Techniques for Macro Finance

Wenhao Li

September 18, 2018

Contents

1 Optimization

1.1 A General Envelope Theorem

This is based on Milgrom & Segal (2002). Denote

$$V(t) = \max_{x \in X} f(x,t), \quad X^*(t) = \arg\max_{x \in X} f(x,t)$$

Theorem 1 (Derivative Part of Envelope Theorem)**.** *Suppose that t is in a bounded interval, say $[0,1]$, $x \in X^(t)$ and the derivative $f_2(x,t)$ exists. Then we have the following results.*

- *If $t < 1$ and $V'(t^+)$ exists, then $V'(t^+) \geq f_2(x,t)$.*
- *If $t > 0$ and $V'(t^-)$ exists, then $V'(t^-) \leq f_2(x,t)$.*
- *If $t \in (0,1)$, and $V'(t)$ exists, then $V'(t) = f_2(x,t)$.*

But we need slightly stronger conditions for the integral formula as follows.

Theorem 2. *If for all (x,t), $f_2(x,t)$ exists, and $|f(x,t)| \leq B(t)$ for some function $B(t)$ with finite integral on $[0,1]$, then for all t,*

$$V(t) - V(0) = \int_0^t f_2(x(s),s)ds$$

Note: for a general optimization problem with constraints, we can first build up an equivalence between the relaxed problem and the original one, and then apply the general envelope theorem on the relaxed problem. Because we only need to study the relation between t and V, we don't need to relax constraints that only involve x.

1.2 Parametric Continuity and Monotonicity

We want to study the properties of the following problem

$$V(\theta) = \max_{a \in \mathcal{D}(\theta)} f(a,\theta)$$

where we view a as "action" and $V(\theta)$ as the value function for type $\theta \in \Theta \subset \mathbb{R}^m$. The constraint set $\mathcal{D}(\theta) \subset \mathbb{R}^n$. Define the solution set as

$$\mathcal{A}^*(\theta) = \{a | f(a,\theta) = V(\theta)\}$$

We are interested in the following questions:

- Continuity of $V(\theta)$ and $\mathcal{A}^*(\theta)$.
- Concavity of $V(\theta)$.
- Differentiability of $V(\theta)$.

- Monotonicity of $\mathcal{A}^*(\theta)$.

The following theorem addresses the first two questions.

Theorem 3 (The Extended Maximum Theorem)**.** *Let $f(a, \theta)$ be a continuous function in both a and θ, and $\mathcal{D}$ be a compact valued, continuous correspondence. Then we have the following results:*

- *$V(\theta)$ is a continuous function on Θ, and $\mathcal{A}^*(\theta)$ is a compact-valued, upper-semi continuous correspondence on Θ.*
- *If $f(a, \theta)$ is (strictly) concave, and $\mathcal{D}(\theta)$ is convex for all $\theta \in \Theta$, then $V(\theta)$ is (strictly) concave.*
- *If $f(\cdot, \theta)$ is concave for each θ, then $\mathcal{A}^*(\theta)$ is convex. When the "concave" is replaced by "strictly concave", then $\mathcal{A}^*(\theta)$ is a continuous single-valued function.*

Proof. Refer the proof in Chapter 9 of Sundaram (1996), A First Course in Optimization Theory. □

Next, we want to study the monotonicity of $\mathcal{A}^*(\theta)$ and differentiability of $V(\theta)$. For monotonicity, we need supermodularity in multi-dimensional space. For differentiability, we need to construct a differentiable and concave supporting function.

For a constant one dimensional action space $\mathcal{D}(\theta) = S$, we have the following results.

Theorem 4 (Parametric Monotonicity for Constant Single-Dimensional Action Space)**.** *Assume $\mathcal{D}(\theta) = S \subset \mathbb{R}$, and $f(a, \theta)$ has strictly increasing differences in (x, θ). Then the optimal actions are monotonically increasing in θ. If $f(a, \theta)$ has weakly increasing differences, then $\mathcal{A}^*(\theta)$ is increasing in the strong set order.*

Proof. See the proof in Chapter 9 of Sundaram (1996), and also the lecture notes of Ilya's microeconomic class. □

Theorem 5 (Parametric Monotonicity)**.** *Theorem 10.7 in Sundaram (1996).*

Next, I state the results in differentiability of $V(\theta)$.

Theorem 6 (Differentiability of $V(\theta)$)**.** *Suppose $V(\theta)$ is concave. Let $\theta \in interior(\Theta)$, and $\mathcal{N}(\theta_0)$ a neighborhood around θ_0. Suppose $w : \mathcal{N}(\theta_0) \to \mathbb{R}$ satisfies*

- *$w(\theta) \leq V(\theta)$*
- *$V(\theta_0) = w(\theta_0)$*
- *$w(\cdot)$ is differentiable at θ_0.*

i.e. $w(\theta)$ is supporting the concave function $V(\theta)$ from below and touches $V(\theta)$ at the point θ_0. Then $V(\theta)$ is differentiable at θ_0.

Proof. Any subgradient of $V(\cdot)$ at θ_0, denoted by $\nabla V(\theta_0)$, must satisfy

$$\nabla V(\theta_0) \cdot (\theta - \theta_0) \geq V(\theta) - V(\theta_0) \geq w(\theta) - V(\theta_0) \geq w(\theta) - w(\theta_0)$$

Since $w(\cdot)$ is differentiable at w_0, we can write

$$w(\theta) - w(\theta_0) = \nabla w(\theta) \cdot (\theta - \theta_0) + o(\theta - \theta_0)$$

Because θ can be arbitrary value in the neighborhood, we must have

$$\nabla V(\theta_0) = \nabla w(\theta)$$

which means the subgradient is unique. Thus $V(\cdot)$ is differentiable at θ_0. □

2 Stochastic Calculus and Stochastic Differential Equations

2.1 Change of Measure

Basic question: Why a change of measure doesn't change the volatility part, but only the drift?

Answer: Change of measure only changes the weight assigned to different paths, but not paths themselves. The volatility of a process could be backed out from a single path, which is not affected by the change of measure.

Two probability measures Q and P are equivalent probability measures on $(\Omega, \mathcal{F})$ if, for any event A, $P(A) = 0$ if and only if $Q(A) = 0$, i.e. the two probability measures have the same zero set. In this case, there is always a strictly positive random variable ξ called the *Radon-Nikodym* derivative of Q w.r.t. P, with the following property: If Z satisfies $E^Q[|Z|] < \infty$, then $E^Q[Z] = E^P[\xi Z]$.

Proposition 1. *If $\mathcal{G} \subset \mathcal{F}$, then*

$$E^Q[Z|\mathcal{G}] = \frac{1}{E^P[\xi|\mathcal{G}]} E^P[\xi Z|\mathcal{G}]$$

Proof. To prove the above claim, we can rewrite it as

$$E^Q[\left(E^P[\xi|\mathcal{G}]Z\right)|\mathcal{G}] = E^P[\xi Z|\mathcal{G}]$$

Denote

$$E^P[\xi|\mathcal{G}] = Y$$

We need to prove that the

$$E^Q[ZE^P[\xi|\mathcal{G}]\mathbf{1}_A] = E^Q[E^P[\xi Z|\mathcal{G}]\mathbf{1}_A]$$

for any $A \in \mathcal{G}$. We note that

$$E^Q[ZY\mathbf{1}_A] = E^Q[E^P[\xi Z|\mathcal{G}]\mathbf{1}_A]$$

which is immediate when we note

$$E^Q[E^P[\xi Z|\mathcal{G}]\mathbf{1}_A] = E^P[E^P[\xi Z|\mathcal{G}]\xi\mathbf{1}_A]$$

$$= E^P[E^P[\xi Z\mathbf{1}_A|\mathcal{G}]Y] = E^P[\xi Z\mathbf{1}_A Y] = E^Q[ZY\mathbf{1}_A]$$

□

Define a Brownian filtration $\mathcal{F}$ and an associated probability space $(\Omega, \mathcal{F}, P)$. Suppose that the probability Q is defined as

$$\frac{dQ}{dP} = \xi_T$$

and

$$\xi_t = E[\xi_T|\mathcal{F}_t]$$

Then Proposition 1 implies the following result: A process $\{X_t, 0 \leq t \leq T\}$ is a Q martingale if and only if $\{\xi_t X_t, 0 \leq t \leq T\}$ is a P martingale. This is because

$$E_t^Q[X_T] = X_t \Rightarrow \quad \frac{E_t^P[\xi_T X_T]}{\xi_t} = X_t \quad \Rightarrow \xi_t X_t = E_t^P[\xi_T X_T]$$

Denote the Brownian motion under Q as B^Q. Then we must have $\xi_t B_t^Q$ a martingale under P, which implies

$$B_t^Q = B_t - \theta_t dt$$

where

$$\frac{d\xi_t}{\xi_t} = \theta_t dB_t$$

We can explicitly write a Radon-Nikodym derivative process ξ_t that satisfies the above SDE as

$$\xi_t = \exp(\int_0^t \theta_s dB_s - \frac{1}{2}\int_0^t \theta_s^2 ds)$$

where we need the Novikov condition

$$E[\exp(\frac{1}{2}\int_0^T \theta_s^2 ds)] < \infty$$

to guarantee that ξ_t is integrable and thus indeed a martingale. The Novikov condition is quite easy to interpret. Suppose that θ_s is a deterministic process, then

$$\begin{aligned} &E[\exp(\textstyle\int_0^t \theta_s dB_s - \frac{1}{2}\int_0^t \theta_s^2 ds)] < \infty \\ &\Leftrightarrow E[\exp(\textstyle\int_0^t \theta_s dB_s)] < \infty \\ &\Leftrightarrow E[\exp(\textstyle\frac{1}{2}\int_0^t \theta_s^2 ds)] < \infty \end{aligned}$$

2.2 Properties of General Stochastic Processes

First, we want to know whether a stochastic process is continuous. It is answered by the following continuity theorem.

Theorem 7 (Kolmogorov's Continuity Theorem)**.** *Suppose that the process* $X = \{X_t\}_{t\geq 0}$ *satisfies the following condition: for all* $T > 0$ *there exist positive constants* α, β, D *such that*

$$E[|X_t - X_s|^\alpha] \leq D \cdot |t-s|^{1+\beta}, \quad 0 \leq s, t \leq T$$

Then there exists a continuous version of X*.*

Then we have the martingale representation theorem for square integrable martingales.

Theorem 8 (The Martingale Representation Theorem for Square Integrable Martingales, Theorem 4.3.4 of Øksendal (2003).)**.** *Let* $B(t) = (B_1(t), \cdots, B_n(t))$ *be* n*-dimensional. Suppose* M_t *is an* $\mathcal{F}_t$ *martingale w.r.t.* $\mathbb{P}$ *and* $M_t \in L^2(\mathbb{P})$ *for all* $t \geq 0$*. Then there exists a unique stochastic process* $g(s,\omega)$ *such that* $g \in \mathcal{H}^2(0,t)$ *for all* $t \geq 0$*, and*

$$M_t(\omega) = E[M_0] + \int_0^t g(s,\omega) dB(s,\omega).$$

Next, we also have a version of martingale representation for local martingales.

Theorem 9 (The Martingale Representation Theorem for Local Martingales[1])**.** *Let* $B(t) = (B_1(t), \cdots, B_n(t))$ *be* n*-dimensional. Suppose* M_t *is an* $\mathcal{F}_t$ *local martingale w.r.t.* $\mathbb{P}$*. Then we can find a predictable process* $g(s,\omega)$ *satisfying*

$$\int_0^t g(s,\omega)^2 ds < \infty \quad a.s.$$

and

$$M_t(\omega) = E[M_0] + \int_0^t g(s,\omega) dB(s,\omega)$$

2.3 Ito's Formula

The quadratic variation is defined as

$$[X,Y] = XY - \int X\_dY - \int Y\_dX$$

Since the process $[X,X]$ is non-decreasing with right continuous paths, we can decompose $[X,X]$ *path-by-path* into

$$[X,X]_t = [X,X]_t^c + \sum_{0\leq s\leq t} (\Delta X_s)^2$$

where ΔX is the jump component and $[X,X]^c$ is the path-by- path continuous part of $[X,X]$.

The most general Ito's formula applies to semimartingales, which is defined as a composable process that has a local martingale component and a finite variation component. The class of semimartingales is the largest that can be used for defining stochastic integral, which can be easily interpreted from the definition in chapter 2 of Protter (2005) that a semimartingale is a process that maps a previsible process into a finite-valued random variable with certain topological properties.

[1]Refer to The Martingale Representation Theorem online.

Remark 1. *Why we want to use semimartingales? We want to have a class of stochastic process that behave nicely against the previsible integrand and has the closure property itself, i.e. remains its property after stochastic integration. As stochastic integration, we in general want two types of properties: (1) Algebraically, it should be a linear mapping. (2) Topologically, it should preserve the monotonic convergence theorem or bounded convergence theorem.*

Next, we have a general Ito's formula on semimartingales.

Theorem 10 (Ito's Formula, Theorem 32 of Chapter 7, Protter (2005))**.** *Let X be a semimartingale and let f be a C^2 real function. Then $f(X)$ is again a semimartingale, and the following formula holds:*

$$f(X_t) - f(X_0) = \underbrace{\left(\int_{0+}^{t} f'(X_{s-}) dX_s - \sum_{0<s\leq t} f'(X_{s-}) \Delta X_s \right)}_{\textit{first order no jump component}}$$

$$+ \underbrace{\frac{1}{2} \int_{0+}^{t} f''(X_{s-}) d[X,X]_s^c}_{\textit{second order no jump component}} + \underbrace{\sum_{0<s\leq t} \left(f(X_s) - f(X_{s-}) \right)}_{\textit{jump component}}$$

The same formula applies when $f \in C^1$ everywhere, and C^2 outside finitely many points with $|f''(x)| \leq M$ at C^2 points.

From the above theorem, we know that semimartingales are preserved after C^2 transformations. This property slightly extends: semimartingales are preserved after convex transformation.

Theorem 11. *Let $f : \mathbb{R} \to \mathbb{R}$ be convex and let X be a semimartingale. Then*

$$f(X_t) - f(X_0) = \int_{0+}^{t} f'(X_{s-}) dX_s + A_t$$

where f' is the left derivative of f and A is an adapted, right continuous, ***increasing*** *process. Moreover, jumps in A_t satisfies*

$$\Delta A_t = f(X_t) - f(X_{t-}) - f'(X_{t-}) \Delta X_t$$

Note that due to the generality of $f(\cdot)$, we cannot explicitly write down the process A_t. The above theorem implies that if X is a semimartingale, then $|X|$, X^+, X^- are all semimartingales.

A commonly used formula: Ito's formula for ratio. Suppose two stochastic processes X_t and Y_t follow

$$\frac{dX_t}{X_{t-}} = \mu_t^X dt + \sigma_t^X dB_t + \kappa_{t-}^X dN_t$$

$$\frac{dY_t}{Y_{t-}} = \mu_t^Y dt + \sigma_t^Y dB_t + \kappa_{t-}^Y dN_t$$

Then the ratio of the two processes follows

$$\frac{d(X_t/Y_t)}{(X_{t-}/Y_{t-})} = (\mu_t^X - \mu_t^Y + (\sigma_t^Y)^2 - \sigma_t^X \sigma_t^Y) dt + (\sigma_t^X - \sigma_t^Y) dB_t + (\frac{1+\kappa_t^X}{1+\kappa_t^Y} - 1) dN_t$$

Without Poisson jumps, we can use a convenient interim result

$$\frac{d(X_t/Y_t)}{(X_{t-}/Y_{t-})} = \frac{dX_t}{X_t} - \frac{dY_t}{Y_t} + (\frac{dY_t}{Y_t})^2 - \frac{dX_t}{X_t}\frac{dY_t}{Y_t}$$

2.4 Stochastic Differential Equations

For stochastic differential equation in the form

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad 0 \le t \le T$$

an existence and uniqueness theorem is as follows.

Theorem 12 (Existence and Uniqueness of Stochastic Differential Equations)**.** *If the coefficients satisfy a space-variable Lipschitz condition*

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \le D|x - y|$$

and a spacial growth condition

$$b(t, x)^2 + \sigma(t, x)^2 \le K(1 + x^2)$$

for some $K > 0$. Then there exists a unique continuous adapted solution solution X_t that is uniformly bounded

$$\sup_{0 \le t \le T} E[X_t^2] < \infty$$

Remark 2. *Q: Is it possible to extend the uniformly boundedness to the following integration:*

$$E[e^{X_t}]$$

A good way is probably just derive the Lipschitz condition and spatial growth condition for $Y_t = f(X_t) = e^{X_t}$. However, the spatial growth condition might be violated due to the exponential.

2.5 Local Time

The computation of local time.

Proposition 2. *Denote the density of X_t as $p(t, x)$, where X_t is a semimartingale. Let the local time of X_t at point x be L_t^x, which is formally defined as*

$$L_t^x = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t I\{x - \varepsilon \le X_s \le x + \varepsilon\} ds$$

Then we have

$$E[L_t^x] = \int_0^t p(s, x) dx$$

2.6 The Filtering Problem

The filtering problem can be generally stated as follows: Given observations of Z_t that satisfies

$$dY_t = c(t, Y_t)dt + \gamma(t, Y_t)dZ_t, \quad Z_0 = 0,$$

what is the best estimate $\hat{X}_t$, measurable w.r.t. $\sigma\{Z_s, s \leq t\}$, for the state X_t in the system

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t$$

where Z_t and B_t are independent Brownian motions.

2.7 Verification of the HJB Equation

$$V(n_0) = E[\int_0^t e^{-\rho t}u(c_s)ds] + e^{-\rho t}E[V(n_t)]$$

Step 1: Under optimal policy.

Then the first convergence can be derived from monotone convergence theorem. The second one can be bounded. Basic question: How to calculate the bounds of the following process:

$$dX_t = \mu_t dt + \sigma_t dB_t$$

Step 2: Under general policy.

$$V(n_0) \geq E[\int_0^t e^{-\rho t}u(c_s)ds] + e^{-\rho t}E[V(n_t)]$$

For the second term, typically people use the "transversality condition, which is imposed in an ad hoc way.

2.8 Kolmogrov Forward Equations

For a stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dB_t,$$

we want to know the density function of X_t. The density function will satisfy a differential equation called Komogrov forward equation. A starting point is to characterize how the density changes over time.

Denote the density of X_t as $g(x, t)$ at time t. Then intuitively, for any C^2 function $f(\cdot)$ with

$$\lim_{x\to\infty} f(x) = \lim_{x\to-\infty} f(x) = 0,$$

we have

$$\frac{d}{dt}E[f(X_t)] = \int f(x)g_t(x, t)dt = \int g(x, t)\left(\mu(x)f'(x) + \frac{\sigma^2(x)}{2}f''(x)\right)dx$$

With integration by part, we have

$$\int g(x,t)\left(\mu(x)f'(x)+\frac{\sigma^2(x)}{2}f''(x)\right)dx$$

$$=-\int f(x)\frac{\partial\left(g(x,t)\mu(x)\right)}{\partial x}dx+\int f(x)\frac{\partial\left(\frac{1}{2}\sigma^2(x)g(x,t)\right)}{\partial x}dt$$

Because $f(\cdot)$ is an arbitrary C^2 function with only limit constraints, we get

$$g_t(x,t)=-\frac{\partial}{\partial x}\left(g(x,t)\mu(x)\right)+\frac{\partial^2}{\partial x^2}\left(\frac{1}{2}\sigma^2(x)g(x,t)\right) \tag{1}$$

Remark 3 (Interpretation)**.** *The above equation has a very natural interpretation of sand piles. First, let's interpret the drift term. When the slope of the sand hill is positive, and the drift to the right $\mu(x)=\mu>0$, then for sure the density at x will decrease. Second, fixing $\sigma(x)=\sigma$, the second part is about the curvature of the density. Suppose we are shaking the sand pile. Then the location with a concave shape will dissipate, which means $g''_{xx}(x,t)<0$ will reduce density at x. For non constant $\sigma(x)$, we can interpret little creatures at different location with large $\sigma(x)\sigma(x)'$, and the sand will accumulate in those locations.*

Suppose a stationary distribution exists. Then it must satisfy

$$-\frac{\partial}{\partial x}\left(g(x)\mu(x)\right)+\frac{\partial^2}{\partial x^2}\left(\frac{1}{2}\sigma^2(x)g(x)\right)=0$$

$$\Rightarrow -g(x)\mu(x)+\frac{\partial}{\partial x}\left(\frac{1}{2}\sigma^2(x)g(x)\right)=F$$

for some constant F. Typically, we should have $\mu(x)\to 0$ for x on the boundary, which implies $F=0$. Then we get

$$g(x)=K\frac{1}{\sigma^2(x)}e^{\int\frac{2\mu(x)}{\sigma^2(x)}dx}$$

Example: OU process

$$dX_t=\theta(\mu-X_t)dt+\sigma dZ_t$$

Then the density is

$$g(x)=Ke^{-\frac{2\theta(x-\mu)^2}{\sigma^2}}$$

which is the density of distribution $N(\mu,\sigma^2/(2\theta))$. For a general Brownian motion

$$dX_t=\mu dt+\sigma dB_t$$

we do not have a stationary distribution, because when $\mu\neq 0$, X_t will go to either ∞ or $-\infty$. When $\mu=0$, the process will converge to a general uniform distribution on the whole real line.

2.9 Viscosity Solutions

Suppose we have a general function $F(x, v, P, X)$, with the interpretation that v is certain value function, $P = Dv(x)$ and $X = D^2v(x)$. The state variable x can be multidimensional, in which case we have $v \in \mathbb{R}$, $P \in \mathbb{R}^n$, and $X \in S^{n\times n}$. We define the monotone property as follows.

Definition 1 (Monotone Property). *$F(x, v, P, X)$ is said to be monotone if it increases in v, but decreases in X, where the order in X is in the semidefinite order.*

Then we would like to define a sub solution and a super solution that are effective lower and upper bounds of the true solution. A solution to the differential equation $F(x, v, Dv, D^2v) = 0$ then is both a sub and super solution.

Definition 2 (Sub-solution). *An upper semi-continuous function $\underline{v}$ is a sub-solution if for any C^2 test functions ϕ, s.t. $\phi(x_0) = \underline{x}_0$ for some x_0, $\phi(x) \geq \underline{v}(x)$ in the neighborhood of x_0,*

$$F(x_0, \phi(x_0), D\phi(x_0), D^2\phi(x_0)) \leq 0$$

Then we have the following uniqueness result.

Proposition 3 (Uniqueness of Boundary Value Problem in Viscosity Solutions). *Suppose $\underline{v}$ is a sub-solution to the problem*

$$\begin{aligned}&F(x, v, Dv, D^2v) = 0\\&v(x|x \in \partial\Omega) = v^{boundary}(x)\end{aligned}$$

and $\bar{v}$ is a super solution. Suppose on the boundary $\underline{v} \leq \bar{v}$, then we have $\underline{v} < \bar{v}$ on the whole region Ω.

The above proposition implies uniqueness of the solution within the class of functions having the required boundary condition.

Uniqueness of viscosity solutions is guaranteed by the following proposition.

Proposition 4 (Uniqueness of Viscosity Solutions). *The pointwise supremum of all sub-solutions is a solution.*

3 Consumption and Portfolio Choice

The general setup of consumption and portfolio choice problem, and how to solve them. Darrell's homework 9.9 and 9.10 are good examples.

Based on the form of utility, we can conjecture different value functions.

- Log utility
$$U(c) = \int_0^\infty e^{-\rho t} \log(c_t)dt$$
Then we can conjecture the value function as
$$V(W_t, X_t) = \frac{\log(W_t)}{\rho} + J(X_t)$$
where X_t is a state variable that determines asset returns, volatilities and so on.

- CRRA utility.

$$U(c) = \int_0^\infty e^{-\rho t} \frac{c_t^{1-\gamma}}{1-\gamma} dt$$

Then we can conjecture the value function as

$$V(W_t, X_t) = \frac{W_t^{1-\gamma}}{1-\gamma} J(X_t)^{1-\gamma}$$

Once we get a good conjecture of the utility function, then we can plug in and get a solution. Note that under log utility, the wealth part and the state variable part are separable, and thus optimal consumption policy is independent from other state variables.

3.1 The Most Basic Example: Merton's Problem

In Merton's problem, the basic setup has N risky securities and a safe asset. Denote $X = (X^{(1)}, \cdots, X^{(N)})$, where

$$dX_t^{(i)} = \mu^i X_t^{(i)} dt + X_t^{(i)} \sigma^{(i)} dB_t$$

and the risk-free rate is r, with price

$$d\beta_t = r\beta_t dt$$

Here μ, σ and r are constants (vectors or matrix). Later we can study general processes.

The Finite Horizon Case

The optimization problem is

$$\max_{c \geq 0, \theta} U(c, Z) = E[\int_0^T u(c_t, t) dt + F(Z)]$$

where θ is the portfolio weight on risky assets. Then

$$\frac{dW_t}{W_t} = (\theta_t \cdot (\mu - r\mathbf{1}) + r - \frac{c_t}{W_t}) dt + \theta_t^T \sigma dB_t$$

and $W_T = Z$. Next, we conjecture

$$J(w, t) = E[\int_t^T u(c_t, t) dt + F(Z) | W_t = w]$$

can solve the Bellman equation

$$\sup_{c, \theta} \{u(c, t) + \mathcal{D}_{c,\theta} J(w, t)\} = 0$$

where

$$D_{c,\theta} J(w, t) = w J_w \left(\theta \cdot (\mu - r\mathbf{1}) + r - \frac{c}{w}\right) + \frac{1}{2} w^2 J_{ww} \theta^T \sigma \sigma^T \theta + J_t$$

This results in the

$$u_c'(c,t) = J_w \frac{1}{w}$$

$$J_w(\mu - r\mathbf{1}) + wJ_{ww}\sigma\sigma^T\theta = 0$$

which indicates that the optimal portfolio choice should be

$$\theta = -\frac{J_w}{wJ_{ww}}\left(\sigma\sigma^T\right)^{-1}(\mu - r\mathbf{1})$$

To fully solve the problem, we need to provide the forms of u and F. Assume $u = 0$, and

$$F(w) = \frac{w^{1-\gamma}}{1-\gamma}$$

which results in $c = 0$, and a conjectured form

$$J(w,t) = k(t)\frac{w^{1-\gamma}}{1-\gamma}$$

This leads to the optimal portfolio and consumption

$$\theta = \frac{1}{\gamma}\left(\sigma\sigma^T\right)^{-1}(\mu - r\mathbf{1})$$

and the equation

$$k'(t) + A_0 k(t) = 0, \quad k(T) = 1$$

where

$$A_0 = (1-\gamma)r - \frac{1-\gamma}{2\gamma}(\mu - r\mathbf{1})^T\left(\sigma\sigma^T\right)^{-1}(\mu - r\mathbf{1})$$

As a result,

$$k(t) = e^{A_0(T-t)}$$

and

$$J(W_t,t) = e^{A_0(T-t)}\frac{W_t^{1-\gamma}}{1-\gamma}$$

Verification Arguments: construct a super martingale for any strategy, and a true martingale for the optimal strategy. We can construct a process

$$M_t = \int_0^t u(c_s,s)ds + J(W_t,t)$$

which is the value of the whole process if we follow an arbitrary strategy until time t, and from t onwards uses the "optimal" strategy. By Ito's formula, we have

$$dM_t = u(c_t,t) + \mathcal{D}_{c,\theta}J(W_t,t) + J_w(W_t,t)W_t\theta_t^T\sigma dB_t \leq J_w(W_t,t)W_t\theta_t^T\sigma dB_t$$

$$\Rightarrow M_t \leq \int_0^t J_w(W_s,s)W_s\theta_s^T\sigma dB_s + M_0$$

for any feasible strategies. Because we have $M_t \geq 0$, the right hand side is a nonnegative local martingale, which is a super martingale. As a result, we have

$$E[M_T] \leq E[\int_0^T J_w(W_s,s)W_s\theta_s^T\sigma dB_s + M_0] \leq M_0$$

$$\Rightarrow E[\int_0^T u(c_t,t)dt + F(W_T)] \leq J(w,0)$$

As a result, the function $J(w,0)$ is a upper bound. Then we can prove that a strategy c^* and θ^* that achieves the Bellman equation should achieve the upper bound. Denote

$$M_t^* = \int_0^t u(c_s^*, s)ds + J(W_t^*, t)$$

Then we have

$$dM_t^* = e^{A_0(T-t)}(W_t^*)^{1-\gamma}(\theta_t^*)^T \sigma dB_t$$

To prove this is indeed a martingale, we can calculate

$$E[M^*, M^*]_T = E[\int_0^T e^{A_0 2(T-t)}(W_t^*)^{2(1-\gamma)}(\theta_t^*)^T \sigma\sigma^T \theta_t^* dt]$$

Note that we only need to show that

$$E[(W_t^*)^{2(1-\gamma)}] \leq K$$

for some constant K, which can be readily proved from the law of motion for W_t. In general, we need a specific argument for each specific problem.

The Infinite Horizon Case

The only complexity from the infinite horizon is that we need to impose the so-called "transversality condition" to the Bellman equation to guarantee the solution is indeed the optimal value function. Denote

$$\begin{aligned}
&\sup_{c_t,\theta_t} E[\textstyle\int_0^\infty e^{-\rho t}u(c_t)dt] \\
&s.t. \\
&\tfrac{dW_t}{W_t} = (\theta_t \cdot (\mu - r\mathbf{1}) + r - \tfrac{c_t}{W_t})dt + \theta_t^T \sigma dB_t
\end{aligned}$$

To get the value function, we solve the Bellman equation

$$\rho J(w) = \sup_{c,\theta}\left\{u(c) + V'(w)w(\theta \cdot (\mu - r\mathbf{1}) + r - \frac{c}{w}) + V''(w)w^2\theta^T\sigma\sigma^T\theta\right\}$$

subject to the transversality condition

$$\lim_{T\to\infty} E[e^{-\rho T}|J(W_T)|] = 0$$

for any admissible control. Note that order to confirm the optimality of this policy, the transversality condition **must be checked**.

The Martingale Approach

Conditions for the martingale approach to work:

- Market is complete.

- There is no portfolio constraint.

However, these two conditions significantly restrict the applicability of this approach. Moreover, the martingale approach doesn't generate the optimal portfolio choice immediately, and the optimal portfolio choice can be solved with Malliavian calculus, but explicitly only under very special cases. Thus this approach is of very limited applicability.

