the only assumption on the candidate policy is that taxes on capital are positive in the steady state, the optimal policy cannot have this feature. Crucial to the argument is the government's ability to save.

The same reasoning can be used to show that the optimal policy cannot feature long-run capital *subsidies*. This is indeed quite trivial: why would be using distortionary labor taxes to finance another distortion, in this case a capital subsidy? That said, it is possible to follow the logic above: in this case, though, the welfare gains follow from back-loading distortions.

3 A General Argument for Zero Capital Taxes

The front-loading principle can be used to derive the optimality of zero capital taxes for more general Ramsey models and indeed in holds for a broad class of public finance problems. In this section, we formalize the argument more generally. To this end, we introduce the primal approach to the Ramsey problem.[11] This approach begins by characterizing the set of competitive equilibrium allocations, that is, the allocations that can be supported in equilibrium for some choice of prices and tax rates, just in terms of constraints on the physical variables. Then, social welfare is maximized by choice of the allocations within this set.

[9] To see this, note that $-\left(u_{ss}^c + \alpha u_{ss}^n\right)\varepsilon$ is the welfare effect at date t of raising distortions. The resource constraint can be satisfied without any change in investment, since it can be shown that (3) implies $dc_t - F_{ss}^n dn_t \leq 0$ whenever $u_{ss}^c + \alpha u_{ss}^n \leq 0$.

[10] The reader may feel that we have concluded the policy reform without actually saying anything about policy. Indeed the details on how policy changes depend quite a bit on the specification. In Section 5 we illustrate the proof with a series of policy experiments in a standard Ramsey model.

[11] The primal approach was pioneered in Atkinson and Stiglitz (1980) and Lucas and Stokey (1983).

For the benchmark Ramsey model in the previous section, any feasible allocation belongs a competitive equilibrium if and only if it satisfies *the implementability constraint:*

$$\sum_{t=0}^{\infty} \beta^t z\left(c_t, n_t\right) \geq a_0, \tag{6}$$

where

$$z\left(c_t, n_t\right) = u_t^c\left(c_t, n_t\right) c_t + u_t^n\left(c_t, n_t\right) n_t,$$

and a_0 is the shadow value of initial assets,

$$a_0 = u_0^c\left\{b_0 + \left(\left(1 - \tau_0^k\right) F_0^k + 1 - \delta\right) k_0\right\}.$$

The implementability constraint (6) can be derived by substituting the equilibrium conditions into the date $t = 0$ household budget constraint.

The primal approach reveals a remarkable similarity across many Ramsey models. The set of competitive equilibrium allocations can usually be characterized by a system of implementability constraints of the form

$$\sum_{t=0}^{\infty} \beta^t Z\left(x_t\right) \geq A_0, \tag{7}$$

where x_t is a vector collecting all consumption and labor decisions in the economy,

$$x_t = \left\{c_{1t}, c_{2t}, ..., n_{1t}, n_{2t}, ...\right\},$$

and Z is a well-behaved function from the allocation space X to $\Re^m$, where m is the finite number of constraints. Stochastic environments with or without complete markets require some further generalization. We will discuss them later.

We now revisit our argument strictly in terms of allocations and the generalized implementability constraint (7). Consider a candidate optimal allocation that features a strictly positive intertemporal wedge in steady state (2). We construct a perturbation of the consumpttion profile at dates t and $t + 1$ as follows:

$$\begin{aligned} dc_t &= -\varepsilon, \\ dc_{t+1} &= \left(F_{ss}^k + 1 - \delta\right) \varepsilon, \end{aligned}$$

where $\varepsilon > 0$ is arbitrarily small vector evaluated at steady state. We now need to adjust labor supply so that the resulting perturbation in the allocation dx_t preserves the competitive equilibrium conditions. The admissible perturbations are define by the implementability constraint (7):

$$D^x Z_{ss} dx_t = 0.$$

The system above identifies the labor supply response to the change in consumption:

$$dn_t = \Gamma dc_t.$$

where Γ is a matrix of the appropriate dimensions.[12] The same relationship works at date $t+1$, $dn_{t+1} = \Gamma dc_{t+1}$. The labor supply distortions have a symmetric effect at both dates and we can check the feasibility of the proposed variation following the same steps as before.

Collecting all the welfare changes at dates t and $t+1$:

$$D^x \left[U\left(x_{ss}\right) + \beta U\left(x_{ss}\right)\right] dx_t = D^c U\left(x_{ss}\right)\left(dc_t + \beta dc_{t+1}\right) + D^n U\left(x_{ss}\right)\left(dn_t + \beta dn_{t+1}\right).$$

Using the restriction:

$$dn_t + \beta dn_{t+1} = \Gamma\left(dc_t + \beta dc_{t+1}\right),$$

we obtain:

$$D^x \left[U\left(x_{ss}\right) + \beta U\left(x_{ss}\right)\right] dx_t = \left(D^c U\left(x_{ss}\right) + D^n U\left(x_{ss}\right)\Gamma\right)\left(dc_t + \beta dc_{t+1}\right).$$

Without loss of generality, the first term can be taken to be strictly positive. The second term is strictly positive if:

$$dc_t + \beta dc_{t+1} = \left(\beta\left(F_{ss}^k + 1 - \delta\right) - 1\right)\varepsilon > 0,$$

that is, as long as the intertemporal wedge is postive. The first-order effect on welfare is thus strictly positive.

The argument merely hinges on the presence of a positive intertemporal wedge at the steady state as well as the fact that we can restore the equilibrium conditions within each period.[13] The front-loading principle thus applies to any Ramsey model or more general public finance problem where allocations are subject to a constraint of the form of (7).

3.1 Stochastic Economies

The counterpart of zero capital tax result for stochastic economies is given in Zhu (1992). If the optimal policy plan converges to a stationary distribution, then either the optimal capital tax is zero or fluctuates around zero. In other words, it is never optimal to tax or subsidize capital with probability one.

Zhu (1992) shows that, if there is a set of complete Arrow-Debreu markets, competitive equilibrium allocations satisfy the following implementability constraint

$$E_0 \left\{\sum_{t=0}^{\infty} \beta^t z\left(c_t, n_t, s_t\right)\right\} \geq a_0,$$

where s_t is a vector of exogenous shocks and E_0 is the expectation operator. The setting is, thus, similar to the non-stochastic economy. Our argument, though, seems to rely heavily on steady-state analysis. Is it possible to adapt it to a stochastic economy?

