Taxation, Debt and Redistribution

Anmol Bhandari, David Evans, Mikhail Golosov, Thomas J. Sargent

September 12, 2013

What do we do ?

We study optimal taxation under commitment with

- **Heterogeneous agents**
  $\rightarrow$ Agents have different productivities
- **Incomplete markets**
  $\rightarrow$ All agents trade a risk-free bond
- **Affine taxes**
  $\rightarrow$ Government can levy a proportional tax on labor earnings + lumpsump (tax or transfer)
- **Aggregate shocks**
  $\rightarrow$ Shocks to productivities, government expenditure etc.

What are we after ?

1. How costly are debt levels ?
2. What are the long run properties of optimal allocations ?
3. How should policy respond to aggregate shocks with heterogeneous cross sectional implications ?

Representative agent with linear taxes

- Higher levels of debt are distortionary
- With incomplete markets the government accumulates assets overtime

Redistribution and optimal transfers

- Representative agent models impose restrictions on transfers
  - Implicit motives for redistribution : Poor people cant pay lumpsump taxes
  - These constraints either *almost* always bind (For eg: Lucas Stokey, AMSS) and are key for long run dynamics
- We begin with explicit redistribution motives but leave transfers to be determined optimally

The prescription for optimal tax-transfers are substantially different with explicit redistribution considerations

Key mechanisms

Two departures from representative agent models:

- **Unrestricted transfers**: Level of debt is not distortionary. What matters is how it is distributed across agents
- **Explicit redistribution motives**: Endogenous costs of fluctuating transfers. Taking away a unit of consumption good affects "rich" and "poor" people differently

One part of heterogeneity (productivities) is exogenous but heterogeneity in assets is endogenous. The optimal policy is both affected and in turn affects the net distribution of assets

- **Absence of agent specific transfers** : Engineer a negative cross correlations in net assets and earnings
- **Absence of state contingent securities**: Exploit the endogenous fluctuations in interest rates

Ingredients

- **Uncertainty** : Markov aggregate shocks s_t
- **Demography** : I types of infinitely lived agents (of mass π_i) and benevolent planner
- **Technology** : Output is linear in labor supply. Agents differ in productivity $\{\theta_i(s_t)\}_{i,t}$
- **Preferences** (Households)

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \bar{\beta}_t U^i \left(c_i(s^t), l_i(s^t) \right),$$

where $\bar{\beta}_t = \left[\Pi_{j=0}^{t-1} \beta(s_j) \right]$ (why ?)
- **Preferences** (Planner) : Given Pareto weights $\{\alpha_i\}$

$$\mathbb{E}_0 \sum_{i=1}^{I} \pi_i \alpha_i \sum_{t=0}^{\infty} \bar{\beta}_t U_t^i \left(c_{i,t}, l_{i,t} \right),$$

- **Asset markets** : All agents trade a risk free bond

Constraints

- **Affine Taxes** : Agent i's tax bill

$$-T_t + \tau_t \theta_{i,t} l_{i,t}$$

- **Budget constraints**
  - Agent i : $c_{i,t} + b_{i,t} = (1 - \tau_t)\,\theta_{i,t} l_{i,t} + R_{t-1} b_{i,t-1} + T_t$
  - Government: $g_t + B_t + T_t = \tau_t \sum_{i=1}^{I} \pi_i \theta_{i,t} l_{i,t} + R_{t-1} B_{t-1}$

- **Market Clearing**
  - Goods: $\sum_{i=1}^{I} \pi_i c_i(s^t) + g\,(s_t) = \sum_{i=1}^{I} \pi_i \theta_i\,(s_t)\, l_i(s^t)$
  - Assets: $\sum_{i=1}^{I} \pi_i b_{i,t} + B_t = 0$

- **Initial conditions**: Distribution of assets $\{b_{i,-1}\}_i$ and B_{-1}

Ramsey Problem

Definition

Competitive equilibrium: Given $\left(\{b_{i,-1}\}_i, B_{-1}\right)$ and $\{\tau_t, T_t\}_{t=0}^{\infty}$ all allocations are chosen optimally, markets clear

Definition

Optimal competitive equilibrium: Welfare-maximizing competitive equilibrium for a given $\left(\{b_{i,-1}\}_i, B_{-1}\right)$

Ricardian Equivalence

- *Result* : There is a **large set** of transfers and asset profiles that support the same competitive allocation
- *Logic* : Taking away a unit of every agent's assets and increasing a unit of transfer leaves budget sets unchanged

