ON THE OPTIMAL TIMING OF CAPITAL TAXES

John Hassler, Per Krusell, Kjetil Storesletten and Fabrizio Zilibotti

November 26, 2004
Summary

- Introduction
- Our modeling strategy
- The standard model
- Generalizing and specializing
- Analysis
- Results
- Interpretations
- Lack of commitment
- Calibration
- Stochastics
- Conclusions
• What is the desired path of taxes to finance an exogenous stream of government expenditure?
Introduction

- What is the desired path of taxes to finance an exogenous stream of government expenditure?

- In this paper we argue that in very natural settings:
  - optimal taxes involve fluctuations in rates and investments,
  - these optimal fluctuations dampens or disappear when future taxes cannot be committed to.

Explanations:
- constant taxes give smooth static distortions e.g., labor supply, but not necessarily smooth dynamic distortions,
- front-loading is a knife-edge result hinging on depreciation being exactly geometric.
Introduction

- What is the desired path of taxes to finance an exogenous stream of government expenditure?

- In this paper we argue that in very natural settings:
  - optimal taxes involve fluctuations in rates and investments,
  - these optimal fluctuations dampens or disappear when future taxes cannot be committed to.

- Extends previous results in the literature:
  - taxes should be constant to smooth distortions (Barro -79),
  - taxes should be front-loaded, high taxes in the first period (only) to take advantage of pre-installed capital that provides an inelastic tax-base (Chamley -86, Judd -85),
  - provides analytical characterization of tax-dynamics while (most) previous results regard steady state.
Introduction

- What is the desired path of taxes to finance an exogenous stream of government expenditure?

- In this paper we argue that in very natural settings:
  - optimal taxes involve fluctuations in rates and investments,
  - these optimal fluctuations dampens or disappear when future taxes cannot be committed to.

- Extends previous results in the literature:
  - taxes should be constant to smooth distortions (Barro -79),
  - taxes should be front-loaded, high taxes in the first period (only) to take advantage of pre-installed capital that provides an inelastic tax-base (Chamley -86, Judd -85),
  - provides analytical characterization of tax-dynamics while (most) previous results regard steady state.

- Explanations:
  - constant taxes give smooth static distortions e.g., labor supply, but not necessarily smooth dynamic distortions,
  - front-loading is a knife-edge result hinging on depreciation being exactly geometric.
Our modeling strategy

• Start with a very standard setup;
  • A benevolent government with commitment choose taxes to finance an exogenous stream of expenditures.
  • Neo-classical production function in labor and capital.
  • Standard time-additive preferences over leisure and consumption.

• We modify the model in some dimensions.
  We generalize in two:
  • We allow “quasi-geometric” depreciation. One depreciation rate the first period, another thereafter.
  • We allow adjustment costs, consumption and investments are not perfect substitutes.

• We specialize production and consumption to be linear. This provides analytical tractability and allows a complete characterization.
Our modeling strategy

- Start with a very standard setup;
  - A benevolent government with commitment choose taxes to finance an exogenous stream of expenditures.
  - Neo-classical production function in labor and capital.
  - Standard time-additive preferences over leisure and consumption.

- Then we modify the model in some dimensions.
  - We generalize in two: we allow "quasi-geometric" depreciation. One depreciation rate the first period, another thereafter.
  - We allow adjustment costs, consumption are not perfect substitutes.
  - We specialize production and consumption to be linear. This provides analytical tractability and allows a complete characterization.
Our modeling strategy

- Start with a very standard setup;
  - A benevolent government with commitment choose taxes to finance an exogenous stream of expenditures.
  - Neo-classical production function in labor and capital.
  - Standard time-additive preferences over leisure and consumption.

- Then we modify the model in some dimensions.

- We generalize in two:
  - we allow "quasi-geometric" depreciation. One depreciation rate the first period, another thereafter.
  - we allow adjustment costs, consumption and investments are not perfect substitutes.
Our modeling strategy

- Start with a very standard setup;
  - A benevolent government with commitment choose taxes to finance an exogenous stream of expenditures.
  - Neo-classical production function in labor and capital.
  - Standard time-additive preferences over leisure and consumption.

- Then we modify the model in some dimensions.

- We generalize in two:
  - we allow ”quasi-geometric” depreciation. One depreciation rate the first period, another thereafter.
  - we allow adjustment costs, consumption and investments are not perfect substitutes.

