Macroeconomía y Costos de Ajuste
Cátedras 3 y 4
6 de agosto de 2008

Eduardo Engel

Yale University

CEA - DII - U. de Chile
Agosto, 2008.

I. Basics

1. Course Overview
2. Evidence of Lumpiness and Non-Convex Adjustment Costs
3. Classic Aggregation
4. Quadratic Adjustment Costs
5. Calvo Model
6. Application: Labor Regulation and Adjustment
7. VARs and Lumpy Adjustment
8. *Ss* policies

7. VARs and Lumpy Adjustment

1. Motivation
2. Single variable
3. Slow Aggregate Convergence
4. Bias Correction
5. Extensions
6. Conclusion

Based on a forthcoming revision of Caballero and Engel (2003)

7.1. Motivation

- Macroeconomic variables: lumpy: prices, investment, employment, target/intended Federal Funds Rate
- Common strategy in applied macro:
  - Linear/VAR approximation
  - Interested in speed of response to shocks

Summary of Results

When you use VAR approximations to estimate the speed of adjustment underlying a macroeconomic variable with lumpy micro adjustment:

- You obtain upward biased estimate of the speed-of-adjustment/IRF
- You infer infinitely fast response to shocks for single variable
- The bias tends to zero as the number of units over which you aggregate tends to infinity, yet convergence is very slow

7.2. Single Variable

Simple Lumpy Adjustment Model: Calvo

- Shock: y^*
- Variable of interest: y
- ξ_t i.i.d. Bernoulli(λ)
- If $\xi_t = 1$ adjust at no cost, if $\xi_t = 0$ don't adjust
- $\mathrm{E}[\xi_t] = \lambda$.

Single Variable

- When adjust, choose y_t that solves:

$$\min_{y_t} \mathrm{E}_t[\sum_{k\geq 0} \beta^k (1-\lambda)^k (y_t - y^*_{t+k})^2]$$

- Assuming y^* random walk, leads to:

$$y_t \;=\; [1-\beta(1-\lambda)]\sum_{k\geq 0}[\beta(1-\lambda)]^k \mathrm{E}_t[y^*_{t+k}] \;=\; c + y^*_t$$

- Hence:

$$\boxed{\Delta y_t \;=\; \xi_t(y^*_t - y_{t-1}).} \qquad (1)$$

- And therefore:

$$\mathrm{E}[\Delta y_t \,|\, y^*_t,\, y_{t-1}] \;=\; \lambda(y^*_t - y_{t-1}).$$

Impulse Response Functions: Review

- The impulse response function of x w.r.t. v is defined as:

$$\mathrm{IRF}_k^{x,v} \equiv \mathrm{E}_t\left[\frac{\partial x_{t+k}}{\partial v_t}\right]; \qquad k = 0, 1, 2, \ldots$$

- For the IRF to be of interest in macroeconomics, we usually require that v be i.i.d. (innovation)
- Sometimes normalize the IRF, various alternatives:
  - response to shock of size σ_v instead of one
  - economic model (e.g., long run neutrality of money), suggests including a multiplicative constant so that $\sum_k \mathrm{IRF}_k = 1$
- Half-life of a shock: L s.t. $\sum_{k\leq L} \mathrm{IRF}_k = \frac{1}{2}\sum_{k\geq 0} \mathrm{IRF}_k$. Meaningful when $\mathrm{IRF}_k \geq 0$ for all k

IRF for Linear Relationships: Review

- In the case of a linear relationship between x and v:

$$x_t = \sum_{k \geq 0} a_k v_{t-k}$$

we have:

$$\mathrm{IRF}_k^{x,v} = a_k.$$

and no averaging is needed

IRF for an AR(1) Process: Review

- Assume that:

$$x_t = \phi x_{t-1} + v_t$$

  with $|\phi| < 1$.

- Applying the above relationship again and again, and assuming v has bounded second moments, so that $\phi^k x_{t-k}$ converges to zero as k tends to infinity, we obtain:

$$x_t = \sum_{k \geq 0} \phi^k v_{t-k}$$

- It then follows from the previous slide that:

$$\mathrm{IRF}_k^{x,v} = \phi^k.$$

- If we want the IRF to add up to one, we have:

$$\mathrm{IRF}_k^{x,v} = (1 - \phi)\phi^k.$$

IRF for an AR Process: Review

- We have:

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + v_t$$

  where v is an innovation (i.i.d. orthogonal to past values of x) and, to ensure stationarity, the roots of the polynomial $1 - \phi_1 z - \ldots - \phi_p z^p$ are all outside the unit circle
- To calculate the IRF we could use the moving average representation (as in the previous slide)
- Or we could use that the IRF satisfies a difference equation analogous to the difference equation that defines x:

$$\text{IRF}_k = \phi_1 \text{IRF}_{k-1} + \ldots + \phi_p \text{IRF}_{k-p}.$$

Vector Autoregression (VAR): Review

- A VAR representation of a stationary stochastic process consists in approximating the process by an AR(p) process
- Even though below we focus on the case where the process being approximated is one-dimensional, VAR are used mostly in macroeconomics in the multi-dimensional case

True Impulse Response Function

- Back to the non-linear Calvo model
- We have:

$$\begin{aligned}
\mathrm{IRF}_k^{\Delta y, \Delta y^*} &= \mathrm{E}_t \left[\frac{\partial \Delta y_{t+k}}{\partial \Delta y_t^*} \right] \\
&= \mathrm{Prob}[\xi_t = \xi_{t+1} = \ldots = \xi_{t+k-1} = 0, \xi_{t+k} = 1] \\
&= \lambda(1-\lambda)^k.
\end{aligned}$$

- Hence:
  - same IRF as for an AR(1) with $\rho = 1 - \lambda$
  - same IRF as corresponding quadratic adjustment model

Proposition 1: Ultra-Fast Estimate

Assumptions:

- ξ_t: see above
- Δy^* i.i.d.$(0, \sigma^2)$

Then:

- Δy i.i.d.
- VAR representation of Δy (i.e., approximating Δy by an AR process) is white noise (i.i.d.)
- Infer IRF corresponding to $\lambda = 1$, independent of true value of λ
- Inferred adjustment: infinitely fast

Intuition

- We have:

$$\Delta y_t = \begin{cases} \sum_{k=0}^{l_t-1} \Delta y^*_{t-k} & \text{if adjust in } t, \\ 0 & \text{otherwise.} \end{cases}$$

  where l_t denotes the number of periods since the last adjustemtn, as of time t

- Why $\mathrm{Cov}(\Delta y_t\ ,\ \Delta y_{t-1}) = 0$ indep. of the true λ?
  - if do not adjust in $t-1$ or t: $\Delta y_t \times \Delta y_{t-1} = 0$.
  - if adjust in $t-1$ and t:

$$\mathrm{Cov}(\Delta y_t, \Delta y_{t-1}) = \mathrm{Cov}(\Delta y^*_t, \Delta y^*_{t-1} + \Delta y^*_{t-2} + \ldots) = 0.$$

Wold's Decompositin and VARs

- Every stationary (non-deterministic) process X has a (unique) MA(∞) representation:

$$X_t = \sum_{j \geq 0} \psi_j \varepsilon_{t-j},$$

  with:
  - $\sum \psi_j^2 < \infty$, $\psi_0 = 1$
  - ε_t uncorrelated (i.e., white noise)
  - $\forall t$ the linear space generated by $\varepsilon_s, s \leq t$ equal to that generated by $X_s, s \leq t$
- IRF w.r.t. ε-shocks: $\psi_0, \psi_1, \ldots, \psi_k, \ldots$
- The multivariate version of this result is the basis for VARs

Wold's Decomposition and Lumpy Adjustment

- For a single agent with lumpy adjustment, the Wold decomposition Δy is a white noise process
- Hence: $\psi_k = 0, k \geq 1$ and $\text{IRF}_k = 0, k \geq 1$
- Yet we saw that: $\text{IRF}_k \propto (1 - \lambda)^k$
- What happened to Wold's Decomposition and to VARs?
  - Wold provides the IRF to the wrong shock
  - Wold only considers first and second moments, with lumpy adjustment higher moments also play a crucial role
  - Relation of Δy between economically interesting shocks (which include the ξ) is highly non-linear

7.3. Slow Aggregate Convergence

Rotemberg's Equivalence Result:

- Simple lumpy adjustment model with $\xi_{i,t}$'s i.i.d.:
- Aggregate over $N = \infty$ agents
- Result: Δy is AR(1) with $\rho = 1 - \lambda$
- Infer the correct theoretical IRF from aggregate ($N = \infty$)
- How fast does convergence take place?