The variation argument goes through with an important modification. Consider a policy which taxes capital with probability one and converges to a stationary distribution. An important implication of stationarity is that allocations can be characterized by a time-invariant

[12] The existence of a solution requires that the optimal plan satisfies non-degenerate constraint qualifications.

[13] There are some additional regularity conditions on Z and allocations. See Albanesi and Armenter (2007) for a full discussion.

function of the state of the economy. As before, we seek to exploit the intertemporal wedge by transferring consumption from date t to a later date. Now, though, we will carry the additional resources forward until the state of the economy at date t is revisited and allocations are identical to those at date t. How long it takes will depend on the realization of shocks and thus the stopping time d is a random variable. Stationarity guarantees that we always revisit any given state so the stopping time d is finite with probability one. Allocations before date t and after any realization of date $t + d$ are left intact.

By construction the labor distortions have a symmetric effect at both dates t and $t+d$. We thus have the crucial feature used in the steady state analysis. Following essentially the same steps as before, it is then possible to prove that the proposed variation is feasible and that the welfare effect is proportional to the accumulated change in consumption, $dc_t + \beta^d dc_{t+d}$. We can then apply the front-loading principle. We reduce consumption at date t by ε. Whenever the sequence of shocks is, we rebate the proceedings

$$dc_{t+d} = \left(\Pi_{j=t+1}^{d} \left(F_j^k + 1 - \delta\right)\right) \varepsilon.$$

The expression above holds for all possible realizations of the stopping time d — different realizations will have a different value for dc_{t+d}. However, if the capital tax is positive at all times, then we know that

$$\beta^d \left(\Pi_{j=t+1}^{d} \left(F_j^k + 1 - \delta\right)\right) > 1$$

for all realizations of d. The effect of labor distortions at date t and $t + d$ are symmetric and cancel. We thus conclude that the variation delivers positive welfare gains.

This reasoning can be extended to incomplete markets. If the only asset traded is a risk-free bond, it is not possible to characterize the set of competitive equilibrium with a single implementability constraint.[14] Instead, there is an implementability constraint for every sequence of shocks $s^t = \{s_0, s_1, ..., s_t\}$,

$$E_{s^t} \left\{ \sum_{j=t}^{\infty} \beta^{j-t} z\left(c_j, n_j, s_j\right) \right\} = b\left(s^{t-1}\right) + V\left(s^t\right),$$

where E_{s^t} is the conditional expectation operator, $b\left(s^{t-1}\right)$ are the holdings of a risk-free one period bond, and $V\left(s^t\right)$ is the present value of capital returns and non-negative government transfers. Our variation argument fails because the shifting of consumption between periods will necessary change debt holdings at date t. In order to satisfy the implementability constraints, it would be necessary to redistribute distortions across states. This gives rise to a possible welfare loss which could overcome the welfare gains from intertemporal redistribution.

However, the necessary conditions for optimality often imply that, eventually, future implementability constraints are no longer binding and can be safely ignored.[15] Intuitively, the government can escape the additional constraints imposed by the incomplete markets by accumulating enough assets and then smoothing distortions across states using non-negative

[14] Farhi (2006) analyzes a similar economy but does not use the primal approach.

[15] It is beyond the scope of this paper to prove this result. See Aiyagari et al. (2002) and Farhi (2006), as well as Albanesi and Armenter(2007), for a proof and discussion of the result.

transfers. Hence once the government has enough assets the problem *from then on* is not different from the Ramsey model with complete markets that we discussed before.

Asset Restrictions. Our last discussion concerns the implications of a balanced-budget constraint on the government. In an otherwise standard Ramsey model, the government is forced to finance its expenditure solely from current tax revenues, that is, it cannot save or borrow. Both Judd (1985) and Chamley (1986) claim that the absence of permanent intertemporal distortions does not depend on assumptions about the government's ability to borrow or lend.[16]

This result may be surprising in light of the crucial role of the ability to front-load distortions for our result. How can the government reallocate distortions across periods without being able to save or borrow? It turns out that the government usually can independently manipulate the path of consumption and capital even if it has no access to debt. It is actually helpful to consider an economy where it *cannot*. Lansing (1999) shows that a zero capital tax cannot be implemented in a two-class economy with logarithmic preferences and a balanced-budget constraint. The key observation is that the present value of future consumption only depends on current consumption under logarithmic preferences. As a result the government cannot induce a change in the intertemporal profile of consumption and an allocation with no permanent intertemporal distortions cannot generically be implemented in equilibrium.

3.2 Distortions and Taxes

In the models analyzed thus far, taxes are the only source of frictions. This need not be a general property of Ramsey models. The solution of the Ramsey problem prescribes a particular pattern of distortions. An optimal tax system implements this pattern—indeed it does not need to be unique. It is clear thus that we need to distinguish between taxes and distortions. The discussion above makes clear that the front-loading logic applies to distortions. So it is possible for the optimal capital tax to be different from zero while intertemporal distortions must be zero.

We illustrate this reasoning with two examples. Aiyagari (1995) presents an economy with idiosyncratic productivity risk, incomplete markets and borrowing constraints. Even if there are no capital taxes, there is an intertemporal wedge—in this case arising from the lack of complete markets and the borrowing constraint. A precautionary motive leads agents to over-accumulate capital.

Erosa and Gervais (2000) study optimal fiscal policy in an overlapping generations economy. As it is well-known there may be a dynamic inefficiency in overlapping generations economies as private agents fail to internalize the benefits of investment for future generations. Therefore, the level of capital may not be at the social optimum.

In Aiyagari (1995) the optimal capital tax is positive yet the level of aggregate capital is undistorted. That is, the positive capital tax offsets the precautionary demand for capital. In Erosa and Gervais (2000) the optimal capital tax may be different from zero, but the return to capital is always equated with the social discount factor. In an over-lapping generations

[16] See also Stockman (2001) for a detailed analysis.

economy the individual discount factor varies with age. With age-specific taxes, private and social discount factors can be equated without resorting to capital taxes. Without age-specific taxes, capital taxes achieve this purpose.