- We specialize production and consumption to be linear. This provides analytical tractability and allows a complete characterization.
The standard model

- Time is discrete and infinite and there is an infinitely lived household with standard time-additive preferences: \( \sum_{t=0}^{\infty} \beta^t u(c_t, 1 - n_t) \)
The standard model

- Time is discrete and infinite and there is an infinitely lived household with standard time-additive preferences: \( \sum_{t=0}^{\infty} \beta^t u(c_t, 1 - n_t) \)

- Production is of the standard, neoclassical variety: \( c_t + i_t = F(k_t, n_t) \) where \( F \) has constant returns to scale.
The standard model

- Time is discrete and infinite and there is an infinitely lived household with standard time-additive preferences: \( \sum_{t=0}^{\infty} \beta^t u(c_t, 1 - n_t) \)

- Production is of the standard, neoclassical variety: \( c_t + i_t = F(k_t, n_t) \) where \( F \) has constant returns to scale.

- Factor markets are perfect.
The standard model

- Time is discrete and infinite and there is an infinitely lived household with standard time-additive preferences: $\sum_{t=0}^{\infty} \beta^t u(c_t, 1 - n_t)$

- Production is of the standard, neoclassical variety: $c_t + i_t = F(k_t, n_t)$ where $F$ has constant returns to scale.

- Factor markets are perfect.

- The stock of capital accumulates according to $k_{t+1} = i_t + (1 - \delta)k_t$.
The standard model

- Time is discrete and infinite and there is an infinitely lived household with standard time-additive preferences: \( \sum_{t=0}^{\infty} \beta^t u(c_t, 1-n_t) \)

- Production is of the standard, neoclassical variety: \( c_t + i_t = F(k_t, n_t) \) where \( F \) has constant returns to scale.

- Factor markets are perfect.

- The stock of capital accumulates according to \( k_{t+1} = i_t + (1-\delta)k_t \).

- The government has to finance a given sequence \( \{g_t\}_{t=0}^{\infty} \) of government expenditures. Or, alternatively, there is a public good with a constant marginal utility, that only the government can provide.
The standard model

- Time is discrete and infinite and there is an infinitely lived household with standard time-additive preferences: \( \sum_{t=0}^{\infty} \beta^t u(c_t, 1 - n_t) \)

- Production is of the standard, neoclassical variety: \( c_t + i_t = F(k_t, n_t) \) where \( F \) has constant returns to scale.

- Factor markets are perfect.

- The stock of capital accumulates according to \( k_{t+1} = i_t + (1 - \delta) k_t \).

- The government has to finance a given sequence \( \{g_t\}_{t=0}^{\infty} \) of government expenditures. Or, alternatively, there is a public good with a constant marginal utility, that only the government can provide.

- The Ramsey Problem is to maximize welfare, subject to the government’s intertemporal budget constraint and the IC constraints (implies that allocation can be decentralized)

\[
\begin{align*}
    u_{i,t} &= u_{c,t} \cdot w_t (1 - \tau_{nt}), \\
    u_{c,t} &= \beta u_{c,t+1} \cdot (r_t (1 - \tau_{kt}) + 1 - \delta),
\end{align*}
\]
The standard model

- Time is discrete and infinite and there is an infinitely lived household with standard time-additive preferences: $\sum_{t=0}^{\infty} \beta^t u(c_t, 1-n_t)$

- Production is of the standard, neoclassical variety: $c_t + i_t = F(k_t, n_t)$ where $F$ has constant returns to scale.

- Factor markets are perfect.

- The stock of capital accumulates according to $k_{t+1} = i_t + (1-\delta)k_t$.

- The government has to finance a given sequence $\{g_t\}_{t=0}^{\infty}$ of government expenditures. Or, alternatively, there is a public good with a constant marginal utility, that only the government can provide.

- The Ramsey Problem is to maximize welfare, subject to the government’s intertemporal budget constraint and the IC constraints (implies that allocation can be decentralized)

  $$u_{i,t} = u_{c,t} \cdot w_t(1-\tau_{nt}),$$
  $$u_{c,t} = \beta u_{c,t+1} \cdot (r_t(1-\tau_{kt}) + 1-\delta),$$

- Capital income taxes are determined one period ahead and are not allowed to be vintage specific.
Generalizing and specializing (1/5)

Generalizations

- Substantial empirical evidence that capital does not depreciate geometrically:
  - Oliner (1996) finds that economic depreciation for machine tools is significantly increasing with age.
  - Some contrary evidence exist for resale value.
  - Human capital does not live for ever.
Generalizing and specializing (2/5)

Generalizations

- We assume *quasi-geometric* depreciation. A unit of investment at time $t$ leads to one unit of productive capital in period $t+1$, $1 - \rho \delta$ units in period $t+2$, and $(1 - \rho \delta)(1 - \delta)^s$ units in period $t + 2 + s$.