Technical Assumptions

- $\Delta y^*_{i,t} \equiv v^A_t + v^I_{i,t}$ with:
  - v^A_t's: i.i.d. (μ_A, σ^2_A),
  - $v^I_{i,t}$'s: i.i.d. $(0, \sigma^2_I)$,
  - $\xi_{i,t}$'s: i.i.d. Bernoulli(λ), $\lambda \in (0, 1]$.

Notation

- Define:

$$\begin{aligned}\Delta y_t &\equiv \sum_i w_i \Delta y_{i,t}, \\ \text{Effective } N &\equiv 1 / \sum_i w_i^2.\end{aligned}$$

- If the true micro-model is Calvo, does estimating an AR(1)

$$\Delta y_t = \rho \Delta y_{t-1} + \epsilon_t, \tag{2}$$

yield a consistent $\rho = 1 - \lambda$?

Proposition 2: Aggregate Bias

- $\hat{\rho}$: OLS estimator of ρ in (2).
- Technical assumptions hold.
- Then:

$$\boxed{\mathrm{plim}_{T\to\infty}\hat{\rho} \;=\; \frac{K}{1+K}(1-\lambda),} \tag{3}$$

with

$$K \;\equiv\; \frac{\frac{\lambda}{2-\lambda}(N-1)\sigma_A^2 - \mu_A^2}{\sigma_A^2 + \sigma_I^2 + \frac{2-\lambda}{\lambda}\mu_A^2}.$$

Intuition

- For simplicity: $\mu_A = 0$.

$$\begin{aligned}
\rho_1 &= \frac{\mathrm{Cov}(\Delta y_t^N, \Delta y_{t-1}^N)}{\mathrm{Var}(\Delta y_t^N)} \\
&= \frac{N\mathrm{Cov}(\Delta y_{1,t}, \Delta y_{1,t-1}) + N(N-1)\mathrm{Cov}(\Delta y_{1,t}, \Delta y_{2,t-1})}{N\mathrm{Var}(\Delta y_{1,t}) + N(N-1)\mathrm{Cov}(\Delta y_{1,t}, \Delta y_{2,t})} \\
&= \frac{0 + N(N-1)\frac{\lambda}{2-\lambda}(1-\lambda)\sigma_A^2}{N(\sigma_A^2+\sigma_I^2) + N(N-1)\frac{\lambda}{2-\lambda}\sigma_A^2}.
\end{aligned}$$

- Hence:

	$\mathrm{Cov}(\Delta y_{1,t}, \Delta y_{1,t-1})$	$\mathrm{Var}(\Delta y_{1,t})$
Micro AR(1):	$\frac{\lambda}{2-\lambda}(1-\lambda)(\sigma_A^2+\sigma_I^2)$	$\frac{\lambda}{2-\lambda}(\sigma_A^2+\sigma_I^2)$
Calvo ($\mu_A = 0$):	0	$\sigma_A^2+\sigma_I^2$

- Bias vanishes as N goes to infinity.
- Convergence slower if:

$$\sigma_I \uparrow, \quad \sigma_A \downarrow, \quad \lambda \downarrow, \quad |\mu_A| \uparrow .$$

Applications

- Employment, prices, investment.
- Parameters:
    - Caballero, Engel, Haltiwanger (1997)
    - Bils and Klenow (2004)

Table: Slow Convergence

Reported: Estimated Half-Life of Shock

	Effective number of agents (N)					
	100	500	1,000	5,000	10,000	∞
Employment (quart.)	0.373	0.723	0.912	1.225	1.287	1.357
Prices (mon.)	0.287	0.595	0.861	1.824	2.225	2.901
Investment (annual)	0.179	0.399	0.582	1.516	2.167	4.596

What is the relevant value of N?

LRD-establishments (1972-88, cont. panel):

- Employment: 2,611
- Investment: 699

US-Non-Farm-Business (2001):

- Employment: 3,683
- Investment: 986

What is the relevant value of N? (Cont.)

Summing up:

- Aggregate data:
  - Employment: maybe significant bias
  - Prices: maybe significant bias
  - Investment: large bias for sure
- Sectoral data: major bias
- Back to tables

7.4. Bias Corrections

- ARMA Correction
- Instrumental Variable Correction

Proposition 3: ARMA Correction

- Same assumptions as in Proposition 2.
- Add an MA(1) term to the standard partial adjustment equation:

$$\Delta y_t^N = (1-\lambda)\Delta y_{t-1}^N + v_t - \theta v_{t-1}.$$

- $\hat{\lambda}^N$: any consistent estimator of the AR-coefficient.

Proposition 3: ARMA Correction

Then:

- Δy_t^N follows an ARMA(1,1) with AR parameter $1-\lambda$.
- $\text{plim}_{T\to\infty}\hat{\lambda}^N = \lambda$.
- MA coefficient, θ: "nuisance" parameter that depends on N, μ_A, σ_A and σ_I.
- $\lim_{N\to\infty} \theta(N, \lambda, \mu_A, \sigma_A, \sigma_I) = 0$.

Implications of Proposition 3

- If estimate an AR(p) with large p you get the wrong IRF
- Must estimate an ARMA(1,1) and drop MA term before calculating the IRF
- Partial solution when N not too small (coincidental reduction) and T not too small.

Proposition 4: IV Correction

- Consider the standard linear equation:

$$\Delta y_t = (1-\lambda)\Delta y_{t-1} + \lambda \Delta y_t^*,$$

- The following are valid instruments:
  - Components of Δy_{t-1}^*, Δy_{t-2}^*, ...
  - $\Delta y_{t-2}, \Delta y_{t-3}, \ldots$
- Not a valid instrument: Δy_{t-1}

Sketch of Proof

From (1)

$$\Delta y_{it} = \xi_{it}(y^*_{it} - y_{i,t-1}) = \lambda(y^*_{it} - y_{i,t-1}) + (\xi_{it} - \lambda)(y^*_{it} - y_{i,t-1}),$$

and therefore

$$\Delta y_{i,t-1} = \lambda(y^*_{i,t-1} - y_{i,t-2}) + (\xi_{i,t-1} - \lambda)(y^*_{i,t-1} - y_{i,t-2}).$$

Subtracting the latter from the former and rearranging:

$$\Delta y_{it} = (1 - \lambda)\Delta y_{i,t-1} + \lambda \Delta y^*_{it} + \varepsilon_t,$$

with

$$\varepsilon_{it} = (\xi_{it} - \lambda)(y^*_{it} - y_{i,t-1}) - (\xi_{i,t-1} - \lambda)(y^*_{i,t-1} - y_{i,t-2}).$$

Sketch of Proof

- A valid instrument needs to be orthogonal to ε_t.
- Δy_t does not satisfy this condition because it is correlated with the second term in the above expression.
- By contrast, Δy^*_{t-k}, $k = 1, 2, 3, \ldots$ is a valid instrument. To see this note that these variables are independent from Δy^*_t, ξ_{it} and $\xi_{i,t-1}$.
- $\Delta y_{t-2}, \Delta y_{t-3}, \ldots$ are also valid instruments

Using IV Correction

- U.S. Private Sector
- Panel with 10 one-digit sectors
- 1987–2004
- Quarterly data
- Instruments: unit labor costs, ...