3.2 General Return Processes

Now we consider general return processes, such that

$$dX_t^{(i)} = X_t^{(i)}\mu^i(s_t)dt + X_t^{(i)}\sigma^i(s_t)dB_t$$

and

$$d\beta_t = r(s_t)\beta_t dt$$

where s_t is the state variable (could be a vector) following Ito's process

$$ds_t = a(s_t)dt + b(s_t)dB_t$$

It is important to note that the market can be dynamically incomplete. Let σ be the volatility matrix for the vector process $X = (X^{(1)}, \cdots, X^{(N)})$, p be the dimension of B_t, and d be the dimension of state variables s_t. We can have $N < p$, while s_t has nonzero loadings on the $p - N$ leftovers of the Brownian motion. Then we don't have instruments to trade on the state s_t, which causes market incompleteness.

The optimization problem is

$$\begin{array}{l} \max\limits_{c_t,\theta_t} \;\; E[\int_0^\infty e^{-\rho t}u(c_t)dt] \\ s.t. \\ \frac{dW}{W} = (\theta_t \cdot (\mu - r\mathbf{1}) + r - \frac{c}{W})dt + \theta^T\sigma dB \end{array}$$

where I have omitted all t subscripts in the constraints for simplicity. Define indirect utility function

$$J(W_t, s_t) = \sup_{\{c_s,\theta_s\}_{s=t}^\infty} E_t[\int_t^\infty e^{-\rho(s-t)}u(c_s)ds]$$

Then it should satisfy the Bellman equation

$$\rho J(w,s) = \sup_{c,\theta}\left\{ \begin{array}{c} u(c) + J_w w(\theta_t \cdot (\mu - r\mathbf{1}) + r - \frac{c}{w}) + J_s \cdot a(s) \\ +\frac{1}{2}J_{ww}w^2\theta^T\sigma\sigma^T\theta + \frac{1}{2}tr(J_{ss}b(s)b^T(s)) + w\theta^T\sigma b(s_t)^T J_{sw} \end{array} \right\}$$

Then we can get the first order conditions for c and θ.

$$u'(c) - J_w = 0$$

$$J_w w(\mu - r\mathbf{1}) + J_{ww}w^2\sigma\sigma^T\theta + w\sigma b(s_t)^T J_{sw} = 0$$

Note that the equation for c is the continuous time envelope condition (need more work to see how the continuous-time envelope condition works). Then we can define the inverse marginal utility function as $G = (U')^{-1}$, and get $c^* = G(J_w)$. Moreover, the optimal portfolio choice should be

$$\theta = -\frac{J_w}{J_{ww}w}\left(\sigma\sigma^T\right)^{-1}(\mu - r\mathbf{1}) - \frac{J_w}{J_{ww}w}\left(\sigma\sigma^T\right)^{-1}\sigma b(s_t)^T\frac{J_{sw}}{J_w}$$

where the first term is the typical risk aversion component, and the second term is on hedging the changes in investment opportunities. The second part can be simplified as

$$\theta_2 = \left(\sigma\sigma^T\right)^{-1}\sigma b(s_t)^T\left(-\frac{J_{sw}}{J_{ww}w}\right)$$

With $J_w = u'(c)$, we can get

$$J_{ww} = u''(c)\frac{\partial c}{\partial w}$$

$$J_{sw} = u''(c)\frac{\partial c}{\partial s}$$

which results in

$$\theta_2 = \left(\sigma\sigma^T\right)^{-1}\sigma b(s_t)^T\left(-\frac{\left(\frac{\partial c}{\partial s}\right)}{w\frac{\partial c}{\partial w}}\right)$$

Suppose the value function J is concave in w. By concavity of u, we can get

$$\frac{\partial c}{\partial w} > 0$$

Suppose the asset returns are positively correlated to the state variable. Thus when $\partial c/\partial s < 0$, we have $\theta_2 > 0$, i.e. if the state is a good hedge (higher asset returns when consumption is lower), there is additional positive hedging demand.

When $u(c)$ is a log utility, we can get indirect utility function in the form of $J(W_t, s_t) = U(W_t) + F(s_t)$, which means that the cross derivative $J_{sw} = 0$. Thus there is no hedging demand for log utility, and portfolio choice can be easily solved, regardless of how many state variables we have. Moreover, consumption is also solved easily. The log-utility investor behaves as myopic.

Solution for CRRA Utilities

To get more intuition from the problem, we assume the following CRRA utility function

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}$$

where γ is the relative risk aversion. By the typical arguments for homothetic utility functions and scalable portfolio choice, we can get

$$J(W_t, s_t) = h(s_t)\frac{W_t^{1-\gamma}}{1-\gamma}$$

for some function $h(\cdot)$, which results in

$$wJ_w = h(s)w^{1-\gamma}, \quad w^2 J_{ww} = -\gamma h(s)w^{1-\gamma}, \quad wJ_{sw} = h'(s)w^{1-\gamma}$$

Then the FOCs for consumption and portfolio choice become

$$c^{-\gamma} - h(s)w^{-\gamma} = 0$$
$$h(s)w^{1-\gamma}(\mu - r\mathbf{1}) - \gamma h(s)w^{1-\gamma}\sigma\sigma^T\theta + \sigma b(s_t)^T h'(s)w^{1-\gamma} = 0$$

which results in

$$c = h(s)^{-1/\gamma}w$$
$$\theta = \frac{1}{\gamma}\left(\sigma\sigma^T\right)^{-1}\left((\mu - r\mathbf{1}) + \sigma b(s)^T\frac{h'(s)}{h(s)}\right)$$

and the Bellman equation

$$\rho h(s)\frac{w^{1-\gamma}}{1-\gamma} = h(s)^{-(1-\gamma)/\gamma}\frac{w^{1-\gamma}}{1-\gamma} + h(s)w^{1-\gamma}(\theta\cdot(\mu - r\mathbf{1}) + r - h(s)^{-1/\gamma}) + \frac{w^{1-\gamma}}{1-\gamma}h'(s)\cdot a(s)$$
$$-\frac{1}{2}\gamma h(s)w^{1-\gamma}\theta^T\sigma\sigma^T\theta + \frac{1}{2}\frac{w^{1-\gamma}}{1-\gamma}tr(Dh(s)b(s)b^T(s)) + \theta^T\sigma b(s)^T h'(s)w^{1-\gamma}$$

which simplifies into

$$\rho h(s) = h(s)^{-(1-\gamma)/\gamma} + h(s)(1-\gamma)(\theta(s)\cdot(\mu - r\mathbf{1}) + r - h(s)^{-1/\gamma}) + h'(s)\cdot a(s)$$
$$-\frac{1}{2}\gamma(1-\gamma)h(s)\theta^T\sigma\sigma^T\theta + \frac{1}{2}tr(Dh(s)b(s)b^T(s)) + (1-\gamma)\theta^T\sigma b(s)^T h'(s)c$$

Thus we get a second order differential equation system for $h(s)$, where the second order term comes from

$$\frac{1}{2}tr(Dh(s)b(s)b^T(s))$$

3.3 The Martingale Approach

In this part, we need market completeness so that we can replicate any final return with the traded assets. Specifically, let $N = p$, i.e. the number of assets equal the number of underlying diffusion processes. The optimization problem can be translated into

$$\begin{array}{l} \max\limits_{c} \quad E[\int_0^\infty e^{-\rho t}u(c_t)dt] \\ s.t. \\ E[\int_0^\infty \xi_t c_t dt] \le w \end{array}$$

from which we can easily get the optimal consumption policy as a function of the Lagrangian multiplier λ, i.e.

$$e^{-\rho t}u'(c_t) = \lambda\xi_t$$

where

$$\xi_t = \exp(-\int_0^t \eta_s dB_s - \frac{1}{2}\int_0^t \eta_s\cdot\eta_s ds)$$

and

$$\eta_t = \sigma(s_t)^{-1}(\mu(s_t) - r(s_t))$$

To get λ, we need to solve the equation that

$$E[\int_0^\infty \xi_t(s_t)c_t(\lambda, s_t)dt] = w$$

which is hard to evaluate. A typical approach is to get a differential equation of the left-hand side as a function of w and the state s,

3.4 Diffusion-Jump Processes with Constant Returns

Now we can introduce Poisson jumps into the model.

An Infinite Horizon Model with Constant Returns

We will start with a very simple model of a stock and a bond with default risk. The return of the stock is

$$\frac{dX_t}{X_t} = \mu dt + \sigma dB_t$$

and the return of the bond is

$$d\beta_t = r\beta_t dt - \beta_t dN_t$$

where N_t is a Poisson process with rate λ. Then the budget equation is

$$\frac{dW}{W} = (\theta(\mu - r) + r - \frac{c}{W})dt + \theta\sigma dB - (1-\theta)dN_t$$

Let the utility function be

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}$$

The property of the homothetic utility functions and scalable investments still holds, resulting in a conjecture of indirect utility

$$J(w) = A^{-\gamma}\frac{w^{1-\gamma}}{1-\gamma}$$

The HJB equation is

$$\rho J(w) = \sup_{c,\theta}\left\{u(c) + J_w w(\theta_t \cdot (\mu - r) + r - \frac{c}{w}) + \lambda\left(J(\theta w) - J(w)\right) + \frac{1}{2}J_{ww}w^2\theta^T\sigma\sigma^T\theta\right\}$$

where the optimization over c and θ can be written as

$$\max_{c>0}\left\{\frac{c^{1-\gamma}}{1-\gamma} - A^{-\gamma}w^{-\gamma}c\right\}$$

$$\max_{\theta}\left\{A^{-\gamma}(\mu - r)\theta + \lambda A^{-\gamma}\frac{\theta^{1-\gamma}}{1-\gamma} - \frac{1}{2}A^{-\gamma}\gamma\sigma^2\theta^2\right\}$$

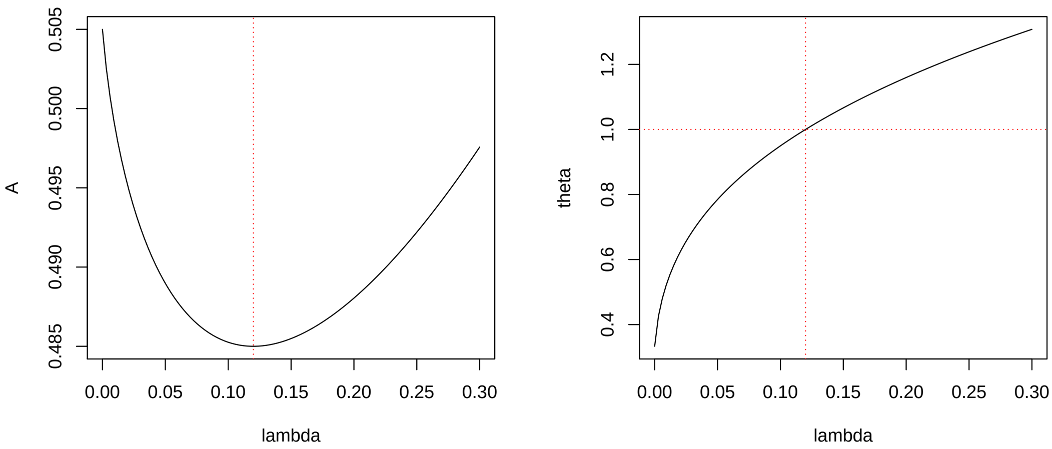
Thus the optimal portfolio choice and consumption are

$$c = Aw$$

$$\theta = \frac{1}{\gamma\sigma^2}\left(\mu - r + \lambda\theta^{-\gamma}\right) \tag{2}$$

With $\gamma > 1$, the right-hand-side of equation for θ decreases with the left-hand-side increases, and it is easy to see that there exists a unique solution θ^*. Moreover, θ^* increases with λ. Thus with higher default risks, we should invest less in bonds and more in stocks. Next, we can plug in the solution to the HJB equation to solve A,

$$\rho = A + (1-\gamma)(\theta^*(\mu - r) + r - A) + \lambda(\theta^{1-\gamma} - 1) - \frac{1}{2}\gamma(1-\gamma)\sigma^2(\theta^*)^2$$



(a) $A(\lambda)$ (b) $\theta(\lambda)$

Figure 1: A and θ as functions of λ

which results in

$$A = \frac{1}{\gamma}\left(\rho - (1-\gamma)(\theta^*(\mu - r) + r) - \lambda(\theta^{1-\gamma} - 1) + \frac{1}{2}\gamma(1-\gamma)\sigma^2(\theta^*)^2\right) \tag{3}$$

We find that λ indirectly enters A from θ^*, which is an increasing function of λ, and also directly enters in the expression. Next, I plot A as a function of λ in Figure 1. It is interesting to see that for λ larger than λ^*, the coefficient $A(\lambda)$ starts to increase with λ, contrary to our intuition that the larger risk associated with assets should result in less consumption and more pre-cautious savings. However, by studying (b) of Figure 1, we find $\theta(\lambda) > 1$ for λ greater than λ^*. This means that the investor is shorting the asset! Thus a default event is actually increasing investment returns, thus boosting consumption. In practice, we should properly restrict portfolio holdings to avoid this wired case.

A Finite Horizon Model with Constant Returns

To solve this puzzle, we can instead study a finite horizon problem. The objective function is

$$\max_{\theta,c} E[\int_0^T e^{-\rho t}u(c_t)dt + u(W_T)]$$

Then we can denote the indirect utility function as

$$J(w,t) = \max_{\{\theta_s,c_s\}_{s=t}^T} E[\int_t^T e^{-\rho(t-s)}u(c_s)ds|W_t = w]$$

Conjecture the indirect utility function as

$$J(w,t) = A(t)^{-\gamma}\frac{w^{1-\gamma}}{1-\gamma}$$

where $A(t) > 0$ for all $t \in [0, T]$. The HJB equation is

$$\sup_{c,\theta} \left\{ u(c) - \rho J + J_w w(\theta_t \cdot (\mu - r) + r - \frac{c}{w}) + \lambda \left(J(\theta w, t) - J(w, t) \right) + \frac{1}{2} J_{ww} w^2 \theta^T \sigma \sigma^T \theta + J_t \right\} = 0$$

which results in

$$u'(c) = J_w \quad \Rightarrow \quad c = A(t) w$$

and the optimization over θ as

$$\max_{\theta} \left\{ A(t)^{-\gamma} (\mu - r)\theta + \lambda A(t)^{-\gamma} \frac{\theta^{1-\gamma}}{1-\gamma} - \frac{1}{2} \gamma A(t)^{-\gamma} \sigma^2 \theta^2 \right\}$$

Then we can get FOC

$$\theta = \frac{\mu - r + \lambda \theta^{-\gamma}}{\gamma \sigma^2}$$

Then plug in the HJB equation to get

$$\frac{(A(t)w)^{1-\gamma}}{1-\gamma} - \rho A(t)^{-\gamma} \frac{w^{1-\gamma}}{1-\gamma} + A(t)^{-\gamma} w^{1-\gamma} (\theta(\mu - r) + r - A(t)) + \lambda(\theta^{1-\gamma} - 1) A(t)^{-\gamma} \frac{w^{1-\gamma}}{1-\gamma}$$

$$- \frac{1}{2} \gamma A(t)^{-\gamma} w^{1-\gamma} \sigma^2 \theta^2 - \gamma A(t)^{-\gamma-1} A'(t) \frac{w^{1-\gamma}}{1-\gamma} = 0$$

which can be simplified into

$$A'(t) = \frac{A(t)}{\gamma} \left(\gamma A(t) - \rho + (1-\gamma)(\theta(\mu - r) + r) + \lambda(\theta^{1-\gamma} - 1) - \frac{1}{2}(1-\gamma)\gamma\sigma^2\theta^2 \right)$$

and the boundary condition is

$$A(T) = 1$$

We can take this solution and let $T \to \infty$, then we find $A(0) \to A$, where A is the value in the infinite horizon case. This confirms the validity of both approaches.

Summary: The problem comes from the fact that as λ goes to infinity, the expected return of the bond goes to -1, and investors have full incentives to short the bond, which means a default event is actually beneficial to the investors. Only when the effective return of the bond remains positive, we have meaningful results. As a result, we should be very clear about whether the investor will "ride on the jump" or "suffer on the jump".

3.5 Diffusion-Jump Processes with Aggregate Uncertainty

Now we can deal with the general case with aggregate uncertainty. Time is infinite. Assume N risky assets and a risk-free asset. The N risky assets have price $X_t = (X_t^1, X_t^2, \cdots, X_t^N)$, with law of motion

$$dX_t^{(i)} = X_t^{(i)} \mu^i(s_t) dt + X_t^{(i)} \sigma^i(s_t) dB_t - X_t^{(i)} \kappa^i(s_t) dN_t, \;\; i \in \{1, \cdots, N\}$$

where N_t is a Poisson process with rate $\lambda(s)$. In vector form, the above is

$$\frac{dX_t}{X_t} = \mu(s_t) dt + \sigma(s_t) dB_t - \kappa(s_t) dN_t$$

where the division if element-wise. Denote the price of the risk-free asset as β_t, and the law of motion

$$d\beta_t = r(s_t)\beta_t dt$$

where s_t is the state variable (could be a vector) following Ito's process

$$ds_t = a(s_t)dt + b(s_t)dB_t + \delta(s_t)dN_t$$

Investors solve the following portfolio choice and consumption problem:

$$\begin{aligned}
&\sup_{\{c_t,\theta_t\}_{t=0}^{\infty}} E[\int_0^\infty e^{-\rho t}u(c_t)dt] \\
&s.t. \\
&\begin{cases} c \geq 0, \quad W \geq 0 \\ \frac{dW}{W} = \left(\theta(\mu - r\mathbf{1}) + r - \frac{c}{W}\right)dt + \theta^T\sigma dB - \theta^T\kappa dN \end{cases}
\end{aligned}$$

where I have omitted subscripts t in all the constraints to simplify notations. We can define the indirect utility function as

$$\begin{aligned}
&J(W_t, s_t) = \sup_{\{c_s,\theta_s\}_{s=t}^{\infty}} E_t[\int_t^\infty e^{-\rho(s-t)}u(c_s)ds] \\
&s.t. \\
&\begin{cases} c \geq 0, \quad W \geq 0 \\ \frac{dW}{W} = \left(\theta(\mu - r\mathbf{1}) + r - \frac{c}{W}\right)dt + \theta^T\sigma dB - \theta^T\kappa dN \end{cases}
\end{aligned}$$

Then the Bellman equation is

$$\rho J(w,s) = \sup_{c,\theta}\{u(c) + \mathcal{L}_{c,\theta}J(w,s)\}$$

where the infinitesimal generator is

$$\begin{aligned}
&\mathcal{L}_{c,\theta}J(w,s) = wJ_w\left(\theta(\mu - r\mathbf{1}) + r - \frac{c}{w}\right) + \frac{1}{2}w^2J_{ww}\theta^T\sigma\sigma^T\theta + J_s \cdot a(s) \\
&+\frac{1}{2}tr(J_{ss}b(s)b^T(s)) + w\theta^T\sigma b(s)^T J_{sw} + \lambda(s)\left(J(w(1-\theta^T\kappa), s+\delta(s)) - J(w,s)\right)
\end{aligned}$$

Thus the optimization over c is

$$\max_c \ \{u(c) - J_w c\}$$

and the optimization over θ is

$$\max_\theta \left\{wJ_w(\mu - r\mathbf{1})\cdot\theta + \frac{1}{2}w^2J_{ww}\theta^T\sigma\sigma^T\theta + w\theta^T\sigma b(s)^T J_{sw} + \lambda(s)J(w(1-\theta^T\kappa), s+\delta(s))\right\}$$

We note that $J(w,s)$ should be a strictly increasing function of w (with more w, we can immediately consume to boost up total utility), thus $J_w < 0$. Moreover, we impose the Inada condition on $u(c)$, so that the optimization over c has an interior solution

$$u'(c) = J_w$$

For the portfolio optimization, we would like to have a convex optimization. A sufficient condition would be $J(w,s)$ concave in w. Then we can solve θ with the following first order condition

$$wJ_w(\mu - r\mathbf{1}) + w\sigma b(s)^T J_{sw} + \frac{1}{2}w^2J_{ww}\sigma\sigma^T\theta = \lambda(s)J_w(w(1-\theta^T\kappa), s+\delta(s))w\kappa$$

The left-hand-side decreases in θ, while the right-hand-side increases in θ. When λ increases, the solution θ should increase. Thus the intuition that assets subject to jump risk should reduce its portfolio weight is robust. Note that when $\theta \cdot \mathbf{1} > 1$, the interpretation will be slightly different as now we are shorting risky assets.

Moreover, we find that the dynamics of the state variable s, including $a(s)$, $b(s)$ and $\delta(s)$ are all in the HJB equation, showing the importance of the aggregate uncertainty. To proceed, we need to impose specific forms of the utility function.

CRRA Utility Functions

Denote

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}$$

Then following the property of homogeneous utility functions, we can get

$$J(w, s) = h(s)\frac{w^{1-\gamma}}{1-\gamma}$$

Then

$$c = h(s)^{-1/\gamma} w$$

and

$$h(s)(\mu - r\mathbf{1}) + \sigma b(s)^T h'(s) - \frac{1}{2}\gamma h(s)\sigma\sigma^T\theta = \lambda(s)h(s+\delta(s))(1-\theta^T\kappa)^{1-\gamma}\kappa$$

The HJB equation is reduced to

$$\rho h(s) = h(s)^{-(1-\gamma)/\gamma} + h(s)(1-\gamma)\left(\theta(\mu - r\mathbf{1}) + r - h(s)^{-1/\gamma}\right) - \tfrac{1}{2}(1-\gamma)\gamma h(s)\theta^T\sigma\sigma^T\theta + h'(s)\cdot a(s)$$
$$+\tfrac{1}{2}tr(D^2h(s)b(s)b^T(s)) + (1-\gamma)\theta^T\sigma b(s)^T h'(s) + \lambda\left(h(s+\delta(s))(1-\theta^T\kappa)^{1-\gamma} - h(s)\right)$$

which is a second order ODE for $h(s)$. **Question again: What is the boundary condition for $h(s)$?**

Log Utility with Single Risky Asset and 2-D State Variable

Now we consider a even more special case, with only one risky asset

$$\frac{dX_t}{X_t} = \mu(s_t)dt + \sigma(s_t)dB_t - \kappa(s_t)dN_t$$

where B_t and N_t are both one dimensional. There are two state variables $s_t^{(1)}$ and $s_t^{(2)}$,

$$\frac{ds_t^{(1)}}{s_t^{(1)}} = a_1(s)dt + b_1(s)dB_t + \delta_1(s)dN_t$$

$$\frac{ds_t^{(2)}}{s_t^{(2)}} = a_2(s)dt + \delta_2(s)dN_t$$

With log utility, we can conjecture

$$J(w, s) = \frac{1}{\rho}\log(w) + h(s)$$

Then the Bellman equation is reduced to

$$\rho J(w, s) = \sup_{c,\theta}\left\{\begin{array}{l} \log(c) + \frac{1}{\rho}\left(\theta(\mu - r) + r - \frac{c}{w}\right) - \frac{1}{2}\frac{1}{\rho}\theta^T\sigma\sigma^T\theta + h'(s)\cdot a(s) \\ +\frac{1}{2}tr(D^2h(s)b(s)b^T(s)) + \lambda(s)\left(\frac{1}{\rho}\log(1-\theta^T\kappa) + h(s+\delta(s)) - h(s)\right) \end{array}\right\}$$

which leads to

$$c = \rho w$$

and

$$\mu - r - \frac{1}{2}\sigma^2\theta = \lambda(s)\frac{\kappa}{1-\theta\kappa}$$

$$\frac{1}{2}\sigma^2\kappa\theta^2 - (\frac{1}{2}\sigma^2 + \kappa(\mu - r))\theta + \mu - r - \kappa\lambda(s) = 0$$

where we need the solution with $\theta < 1/\kappa$. We can solve this second order equation to get

$$\theta = \frac{\frac{1}{2}\sigma^2 + \kappa(\mu - r) - \sqrt{\left(\frac{1}{2}\sigma^2 + \kappa(\mu - r)\right)^2 - 2\left(\mu - r - \kappa\lambda(s)\right)\sigma^2\kappa}}{\sigma^2\kappa}$$

Then plug in the Bellman equation, we get

$$\begin{array}{l} \rho h(s) = \log(\rho) + \frac{1}{\rho}\left(\theta(\mu - r) + r - \rho\right) - \frac{1}{2}\frac{1}{\rho}\sigma^2\theta^2 + h'(s)\cdot a(s) \\ +\frac{1}{2}tr(D^2h(s)b(s)b^T(s)) + \lambda(s)\left(\frac{1}{\rho}\log(1-\theta^T\kappa) + h(s+\delta(s)) - h(s)\right) \end{array}$$

With $b_1(s) = 0$, we can simplify the above into

$$\begin{array}{l} \rho h(s) = \log(\rho) + \frac{1}{\rho}\left(\theta(\mu - r) + r - \rho\right) - \frac{1}{2}\frac{1}{\rho}\sigma^2\theta^2 + \frac{\partial h(s)}{\partial s_1}a_1(s) + \frac{\partial h(s)}{\partial s_2}a_2(s) \\ +\frac{1}{2}\frac{\partial^2 h(s)}{\partial (s_1)^2}b_1^2(s) + \lambda(s)\left(\frac{1}{\rho}\log(1-\theta^T\kappa) + h(s+\delta(s)) - h(s)\right) \end{array}$$

which is a second order partial differential equation. Fortunately, it seems that $h(s)$ should be separable in states, because there is no cross-partial derivative.

3.6 The Utility Gradient Approach

A intuitive approach in asset pricing is to use the gradient of utility function to price assets. For example, for a consumption portfolio choice problem in discrete time, the pricing kernel from t to $t+1$ is

$$M_{t,t+1} = \beta\frac{u'(c_{t+1})}{u'(c_t)}$$

for utility

$$E[\sum_{t=0}^{\infty}\beta^t u(c_t)]$$

For a continuous time version, the pricing kernel at time t is

$$\xi_t = e^{-\rho t} u'(c_t)$$

for utility

$$E[\int_0^\infty e^{-\rho t} u(c_t) dt]$$

See Section C, Chapter 2 in Duffie (2010) for a proof of the discrete time case, and Appendix G of Duffie (2010) for a proof of the continuous time case

With the above results, we can use the pricing equation for the risk free asset to get

$$E_t[M_{t,t+1} R_t] = 1$$

in the discrete time model, or

$$E_t[d(\xi_t e^{\int_0^t r_s ds})] = 0$$

in the continuous time case. We can also write the above as

$$r_t dt = -E_t[\frac{d\xi_t}{\xi_t}]$$

In a model with money in the utility

$$E[\sum_{t=0}^{\infty} \beta^t u(c_t, m_t)] \quad \text{or} \quad E[\int_0^\infty e^{-\rho t} u(c_t, m_t)]$$

a similar result hold only if the utility on consumption and money follows the Cobb-Douglas form

$$u(c_t, m_t) = c_t^\alpha m_t^{1-\alpha}$$

because in this case the budget equation of

$$dw = (... - c - i \cdot m) dt + ...$$

implies a constant ratio

$$\frac{c}{i \cdot m} = const$$

Thus the projection of changes in portfolio onto the consumption is proportional to the chagnes in portfolio, which makes the proof get through.

4 Models of Bank Debt

The banking literature is huge and dates back more than a century ago. However, the rich finance literature on banking and bank debt is highly disconnected with the most recent macro finance literature that studies how financial frictions affect the real economy.

4.1 Market Liquidity and Bank Debt

Liquidity Creation and Banks (Diamond and Dybvig, 1983)

Two period model with liquidity service. Banks help households/firms to smooth consumption, since households/firms might have liquidity shocks that prefer early consumption. Without trade, each household does the following optimization

$$\max_{c_1(\tilde{\theta}),c_2(\tilde{\theta})} E[\tilde{\theta}u(c_1) + (1-\tilde{\theta})u(c_2)]$$

$$c_1 + \frac{c_2}{R} \leq 1$$

which results in

$$c_1(\tilde{\theta}=1) = 1, \quad c_1(\tilde{\theta}=0) = 0$$

$$c_2(\tilde{\theta}=1) = 0, \quad c_1(\tilde{\theta}=0) = R$$

Denote $P(\tilde{\theta}=1) = \phi$. Then we have

$$V^A = \phi u(1) + (1-\phi)u(R)$$

which denotes the expected utility in an autarky state.

However, if we have a continuum of households, then they can be better since some prefer early consumption while others prefer late consumption. Households are able to do cross insurance to improve social welfare.

Assume now households can trade AD securities at the beginning of period 0. Then the equilibrium should achieve Pareto optimality. Thus

$$\begin{array}{l} \max \phi u(c_1) + (1-\phi)u(c_2) \\ s.t. \\ \phi(c_1) + (1-\phi)\frac{c_2}{R} = 1 \end{array}$$

The the FOC is

$$\frac{u'(c_2)}{u'(c_1)} = \frac{1}{R}$$

Denote the optimal solution as c_1^* and c_2^*, which satisfy

$$\begin{cases} \phi c_1^* + (1-\phi)\frac{c_2^*}{R} = 1 \\ Ru'(c_2^*) = u'(c_1^*) \end{cases}$$

We would like to compare the above solution to $c_1 = 1$ and $c_2 = R$.

$$Ru'(R) - u'(1) = \int_1^R (xu''(x) + u'(x))dx$$

$$= \int_1^R u'(x)\left(x\frac{u''(x)}{u'(x)} + 1\right)dx$$

Note that when relative risk aversion of utility function is greater than 1, then

$$x\frac{u''(x)}{u'(x)} + 1 < 0$$

and

$$Ru'(R) - u'(1) < 0 = Ru'(c_2^*) - u'(c_1^*)$$

Denote the function

$$g(x) = Ru'(x) - u'(\frac{R - (1-\phi)x}{R\phi})$$

which decreases with x. We now know that

$$g(R) < g(c_2^*) = 0$$

Then we should have $R > c_2^*$. We also know that $1 < c_1^* < c_2^*$ from budget constraint. Then

$$R > c_2^* > c_1^* > 1$$

The Pareto optimality allocation has smaller difference between early consumption and later consumption, thus providing insurance to households.

Debt Optimizes Market Liquidity (DeMarzo and Duffie, 1999)

This paper uses a contract design approach to explain why debt is the optimal security. When we have information asymmetry, debt is the optimal security to issue. The retained equity is used to maximize incentives. Debt maximizes primary market liquidity.

The similar thought can be applied to the maturity of debt. Longer term debts are more subject to information asymmetry and moral hazards.

Debt Maximizes Secondary Market Liquidity (Gorton and Pennacchi, 1990)

Debt security maximizes secondary market liquidity.

4.2 Market Power of Banks

A classical paper by Keeley (1990) discusses the interaction of bank market power and the deposit insurance. Larger market power produces higher rents for banks and thus higher charter value, which makes the banks less likely to take risky strategies to explore the deposit insurance. However, when the market becomes more competitive, then banks have more incentives to explore the deposit insurance and thus take more risks. This paper thus has predictions on bank monopoly power and bank risk taking. The key implication is that banks with more monopoly power take lower risks. Question: Is this intuition still true in today's world?