- We note that now old and new capital ($k^o_t$ and $k^n_t$) are not identical. Specifically,

$$
\begin{align*}
  k_{t+1}^o &= (1 - \delta)k_t^o + (1 - \rho \delta)k_t^n, \\
  k_{t+1}^n &= i_t.
\end{align*}
$$

Graph

- $\rho \in (0, 1), \delta \in (0, 1)$ is an accelerated depreciation structure, first a depreciation rate of $\rho \delta$, then $\delta$ and with $\rho = 1$, we have geometric depreciation,

- $\rho = 0, \delta = 1$ corresponds to a 2-period OLG structure.

- Like in standard model, we maintain that taxes cannot be vintage specific.
Generalizations

- We also assume that investments and consumption goods are imperfect substitutes, the resource constraint being,

\[ c_t + G(i_t) = F(k_t, n_t) \]

where \( G \) is convex.

- The convex \( G \) can be interpreted as
  - a concavity in production of an investment good,
  - literally as an installation costs, or,
  - as an effort cost of producing human capital.

- To maintain a non-trivial analysis, we assume that investment cannot be taxes/subsidized directly. Only factor income is taxed.
Replicating some old results

Proposition 1 and 2

- Assume that $\rho = 1$ and that $G(i)$ is linear. Then if the Ramsey problem above has a solution where the tax rates converge to a limit, it has to have $\tau_{kt} \to 0$. Moreover, in the special case where $u(c, n) = \frac{c^{1-\sigma} - 1}{1-\sigma} + v(n)$, the Ramsey solution has $\tau_{kt} = 0$ for all $t \geq 2$.


- Assume that $\rho = 1$ and that $G(i)$ is strictly convex. Then the Ramsey problem does not have a solution in which $\tau_{kt} \to 0$.

  This results is closely related to in Correia (1996).
Specializing the setup

We assume:

- Utility is linear.
- Output is linear in capital.
- The adjustment cost is quadratic.
- We can then analytically characterize the optimal path of taxes and investments.
We show that the Ramsey problem can be written:

$$\max_{\{T_0, T\}_{t=0}^{\infty}} (A - 1)\hat{T}_0 k_1^0 + \sum_{t=0}^{\infty} \beta^t y(T_t).$$

where $y(T_t) \equiv A\kappa i(T_t) - i(T_t)^2 (2A - 1)$ and $i(T_t) = \frac{1}{2} (\kappa - T_t)$.

- $A$ can be interpreted either as the Lagrange multiplier on the budget constraint of the government or as the marginal value of public good provision.

- $T_t \equiv \beta \tau_{t+1} + (1 - \rho\delta) \sum_{s=2}^{\infty} \beta^s (1 - \delta)^{s-2} \tau_{t+s}$, denotes the effective discounted sum of taxes falling on investments in period $t$ and $\hat{T}_0$ is the corresponding discounted sum of taxes levied on pre-installed capital $k_1^0$.

- $i(T_t)$ denotes privately optimal investments and results by solving the Euler equation forward.

- The function $y(T_t)$ is the present value of the contribution to the planner’s utility of the investment vintage installed at $t$. Each vintage contributes to the planner’s utility via private consumption ($i_t (\kappa - T_t)$), by paying taxes ($AT_t i_t$), and the investment cost ($-i_t^2$).
Results

- Under geometric depreciation ($\rho = 1$), tax rates are constant after one period: 
  \[ \tau_t = \tau^* \equiv \frac{(A - 1)}{(2A - 1)} < \frac{1}{2} \] from $t = 2$ and onwards and \[ \tau_1 = \tau^* \left( 1 + \frac{2k_1}{\beta} \right). \]

- Period 0 investments are more distorted than later distortions. Sketch of proof

---

Graph, Sketch of proof
• Under geometric depreciation ($\rho = 1$), tax rates are constant after one period: $\tau_t = \tau^* \equiv (A - 1) / (2A - 1) < \frac{1}{2}$ from $t = 2$ and onwards and $\tau_1 = \tau^* \left(1 + \frac{2k_1^o}{\beta}\right)$.