Table: IV Correction for Employment

Reported: Estimated Half-Life of Shock

AR(1)	With IV
2.40	4.95
(0.25)	(1.22)

IV Correction and Bils-Klenow Finding

- Calvo model predicted:

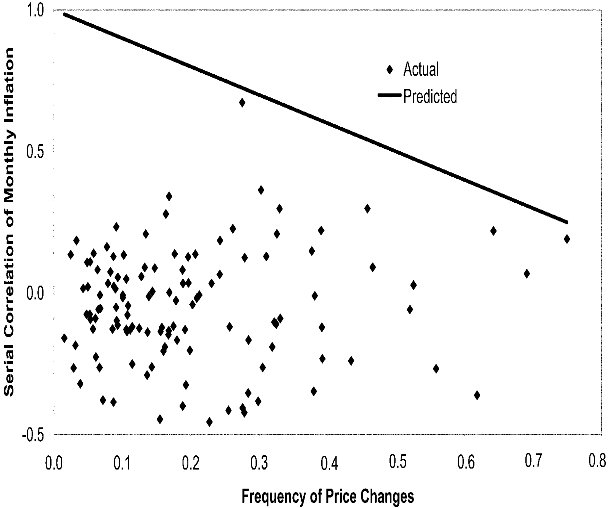
$$\hat{\rho}_1 \cong 1 - \hat{\lambda}$$

  where ρ_1 is obtained regressing π_t on π_{t-1} and λ from micro adjustment frequency

- Using 123 ELIs, Bils-Klenow obtained:

$$\hat{\rho}_1 \ll 1 - \hat{\lambda}.$$

Testing the Calvo Model



IV Correction and Bils-Klenow Finding

- Results arrived today!
- Instruments: prices from other goods (brothers and sisters, uncles and aunts), energy price index, producer price index
- For every ELI we estimate the adjustment speed λ via three alternatives:
  - Micro frequency of price adjustments: $\hat{\lambda}_{\text{micro}}$
  - OLS estimate for ρ_1: $\hat{\lambda}_{\text{OLS}} = 1 - \hat{\rho}_{1,\text{OLS}}$
  - IV estimate for ρ_1: $\hat{\lambda}_{\text{IV}} = 1 - \hat{\rho}_{1,\text{IV}}$

IV Correction and Bils-Klenow Finding

- We then estimate the regression:

$$\hat{\lambda}_{i,\mathrm{IV}} - \hat{\lambda}_{i,\mathrm{OLS}} = \beta(\hat{\lambda}_{i,\mathrm{micro}} - \hat{\lambda}_{i,\mathrm{OLS}})$$

- It follows that

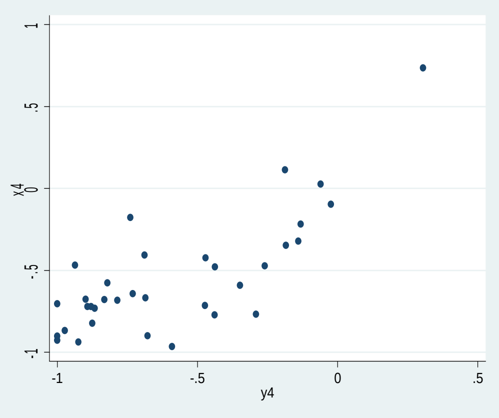
$$\hat{\lambda}_{i,\mathrm{IV}} = \beta\hat{\lambda}_{i,\mathrm{micro}} + (1-\beta)\hat{\lambda}_{i,\mathrm{OLS}}$$

- Hence β captures the fraction of the gap that is closed on average by IV between OLS and the “true” micro estimates

IV Correction and Bils-Klenow Finding

- We obtain $\hat{\beta}$ close to zero if we work with all ELIs
- We obtain $\hat{\beta}$ close to 1 if we work with the subset of ELIs (34/180) with *t*-stats for $\hat{\lambda}_{\text{IV}} \geq 1.65$
- Next slide: plots $(\hat{\lambda}_{i,\text{IV}} - \hat{\lambda}_{i,\text{OLS}})$ against $(\hat{\lambda}_{i,\text{micro}} - \hat{\lambda}_{i,\text{OLS}})$ for the subset of ELIs where the IV-estimates are precise

IV Correction and Bils-Klenow Finding



7.5. Extensions

- Relaxing random walk assumption:
  - y^*: AR(1) with autocorrel. ϕ
  - Bias is zero if $\phi = 0$
  - Bias similar to r.walk case if $\phi > 0.80$
- Adding smooth adjustment:
  - Lumpy and smooth micro adjustment:

$$y_t^* \quad \overset{Calvo}{\longrightarrow} \quad \tilde{y}_t \quad \overset{AR(p)}{\longrightarrow} \quad y_t.$$

  - Bias for lumpy component

7.5. Extensions

- Relaxing Poisson/Calvo assumption:
  - Adjustment-shocks follow any arrival process
  - Generates rich class of potential theoretical IRF
  - Rotemberg result generalizes to this case
  - Similar bias
- Aggregate hazard shocks:
  - λ time-varying
  - Bias larger

7.5. Extensions

- Increasing hazard (generalized *Ss*) models:
  - Same result (and intuition) for $N = 1$
  - Bias larger (similar to case with aggregate hazard shocks)
- Time-aggregation:
  - Smoothing via time-aggregation
  - Same bias for $N = 1$ and slow convergence
  - Application: interest rates

8. *Ss* Policies

1. Optimality of *Ss* policies: a one-sided example
2. Generalized *Ss* policies
3. Brownian motion: a brief review
4. *Ss* policies in continuous time
5. Invariant and ergodic distributions

8.1. Optimality of *Ss* policies: an example

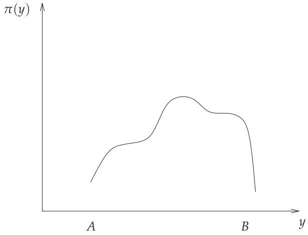
- Fixed adjustment costs are the basis to rationalize lumpy behavior.
- Proving that micro-optimal policies are of the *Ss* type is very difficult, even if you consider an individual agent in partial equilibrium. Proving existence and uniqueness is usually straightforward (thanks to Blackwell's conditions). Yet proving that the optimal policy is of the *Ss* type can be very hard.
- In this section we give the flavor of such a proof, by looking at a simple (one-sided policy) case.
- Later we will use (but won't prove) more general *Ss* policies.

A Particular Case

Problem Formulation

- A monopolist sets the nominal price of a good in an inflationary environment.
- Notation:
  - β: Monopolist's discount factor
  - x_t: Log of real price at t, before deciding whether to adjust.
  - y_t: Log of the real price chosen in t.
  - K: Cost of adjusting the price (assumed constant).
- y takes values in $[A, B]$: if the real price is below A it is optimal for the monopolist not to produce, if it is above B there is no demand.

Optimality of *Ss* policies: an example



$\pi : [A, B] \rightarrow \mathbb{R}$ denotes the monopolist's profit as a function of the log-price. This function is assumed quasiconcave (see the figure above). It attains its maximum at $\hat{y}$ (quasiconcavity ensures that $\hat{y}$ is the only local maximum).

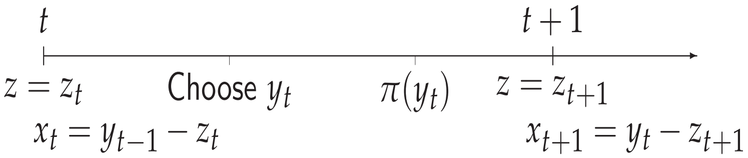
Optimality of *Ss* policies: an example

- Inflationary shocks in period t, before the monopolist decides whether to adjust, are denoted by z_t (in the notation of previous lectures, they are the $\Delta\widehat{y}$) and assumed i.i.d. with common density $q(z)$.
- We also assume $z_t \geq 0$. This non-negativity assumption rules out the possibility of ever wanting to decrease the nominal price.
- The crucial assumptions: fixed adjustment cost and non-negative inflationary shocks.

Optimality of Ss policies: an example

We adopt the following timing convention (depicted below):

- The monopolist begins period t with real log-price y_{t-1}.
- The inflationary shock, z_t, takes place, lowering the real log-price to x_t.
- The monopolist decides whether to adjust or not, resulting in a nominal log-price y_t.
- Period t profits are realized.



Optimality of *Ss* policies: an example

- Denote: $v(x) \equiv$ Present value of profits, net of adjustment costs, when the current real price is x, discount factor: β.
- The Bellman equation for this problem the is:

$$v(x) = \max \Big\{ \underbrace{\pi(x) + \beta \int v(x-z)q(z)\,dz}_{\text{no adjustment}}\,,$$

$$\underbrace{\sup_y \left[\pi(y) - K + \beta \int v(y-z)q(z)\,dz\right]}_{\text{adjust}} \Big\}. \tag{4}$$

Existence and Uniqueness of a Solution

- Define:

$$(Tv)(x) = \max\{\pi(x) + \beta \int v(x-z)q(z)\,dz,$$

$$\sup_{y\in[A,B]}[\pi(y) - K + \beta \int v(y-z)q(z)\,dz]\}.$$

- If $|v(x)| \leq C$, then $|(Tv)(x)| \leq \max_u |\pi(u)| + \beta C$ and hence T preserves boundedness (i.e., $T : \mathcal{B}([A,B]) \to \mathcal{B}([A,B])$) and we can apply Blackwell's Theorem.
- Hence must prove that Monotonicity and Discounting conditions hold.