Theorem

Given $\left(\{b_{i,-1}\}_i, B_{-1}\right)$, *let* $\left\{\{c_{i,t}, l_{i,t}, b_{i,t}\}_i, B_t, R_t\right\}_t$ *and* $\{\tau_t, T_t\}_t$ *be a competitive equilibrium. For any bounded sequences* $\left\{\hat{b}_{i,t}\right\}_{i,t\geq -1}$ *that satisfy*

$$\hat{b}_{i,t} - \hat{b}_{1,t} = \tilde{b}_{i,t} \equiv b_{i,t} - b_{1,t} \text{ for all } t \geq -1, i \geq 2,$$

there exist sequences $\left\{\hat{T}_t\right\}_t$ *and* $\left\{\hat{B}_t\right\}_{t\geq -1}$ *that satisfy the market clearing such that* $\left\{\left\{c_{i,t}, l_{i,t}, \hat{b}_{i,t}\right\}_i, \hat{B}_t, R_t\right\}_t$ *and* $\left\{\tau_t, \hat{T}_t\right\}_t$ *constitute a competitive equilibrium given* $\left(\left\{\hat{b}_{i,-1}\right\}_i, \hat{B}_{-1}\right)$.

Ricardian Equivalence: Implications

- No precautionary motive : WLOG normalize government assets B_t can be set to zero
- Exogenous borrowing constraints are not restrictive

Theorem

For every competitive equilibrium (allocation and interest rate sequence) in an economy without exogenous borrowing constraints there is a government tax policy such the same allocation and interest is a part of a competitive equilibrium in an economy with exogenous borrowing constraints of the form $b_{i,t} > \underline{b}_i$

Thus Ricardian equivalence holds with distortionary taxes and borrowing limits

Optimal allocations : Primal Approach

We focus on interior equilibria. First-order necessary conditions for the consumer's problem are

1. Eliminate taxes: τ_t

$$(1-\tau_t)\,\theta_{i,t}U^i_{c,t} = -U^i_{l,t},$$

2. Eliminate prices: R_t

$$U^i_{c,t} = \beta_t R_t \mathbb{E}_t U^i_{c,t+1}.$$

3. Eliminate transfers: T_t

$$(c_{i,t}-c_{1,t})+\tilde{b}_{i,t} = -\frac{U^i_{l,t}}{U^i_{c,t}}l_{i,t}+\frac{U^1_{l,t}}{U^1_{c,t}}l_{1,t}+\frac{U^i_{c,t-1}}{\beta_{t-1}\mathbb{E}_{t-1}U^i_{c,t}}\tilde{b}_{i,t-1}\ \forall i \geq 2, t$$

This yields "implementability constraints"

Notation: $\tilde{b}_{i,t} = b_{i,t} - b_{1,t}$ or the "net assets" of Agent i

Optimal allocations : Sequential Problem

Denote $Z_t^i = U_{c,t}^i c_{i,t} + U_{l,t}^i l_{i,t} - \frac{U_{c,t}^i}{U_{c,t}^1}\left[U_{c,t}^1 c_{1,t} + U_{l,t}^1 l_{1,t}\right]$. The optimal policy solves,

$$\max_{c_{i,t}, l_{i,t}, \tilde{b}_{i,t}} \mathbb{E}_0 \sum_{i=1}^{I} \pi_i \alpha_i \sum_{t=0}^{\infty} \bar{\beta}_t U_t^i \left(c_{i,t}, l_{i,t}\right),$$

subject to

$$\tilde{b}_{t-1} \frac{U_{c,t-1}^i}{\beta_{t-1}} = \left(\frac{\mathbb{E}_{t-1} U_{c,t}^i}{U_{c,t}^i}\right) \mathbb{E}_t \sum_{k=t}^{\infty} \left[\prod_{j=t}^{k-1} \beta_j\right] Z_k^i \quad \forall t \geq 1$$

$$\tilde{b}_{-1} = \mathbb{E}_{-1} \sum_{k=0}^{\infty} \left[\prod_{j=0}^{k-1} \beta_j\right] Z_k^i$$

$$\frac{\mathbb{E}_{t-1} U_{c,t}^i}{U_{c,t-1}^i} = \frac{\mathbb{E}_{t-1} U_{c,t}^j}{U_{c,t-1}^j}$$