• Period 0 investments are more distorted than later distortions. **Sketch of proof**

• In the general case with quasi-geometric depreciation, the optimal sequence of tax rates is:

$$\tau_{t+1} = \tau^* - \delta (1 - \rho) (\tau_t - \tau^*) \text{ for } t \geq 1,$$

$$\tau_1 = \tau^* \left(1 + 2k_1^o \frac{1 + \beta \delta (1 - \delta) (1 - \rho)}{\beta (1 - \beta \delta^2 (1 - \rho)^2)}\right).$$

• **Graph, Sketch of proof**
Interpretations

- First, we note that with cohort-specific taxes, the planner would set 100% in each period on preinstalled capital and a constant tax equal to $\tau^*$ on all investments.
Interpretations

• First, we note that with cohort-specific taxes, the planner would set 100% in each period on preinstalled capital and a constant tax equal to $\tau^*$ on all investments.

• When $\rho=1$, future tax paths cannot be used to substitute for the non-existence of cohort-specific taxes. $\hat{T}_0$ is identical to $T_0$. The trade-off between distortions and taxing pre-installed capital is worse for taxes later than $\tau_1$. 
Interpretations

• First, we note that with cohort-specific taxes, the planner would set 100% in each period on preinstalled capital and a constant tax equal to $\tau^*$ on all investments.

• When $\rho=1$, future tax paths cannot be used to substitute for the non-existence of cohort-specific taxes. $\hat{T}_0$ is identical to $T_0$. The trade-off between distortions and taxing pre-installed capital is worse for taxes later than $\tau_1$.

• When $\rho < 1$ oscillating taxes reduces distortions on investments while maintaining taxation on pre-installed capital. $\hat{T}_0$ is not identical to $T_0$. Oscillations acts as a substitute for cohort-specific taxes.
Interpretations

• First, we note that with cohort-specific taxes, the planner would set 100% in each period on preinstalled capital and a constant tax equal to $\tau^*$ on all investments.

• When $\rho=1$, future tax paths cannot be used to substitute for the non-existence of cohort-specific taxes. $\hat{T}_0$ is identical to $T_0$. The trade-off between distortions and taxing pre-installed capital is worse for taxes later than $\tau_1$.

• When $\rho < 1$ oscillating taxes reduces distortions on investments while maintaining taxation on pre-installed capital. $\hat{T}_0$ is not identical to $T_0$. Oscillations acts as a substitute for cohort-specific taxes.

• Specifically, by promising low taxes in period 2, investments in period 0 can be stimulated while the tax-revenue from the inelastic tax-base $k^0_1$ is maintained.
Interpretations

• First, we note that with cohort-specific taxes, the planner would set 100% in each period on preinstalled capital and a constant tax equal to $\tau^*$ on all investments.

• When $\rho = 1$, future tax paths cannot be used to substitute for the non-existence of cohort-specific taxes. $\hat{T}_0$ is identical to $T_0$. The trade-off between distortions and taxing pre-installed capital is worse for taxes later than $\tau_1$.

• When $\rho < 1$ oscillating taxes reduces distortions on investments while maintaining taxation on pre-installed capital. $\hat{T}_0$ is not identical to $T_0$. Oscillations acts as a substitute for cohort-specific taxes.

• Specifically, by promising low taxes in period 2, investments in period 0 can be stimulated while the tax-revenue from the inelastic tax-base $k^0_1$ is maintained.

• Then, it becomes more favorable to invest in period 1, which leads to an optimal increase in $\tau_3$ relative to the steady state $\Rightarrow$ oscillations.
The tax-system must balance the marginal distortionary cost of raising revenues by any $\tau_s$ (denoted $D_s$) to the excess value of the marginal revenue $R_s$ of the tax. $D_s = (A - 1) R_s$. 
The tax-system must balance the marginal distortionary cost of raising revenues by any $\tau_s$ (denoted $D_s$) to the excess value of the marginal revenue $R_s$ of the tax. $D_s = (A - 1) R_s$.

Any tax $\tau_t$ affects all previous investments. The marginal distortionary cost of $\tau_t$ therefore depends on all previous investment wedges, $\kappa/2 - i_s$, and accumulates over time.
Interpretations - Distortion smoothing

- The tax-system must balance the marginal distortionary cost of raising revenues by any $\tau_s$ (denoted $D_s$) to the excess value of the marginal revenue $R_s$ of the tax. $D_s = (A - 1)R_s$.