3.3 Beyond Ramsey Models

The idea that front-loading rules out permanent intertemporal distortions applies to a large variety of public finance problems beyond the class of Ramsey models. Albanesi and Armenter (2007) consider a large class of public finance problems that can be represented as a choice of allocations subject to feasibility and a set of additional constraints, which are called *admissibility constraints*. The Ramsey problems discussed above are examples of this class: the implementability constraints are a special case of the admissibility constraints. The general framework also encompasses problems with incentive compatibility constraints due to limited commitment or private information, as well as political economy frictions and possibly arbitrary constraints on assets.[17]

Albanesi and Armenter (2007) identify a sufficient condition that rules out permanent intertemporal distortions in the optimum. If there exists an allocation that satisfies all constraints and eventually converges to the same steady state as the first best allocations, then there are no permanent intertemporal distortions in the second best. Returning to the Ramsey model, if we can find a policy that leads the economy to the same steady state as the first best allocations, then the optimal capital tax will be zero in the long run. The first best steady state requires no distortionary taxation. It is easy to find a policy that eventually achieves it. With an initial high taxation phase, the government can accumulatie assets to the point where all public expenditures can be financed from the return to these assets. This eliminates the need to levy distortionary taxes. Such a policy will rarely be optimal but for our result to apply it is only necessary that such a policy exists. This makes the sufficient condition typically straightforward to verify.

What is the relationship with front-loading? If it is possible to find an allocation that converges to the first best steady state, then the admissibility constraints allow for all distortions to be front-loaded. This means the variational argument above can be applied and we can disprove that an allocation featuring permanent intertemporal distortions is optimal.

Albanesi and Armenter (2007) also discuss a weaker sufficient condition which is particularly relevant for Ramsey models. If the level of aggregate capital can be decoupled of its distribution among agents, then as long as there is an admissible plan that achieves the first best level of capital, there would be no permanent intertemporal distortions in the second best. The condition is weaker, thus, since it does not require that labor or consumption is efficiently allocated. However it only applies into a subset of public policy problems.[18]

[17] Some nomenclature is needed. We refer to the solution subject to feasibility and admissibility constraints as the second best allocation. The choice of allocations subject only to resource feasibility corresponds to the notion of the first best.

[18] We have discussed one of them: Aiygari (1995).

4 When the Front-loading Argument Fails

In this section we briefly discuss several instances where the front-loading principle fails. First, the government may not be able to shift intratemporal distortions across periods. This may be the case because the government does not have access to the necessary assets or because the distortions themselves cannot be redistributed across periods. Asset restrictions could give rise to the first example. Economies with private information typically display perpertually binding incentive compatibility constraints and fall in the second category.

A second possibility is that there does not exist a perturbation that front-loads distortions while preserving the competitive equilibrium conditions. Ramsey models with incomplete factor taxation exemplify this problem. Finally, we discuss the case where the front-loading principle fails due to restrictions on the tax choices that induce history dependence.

Incomplete Factor Taxation. As first established in Correia (1996) and Jones, Manuelli, and Rossi (1997), optimal capital taxes may be positive in the steady state if the government cannot tax *every* factor of production at the rate of choice. It must be thus that our variation argument cannot be applied. This is perhaps surprising because there is no clear impediment to front-loading like in the previous examples.

Under incomplete factor taxation, our proposed variation does not restore a competitive equilibrium. The change in investment directly affects the household's income through the rents from the untaxed factor. The labor supply must then adjust further in order to satisfy the equilibrium conditions.[19] Unlike consumption, there is no symmetric reduction in investment to reverse the welfare impact of the additional change in labor. We are left, thus, with an additional first order welfare effect which can counter the gains from intertemporal redistribution.

History Dependence. Our argument may also fail if, for whatever reason, taxes display some form of history dependence. For example, the government may be restricted to pick a constant capital tax for all periods. In this case our variation argument fails because it generally requires adjusting the capital tax. Assuming a constant capital tax is common in applied work on Ramsey models — see, for example, Domeij and Heathcote (2004). In more general terms, any model with a cost of adjusting policy may feature a positive capital tax in steady state.

Private Information. Consider an economy where individual productivity is subject to idiosyncratic shocks that are not publicly observable. Private information gives rise to incentive compatibility constraints. This friction is a source of distortions. The government cannot fully insure idiosyncratic risk and simultaneously induce agents to work according to their productivity. This generates a trade-off between insurance and efficiency.

Crucially, the distortions arising from private information cannot be fully front-loaded is agents are risk averse as long as idiosyncratic uncertainty persists. This gives rise to a permanent intertemporal wedge. That is not to say that the optimal allocation does not

[19] Armenter (2007) shows that the result depend on the additional constraints imposed on policy at date $t = 0$. If the government is barred from manipulating the value of initial assets, then the Chamley-Judd result re-appears.

engage in an intertemporal distribution of resources over time. Indeed, the optimal plan often calls for downward path of wealth over time in order to ameliorate the provision of incentives. Our argument fails, thus, because an increase in distortions at date t does not map into a proportional reduction at a later date.

5 Policy Reforms

We derived the Chamley-Judd by proposing a welfare-improving policy reform to a candidate optimal policy which was taxing capital. The argument is perhaps surprising in that we did not rely on a reform of the long-run capital tax. Here we evaluate a series of "policy reform" based on our derivation of the Chamley-Judd result. This serves a double purpose. First it illustrates the front-loading logic by providing the path for taxes and debt; second, it shows that front-loading taxes can deliver welfare gains beyond the variational argument we used.

We work with a model with heterogeneous agents similar to Judd (1986). The status-quo policy is set to replicate some stylized facts of U.S. policy — a relatively high capital tax rate among them. Under the policy reform, labor taxes are raised in the early periods and the proceeds used to finance a tax cut over a long horizon. The policy (and the economy) eventually returns to the initial status quo so all welfare changes are due to the front-loading and not to different steady states. The policy reform is also designed to preserve the after-tax capital income under the status-quo policy.

5.1 A Two-Class Economy

We use a two-class economy similar to Judd (1986). Time is discrete and infinite, $t = 0, 1, ...$ There is no uncertainty.

There are two types of agents: capitalist households, denoted by subscript $i = 1$, who own all assets in the economy but have no labor endowment; and worker households, denoted by a subscript $i = 2$, who get all their income from renting labor and have no access to any means of savings. The extreme distribution of factor endowment is an obvious simplification, but allows us to keep track of the redistributive implications of policy reforms.