$$\sum_{i=1}^{I} \pi_i c_i(s^t) + g\left(s_t\right) = \sum_{i=1}^{I} \pi_i \theta_i \left(s_t\right) l_i(s^t),$$

$$\frac{U_{l,t}^i}{\theta_{i,t} U_{c,t}^i} = \frac{U_{l,t}^1}{\theta_{1,t} U_{c,t}^1}$$

$$\tilde{b}_{t-1} \frac{U_{c,t-1}^i}{\beta_{t-1}} \text{ is bounded}$$

Ramsey Problem : Recursive

We can split the problem into two steps

1. $\mathbf{t \geq 1}$:Ex-ante continuation problem with state variables $(\mathbf{x}, \boldsymbol{\rho}, s_{\_})$ defined as

$$\mathbf{x} = \beta^{-1} \left(U^2_{c,t-1} \tilde{b}_{2,t-1}, ..., U^I_{c,t-1} \tilde{b}_{I,t-1} \right)$$

$$\boldsymbol{\rho} = \left(U^2_{c,t-1} / U^1_{c,t-1}, ..., U^I_{c,t-1} / U^1_{c,t-1} \right)$$

2. $\mathbf{t = 0}$: Ex-post initial problem with state variables $(\tilde{\mathbf{b}}_{-\mathbf{1}}, s_0)$

Optimal allocations : Recursive Problem $t \geq 1$

$$V(\mathbf{x}, \boldsymbol{\rho}, s_-) = \max_{c_i(s), l_i(s), x'(s), \rho'(s)} \sum_s \Pr(s|s_-) \left(\left[\sum_i \pi_i \alpha_i U^i(s) \right] + \beta(s) V(\mathbf{x}'(s), \boldsymbol{\rho}'(s), s) \right)$$

where the maximization is subject to

$$U_c^i(s)\left[c_i(s) - c_1(s)\right] + U_c^i(s) \left(\frac{U_l^i(s)}{U_c^i(s)} l_i(s) - \frac{U_l^1(s)}{U_c^1(s)} l_1(s) \right) + \beta(s) x_i'(s) = \frac{x U_c^i(s)}{\mathbb{E}_{s_-} \mathbf{U}_c^i} \text{ for all } s, i \geq 2$$

$$\frac{\mathbb{E}_{s_-} \mathbf{U}_c^i}{\mathbb{E}_{s_-} \mathbf{U}_c^1} = \rho_i \text{ for all } i \geq 2$$

$$\frac{U_l^i(s)}{\theta_i(s) U_c^i(s)} = \frac{U_l^1(s)}{\theta_1(s) U_c^1(s)} \text{ for all } s, i \geq 2$$

$$\sum_i \pi_i c_i(s) + g(s) = \sum_i \pi_i \theta_i(s) l_i(s) \quad \forall s$$

$$\rho_i'(s) = \frac{U_c^i(s)}{U_c^1(s)} \text{ for all } s, i \geq 2$$

$$\underline{x}_i(s; \mathbf{x}, \boldsymbol{\rho}, s_-) \leq x_i(s) \leq \bar{x}_i(s; \mathbf{x}, \boldsymbol{\rho}, s_-)$$

Optimal allocations : Recursive Problem $t = 0$

$$V_0\left(\{\tilde{b}_{i,-1}\}_{i=2}^{I}, s_0\right) = \max_{c_{i,0} l_{i,0}, x_0, \rho_0} \sum_i \pi_i \alpha_i U^i(c_{i,0}, l_{i,0}) + \beta(s_0) V\left(x_0, \rho_0, s_0\right)$$

where the maximization is subject to

$$U^i_{c,0}\left[c_{i,0} - c_{1,0}\right] + U^i_{c,0}\left(\frac{U^i_{l,0}}{U^i_{c,0}} l_{i,0} - \frac{U^1_{l,0}}{U^1_{c,0}} l_{1,0}\right) + \beta(s_0) x_{i,0} = U^i_{c,0}\tilde{b}_{i,-1} \text{ for all } i \geq 2$$

$$\frac{U^i_{l,0}}{\theta_{i,0} U^i_{c,0}} = \frac{U^1_{l,0}}{\theta_{1,0} U^{1,0}_c} \text{ for all } i \geq 2$$

$$\sum_i \pi_i c_{i,0} + g_0 = \sum_i \pi_i \theta_{i,0} l_{i,0}$$

$$\rho_{i,0} = \frac{U^i_{c,0}}{U^1_{c,0}} \text{ for all } i \geq 2$$

Steady States

Let $\Psi(s; \mathbf{x}, \boldsymbol{\rho}, s_{-})$ be an optimal law of motion for the state variables for the $t \geq 1$ recursive problem, i.e.,

$$\Psi(s; \mathbf{x}, \boldsymbol{\rho}, s_{-}) = \left(x'(s), \rho'(s)\right)$$

solves $t \geq 1$ Bellman equation given state $(\mathbf{x}, \boldsymbol{\rho}, s_{-})$