Following the idea in Keeley (1990), Hellmann et al. (2000) argues that while capital-requirement regulation can induce prudent behavior, the policy yields Pareto-inefficient outcomes. Capital requirements reduce gambling incentives by putting bank equity at risk. However, they also have a perverse effect of harming banks' franchise values, thus encouraging gambling. Pareto-efficient outcomes can be achieved by adding deposit-rate controls as a regulatory instrument, since they facilitate prudent investment by increasing franchise values.

Next, Boyd and De Nicolo (2005) points out that the previous literature misses a point that the asset side riskiness will change with competition. When banks have less competition, although the higher charter value reduces risk taking incentives, the higher loan rate charged on their borrowers make borrower default more likely and thus riskier loans. This channel potentially reverses the effect of the charter value channel. A following up paper, Martinez-Miera and Repullo (2010) shows that when the effect pointed out by Boyd and De Nicolo (2005) is taken into account, a U-shaped relationship between competition and the risk of bank failure generally obtains.

As a first paper with international evidence, Beck et al. (2006) find that crises are less likely in economies with more concentrated banking systems, using data on 69 countries from 1980 to 1997. Thus regulatory policies and institutions that thwart competition are associated with greater banking system fragility. However, with a maximum of 279 banks across 48 countries, Laeven and Levine (2009) show that the same regulation has different effects on bank risk taking depending on the bank's corporate governance structure.

4.3 Equity Issuance and Equity Capital of Banks

Empirically, banks hold excess capital above regulatory requirements, which is inconsistent with the story for maximizing deposit insurance value. Allen et al. (2011) show that when credit markets are competitive, market discipline coming from the asset side induces banks to hold positive levels of capital as a way to commit to monitor and attract borrowers. When banks have more skin in the game, they have more incentives to monitor, which improves the lending performance of banks.

Remark 4. *The "competition-fragility" view of banking has opposite predictions to the monopoly power – intermediary asset pricing intuition. Under the "competition-fragility" view, banks with more monopoly power should have a higher charter value, which increases their incentives to issue new equity in time of distress. However, under the monopoly power – intermediary asset pricing intuition, banks with more monopoly power should have less incentives to issue new equity in time of distress, due to the conflict of new and old equity holders. These two interesting interactions might be useful to together provide a complete view on bank equity issuance, especially during a financial crises. In order to connect to empirics, I should collect a dataset on bank equity issuance, and relate it to banks' monopoly power.*

5 Models of Money

The basic models of money include three categories: (1) Money in the utility function, e.g. Sidrauski (1967). (2) Shopping time. (3) Cash in advance. Other less frequently used models might include money as storage, e.g. the Turnpike model by Townsend (1980) and Samuelson (1958).

5.1 Money in the Utility Models

The setup in this section is from Chapter 2 of Walsh (2017), which is called the Sidrauski (1967) model. Denote the utility function of the representative household as

$$U_t = u(c_t, z_t)$$

where z_t is the flow of services yielded by money holdings and c_t is the per capita consumption. Assume the utility function $u(c, z)$ is strictly increasing in both arguments, strictly concave, and continuously differentiable. A reasonable specification is

$$z_t = \frac{M_t}{P_t N_t} \equiv m_t$$

which is the per capita real holdings of money. The household is maximizing utility

$$\sum_{t=0}^{\infty} \beta^t u(c_t, m_t)$$

where $\beta \in (0, 1)$. The economy has money, bonds with nominal interest rate i_t, and physical capital with depreciation rate δ. Then the budget equation of the whole household sector at time t is

$$\underbrace{Y_t}_{\text{production}} + \underbrace{\tau_t N_t}_{\text{tax rebate}} + \underbrace{(1-\delta)K_{t-1}}_{\text{capital after depreciation}} + \underbrace{\frac{(1+i_{t-1})B_{t-1}}{P_t}}_{\text{bond payoff}} = C_t + K_t + \underbrace{\frac{M_t - M_{t-1}}{P_t}}_{\text{increase in money holding}} + \underbrace{\frac{B_t}{P_t}}_{\text{bond}}$$

Assume the total population grows at rate

$$\frac{N_{t+1}}{N_t} = 1 + n$$

To get a stationary equilibrium, we can normalize everything with population. Denote all variables in per capita unit in small letters. Then we can rewrite the budget equation as

$$y_t + \tau_t + (1-\delta)\frac{k_t}{1+n} + \frac{(1+i_{t-1})b_{t-1} + m_{t-1}}{(1+n)(1+\pi_t)} = c_t + k_t + m_t + b_t$$

where $b_t = B_t/(P_t N_t)$ is the per capita real bond holding and $m_t = M_t/(P_t N_t)$ is the per capita real money holding.

Thus the household problem is to optimize over c_t, k_t, b_t, and m_t subject ot the budget constraint. The Lagrangian of the household is

$$L = \sum_{t=0}^{\infty} \beta^t \left(u(c_t, m_t) + \lambda_t \left(y_t + \tau_t + (1-\delta)\frac{k_{t-1}}{1+n} + \frac{(1+i_{t-1})b_{t-1} + m_{t-1}}{(1+n)(1+\pi_t)} - c_t - k_t - m_t - b_t \right) \right)$$

$$
= \beta^t \left(\begin{array}{c} u(c_t, m_t) + \lambda_t \left(f(\frac{k_{t-1}}{1+n}) + \tau_t + (1-\delta)\frac{k_{t-1}}{1+n} + \frac{(1+i_{t-1})b_{t-1}+m_{t-1}}{(1+n)(1+\pi_t)} - c_t - k_t - m_t - b_t \right) \\ +\beta \left(u(c_{t+1}, m_{t+1}) + \lambda_{t+1} \left(\begin{array}{c} f(\frac{k_t}{1+n}) + \tau_{t+1} + (1-\delta)\frac{k_t}{1+n} + \frac{(1+i_t)b_t+m_t}{(1+n)(1+\pi_t)} \\ -c_{t+1} - k_{t+1} - m_{t+1} - b_{t+1} \end{array} \right) \right) \\ + \cdots \end{array} \right)
$$

for any interior solution, which can be guaranteed by assuming the Inada condition for both consumption and money holding. Note that the per capital production is

$$
y_t = f(\frac{k_{t-1}}{1+n})
$$

if we assume a production function that is constant return to scale,

$$
F(\lambda K_t, \lambda N_{t-1}) = \lambda F(K_t, N_{t-1})
$$

and define

$$
f(k) = F(k, 1)
$$

With the Lagrangian, first order conditions for c_t, m_t, b_t, and k_t are:

$$
u_c(c_t, m_t) = \lambda_t \qquad \text{(FOC of } c_t\text{)}
$$

$$
-\lambda_t + u_m(c_t, m_t) + \beta\lambda_{t+1}\frac{1}{(1+n)(1+\pi_{t+1})} = 0 \qquad \text{(FOC of } m_t\text{)}
$$

$$
-\lambda_t + \beta\lambda_{t+1}\frac{(1+i_t)}{(1+n)(1+\pi_{t+1})} = 0 \qquad \text{(FOC of } b_t\text{)}
$$

$$
-\lambda_t + \beta\lambda_{t+1}\left((1-\delta)\frac{1}{1+n} + \frac{1}{1+n}f'(\frac{k_t}{1+n})\right) = 0 \qquad \text{(FOC of } k_t\text{)}
$$

We can replace λ_t in the FOC of m_t by FOC of c_t to get

$$
\underbrace{u_m(c_t, m_t)}_{\text{direct benefit}} + \underbrace{\beta\frac{u_c(c_{t+1}, m_{t+1})}{(1+n)(1+\pi_{t+1})}}_{\text{indirect benefit}} = u_c(c_t, m_t)
$$

where the left hand side is the benefit of having additional unit of money at time t that includes both the direct benefit of money holding and the indirect benefit of increasing next period consumption, and the right hand side is the cost of having additional unit of money at time t, which is the cost of reducing consumption at time t. The above equation can be rewritten as

$$
\frac{u_m(c_t, m_t)}{u_c(c_t, m_t)} = 1 - \beta\frac{1}{(1+n)(1+\pi_{t+1})}\frac{u_c(c_{t+1}, m_{t+1})}{u_c(c_t, m_t)}
$$

where the ratio of $u_c(t+1)/u_c(t)$ can be derived from the FOC of b_t as

$$
\beta\frac{u_c(c_{t+1}, m_{t+1})}{u_c(c_t, m_t)}\frac{(1+i_t)}{(1+n)(1+\pi_{t+1})} = 1
$$

Combining the two, we have

$$
\frac{u_m(c_t, m_t)}{u_c(c_t, m_t)} = \frac{i_t}{1+i_t} \tag{4}
$$

Thus $i_t/(1+i_t)$ can be interpreted as the relative price of real money holding over the consumption goods. Equation (4) characterizes the demand for money as a function of consumption and nominal interest rate.

Remark 5. *To see how money demand changes with consumption and nominal interest rate, we use the following CES form of utility*

$$u(c,m) = \left(\xi c^{1-e} + (1-\xi)m^{1-e}\right)^{\frac{1}{1-e}}$$

As a result, we have

$$\frac{(1-\xi)}{\xi}\left(\frac{m}{c}\right)^{-e} = \frac{i}{1+i}$$

$$\Rightarrow m = \left(\frac{i}{1+i}\cdot\frac{\xi}{1-\xi}\right)^{-\frac{1}{e}} c$$

which decreases with interest rate i but increases with consumption c. Furthermore, when consumption and money are more substitutable, money demand decreases.

Finally, the government budget constraint is

$$\frac{B_t - B_{t-1}}{P_t} + \frac{M_t - M_{t-1}}{P_t} = \underbrace{\tau_t N_t}_{\text{totaltaxrebate}}$$

Steady State Equilibrium

Consider a steady state where the money growth rate is at rate θ and population growth $n = 0$. In the steady state, the real money balance should be a constant. Thus the inflation rate is just the money growth rate, i.e. $\pi^{ss} = \theta$. The equilibrium conditions can be written as

$$\frac{u_m(c^{ss}, m^{ss})}{u_c(c^{ss}, m^{ss})} = 1 - \beta\frac{1}{1+\pi^{ss}}$$

$$\beta\frac{1+i^{ss}}{1+\pi^{ss}} = 1$$

$$-1 + \beta f'(k^{ss}) + \beta(1-\delta) = 0$$

Thus the stationary nominal interest rate is

$$i^{ss} = \frac{1}{\beta}(1+\theta) - 1$$

and capital stock is

$$f'(k^{ss}) = \frac{1}{\beta} - 1 + \delta$$

With decreasing marginal return, we can get k^{ss} decrease in δ and increases in β. Thus more capital depreciation, less stationary capital stock. But more discounting, more capital stock. By the government budget constraint, we have

$$\tau^{ss} = \frac{\theta}{1+\pi^{ss}}m^{ss} = \frac{\theta}{1+\theta}m^{ss}$$

The budget constraint of the household implies

$$y^{ss} + (1-\delta)k^{ss} = c^{ss} + k^{ss}$$

$$\Rightarrow c^{ss} = f(k^{ss}) - \delta k^{ss}$$

Remark 6. *We find that only the real money balance enters all equations, which means an increase in the money supply will result in a one-to-one change in price level, the so-called "money neutrality". Moreover, the model has "super neutrality", since the growth rate of money θ has no effect on equilibrium capital stock and consumption. Similar super neutrality is present also in the shopping time models, where by assumption consumption equals to output that does not change over time.*

Existence of A Steady State Equilibrium

The equation that determines the existence of a steady state equilibrium is

$$\frac{u_m(c^{ss}, m^{ss})}{u_c(c^{ss}, m^{ss})} = \frac{i^{ss}}{1 + i^{ss}}$$

which is guaranteed if u is separable and u_m has a range of $(0, 1)$.

Next, we want to study the path that reaches the steady state equilibrium. Assume separable utility

$$u(c, m) = v(c) + \phi(m)$$

Then it is possible to have $c_t = c^{ss}$, $k_t = k^{ss}$, but real money balance m_t is changing. We want to discuss that scenario, where

$$\frac{\phi'(m_t)}{v'(c^{ss})} = \frac{i_t}{1 + i_t}$$

$$1 + i_t = \frac{1}{\beta}(1 + \pi_{t+1})$$

Assume money growth is constant such that $M_{t+1}/M_t = \theta$, then we have

$$\pi_{t+1} = \frac{P_{t+1}}{P_t} = \frac{M_t}{P_t}\frac{P_{t+1}}{M_{t+1}}\frac{M_{t+1}}{M_t} = \frac{m_t}{m_{t+1}}(1 + \theta)$$

As a result,

$$\frac{\phi'(m_t)}{v'(c^{ss})} = 1 - \frac{\beta}{1 + \frac{m_t}{m_{t+1}}(1 + \theta)}$$

$$\Rightarrow \frac{m_t}{\beta v'(c^{ss}) - v'(c^{ss}) + \phi'(m_t)}\left(v'(c^{ss}) - \phi'(m_t)\right) = \frac{m_{t+1}}{(1 + \theta)}$$

Optimal Monetary Policy and the Cost of Inflation

Since the steady state consumption is not affected by monetary policy, we want to maximize the real money holding. Suppose there exists a satiation point such that $u_m(c^{ss}, m^*) = 0$. Then the optimal monetary policy is to satiate money demand, which implies $i^{ss} = 0$. Thus again in this model, the Friedman rule describes the optimal monetary policy. Zero nominal interest rate and positive real interest rate implies deflation, i.e. inflation is negative.

To measure the welfare lost of monetary policy, we can use the demand curve of real money balance. The shade below the money demand curve is a measure of welfare loss due to positive nominal interest rate.

Remark 7. *I think the key reason of not going to a deflationary scenario in reality is because of disintermediation of the banking sector. Imagine the following case: the nominal interest rate is zero so that holding money has zero cost. In that case, households might just prefer holding money instead of having deposits in the banks. That will cause disintermediation of the financial sector, which dramatically reduces credit supply.*

Robustness of Superneutrality

In this model, the superneutrality seems a very strong result. Indeed, the empirical evidence of Barro (1995) implies that inflation has a negative effect on growth.

To account for the empirical evidence, we can add a labor choice. If the real money holding can affect the marginal utility on leisure, then the long-run labor supply and thus output will be affected. However, due to the lack of microfoundation in the MIU model, it is hard to argue why money should affect the marginal utility of leisure. The shopping time model has an answer: with larger real money holding, the shopping time is reduced and thus more time is devoted to working. But this channel is unlikely to be very significant in reality.

Another channel through which inflation can affect the stead-state stock of capital occurs if money enters directly into the production function (Fischer, 1974). Since the marginal product of capital (MPK) is determined by $1/\beta - 1 + \delta$, if $\partial \text{MPK}/\partial m > 0$, then with a larger real money balance, we should have a lower capital stock to achieve the same MPK. Thus more money supply reduces capital stock, which is to the contrary of the *Tobin effect*, proposed by Tobin (1965). Tobin (1965) argued that higher inflation will induce a portfolio substitution towards capital that would increase the stead-state capital-labor ratio.

A Stochastic Model with Labor

In the baseline MIU model, money is neutral and the growth rate of money affects the economy through real money holding, while the consumption and production are not affected. Next, we can introduce shocks into the model and study the impulse responses to monetary policy changes. To generate a connection between monetary policy and production, we have to include labor. Moreover, we assume zero population growth to simplify the model. Then the production function is

$$y_t = e^{z_t} k_{t-1}^{\alpha} n_t^{1-\alpha}$$

where z_t is a productivity shock that satisfies

$$z_t = \rho_z z_{t-1} + \zeta_t$$

Household utility function is

$$u(c_t, m_t, 1 - n_t) = \frac{\left(a c_t^{1-b} + (1-a) m_t^{1-b}\right)^{\frac{1-\phi}{1-b}}}{1-\phi} + \psi \frac{(1-n_t)^{1-\eta}}{1-\eta}$$

where each household has a unit of time to allocate between labor and leisure. The money growth satisfies

$$\theta_{t+1} - \theta^{ss} = \rho(\theta_t - \theta^{ss}) + \varepsilon_{t+1}$$

In this model, a positive money growth shock reduces output and employment, but increases nominal interest rate. However, the magnitude generated by this model is very small. A one standard deviation decrease in money growth implies less than 50 bps increase in output. Even with shocks, the model has money neutrality such that an instantaneous increase in money supply increases price level, but has no effects on the real economy. The adjustment on taxation or government bonds to balance government budget has no real effects as well. As a result, the basic model is not suitable to consider real impacts of short-term monetary policy.

Next, we can collect equations that determine the system. The household optimization problem will generate five first order conditions, on c_t, m_t, b_t, k_t, and n_t. We also have a resource constraint and production technology equation. Notice that due to lump-sum taxation and government budget balance, we do not need to consider taxation in any of the equations. The first oder conditions are

$$u_c(t) = \lambda_t \qquad \text{(FOC on } c_t\text{)}$$

$$-\lambda_t + u_m(t) + \beta \mathbb{E}_t[\lambda_{t+1} \frac{1}{(1+\pi_{t,t+1})}] = 0 \qquad \text{(FOC on } m_t\text{)}$$

$$-\lambda_t + \beta \mathbb{E}_t[\lambda_{t+1} \frac{1+i_{t,t+1}}{1+\pi_{t,t+1}}] = 0 \qquad \text{(FOC on } b_t\text{)}$$

$$-\psi(1-n_t)^{-\eta} + \lambda_t e^{z_t} k_{t-1}^{\alpha} n_t^{-\alpha} = 0 \qquad \text{(FOC on } n_t\text{)}$$

$$-\lambda_t + \beta \mathbb{E}_t[\lambda_{t+1}\left(1-\delta + e^{z_{t+1}} k_t^{\alpha} n_{t+1}^{1-\alpha}\right)] = 0 \qquad \text{(FOC on } k_t\text{)}$$

The next step is to log linearize the system to study its properties around the steady state. The basic rule is

$$f(x) = f(x^{ss}) + f'(x^{ss}) x^{ss} \hat{x}$$

where $\hat{x} = \log(x/x^{ss})$ is the log deviation. The explicit FOC on c_t is

$$u_c = \left(a c_t^{1-b} + (1-a) m_t^{1-b}\right)^{\frac{b-\phi}{1-b}} a c_t^{-b}$$

With linearizion, we have

$$\hat{u}_t = -b\hat{c}_t + (b-\phi) \frac{a (c^{ss})^{1-b} \hat{c}_t + (1-a)(m^{ss})^{1-b} \hat{m}_t}{a (c^{ss})^{1-b} + (1-a)(m^{ss})^{1-b}}$$

Define

$$\gamma = \frac{a (c^{ss})^{1-b}}{a (c^{ss})^{1-b} + (1-a)(m^{ss})^{1-b}}$$

Then

$$\hat{\lambda}_t = \hat{u}_t = \left(\gamma(b-\phi) - b\right) \hat{c}_t + (1-\gamma)(b-\phi)\hat{m}_t \qquad (5)$$

When the risk aversion is the same as the elasticity of substitution between real money holding and consumption, i.e. $b = \phi$, then consumption is not related to real money holdings. Next, the resource constraint implies

$$y^{ss}\hat{y}_t = c^{ss}\hat{c}_t + k^{ss}\hat{k}_t - (1-\delta)k^{ss}\hat{k}_{t-1}$$

which means production equals consumption plus investment. The linearized production function is

$$\hat{y}_t = z_t + \alpha \hat{k}_{t-1} + (1-\alpha)\hat{n}_t \qquad (6)$$

To facilitate discussions, we introduce the real interest rate $r_{t,t+1}$ and its log deviation $\hat{r}_{t,t+1}$. Then by definition,

$$\hat{r}_{t,t+1} = \hat{i}_{t,t+1} - \mathbb{E}_t[\hat{\pi}_{t,t+1}] \tag{7}$$

Next, we proceed to log-linearize the other four FOCs. The log-linearization of conditional expectation of λ_{t+1} is

$$\mathbb{E}_t[\lambda_{t+1}] = \mathbb{E}_t[\lambda^{ss}e^{\hat{\lambda}_{t+1}}] = \mathbb{E}_t[\lambda^{ss}(1+\hat{\lambda}_{t+1})] = \lambda^{ss} + \lambda^{ss}\mathbb{E}_t[\hat{\lambda}_{t+1}]$$

Thus the FOC on b_t implies

$$\lambda^{ss}(1+\hat{\lambda}_t) = \beta\lambda^{ss}(1+E_t[\hat{\lambda}_{t+1}])(1+r^{ss})(1+\hat{r}_{t,t+1})$$

$$\Rightarrow \hat{\lambda}_t = E_t[\hat{\lambda}_{t+1}] + \hat{r}_{t,t+1} \tag{8}$$

where I have used $1 + r^{ss} = 1/\beta$. Next, with both FOC on m_t and b_t we have

$$\frac{u_m(t)}{u_c(t)} = \frac{i_{t,t+1}}{1+i_{t,t+1}} \Rightarrow \frac{1-a}{a}\frac{m_t^{-b}}{c_t^{-b}} = \frac{i_{t,t+1}}{1+i_{t,t+1}}$$

$$\Rightarrow \log(\frac{1-a}{a}) - b\log(m_t) + b\log(c_t) = \log(1 - \frac{1}{1+i_{t,t+1}})$$

$$\Rightarrow -b\hat{m}_t + b\hat{c}_t = \frac{1}{i^{ss}(1+i^{ss})}(1+i^{ss})\hat{i}_t$$

$$\Rightarrow \hat{m}_t = \hat{c}_t - \frac{1}{b}\frac{1}{i^{ss}}\hat{i}_t \tag{9}$$

Thus the money holding is negatively related to the nominal interest rate. The sensitivity of money holding to nominal interest rate also depends on b, the substitution between consumption and real money holding.

Next, the log-linearized FOC on n_t is

$$-\eta\hat{\ell}_t = \hat{\lambda}_t + z_t + \alpha\hat{k}_{t-1} - \alpha\hat{n}_t$$

With $z_t + \alpha\hat{k}_{t-1} = \hat{y}_t - (1-\alpha)\hat{n}_t$ and $\ell^{ss}\hat{\ell}_t + n^{ss}\hat{n}_t = 0$, we get

$$(1 + \frac{n^{ss}}{\ell^{ss}}\eta)\hat{n}_t = \hat{\lambda}_t + \hat{y}_t \tag{10}$$

Then we log linearize FOC on k_t to get

$$r^{ss}(1+\hat{r}_{t,t+1}) = \alpha\frac{y^{ss}}{k^{ss}}\mathbb{E}_t[1 + \hat{y}_{t+1} - \hat{k}_t] - \delta$$

$$\Rightarrow r^{ss}\hat{r}_{t,t+1} = \alpha\frac{y^{ss}}{k^{ss}}\mathbb{E}_t[\hat{y}_{t+1} - \hat{k}_t]$$

The output deviation is

$$\hat{y}_{t+1} = z_{t+1} + \alpha\hat{k}_t + (1-\alpha)\hat{n}_{t+1}$$

Thus

$$r^{ss}\hat{r}_{t,t+1} = \alpha\frac{y^{ss}}{k^{ss}}\left(\rho_z z_t - (1-\alpha)\hat{k}_t + (1-\alpha)\mathbb{E}_t[\hat{n}_{t+1}]\right) \tag{11}$$

Finally, the real money growth is

$$\frac{m_t}{m_{t-1}} = \frac{1+\theta_t}{1+\pi_{t-1,t}}$$

which implies

$$\hat{m}_t - \hat{m}_{t-1} = \theta_t - \theta^{ss} - \hat{\pi}_{t-1,t} \tag{12}$$

5.1* A Continuous Time Money in the Utility Model

Money in the utility models are the workhorse of a broad class of monetary policy problems, and they are inclusive since transaction cost type models can be embedded in MIU models, and many results are quite similar under CIA and MIU models. In this extended section, I will set up a basic continuous time MIU model, as the baseline model in Tella (2018).

The economy is populated by a continuum of agents with log preferences over consumption c and real money holding $m = M/P$.

$$U(c, m) = E[\int_0^\infty e^{-\rho t} ((1-\beta)\log(c_t) + \beta \log(m_t))\, dt]$$

Money and consumption enter separately, so money will be superneutral, meaning that both the growth rate of money and the quantity of money will have no impact on consumption. Agents can continuously trade capital and use it to produce consumption $y_t = ak_t$, but with idiosyncratic capital quality shock $k_{i,t}\sigma dW_{i,t}$, where $W_{i,t}$ is an idiosyncratic Brownian motion. Here I include idiosyncratic risks as Tella (2018) because it is helpful to understand money as improving risk sharing. But we can just set $\sigma = 0$ to get back to the most basic model. Since idiosyncratic risks in the aggregate are washed away, the aggregate capital stock evolves according to

$$dk_t = (x_t - \delta k_t)dt$$

where x_t is investment. Then the aggregate resource constraint is

$$c_t + x_t = ak_t$$

We note that because investment transfers consumption with capital one-to-one, the price of capital is equal to 1. The total money stock evolves as

$$\frac{dM_t}{M_t} = \mu_t^M dt$$

and the central bank will target a constant inflation rate π through the choice of μ_t^M, which on the balanced growth path implies $\mu_M = \pi + \text{growth rate}$. The real interest rate will be $r_t = i_t - \pi$.

The total wealth is $w_t = k_t + m_t + h_t$, which includes the real value of future tax transfers

$$h_t = \int_t^\infty e^{-\int_t^s r_u du} \frac{dM_s}{p_s} ds$$

where we do not take expectation since there is no aggregate uncertainty in the model.

Remark 8. *In this model, we do not have government bonds, but households can borrow and lend to each other in a bond market. The government has a balanced budget such that*

$$\frac{dM_t}{p_t} = \textit{tax rebate at time } t$$

which shows the speciality of money that the government does not need to pay it back, but rather use the resources raised by money to rebate the agents. Since money is in the utility, it becomes a real wealth. We should treat it similarly as consumption. Money as a "real" asset that generates utility thus enters the aggregate wealth, just like capital. It is different from the "bubble money" version of Brunnermeier and Sannikov (2016).

Then the dynamics of the budget constraint of the agent is

$$dw_{i,t} = ((\alpha - \delta)k_{i,t} - \pi_t m_{i,t} + (w_{i,t} - k_{i,t} - m_{i,t})r_t - c_t)\, dt + k_{i,t}\sigma dW_{i,t}$$
$$= (r_t w_{i,t} + (\alpha - \delta - r_t)k_{i,t} - i_t m_{i,t} - c_t)\, dt + k_{i,t}\sigma dW_{i,t}$$

where in the second equation we have used the fisher equation $i_t = \pi + r_t$. To facilitate derivations, we can write the growth of wealth as

$$\frac{dw_{i,t}}{w_{i,t}} = \left(r_t + (\alpha - \delta - r_t)\frac{k_{i,t}}{w_{i,t}} - i_t\frac{m_{i,t}}{w_{i,t}} - \frac{c_{i,t}}{w_{i,t}}\right) dt + \frac{k_{i,t}}{w_{i,t}}\sigma dW_{i,t}$$

Note that this way of defining wealth has maintained the proportionality property. Another approach would be defining $\tilde{w}_t = k_t + m_t + d_t$, where the net debt across households is $d_t = 0$. Thus we effectively have $\tilde{w}_t = k_t + m_t$. The the budget dynamics of the households will be the exact same except a fixed tax rebate term added to the dt component, i.e.

$$d\tilde{w}_{i,t} = \cdots + \tau_t dt$$

subject to the natural debt limit

$$\underline{w}_{i,t} = -h_t$$

so that

$$\tilde{w}_{i,t} \geq \underline{w}_{i,t}$$

which destroys the nice property that everything is scalable with respect to wealth, since $\underline{w}_{i,t}$ is not proportional to $w_{i,t}$. Moreover, the added τ_t term also destroys the proportionality property.

Balanced Growth Path Equilibrium

Then household FOC on capital implies

$$\frac{\alpha - \delta - r_t}{\sigma^2} = \frac{k_{i,t}}{w_{i,t}}$$

FOC on money holding implies

$$\frac{m_{i,t}}{w_{i,t}} = \frac{\rho\beta}{i_t}$$

The FOC on consumption results in

$$\frac{c_{i,t}}{w_{i,t}} = (1 - \beta)\rho$$

Denote the volatility of wealth growth as σ_c, which in equilibrium is

$$\sigma_c = \frac{k_t}{k_t + m_t + h_t}\sigma = \frac{k_{i,t}}{w_{i,t}}\sigma = (1 - \lambda)\sigma$$

where we define the wealth share in money as λ.

A balanced growth path equilibrium with be scale invariant to the aggregate capital k_t. Thus we can normalize all aggregate variables by k_t, and define hat variables as the normalized ones, e.g. $\hat{m}_t = m_t/k_t$. The balanced growth path implies that the growth rate of individual wealth and consumption will be the same as growth of aggregate capital. Then we have

$$\frac{dk_t}{k_t} = (\hat{x}_t - \delta)dt$$

$$\frac{dw_{i,t}}{w_{i,t}} = \frac{dc_{i,t}}{c_{i,t}} = (\hat{x}_t - \delta)dt + \sigma_c dW_{i,t}$$

On the other hand, plugging in the money holding, consumption policy, and FOC on capital, we have

$$\frac{dw_{i,t}}{w_{i,t}} = \left(r_t + (\frac{k_{i,t}}{w_{i,t}}\sigma)^2 - \rho\beta - \rho(1-\beta)\right) dt + \sigma_c dW_{i,t}$$

$$= \left(r_t + \sigma_c^2 - \rho\right) dt + \sigma_c dW_{i,t}$$

Equating this with the balanced growth implied results, we have

$$r_t + \sigma_c^2 - \rho = \hat{x}_t - \delta \quad \Rightarrow \quad r_t = \rho + (\hat{x}_t - \delta) - \sigma_c^2$$

As a result, the typical Euler equation is not an independent equation for the model, but a equation coming from the wealth equation. It does not bring additional information since we do not know the growth rate of consumption in the equation. The additional information comes from the growth rate of consumption = growth rate of capital, which comes from equilibrium conditions. Furthermore, we notice that the Cobb-Douglas utility function is important in getting such a form because it implies a constant share of money and consumption, shown in the $-\rho\beta - \rho(1-\beta)$ part of dw/w.