Verifying Blackwell's Conditions

Monotonicity

We have that:

$$f \leq g \;\Rightarrow\; \int f(x-z)q(z)\,dz \leq \int g(x-z)q(z)\,dz$$

$$\Rightarrow\; \pi(x) + \beta \int f(x-z)q(z)dz \leq \pi(x) + \beta \int g(x-z)q(z)dz \qquad (5)$$

Optimality of *Ss* policies: an example

We also have that:

$$f \le g \Rightarrow \pi(y) - k + \beta \int f(y-z)q(z)\,dz \le$$

$$\pi(y) - k + \beta \int g(y-z)q(z)\,dz$$

$$\Rightarrow \sup_{y\in[A,B]} \left[\pi(y) - K + \beta \int f(y-z)q(z)\,dz\right] \le$$

$$\sup_{y\in[A,B]} \left[\pi(y) - K + \beta \int g(y-z)q(z)\,dz\right] \tag{6}$$

Combining (5) and (6) yields:

$$f \le g \Rightarrow Tf \le Tg,$$

thus establishing monotonicity.

Optimality of *Ss* policies: an example

Discounting

We have that:

$$
\begin{aligned}
(T[f+a])(x) = \max\Big\{&\pi(x) + \beta \int [f+a](x-z)q(z)\,dz, \\
&\sup_{y\in[A,B]} \Big[\pi(y) - K + \beta \int [f+a](y-z)q(z)\,dz\Big]\Big\} \\
= \max\Big\{&\pi(x) + \beta a + \beta \int f(x-z)q(z)\,dz, \\
&\sup_{y\in[A,B]} \Big[\pi(y) - K + \beta a + \beta \int f(y-z)q(z)\,dz\Big]\Big\} \\
= \beta a &+ (Tf)(x),
\end{aligned}
$$

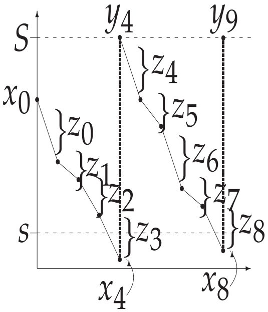
The existence and uniqueness of a solution to the Bellman equation now follows from Blackwell's Theorem.

Characterizing the Optimal Policy

- The proof of existence and uniqueness was trivial, the challenge is to prove that, for some constants s and S, the optimal policy is of the form (see the figure on the next slide):

$$\begin{aligned} x_t &\leq s \Rightarrow \text{adjust } y_t \text{ to } S, \\ x_t &> s \Rightarrow y_t = x_t \quad \text{that is: no adjustment.} \end{aligned}$$

Characterizing the Optimal Policy



The Auxiliary Value Function

Denote by $(Jv)(x) \equiv \pi(x) + \beta \int v(x-z)q(z)dz$ the present value of profits, net of adjustment costs, given a current real log-price of x, assuming you do *not* adjust your price in the current period.

The Bellman equation may then be written as:

$$v(x) = \max \left\{ (Jv)(x), \sup_{y \in [A,B]} [(Jv)(y) - K] \right\}. \tag{7}$$

Next we show that if v is continuous and bounded function, then so is the function (Jv).

The Auxiliary Value Function

- If $x_n \longrightarrow x$, then:

$$|(Jv)(x_n) - (Jv)(x)| \leq$$

$$|\pi(x_n) - \pi(x)| + \beta \int |v(x_n - z) - v(x - z)|\, q(z)\, dz$$

- Next we let $n \longrightarrow \infty$. We can exchange exchange $\int$ and lim because of Lebesgue's Dominated Convergence Theorem (here we use that v is bounded).

The Auxiliary Value Function

- We obtain:

$$\lim_{n\to\infty} |(Jv)(x_n) - (Jv)(x)| \leq \lim_{n\to\infty} |\pi(x_n) - \pi(x)|$$

$$+\beta \int_A^B \lim_{n\to\infty} |v(x_n - z) - v(x - z)|\, q(z)\, dz$$

- And since v and π are continuous, we conclude that:

$$(Jv)(x_n) \to (Jv)(x)$$

  and therefore (Jv) is continuous.

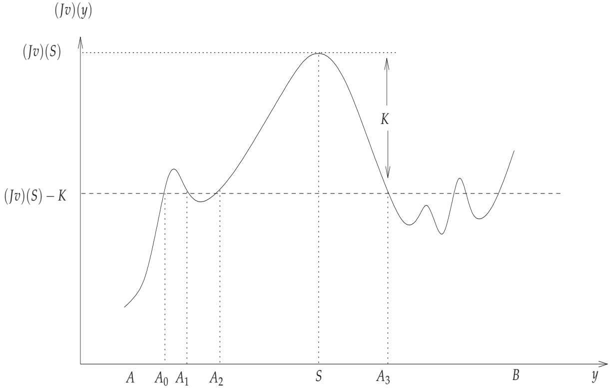
The Auxiliary Value Function

- Since (Jv) is continuous and $[A, B]$ compact, $\sup_y (Jv)(y)$ is attained at some point in $[A, B]$, denote this point by S.
- Then (7) can be written as:

$$v(x) = \max \{(Jv)(x),\ (Jv)(S) - K\} .$$

- In what follows we may focus on values of y in $[A, S]$, since, assuming initially $y \leq S$, we have that the value y (and, due to the monotonicity of shocks, also x) will never exceed S.
- Next comes the non-trivial part. If Jv looks as in the figure on the next page, the optimal rule is not of the Ss type.

The Auxiliary Value Function



The Auxiliary Value Function

- Indeed, in the case depicted in the preceding figure, we have that the optimal policy is:
  - $x \in [A, A_0] \cup [A_1, A_2] \Rightarrow$ adjust.
  - $x \in [A_0, A_1] \cup [A_2, S] \Rightarrow$ do *not* adjust.
  - $x > S \Rightarrow$ irrelevant.
- It follows that to prove that the optimal policy is *Ss*, we must find sufficient conditions to rule out situations like the one depicted in the preceding figure, namely conditions that ensure that $(Jv)(y)$ crosses the horizontal line at $y = (Jv)(S) - K$ only once for $y \in [A, S]$.

The Auxiliary Value Function

- To do this we show that, denoting by $\hat{y}$ the largest value attained by π:
  - (a) (Jv) is strictly increasing in $[A, \hat{y}]$.
  - (b) (Jv) grows strictly less than K in any interval contained in $[\hat{y}, S]$.
- In what follows we use the assumption that π is quasiconcave in $[A, B]$, that is, that it has a unique local (and therefore global) maximum, $\hat{y}$.

Proof of (a)

- We first show that:
  - w strictly increasing in $[A, \hat{y}] \Rightarrow Tw$ strictly increasing in $[A, \hat{y}]$.
- Assume $w(x) < w(x'), \forall x, x'$ with $A \leq x < x' \leq \hat{y}$.
- Then:

$$
\begin{array}{ccc}
\pi(x) & \overset{\pi \text{ quasiconcave}}{<} & \pi(x') \\
\int w(x-z)q(z)\,dz & < & \int w(x'-z)q(z)\,dz,
\end{array}
$$

and it follows that $(Tw)(x) < (Tw)(x')$ and Tw is is strictly increasing in $[A, \hat{y}]$.

Proof of (a)

- Next we apply the "Garbage In–Garbage Out Result", noting that the set of non-decreasing (i.e., weakly increasing) functions defined on $[A, \hat{y}]$ is closed and includes those which are strictly increasing in that interval.
- It then follows that v is strictly increasing in $[A, \hat{y}]$

Proof of (b)

- First note that, since adjusting real prices from x to x' costs K, we must have, for all $x, x' \in [A, S]$:

$$|v(x) - v(x')| \leq K. \tag{8}$$

- Hence, for $y, y' \in [\hat{y}, S)$, with $(Jv)(y') > (Jv)(y)$:

$$\begin{aligned}
0 &\leq (Jv)(y') - (Jv)(y) \\
&\leq (Jv)(S) - (Jv)(y) \\
&= \pi(S) - \pi(y) + \beta \int [v(S-z) - v(y-z)]\, q(z)\, dz \\
&\leq \beta \int |v(S-z) - v(y-z)|\, q(z)\, dz \\
&\leq \beta K \\
&< K,
\end{aligned}$$

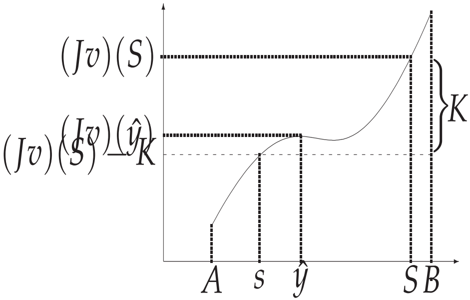
Proof of (b)

where:

- Second step: (Jv) attains its maximum at S.
- Fourth step: since y and y' are to the right of the point where π attains its maximum and π is quasiconcave, we have $\pi(y) > \pi(S)$.
- Fifth step: from (8).