There is a unit measure of capitalists and μ of worker households. The capitalist household's preferences over consumption sequence $\{c_{1t}\}_{t=0}^{\infty}$ are

$$U_1\left(\{c_{1t}\}_{t=0}^{\infty}\right) = \sum_{t=0}^{\infty} \beta^t \frac{c_{1t}^{1-\sigma}}{1-\sigma},$$

with $0 < \beta < 1$, and $\sigma \geq 0$, with $u^1(c_{1t}) = \log(c_{1t})$ for $\sigma = 1$. The capitalist household problem consists of choosing sequence for consumption c_{1t}, capital k_t and debt holdings b_t to maximize welfare subject to a sequence of budget constraints,

$$c_{1t} + q_t b_{t+1} + k_{t+1} \leq \left(\left(1 - \tau_t^k\right) r_t^k + 1 - \delta\right) k_t + b_t$$

for all $t \geq 0$, where b_t are bonds sold at discount rate q_t, r_t^k is the capital rental rate, and τ_t^k is the capital tax rate. The depreciation rate is non-negative and strictly less than one,

$0 \leq \delta < 1$. There are also non-negativity constraints and natural debt limits which we ignore for expositional purposes.

The worker household's preferences over consumption c_2 and labor n are given by

$$U_2\left(\{c_{2t}, n_t\}_{t=0}^{\infty}\right) = \sum_{t=0}^{\infty} \beta^t \frac{\left(c_{2t} + \psi_0 \frac{(1-n_t)^{1-\psi}}{1-\psi}\right)^{1-\sigma}}{1-\sigma}$$

with $\psi_0 > 0$ and $\psi > 0$. The worker household problem is to choose consumption and labor to maximize welfare subject to the sequence of budget constraints

$$c_{2t} \leq \left(1 - \tau_t^n\right) w_t n_t$$

for all $t \geq 0$ where w_t is the wage rate. Note the worker households have no means of saving, making it impossible for them to smooth variation in their income.

There is a representative firm with a constant returns to scale production function

$$F\left(k_t, \mu n_t\right) = A k_t^{\alpha} \left(\mu n_t\right)^{1-\alpha}$$

where $A > 0$ and $\alpha \in (0, 1)$. Profit-maximization leads to equate factor prices with their respective marginal products,

$$\begin{aligned} r_t^k &= \alpha A \left(\frac{k_t}{\mu n_t}\right)^{\alpha-1}, \\ w_t &= (1-\alpha) A \left(\frac{k_t}{\mu n_t}\right)^{\alpha}. \end{aligned}$$

The government budget constraint needs to hold at every date $t \geq 0$

$$g + b_t - q_t b_{t+1} \leq \tau_t^n w_t n_t + \tau_t^k r_t^k k_t.$$

The government expenditure g is taken exogenously.

Finally we close the description of the environment with the resource constraint

$$c_{1t} + \mu c_{2t} + g + k_{t+1} \leq A k_t^{\alpha} \left(\mu n_t\right)^{1-\alpha} + (1-\delta) k_t$$

which must hold at all dates t.

A competitive equilibrium in this economy is given by a set of allocations, prices, and a policy such that both households solve their problem given prices and policy, firms maximize profits, the government budget constraint holds, and all markets clear.

5.2 Status-Quo Policy and Calibration

We will evaluate our policy reforms against a status-quo policy given by constant tax rates, $\tau_t^k = \tau^k$ and $\tau_t^n = \tau^n$ for all dates $t \geq 0$. We would like to set the tax rates and government spending $\left\{\tau^k, \tau^n, g\right\}$ as representative of the level of taxation of the U.S. Unfortunately, there is a considerable dispersion in the estimates of the average tax rates. Lucas (1990) estimates a capital and labor taxes of 36% and 40% respectively. Chari, Christiano, and Kehoe (1995) have lower estimates and choose to set the capital tax at 28% percent and the labor tax at 24% percent in their policy experiments. Some more recent work suggest capital tax rates as high as 51% —see Carey and Tchilinguirian (2000). Our baseline parametrization will use Lucas (1990) estimates and we will then check our results for a low and high taxation economies, using Chari et al. (1995) and the more recent estimates respectively. We set g such that it represents 20% of total output. We always evaluate the status-quo policy at steady state.

Our choice of the remaining parameters is pretty standard. We set the period to be one year, with the discount rate $\beta = (1.02)^{-1}$ and the depreciation rate 5%, $\delta = .05$. We set $\alpha = .36$ so capital income is roughly one third of total income. The resulting capital to output ratio is just above 3.2. The intertemporal elasticity of substitution is $\sigma = 2$. The parameters governing the labor supply are $\psi = 2$ and ψ_0 is set such that, in the steady state, labor is 60% of the time endowment. Parameters μ and A are completely irrelevant to the results, and are just normalized.

5.3 A Pareto-Improving Reform with Front-Loading

First we design the policy reform such that the after-tax capital income at all periods is the same than under the status-quo policy. The reason for this is twofold. First, it ensures the reform is welfare-neutral for the capitalist household. Second, by guaranteeing the after-tax return to all assets to be the same than under the status-quo policy, we make sure the welfare change does not arise from some form of indirect taxation of assets. Interestingly, the guarantee implies that the policy reform can be pre-announced without any change.[20]

We construct the policy reform $\left\{\tilde{\tau}_t^k, \tilde{\tau}_t^n\right\}_{t=0}^{\infty}$ as follows. Initial conditions at date $t = 0$ are given by the steady state level of capital and debt under the status-quo policy. At date $t = 0$ the labor tax is raised two percentage points to 42%. The resulting primary surplus is used to finance a labor tax cut over a long horizon — we target the tax cut to have a half-life of about 12 years.[21] At every period, the capital tax adjusts such that the after-tax return to capital is constant. Eventually the policy (and the economy) converges back to the status quo.

We find that the policy reform delivers welfare gains to the worker household despite sharps movements in their income. Figure 1 displays the allocations in terms of percentage deviations from the steady state. As expected, the policy reform first contracts the labor supply and output as we raised taxes at date $t = 0$. There is a surge in investment despite the contraction: the policy reform is effectively reallocating resources from date $t = 0$ to later dates via the front-loading of taxes. Both output and labor are sustained above their steady states for a long horizon from $t = 1$ onwards.

Figure 1 also contain the profile of consumption for both capitalist and worker households. By construction, the capitalist household consumption is constant and equal to its steady state value. The worker consumption experiences a sharp fall (about 4%) at date $t = 0$ to be above

[20] We think this is important because, in practice, policy reforms will never be unanticipated. See Klein and Domeij (2005) for a discussion of preannounced fiscal reforms.