Remark 9. *Why I do not need this assumption in a model without money? This is because in that scenario we only have capital as the aggregate wealth, which immediately implies that growth in wealth is the same as growth in capital. However, when we have both money and real wealth, there might be an unbalanced growth path (although it should converge to the balanced growth one) where wealth is not proportional to capital due to the non proportional real value of money. As a result, the "balanced growth equilibrium" restriction serves as an additional assumption that helps pin down the equilibrium.*

Next, from the FOC on money holding and consumption holding, we can eliminate $w_{i,t}$ and get

$$\frac{m}{c} = \frac{\beta}{1-\beta}\frac{1}{r+\pi}$$

We notice that once we solve λ, then all the ratios and interest rates will be solved. To get an equation for λ, we should use the last condition that

$$h_t = \int_t^\infty e^{-r_s(s-t)}\frac{dM_s}{p_s} = \int_t^\infty e^{-r_s(s-t)}\frac{d(p_s m_s)}{p_s}$$

$$= \int_t^\infty e^{-r_s(s-t)} m_s \pi_s ds + \int_t^\infty e^{-r_s(s-t)} dm_s$$

$$= \int_t^\infty e^{-r_s(s-t)} m_s \pi_s ds + \int_t^\infty r_s m_s e^{-r_s(s-t)} ds + \lim_{T\to\infty} m_T e^{-r_T(T-t)} - m_t$$

$$= \int_t^\infty e^{-r_s(s-t)} m_s i_s ds - m_t$$

On the balanced growth path, interest rates are not time varying. As a result, we should have

$$m_t + h_t = \int_t^{\infty} e^{-r(s-t)} e^{(\hat{x}-\delta)(s-t)} m_t i ds = \frac{m_t i}{r - (\hat{x} - \delta)}$$

Using the FOC on money

$$m_t i = \rho \beta w_t$$

We arrive at

$$\frac{m_t + h_t}{w_t} = \frac{\rho\beta}{r - (\hat{x} - \delta)} = \lambda$$

Finally, we use the expression for r and σ_c to get

$$\lambda = \frac{\rho\beta}{\rho - \sigma_c^2} = \frac{\rho\beta}{\rho - ((1-\lambda)\sigma)^2}$$

which can be used to solve λ, and thus other quantities. A brief discussion of the liquidity ratio λ is as follows.

- First, an increase in the utility share β in money will increase liquidity ratio λ.
- Second, λ decreases with discount ρ. Thus if agents discount future more, the liquidity ratio will be lower.
- Third, λ increases with idiosyncratic volatility σ.

Next, we note that if $\sigma = 0$, then we always have

$$\lambda = \beta$$

and interest rate is

$$r = a - \delta - (1-\beta)\sigma^2 = a - \delta$$

Investment is thus

$$\hat{x} = a - \rho + (1-\beta)\sigma^2 - (1-\beta)^2\sigma^2 = a - \rho$$

As a result, without idiosyncratic risks, investment and interest rates are not affected by money. Even with idiosyncratic risks, they are not affected by money and monetary policy as well. But they will respond to an increase in undiversifiable risks.

From the above discussions, we find that the key variable that makes a monetary economy ($\beta > 0$) different from a non-monetary economy is that the liquidity ratio is positive, i.e. $\lambda > 0$, which is associated with a positive interest rate $i > 0$, which is the liquidity premium of money. Suppose the liquidity premium is zero, i.e. $i_t = 0$, then we have $m_t + h_t = 0$, and thus $\lambda = 0$, which means that money does not matter at all.

Government Debt, Deposits and Ricardian Equivalence

We can change the model setup to include government bonds. Let b_t be the real value of government debt, and $d\tau$ be lump-sum tax transfers to households, where $\tau > 0$ means rebating the households. Then the government budget constraint is

$$db_t + \frac{dM_t}{p_t} = b_t(i_t^b - \pi) + d\tau_t$$

$$\frac{dM_t}{p_t} = dm_t + \pi m_t dt$$

where i_t^b is the nominal interest rate on government debt. The above equations implies

$$d\left(e^{-r_t t} b_t\right) = e^{-r_t t}\left(b_t(i_t^b - i_t)dt - dm_t - \pi m_t dt)\right) + e^{-r_t t} d\tau_t$$

$$= e^{-r_t t} b_t(i_t^b - i_t)dt - d(e^{-r_t t} m_t) - e^{-r_t t} i_t m_t dt + e^{-r_t t} d\tau_t$$

$$\Rightarrow -b_t = \int_t^\infty e^{-\int_t^s r_u du}(i_s^b - i_s)b_s ds + m_t - \int_t^\infty e^{-\int_t^s r_u du} i_s m_s ds + \int_t^\infty e^{-\int_t^s r_u du} d\tau_s$$

$$\Rightarrow b_t + m_t + \int_t^\infty e^{-\int_t^s r_u du} d\tau_s = \int_t^\infty e^{-\int_t^s r_u du}\left((i_s - i_s^b)b_s + i_s m_s\right) ds$$

Then the total wealth is

$$w_t = k_t + b_t + m_t + \int_t^\infty e^{-\int_t^s r_u du} d\tau_s = k_t + \int_t^\infty e^{-\int_t^s r_u du}\left((i_s - i_s^b)b_s + i_s m_s\right) ds$$

which clearly shows that the wealth gains from money and bonds come from the liquidity premium, which in equilibrium will be positive if and only if they are in the utility.

Next, we can add a banking sector, and denote the wealth of banking sector as w_t^b and the wealth of the household sector as w_t^h. Then the total wealth is

$$w_t = w_t^b + w_t^h$$

The banks pay dividend to the households. The budget of bankers is

$$dn_t = n_t r_t + d_t(i_t - i_t^d)dt - df_t$$

where df_t is the dividend payment to households. Then we get the present value of dividend as

$$v_t = \int_t^\infty e^{-\int_t^s r_u du} df_s = \int_t^\infty e^{-\int_t^s r_u du} d_s(i_s - i_s^d) ds$$

Thus the total wealth in this scenario is

$$w_t = k_t + b_t + m_t + v_t + \int_t^\infty e^{-\int_t^s r_u du} d\tau_s$$

$$= k_t + \int_t^\infty e^{-\int_t^s r_u du}\left((i_s - i_s^b)b_s + i_s m_s + d_s(i_s - i_s^d)\right) ds$$

5.2 Sticky Prices and Wages

The models in this subsection is from chapter 6 of Walsh (2017). We start with a MIU model with fixed capital stock and no price stickiness. Furthermore, we assume $b = \phi$ so that real money holding and consumption are separable and thus the only real impact of money growth should come from stickiness. The utility function is now

$$u(c, m, 1-n) = \frac{a}{1-\phi}c^{1-\phi} + \frac{(1-a)}{1-\phi}(\frac{m}{p})^{1-\phi} + \psi\frac{(1-n)^{1-\eta}}{1-\eta}$$

where I have used m to denote the nominal holding of money in order to study the price level determination later more conveniently. Then we have the following log-linearized system:

$$\hat{y}_t = (1-\alpha)\hat{n}_t + z_t$$

$$\hat{y}_t = \hat{c}_t$$

$$\hat{y}_t - \hat{n}_t = \underbrace{\hat{w}_t - \hat{p}_t}_{\text{real wage}}$$

which is the producer FOC on n_t. Then the household FOC on n_t is

$$\eta\frac{n^{ss}}{\ell^{ss}}\hat{n}_t + \phi\hat{c}_t = \hat{w}_t - \hat{p}_t$$

The consumption Euler equation is simplified into

$$\phi\mathbb{E}_t[\hat{c}_{t+1} - \hat{c}_t] = r_t$$

The demand for money is

$$\hat{m}_t - \hat{p}_t = \hat{c}_t - \frac{1}{bi^{ss}}\hat{i}_t$$

The Fischer equation is

$$\hat{i}_t = \hat{r}_t + \mathbb{E}_t[\hat{p}_{t+1}] - \hat{p}_t$$

The flexible optimal labor can be solved from the first five equations

$$\hat{n}^*_t = \frac{1-\phi}{1+\bar{\eta}+(1-\alpha)(\phi-1)}z_t = b_0 z_t$$

and optimal real wage is

$$\hat{\omega}^*_t = \frac{\bar{\eta}+\phi}{1+\bar{\eta}+(1-\alpha)(\phi-1)}z_t = b_1 z_t$$

with

$$\bar{\eta} = \eta\frac{n^{ss}}{1-n^{ss}}$$

We find that both labor supply and real wage increases with productivity shock z_t.

Next, let's change the wage setting into one period stickiness. The specific contract is that households demand a contract wage one period before in a competitive way (think of households as in competitive wage unions that ask for pre-set wage for the next period). Then the next period firms can choose employment (abundance of workers). The contract nominal wage thus satisfies

$$E_{t-1}[\hat{w}^c_t - \hat{p}_t] = \mathbb{E}_{t-1}[\hat{\omega}^*_t]$$

$$\Rightarrow \hat{w}_t^c = \mathbb{E}_{t-1}[\hat{\omega}_t^* + \hat{p}_t]$$

Next period employment will equal the marginal value of productivity, which implies

$$\hat{y}_t - \hat{n}_t = \hat{w}_t^c - \hat{p}_t$$

$$\Rightarrow \hat{n}_t = \hat{y}_t + (\hat{p}_t - \mathbb{E}_{t-1}[\hat{p}_t]) - \mathbb{E}_{t-1}[\hat{\omega}_t^*]$$

We note that the flexible-price real wage satisfies

$$\hat{\omega}_t^* = \hat{y}_t^* - \hat{n}_t^*$$

Thus

$$\hat{n}_t - E_{t-1}[\hat{n}_t^*] = \hat{y}_t - E_{t-1}[\hat{y}_t^*] + \hat{p}_t - E_{t-1}[\hat{p}_t]$$

Since

$$\hat{y}_t - E_{t-1}[\hat{y}_t^*] = (1-\alpha)(\hat{n}_t - E_{t-1}[\hat{n}_t^*]) + z_t - E_{t-1}[z_t]$$

We finally get

$$\hat{n}_t - E_{t-1}[\hat{n}_t^*] = \frac{1}{\alpha}(\hat{p}_t - E_{t-1}[\hat{p}_t]) + \frac{1}{\alpha}(z_t - E_{t-1}[z_t])$$

Thus an unexpected increase in prices makes the real wage lower than optimal, causing firms to expand employment. Furthermore, plugging this expression back into the productivity equation, we have

$$\hat{y}_t - E_{t-1}[\hat{y}_t^*] = \frac{(1-\alpha)}{\alpha}(\hat{p}_t - E_{t-1}[\hat{p}_t]) + \frac{1}{\alpha}(z_t - E_{t-1}[z_t])$$

Thus the output responses to monetary shocks.

5.3 Shopping Time Models

Compared with the MIU model, shopping time models are more primitive. Moreover, shopping time models provide microfundations for the MIU models and pin down the cross partial derivatives of labor and money. The setup in this subsection is from the chapter on fiscal-monetary theories of inflation in Ljungqvist and Sargent (2012), which adapts from Baumol (1952) and Tobin (1956).

Consider an endowment economy without uncertainty. A representative household has a unit of time distributed among labor and shopping time. A single good is produced at a constant amount $y > 0$ each period $t \geq 0$, divided between private consumption c_t and government spending g_t,

$$y = g_t + c_t$$

The preference of households is

$$\sum_{t=0}^{\infty} \beta^t u(c_t, \ell_t)$$

where $\beta \in (0,1)$, $c_t \geq 0$, $\ell_t \geq 0$, and $u(\cdot,\cdot)$ is increasing and concave in both arguments, with cross derivative $u''_{c,\ell} \geq 0$. The shopping time needed is

$$s_t = H(c_t, \frac{m_t}{p_t})$$

and the total time constraint is

$$\ell_t + s_t \leq 1$$

We assume $H, H_c, H_{cc}, H_{m/p,m/p} \geq 0$ while $H_{m/p}, H_{c,m/p} \leq 0$.

Household budget constraint

$$\underbrace{c_t}_{\text{consumption}} + \underbrace{\frac{b_{t+1}}{R_{t,t+1}}}_{\text{new bond holding}} + \underbrace{\frac{m_t - m_{t-1}}{p_t}}_{\text{increase in money holding}} = \underbrace{y}_{\text{endowment}} - \underbrace{\tau_t}_{\text{govt taxation}} + \underbrace{b_t}_{\text{bond payment}}$$

with additional restriction

$$m_t \geq 0$$

which means households cannot issue money. We do not constrain b_t and thus households can borrow and lend freely at rate $R_{t,t+1}$. The government has a budget constraint

$$g_t = \underbrace{\tau_t}_{\text{taxation}} + \underbrace{\frac{B_{t+1}}{R_{t,t+1}} - B_t}_{\text{new borrowing}} + \underbrace{\frac{M_t - M_{t-1}}{p_t}}_{\text{seigniorage}}$$

The Money Demand Function

Next, we will solve the money demand function from households. Denote the Lagrangian multiplier on the budget constraint as λ_t and on the time constraint as μ_t. For an interior solution (i.e. $m_t > 0$), we have

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \left\{ u(c_t, \ell_t) + \lambda_t \left(y - \tau_t + b_t - c_t - \frac{b_{t+1}}{R_{t,t+1}} - \frac{m_t - m_{t-1}}{p_t} \right) + \mu_t \left(1 - \ell_t - H(c_t, \frac{m_t}{p_t}) \right) \right\}$$

The first-order conditions with respect to c_t, ℓ_t, b_{t+1}, and m_t are

$$u_c(c_t, \ell_t) = \lambda_t + \mu_t H_c(c_t, \frac{m_t}{p_t})$$

$$u_\ell(c_t, \ell_t) = \mu_t$$

$$-\lambda_t \frac{1}{R_{t,t+1}} + \beta \lambda_{t+1} = 0$$

$$-\lambda_t \frac{1}{p_t} - \mu_t H_{m/p}(c_t, \frac{m_t}{p_t}) \frac{1}{p_t} + \beta \lambda_{t+1} \frac{1}{p_{t+1}} = 0$$

First, we can replace λ_t and get the following pricing equation

$$\beta \frac{u_c(t+1) - u_\ell(t+1) H_c(t+1)}{u_c(t) - u_\ell(t) H_c(t)} R_{t,t+1} = 1$$

Compared to a typical model without money, the marginal value of consumption is lower because there is a cost of consumption: consumption demands time. Then we can combine FOC on b_t and m_t to get

$$\lambda_t \left(\frac{p_t}{R_{t,t+1} p_{t+1}} - 1 \right) = \mu_t H_{m/p}(t)$$

We denote inflation as

$$\pi_{t,t+1} = \frac{p_{t+1}}{p_t}$$

and nominal interest rate as

$$R^{\$}_{t,t+1} = R_{t,t+1}\pi_{t,t+1}$$

Then we have

$$\lambda_t \left(1 - \frac{1}{R^{\$}_{t,t+1}}\right) = -\mu_t H_{m/p}(t) \tag{13}$$

which sets the cost of holding additional money balance to the benefit of holding additional money balance. Next, we can plug in the expressions for λ_t and μ_t to get

$$\left(\frac{u_c(t)}{u_\ell(t)} - H_c(t)\right)\left(1 - \frac{1}{R^{\$}_{t,t+1}}\right) + H_{m/p}(t) = 0$$

Expressing the above explicitly, we have

$$\left(\frac{u_c(c_t, \ell_t)}{u_\ell(c_t, \ell_t)} - H_c(c_t, \frac{m_t}{p_t})\right)\left(1 - \frac{1}{R^{\$}_{t,t+1}}\right) + H_{m/p}(c_t, \frac{m_t}{p_t}) = 0$$

We want to study how money demand m_t/p_t changes with consumption c_t and nominal interest rate $R^{\$}_{t,t+1}$.

We notice from (13) that the gross nominal interest rate is bounded below by 1, i.e. $R^{\$}_{t,t+1} \geq 1$. Next, by assumption, H_c is a decreasing function in m_t/p_t, while $H_{m/p}$ is an increasing function in m_t/p_t. Thus

$$h(\frac{m_t}{p_t}; c_t, R^{\$}_{t,t+1}) = \left(\frac{u_c(c_t, \ell_t)}{u_\ell(c_t, \ell_t)} - H_c(c_t, \frac{m_t}{p_t})\right)\left(1 - \frac{1}{R^{\$}_{t,t+1}}\right) + H_{m/p}(c_t, \frac{m_t}{p_t})$$

increases in m_t/p_t. Moreover, since H_c increases in c_t and $H_{m/p}$ decreases in c_t, we get h decrease in c_t. As a result, an increase in consumption c_t should increase the money demand m_t/p_t solved by $h = 0$. Furthermore, since h is an increasing function in $R^{\$}_{t,t+1}$, an increase in interest rate decreases money demand m_t/p_t. Define the money demand function as

$$\frac{m_t}{p_t} = F(c_t, R^{\$}_{t,t+1})$$

Then from above discussion, $F_c \geq 0$ but $F_{R^{\$}} \leq 0$.

Definition 3 (Equilibrium)**.** *Given an exogenous sequences $\{g_t, \tau_t\}_{t=0}^{\infty}$, initial values $B_0 = b_0$ and $M_{-1} = m_{-1} > 0$, an equilibrium is a price system $\{R_t, p_t\}_{t=0}^{\infty}$, a sequence of consumption $\{c_t\}_{t=0}^{\infty}$, a sequence of government borrowing $\{B_t\}_{t=0}^{\infty}$ and a positive sequence for money supply $\{M_t\}_{t=0}^{\infty}$, such that (1) Given the price system and taxation, the households optimum problem is solved with $b_t = B_t$, $m_t = M_t$, subject to household budget constraint and time constraint. (2) The government budget constraint is satisfied. (3) Resource constraint is satisfied $c_t + g_t = y$.*

In such a deterministic economy, we can discuss its "long-run" properties, i.e. the properties of a stationary equilibrium, with fixed interest rate, inflation, consumption, leisure and borrowing. Stationarity implies $\lambda_t = \lambda$, which leads to the real return on government bond

$$R_{t,t+1} = R = \frac{1}{\beta}$$

and money holding

$$m^* = \frac{m_t}{p_t} = F(c, R^*) = F(y - g, R^*) = f(\pi R)$$

where we have suppressed the constants c in the last step. Using government budget equation, we have

$$g = \tau + \frac{B}{R} - B + \frac{\pi - 1}{\pi} m^*$$

Thus

$$g - \tau + B(1 - \beta) = \frac{\pi - 1}{\pi} f(\frac{\pi}{\beta}) \tag{14}$$

Note that since we do not have production side of the economy, aggregate consumptions are the same. Thus variations in welfare come from the time devoted to shopping. The equation might result in multiple equilibria because $(\pi - 1)/\pi$ increases in π while $f(\pi/\beta)$ decreases in π. The least inflation equilibrium is preferred because in that case real money holding is maximized, which reduces shopping time and increases leisure time.

Finally, we need to consider the starting value of the equilibrium. The initial money holding should satisfy

$$g = \tau_0 + \beta B - B_0 + f(\frac{\pi}{\beta}) - \frac{M_{-1}}{p_0} \tag{15}$$

which should pin down initial price level p_0.

Monetary Doctrines

- Sustained deficits cause inflation.

  Suppose we are at a low inflation equilibrium. Then inflation increases with government spending. However, if we are at a high inflation equilibrium, then inflation might decrease with government spending.

- Zero inflation policy.

  Putting $\pi = 1$, we get

  $$B = \frac{R}{R - 1}(g - \tau) = \sum_{t=0}^{\infty} R^{-t}(g - t)$$

  As a result, the current value of government borrowing should equal to its discounted value of payoffs. Note: we do not need to consider future borrowing because they will cancel out if we combine all the budget constraints. With transversality condition, the only thing that matters if repayment in each period. The above exercise also shows that inflation acts as a taxation. Without inflation taxation, the government has to repay its debt purely from taxation.

- Unpleasant monetarist arithmetic.

  Suppose we do an open market operation at period 1 that increases bonds B, but reduces money M_t. Then from equation (14), we find that inflation is increased by such a open market operation, in a low inflation equilibrium. However, the initial price level effect could go either way.

  However, we notice that the above claim depends on the fact that the government should keep the amount of government borrowing fixed, which implies a decreasing of borrowing over time. If B increases, that means the decrease in government borrowing should be larger, which has to be supplemented by higher inflation. Once we have a growing government borrowing, e.g. $Be^{\gamma t}$, a permanent increase in government borrowing should reduce inflation.

- Quantity theory of money.

  By changing initial money supply M_{-1}, the only adjustment comes from initial price level p_0, by the same multiple.

- Optimal monetary policy and the Friedman rule.

  The basic idea is to satiate money demand until the maximum level. This is not equivalent to "increasing money supply to a satiated point" because what matters is the real money supply. Suppose we have a satiation point such that $H_{m/p}(c, m/p) = 0$. Then from the first order condition that sets the cost of holding money to the benefit of holding money, the cost of holding money should be zero, i.e.

  $$\lambda_t \left(1 - \frac{1}{R^{\$}}\right) = 0$$

  With utility strictly increasing in consumption, we always have $\lambda_t > 0$, and thus the nominal interest rate $R^{\$} = 1$, which implies an inflation rate

  $$\pi = \beta < 1$$

  Thus the economy has strict deflation. This is so-called "Friedman rule". The key idea is to satiate real money demand.

- A fiscal theory of the price level.

  This is a different interpretation of the equations. Previously we assume that the government control the money supply process and adjust its borrowing to satisfy government spending. Alternatively, we can interpret the government as directly setting the nominal interest rate $R^{\$}$, which avoids the equilibrium multiplicity problem. Given government spending and taxation, the value of government bond is just its discounted direct taxation income plus inflation taxation, minus discounted government spending. From the initial price equation, we have

  $$\frac{M_{-1}}{p_0} = \tau_0 + \beta B - B_0 + f(\frac{\pi}{\beta}) - g$$

  Given money supply, the price level p_0 has to adjust to satisfy the equation.

- Exchange rate (in)determinacy.

  Consider a two country version of the model. Denote $M_{i,t}$ as the money supply in country $i \in \{1,2\}$ at time t. Both countries have the same consumption goods and there is no barriers of trade. Thus the real exchange rate is pinned down by purchasing power parity, which implies

  $$p_{1,t} = e_t p_{2,t}$$

  We note that by no arbitrage, different currencies should have the same return,

  $$\frac{p_{1,t+1}}{p_{1,t}} = \frac{p_{2,t+1}}{p_{2,t}} \tag{16}$$

  which implies

  $$e_{t+1} = e_t = e$$

  is a constant. Because of international trade, we cannot use independent budget constraints of each government. Instead, we have to use an aggregate budget constraint,

  $$g - \tau + B(1-\beta) = \frac{\pi - 1}{\pi} f(\frac{\pi}{\beta})$$

  where I have already used the result that $R = 1/\beta$ in a stationary equilibrium and inflation rates of two countries are the same. At time 0, the aggregate budget constraint is

  $$g = \tau + \beta B - B_0 + f(\pi R) - \frac{M_{1,-1} + eM_{2,-1}}{p_{1,-1}}$$

  With this equation alone, we cannot pin down both $p_{1,-1}$ and e, which means the exchange rate is indeterminate.

5.4 Cash-in-Advance (CIA) Models

We have two different timing conventions, the Lucas (1982) timing, and the Svensson (1985) timing.

- Lucas timing: Shocks realized. Then decide money and consumption together.
- Svensson: Shocks are realized. However, only beginning of period money holding can be used for consumption. After consumption, new money holding can be decided. This setup can be motivated by the "shopping time" example, where a household is divided into a worker and a shopper. Shopper can only use predetermined money level to make consumption.

Lucas Timing

The basic setup is as follows

$$\max_{\{C_t, M_t, B_t\}} E[\sum_{t=0}^{\infty} u(C_t)]$$
$$s.t.$$
$$\begin{cases} \underbrace{P_t Y_t}_{\text{endowment}} + \underbrace{(M_{t-1} - P_{t-1}C_{t-1})}_{\text{money leftover}} + \underbrace{B_{t-1}}_{\text{bond payment}} = \underbrace{M_t}_{\text{new money holding}} + \underbrace{Q_t B_t}_{\text{new bond holding}} \\ P_t C_t \leq M_t \end{cases}$$

Note: Consumption should only appear once, either in today's immediate consumption or tomorrow. The notation B_t means holding of $t \to t+1$ bond with total face value of B_t, and $B_t \in m\mathcal{F}_t$. Furthermore, the money holding $M_t \in m\mathcal{F}_t$, where

$$\mathcal{F}_t = \{\sigma(Y_s),\ s \leq t\}$$

Svensson Timing

The basic setup is as follows

$$\max_{\{C_t, M_{t+1}, B_t\}} E[\sum_{t=0}^{\infty} u(C_t)]$$
$$s.t.$$
$$\begin{cases} \underbrace{P_t Y_t}_{\text{endowment}} + \underbrace{(M_t - P_t C_t)}_{\text{money leftover}} + \underbrace{B_{t-1}}_{\text{bond payment}} = \underbrace{M_{t+1}}_{\text{new money}} + \underbrace{Q_t B_t}_{\text{new bond holding}} \\ P_t C_t \leq M_t \end{cases}$$

Notice that the notation is very bad. At time t, we should make choices that are measurable with respect to $\mathcal{F}_t$. Let me change the formulation into the following by replacing M_{t+1} into M_t,

$$\max_{\{C_t, M_t, B_t\}} E[\sum_{t=0}^{\infty} u(C_t)]$$
$$s.t.$$
$$\begin{cases} \underbrace{P_t Y_t}_{\text{endowment}} + \underbrace{(M_{t-1} - P_t C_t)}_{\text{money leftover}} + \underbrace{B_{t-1}}_{\text{bond payment}} = \underbrace{M_t}_{\text{new money}} + \underbrace{Q_t B_t}_{\text{new bond holding}} \\ P_t C_t \leq M_{t-1} \end{cases}$$

Even if the CIA constraint is binding, we can still get time varying money multiplier.

Remark 10. *A trick to remember the timing difference between Lucas timing and Svensson timing is that the only difference comes from the consumption constraint, where Svensson timing uses money decided in the last period for consumption.*

A Deterministic Model

I will set up a deterministic general equilibrium model with only consumption in the utility as a benchmark. The timing follows Svensson (1985). Consider a representative household that maximizes

$$\sum_{t=0}^{\infty} \beta^t u(c_t)$$

subject to the budget equation

$$P_t w_t \equiv (1 + i_{t-1,t}) B_{t-1} + (1-\delta) P_t k_{t-1} + P_t f(k_{t-1}) + M_{t-1} + T_t \geq P_t c_t + P_t k_t + B_t + M_t$$

Define inflation

$$1 + \pi_{t-1,t} = \frac{P_t}{P_{t-1}}$$

Then divide both sides of the budget equation by the price level P_t to get

$$w_t \equiv \frac{(1 + i_{t-1,t}) b_{t-1}}{1 + \pi_{t-1,t}} + (1-\delta) k_{t-1} + f(k_{t-1}) + \frac{m_{t-1}}{1 + \pi_{t-1,t}} + \tau_t \geq c_t + k_t + b_t + m_t$$

where w_t is the period t resources for spending. With the Svensson timing, we have the following transaction constraint

$$c_t \leq \frac{M_{t-1}}{P_t} + \frac{T_t}{P_t} = \frac{m_{t-1}}{1+\pi_{t-1,t}} + \tau_t$$

To have a sense on the holding cost of money, we can define nominal financial asset holding as $a_t = m_t + b_t$. Then

$$w_t = \frac{(1+i_{t-1,t})a_{t-1}}{1+\pi_{t-1,t}} - \frac{i_{t-1,t}m_{t-1}}{1+\pi_{t-1,t}} + (1-\delta)k_{t-1} + f(k_{t-1}) + \frac{m_{t-1}}{1+\pi_{t-1,t}} + \tau_t$$

In terms of $t-1$ resources, the holding cost is

$$\frac{i_{t-1,t}m_{t-1}}{1+\pi_{t-1,t}}\frac{1}{1+r_{t-1,t}} = m_{t-1}\frac{i_{t-1,t}}{1+i_{t-1,t}}$$

which is the same as MIU models. Next, let the Lagrangian multiplier on budget equation be $\beta^t\lambda_t$ and the Lagrangian multiplier on the CIA constraint as $\beta_t\mu_t$. Then the Lagrangian is

$$L = \beta^t \begin{pmatrix} u(c_t) + \lambda_t\left(\cdots + \tau_t - c_t - k_t - b_t - m_t\right) \\ +\mu_t\left(\frac{m_{t-1}}{1+\pi_{t-1,t}} + \tau_t - c_t\right) \\ +\beta\lambda_{t+1}\left(\frac{(1+i_{t,t+1})b_t}{1+\pi_{t,t+1}} + (1-\delta)k_t + f(k_t) + \frac{m_t}{1+\pi_{t,t+1}} + \cdots\right) \\ +\beta\mu_{t+1}\frac{m_t}{1+\pi_{t,t+1}} + \cdots \end{pmatrix} + \cdots$$

The first order conditions over c_t, k_t, b_t and m_t are

$$u'(c_t) - \lambda_t - \mu_t = 0$$

$$-\lambda_t + \beta\lambda_{t+1}\left(1-\delta+f'(k_t)\right) = 0$$

$$-\lambda_t + \beta\lambda_{t+1}\frac{1+i_{t,t+1}}{1+\pi_{t,t+1}} = 0$$

$$-\lambda_t + \beta\lambda_{t+1}\frac{1}{1+\pi_{t,t+1}} + \beta\mu_{t+1}\frac{1}{1+\pi_{t,t+1}} = 0$$

With the first and the fourth equation, we have

$$\beta\frac{u'(c_{t+1})}{u'(c_t)}\frac{1}{1+\pi_{t,t+1}} + \frac{\mu_t}{u'(c_t)} = 1$$

which is the pricing equation for money. We notice that the second term is due to the relaxation effect on the transaction constraint. The fourth equation can be also written as

$$\beta\frac{\lambda_{t+1}+\mu_{t+1}}{\lambda_t} = 1 + \pi_{t,t+1}$$

With the the third equation

$$\lambda_t = \beta\lambda_{t+1}\frac{1+i_{t,t+1}}{1+\pi_{t,t+1}}$$

We finally get

$$i_{t,t+1} = \frac{\mu_{t+1}}{\lambda_{t+1}}$$

Thus the nominal interest rate is the ratio between the shadow value of the CIA constraint and the shadow value of the budget equation. Alternatively, we should interpret that the interest rate is positive if and only if the CIA constraint is binding. We can rewrite the marginal utility equation as

$$u'(c_t) = \lambda_t + \mu_t = \lambda_t(1 + i_{t-1,t}) > \lambda_t$$

Thus the marginal utility of consumption exceeds the social cost of producing it.

Next, we want to study the steady state properties. Again, we have

$$f'(k^{ss}) = \frac{1}{\beta} - 1 + \delta$$

$$c^{ss} = f(k^{ss}) - \delta k^{ss}$$

which implies that money has nothing to do with consumption and capital stock in the steady state. Furthermore, the CIA constraint is always binding, which means $m^{ss} = c^{ss}$. Thus the steady state money holding is not related to money growth. I call this scenario as "hyper neutrality of money".