This concludes our proof. We have shown that (Jv) is of the form depicted in the figure on the following page. In doing so we have also shown that $s < \hat{y} < S$.

Proof of (b)



Comments on the Optimal Policy

- At the moment of adjusting its nominal price, the firm's real price is above its frictionless optimal price, $\hat{y}$: since the firm is forward looking, it chooses a high real price when adjusting, aware of the fact that this price will decrease until it adjusts its price again
- By the end of a pricing cycle, the firm's real price is below its frictionless optimal price $\hat{y}$
- The length of price cycles varies from one cycle to the next, depending on realizations of z_t
- Even if the cost of deviating from the static target price is symmetric, we will not have $S = -s$, in fact, in this case $S < -s$. Why?

General Comments

- The first proof of optimality of *Ss* rules is due to Yale's Herb Scarf, in 1959. This result was a major breakthrough. Many bright mathematical/statistical/operation research minds worked on the problem during the 1950s. For the history behind the proof, see Scarf (2002, CFP 1036).
- To prove optimality of *Ss* rules, Scarf introduced the concept of $K-$convexity of a function, which is closely related to the final part (part (b)) of the proof.
- The original setting for this problem was inventory theory, the first applications to pricing were in the late 1970's, while other macro variables (durable consumption, employment, investment) had to wait until the late 1980s/early 1990s.
- The proof presented in this lecture is due to Andrew Caplin.

8.2. Generalized *Ss* Policies

Ss policies are characterized by:

- One or more state-variables: variables that determines whether you adjust or not. In the example above there is one state-variable, x, which is the real log-price
- An inaction range, $\mathcal{I}$. As long as $x \in \mathcal{I}$ you do not adjust. In the example above: $\mathcal{I} = (s, S]$.
- One or more adjustment triggers: values of x that trigger an adjustment. In the example above, any value of $x \leq s$ is a trigger.
- One or more adjustment targets: values to which when you reach an adjustment trigger. In the example above, S is the only adjustment target

State– vs. Time–dependent Rules

Ss policies are called state-dependent policies, since whether you act or not depends on your state-variable (in the above example, the real price).

By contrast, policies considered in earlier papers within the Neo-Keynesian tradition (Fischer and Taylor) are time-dependent, since the moments where you act are separated by a fixed interval of time.

Calvo-type adjustments are clearly not state-dependent. Are they time-dependent? If we take a broader view of what "time-dependent" means, they are. For this we define a time-dependent rule as a policy such that whether you adjust or not in a given period depends only on the time that has elapsed since the last time you adjusted, not on your (or the economy's) particular state-variables

Generalized *Ss* Policies

- We relax the assumption that the z_t are nonnegative
- In some cases we also relax the fixed-cost-of-adjustment assumption
- Assuming that the static real log-price follows a random walk simplifies things further, much of what follows can be extended to the case where this price follows an AR(1)
- Formal proofs for many of the “approximate results” that follow require us to work in continuous time. We briefly discuss later in this lecture.

(L, C, U) Policies

- We relax the $z_t \geq 0$ assumption in our benchmark model, that is, we allow for negative inflationary shocks.
- By contrast with the benchmark model, now there are times when it may be optimal to decrease the nominal price.
- The z_t shocks continue being i.i.d.
- The cost of adjusting the nominal price may be the same for increases and decreases, or may differ. What matters is that we have a fixed cost in both cases.

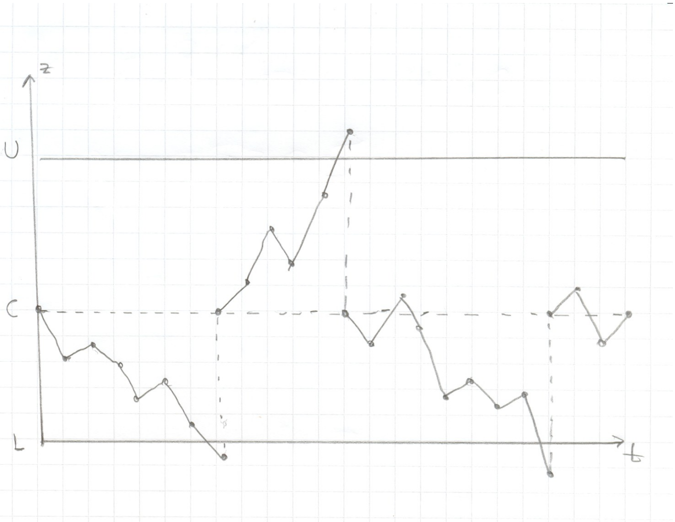
(L, C, U) Policies

Result

The the optimal policy is two-sided *Ss*, characterized by three thresholds: $L < C < U$:

- when the real price reaches the lower trigger L, the nominal price is increased by $C - L$ and a new price cycle begins, with the real price at C
- when the real price reaches the upper trigger U, the nominal price is decreased by $U - C$ and a new price cycle begins, with the real price at C

(L, C, U) Policies



(L, C, U) Policies

- Inaction range: $L < x < U$
- Target real price: C
- Triggers: L and U
- *Ss* policies where the control variable can both increase and decrease are referred to as two-sided *Ss* policies, by contrast with the case considered in our benchmark model, which is a one-sided *Ss* policy

(d, D, U, u) Policies

As in the (L, C, U) case, but now the adjustment cost has a variable component (as well as the fixed component we considered earlier).

That is, the firm pays $K^+ + k^+ \Delta x$ for a nominal price increase of Δx and pays $K^- - k^- \Delta x$ for a nominal price decrease of Δx. The constants K^+, K^-, k^+, k^- are all strictly positive.

(d, D, U, u) Policies

Result

The the optimal policy is (approximately) of the (d, D, U, u) type:

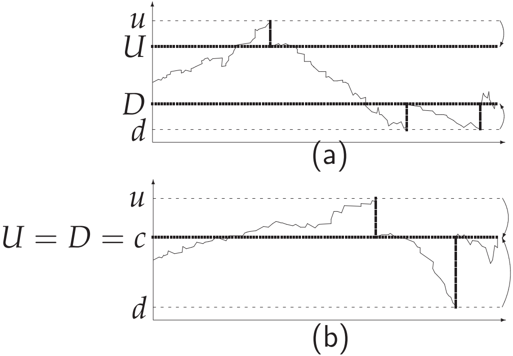
- When x reaches d, the nominal price is increased so that the real price reaches D
- When x reaches u, the nominal price is decreased so that the real price reaches U
- We have: $d < D < U < u$.

(d, D, U, u) Policies

- Triggers: d and u
- Targets: D and U
- Inaction range: $d < x < u$
- Approximation because of discrete time: the more x falls below d, the smaller the real price you choose after adjusting. This is not an issue in continuous time.

The first panel on the following slide depicts the generic (d, D, U, u) policy, the second panel the particular case of a (L, C, U) policy.

Shape of the Optimal Policy



Adjustment Hazard Model

- Same assumptions as in (L, C, U) case, but now the fixed adjustment cost in a given period are i.i.d. draws from a known distribution $G(\omega)$, with probability density $g(\omega)$
- The optimal policy now depends not only on the firm's real price immediately before adjusting, x, but also on the current adjustment cost draw, ω

Adjustment Hazard Model

Result

Conditional on ω, the optimal policy is of the (L, C, U) type, with $L = L(\omega)$ and $U = U(\omega)$. C does not depend on ω. We also have: $L'(\omega) < 0$, $U'(\omega) > 0$ and $L(\omega) < C < U(\omega)$.