- Any tax $\tau_l$ affects all previous investments. The marginal distortionary cost of $\tau_l$ therefore depends on all previous investment wedges, $\kappa/2 - i_s$, and accumulates over time.

- Consider an increase in $k^0_1$ from e.g., zero. Then, the marginal revenue of $\tau_1$ increases so that tax should go up. The increase in $k^0_1$ increases the marginal revenue also of $\tau_2$. However, the increase in $\tau_1$ increase also the marginal distortionary cost of $\tau_2$ since it affects also period zero investments.
The tax-system must balance the marginal distortionary cost of raising revenues by any $\tau_s$ (denoted $D_s$) to the excess value of the marginal revenue $R_s$ of the tax. $D_s = (A - 1)R_s$.

Any tax $\tau_t$ affects all previous investments. The marginal distortionary cost of $\tau_t$ therefore depends on all previous investment wedges, $\kappa/2 - i_s$, and accumulates over time.

Consider an increase in $k_1^0$ from e.g., zero. Then, the marginal revenue of $\tau_1$ increases so that tax should go up. The increase in $k_1^0$ increases the marginal revenue also of $\tau_2$. However, the increase in $\tau_1$ increase also the marginal distortionary cost of $\tau_2$ since it affects also period zero investments.

The higher marginal cost of taxes in period 2 is:

- exactly as large the increase in the marginal revenue when $\rho = 1$. $\Rightarrow$ $\tau_2$ should not change and remain at $\tau^*$.
- larger than the increase in the marginal revenue when $\rho < 1$ (new capital depreciates less). $\Rightarrow$ $\tau_2$ should be reduced relative to $\tau^*$, producing oscillatory dynamics.
- smaller than the increase in the marginal revenue when $\rho > 1$ (new capital depreciates more). $\Rightarrow$ $\tau_2$ should also be increased relative to $\tau^*$, producing monotone dynamics. **Formal analysis**
Lack of commitment

- Suppose now that the benevolent planner can set taxes only for the next period. This case can alternatively be interpreted as a political economy outcome with voters caring altruistically about future cohorts.
Lack of commitment

- Suppose now that the benevolent planner can set taxes only for the next period. This case can alternatively be interpreted as a political economy outcome with voters caring altruistically about future cohorts.

- Note that in this case, financing of a given stream of expenditures and providing a public good, are not equivalent. In the former case, the debt, or alternatively, the shadow value of public funds become an additional state variable. We choose the public good formulation for transparency.
Lack of commitment

- Suppose now that the benevolent planner can set taxes only for the next period. This case can alternatively be interpreted as a political economy outcome with voters caring altruistically about future cohorts.

- Note that in this case, financing of a given stream of expenditures and providing a public good, are not equivalent. In the former case, the debt, or alternatively, the shadow value of public funds become an additional state variable. We choose the public good formulation for transparency.

- We solve for the Markov equilibrium, defined as:
  1. A time invariant policy rule, $\tau_t = T(k^o_t)$ determining taxes as a function of the state variable $k^o_t$ that maximizes the government's Bellman equation under the expectation that the same tax function is used in the future.

  2. An investment rule such that investments are done privately optimally given the expectations that future taxes are set according to $\tau_{t+1} = T(k^o_{t+1})$. 
Lack of commitment - results

The results are:

- Steady state taxes are higher, they might even be above the maximum of the dynamic Laffer curve.
The results are:

- Steady state taxes are higher, they might even be above the maximum of the dynamic Laffer curve.

- Dynamics are monotone or show more dampened oscillations than under commitment. This is easy to understand. Under commitment, the planner promises low taxes in period 2, but *ex post*, the planner wants to renege on this promise.
Lack of commitment - results

The results are:

- Steady state taxes are higher, they might even be above the maximum of the dynamic Laffer curve.

- Dynamics are monotone or show more dampened oscillations than under commitment. This is easy to understand. Under commitment, the planner promises low taxes in period 2, but ex post, the planner wants to renege on this promise.

- This is opposite of political business cycles.
Let us now parametrize/calibrate to get a feel for quantities. We set;

- the period length to 4 years (electional cycle).
- the long-run tax rate under no commitment to 50% (as in Klein and Ríos-Rull (2003)),
- long run depreciation rates is 10% per year,
- $\rho = 0.5$, half as fast depreciation during the first (4 year) period,
- annual interest rates to 5%.
Quantitative effects

The findings are as follows.