However, the "hyper neutrality of money" will not survive if we have both cash goods and credit goods, as proposed by Lucas and Stokey (1983). The cash good is subject to the CIA constraint but the credit good is not. As a result, inflation distorts the relative margin of credit good and cash good.

A Stochastic Model with Labor

The "hyper neutrality of money" in the simple deterministic CIA framework is not robust when we add labor and randomness. In general, we should view inflation as taxation on money. Once other factors enter household utility and money holding affects the relative marginal utility of different factors, inflation should have an impact on utility. The stochastic setting makes money velocity time-varying and allows us to study the impulse response to shocks.

The representative household sector has the following utility function

$$E[\sum_{t=0}^{\infty} \beta^t \left(\frac{c_t^{1-\phi}}{1-\phi} + \psi \frac{(1-n_t)^{1-\eta}}{1-\eta} \right)]$$

subject to the same budget constraint as in the deterministic model. But now the production function becomes

$$y_t = e^{z_t} k_{t-1}^{\alpha} n_t^{1-\alpha}$$

where

$$z_t = \rho_z z_{t-1} + \zeta_t$$

is an AR(1) process and the money growth rate $\theta_t = M_t/M_{t-1}$ satisfies

$$\theta_t - \theta^{ss} = \rho_\theta(\theta_{t-1} - \theta^{ss}) + \phi z_{t-1} + \varepsilon_t$$

Then we can solve for the steady state:

$$y^{ss} = (k^{ss})^{\alpha}(n^{ss})^{1-\alpha}$$

$$\alpha \frac{y^{ss}}{k^{ss}} = \frac{1}{\beta} - 1 + \delta$$
$$c^{ss} = y^{ss} - \delta k^{ss}$$
$$c^{ss} = m^{ss}$$

The first order condition for labor implies

$$\psi(1 - n^{ss})^{-\eta} = (1 - \alpha)\frac{y^{ss}}{n^{ss}}\lambda^{ss}$$

The interest rate satisfies

$$(c^{ss})^{-\phi} = \lambda^{ss} + \mu^{ss} = (1 + i^{ss})\lambda^{ss}$$

and

$$1 + i^{ss} = \frac{1 + \theta^{ss}}{\beta}$$

where I have used the result $i = \mu/\lambda$. The above equations are then enough to solve steady state variables $\{y^{ss}, k^{ss}, c^{ss}, m^{ss}, n^{ss}, \lambda^{ss}, i^{ss}\}$. After solving the steady state, we will find that the elasticity of labor supply w.r.t. money growth rate

5.5 Summary of Money Neutrality

In all of the money models with flexible price, including MIU, CIA, shopping time etc, we have money **neutrality**, where *the level of the money supply* does not influence real variables such as consumption, investment and production.

Superneutrality of money is a stronger property than neutrality of money. It holds that not only is the real economy unaffected by the level of the money supply but also that the *rate of money supply growth* has no effect on real variables. In the MIU model, if consumption, labor and money holding are separable in the utility function, then superneutrality holds. In CIA models and shopping time models, if labor supply is fixed, then we also have superneutrality.

The classical dichotomy is the idea, attributed to classical and pre-Keynesian economics, that real and nominal variables can be analyzed separately. This statement is much stronger than superneutrality, and it means we can separate the equation system that determines the real variables completely from money.

5.6 New Keynesian Monetary Models

The monetary policy channel of New Keynesian monetary models mainly comes from two aspects: (1) Sticky prices. (2) Sticky wages. A typical New Keynesian model has sticky price, where firms engage in monopolistic competition and set prices that are sticky. The sticky price implies sticky inflation. Thus a change in nominal interest rate implies changes in real interest rate, which connects to investment and consumption. A smaller interest rate will result more consumption today in the expense of future consumption.

The sticky wage assumption also has a different implication on the optimal inflation target. According to the Friedman rule, the optimal inflation is actually deflation that yields a zero nominal interest rate. However, in models with sticky prices, the inefficiency comes from price dispersion or price level changes that leads to the obsolesce of existing prices. Thus an optimal inflation rate is zero, implying nominal interest rate equal to real interest rate.

The Basic Model

The basic model is adding price stickiness and monopolistic competition to an MIU model. To put all effects into the price stickiness, the baseline MIU model has separable utility function in consumption, money and labor. Thus without price stickiness, the model exhibits the classical dichotomy.

Household maximize the expected discounted utility:

$$E[\sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\sigma}}{1-\sigma} + \frac{\gamma}{1-b} (\frac{M_t}{P_t})^{1-b} - \chi \frac{N_t^{1+\eta}}{1+\eta} \right)]$$

The consumption good is composite of different types of consumption produced by monopolistically competitive final final goods producers.

$$C_t = \left(\int_0^1 c_{j,t}^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}}$$

where the elasticity of substitution $\theta > 1$ so that the goods are substitutes. As typical in solving monopolistic competition problems, we can solve the consumption choice in two steps. First, given a level of C_t, households choose the allocation among different consumption to minimize the total cost. Second, given the cost of achieving any level C_t, the household chooses C_t, N_t and M_t optimally.

5.7 Monetary Policy Transmission

The Keynesian school argues that monetary policy changes "aggregate demand". The money demand curve shifts with different monetary policy, changing the real interest rate of the economy, and finally investment and consumption. In this sense, the New Keynesian model and the old Keynesian school have the same basic arguments. However, the old Keynesian school has a very reduced form "money demand" function which best corresponds to the MIU, shopping time, or CIA models. But we notice that the transaction side of monetary policy is becoming less favorable and struggles to generate a large impact. The New Keynesian model instead builds its channel through sticky prices or wages. Even if we have no transaction benefit of money, the sticky price or wage will deliver monetary policy impacts.

The bank credit view

The "money view" is named in

6 Models of Open Market Operations

In this section, I will review models of open market operations. I do not use the name of “monetary policy” because monetary policy also includes other aspects, such as discount window lending and reserve requirements, although open market operations is the most common one.

Here is my personal view of open market operations. The world is composed of five sectors: The Federal reserve, non-commercial banks, commercial banks, households, and the government, where the government is taken as exogenous. The federal reserve is doing open market operations to inject or withdraw liquidity from the commercial banking sector.

Mechanically, any open market sell of treasury securities should decrease deposits at least by the same amount. This is indeed true when the banking sector is fully competitive. This sets the lower bound on deposit outflows. Furthermore, according to the traditional theory on bank reserve requirement channel, the deposit outflow sensitivity should be at most the inverse reserve ratio. I think in a reasonable model, I can have sensitivities ranging from 1 to inverse reserve ration. Because of bank monopoly power, the sensitivity should be greater than 1.

On the other hand, government bonds are liquid securities close to money with only government default risks. Typically bank deposits in large denominations have higher returns (measure the wholesales short-term funding rate of banks, or commercial paper rates. We already know that banker acceptance has a higher interest rate than bonds, implying a higher transactional costs). The noncommercial banking sector uses government bonds effectively as reserves to generate money to households. For each unit of government bond increase in the noncommercial banking sector, we should observe a multiple increase in their “shadow money”.

Compare commercial bank wholesales funding rate and the rate on government bond. In a full competitive market, this will reflect whether bank deposits provide better liquidity than government bonds in the broader economy.

On deposit outflow sensitivity on government bonds: If all variations come from open market operations, then we should observe that the sensitivity of deposits on government bonds must be greater than 1. However, since major variations from government bonds do not come from monetary policy, we are not sure about the sensitivity. Suppose all new holdings of newly issued government bonds come from investors outside the commercial banking system, then we should observe zero funding outflow from the commercial banks.

Question: What is monetary policy “passthrough”? Is it about loan rates? Or rates in other markets? Suppose banks have monopoly power and the world only has commercial banks. Then it seems possible to have banks increase loan rate by 2% for each 1% increase in deposit rate, which means 200% passthrough? Thus it is unclear whether monopoly power should increase or decrease monetary policy passthrough. Another point about passthrough is whether we should only care about different interest rates. Depending on demand or supply effects, the same rate might indicate different things.

6.1 Review of Literature

As a benchmark, Wallace (1981) provides an irrelevance result for open market operations.

Literature suggests the following channels of monetary policy transmission.

- The required reserve mechanism, which is also called the "money view", that describes changes caused by the monetary policy as mainly through changes in deposits because of a binding reserve requirement.
- New Keynesian channel, where long the long-term interest rate matters. A short-term change in interest rate matters very little.
- Bank lending channel, or the "credit view". Classical underpinning on the lending channel is the required reserve mechanism.
- Bank balance sheet channel, which works through "surprise" in interest rate changes and a mismatch of maturities on liability side and asset side.

6.2 Bank Monopoly Power and the Deposit Channel

For simplicity, consider a two period model. The economy lasts for one period and there is no risk. The representative household maximizes utility over final wealth W and liquidity services l, according to a CES aggregator

$$u(W_0) = \max \left(\underbrace{W}_{\text{wealth}}{}^{\frac{\rho-1}{\rho}} + \lambda \cdot \underbrace{l}_{\text{liquidity}}{}^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}}$$

where λ is a share parameter and ρ is the elasticity of substitution between wealth and liquidity services. It is natural to consider wealth and liquidity as complements, and thus we set $\rho < 1$.

Liquidity services are themselves an aggregation

$$l(M, D) = \left(\underbrace{M}_{\text{cash}}{}^{\frac{\varepsilon-1}{\varepsilon}} + \delta \cdot \underbrace{D}_{\text{deposits}}{}^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

We think cash and deposits are substitutes, and thus set $\varepsilon > 1$.

Deposits are also a composition of different banks,

$$D = \left(\frac{1}{N} \sum_{i=1}^{N} D_i^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}}$$

Deposits are not perfect substitutes, i.e. $\eta > 1$. which gives banks monopoly power and sustain nonzero profits. Implicitly, each bank is assumed to have a mass of $1/N$. When each bank has the same deposit D_i, we have $D_i = D$.

For now, assume that banks can only invest in bonds, with return f. We should interpret this as including different assets such as stocks, mutual funds, and others with a common risk-adjusted rate of return, which is set by the central bank. On the deposit side, each bank i charges a deposit spread s_i and pays a deposit rate $f - s_i$, to maximize profit $D_i s_i$, with the first order condition

$$\frac{\partial D_i}{\partial s_i} s_i + D_i s_i = 0$$

$$\Rightarrow \frac{\partial D_i}{\partial s_i}\frac{s_i}{D_i} = -1$$

which means that the bank will increase deposit spread until the elasticity of demand on deposit spread is -1. The household budget equation is

$$\begin{aligned} W &= (W_0 - M - D)(1+f) + M + D(f-s) \\ &= W_0(1+f) - Mf - Ds \end{aligned}$$

where W_0 is the initial wealth, M is the money holding, and D is the deposit holding. Thus the (representative) household optimization problem is

$$\begin{aligned} &\max_{D_i,M} \left(W^{\frac{\rho-1}{\rho}} + \lambda \cdot l^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}} \\ &s.t. \\ &l(M,D) = \left(M^{\frac{\varepsilon-1}{\varepsilon}} + \delta \cdot D^{\frac{\varepsilon-1}{e}}\right)^{\frac{\varepsilon}{\varepsilon-1}} \\ &W = W_0(1+f) - Mf - Ds \\ &D = \left(\tfrac{1}{N}\sum_{i=1}^{N} D_i^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} \\ &s = \tfrac{1}{N}\sum_{i=1}^{N} \tfrac{D_i}{D} s_i \end{aligned}$$

Then the FOC on D_i is

$$U^{\frac{1}{\rho-1}} W^{-\frac{1}{\rho}} \left(-\frac{s_i}{N}\right) + U^{\frac{1}{\rho-1}} \lambda l^{-\frac{1}{\rho}} \frac{\partial l}{\partial D} D^{\frac{1}{\eta-1}} \frac{1}{N} D_i^{-1/\eta} = 0$$

$$\Rightarrow W^{-1/\rho}\frac{s_i}{N} = \lambda l^{-1/\rho} \frac{\partial l}{\partial D} D^{\frac{1}{\eta-1}} \frac{1}{N} D_i^{-1/\eta}$$

Dividing the equation for different D_i, we have

$$\frac{s_i}{s_j} = (\frac{D_i}{D_j})^{-1/\eta}$$

$$\Rightarrow \frac{D_i}{D_j} = (\frac{s_i}{s_j})^{-\eta}$$

which is the equation in the appendix of the paper. To get the sensitivity of aggregate demand over the aggregate spread index, we use an alternative form of the above equation to get

$$s_i D_j^{1-\frac{1}{\eta}} = s_j D_j (D_i)^{-1/\eta}$$

$$\Rightarrow s_i \frac{1}{D}\frac{1}{N}\sum_{j=1}^{N} D_j^{1-\frac{1}{\eta}} = (D_i)^{-1/\eta} \frac{1}{N}\sum_{j=1}^{N} s_j \frac{D_j}{D}$$

$$\Rightarrow s_i D^{-\frac{1}{\eta}} = (D_i)^{-1/\eta} s$$
$$\Rightarrow \frac{D_i}{D} = (\frac{s_i}{s})^{-\eta}$$

Then we can take log on both sides to get

$$\log(D_i) - \log(D) = -\eta \log(s_i) + \eta \log(s)$$

Taking derivative over s_i, we have

$$\frac{1}{D_i}\frac{\partial D_i}{\partial s_i} - \frac{1}{D}\frac{\partial D}{\partial s}\frac{\partial s}{\partial s_i} = -\eta\frac{1}{s_i} + \eta\frac{1}{s}\frac{\partial s}{\partial s_i}$$
$$\Rightarrow \frac{s_i}{D_i}\frac{\partial D_i}{\partial s_i} = \frac{s_i}{D}\frac{\partial D}{\partial s}\frac{\partial s}{\partial s_i} - \eta(1 - \frac{s_i}{s}\frac{\partial s}{\partial s_i})$$

At the symmetry equilibrium point, we should have

$$\frac{s_i}{D_i}\frac{\partial D_i}{\partial s_i} = \frac{\partial s}{\partial s_i}\frac{s}{D}\frac{\partial D}{\partial s} - \eta(1 - \frac{\partial s}{\partial s_i})$$

Thus the key is to calculate $\partial s/\partial s_i$ in order to get a relation between the individual deposit sensitivity $\partial s_i/\partial$. With the ratio formula D_i/D_j and definition of s, we have

$$s = \frac{1}{N}\sum_{i=1}^{N}\frac{D_i}{D}s_i = \frac{1}{N}\sum_{i=1}^{N}(\frac{s_i}{s})^{-\eta} s_i = \frac{1}{N}s^{\eta}\sum_{i=1}^{N}s_i^{1-\eta}$$
$$\Rightarrow s = \left(\frac{1}{N}\sum_{i=1}^{N}s_i^{1-\eta}\right)^{\frac{1}{1-\eta}}$$
$$\frac{\partial s}{\partial s_i} = s^{\frac{\eta}{1-\eta}}\frac{1}{N}s_i^{-\eta}$$

which is different from the paper that says $\partial s/\partial s_i = 1/N$. I think the paper has made a mistake in this part. But if indeed the partial derivative is $1/N$, with bank FOC, we get

$$-\frac{s}{D}\frac{\partial D}{\partial s} = 1 - (\eta - 1)(N - 1)$$

Another way to derive the aggregation results

Another way to derive the aggregation results is to replace every D_i with a specific D_{i0}, and then get the relation between D and D_{i0}. Next, we can write $\sum_i D_i s_i$ as $D \cdot s$ for some aggregated price index s. This is a useful way, because we might not guess the correct form of s before hand. Following the above procedure, we have

$$D = \left(\frac{1}{N}\sum_{i=1}^{N}D_i^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} = \left(\frac{1}{N}\sum_{i=1}^{N}(D_{i0}(\frac{s_i}{s_{i0}})^{-\eta})^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} = D_{i0}s_{i0}^{\eta}\left(\frac{1}{N}\sum_{i=1}^{N}s_i^{1-\eta}\right)^{\frac{\eta}{\eta-1}}$$

which implies

$$D_i = \frac{D}{s_i^{\eta}\left(\frac{1}{N}\sum_{i=1}^{N}s_i^{1-\eta}\right)^{\frac{\eta}{\eta-1}}} = Ds_i^{-\eta}\left(\frac{1}{N}\sum_{i=1}^{N}s_i^{1-\eta}\right)^{-\frac{\eta}{\eta-1}}$$

Thus the total cost is

$$\sum_{i=1}^{N}\frac{1}{N}D_i s_i=\sum_{i=1}^{N}\frac{1}{N}Ds_i^{-\eta}\left(\frac{1}{N}\sum_{i=1}^{N}s_i^{1-\eta}\right)^{-\frac{\eta}{\eta-1}}$$

$$=D\left(\frac{1}{N}\sum_{i=1}^{N}s_i^{-\eta}\right)\left(\frac{1}{N}\sum_{i=1}^{N}s_i^{1-\eta}\right)^{-\frac{\eta}{\eta-1}}=D\left(\frac{1}{N}\sum_{i=1}^{N}s_i^{1-\eta}\right)^{\frac{1}{1-\eta}}$$

Thus we can define the price index as

$$s=\left(\frac{1}{N}\sum_{i=1}^{N}s_i^{1-\eta}\right)^{\frac{1}{1-\eta}}$$

Remark 11. *A better way to deal with CES aggregators in multiple layers is to aggregate from bottom up. For example, in this problem, we can first solve the individual bank deposit problem*

$$\begin{aligned}&\max\left(\tfrac{1}{N}\sum_{i=1}^{N}D_i^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}\\ &s.t.\\ &\tfrac{1}{N}\sum_{i=1}^{N}s_iD_i\le c\end{aligned}$$

Then by the property of the CES aggregator, we immediately have a substitution result

$$(D_i/D_j)^{-\frac{1}{\eta}}=s_i/s_j \qquad \text{(Cross-FOC)}$$

To get the aggregate result, we can define price index

$$s=\frac{1}{N}\sum_{i=1}^{N}s_i\frac{D_i}{D}$$

and aggregate demand

$$D=\left(\frac{1}{N}\sum_{i=1}^{N}D_i^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}$$

As long as the relative allocation of different deposits satisfy the first order conditions (Cross-FOC), we should write the budget equation as only a function of s.

Using the above aggregation results, we can easily derive the substitution between money and deposits as

$$\left(\frac{M}{D}\right)^{-\frac{1}{\varepsilon}}=\delta\frac{f}{s}$$

$$\Rightarrow\frac{D}{M}=\delta^{\varepsilon}(\frac{s}{f})^{-\varepsilon}$$

Thus a higher substitutability between money and deposits makes larger deposits outflow with higher deposit spread s. Using this relationship, we can aggregate them again into an aggregate liquidity index,

$$l=\left(M^{\frac{\varepsilon-1}{\varepsilon}}+\delta^{\varepsilon}\left(\frac{s}{f}\right)^{-(\varepsilon-1)}M^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}}=\left(1+\delta^{\varepsilon}\left(\frac{s}{f}\right)^{-(\varepsilon-1)}\right)^{\frac{\varepsilon}{\varepsilon-1}}M$$

Then the cost can be expressed as

$$s_l \cdot l = Mf + Ds = Mf + M\delta^{\varepsilon}(\frac{s}{f})^{-\varepsilon} s = \frac{f}{\left(1+\delta^{\varepsilon}\left(\frac{f}{s}\right)^{(\varepsilon-1)}\right)^{\frac{1}{\varepsilon-1}}} l$$

which means

$$s_l = \frac{1}{\left(1+\delta^{\varepsilon}\left(\frac{f}{s}\right)^{(\varepsilon-1)}\right)^{\frac{1}{\varepsilon-1}}} f$$

The substitution between liquidity and bonds is a result of the following problem

$$\begin{aligned} &\max \quad \left(W^{\frac{\rho-1}{\rho}} + \lambda l^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}} \\ &s.t. \\ &W = W_0(1+f) - s_l l \end{aligned}$$

which implies

$$\frac{l}{W} = \lambda^{\rho} s_l^{-\rho}$$

With the budget equation

$$W + s_l l = W_0(1+f)$$

we can solve for liquidity as

$$l = \frac{W_0(1+f)}{s_l^{\rho}\lambda^{-\rho} + s_l}$$

Then we can solve for demand function

$$D = \frac{\delta^{\varepsilon}\left(\frac{s}{f}\right)^{-\varepsilon}}{\left(1+\delta^{\varepsilon}\left(\frac{s}{f}\right)^{-(\varepsilon-1)}\right)^{\frac{\varepsilon}{\varepsilon-1}}} l = \frac{\delta^{\varepsilon}\left(\frac{s}{f}\right)^{-\varepsilon}}{\left(1+\delta^{\varepsilon}\left(\frac{s}{f}\right)^{-(\varepsilon-1)}\right)^{\frac{\varepsilon}{\varepsilon-1}}} \frac{W_0(1+f)}{s_l^{\rho}\lambda^{-\rho} + s_l}$$

Remark 12. *I call the approach of this type of model as* ***"bottom-up-bottom"*** *approach, which first calculate cost index from bottom to up. Then it distributes the allocation from up to bottom. In this example, if we fix s_i, then we get s a function of s_i. Then the up-bottom will go down to each D_i, which could be different.*

If we take $\lambda \to 0$, then we can explicitly calculate the demand sensitivity to s. The explicit calculation is just a matter of algebra. In the paper, the calculated sensitivity is

$$-\frac{\partial D}{\partial s}\frac{s}{D} = \frac{1}{1+\delta^{\varepsilon}\left(\frac{f}{s}\right)^{\varepsilon-1}}\varepsilon + \frac{\delta^{\varepsilon}\left(\frac{f}{s}\right)^{\varepsilon-1}}{1+\delta^{\varepsilon}\left(\frac{f}{s}\right)^{\varepsilon-1}}\rho$$

Plug this into the aggregate equation, we can solve for f/s. We note that without $\partial s/\partial s_i$, the whole procedure breaks down.

7 No Arbitrage Pricing

7.1 Pricing Kernel and Interest Rate in Discrete Time

Consider a representative agent economy with CRRA utility functions of risk aversion γ. Then we have

$$M_{t,t+1} = \beta(\frac{c_{t+1}}{c_t})^{-\gamma}$$

and log pricing kernel

$$m_{t,t+1} = \log(\beta) - \gamma\Delta c_{t+1}$$

Assume consumption growth is log-normal. Then the bond price is

$$P_t^{(1)} = E[M_{t,t+1}] = E[e^{m_{t,t+1}}] = \exp(\log(\beta) - \gamma E_t[\Delta c_{t+1}] + \frac{1}{2}\gamma^2\sigma_t^2(\Delta_{t+1}))$$

and the interest rate is

$$i_t^{(1)} = -\log(P_t^{(1)}) = -\log(\beta) + \gamma E_t[\Delta c_{t+1}] - \frac{1}{2}\gamma^2\sigma_t^2(\Delta c_{t+1})$$

The nominal interest rate is

$$i_t^{(1)\$} = -\log(\beta) + \gamma E_t[\Delta c_{t+1}] - \frac{1}{2}\gamma^2\sigma_t^2(\Delta c_{t+1}) + \pi_{t+1}$$

Implication: Higher discount rate, higher consumption growth, and higher consumption growth volatility (more precautious saving drives down interest rate) result in lower interest rates.

The n-period nominal interest rate is

$$\exp(-ni_t^{(n)\$}) = E_t[\beta^n \exp\left(-\sum_{i=1}^{n}\gamma\Delta c_{t+i} + \pi_{t+i}\right)]$$

Suppose both the conditional consumption growth and conditional inflation have constant volatility and they are **uncorrelated**. Then we can get

$$i_t^{(n)\$} - \mu^{(n)\$} = \frac{1}{n}E_t[\sum_{i=1}^{n}(\gamma\Delta c_{t+i} + \pi_{t+i} - (\gamma\mu_c + \mu_\pi))]$$

where $\mu^{(n)\$}$ is the unconditional expectation of interest rates. Then we find that the variation over interest rate comes from consumption growth and inflation. However, in general the correlation will matter.

In general, we can express interest rates and bond risk premium in terms of the log nominal pricing kernel $m_t^{\$}$. Denote n period nominal bond holding period return as

$$hpr_{t\to t+1}^{(n)\$} = \log(\frac{P_{t+1}^{(n-1)\$}}{P_t^{(n)\$}})$$

The basic pricing formula is

$$E_t[e^{m_{t,t+1}}e^{hpr_{t\to t+1}^{(n)\$}}] = 0$$

and we are interested in the bond risk premium, i.e.

$$E_t[hpr_{t\to t+1}^{(n)\$}] + \frac{1}{2}\text{var}_t(hpr_{t\to t+1}^{(n)\$}) - i_t^{(1)\$}$$

We can derive a general formula for the excess return of any assets. Denote the nominal return of the asset as

$$R_{t+1} = e^{r_{t+1}}$$

Then we should have

$$E_t[e^{m_{t,t+1}^{\$}} e^{r_{t+1}}] = 1$$

$$\Rightarrow E_t[m_{t,t+1}^{\$} + r_{t+1}] + \frac{1}{2}\text{var}_t[m_{t,t+1}^{\$} + r_{t+1}] = 0$$

$$\Rightarrow E_t[m_{t,t+1}^{\$}] + E_t[r_{t+1}] + \frac{1}{2}\text{var}_t[m_{t,t+1}^{\$}] + \frac{1}{2}\text{var}_t[r_{t+1}] + \text{cov}_t(m_{t,t+1}^{\$}, r_{t+1}) = 0$$

Note that

$$E_t[m_{t,t+1}^{\$}] + \frac{1}{2}\text{var}_t[r_{t+1}] = -i_t^{(1)\$}$$

Thus excess return is

$$rp_t = \log(E_t[\frac{R_{t+1}}{\exp(i_t^{(1)\$})}]) = E_t[r_{t+1}] + \frac{1}{2}\text{var}_t[r_{t+1}] - i_t^{(1)\$} = -\text{cov}_t(m_{t,t+1}^{\$}, r_{t+1})$$

Then we get a general statement that for any asset with nominal return $R_{t+1} = e^{r_{t+1}}$, the excess return of this asset is the negative of its covariance with log nominal pricing kernel $m_{t,t+1}^{\$}$.

Upward Sloping or Downward Sloping Term Structure?

It is important to have a model that fits the term structure of bonds. It depends on the consumption process.

Case 1: Trend Stationary

We can denote the consumption process as

$$\log(c_t) = a + bt + u_t$$

Then consumption growth should follow

$$\Delta c_{t+1} = b + u_{t+1} - u_t$$

which results in

$$\text{cov}(\Delta c_{t+1}, \Delta c_t) = -\sigma^2(u_t) < 0$$

As a result, low consumption growth will predicts high consumption growth in the future. In bad time, expectation of future interest rate is high. Interest rates are countercyclical, which is not consistent with U.S. data, where in a crisis interest rates are much lower.

Furthermore, in good times, people expect lower consumption growth in the future, which implies a downward sloping term structure. But in bad times, people expect higher consumption growth in the future, which implies a upward sloping term structure. Thus whether the term structure is upward or downward sloping will crucially depend on whether consumption is above the trend line or below the trend line, which is again counterfactual.

Case 2: Difference Stationary

Assume consumption process follows the following difference stationary process

$$\Delta c_t = x_t$$

$$x_{t+1} = \phi x_t + \varepsilon_{t+1}$$

In the data, we get a positive autocorrelation of consumption growth. As a result, in bad times, we will expect cumulative effects of low consumption growth rates and a lower interest rates. Thus interest rate should be procyclical, which is close to data.

Furthermore, the consumption growth will persistently be low in bad times, resulting in low interest rates for a while. Thus long-term bonds are good hedge, and demand lower yield. This results in downward sloping yield curve.

Rule of the game: With positive autocorrelation of consumption growth, find a reasonable explanation of the upward sloping yield curve (both nominal and real) we observe in reality.

An explanation provided by Piazzesi et al. (2006) is that inflation is bad news for future consumption growth. Because inflation is volatile, long term nominal bonds requires higher risk premia. Furthermore, since in bad states, typically inflation is low, we expect higher consumption growth in the future, which results in upward sloping real term structure as well.

7.2 Affine Term Structure Models in Discrete Time

Now let's study the model of Ang and Piazzesi (2003), which is a famous paper that builds a no-arbitrage VAR of term structure dynamics with both macroeconomic and latent variables. Compared to the macro literature on term structure, it has much more discipline from no arbitrage and is a complete description of the term structure. Compared to the finance literature on no arbitrage pricing, it has macro variables and connects its approach to VAR. Thus this is a classical paper that belongs to the emerging macro finance literature.

Describe the pricing kernel as

$$M_{t+1} = \exp(-y_t^{(1)} - \frac{1}{2}\lambda_t^T \lambda_t - \lambda_t^T \varepsilon_{t+1})$$

where $y_t^{(1)}$ is the short rate

$$y_t^{(1)} = \delta_0 + \delta_1^T x_t$$

and the price of risk is

$$\lambda_t = l_0 + l_1 x_t$$

Assume x_t has affine dynamics

$$x_{t+1} = \mu + \phi x_t + \sigma \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim N(0, I)$$

Results: the log bond price $p_t^{(n)} = \log(P_t^{(n)})$ is affine in x_t,

$$p_t^{(n)} = A_n + B_n^T x_t$$

where

$$A_{t+1} = A_n + B_n^T(\mu - \sigma l_0) + \frac{1}{2} B_n^T \sigma \sigma^T B_n - \delta_0$$
$$B_{n+1}^T = B_n^T(\phi - \sigma l_1) - \delta_1^T$$

7.3 Pricing Kernel and Interest Rate in Continuous Time

In continuous time, a pricing kernel w.r.t. probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is defined as an adapted process Λ_t, such that $\Lambda_t P_t$ is a martingale for any asset with price P_t, where we do not consider dividends[2]. Then we have

$$E_t[d(\Lambda_t P_t)] = 0$$

In practice, it is easier to work with growth rate, so we can express the above into

$$E_t[\frac{d\Lambda_t}{\Lambda_t} + \frac{dP_t}{P_t} + \frac{d\Lambda_t}{\Lambda_t} \cdot \frac{dP_t}{P_t}] = 0$$

For risk free rate asset $P_t = P_0 e^{rt}$, we get

$$r_t = -E_t[\frac{d\Lambda_t}{\Lambda_t}]/dt$$

For risky assets with return

$$\frac{dP_t}{P_t} \triangleq dR_t = \mu_t dt + \sigma_t dB_t$$

We get

$$\mu_t - r_t = -E_t[\frac{d\Lambda_t}{\Lambda_t} \cdot \frac{dP_t}{P_t}]/dt = -Cov_t(\frac{d\Lambda_t}{\Lambda_t}, \frac{dP_t}{P_t})/dt$$

where the second equality comes from the fact that the term

$$E_t[\frac{d\Lambda_t}{\Lambda_t}] \cdot E_t[\frac{dP_t}{P_t}] = o(t)$$

Thus if an asset has higher return correlation with the SDF, the required expected return should be lower. This is intuitive because higher SDF value means it is more valuable to have a good return at that time. Any asset that caters to a higher SDF value should demand lower expected return since it is a good hedge.