[21] Tax provisions usually phase out slowly. We discuss some alternative policy reform designs later.

steady state from date $t = 1$ onwards. Recall the tilted consumption profile is the *source* of welfare gains, because the intertemporal wedge induced by the capital tax implies consumption in future dates is more valuable than today.

The policy variables and factor prices are in Figure 2. The capital tax rate falls by a small amount to compensate the fall in the marginal product of capital brought in by the contraction in labor and the subsequent increase in capital. It is no accident that the dynamics of debt are the mirror image of investment. Since the after-tax return to assets is unchanged, the policy reform act by changing the composition of the capitalist's portfolio, away from government debt and into capital.

5.4 An Alternative Policy Reform with a Constant Capital Tax

We also investigate a second policy reform with a constant capital tax $\tilde{\tau}_t^k = \tau^k$. The labor tax is raised by x percent for T periods, and the proceeds are used to finance a labor tax cut over a long horizon. As before, the economy converges back to the initial steady state. It is important to emphasize that such a reform effectively cuts in the value of assets even if the capital tax rate is not changed.

We choose to raise the labor tax by one percentage point over four years, and target that the subsequent cut has a half life of about twelve years. We find large welfare gains for the workers but welfare loses for capitalists. Hence this is not a Pareto-improving reform. This is quite evident from Figure 3 which depicts allocations. Capitalist consumption drops a little—not even a tenth of a percent—from steady state, but only recovers very slowly. In contrast, worker consumption falls on impact but rebounds to 3% above the steady state.

The difference in the consumption profile with respect to the previous reform is easily accounted. As before, investment increases despite the output contraction. However, this time the rise in investment also comes from a cut in capitalist consumption, easing the fall in worker consumption. The subsequent output expansion benefits only workers, as the additional tax revenues are used to cut labor taxes only. Figure 4 displays policy variables and factor prices. The resulting dynamics are very similar to the previous reform.

5.5 Robustness Analysis

We have performed both policy reforms for an array of different parametrizations and found the same pattern: worker households always have welfare gains; capitalist households have small loses if they are ot compensated with a temporary cut in the capital tax. Figures 5 and 6 depict allocations, tax rates, and prices for the policy reform 1 when the status quo policy is set to replicate the capital and labor taxes in Chari, Christiano, and Kehoe (1995), $\tau^k = .28$ and $\tau^n = .24$. The response in allocations is very similar with only small changes in magnitude. We also rerun policy reform 1 against a status quo policy of zero labor taxes and zero government spending. The capital tax is set to 36% and all its proceed rebated to the capitalist household. This extreme parametrization of policy is interesting because the intertemporal wedge is the sole distortion in steady state. Not surprisingly, we find large welfare gains for the worker household in this case.

We also explored different preference parametrizations without overturning the welfare results. Interestingly, the intertemporal discount rate can be really low (about $\beta = .5$) and yet the front-loading is welfare-improving for workers. Other parameters, within their usual range, did not affect the results.

Finally we evaluated different versions of the policy reforms discussed above. We found that there are welfare gains even for large policy reforms as long as the tax cuts are smoothly faded out. If the additional tax revenues are used to finance a large, but short-lived cut in the labor tax then the second order effects can dominate and lead to welfare losses for both households.

6 Conclusions

In this paper, we have emphasized that the ability to front-load distortions is critical for the Chamley-Judd result. The main advantage of our argument lays in its generality. The government can front-load all distortions in most Ramsey models. We are thus not surprised to learn that all intertemporal distortions under the optimal policy are temporary in these models. Our discussion can also provide guidance on how to pose normative questions in fiscal policy more generally. The recent research on optimal taxation with private information has emphasized that an approach based on arbitrary fiscal instruments may unknowingly leave the relevant trade-offs out.[22] And, indeed, second best allocations under private information generally feature permanent intertemporal distortions. The literature on private information, however, does not isolate the relevant trade-off missing in the Ramsey model. Our discussion makes it clear that the ability to front-load all distortions in the Ramsey model that generate the different policy prescription — as the distortion associated with private information cannot be fully front-loaded ahead of the realization of the uncertainty.

Our argument for the Chamley-Judd result also contains a lesson for policy design. We know that the Ricardian equivalence does not hold in Ramsey models as taxes are distortionary. Hence we are to expect that changing the timing of taxes has welfare effects. However we lack a clear view of the direction of these welfare changes. Our discussion of the Chamley-Judd result makes clear that postponing taxes is welfare-reducing as long as there are intertemporal distortions, either because of a positive capital tax or some other distortion on capital accumulation.

[22] See Werning (2007) for an extensive discussion on this point.

References

[1] Aiyagari, S Rao. 1994. Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics* 109(3): 659-84.

[2] Aiyagari, S. Rao. 1995. Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting. *Journal of Political Economy* 103(6): 1158-75,

[3] Aiyagari, S. Rao, Albert Marcet, Thomas J. Sargent, and Juha Seppala. 2002. Optimal Taxation without State-Contingent Debt. *Journal of Political Economy* 110 (6): 1220–1254.

[4] Atkeson, Andrew, V.V. Chari, and Patrick J. Kehoe. 1999. Taxing Capital Income: A Bad Idea. Federal Reserve Bank of Minneapolis Quarterly Review 23 (3): 3–17.

[5] Chamley, Christophe, 1986. Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives. *Econometrica* 54 (3): 607–622.

[6] Atkinson, Anthony B. and Joseph E. Stiglitz. 1980. Lectures on Public Economics. McGraw-Hill, New York.

[7] Chari, V.V., and Patrick J. Kehoe. 1999. Optimal fiscal and monetary policy. In J. B. Taylor, and M. Woodford (ed.), Handbook of Macroeconomics.

[8] Chari, V.V., Lawrence J.Christiano, and Patrick J. Kehoe. 1995. Policy Analysis in Business Cycle Models. In T.F. Cooley (ed.), Frontiers of Business Cycle Research.

[9] Conesa, Juan Carlos, Sagiri Kitao and Dirk Krueger. 2006. Taxing Capital? Not a Bad Idea After All! NBER WP 12880.

[10] Correia, Isabel H. 1996. Should Capital Income Be Taxed in the Steady State? *Journal of Public Economics* 60 (1): 147-151.

[11] Erosa, Andres and Martin Gervais. 2002. Optimal Taxation in Life-Cycle Economies. *Journal of Economic Theory* 105(2): 338-369.