Definition

A steady state $\left(\mathbf{x}^{SS}, \boldsymbol{\rho}^{SS}\right)$ satisfies $\left(\mathbf{x}^{SS}, \boldsymbol{\rho}^{SS}\right) = \Psi\left(s; \mathbf{x}^{SS}, \boldsymbol{\rho}^{SS}, s_{-}\right)$ for all s, s_{-}

The steady state is a "node" such that continuation allocation has no further history dependence.

Existence

- Quasi Linear : For quasi-linear preferences it exists for a wide range of parameters and shocks. Further the economy goes in steady state in one period. Output and taxes are constant thereafter.
- For general preferences a degenerate SS exists if shocks are IID and take two values. The economy converges to this for all initial conditions.
- Outside the binary - IID case there exists an ergodic region where $(\mathbf{x}, \boldsymbol{\rho})$ no longer constant, but their fluctuations tend to be markedly reduced relative to the transient fluctuations

Intuition

- Consider 2 agents with $\theta_1(s) > \theta_2 = 0$.
- The state variable x is marginal utility scaled relative assets of the unproductive agent : $U_c^2(s)[b_2(s) - b_1(s)]$.
- One can normalize $b_2(s) = 0$ and x also can be interpreted to be scaled assets of the government

We can parse the two main forces that determine the dynamics of taxes and assets:

- **Fluctuations in inequality**, measured by spreads in marginal utilities
- **Fluctuations in the interest rates**

For quasi linear preferences both these forces are absent

Inequality distortions

Start with a spread in discount factors such that interest rates are equalized across i.e $R(s_l) = R(s_h)$. One can show that steady state $x > 0$

TFP (θ_1) : Adjust taxes τ or transfers T

Suppose $x = 0$ or $b_2(s) = b_1(s)$,

Transfers hurts the low productivity agent more

A fall in transfers that increases inequality gives rise to a cost not present in representative agent economies. To reduce the costs of inequality distortions optimal policy

x

By reducing the relative asset holdings of the productive agent the after tax, after-interest incomes of both agents are closer

Interest rates fluctuations

1. Suppose discount factors are constant. In our simple example, the implied interest rates are countercyclical and we still have $x > 0$. Consider again,
   - ↓ TFP

     If the tax rate τ is left unchanged, the government faces a shortfall of revenues. The optimal policy recommends

     ↑ x. This is same as the government accumulating assets [1]
   - It can use higher interest income to offset some of the revenue losses from taxes on labor
   - This force is similar to representative agent economies with endogenous fluctuations in interest rates
2. If discount factors spread is large, such that we have pro-cyclical interest rates one can have $x < 0$

[1] Normalize $b_2(s) = 0$

Comparative statics with Pareto weights

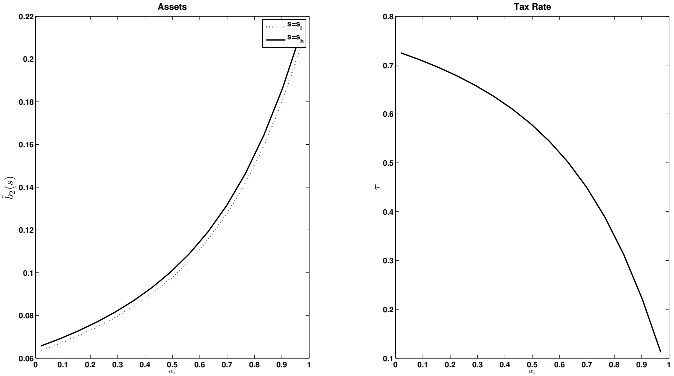

Figure: Stead state assets : $\tilde{b}_2(s) = \frac{\beta x^{SS}}{U_c^2(s)}$ and taxes : τ^{SS} as a function of Agent 1's (high skilled) Pareto weight

Numerical Example

Use a calibrated version of the economy to

- Revisit the magnitude of these forces and
- Study optimal policy responses at business cycle frequencies when the economy is possibly far away from the steady state

Numerical Example : Calibration

Take a 2 shock 2 type economy with preferences $U(c, l) = \psi \log(c) + (1 - \psi) \log(1 - l)$ and allow $\theta_i(s), \beta(s), g(s)$ to depend to shocks.