To connect the above to consumption, we can work with CRRA utility

$$u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$$

[2]With dividend, we should define cum-dividend asset price, which can be operated in the same way.

with discount rate ρ, which implies

$$\Lambda_t = e^{-\rho t} c_t^{-\gamma}$$

is a pricing kernel. Thus the risk free rate is

$$r = \rho + \gamma E_t[\frac{dc_t}{c_t}]/dt - \frac{1}{2}\gamma(\gamma+1)(\frac{dc_t}{c_t})^2/dt$$

Thus with higher consumption growth rate, the interest rate will be higher, and but the interest rate is lower with more consumption growth volatility. With log utility, we can use the above formula with $\gamma = 1$, which yields

$$r = \rho + E_t[\frac{dc_t}{c_t}]/dt - (\frac{dc_t}{c_t})^2/dt$$

To get a term structure, we need to price longer term asset. For example, denote $r_{t,t+\Delta}$ as the interest rate of a bond that matures at $t + \Delta$, starting from t. Then we should have

$$E_t[e^{\Delta r_{t,t+\Delta}} e^{-\rho\Delta}(\frac{c_{t+\Delta}}{c_t})^{-\gamma}] = 1$$

which should imply the long-term interest rate.

8 Continuous Dynamic General Equilibrium

Typically, this type of model involves the following steps:

- Solve the individual optimization problem, and get optimal consumption and portfolio choices.
- Use equilibrium condition to get differential equations.

8.1 Solution to Brunnermeier and Sannikov (2014)

In the paper, both households and bankers are risk neutral. Bankers have better investment technology, higher discounting (thus less patient), and nonnegative consumption, while households can have negative consumption to guarantee the risk-free rate is just the discounting rate, which greatly simplifies the model. Basic notations: (1) Price of capital $q(\eta)$. (2) Value to wealth ratio $\theta(\eta)$. (3) Fraction of risky assets held by bankers $\psi(\eta)$.

- Households.
  - If households ever hold a positive fraction of the risky asset, the expected return of the risky asset held by households and the risk free asset should be the same, i.e.

    $$E_t[d\underline{r}_t^k] = rdt$$

    Otherwise, households hold zero fraction of the risky asset, and $\psi_t = 1$ (bankers hold all the risky asset). Household consumption in this model is flexible to pin down risk free rate $r_t = \rho^h$, and clears both the consumption and risk free asset markets, because households can transform between consumption and risk free assets.

- Bankers
  - If bankers ever hold a positive fraction of the risky asset, then the expected return of the risky asset should satisfy the risk premium equation
  $$\frac{E_t[dr_t^k]}{dt} - r = -\sigma_t^\theta(\sigma + \sigma_t^q)$$
  where $\sigma_t^\theta \leq 0$. The risk premium comes from precautionary motive because bankers suffer losses exactly when the investment opportunities are good, i.e. $Z_t \uparrow$, $\eta_t \downarrow$, and $\theta_t \uparrow$ (marginal value of wealth increases). Thus either $\psi_t = 0$, or the above equation holds. When either $\psi_t = 0$ or $\psi_t = 1$, we know the portfolio choice for both households and bankers. When $\psi_t \in (0, 1)$, both return equations should hold, and thus we can put them together and get
  $$\frac{a - \underline{a}}{q(\eta)} + \underline{\delta} - \delta + (\sigma + \sigma_t^q)\sigma_t^\theta = 0$$
  which can be used to solve ψ. After solving ψ, we can get q'' and θ'', and thus a system of ordinary differential equations.
- Boundary conditions.
  - First, when $\eta = 0$, the economy will stay at $\eta = 0$ forever, and thus only households price the assets. This will result in an asset price $q(0) = \underline{q}$.
  - Second, when $\eta = \eta^*$, by definition the household will consume, and the Bellman equation implies $\theta(\eta^*) = 1$.
  - Third, we note that the optimal instantaneous consumption at η^* is to push η_t back to η^*, because otherwise the slope of value function is greater than 1 and the additional consumption makes bankers lose value. This means that the point η^* is a reflection boundary of the system. By the standard arguments for reflection boundary, we should have $\theta'(\eta^*) = q'(\eta^*) = 0$.[3] The intuition is as follows. If $\eta_t = \eta^*$, then $\eta_{t+\varepsilon}$ can be approximated by $\eta^* - A\sigma_t^\eta\sqrt{dt}$, where $A = \sqrt{2/\pi}$, because the change is the absolute value of a normal distribution with mean 0 and variance $(\sigma_t^\eta)^2 dt$. Then starting from $\eta_t = \eta^*$, we have $q(\eta_{t+dt}) \sim q(\eta^*) + q'(\eta^*)A\sigma_t^\eta\sqrt{dt}$. Thus the loss pert unit of time dt is $q'(\eta^*)A\sigma_t^\eta\sqrt{dt}$, which means the average loss rate goes to infinity if $q'(\eta^*) \neq 0$. Similarly, the drift of $\theta(\eta_t)$ will be infinite at η^* if $\theta'(\eta^*)$, violating the finite drift solution in the paper.
  - Finally, when $\eta \to 0$, the asset price is absorbed at $\underline{q}$. Thus bankers can take infinite leverage with small wealth and generate an infinite rate of return, leading to $\theta(0) = \infty$.
  - The above five boundary conditions are used to solve two second order ODEs, with one endogenous boundary.

Explanation on "hedging demand": In the future, there are states where marginal utility of wealth is high or low. Return on capital is correlated with the marginal utility

[3] We don't need $\psi'(\eta^*) = 0$ because it is not differentiable in the first place.

of wealth. If return on capital is high when marginal utility on wealth is low, then intermediaries are demanding higher risk premium. This hedging demand will endogenously restrict the leverage of the intermediary.

Why household discount is the same as the interest rate? Technically, the HJB equation results in the marginal utility of wealth to households always equal to 1, which implies the interest rate equals the discount rate. Intuitively, households can freely transform from consumption to wealth, thus making sure the discount rate is the same as the interest rate. It is thus very important to have intermediaries with nonnegative consumption, which prevents the "first best" solution, where intermediaries consume minus infinity only at the beginning by borrowing from households, while households consume continuously into the future.

As shown in the paper, with log utility, the solution is much simpler. The main reason is that portfolio choice and consumption decisions are no longer dependent on individual indirect utility functions, which greatly simplifies discussions. Without log utility, even under risk neutrality, we need to describe the marginal utility process for wealth, which requires solving differential equations.

In my paper, if I allow bankers to be risk neutral, then the model has bankruptcy, which could cause riskiness of bank debt, and closer to reality. However, the cost is to introduce another function $\theta(w, \lambda)$ to be solved.

8.2 Solution Techniques for Continuous Time Dynamic General Equilibrium

Several important observations:

- The main simplification from continuous time technique is the representation of uncertainty. Because all uncertainties are local, in the derived functional equation system, we do not have any expectation terms.
- The continuous time general equilibrium has a natural connection to the projection methods with Chebyshev polynomials because differentiation is quite easy with projection methods.
- In general, solving a PDE with multiple state variables is very difficulty. Once we have a general equilibrium with a lot of different state variables, we need to resort to perturbation techniques that approximate the solution. However, the perturbation method will results in giant sparse matrix when solving the differential equations. We can resort to the method proposed by Ahn et al. (2017) that is able to tackle large scale problems, even with heterogeneous agents.
- With jumps, we have two ways to go. (1) Add an outside loop that uses last step solution to solve for the jump equations. (2) Directly use projection method. But instead of using collocation, use an optimization with constraint algorithm. Make sure to supply the first order and second order derivatives to the solver to speed up the algorithm.

8.3 Monopolistic Competition

Two ways of modeling monopolistic competition:

- Assume that firms sell differentiated varieties of a good to consumers who aggregate these according to a CES index.
- Assume a continuum of intermediate producers with market power who each sell a different variety to a competitive final goods producer whose production function is a CES aggregate of intermediate varieties.

The second assumption is more common. But I think we should have an equivalence between the two?

8.4 A Deterministic RBC Model in Continuous Time

A representative households maximize

$$\int_0^\infty e^{-\rho t}\left\{\log(C(t)) - \frac{N(t)^{1+\phi}}{1+\phi}\right\}dt$$

subject to

$$P(t)\cdot C(t) + B'(t) = i(t)\cdot B(t) + W(t)N(t)$$

where $B(t)$ is bond holding, $P(t)$ is price level, $W(t)$ is nominal wage, and $i(t)$ is nominal interest rate. I have used capital letter notations since here the household is representative. Denote the Lagrangian multiplier on the budget constraint as $e^{-\rho t}\lambda(t)$. Then the principal of optimality indicates

$$\begin{cases} \frac{1}{C(t)} = \lambda(t)P(t) \\ N(t)^\phi = \lambda(t)W(t) \\ \lambda'(t) = (\rho - i(t))\lambda(t) \end{cases}$$

which implies

$$\frac{C'(t)}{C(t)} = i(t) - \frac{P'(t)}{P(t)} - \rho$$

Define inflation rate as

$$\pi(t) = \frac{P'(t)}{P(t)}$$

Then we have

$$\frac{C'(t)}{C(t)} = i(t) - \pi(t) - \rho$$

Consumption growth is equal to nominal interest rate minus discount rate. Or in the other way around, we have

$$i(t) = \pi(t) + \frac{C'(t)}{C(t)} + \rho$$

which means interest rate equals the sum of inflation rate, consumption growth rate, and discount rate. In a stochastic world, we will have a volatility term at the end. The real interest rate is

$$r(t) \triangleq i(t) - \pi(t) = \frac{C'(t)}{C(t)} + \rho$$

A competitive final good producer aggregates a continuum of intermediate inputs

$$Y = (\int_0^1 y_j^{\frac{\varepsilon-1}{\varepsilon}} dj)^{\frac{\varepsilon}{\varepsilon-1}}$$

where ε is the elasticity of substitution. The cost minimization of final good producer implies the following demand of intermediate good of type j

$$y_j^D(p_j) = (\frac{p_j}{P})^{-\varepsilon} Y$$

Intermediate good producers have monopolistic competition and production uses only labor,

$$y_j(t) = A(t) n_j(t)$$

Thus each intermediate good producer solves

$$\max_{p_j} \left\{ p_j \cdot y_j^D(p_j) - \frac{y_j^D(p_j)}{A(t)} W(t) \right\}$$

which yields

$$p_j(t) = P(t) = \frac{\varepsilon}{\varepsilon - 1} \frac{W(t)}{A(t)}$$

We note that when $\varepsilon \to \infty$, i.e. all intermediate goods are perfectly substitutable and thus the intermediate good market is fully competitive, the price of each intermediate good is just equal to $W(t)/A(t)$, the marginal cost of producing that good. When $\varepsilon \downarrow 1$, the competitive is much weaker and price is much higher than the marginal production cost.

Next, market clearing implies

$$C(t) = A(t) N(t)$$

Combining household FOCs on labor and consumption, we have

$$C(t) = A(t) N(t)$$

$$\Rightarrow C(t) N(t)^{\phi} = A(t) N(t)^{\phi+1}$$

Plugging in household FOC and price equation from production side, we have

$$N(t) = \left(\frac{\varepsilon}{\varepsilon - 1} \right)^{-\frac{1}{\phi+1}}$$

Thus labor supply is constant. Then consumption is just

$$C(t) = A(t) \left(\frac{\varepsilon}{\varepsilon - 1} \right)^{-\frac{1}{\phi+1}}$$

Denote productivity growth rate as

$$g(t) = \frac{A'(t)}{A(t)}$$

Then consumption grows at the same rate, and real interest rate is

$$r(t) = g(t) + \rho$$

Price level $P(t)$ and inflation $\pi(t) = P'(t)/P(t)$ has no impact on the economy at all.

With this simple model, we can measure welfare. Assume a constant productivity growth rate $g < \rho$. Then the social welfare is

$$\int_0^\infty e^{-\rho t}\left\{\log(A(0)e^{gt}N) - \frac{N^{1+\phi}}{1+\phi}\right\}dt$$

$$= \frac{1}{\rho}\left(\log(A(0)) + \frac{g}{\rho} + \log(N) - \frac{N^{1+\phi}}{1+\phi}\right)$$

$$= \frac{1}{\rho}\left(\log(A(0)) + \frac{g}{\rho} - \frac{1}{\phi+1}\left(\log(\frac{\varepsilon}{\varepsilon-1}) + \frac{\varepsilon}{\varepsilon-1}\right)\right)$$

Thus an increase in competition ε increases welfare.

8.5 A Deterministic New Keynesian Model in Continuous Time

This model borrows from Benjamin Moll's lecture notes on advanced macroeconomics. New Keynesian model = RBC model with sticky prices. The setup is the same as the last section, except sticky price. Intermediate good producers are subject to quadratic price adjustment cost

$$\Theta_t(\frac{p'}{P}) = \frac{\theta}{2}\left(\frac{p'}{P}\right)^2 P(t)Y(t)$$

The per period profit is still

$$\Pi_t(p) = p\left(\frac{p}{P(t)}\right)^{-\varepsilon} Y(t) - \frac{W(t)}{A(t)}\left(\frac{p}{P(t)}\right)^{-\varepsilon} Y(t)$$

Adjustment costs are paid as a transfer to consumers to avoid real resource costs of inflation. Then the optimal control problem of each intermediate producer is

$$\max_{p(t),t\geq 0}\int_0^\infty e^{-\int_0^t i(s)ds}\left(\Pi_t(p) - \Theta_t(\frac{p'}{P})\right)$$

This is a typical calculus of variation problem, which can be easily solved.

9 General Equilibrium Effects of Government Bonds and Taxation

9.1 Government Bond Effects

In this section, I will study how government bond affects the economy through a simple two-period model. To keep things simple, I start with an exchange economy with homogeneous agents. Then I switch to a production economy with homogeneous agents. Finally, I study a heterogeneous agent economy with two type of agents.

An Exchange Economy

A unit mass of households, each with endowment y_1 and y_2 in period 1 and 2. Households could freely borrow from and lend to each other, with endogenous interest rate r. All households have separate utility

$$u(c_1) + \beta u(c_2)$$

on consumptions. Government spends G_1 and G_2 at period 1 and 2. To fund government spending, the government can impose a lump-sum tax on all agents, with rates τ_1 and τ_2 at period 1 and 2. Government can also issue government bonds, with quantity Q in the first period (it is meaningless to issue second period bond, because no one will hold it). By no arbitrage, the interest rate paid on government bond should also be r. We will fix the amount of government spending, but vary quantity of government bonds and taxation to see if they affect the total social welfare.

Household problem:

$$\max_{c_1,c_2} u(c_1) + \beta u(c_2)$$
$$s.t.$$
$$\begin{cases} y_1 = c_1 + \tau_1 + b_1 + a \\ y_2 + (1+r)a + (1+r)b_1 = c_2 + \tau_2 \\ b_1 \geq 0 \end{cases}$$

Government budget constraint:

$$\tau_1 + Q = G_1$$

$$\tau_2 = G_2 + (1+r)Q$$

Asset market clearing

$$a = 0$$

$$b_1 = Q$$

Households are indifferent between a and b_1, and only $a + b_1$ matters for individual households. As a result, we can solve the individual optimization, and then clear the market with $a + b_1 = Q$. Household FOC implies

$$u'(c_1) = \beta(1+r)u'(c_2)$$

Cancelling out a in household budget constraints, we get

$$y_1 + \frac{y_2}{(1+r)} = \frac{c_2 + \tau_2}{(1+r)} + c_1 + \tau_1$$

Observation: In a deterministic world, when borrowing and lending are unrestricted and at the same rate, we can write an aggregate budget constraint, instead of an constraint for each period.

With the above expressed by r, we can get an equation for r by setting

$$a + b_1 = y_1 - c_1 - \tau_1 = Q$$

As a result,

$$c_1 = y_1 - Q - \tau_1$$

$$c_2 = \frac{y_2 - \tau_2}{(1+r)} + Q$$

In this model, government has the "special technology" to transform goods from period 1 to period 2, with either taxation and government bonds. Now let's shut down the government spending channel by assuming $G_1 = G_2 = 0$. Then

$$c_1 = y_1, \quad c_2 = y_2$$

The interest rate is solved from

$$u'(y_1) = \beta(1 + r^*)u'(y_2)$$

Thus everything is the same regardless the quantity of government bond when we adjust taxation so that the total government spending is kept at zero. Ricardian equivalence holds in this case: How government funds its spending doesn't matter with lump-sum tax and fixed government spending.

Non-zero Government Spending

Now assume instead $\tau_1 = \tau_2 = 0$, but government spending could adjust. In equilibrium, we have

$$\begin{aligned} c_1 &= y_1 - Q \\ c_2 &= y_2 + (1+r)Q \end{aligned}$$

Thus the government could shift resources from period 1 to period 2 and affect the aggregate utility. We note that the government spending is $G_1 = Q$ and $G_2 = -(1+r)Q$. Adding both government spending and household spending, we get

$$\begin{aligned} c_1 + G_1 &= y_1 \\ c_2 + G_2 &= y_2 \end{aligned}$$

which satisfies the typical resource constraint.

With positive Q, the consumption at period 1 will decline to $y_1 - Q$, while the consumption at period 2 will increase to $y_2 + (1+r)Q$. The interest rate changes, and could be solved from the following equation

$$u'(y_1 - Q) = \beta(1+r)u'(y_2 + (1+r)Q)$$

Thus we have $r > r^*$ when $Q > 0$, i.e. interest rate rises when the government borrows from today, which distorts the consumption-saving margin of households. In this case, an increase of Q is the same as increase in government spending G_1, and thus we cannot separate the pure effects of government bond quantity.

Wealth Tax

Suppose the taxation is on wealth, with rate τ_1 and τ_2. Then household optimization problem is

$$\begin{aligned} &\max_{c_1,c_2} u(c_1) + \beta u(c_2) \\ &s.t. \\ &\begin{cases} y_1 = c_1 + b + a \\ (1-\tau_2)\left(y_2 + (1+r)(a+b)\right) = c_2 + \tau_2 \\ b \geq 0 \end{cases} \end{aligned}$$

Tax rate is solved from

$$(1-\tau_2)\left(y_2+(1+r)(a+b)\right)=(1+r)Q$$

$$\Rightarrow 1-\tau_2=\frac{(1+r)Q}{y_2+(1+r)Q}$$

Consumption FOC

$$u'(c_1)=\beta u'(c_2)(1-\tau_2)(1+r)$$

Resource constraint

$$c_1=y_1$$

$$c_2=y_2$$

When we have more Q, the interest rate decreases.

A Production Economy

Now let's add production into the economy. We assume that all households start with equal holdings of capital K_1. Firms rent capital from households at rental rate r^K, and produce output $F(K)$ with capital K, where $F(K)=K\cdot F(1)$ is homogeneous of degree 1. Then the household problem is

$$\max_{c_1,c_2} u(c_1)+\beta u(c_2)$$
$$s.t.$$
$$\begin{cases} w_1=c_1+b_1+a+l+\tau_1 \\ (1+r)(b_1+a)+(1+r^K-\delta)l=c_2+\tau_2 \\ b_1\geq 0 \end{cases}$$

where a is investment in the non-government bond market, b is investment in government bonds, and l is the amount of capital lending to firms. By no arbitrage from the households, we need

$$1+r^K-\delta=1+r$$

The production sector FOC implies

$$r^K=F_K(K_2)$$

Combining investment and production equations, we have

$$F_K(K_2)-\delta=r$$

Government budget constraint with zero government spending:

$$\tau_1+Q=0$$

$$\tau_2=(1+r)Q$$

With $Q>0$, at period 1, government is distributing all the wealth raised by government bonds to the agents in the economy. Asset market clearing implies

$$a=0,\quad b_1=Q$$

Resource constraint:

$$F(K_1) = c_1 + K_2 - (1-\delta)K_1$$
$$F(K_2) = c_2 - (1-\delta)K_2$$

where K_1 should be interpreted as the capital lent to firms in the last period. Firms understand that the capital they rent will depreciate by δ, but they still pay r^K per pre-depreciation unit of capital. Note that we allow negative investment, i.e. capital is transformed into consumption, and we require $K_3 \geq 0$. We note that interest rate is not affected by government bonds. To solve for consumptions, K_2 and interest rate, we have

$$\begin{cases} F_K(K_2) - \delta = r \\ F(K_1) = c_1 + K_2 - (1-\delta)K_1 \\ F(K_2) = c_2 - (1-\delta)K_2 \\ u'(c_1) = \beta(1+r)u'(c_2) \end{cases}$$

which is not affected by government bond as well. When government spending is nonzero, consumption will be affected. Note that if we take household constraint to the aggregate level, they are consistent with the resource constraints.

An Infinite Horizon Version

Now let's move one step further to the continuous time analog. Fixed amount of capital and fixed productivity $F(K) = AK$. The government supplies a fixed amount of government bond Q. Endogenous interest rate r. The simple bond market clears at quantity zero, while the government bond market clears at the exogenous quantity QK. Assume log utility function $u(c) = \log(c)$, and discount rate β.

I will first study a model without existing government bond. Then I will study a model with existing government bond and the government will pay the same amount of matured bond starting from the initial period.

(1) Lump-sum Taxation

Household optimization problem:

$$\begin{aligned} &\max \sum_{t=0}^{\infty} \beta^t u(c_t) \\ &s.t. \\ &(1+r_t)(b_t + a_t + l_t) = (b_{t+1} + a_{t+1} + l_{t+1}) + c_t + \tau_t \\ &b_t \geq 0, \quad a_t \geq \underline{a}, \quad l_t \geq \underline{l} \end{aligned}$$

The wealth at time zero

$$w_0 = b_0 + a_0 + l_0$$

with

$$b_0 = 0$$

since we assume the government bond starts issuance from the beginning of period 1. Wealth at the beginning of period t is

$$w_t = a_t + b_t + l_t$$

All wealth variables denote after-tax wealth. Firm FOC

$$r^K = A$$

Investment project

$$r = r^K - \delta$$

Government bond budget constraint with zero government spending

$$\tau_t = \begin{cases} rQ & t \geq 2 \\ -Q & t = 1 \end{cases}$$

Asset market clearing

$$a_t = 0$$

$$b_t = 0 \text{ for } t \geq 1, \text{ and } b_0 = 0$$

Resource constraint:

$$F(K_t) = c_t + K_{t+1} - (1-\delta)K_t$$

In equilibrium, the period t wealth is

$$w_t = Q + K_t$$

Note that the resource constraint is just an aggregate household wealth balance constraint.

(2) Wealth Tax

With wealth tax, the budget equation for households become

$$(1 + r_t - \tau_t)(b_t + a_t + l_t) = (b_{t+1} + a_{t+1} + l_{t+1}) + c_t$$

which results in the following intertemporal equation

$$\frac{\beta u'(c_{t+1})}{u'(c_t)}(1 + r_t - \tau_t) = 1$$

which will result in higher interest rate,

Continuous Homogeneous Production Economy with Wealth Tax

Assume homogeneous agents with capital K. Production is in AK form. Government bond has total value QK. Then we get

$$\begin{cases} \rho W = AK \\ W = pK + QK \end{cases}$$

from which we can solve W and p.

$$W = \frac{A}{\rho}K \ , \quad p = \frac{A}{\rho} - Q$$

Note that the total wealth is just the discounted value of capital, because any government bond is just a transfer of resource, with the same discounting rate. The wealth equation can be interpreted as follows

$$W = \underbrace{pK}_{\text{value of capital}} + \underbrace{QK}_{\text{value of claim to future taxation}}$$

When Q increases, the tax burden in the future increases, and capital is worth less to households, i.e. p declines. But because households now hold the claim to future taxation, total household wealth doesn't change, and thus household consumption remains the same. The interest rate is

$$r = \frac{A}{p}$$

which jumps up when Q increases.

When we have investment $i(p)$ which increases in p, we will get a change, because the change in value of capital will affect investment. When future is worth less, investment into the future is less, thus reducing total investment and productivity. Now the system of equation is

$$\begin{cases} \rho W = (A - i(p))K \\ W = pK + QK \end{cases}$$

We find p as a function of Q is still decreasing in Q, which results in investment $i(Q)$ decreasing in Q. Thus the total wealth W increases in Q, i.e. more government bond now increases today's wealth. This is reasonable: With less investment, we are moving tomorrow's wealth into today's wealth. The interest rate is

$$r = \frac{A - i(p)}{p}$$

The marginal response of price to quantity of government bond could be solved as

$$p'(Q) = -\frac{\rho}{i'(p) + \rho} \in (-1, 0)$$

The marginal response of investment to quantity of government bond is

$$i'(Q) = -\frac{\rho i'(p)}{i'(p) + \rho}$$

$$\Rightarrow |i'(Q)| < \min\{\rho, i'(p)\}$$

which is quite small numerically.

Continuous-Time Homogeneous Production Economy with Lump-Sum Tax

Basic question: Does Ricardian equivalence helps solve the model?

In a homogeneous agent model with government bond and lump-sum taxation, the Ricardian equivalence holds. As a result,

$$c = (A - i(p))K$$

$$W = pK + QK$$

while the price is solved from

$$\frac{A - i(p)}{\rho} = p$$

Increase in government bond Q will increase total wealth of households, but not consumption or asset returns. The interest rate will be

$$r = \frac{A - i(p)}{p} = \rho$$

Continuous Homogeneous Production Economy with Capital Tax

With capital tax, the portfolio choice problem changes. However, the consumption policy doesn't change, and the wealth equation doesn't change. Thus price of capital and household wealth are the same as a production economy with wealth tax. However, the interest rate has to change to

$$r = \frac{(A - i(p))(1-\tau)}{p}$$

In the full model of my paper, capital tax will introduce additional complexity since now I have to solve the interest rate r^g and capital tax in the solution procedure, which makes the functional equation into second order delayed PDE instead of first oder delayed PDE.

In the full model, I conjecture that capital tax reduces bank holding of capital, which decreases productivity of the economy.

9.2 Comparison of Different Taxation in a Two-Period Economy

In this subsection, I will compare how different taxations affect the economy in a two-period model with production. Specifically, I will study the following taxation:

- Lump-sum tax.
- Income tax (tax on capital income).
- Consumption tax.
- Wealth tax.

$$\begin{aligned} &\max\ u(c_1) + \beta u(c_t) \\ &s.t. \\ &\begin{cases} w_1 = (1+\tau_1^c)c_1 + b_1 + a + l + \tau_1 \\ (1-\tau^w)\big((1+r)(b_1+a) + (1+(r^K-\delta)(1-\tau^K))l\big) = (1+\tau_2^c)c_2 + \tau_2 \end{cases} \end{aligned}$$

where τ^w is wealth tax, τ^c is consumption tax, τ^K is the income tax on capital, and τ_1, τ_2 are lump-sum taxation at period 1 and 2. By no arbitrage, we have

$$r = (r^K - \delta)(1-\tau^K)$$

Firm FOC implies

$$r^K = F_K(K_2)$$

As a result, any taxation on capital would distort the investment in capital. We can combine the two constraint to get the aggregate household budget

$$(1-\tau^w)w_1 = \frac{(1+\tau_2^c)}{(1+r)}c_2 + (1-\tau^w)(1+\tau_1^c)c_1 + \frac{\tau_2}{(1+r)} + (1-\tau^w)\tau_1$$

As a result, the inter-temporal consumption FOC is

$$\frac{\beta u'(c_2)}{u'(c_1)} = \frac{(1+\tau_2^c)}{(1+r)(1-\tau^w)(1+\tau_1^c)}$$

We find that wealth tax and non-stationary consumption tax both have distorting effects on the economy. To solve the model, we use the following equilibrium equation

$$F(K_1) = c_1 + K_2 - (1-\delta)K_1$$

$$F(K_2) = c_2 - (1-\delta)K_2$$

The following equation shows clearly how different taxation affects equilibrium

$$\frac{\beta u'(c_2)}{u'(c_1)} = \frac{(1+\tau_2^c)}{(1+(r^K-\delta)(1-\tau^K))\,(1-\tau^w)(1+\tau_1^c)} \tag{17}$$

We have only four unknowns c_1, c_2, K_2, and r, it is easy to see how taxation distorts the equilibrium. By the first welfare theorem, without any taxation, the solution is Pareto optimal. As a result, any distortion by taxation is detrimental to social welfare. Based on the monotonicity of $u'(c)$ and counterarguments, we can get the following results:

- An increase in period i consumption taxation τ_i^c reduces period i consumption. If $F(K)$ is concave, then the interest rate is reduced when τ_1^c increases. When $\tau_1^c = \tau_2^c$, consumption taxation doesn't distort the equilibrium.
- An increase in τ^w increases period 1 consumption c_1, because any wealth leftover will be taxed.
- An increase in τ^K reduces the interest rate as capital today can be translated into less capital tomorrow. This increases period 1 consumption c_1. But in equilibrium, the interest rate should decline.

In general, taxation on future consumption or taxation on wealth will increase interest rate, but taxation on capital income will decrease interest rate. Lump-sum taxation has no effect to the equilibrium.

9.3 Comparison of Different Taxation in a Continuous-Time Deterministic Economy

We only have households in the economy that maximizes

$$E[\int_0^\infty e^{-\rho t}\ln(c_t)dt]$$

Capital produce output at AK_t. The quantity of capital is subject to

$$\frac{dK_t}{K_t} = (\mu^K + \phi(i_t) - \delta)dt + \sigma^K dB_t$$

where $\phi(i_t)$ is the capital accumulation that reflects a concave adjustment cost, i.e. $\phi''(i_t) < 0$. Bankers and households operate capital at different productivity A^H and A^L. Because we don't have any uncertainty, the price of capital is constant. The gain of holding capital is from its productivity and investment growth, while the loss is by investment cost and depreciation. To solve the equilibrium, we note that with price p at banker wealth share w, we can calculate the fraction of banker capital holding ψ from

$$\psi = \frac{i(p) + \rho p - A^L}{A^H - A^L}$$

With

$$\frac{x^K}{y^K} = \frac{\psi(1-w)}{(1-\psi)w}$$
$$wx^K + (1-w)y^K = 1$$

we get

$$x^K = \frac{\psi}{w}$$
$$y^K = \frac{1-\psi}{1-w}$$

Denote

$$\mu^R = \mu^p + \mu^K + \phi(i(p)) - \delta + \sigma^K \sigma^p - \frac{i(p)}{p} - r$$

Then volatilities can be solved from FOCs of x^K and y^K

$$x^K = \frac{\mu^R + \frac{A^H}{p}}{(\sigma^K + \sigma^p)^2}, \quad y^K = \frac{\mu^R + \frac{A^L}{p}}{(\sigma^K + \sigma^p)^2}$$

$$\Rightarrow (\sigma^K + \sigma^p)^2 = \frac{(A^H - A^L)}{(y^K - x^K)p}$$

With volatility σ^p, we can solve p_w as follows

$$p_w = \frac{\sigma^p}{w(1-w)(x^K - y^K)(\sigma^K + \sigma^p)}$$

Lump-Sum Taxation

Lump-sum taxation doesn't affect the individual optimization problem, but might influence the ratio of banker wealth to household wealth. If the aggregate taxation on banker over household is proportional to their total wealth ratio, then Ricardian equivalence should hold in this economy, which implies the same consumption process with and without government bonds and taxation. To see why Ricardian equivalence holds, we can solve the individual household and banker's optimization problem in a different way. As long as the non-negativity constraint of wealth is never hit, each individual should consume exactly the same way as without taxation. Because state variables are not affected by the taxation as well, we should have exactly the same consumption stream for household and bankers, which implies the same interest rate and same asset prices.