1. In the economy with commitment, the steady-state tax rate is \( \tau^* = 6\% \), i.e., significantly below the 50\% rate under lack of commitment.

2. In the economy with commitment, the AR1 coefficient of taxes is 0.21 (on a 4-year basis): taxes oscillate.

3. In the economy without commitment, the AR1 coefficient of taxes is 0.4: taxes are highly persistent!
So far, oscillations are due to initial conditions, if $k_1^0 = 0$ no oscillations get going. What about variations in spending requirements or marginal public good valuation?
• So far, oscillations are due to initial conditions, if $k_1^o = 0$ no oscillations get going. What about variations in spending requirements or marginal public good valuation?

• The general result is that oscillations occur if markets are incomplete. For example;
• So far, oscillations are due to initial conditions, if $k_1^o = 0$ no oscillations get going. What about variations in spending requirements or marginal public good valuation?

• The general result is that oscillations occur if markets are incomplete. For example;
  1. A perfectly anticipated increase in spending requirements or marginal public good valuation leads to oscillations if there is no perfect market for safe bonds.
So far, oscillations are due to initial conditions, if $k_1^0 = 0$ no oscillations get going. What about variations in spending requirements or marginal public good valuation?

The general result is that oscillations occur if markets are incomplete. For example;

1. A perfectly anticipated increase in spending requirements or marginal public good valuation leads to oscillations if there is no perfect market for safe bonds.

2. Stochastic shocks to spending requirements or marginal public good valuation leads to oscillations if there is no perfect state contingent market. For oscillations not to occur there has to be contingent claims that span all possible realizations of the spending requirement.
Conclusions

We have constructed a model that generalizes some important previous results, showing that
Conclusions

We have constructed a model that generalizes some important previous results, showing that

- taxes should oscillate if depreciation is accelerating, immediate convergence is a knife-edge property of geometric discounting,
Conclusions

We have constructed a model that generalizes some important previous results, showing that

- taxes should oscillate if depreciation is accelerating, immediate convergence is a knife-edge property of geometric discounting,

- lack of commitment dampens the oscillations,
Conclusions

We have constructed a model that generalizes some important previous results, showing that

- taxes should oscillate if depreciation is accelerating, immediate convergence is a knife-edge property of geometric discounting,
- lack of commitment dampens the oscillations,
- in realistic settings, the difference can be very large and the degree of persistence can be used to get a measure of the degree of commitment in politics.

Numerical models with more realistic production and preferences should be used for future quantitative work.
We have constructed a model that generalizes some important previous results, showing that

- taxes should oscillate if depreciation is accelerating, immediate convergence is a knife-edge property of geometric discounting,
- lack of commitment dampens the oscillations,
- in realistic settings, the difference can be very large and the degree of persistence can be used to get a measure of the degree of commitment in politics.
- Numerical models with more realistic production and preferences should be used for future quantitative work.
END
Accelerating depreciation, \( \rho \in (0, 1), \delta \in (0, 1) \)
Ramsey Solution

![Graphs showing time series of tax rate, Investment, Output, and T over time.](image)
The Markov tax sequences (dashed lines) start below the Ramsey tax sequences (continuous lines), are smoother, and converge to steady-states with higher taxation.
When $\rho = 1$, old capital in period 1 is identical to the capital installed in period zero. In particular, the total tax-burden from any sequence of taxes is identical for the two types of capital, i.e.,

$$\hat{T}_0 = T_0.$$ 

Thus, the Ramsey planner faces the following problem:

$$\max_{\{T_t\}_{t=1}^\infty} (A - 1)T_0k^o_1 + \sum_{t=0}^\infty \beta^t y(T_t).$$

First-order conditions are

$$y'(T^*_t) = 0, \forall t > 1.$$ 

and

$$(A - 1)k^o_1 + y'(T^*_0) = 0,$$

Backing out tax rates, this implies that tax rates are also constant after one period: $\tau_t = \tau^* \equiv (A - 1)/(2A - 1) < \frac{1}{2}$ from $t = 2$ and onwards and $\tau_t = \tau^* \left(1 + \frac{2k^o_1}{\beta}\right)$. 
Quasi-geometric depreciation