[12] Domeij, David, and Heathcote, Jonathan. On the Distributional Effects of Reducing Capital Taxes. *International Economic Review*, 2004, 45(2), 523-554

[13] Farhi, Emmanuel. 2006. Capital Taxation and Ownership When Markets are Incomplete. Manuscript, MIT.

[14] Garriga, Carlos. 2003. Optimal Fiscal Policy in Overlapping Generation Models. Manuscript. Florida International University.

[15] Huggett, Mark. 1993. The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies. *Journal of Economic Dynamics and Control* 17: 953- 969.

[16] Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi. 1997. On the Optimal Taxation of Capital Income. *Journal of Economic Theory* 73(1): 93-117.

[17] Judd, Kenneth. 1985. Redistributive Taxation in a Perfect Foresight Model. *Journal of Public Economics* 28, 59-84.

[18] Lansing, Kevin J. 1999. Optimal Redistributive Capital Taxation in a Neoclassical Growth Model. *Journal of Public Economics* 73, 423-453.

[19] Lucas, Robert E., Jr., and Nancy L. Stokey. 1983. Optimal Fiscal and Monetary Policy in an Economy without Capital. *Journal of Monetary Economics* 12: 55–93.

[20] Mirrlees, James. 1971. An exploration in the theory of optimum income taxation. *The Review of Economic Studies* 38: 175-208.

[21] Reis, Catarina. 2006. Taxation without Commitment. Manuscript, MIT.

[22] Rogerson, William. 1985. Repeated moral hazard. *Econometrica* 53:69-76.

[23] Zhu, Xiaodong. 1992. Optimal Fiscal Policy in a Stochastic Growth Model. *Journal of Economic Theory*. 58: 250–89.

7 Figures

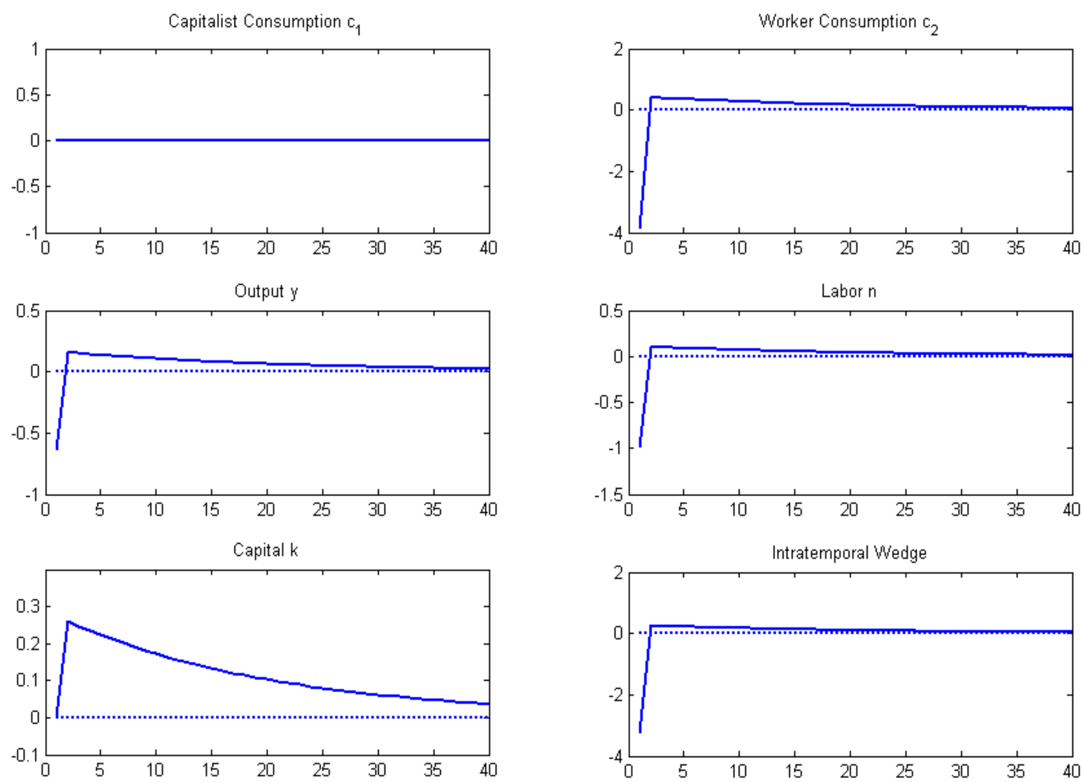

Figure 1: **Policy Experiment 1: Allocations.** Baseline economy. All variables expressed in percentage deviations from the steady state values.