However, I cannot solve the problem using dynamic programming, because now the value function is no longer in the log form due to lump-sum taxation.

Capital Income Taxation

Suppose the government imposes capital taxation τ_t^K for excess capital returns. For bankers, the taxation is on $\bar{R}^K - r$. For households, the taxation is on $\underline{R}^K - r$. This taxation maintains the property of log value function, but changes the portfolio choices into

$$x^K = (1-\tau^K)\frac{\mu^R + \frac{A^H}{p}}{(\sigma^K + \sigma^p)^2}$$

$$y^K = (1-\tau^K)\frac{\mu^R + \frac{A^L}{p}}{(\sigma^K + \sigma^p)^2}$$

where τ^K is solved from

$$\tau^K\left(x^K w(\mu^R + \frac{A^H}{p}) + y^K(1-w)(\mu^R + \frac{A^L}{p})\right) = \frac{Q}{p+Q}r$$

and the volatility equation becomes

$$(\sigma^K + \sigma^p)^2 = \frac{(1-\tau^K)(A^H - A^L)}{(y^K - x^K)p}$$

With positive capital taxation, bankers suffer more, and the equilibrium w tends to be smaller. To solve this problem, I have to solve a second order PDE with the following sequence.

First, from p and w, solve x^K, y^K.

Second, by volatility condition and p_w, we can solve price volatility

$$\sigma^p = \frac{p_w w(1-w)(x^K - y^K)}{1 - p_w w(1-w)(x^K - y^K)}\sigma^K$$

Third, we solve the capital taxation from

$$(\sigma^K + \sigma^p)^2 = \frac{(1-\tau^K)(A^H - A^L)}{(y^K - x^K)p}$$

Fourth, we can solve μ^R from the portfolio choice

$$x^K = (1-\tau^K)\frac{\mu^R + \frac{A^H}{p}}{(\sigma^K + \sigma^p)^2}$$

Fifith, the interest rate r can be solved from

$$\tau^K\left(x^K w(\mu^R + \frac{A^H}{p}) + y^K(1-w)(\mu^R + \frac{A^L}{p})\right) = \frac{Q}{p+Q}r$$

which in combination with definition of μ^R gives us the growth rate of price

$$\mu^p = \mu^R - \left(\mu^K + \phi(i(p)) - \delta + \sigma^K\sigma^p - \frac{i(p)}{p} - r\right)$$

Finally, with interest rate solved, we can plug in the equilibrium condition for μ^p, which gives us p_{ww}.

In summary, with capital taxation, we not only have distortion "holding claim to the future reduces capital price", but also on the relative wealth between bankers and households. Because bankers hold more capital than households and have higher returns, capital taxation will reduce the relative portfolio share of bank-held capital, which reduces productivity and welfare.

Consumption Taxation

Consumption tax will not affect the log form of the value function. However, it affects the relation between consumption and wealth. With consumption tax τ^c, we have

$$C^h = \frac{\rho W^h}{1-\tau^c}, \quad C = \frac{\rho W}{1-\tau^c}$$

As a result, the equilibrium price equation becomes

$$\rho(W^h + W)\frac{1}{1-\tau^c} = (\psi A^H + (1-\psi)A^L)K - iK$$

which implies

$$\rho\frac{p+Q}{1-\tau^c} = (\psi A^H + (1-\psi)A^L) - i(p)$$

To solve consumption tax, we have

$$\tau^c(C^h + C) = rQK$$

$$\Rightarrow \frac{\rho\tau^c}{1-\tau^c} = r\frac{Q}{(p+Q)}$$

I cannot find a solution method to this system. But we can analyze the effects of consumption taxation. In addition to the typical channel "holding claim to the future reduces capital price", the taxation on price has an additional force to put down price further. Thus it has "double" distortion between today and the future.

Wealth Taxation

This is the least distorting way of taxation, when lump-sum is not allowed. This is the current taxation scheme I am using. With investment, the interest rate is affected.

Jumps in Taxation

In the full model, we have to allow jumps in taxation, because states of the model could jump. However, taxations based on flow variables, such as capital income taxation and consumption taxation, are not able to achieve jumps.

The Difference Between the Discrete Time Model and the Continuous Time Model

The rental cost of capital is

$$r^K = A - \delta$$

As a result, the interest rate is fixed. Market clearing implies

$$\rho W = (A - i)K$$

$$W = (p + Q)K$$

The investment sector could translates between consumption and capital 1 by 1, thus making $p = 1$. Wealth growth is thus growing at the same rate as capital growth. Wealth growth is

$$\frac{dW}{W} = -\tau^w dt + x^K r^K dt + x^K \sigma^K dB_t + x^g r dt + (1 - x^K - x^g) r dt - \rho$$

$$\frac{dW}{W} = -\tau^w dt + x^K (A - \delta) dt + x^K \sigma^K dB_t + x^g r dt + (1 - x^K - x^g) r dt - \rho$$

where

$$\tau^w W = rQK$$

Because

$$x^K = \frac{p}{p + Q}$$

We have

$$\frac{dW}{W} = \frac{p}{p + Q}(A - \delta) dt + \frac{p}{p + Q}\sigma^K dB_t - \rho$$

As a result, higher quantity of government bond Q, higher wealth and thus consumption today, but lower investment and lower future capital.

$$dK = (i - \delta) K dt$$

The interest rate can be solved from

$$\frac{A - \delta - r}{(\sigma^K)^2} = \frac{p}{p + Q}$$

which implies that an increase in Q raises interest rate r. Under the general setting with investment tied to prices, an increase in government bond with wealth taxation decreases future investment and wealth growth. In the continuous-time model, consumption is a flow variable, while capital is a stock variable. Depending on the state of the world, the productivity of capital will change, which causes price change and other fluctuations. It is much more reasonable to have a price of capital.

If instead we have consumption taxation,

$$\rho \frac{p + Q}{1 - \tau^c} = A^H - i(p)$$

$$\frac{A - \delta - r}{(\sigma^K)^2} = \frac{p}{p + Q}$$

$$\frac{\rho\tau^c}{1-\tau^c} = r\frac{Q}{(p+Q)}$$

The first equation implies that p decreases in τ^c, while the second and third results in

$$\frac{\rho\tau^c}{1-\tau^c} = \left(A - \delta - \frac{p}{p+Q}(\sigma^K)^2\right)\frac{Q}{(p+Q)}$$

which also implies an decrease of price p in τ^c. Note that $\tau^c \in [0,1)$, thus $r > 0$ and we always get a decline in price for any positive τ^c.

10 Credible Government Policies

In this section, I will first set up a simple first period economy to illustrate concepts. Then I will introduce a fully dynamic model to study credible government policies. The content closely follows the Credible Government Policies chapter in Ljungqvist and Sargent. The main technique is about whether the government can have credible policies, and how a social planner can solve the optimal planning problem recursively. We know that the Ramsey plan is quite limited and hard to solve. By using a dynamic programming approach, we can solve much more complicated models, including time inconsistency, asymmetry information, repudiation, etc.

10.1 The One-Period Economy

A unit mass of households live in the economy, each of whom chooses an action $\xi \in \mathcal{X}$. A government chooses an action $y \in \mathcal{Y}$. The set $\mathcal{X}$ and $\mathcal{Y}$ are compact sets so that supremum is within the range. Let the aggregate level of ξ among households be x, which also belong to $\mathcal{X}$. Household utility function is $u(\xi, x, y)$, and the government is benevolent, which means government wants to maximize the sum of household utility. The payoff function $u(\xi, x, y)$ is strictly concave and continuously differentiable. We note that if each household chooses the same ξ, then in equilibrium we should have $\xi = x$, and the aggregate utility is just $u(x, x, y)$.

Then I will introduce and compare three types of concepts: (1) A dictatorial economy. (2) A competitive economy with credible government. The the government optimization problem is called the Ramsey problem. (3) A Markov or Nash economy, where the government has no commitment. The government will choose its optimal policy in response to household choices.

Pareto rank: (1) > (2) > (3). The basic idea is that from (1) to (3), we are adding more and more constraints to the government optimization problem.

First, the dictatorial problem is

$$\max_{x,y}\ u(x, x, y)$$

which is an unconstrained optimization problem, except for the domain requirement $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.

Second, the Ramsey problem is

$$\max_{(x,y)\in C}\ u(x, x, y)$$

where

$$C = \{(x, y) | \max_{\xi} u(\xi, x, y) = u(x, x, y)\}$$

Third, the non-commitment problem is

$$\max_{(x,y)\in C\cap H} \quad u(x, x, y)$$

where

$$H = \{(x, y) | \max_{\eta} u(x, x, \eta) = u(x, x, y)\}$$

Because restrictions are more strict from (1) to (3), the objective functions should be smaller.

10.2 An Infinitely Repeated Economy

One-shot Deviation

First, let's review the intuitions of one-shot deviation. In a finite repeated game, we can start from the last period. From any subgame, by no one-shot deviation, the last period has no deviation at all. Then in the second to the last period, we only need to check that period deviation because any subgame in the next period is going to have no deviation. Recursively, we find that we only need to check the deviation in the first period. Because of the repeated game structure, checking the current period deviation is the same for all periods.

Next, let's take the horizon to infinity. Because of the discounting factor $\beta \in (0, 1)$, once we take the time horizon long enough, deviations in the tail part is going to be dominated by any deviations in the current period. For any deviation that generates a strictly positive difference in the current period, it will dominate the tail part. As a result, one shot deviation is enough for any deviation.

Remark 13. *For any proof, I should get a verbal proof first to fully digest the intuitions. Then go to the math rigorously. The one shot deviation is only applicable to repeated games because of the same structure.*

Why We have Better Outcomes with "Carrot and Stick"?

In all the APS machinery, we have the so-called carrot and stick element. When the government goes off path, the punishment is pushed to the maximum. The stick is bigger than Nash punishment.

The value function is

$$V(\sigma_t) = (1 - \delta)r(x_t, y_t) + \delta V(\sigma_{t+1})$$

(1) Infinite repetition of Nash. Denote the Nash strategy as σ^N, and stage Nash return as $v^N = r(x^N, y^N)$. Then the total expected value is just v^N. Sum it up and then multiply by $1 - \delta$.

(2) Better outcome with trigger strategies. A trigger strategy means the punishment is triggered when a certain strategy is not played. Then to sustain $(\tilde{x}, \tilde{y})$, we can define the following strategy: Play $\tilde{\sigma}$ if $(x, y) = (\tilde{x}, \tilde{y})$, and otherwise play the maximum Nash punishment σ^N. The IC constraint becomes

$$(1-\delta)r(\tilde{x}, \tilde{y}) + \delta\tilde{v} \geq \max_{\eta} \left\{(1-\delta)r(\tilde{x}, \eta) + \delta v^N\right\}$$

which is equivalent to

$$\underbrace{(1-\delta)\left(r(\tilde{x}, \eta) - r(\tilde{x}, \tilde{y})\right)}_{\text{temptation to deviate}} \geq \underbrace{\delta(\tilde{v} - v^N)}_{\text{maximum punishment}} \qquad \text{for all } \eta$$

where the maximum value is sustained by the Nash reversion.

(2) Even Better Outcome with "Carrot and Stick".

In general, we can get a lower continuation value than the Nash, which itself is sustained by continuation subgames. Then the incentive constrained is further relaxed. Thus the value attainable is larger.

Remark 14. *We might recall the "folk theorem" intuition that in repeated games, almost anything is sustainable. However, it is under the premise that the discount rate δ is very close to 1. In general, for $\delta < 1$, we can only sustain a subset, which makes the problem still interesting.*

The APS Machinery – Finding the Set of Values

The APS machinery starts with a set of continuation values for all possible SPE (subgame perfect equilibrium).

$$V = \{V_g(\sigma) | \sigma \text{ is an SPE}\}$$

where

$$V_g(\sigma) = (1-\delta)\sum_{t=0}^{\infty} \delta^t r(x_\sigma, y_\sigma)$$

By definition, for any $v \in V$ and $v_1 \in V$, $v_2 \in V$ and $(x, y) \in C$, we have

$$v = (1-\delta)r(x, y) + \delta v_1 \geq (1-\delta)r(x, \eta) + \delta v_2$$

for any $\eta \in \mathcal{Y}$.

Then we define a tuple (x, y, w_1, w_2) as admissible strategies w.r.t. a set $W \subset \mathbb{R}$, such that $w_1 \in W$, $w_2 \in W$,

$$(x, y) \in C$$

and

$$(1-\delta)r(x, y) + \delta w_1 \geq (1-\delta)r(x, \eta) + \delta w_2$$

for any $\eta \in \mathcal{Y}$.

The machinery works as follows. We are going to prove that V is the largest fixed point under certain operator. As a result, once we have a decreasing property of the operator, the algorithm is going to generate V. Knowing the set V, we can design algorithms that deliver the optimal strategies corresponding to each point in the value set V.

Define an operator B on any subset of $W \subset \mathbb{R}$ as the set of value function

$$(1-\delta)r(x,y) + \delta w_1$$

for any admissible strategies (x, y, w_1, w_2) w.r.t. W. Then we are going to have the following property.

Monotonicity: If $W \subset W' \subset \mathbb{R}$, then $B(W) \subset W'$.

Compactness: B maps compact sets W into compact sets $B(W)$.

Self-Generation Sets belong to Continuation Value Set. For any W with the self-generation property, i.e. $W \subset B(W)$, we have $B(W) \subset V$.

Proof. For any self-generating set, we can construct strategies sequentially, for both on and off the equilibrium path (note: it is important to consider both). This is because for any $w \in B(W)$, we can get $w_1 \in W$ and $w_2 \in W$, and strategies (x, y) such that the tuple (x, y, w_1, w_2) are admissible w.r.t. W. Then using $W \subset B(W)$, we have $w_1 \in B(W)$ and $w_2 \in B(W)$, which implies that we can continue constructing strategies further. Repeat the above infinitely and we get the whole strategy profile for the subgame. Thus the above construction results in a subgame perfect equilibrium starting with value w, which implies $w \in V$. Thus $B(W) \subset V$. □

Next, from the fact that the continuation of any subgame is also a subgame, we have

$$V \subset B(V)$$

Then the above property implies $B(V) \subset V$, which results in $V = B(V)$. Thus V is a fixed point of B. Next, it is easy to see that V is the largest fixed point. Assume not, then we have V' that $B(V') = V'$ but $V \subset V'$. However, since V' is self-generating, we must have $V' \subset V$, which implies $V' = V$. Thus V is the largest fixed point.

The fact that V is the largest fixed point means that if starting from some large set $W \supset V$, and iteratively apply B, we will finally converge to V, not stopping at an interim stage, since otherwise it violates the maximum fixed point property.

Starting from a compact set $W_0 = (\underline{w}_0, \bar{w}_0) \supset V$, because B maps compact set to another compact set and $B(W_0) \subsetneq W_0$ unless $W_0 = V$, the first step will generate a strictly smaller set with boundaries inside $(\underline{w}_0, \bar{w}_0)$. The same property holds for the next step iteration. Thus the algorithm will finally converge to V. Furthermore, because the iteration preserves compactness, it implies that V is also a compact set, which means we can find a best point

$$\bar{v} = \max_{v \in V} v$$

Remark 15. *The above applies to an agent without commitment. With commitment, we need to*

The APS Machinery – Finding Strategies

Next, after we have got the set V, we want to find strategies associated with the best and worst SPE. The worse SPE value $\underline{v}$ is satisfies

$$\begin{aligned}&\underline{v} = \min_{x,y,v\in V}\{(1-\delta)r(x,y)+\delta v\}\\&s.t.\\&(x,y)\in C\\&(1-\delta)r(x,y)+\delta v \geq (1-\delta)r(x,\eta)+\delta\underline{v}\end{aligned}$$

The second contraint must be binding. Thus we have

$$\begin{aligned}&\underline{v} = \min_{x,y}\left\{\max_{\eta}\{r(x,\eta)\}\right\}\\&s.t.\\&(x,y)\in C\\&(1-\delta)r(x,y)+\delta\bar{v} \geq \max_{\eta}\{(1-\delta)r(x,\eta)+\delta\underline{v}\}\end{aligned}$$

Using the objective function, we can simplify the above further into

$$\begin{aligned}&\underline{v} = \min_{x,y}\left\{\max_{\eta}\{r(x,\eta)\}\right\}\\&s.t.\\&(x,y)\in C\\&(1-\delta)r(x,y)+\delta\bar{v} \geq \underline{v}\end{aligned}$$

Thus we are optimizing over x such that x is within the maximum sustainable set by the best and worse continuation values.

Next, for the best scenario $\bar{v}$, we have

$$\begin{aligned}&\bar{v} = \max_{x,y,v\in V}\{(1-\delta)r(x,y)+\delta v\}\\&s.t.\\&(x,y)\in C\\&(1-\delta)r(x,y)+\delta v \geq (1-\delta)r(x,\eta)+\delta\underline{v} \quad \text{for any } \eta\in\mathcal{Y}\end{aligned}$$

It is easy to see that we should choose $v = \bar{v}$, and thus the problem is transformed into

$$\begin{aligned}&\bar{v} = \max_{x,y}\{r(x,y)\}\\&s.t.\\&\begin{cases}(x,y)\in C\\ \bar{v}\geq (1-\delta)r(x,\eta)+\delta\underline{v} \ \text{ for any } \eta\in Y\end{cases}\end{aligned}$$

11 Principle Agency Problems

11.1 Hidden Type and Screening

A principal is a seller and the agent is a buyer. The principal has all the bargaining power, and designs a contract for the agent with hidden type.

- Principal: $t - c(x)$, where $c(x)$ is the cost of producing x units of good.
- Agency: $v(x, \theta) - t$.

The space of all types is Θ, an interval in $\mathbb{R}$, and the action space is X. The transfer t can depend on x, but not θ, since θ is the hidden information. For each schedule $t(x)$, the agent problem is

$$x(\theta) \in \arg\max_{x \in \mathcal{X}} \{v(x, \theta) - t(\theta)\}$$

Full Separation

Assuming that $v'_x(x, \theta)$ strictly increases in θ (or if both are differentiable, $v''_{x\theta}(x, \theta) > 0$), and X is an interval, and $t(\cdot)$ is differentiable, then $x(\theta)$ strictly increases for $x \in interior(X)$.

A simpler rule is supermodularity and Topkis' theorem. By only requiring that $v'_x(x, \theta)$ increases with θ, the theorem delivers a strong set order, but not full separation.

Full Information Benchmark

In the full information benchmark, the optimization problem is

$$\max_x \{v(x, \theta) - c(x)\}$$

Formulation

We can formulate the problem as follows

$$\begin{array}{l} \max \int_{\underline{\theta}}^{\bar{\theta}} (t(\theta) - c(x(\theta)))f(\theta)d\theta \\ s.t. \\ \begin{cases} v(x(\theta), \theta) - t(\theta) \geq v(x(\hat{\theta}), \theta) - t(\hat{\theta}) \text{ for any } \hat{\theta} \\ v(x(\theta), \theta) - t(\theta) \geq 0 \end{cases} \end{array}$$

The key difficulty is the continuum of IC restrictions.

Remark 1. *The above formulation relies on the revelation principle in mechanism design.*

Downward Binding IC

To get intuitions why IC condition is typically downward binding, we can start with the first-best allocation where $t(\theta) = 0$, and $v'_x(x, \theta) = c'(x)$. The IR constraint is binding, which indicates $v(x(\theta), \theta) - t(\theta) = 0$ for all θ.

We assume that $x = 0$ is the outside option that yields the same output regardless the type of the agent, i.e. $v(0, \theta) = 0$ for all θ. Then from the single cross property of the value function $v(x, \theta)$, we find that

$$v(x, \theta) > v(x, \theta')$$

for $\theta > \theta'$. As a result,

$$v(x(\theta'), \theta) > v(x(\theta'), \theta') = t(\theta')$$

which means the deviation value is

$$v(x(\theta'), \theta) - t(\theta') > 0$$

Thus the agent has an incentive to deviate downward and report a lower type. Intuitively, because of the increasing difference property, the same action has higher value for higher types, which means that higher types want to imitate lower types. But this is not the other way around, since

$$v(x(\theta'), \theta'') < v(x(\theta'), \theta') = t(\theta')$$

for $\theta'' < \theta'$, which implies the deviation to a higher type provides value

$$v(x(\theta'), \theta'') - t(\theta') < 0$$

Thus the agent has no incentive of reporting af higher type.

With the same intuition, we find that the IR condition only binds to the lowest type, because for any $\theta \in \Theta$,

$$v(x(\theta), \theta) - t(\theta) > v(x(\underline{\theta}), \theta) - t(\underline{\theta}) > v(x(\underline{\theta}), \underline{\theta}) - t(\underline{\theta}) \geq 0$$

Thus once the IR constraint is binding for the lowest type.

Characterizing the IC Constraint with FOC

A major breakthrough is by Mirrlees (1971) in terms of solving the principal agency problem with hidden types. The main technique is to express the incentive constraints as first order conditions, i.e. transforming a global constraint into local constraints.

Define the utility of type θ reporting $\hat{\theta}$ as

$$\Phi(\hat{\theta}, \theta) = v(x(\hat{\theta}), \theta) - t(\hat{\theta})$$

The IC constraint is equivalently defined as

$$\Phi(\theta, \theta) \geq \Phi(\hat{\theta}, \theta), \text{ for any } \hat{\theta} \in \Theta \tag{IC}$$

Q: Can we use FOC now? A: No. Because we even don't know whether $\phi(\hat{\theta}, \theta)$ is differentiable over $\hat{\theta}$. It is possible that $x(\theta)$ is not differentiable at all.

Important observation: $\phi(\hat{\theta}, \theta)$ is differentiable w.r.t. θ. Can we transform the above problem into variations into θ?

Yes! Just look at the constraint in a different way. Define $U(\theta) = \phi(\theta, \theta)$. For any θ, the constraint (IC) holds. Thus we have

$$\Phi(\hat{\theta}, \theta) - U(\theta) \leq 0$$

for any $\theta \in \Theta$. We know that the equality is reached at $\theta = \hat{\theta}$, which is the maximum value of

$$\Phi(\hat{\theta}, \theta) - U(\theta)$$

Thus at $\theta = \hat{\theta}$, we should have the first order condition

$$\Phi_2'(\hat{\theta}, \hat{\theta}) = U(\hat{\theta})$$
$$\Rightarrow v_\theta(x(\theta), \theta) = U_\theta(\theta)$$

With the assumption that v_θ is uniformly bounded, we know that $U(\theta)$ is Lipshitz continuous (we need only a weaker condition of absolute continuity), which implies the following integral form

$$U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} v_\theta(x(u), u)du \tag{ICFOC}$$

Remark 2. *This can be also derived from the generalized Envelope condition, because we know that*

$$U(\theta) = \sup_{\hat{\theta}}\{v(x(\hat{\theta}), \theta) - t(\hat{\theta})\}$$

The generalized Envelope theorem in Milgrom and Segal (2002) implies that $U'(\theta) = v_\theta(x(\theta), \theta)$.

We want to fully characterize the (IC) constraint. The first order condition is not enough, because so far we have no restriction on $x(\theta)$. By Topkis theorem, we know that $x(\theta)$ is nondecreasing in θ. Denote this as the monotonicity constraint

$$x(\theta) \text{ is nondecreasing in } \theta \tag{M}$$

Proposition 1 (Equivalent Representation of Incentive Constraints)**.** *A pair of strategy and transfer* $(x(\cdot), t(\cdot))$ *is incentive compatible, i.e. satisfies (IC), if and only if both (ICFOC) and (M) hold.*

Proof. The "only if" part has already been proved.

To get the "if" part, we note that the IC is equivalent to

$$\phi(\hat{\theta}, \theta) = \Phi(\hat{\theta}, \theta) - U(\theta)$$

maximized at $\theta = \hat{\theta}$, which can be derived by studying the FOC of $\phi(\hat{\theta}, \theta)$.

$$\frac{\partial \phi(\hat{\theta}, \theta)}{\partial \theta} = \Phi_\theta(\hat{\theta}, \theta) - U_\theta(\theta) = v_\theta(x(\hat{\theta}), \theta) - v_\theta(x(\theta), \theta)$$

Because $x(\cdot)$ is monotone and $v_\theta(x, \theta)$ increases in x, the above derivative reaches 0 at $\theta = \hat{\theta}$ and increases with $\hat{\theta}$, which means the maximum is reached at $\theta = \hat{\theta}$. □

Solving the Reformulated Problem

In general we have to use the Pontryagin's maximum principal for the contract design problem. However, we are able to represent the global incentive restriction into two local restrictions (1) First order condition. (2) Choice monotone in type. Thus we actually only need to solve a much simpler problem. Reformulate the problem into the following.

$$\max_{x(\cdot)} \int_{\underline{\theta}}^{\bar{\theta}} (v(x(\theta), \theta) - U(\theta) - c(x(\theta)))f(\theta)d\theta$$
$$s.t.$$
$$\begin{cases} x(\theta) \text{increasing} \\ U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} v_\theta(x(u), u)du \\ U(\underline{\theta}) \geq 0 \end{cases}$$

Note that the IR constraint is binding, and thus $U(\underline{\theta}) = 0$. With the second equation, we have

$$\begin{aligned}
&\int_{\underline{\theta}}^{\bar{\theta}} U(\theta)f(\theta)d\theta = \int_{\underline{\theta}}^{\bar{\theta}}\int_{\underline{\theta}}^{\theta} f(\theta)v_\theta(x(u),u)dud\theta\\
&= \int_{\underline{\theta}}^{\bar{\theta}}\int_{u}^{\bar{\theta}} f(\theta)v_\theta(x(u),u)d\theta du\\
&= \int_{\underline{\theta}}^{\bar{\theta}} (1-F(u))v_\theta(x(u),u)du\\
&= \int_{\underline{\theta}}^{\bar{\theta}} \frac{(1-F(\theta))}{f(\theta)}v_\theta(x(\theta),\theta)\cdot f(\theta)d\theta
\end{aligned}$$

Let's first drop the monotonicity constraint on x. Then the original problem is transformed into

$$\max_{x(\cdot)} \int_{\underline{\theta}}^{\bar{\theta}} \left(v(x(\theta),\theta) - \frac{(1-F(\theta))}{f(\theta)}v_\theta(x(\theta),\theta) - c(x(\theta))\right)f(\theta)d\theta$$

which is a pointwise optimization over x. We can define the *virtual cost* as

$$\tilde{c}(x) = \underbrace{v_\theta(x,\theta)}_{\substack{\text{information rents by each}\\ \text{agent with binding IC at } \theta}} \cdot \underbrace{\frac{(1-F(\theta))}{f(\theta)}}_{\text{relative mass of agents above } \theta} + c(x)$$

Intuitively, $v_\theta(x,\theta)$ measures the information rent of each agent with a binding IC by imitating θ, and the derivative is a measure of marginal benefit of misreporting types. As we have already seen, the IC constraint only binds downward, and thus imitating θ is only binding for types higher than θ. Consequently, we should measure the relative mass of agents above θ. Multiplying the unit gain v_θ with the relative mass, we get the total incentive cost of choice x.

Remark 3 (Interaction of Choices and Types)**.** *In this setup, we do not have risks. Thus risk aversion is not important. However, the interaction between choice x and type θ is the key. For example, if the agent has utility $v(x,\theta) = h(x) + g(\theta)$, then the optimization for the agent becomes*

$$\max_{\hat{\theta}} \left\{h(x(\hat{\theta})) + g(\theta) - t(\hat{\theta})\right\}$$

which results in the same x for all types θ. Thus agent type is not distinguishable at all. Then the optimal contract will be the same as optimizing over x,

$$\max_{x} \left\{E_{\tilde{\theta}}[g(\tilde{\theta})] + h(x) - c(x)\right\}$$

Another consideration is about which agent should become the principal.

Remark 4 (Who Should be the Principal?)**.** *The choice of who to become principal is not important in this setup, since we only care about socially efficient allocations, i.e. Pareto optimal allocations. For two agents with concave utility, the Pareto optimization problem*

$$\begin{aligned}
&\max_{c_1,c_2} \{\lambda u_1(c_1) + (1-\lambda)u_2(c_2)\}\\
&s.t. \quad c_1 + c_2 \leq y
\end{aligned}$$

is the same as a constrained problem

$$\begin{aligned}
&\max_{c_1,c_2} u_1(c_1)\\
&s.t.\\
&\begin{cases} u_2(c_2) \geq \underline{u}_2\\ c_1 + c_2 \leq y\end{cases}
\end{aligned}$$

with appropriate $\underline{u}_2$. Thus which agent should be the principal only matters for the relative distribution of surplus, not about social efficiency.

Remark 5 (Asymmetric Information or Contractability?)**.** *In the hidden information model, is it important to have asymmetric information, or contractability? Which one is more fundamental? In these contracting problems, as long as the principal does not have an action that influences the agent outside the contract, the only thing that matters is contractability. Even if the principal observes all information, if type is not contractible, results are the same. Thus the fundamental friction is contractability, not asymmetric information. A similar argument holds for hidden action models.*

11.2 Hidden Action

This scenario is more common, especially in macro models. The agent is risk averse, with a Bernoulli utility function defined on action a and payment t,

$$U(a,t) = v(t) - g(a)$$

where $v' > 0$ and $v'' < 0$. The principal is risk neutral with Bernoulli utility function

$$\tilde{x}(a) - t$$

where $\tilde{x}(a)$ is a random variable that depends on agent action a. Importantly, we assume that action a is not contractible, but only output x is contractible.