- Pick the baseline parameters to match some low frequency moments
- Calibrate the fluctuations to match recent US recessions (i.e., 1991-92, 2001-02 and 2008-10):
  1. The left tail of the cross-section distribution of labor income falls by more than right tail
  2. Short term interest rates fall
  3. Recessions last longer than booms

Calibration

Parameter	Value	Description	Target
ψ	0.6994	Frisch elasticity of labor supply	0.5
$\bar{\theta}_1$	4	Log 90-10 wage ratio (Autor et all)	4
$\bar{\theta}_2$	1	Normalize to 1	1
β	0.98	Average (annual) risk free interest rate	2%
α_1	0.69	Marginal tax rate in the economy with no shocks	20%
g	12%	Average pre-transfer expenditure- output ratio	12 %
$\frac{\hat{\theta}_2}{\hat{\theta}_1}$	2.5	Relative drop in wage income of 10th percentile	2.5
$\hat{\theta}_1$	1.2%	Average output loss	3%
$\hat{\beta}(s)$	1.96%	Difference in real interest rates between booms and recession	1.96%
$P(r\|r)$	0.63	Duration of recessions	2.33 years
$P(b\|b)$	0.84	Duration of booms	7 years

Table: Benchmark calibration

We initialize the economy with initial conditions such that implied debt to GDP ratio is 60%

Results: Some variants

From the Benchmark calibration we will study the following variants

1. Acyclical interest rates : Smaller spread in discount factor shocks
2. Countercyclical interest rates: No discount factor shocks
3. No inequality: Equal fall in all agents productivities (TFP shock) and no discount factor shocks
4. Government expenditure shocks: A fall in g that produces a comparable fall in output

Results : Long run

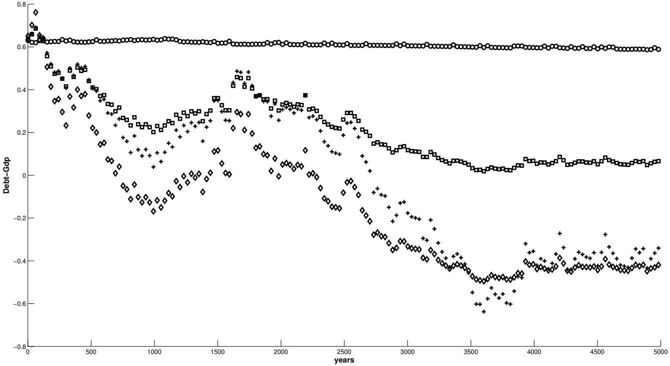

Figure: Debt benchmark (o), acyclical interest rates (+), countercyclical interest rates ($\diamond$) and no inequality shocks ($\square$)

Remarks

- Long run tendency to converge to some ergodic set. But convergence is very slow more details on speed of convergence.
- With low discount factor shocks the trend is towards positive assets
- With high discount factor shocks that produce procyclical real interest rates there is no tendency to reduce debt even after 5000 years

Short Run

To understand the shor run responses,

- We set the exogenous state s_0 so that we are at the outset of a recession
- Solve the time 0 problem with identical initial conditions across different settings. This pins down the initial state vector x_0, ρ_0 that appears in our time 0 Bellman equation
- We then use the policy rules to compute fluctuations of different components in the government budget constraint across states

For each variable z in the table we report in the form $\Delta z \equiv \left(z\left(s_l | x_0, \rho_0, s_0\right) - z\left(s_h | x_0, \rho_0, s_0\right)\right) / \bar{Y}$ where $\bar{Y}$ is average undistorted GDP in percentages

Results : Short run

	Δg	ΔB	ΔT	$\Delta[\tau\theta_1 l_1]$	$\Delta[\tau\theta_2 l_2]$	ΔY	$\Delta\tau$
Benchmark	0.0000	-1.1561	0.6871	-0.1593	-0.3096	-2.8536	0.3732
Acyclical Interest Rates	0.0000	-1.1126	0.6591	-0.1497	-0.3038	-2.8613	0.3879
Countercyclical Interest Rates	0.0000	-1.0794	0.6387	-0.1415	-0.2992	-2.8677	0.3997
No Inequality	0.0000	-0.1380	**-0.5459**	-0.5635	-0.1204	-2.6294	0.0622
Expenditure Shocks	-7.5037	2.9137	2.8612	-1.3759	-0.3530	-2.3443	-1.1598

Table: The tables summarizes the changes in the different components of the government budget as we transit from "boom" to a "recession". All numbers are normalized by un-distorted GDP except τ and reported in percentages.