Adjustment Hazard Model

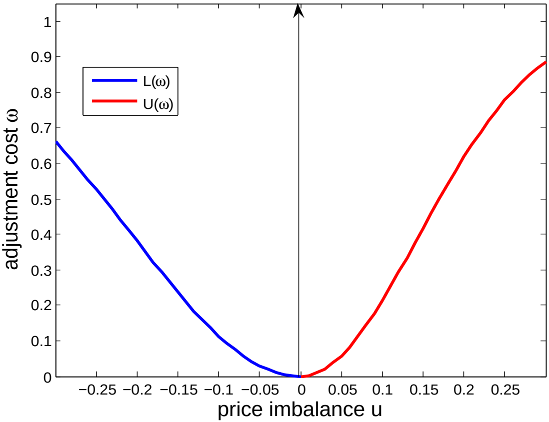
To further describe the optimal policy, it is better to work with the real price in deviation from C:

$$u \equiv x - C.$$

We then denote by $\Omega(u)$ the threshold in the realization of ω, for which the firm is indifferent between adjusting and not adjusting, conditional on a real price deviation of u. That is:

$$\Omega(u) = \begin{cases} L^{-1}(u), & \text{if } u < 0, \\ U^{-1}(u), & \text{if } u > 0. \end{cases}$$

Adjustment Hazard Model



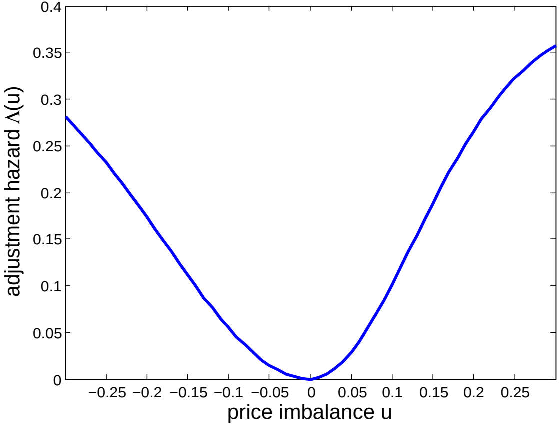
Adjustment Hazard Model

We can now summarize the optimal policy via a state-dependent adjustment hazard, $\Lambda(u)$.

Given a price deviation u, we denote by $\Lambda(u)$ the probability of adjusting the price before observing the current draw of the adjustment cost. Therefore:

$$\Lambda(u) \equiv G(\Omega(u)).$$

Adjustment Hazard Model



Adjustment Hazard Model

- It is straightforward to prove that $\Lambda(u)$ is decreasing for $u < 0$ and increasing for $u > 0$.
  This is referred to as the increasing hazard property: the hazard increases with the absolute value of the price-deviation u
- The standard (L, C, U) policy is obtained when $G(\omega)$ has all its mass at one point
- The Calvo model ($\Lambda(u) = \lambda$ for all u) is obtained when $G(\omega)$ has mass λ at $\omega = 0$ and mass $1 - \lambda$ at a very large value of ω

Adjustment Hazard Model

- This is a state-dependent adjustment hazard, to distinguish it from the usual, time-dependent hazard:
  - time-dependent: probability of adjusting, conditional on having last adjusted t periods ago
  - state-dependent: probability of adjusting given current state (in our case: real price and current adjustment cost)
- In an adjustment hazard model, the size of upward and downward adjustments can vary over time, which does not happen with the standard (L, C, U) model
- This added realism also is useful when estimating the model
- More on adjustment hazards: later in the course.

Adjustment Hazard and Imperfect Information

- So far: assumed the firm observes its real price x at no cost
- In recent work, Mike Woodford relaxes this assumption:
  - the firm is uncertain about its real price, its uncertainty is described by a probability density $h(x)$ that evolves over time
  - every period the firm can pay to reduce its uncertainty; the more it pays, the smaller the variance of the new $h(x)$, yet it cannot eliminate uncertainty completely
- The adjustment cost is always the same, as in the (L, C, U) model

Adjustment Hazard and Imperfect Information

Yet imperfect information implies that the optimal policy leads to an increasing adjustment hazard:

- with perfect information, the optimal policy would be a standard (L, C, U) policy, yet the firm does not know whether its x is in the inaction range or not
- as $h(u)$ moves closer to one of the triggers it pays more to pay for information that makes this distribution more precise; thus as your true value of u moves closer to one of the triggers, the chances that you adjust increase smoothly
- this leads to a hazard that increases smoothly with $|u|$

8.3. Brownian motion: a brief review

A continuous time process $(X_t)_{t \geq 0}$ is a Brownian motion with parameters (μ, σ^2), which we denote $BM(\mu, \sigma^2)$, if:

- (a) $X_0 \equiv 0$.
- (b) Independent increments:

$$t_0 < t_1 < \ldots < t_n \Rightarrow X_{t_n} - X_{t_{n-1}}, X_{t_{n-1}} - X_{t_{n-2}}, \ldots, X_{t_1} - X_{t_0}$$

  are independent.
- (c) $s < t \Rightarrow X_t - X_s \rightsquigarrow N(\mu(t-s), \sigma^2(t-s))$.
- (d) $\Pr\{\omega : t \mapsto X_t(\omega) \text{ continuous}\} = 1$.

Note: it can be shown that (a), (b) and (c) $\Rightarrow$ (d). Also that (a), (b) and (d) $\Rightarrow$ (c).

Brownian motion: a brief review

For a BM:

- $\Pr\{\omega : t \mapsto X_t(\omega)$ is differentiable at no point$\} = 1$.
- Martingale Property: $s < t \Rightarrow E_s[X_t] = X_s$.

Comments:

- The BM is the continuous time version of a random walk. Its sample paths are everywhere continuous and nowhere differentiable.
- A BM(0,1) is also known as a Wiener process or standard BM.
- μ is the drift and σ the instantaneous standard deviation.
- X follows a Geometric Brownian Motino if $\log X$ follows a Brownian Motion

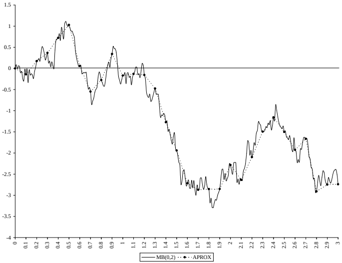
Some Curiosities

- If we divide the interval $[0, t]$ into 2^n segments of length $t/2^n$, so that the k-th segment is

$$\left[\frac{k-1}{2^n}t, \frac{k}{2^n}t\right],$$

we have that the expected value of the sum of lengths of the line segments that join the points of the BM at the preceding points in time tends to infinity when $n \to \infty$ (see the figure on the next slide).

Some Curiosities



Some Curiosities

- Indeed:

$$\begin{aligned}\sum_{k=1}^{2^n} \mathrm{E}\left|X_{\frac{k}{2^n}t} - X_{\frac{k-1}{2^n}t}\right| &= \sum_{k=1}^{2^n} \mathrm{E}\left|N\left(\frac{\mu}{2^n}t, \frac{\sigma^2}{2^n}t\right)\right| \\ &\leq \sum_{k=1}^{2^n} \mathrm{E}\left|N\left(0, \frac{\sigma^2}{2^n}t\right)\right| \\ &= 2^n\frac{C_0\sigma}{2^{n/2}}\sqrt{t} = C_0\sigma\sqrt{t}2^{n/2} \stackrel{n\to\infty}{\longrightarrow} \infty,\end{aligned}$$

where we used $\mathrm{E}|N(\mu,\sigma^2)| \leq \mathrm{E}|N(0,\sigma^2)|$, $C_0 \equiv \mathrm{E}|N(0,1)|$ and $\mathrm{E}|N(0,\tau^2)| = \tau C_0$.

Some Curiosities

- By contrast, the expected quadratic variation is bounded:

$$
\begin{aligned}
\mathsf{E}(\text{quadratic variation}) &\equiv \mathsf{E}\sum_{k=1}^{2^n}\left(X_{\frac{k}{2^n}t} - X_{\frac{k-1}{2^n}t}\right)^2 \\
&= \sum_{k=1}^{2^n}\mathsf{E}\left[N\left(\frac{\mu}{2^n t}, \frac{\sigma^2}{2^n}t\right)^2\right] \\
&\leq \sum_{k=1}^{2^n}\mathsf{E}\left[N\left(0, \frac{\sigma^2}{2^n}t\right)^2\right] \\
&= \sum_{k=1}^{2^n}\frac{\sigma^2}{2^n}t = 2^n\frac{\sigma^2}{2^n}t = \sigma^2 t < \infty.
\end{aligned}
$$

- We conclude that a BM moves a great distance, without getting anywhere because it changes direction often and abruptly.