Remark 6 (Importance of Contractibility)**.** *Again, in this setup, what matters is contractability, not asymmetric information. We can have a setup where the principal do not observe either action or outcome, but the contract can depend on agent's action and enforcible. Then the first best result follows. The fact that some information is not observable by the principal but is contractible can be interpreted as some "law enforcement" mechanism outside the model. Thus regardless of information asymmetry, contractibility determines whether we can achieve first best results.*

The First-Best Benchmark

To understand the model, we start with a benchmark where the effort a is contractible. Then the problem is

$$\begin{aligned} &\max_{t,a} E[x(a) - t] \\ &s.t. \\ &v(t) - g(a) \geq \bar{u} \end{aligned}$$

The second constraint is binding, and we get

$$\max_{a} E[\tilde{x}(a)] - v^{-1}(\bar{u} + g(a))$$

Since the inverse of a strictly increasing and concave function is convex, if $g(a)$ is increasing, then $v^{-1}(\bar{u} + g(a))$ is convex, and the optimization is convex optimization (objective value is concave for the maximization). Thus we can directly take first order condition and get the best action, denoted as a^*. Uniqueness is guaranteed by the strict concavity of v.

The Full Model

In the full model where contracts can only depend on outcome but not action, the optimization problem of the principal is

$$\max_{t(\cdot),a} E[x(a) - t(\tilde{x}(a))]$$
$$s.t.$$
$$E[v(t(\tilde{x}(a))) - g(a)] \geq \bar{u} \tag{IR}$$
$$E[v(t(\tilde{x}(a))) - g(a)] \geq E[v(t(\tilde{x}(a'))) - g(a')] \text{ for any } a' \tag{IC}$$

When the First-Best is Implementable in the Full Model?

Under the following three cases, the first-best action a^* is implementable.

- The first best action a^* is the least cost action, i.e.

  $$g(a^*) = \min_a g(a)$$

  Then by setting $t(x) = t^*$, the incentive constraints are trivially satisfied. The first best is achieved.

- The agent is risk neutral. Then the utilities of principal and agent are consistent. The principal can use a sell-out contract $t(x)$ that is linear in x to generate consistent incentive.

- Shifting support. Suppose that there exists a set S such that $\mathbb{P}(\tilde{x}(a^*) \in S) = 0$, but $\mathbb{P}(\tilde{x}(a) \in S) > 0$ for any $a \neq a^*$. Then the principal can punish the agent harshly outside S and reward highly inside S. Because on equilibrium path, S is never reached, the contract is *effectively constant*, or constant almost surely.

Remark 7. *In all above scenarios, both first best action and first best utilities are achieved.*

Two Actions

Let's get some intuitions from a simple version, where action space $A = \{H, L\}$, with $g(H) > g(L)$, but it is preferable to implement $a^* = H$. The problem can be written as

$$\max_{t(\cdot)} E[\tilde{x}(a^H) - t(\tilde{x}(a^H))]$$
$$s.t.$$
$$E[v(t(\tilde{x}(a^H)))] - g(a^H) \geq E[v(t(\tilde{x}(a^L)))] - g(a^L)$$
$$E[v(t(\tilde{x}(a^H)))] - g(a^H) \geq 0$$

Both IR and IC are binding. If IC is not binding, then the optimal contract is a constant payoff $\bar{t}$, which violates the IC constraint. As a result, the IC constraint is binding. For the IR constraint, if it is not binding, then we can reduce the transfer $t(\cdot)$ uniformly by a small number $\varepsilon > 0$ and strictly increase principal's payoff. Thus IR is also binding.

Denote the Lagrangian on the IC constraint as μ and λ. The Lagrangian is

$$E[\tilde{x}(a^H) - t(\tilde{x}(a^H))] + \mu(E[v(t(\tilde{x}(a^H))) - v(t(\tilde{x}(a^L)))] - g(a^H) + g(a^L))$$
$$+\lambda(E[v(t(\tilde{x}(a^H)))] - g(a^H))$$

where $\mu > 0$ and $\lambda > 0$ because of binding constraints.

We would like to take first order derivative w.r.t. $t(\cdot)$. To do so, we need to explicitly write the above expectations with density function $f(x|a)$, where $a \in \{A^H, A^L\}$. Dropping the irrelevant terms, we get

$$-\int t(x)f(x|a^H)dx + \mu\int \big(v(t(x))f(x|a^H) - v(t(x))f(x|a^L)\big)dx + \lambda\int v(t(x))f(x|a^H)dx$$
$$= \int f(x|a^H)\left(v(t(x))\left(\mu - \mu\frac{f(x|a^L)}{f(x|a^H)} + \lambda\right) - t(x)\right)dx$$

Thus the FOC on $t(\cdot)$ can be studied pointwise.

$$v'(t(x))\left(\mu - \mu\frac{f(x|a^L)}{f(x|a^H)} + \lambda\right) = 1 \tag{18}$$

Since $v' > 0$, we must have

$$\mu - \mu\frac{f(x|a^L)}{f(x|a^H)} + \lambda > 0$$

which implies that in the relaxed problem, the FOC (18) yields an optimal solution, since $v(t)$ is concave. Thus we can use Topkis theorem to study how t changes w.r.t. the likelihood ratio

$$\mathcal{L} = \frac{f(x|a^L)}{f(x|a^H)}$$

With this notation, we can rewrite the optimization as

$$\max_t h(t, L) = v(t)\left(\mu - \mu\mathcal{L} + \lambda\right) - t$$

Because

$$\frac{\partial^2 h(t, \mathcal{L})}{\partial t \partial \mathcal{L}} = -\mu v'(t) < 0$$

the transfer $t(x)$ is smaller when $\mathcal{L}$ is larger. Intuitively, the likelihood ratio $\mathcal{L}$ contains information about the likelihood of the undesired action a^L and thus higher likelihood ration should result in lower transfer payment.

Remark 8. *Rigorously, the likelihood ratio should not be interpreted the same as in statistics, because only a^H is taken in equilibrium. We should interpret the ratio $\mathcal{L}$ as an incentive for deviation, which is punished by a lower payment t when $\mathcal{L}$ is larger.*

To further relate the problem to statistical principles, we can assume that principal have another signal $\tilde{y}$ with the same support under different action a (no shifting support) and the contract can depend on $\tilde{y}$. Then if $\tilde{x}$ is a sufficient statistic for parameter a given $(\tilde{x}, \tilde{y})$, i.e. $f(y|x, a) = f(y|x)$, then the optimal contract should only depend on $\tilde{x}$.

This immediately implies that the principal will not use randomized compensation scheme, because the added randomness does not help distinguish different actions, and all information is contained in $\tilde{x}$.

A Continuum of Actions

With a continuum of actions, the problem is harder because we need to satisfy an uncountable number of constraints. In the literature, we have different assumptions to make sure the IC constraint is well satisfied.

Short-Term Contracts v.s. and Long-Term Contracts

Long-term contract is a contract that can depend on long-term events. Compared with short-term contracts, long-term contracts have a strictly larger contract space. Then by the property of maximization problems, the objective function associated with long-term contracts is at least as large as that of short-term contracts. Thus long-term contracts **dominates** short-term contracts.

This should not be confused with long-term debt and short-term debt, which are fixed instruments, and they have their own relative advantages built in their definitions.

11.3 Hidden Action in Continuous Time

When moving to dynamic contract with hidden actions in continuous time, the hidden action problem actually becomes easier to solve, because of the local nature of uncertainty. Thus we will not solve anything in expectation. Moreover, a two action model is easier to solve compared to a model with a continuum of actions, where we need more assumptions and have to carefully deal with boundary conditions. Refer to Sannikov (2008) for a model with a continuum of actions, and DeMarzo and Sannikov (2006) for a model with two actions.

Basic Setup

The agent produces output $X = \{X_t, t \geq 0\}$.

$$X_t = \int_0^t a_s ds + Z_t$$

where Z is a standard Brownian motion under measure $\mathbb{P}$, and a is "effort". The agent likes consumption but doesn't like effort

$$U_0^{c,a} = \mathbb{E}[\int_0^\infty re^{-rt}(u(c_t) - h(a_t))dt]$$

The principal has an objective

$$P_0^{a,c} = E[\int_0^\infty re^{-rt}dX_t - \int_0^\infty e^{-rt}c_t dt]$$

$$= E[\int_0^\infty e^{-rt}(a_t - c_t)dt]$$

A contract specifies consumption c and recommended effort a, as functions of the history of observed output

$$c_t(\{X_s, 0 \leq s \leq t\}), a_t(\{X_s, 0 \leq s \leq t\})$$

After assigned by the contract $\mathcal{C} = (c, a)$, the agent chooses how much effort to do

$$\tilde{a}_t(\{X_s, 0 \leq s \leq t\})$$

in order to maximize

$$U_0^{c,\tilde{a}} = \mathbb{E}[\int_0^\infty re^{-rt}(u(\tilde{c}_t) - h(\tilde{a}_t))dt]$$

where the consumption process c is affected because X is affected by action a.

Strong and Weak Formulation

We can formulate the problem in different ways. In the basic setup, we have used a strong formulation, which says the hidden action a affects output process $X^a : (\omega, t) \to \mathbb{R}$,

$$X_t^a = \int_0^t a_s ds + Z_t$$

However, this is quite inconvenient, because the effort choice $a_s, s \leq t$ determines the path $X_s^a, s \leq t$, which influences $a_s, s > t$. This double mapping makes the problem intractable.

To work on this setting, we can use a weak formulation, assuming output as a fixed stochastic process $X : (\omega, t) \to \mathbb{R}$ such that $X_t = Z_t$ under $\mathbb{P}$, and action is changing the probability measure with which agents use to evaluate outcomes. Define a measure $\mathbb{P}^a$ generated by action a as an equivalent measure to $\mathbb{P}$ such that Z_t^a is a Brownian motion under $\mathbb{P}^a$, where

$$X_t = \int_0^t a_s ds + Z_t^a$$

In the end, we will find a probability measure $\mathbb{P}^a$ such that Z_t^a is a Brownian motion under that measure, and each agent is evaluating under this measure,

$$U_0^{c,\tilde{a}} = \mathbb{E}^a[\int_0^\infty re^{-rt}(u(\tilde{c}_t) - h(\tilde{a}_t))dt]$$

$$P_0^{a,c} = \mathbb{E}^a[\int_0^\infty e^{-rt}(a_t - c_t)dt]$$

which quantitatively is the same as the strong formulation but is much simpler.

Remark 9 (Why a change of measure only affects the drift term, not the volatility term?)**.** *For any Ito's process $X(\omega, t)$, by taking the quadratic variation of the path $\{X(\omega, t)|0 \leq t \leq T\}$ under $\omega \in \Omega$, we immediately get the process for volatility. As a result, change of measure cannot affect the volatility process, which can be derived from the shape of a path. On the other hand, given a path $\{X(\omega, t)|0 \leq t \leq T\}$ under $\omega \in \Omega$, we are not sure which $\omega' \in \Omega$ generates it. Thus a change of measure will change the interpretation of the drift process.*

Remark 10 (Weak and Strong SDE)**.** *The weak and strong formulations of the problem mimic the weak and strong formulations of SDE. When we say that X is defined by some SDE, e.g.*

$$X_0 = x_0, dX_t = f(X_t)dt + g(X_t)dZ_t$$

where Z denotes a Standard Brownian motion, there are two interpretations

- *In the* <u>*strong formulation*</u>*, we fix a probability space* $(\Omega, \mathbb{P}, \mathcal{F})$*, fix a Brownian motion* Z *on that space, and seek to create a process* X*, measurable w.r.t.* $\sigma(\{Z_t, t \geq 0\})$*, that satisfies the SDE.*

- *In the* <u>*weak formulation*</u>*, we want to find a probability space* $(\Omega, \mathbb{P}, \mathcal{F})$*, a Brownian motion* Z *on that space, and a process* X *measurable w.r.t.* $\sigma(\{Z_t, t \geq 0\})$ *and satisfies the SDE.*

The weak formulation gives us more flexibility since it does not restrict our probability space. There are examples of SDEs where the strong formulation does not have a solution but the weak formulation yields a solution. The weak formulation is more constructive. The distinction between weak and strong formulations also appear in other fields, such as finite difference analysis in physics.

12 Classical Macro Finance Models

In this section, I summarize classical macro finance models that study financial frictions in a macro framework.

12.1 Kiyotaki and Moore (1997)

This paper is a pioneer in studying the collateral channel of financial amplification. The collateral requirement amplifies business cycle fluctuations because in a recession, the income from capital falls, causing the price of capital to fall, which makes capital less valuable as collateral, which limits firms' investment by forcing them to reduce their borrowing, and thereby worsens the recession.

Two key assumptions limit the effectiveness of the credit market in the model.

- First, the knowledge of the "farmers" is an essential input to their own investment projectsthat is, a project becomes worthless if the farmer who made the investment chooses to abandon it.

- Second, farmers cannot be forced to work, and therefore they cannot sell off their future labor to guarantee their debts.

Together, these assumptions imply that even though farmers' investment projects are potentially very valuable, lenders have no way to confiscate this value if farmers choose not to pay back their debts. This endogenously restricts the amount that farmers can borrow, and the constraint is tighter when the price of land declines, further reducing the value of land since farmers are the efficient holders of capital.

The model has two types of players: 1 unit mass of farmers (interpreted as entrepreneurs or firms), and m mass of gatherers (interpreted as households). The farmer's problem is as follows:

$$
\begin{aligned}
&\max_{\{k_t,x_t,b_t\}} E[\sum_{t=0}^{\infty}\beta^t x_t]\\
&s.t.\\
&\begin{cases}
\underbrace{q_t(k_t-k_{t-1})}_{\text{new capital}}+\underbrace{x_t}_{\text{consumption}} = \underbrace{(a+c)k_{t-1}}_{\text{capital production}}+\underbrace{b_t-Rb_{t-1}}_{\text{new borrowing}}\\
x_t \geq ck_{t-1}\\
Rb_t \leq q_{t+1}k_t
\end{cases}
\end{aligned}
$$

where the consumption x_t is at least the nontradable part because it is not storable. Furthermore, the collateral constraint is that the maximum amount of borrowing should be the value of assets. Then the gatherer's problem is

$$
\begin{aligned}
&\max_{\{k_t',x_t',b_t'\}} E[\sum_{t=0}^{\infty}\beta^t x_t']\\
&s.t.\\
&\underbrace{q_t(k_t'-k_{t-1}')}_{\text{new capital}}+\underbrace{x_t'}_{\text{consumption}} = \underbrace{G(k_{t-1})}_{\text{capital production}}+\underbrace{b_t'-Rb_{t-1}'}_{\text{new borrowing}}
\end{aligned}
$$

We notice that the gatherer has no constraint on borrowing. But due to technology disadvantage, they will not borrow as much. We will assume $\beta < \beta'$ so that farmers are more impatient and want to borrow and consume now.

Market clearing conditions include bond market clearing

$$b_t + mb_t' = 0,$$

capital market clearing

$$k_t + mk_t' = K$$

and consumption goods market clearing

$$(a+c)k_t + mG(k_t') = x_t + mx_t'$$

Next, we wan to know the interest rate. Because gatherer has no constraint on borrowing and gatherers are risk neutral, gatherer should pin down the interest rate as

$$R = \frac{1}{\beta'}$$

Moreover, the borrowing constraint of farmers will be binding because otherwise there is an arbitrage opportunity. Then we have

$$Rb_t = q_{t+1}k_t$$

Then the problem is simplified into

$$
\begin{aligned}
&\max_{\{k_t,x_t,b_t\}} E[\sum_{t=0}^{\infty}\beta^t x_t]\\
&s.t.\\
&\left(q_t - \tfrac{q_{t+1}}{R}\right)k_t = (a+c)k_{t-1} - x_t\\
&x_t \geq ck_{t-1}
\end{aligned}
$$

To further simplify the problem, we also want the constraint $x_t = ck_{t-1}$ binding. Intuitively, this means that the banker actually wants to consume as little as possible, when the borrowing constraint is binding, which implies that investment in capital is more lucrative than consuming more now.

Denote the Lagrangian multiplier on the budget constraint as λ_t and the Lagrangian multiplier on the $x_t \geq ck_{t-1}$ constraint as μ_t. Plugging in the binding borrowing constraint, the farmer objective function is

$$\sum_{t=0}^{\infty} \beta^t \left(x_t + \lambda_t \left(\left(q_t - \frac{q_{t+1}}{R} \right) k_t + x_t - (a+c)k_{t-1} \right) + \mu_t \left(x_t - ck_{t-1} \right) \right)$$

The FOC on capital is

$$\lambda_t \left(q_t - \frac{q_{t+1}}{R} \right) - \lambda_{t+1}\beta(a+c) - \mu_{t+1}\beta c = 0$$

and the FOC on consumption is

$$1 + \lambda_t + \mu_t = 0$$

In a stationary equilibrium, we should have $\lambda_t = \lambda$, $\mu_t = \mu$, $q_t = q$, which implies

$$\begin{aligned} &\lambda q \left(1 - \tfrac{1}{R}\right) - \lambda\beta(a+c) - \mu\beta c = 0 \\ &1 + \lambda + \mu = 0 \end{aligned}$$

Thus

$$\mu = \frac{\beta(a+c) - q\left(1 - \frac{1}{R}\right)}{q\left(1 - \frac{1}{R}\right) - \beta(a+c) + \beta c}$$

If the constraint is not binding, we have

$$x > ck$$

and

$$q(1 - \frac{1}{R})k = (a+c)k - x < ak$$

$$\Rightarrow q(1 - \frac{1}{R}) < a$$

However, this implies

$$\mu = \frac{\beta(a+c) - q\left(1 - \frac{1}{R}\right)}{q\left(1 - \frac{1}{R}\right) - \beta(a+c) + \beta c} > \frac{\beta(a+c) - a}{q\left(1 - \frac{1}{R}\right) - \beta(a+c) + \beta c}$$

When we have

$$\beta(a+c) - a > 0$$

The above implies $\mu > 0$, which contradicts to the assumption that the constraint is not binding. Thus we must have a binding constraint in the stationary equilibrium. When the solution is around the stationary equilibrium, we can use essentially the same arguments to arrive at a contradiction.

The gatherer has two degrees of freedom and thus effectively two FOCs:

$$R = \frac{1}{\beta'}$$

$$\frac{1}{R}G'(k_t{}') = q_t - \frac{q_{t+1}}{R}$$

where the second FOC is immediate from the combined two-period budget equation that eliminates time t borrowing

$$q_t(k_t{}' - k_{t-1}{}') + x_t{}' + \frac{1}{R}q_{t+1}(k_{t+1}{}' - k_t{}') + \frac{1}{R}x_{t+1}{}' = G(k_{t-1}) + \frac{1}{R}G(k_t) + \frac{1}{R}b_{t+1}{}' - Rb_{t-1}{}'$$

Since k_t is now the only intertemporal link, the marginal benefit should be the same as marginal cost, which implies the FOC on k_t' above.

Summary of Optimality Conditions

Now we can summarize individual optimality conditions as follows

$$\begin{cases} x_t = ck_{t-1} \\ b_t = \frac{1}{R}q_{t+1}k_t \\ R = \frac{1}{\beta'} \\ \frac{1}{R}G'(k_t{}') = q_t - \frac{q_{t+1}}{R} \end{cases}$$

with the farmer budget equation

$$\left(q_t - \frac{q_{t+1}}{R}\right)k_t = ak_{t-1}$$

We can omit gatherer's budget equation due to the perfect substitution between gatherer consumption and borrowing. With capital market clearing and gatherer FOC on k_t', we have

$$\frac{1}{R}G'(\frac{K-k_t}{m}) = q_t - \frac{q_{t+1}}{R}$$

To study the effect on current price, we can write

$$q_t = \sum_{s=t}^{\infty} \frac{1}{R^{s-t}}(q_t - \frac{q_{t+1}}{R}) = \sum_{s=t}^{\infty} \frac{1}{R}G'(\frac{K-k_s}{m})$$

Suppose we have a deviation from the steady state by increasing farmer capital share k_t. Then from the recovery path, we know that each individual term of $G'(\cdot)$ is higher, which implies current price of capital q_t is higher.

Steady State

First, the farmer budget equation implies

$$q(1-\frac{1}{R}) = a \qquad \Rightarrow \qquad q = \frac{a}{1-\beta'}$$

Furthermore, we have

$$\beta' G'(\frac{K-k}{m}) = a$$

which pins down the steady state farmer operated capital k. Then farmer consumption is

$$x = ck$$

and farmer borrowing is

$$b = \beta' qk = \frac{\beta'}{1-\beta'}ak$$

Dynamics

We can perturb around the steady state. We introduce an unexpected productivity shock to a at time 0 so that a temporarily becomes $a + a\varepsilon > a$ at time 0, before all decisions at time 0. To deal with this new timing, we have to use the original version of the farmer's budget constraint

$$q_0 k_0 = (a + a\varepsilon + c)k_{-1} + b_0 - Rb^* - x_0 + q_0 k^*$$

where $E_0[q_1]$ denotes the price without the unexpected shock. Then the consumption still satisfies

$$x_0 = ck_{-1}$$

and the new borrowing constraint is binding

$$b_0 = \frac{q_1}{R} k_0$$

Thus we have

$$\left(q_0 - \frac{q_1}{R}\right) k_0 = a(1+\varepsilon)k^* - Rb^* + q_0 k^*$$

We notice that the productivity starting from period 1 comes back. From gartherer FOC,

$$\frac{1}{R} G'\left(\frac{K - k_0}{m}\right) = q_0 - \frac{q_1}{R}$$

Denote user cost of capital as

$$u(k) = \frac{1}{R} G'\left(\frac{K - k}{m}\right)$$

Then we have

$$u(k_0)k_0 = (a(1+\varepsilon) - q^* + q_0)k^* \tag{19}$$

We also know that the price is

$$q_0 = \sum_{t=0}^{\infty} \frac{1}{R^t} u(k_t) \tag{20}$$

and the evolution of k_t is

$$k_t u(k_t) = ak_{t-1} \tag{21}$$

for $k \geq 1$. Around the steady state, we can log linearize the above three equations (19), (20), (21) and denote $u'(k^*) = \eta$, which is positive. With log linearization and denoting all log deviations with a hat, we have

$$\begin{cases} \hat{q}_0 = \frac{1}{\eta}\varepsilon \\ \hat{k}_0 = \frac{\eta}{\eta+1}\left(1 + \frac{R}{R-1}\frac{1}{\eta}\right)\varepsilon \end{cases}$$

We note that the effect on price is of same order of magnitude as shock, although the shock is temporary. In comparison, if the user cost of capital comes back from period 1 onwards, then we have

$$\begin{cases} \hat{q}_0 = \frac{1}{\eta}\frac{R-1}{R}\varepsilon \\ \hat{k}_0 = \varepsilon \end{cases}$$

When $R \to 1$, we note that $\hat{q}_0$ in the later case is close to zero. However, in the former case with persistence, the change in $\hat{q}_0$ is still positive, and the change in $\hat{q}_0$ becomes very large.

The effect on output is

$$Y_t = (a + c)k_t + mG(\frac{K - k}{m})$$

$$\hat{y}y^* = (a + c)k^*\hat{k} - G'(\frac{K - k^*}{m})k^*\hat{k}$$

With

$$G'(\frac{K - k^*}{m}) = aR$$

we get

$$\hat{y} = \frac{k^*}{y^*}(a + c - Ra)\,\hat{k}$$

Thus productivity increases with $\hat{k}$ around the steady state.

Others papers to be added: Bernanke and Gertler (1989), Kiyotaki and Moore (2012) etc.

12.2 Summary

In general, we have two types of frictions, equity issuance friction and collateral constraints. They provide similar implications, but are actually quite different, and microfounded in very different ways. One claim is that equity issuance friction is more fundamental. Without equity issuance friction, collateral constraint is never binding. However, without collateral constraint, equity issuance friction still matters.

In the classical model Kiyotaki and Moore (1997), farmers have both equity issuance friction and collateral constraints, and it seems the collateral constraints are essential to the existence of an equilibrium. This is because of risk neutrality and different discount rates. Without a collateral constraints, farmers are going to borrow an infinite amount. However, if both farmers and gartherers are risk averse, and capital is risky, without an collateral constraint, the model will still exhibit a balance sheet channel. When farmers have less net wealth, the risk bearing capacity goes down and capital is inefficiently held more by gatherers, which decreases productivity and asset prices. Since net worth recovers slowly, the impact is long-lasting. This feature is present in Bernanke and Gertler (1989), where without a collateral constraint, the balance sheet channel is still present.

It is important to consider the specific contracting environment and what underlying assumptions derive an equity issuance constraint and what assumptions derive both equity issuance and collateral constraint.

Below we consider the environment where collaterals are needed but equity is not feasible in the optimal contract.

- Inalienable human capital, or freedom of repudiation, which results in bargaining power of the borrowers (Hart and Moore, 1994). Thus lenders only want to lend to the extend that repayment can be secured, because otherwise they will lose. Assume we can collateralize on lands, then lenders only lend up to the collateral value of land.

- Lack of enforcement due to asymmetric information, such as the costly state verification framework in Townsend (1979). In this case, the optimal contract is debt with costly verification whenever the reported value is below the amount of debt. If we allow collateral, then lenders will always require collateral and only verify the noncollateralized part of the output. Thus this also gives rise of collateral.

In other environments, we have no debt constraint but an equity issuance friction.

- Assume that the banker can steal output above and beyond debt payment and resale it on the market to get a smaller fraction $\phi \in (0, 1)$ of it. Then the banker cannot raise more equity than $1 - \phi$ fraction of the total equity due to this constraint.

References

Ahn, S., Kaplan, G., Moll, B., Winberry, T., and Wolf, C. (2017). When inequality matters for macro and macro matters for inequality. In *NBER Macroeconomics Annual 2017, volume 32*. University of Chicago Press.

Allen, F., Carletti, E., and Marquez, R. (2011). Credit market competition and capital regulation. *The Review of Financial Studies*, 24(4):983–1018.

Ang, A. and Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary economics*, 50(4):745–787.

Barro, R. J. (1995). Inflation and economic growth.

Baumol, W. J. (1952). The transactions demand for cash: An inventory theoretic approach. *The Quarterly Journal of Economics*, pages 545–556.

Beck, T., Demirgüç-Kunt, A., and Levine, R. (2006). Bank concentration, competition, and crises: First results. *Journal of Banking & Finance*, 30(5):1581–1603.

Bernanke, B. S. and Gertler, M. (1989). Agency costs, net worth, and business fluctuations. *American Economic Review*, 79(1):14–31.

Boyd, J. H. and De Nicolo, G. (2005). The theory of bank risk taking and competition revisited. *The Journal of finance*, 60(3):1329–1343.

Brunnermeier, M. K. and Sannikov, Y. (2014). A macroeconomic model with a financial sector. *The American Economic Review*, 104(2):379–421.

Brunnermeier, M. K. and Sannikov, Y. (2016). The I theory of money. *National Bureau of Economic Research*.

DeMarzo, P. and Duffie, D. (1999). A liquidity-based model of security design. *Econometrica*, 67(1):65–99.

DeMarzo, P. M. and Sannikov, Y. (2006). Optimal security design and dynamic capital structure in a continuous-time agency model. *The Journal of Finance*, 61(6):2681–2724.

Diamond, D. W. and Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of political economy*, 91(3):401–419.

Drechsler, I., Savov, A., and Schnabl, P. (2017). The deposits channel of monetary policy. *The Quarterly Journal of Economics*, 132(4):1819–1876.

Duffie, D. (2010). *Dynamic asset pricing theory*. Princeton University Press.

Fischer, S. (1974). Money and the production function. *Economic Inquiry*, 12(4):517–533.

Gorton, G. and Pennacchi, G. (1990). Financial intermediaries and liquidity creation. *The Journal of Finance*, 45(1):49–71.

Hart, O. and Moore, J. (1994). A theory of debt based on the inalienability of human capital. *The Quarterly Journal of Economics*, 109(4):841–879.

Hellmann, T. F., Murdock, K. C., and Stiglitz, J. E. (2000). Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough? *American economic review*, pages 147–165.

Keeley, M. C. (1990). Deposit insurance, risk, and market power in banking. *The American Economic Review*, pages 1183–1200.

Kiyotaki, N. and Moore, J. (1997). Credit cycles. *Journal of political economy*, 105(2):211–248.

Kiyotaki, N. and Moore, J. (2012). Liquidity, business cycles, and monetary policy. *National Bureau of Economic Research*.

Laeven, L. and Levine, R. (2009). Bank governance, regulation and risk taking. *Journal of financial economics*, 93(2):259–275.

Ljungqvist, L. and Sargent, T. J. (2012). *Recursive macroeconomic theory*. MIT press.

Lucas, R. E. (1982). Interest rates and currency prices in a two-country world. *Journal of Monetary Economics*, 10(3):335–359.

Lucas, R. E. and Stokey, N. L. (1983). Optimal fiscal and monetary policy in an economy without capital. *Journal of monetary Economics*, 12(1):55–93.

Martinez-Miera, D. and Repullo, R. (2010). Does competition reduce the risk of bank failure? *The Review of Financial Studies*, 23(10):3638–3664.

Milgrom, P. and Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The review of economic studies*, 38(2):175–208.

Øksendal, B. (2003). Stochastic differential equations. In *Stochastic differential equations*, pages 65–84. Springer.

Piazzesi, M., Schneider, M., Benigno, P., and Campbell, J. Y. (2006). Equilibrium yield curves [with comments and discussion]. *NBER Macroeconomics Annual*, 21:389–472.

Protter, P. E. (2005). Stochastic differential equations. In *Stochastic Integration and Differential Equations*, pages 249–361. Springer.

Samuelson, P. A. (1958). An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of political economy*, 66(6):467–482.

Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *The Review of Economic Studies*, 75(3):957–984.

Sidrauski, M. (1967). Rational choice and patterns of growth in a monetary economy. *The American Economic Review*, 57(2):534–544.

Sundaram, R. K. (1996). *A first course in optimization theory*. Cambridge university press.

Svensson, L. E. (1985). Money and asset prices in a cash-in-advance economy. *Journal of Political Economy*, 93(5):919–944.

Tella, S. D. (2018). A neoclassical theory of liquidity traps.

Tobin, J. (1956). The interest-elasticity of transactions demand for cash. *The review of Economics and Statistics*, pages 241–247.

Tobin, J. (1965). Money and economic growth. *Econometrica: Journal of the Econometric Society*, pages 671–684.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic theory*, 21(2):265–293.

Townsend, R. M. (1980). Models of money with spatially separated agents. *Models of monetary economies*, pages 265–303.

Wallace, N. (1981). A modigliani-miller theorem for open-market operations. *The American Economic Review*, 71(3):267–274.

Walsh, C. E. (2017). *Monetary theory and policy*. MIT press.

Xiao, K. (2017). Monetary transmission in shadow banks.