Note that predetermined variables like repayment on existing debt drop out of the accounting and we have

$$\Delta[g] + \Delta[T] + \Delta[B] = \Delta[\tau\theta_1 l_1] + \Delta[\tau\theta_2 l_2]$$

Conclusions

- Size of government debt alone is not informative $\Longrightarrow$ need to know the net distribution of assets in the economy
- The optimal tax and transfer scheme balances
  1. welfare losses from fluctuating taxes
  2. welfare losses from fluctuating transfer
- Since the welfare costs depend on the how debt is distributed, there are incentives to affect the net assets overtime
- With incomplete markets, interest fluctuations turn out to be key for the long run correlations between productivities and net assets
- Ignoring heterogeneity produces misleading results about size and direction of short run optimal policy response

Speed of convergence (I)

Suppose we are in the binary-IID world where steady states are deterministic.

- The optimal policy induces two risk adjusted martingales $\{\boldsymbol{\mu}_t, \boldsymbol{\rho}_t\}$.
- One can represent the optimal allocation recursively in terms of $\{\boldsymbol{\mu}(s^{t-1}), \boldsymbol{\rho}(s^{t-1})\}$ and s_t.
- Why $(\boldsymbol{\mu}, \boldsymbol{\rho})$ instead of $(\mathbf{x}, \boldsymbol{\rho})$?
- Linearize optimal policies for each s_t around the degenerate steady state.
- Study the eigenvalues of the conditional mean and variance dynamics (there are deterministic linear systems)

Speed of convergence (II)

$\hat{\Psi}_t = \begin{bmatrix} \boldsymbol{\mu}_t - \boldsymbol{\mu}^{SS} \\ \boldsymbol{\rho}_t - \boldsymbol{\rho}^{SS} \end{bmatrix}$ be deviations from a steady states

$$\hat{\Psi}_{t+1} = B(s_{t+1})\hat{\Psi}_t$$

This linearized system has coefficients that are functions of the shock. We have

Proposition

If the (real part) of eigenvalues of $\mathbb{E}B(s)$ are less than 1, the system converges to zero in mean. Further for large t, the conditional variance of $\hat{\Psi}$, denoted by $\Sigma_{\Psi,t}$, follows a deterministic process governed by

$$vec(\Sigma_{\Psi,t}) = \hat{B}vec(\Sigma_{\Psi,t-1}),$$

where $\hat{B}$ is a square matrix of dimension $(2N-2)^2$. In addition, if the (real part) of eigenvalues of $\hat{B}$ are less than 1, the system converges in probability.

The eigenvalues (in particular the largest one) are instructive not only for whether the system is locally stable but also how quickly the steady state is reached

Speed of convergence : Size of shocks and risk aversion

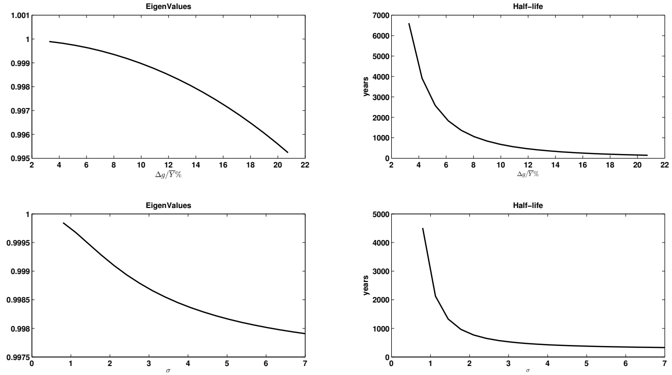

Figure: The top (bottom) panel plots the dominant eigenvalue of $\hat{B}$ and the associated half life as we increase the spread between the expenditure levels (risk aversion).

Back to main.