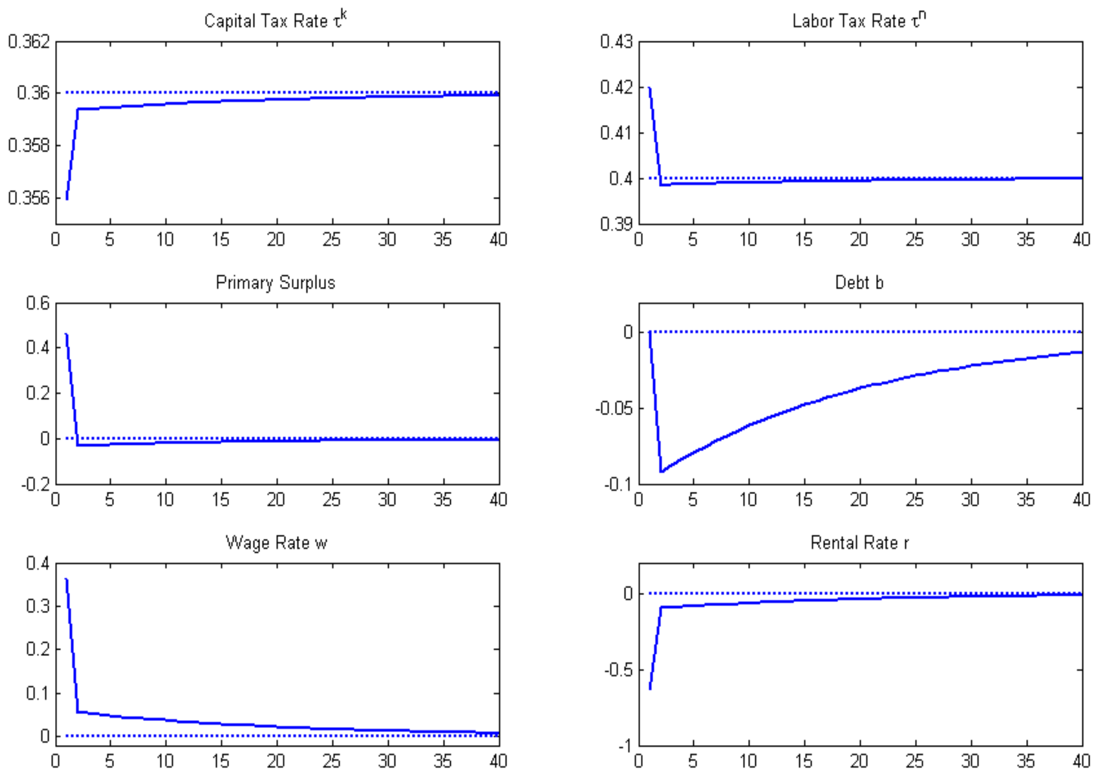

Figure 2: **Policy Experiment 1: Policy and Prices.** Baseline economy. Capital and labor taxes are in rates. All remaining variables expressed in percentage deviations from the steady state values.

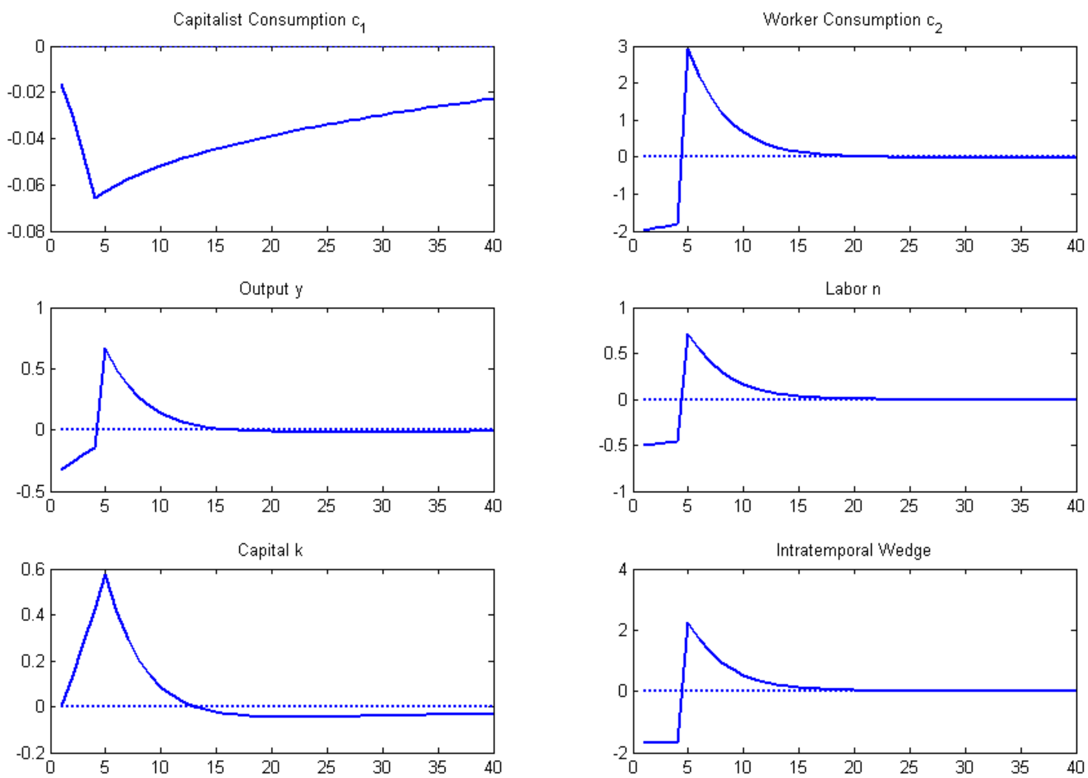

Figure 3: **Policy Experiment 2: Allocations.** Baseline economy. All variables expressed in percentage deviations from the steady state values.

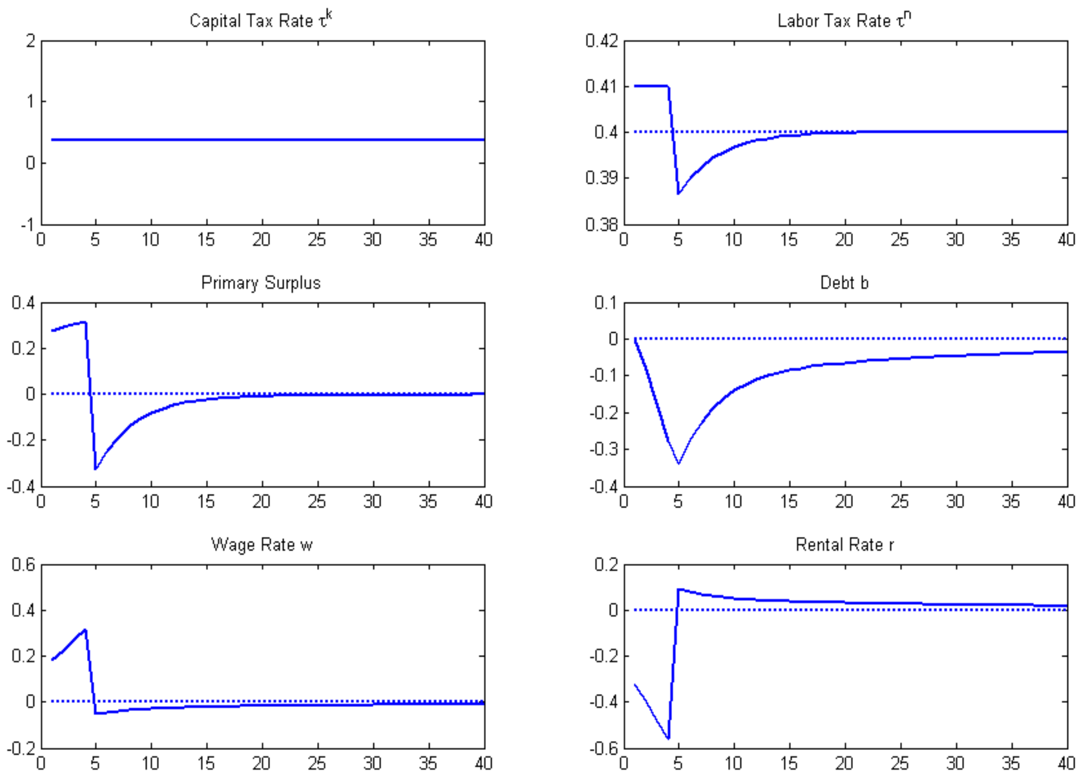

Figure 4: **Policy Experiment 2: Policy and Prices.** Baseline economy. Capital and labor taxes are in rates. All remaining variables expressed in percentage deviations from the steady state values.