In the general case with quasi-geometric depreciation, $\hat{T}_0$ is no longer equal to $T_0$. In particular, increasing $\tau_1$ and lowering $\tau_2$ hits the old installed capital more than period 0 investments. $\hat{T}_0$ becomes larger than $T_0$. However, $\hat{T}_0$ is not independent of the $T_t$'s,

$$\hat{T}_0 = \sum_{t=0}^{\infty} (-\delta \beta (1 - \rho))^t T_t.$$

The Ramsey problem now reads

$$\max_{\{T_t\}_{t=1}^{\infty}} (A - 1) \left( \sum_{t=0}^{\infty} (-\delta \beta (1 - \rho))^t T_t \right) k_1^o + \sum_{t=0}^{\infty} \beta^t y(T_t),$$

and the first-order condition for the $T_t$'s is $(A - 1) k_1^o (-\delta (1 - \rho))^t + y'(T_t) = 0$, which uniquely pins down the $T_t$'s, in turn pinning down the unique feasible sequence of tax rates to:

$$\tau_{t+1} = \tau^* - \delta (1 - \rho) (\tau_t - \tau^*) \text{ for } t \geq 1,$$

$$\tau_1 = \tau^* \left( 1 + 2k_1^o \frac{1 + \beta \delta (1 - \delta) (1 - \rho)}{\beta (1 - \beta \delta^2 (1 - \rho)^2)} \right),$$

where $\tau^* \equiv (A - 1) / (2A - 1) < \frac{1}{2}$. 
The marginal distortionary cost of $\tau_s$, denoted $D_s$ follows
\begin{align*}
D_1 &= \frac{\kappa}{2} - i_0 \\
D_s &= (1 - \delta) D_{s-1} + \delta (1 - \rho) \left( \frac{\kappa}{2} - i_{s-2} \right) + \frac{\kappa}{2} - i_{s-1}.
\end{align*}

Similarly, since the tax base of $\tau_s$ is all the investments done before $s$ that remains at $s$, also the marginal revenue $R_s$ accumulates over time. Specifically,
\begin{align*}
R_1 &= k_1^o + 2i_0 - \frac{\kappa}{2} \\
R_s &= (1 - \delta) R_{s-1} + \delta (1 - \rho) \left( 2i_{s-2} - \frac{\kappa}{2} \right) + 2i_{s-1} - \frac{\kappa}{2} \forall s > 1.
\end{align*}

The condition $D_s = (A - 1) R_s$ can now be written
\begin{align*}
(1 - \delta) D_{s-1} + \delta (1 - \rho) \left( \frac{\kappa}{2} - i_{s-2} \right) + \frac{\kappa}{2} - i_{s-1} \\
= (A - 1) (1 - \delta) R_{s-1} + (A - 1) \delta (1 - \rho) \left( 2i_{s-2} - \frac{\kappa}{2} \right) + (A - 1) \left( 2i_{s-1} - \frac{\kappa}{2} \right).
\end{align*}

Now, let us increase $k_1^o$ from 0. First, we note that the first order condition for $\tau_1$, i.e., $D_1 = (A - 1) R_s$ yields $\frac{\kappa}{2} - i_0 = (A - 1) \left( k_1^o + 2i_0 - \frac{\kappa}{2} \right)$. So higher $k_1^o$ leads to higher marginal revenue of $\tau_1$ and thus increased $\tau_1$ and lower investments $i_0$. 

Distortion smoothing (1/2)
• The higher \( k_1^o \) increases marginal revenues also of \( \tau_2 \). Similarly, the higher investment wedge in period 0 increases the marginal distortionary cost also of taxes in period 2.

• The higher marginal cost of taxes in period 2 is:
  - exactly as large the increase in the marginal revenue when \( \rho = 1 \). \( \Rightarrow \) \( \tau_2 \) should not change and remain at \( \tau^* \).
  - larger than the increase in the marginal revenue when \( \rho < 1 \) (new capital depreciates less). \( \Rightarrow \) \( \tau_2 \) should be reduced relative to \( \tau^* \), producing oscillatory dynamics.
  - smaller than the increase in the marginal revenue when \( \rho > 1 \) (new capital depreciates more). \( \Rightarrow \) \( \tau_2 \) should also be increased relative to \( \tau^* \), producing monotone dynamics.
ON THE OPTIMAL TIMING OF CAPITAL TAXES

John Hassler, Per Krusell, Kjetil Storesletten and Fabrizio Zilibotti

November 26, 2004