Discretizing a BM

- We consider Δt and Δx such that:

$$X(0), X(\Delta t), X(2\Delta t), X(3\Delta t), \ldots$$

is a Markov chain that satisfies:

$$X(t+\Delta t) = \begin{cases} X(t) + \Delta x, & \text{with probability } p, \\ X(t) - \Delta x, & \text{with probability } 1-p. \end{cases}$$

Discretizing a BM

- We choose p, Δt and Δx so as to match the first two moments of the BM:

$$E[\Delta X(t)] = p\Delta x + (1-p)(-\Delta x) = (2p-1)\Delta x = \mu \Delta t.$$

$$\begin{aligned} E[(\Delta X(t))^2] &= (\Delta x)^2 \\ &= E\left[N(\mu\Delta t, \sigma^2 \Delta t)^2\right] = \mu^2(\Delta t)^2 + \sigma^2 \Delta t. \end{aligned}$$

Discretizing a BM

- It follows that:

$$p = \frac{1}{2}\left[1 + \frac{\mu}{\sigma}\sqrt{\Delta t}\right], \tag{9}$$

$$\Delta x = \sigma\sqrt{\Delta t}\sqrt{1 + \frac{\mu^2}{\sigma^2}\Delta t} \simeq \sigma\sqrt{\Delta t}, \tag{10}$$

  where we used the assumption $\Delta t \ll 1$ in the last step, or more precisely, $\frac{\mu^2}{\sigma^2}\Delta t \ll 1$).
- Hence, we can fix one of the parameters p, Δx and Δt (usually you fix Δt), and determine the remaining two via (10) and (9). It can then be shown that as $\Delta t \to 0$ the discrete process $[X(0), X(\Delta t), X(2\Delta t), X(3\Delta t), \ldots]$ converges, in a precise sense, to the BM(μ, σ) $(X_t)_{t\geq 0}$.

An Alternative Discretization

- We'll refer to the discretization described above as the traditional discretization. Here we consider an alternative discretization, which we will use later to provide heuristic derivations of some results.
- We consider a $BM(\mu\,,\,\sigma^2)$, $(X_t)_{t\geq 0}$ and approximate it by a Markov chain:

$$X(0), X(\Delta t), X(2\Delta t), X(3\Delta t), \ldots$$

where the X take values $\{\ldots, -2\Delta x, -\Delta x, 0, \Delta x, 2\Delta x, \ldots\}$ so that:

$$X(t+\Delta t) = \begin{cases} X(t) + \mu\Delta t + \sigma\sqrt{\Delta t} & \text{with prob. } 1/2, \\ \\ X(t) + \mu\Delta t - \sigma\sqrt{\Delta t} & \text{with prob. } 1/2, \end{cases}$$

where $\mu\Delta t + \sigma\sqrt{\Delta t}$ and $\mu\Delta t - \sigma\sqrt{\Delta t}$ are multiples of Δx.

An Alternative Discretization

- Note that:

$$\begin{aligned} \mathsf{E}[\Delta X_t] &= \mu \Delta t, \\ \mathsf{Var}[\Delta X_t] &= \sigma^2 \Delta t, \end{aligned}$$

thus matching the first two moments of the BM.

Why a Brownian Motion?

- Often the true underlying process for y^* is close to a BM.
- Optimal policies are often simpler when you work in continuous time. And the BM is the continuous time version of a random walk.
  - Look at the proof of optimality of one-sided *Ss* policies and note that, in the case of continuous time, the proof is trivial, since z_t cannot jump.
  - The same holds with the standard discretization of a BM described above, since the process changes by $\pm\Delta X$ each period.
  - In both cases, the process cannot 'skip' over states

Why a Brownian Motion?

- Small number of state variables:
  - The BM is one of the simplest continuous time stochastic process. In particular, often the optimal policy does not depend on the current realization of the BM-shock (i.e., *Ss* bands do not vary over time).
  - Next simplest is Ornstein-Uhlenbeck, continuous-time version of an AR(1). In this case the optimal policy depends on the latest realization of the BM (time-varying *Ss* bands).
  - Related to the curse of dimensionality:
    - Important in partial-equilibrium models
    - Crucial in general-equilibrium models

8.4. *Ss* policies in continuous time

A. An Example

B. (More) General Case

A. Ss policies in continuous time: An Example

- Dixit (JPE, 1989)
- Firm's entry and exit decision:
  - fixed entry and exit costs
  - output price follows a random walk
- Optimal policy: exit and entry triggers

Hysteresis

- Definition: failure of an effect to reverse itself as its underlying cause is reversed
- Important in entry-exit decisions
- Example: foreign firms that entered the U.S. market appreciated (mid 1980s) did not exit when the dollar fell back to its original value

Simple Cases

- Single discrete project
- "Firm" (could be worker, consumer, etc.) has discount rate ρ
- Sunk investment: k
- No depreciation
- Immediate rusting if unused
- Avoidable operating cost w per unit of time (fixed)
- Output flow of the project is one unit, thus revenue from the project equals the price P
- In all cases the optimal decision rule is described by two triggers: P_H, P_L, with $P_H > P_L$ and such that the investment should be made if $P > P_L$ and it should be abandoned if P falls below P_L

Simple Case 1

- Firm has no investment in place and it believes current P will persist forever
- Then it will make the investment if

$$P > w + \rho k \equiv P_H$$

where the r.h.s. is the annualized cost of making and operating the investment.

- Conversely, firm has such a project in place and the price falls to P_1, where the firm believes it will stay forever
- Then the firm abandons the project if

$$P_1 < w \equiv P_L$$

- Full cost defines entry trigger P_H
- Variable cost defines exit trigger P_L
- Standard Marshalian theory of the long run vs. the short run

Simple Case 1

- Hysteresis:
  - Initially: $P_L < P < P_H$
  - Then P rose above P_H and the investment was made
  - Then the price fell to its original level, which was insufficient to induce abandonment
- Two problems with above story:
  - Irrational (myopic) expectations
  - Quantitatively small effect

Simple Case 2

- The “usual” value P^* of P is in the range $(w, w + \rho k)$
- P has risen to a higher value but is expected to revert to P^*
- Due to mean-reversion, a price above $w + \rho k$ does not suffice to induce investment. The trigger P_H will be above $w + \rho k$
- Similarly, with mean reversion we have that $P_L < w$

Simple Case 3

- Ongoing uncertainty but no mean-reversion
- Current price: $w + \rho k$
- From here on it will move up or down by h, with equal probabilities
- If the firm invests right away and continues active forever, its expected present value net of investment cost is zero
- If it waits one period:
  - $P \uparrow$: invests next period and obtains positive present value
  - $P \downarrow$: does not invest, present value is zero
- Expected value of waiting one period: positive: option value feature
- In this case: $P_H > w + \rho k$
- Similarly: $P_L < w$

The Model

- Firm can become active by investing a lump sum *k*. Then produces a unit of output at a variable cost *w*.
- It can suspend operations by paying a lump sum exit cost *l*
- It must repay *k* if it decides to reenter at some future time
- Cost of capital $=$ interest rate $= \rho$
- w, k, l, ρ are constant and nonstochastic
- Uncertainty: *P* follows a Geometric Brownian Motion
- The firm is risk neutral and maximizes its expected present value

Price Process

- We assume that $\log P$ follows a BM(μ, σ^2)
- This is often written as:

$$\frac{dP}{P} = \mu dt + \sigma dz$$

  with z_t a standard BM(0,1).
- Thus $\log P_t - \log P_0$ is $N(\mu, \sigma^2)$
- Thus $\mathrm{E}[dz] = 0$ and $\mathrm{E}[dz^2] = dt$

Solving the firm's problem

Two state variables: (P, k) where $k = 1$ if the firm is active and $k = 0$ if it is inactive Denote the corresponding value functions by $V_1(P)$ and $V_0(P)$