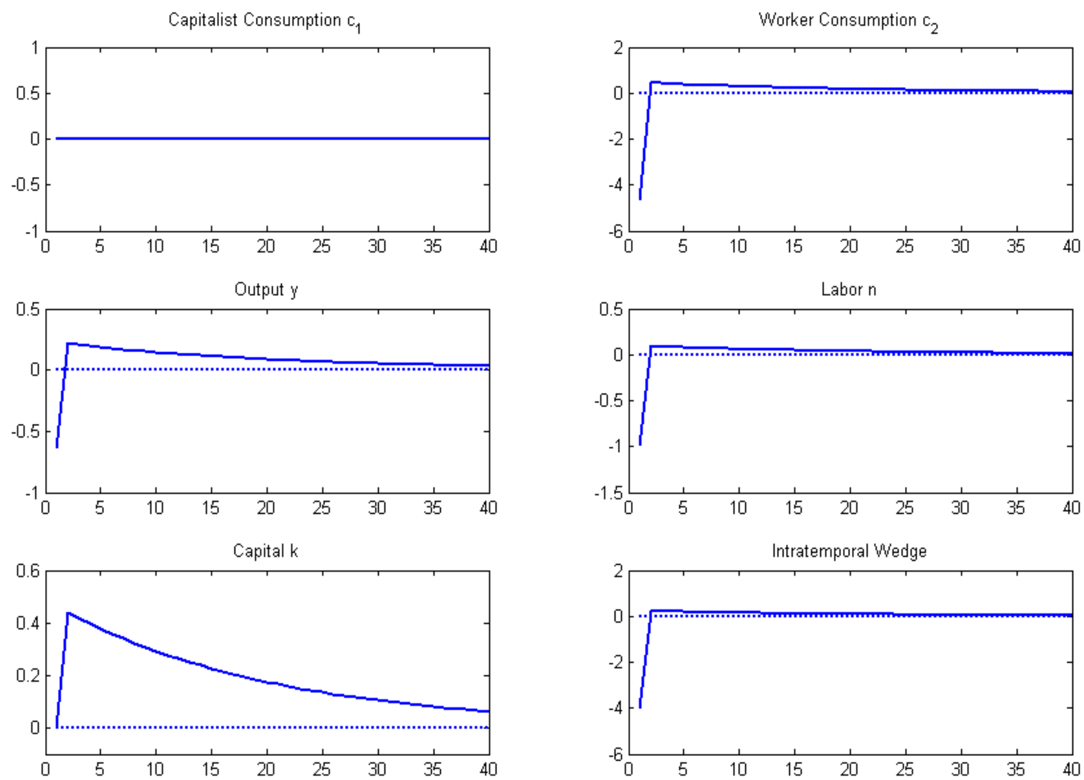

Figure 5: **Policy Experiment 1: Allocations.** Low-taxation economy. All variables expressed in percentage deviations from the steady state values.

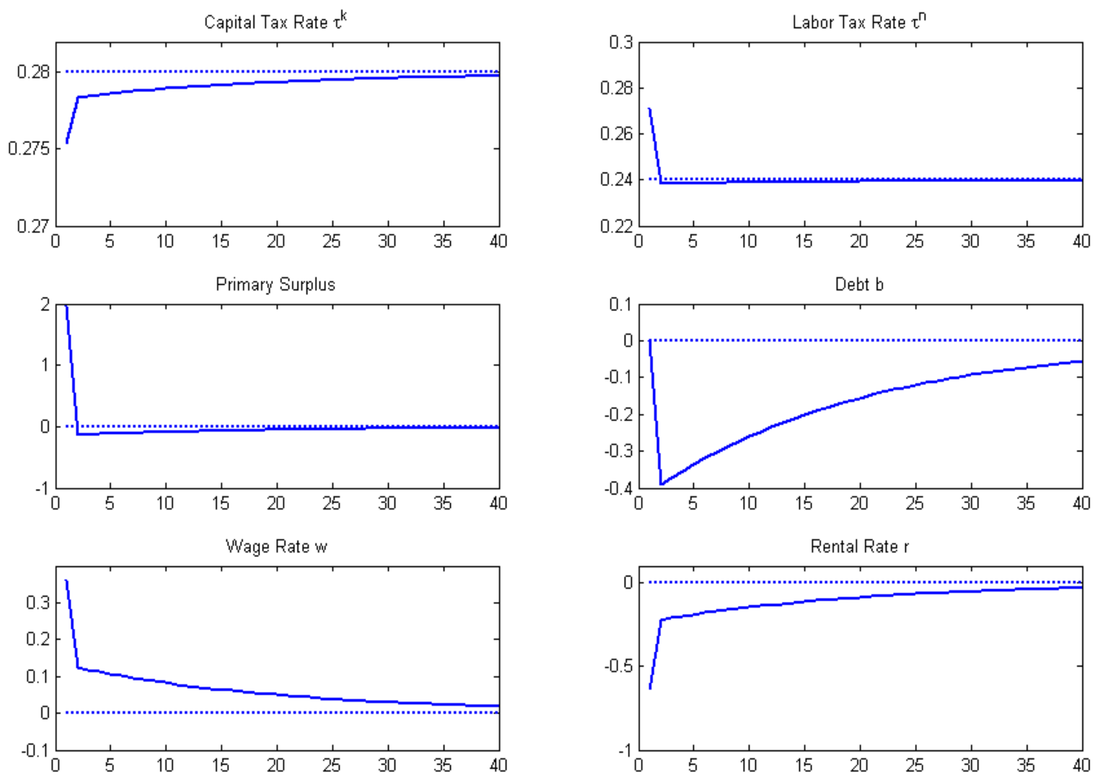

Figure 6: **Policy Experiment 1: Policy and Prices.** Low-taxation economy. Capital and labor taxes are in rates. All remaining variables expressed in percentage deviations from the steady state values.