Assume the price is P at time t, the firm is active at t, and the firm does not become inactive in $(t, t + \Delta t)$. Then, denoting by $\mathrm{E}[dV]$ the expected change in the value of the firm during $(t, t + \Delta t)$ we have:

$$\mathrm{E}[dV_1] = \tfrac{1}{2}[V_1(Pe^{\mu\Delta t+\sigma\sqrt{\Delta t}}) + V_1(Pe^{\mu\Delta t-\sigma\sqrt{\Delta t}})] - V_1(P)$$

which, via Taylor expansions of 2nd order leads to

$$\mathrm{E}[dV_1] = \mu P V_1'(P)\Delta t + \tfrac{1}{2}\sigma^2 P^2 V_1''(P)$$

Solving the firm's problem

From the basic finance identity:

$$\rho V(P) = \text{Dividends} + \text{Capital Gains}$$

we therefore have:

$$\rho V_1(P) = P - w + \mu P V_1'(P)\Delta t + \tfrac{1}{2}\sigma^2 P^2 V_1''(P)$$

and therefore

$$\tfrac{1}{2}\sigma^2 P^2 V_1''(P) + \mu P V_1'(P) - \rho V_1(P) = w - P.$$

A similar argument leads to:

$$\tfrac{1}{2}\sigma^2 P^2 V_0''(P) + \mu P V_0'(P) - \rho V_0(P) = 0.$$

Solving the firm's problem

Thus two pretty standard differential equations, that can be solved explicitly:

$$V_0(P) = A_0 P^{-\alpha} + B_0 P^{\beta}, \tag{11}$$

$$V_1(P) = A_1 P^{-\alpha} + B_1 P^{\beta} + \left(\frac{P}{\rho - \mu} - \frac{w}{\rho} \right), \tag{12}$$

with

$$\beta = \frac{1 - m + \sqrt{(1-m)^2 + 4r}}{2} > 1,$$

$$-\alpha = \frac{1 - m + \sqrt{(1-m)^2 + 4r}}{2} < 0,$$

with $m = 2\mu/\sigma^2$ and $r = 2\rho/\sigma^2$.

Interpretation

- The term in parenthesis on the r.h.s. of (12) is the value obtained if the project is active forever. Therefore the remaining part of the solution must be the value of the option to close down optimally.
- Similarly, the whole expression for $V_0(P)$ must be the valued of becoming active for an idle firm.

Endpoint conditions

- If P very small, the option value of activating should be small, thus $A_0 = 0$
- Similarly $B_1 = 0$
- Hence:

$$\begin{aligned} V_0(P) &= BP^{\beta}, \\ V_1(P) &= AP^{-\alpha} + \left(\frac{P}{\rho - \mu} - \frac{w}{\rho} \right), \end{aligned}$$

Endpoint conditions

- Value matching conditions:

$$\begin{aligned} V_0(P_H) &= V_1(P_H) - k, \\ V_1(P_L) &= V_0(P_L) - l. \end{aligned}$$

- Smooth pasting conditions:

$$\begin{aligned} V_0'(P_H) &= V_1'(P_H), \\ V_0'(P_L) &= V_1'(P_L). \end{aligned}$$

- Hence we have four equations for four unknowns: A, B, P_H and P_L

Implications

Based on the optimal solution characterized above Dixit shows that:

$$
\begin{aligned}
P_H &> w + \rho k \equiv W_H, \\
P_L &< w - \rho l \equiv W_L.
\end{aligned}
$$

Comparative statics: the figure on the following slide shows that P_H/W_H is increasing in σ while P_L/W_L is decreasing in σ. It also shows that both are decreasing in μ. Most important, the differences between the optimal and the Marshalian triggers are large for reasonable parameter values.

Comparative Statics

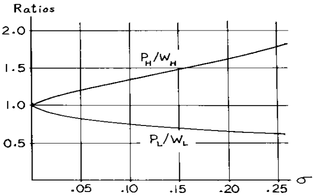


Fig. 3.—Effects of changes in σ

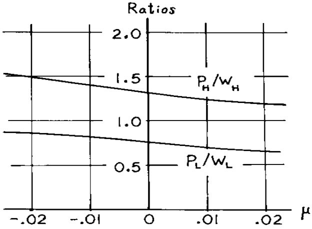


Fig. 4.—Effects of changes in μ

B. *Ss* policies in continuous time: General Case

- Based on Dixit (1991).
- The problem::
  - A state variable (e.g., price gap $p - p^*$), z, follows a $BM(\mu\,,\,\sigma^2)$ when no control is exerted.
  - Exerting control (adjusting) is costly, entailing both a fixed and a variable component:

$$B(\Delta z) = \begin{cases} K^+ + k^+\Delta z, & \Delta z > 0; \\ K^- - k^-\Delta z, & \Delta z < 0, \end{cases}$$

where K^+, K^-, k^+ and $k^- \geq 0$.

Ss policies in continuous time: General Case

- The case of one-sided *Ss* policies corresponds to $K^+ > 0$ and $k^+ = 0$; K^- and k^- are irrelevant in that case.
- The case of (L, c, U) policies corresponds to $k^+ = k^- = 0$ and $K^+ = K^- > 0$.
- The utility/profit flow of having z is $\pi(z)$, with π quasiconcave with maximum at $\hat{z}$.
- The agent faces the following problem:

$$\max_{z_t} \mathrm{E}\left[PV(\pi - \text{Adjustment costs})\right].$$

Examples

- All variables that follow are in logs.
- Prices:

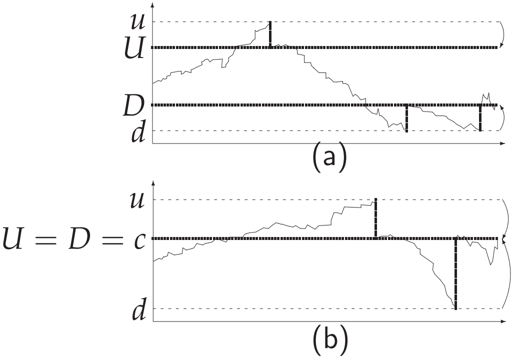
  x_t = Nominal price at t,
  x_t^* = Nominal price if there were no frictions,
  $z_t \equiv x_t - x_t^*$,
  $x_t^* \rightsquigarrow BM(\mu\,,\,\sigma^2)$,
  z_t when not adjusting $\rightsquigarrow BM(-\mu\,,\,\sigma^2)$.
- Durable good:

  x_t = Durable stock at t, depreciates at rate δ,
  x_t^* = Optimal durable stock in the absence of frictions,
  $z_t \equiv x_t - x_t^*$,
  $x_t^* \rightsquigarrow BM(\mu\,,\,\sigma^2)$,
  z_t when not adjusting $\rightsquigarrow BM(-\mu - \delta\,,\,\sigma^2)$.

Examples

- Investment:
  x_t = Capital stock in t, state variable, depreciates at rate δ,
  x_t^* = Desired capital stock at t,
  $z_t \equiv x_t - x_t^*$,
  $x_t^* \rightsquigarrow BM(\mu\,,\,\sigma^2)$,
  z_t when not adjusting $\rightsquigarrow BM(-\mu-\delta\,,\,\sigma^2)$.

Shape of the Optimal Policy



Shape of the Optimal Policy

- Panel (a) in the previous Figure presents the optimal policy when we have both fixed and proportional adjustment costs
- Panel (b) considers the case with no proportional adjustment costs.

Some particular cases

1. $U = D$ if $k^+ = k^- = 0$, that is, if we have a fixed adjustment cost. The corresponding optimal policy is illustrated in Panel (b) of the preceding figure.
2. If $K^- \to \infty$ or $k^- \to \infty$, the agent will never adjust downwards and we have a one-sided Ss policy.

Finding the Bands of the Optimal Policy

- Assuming (it can be proved formally) that the optimal policy is of the sort described above, how do we determine the optimal values of u, d, U and D?
- Let $J(x)$ denote the value function evaluated at values of x where no adjustment takes place, that is, $d < x < u$ (we are excluding only two feasible values of x: d and u).
- Then, denoting by r the discount rate and using the alternative discretization of a BM:

$$
\begin{aligned}
J(x) &\overset{\Delta t \ll 1}{\cong} \pi(x)\Delta t + \\
&+ \frac{1}{2} e^{-r\Delta t} \left\{ J(x - \mu\Delta t + \sigma\sqrt{\Delta t}) + J(x - \mu\Delta t - \sigma\sqrt{\Delta t}) \right\}.
\end{aligned}
$$

Finding the Bands of the Optimal Policy

- Taking a second order Taylor expansion (formally we are applying Ito's Lemma) and doing some algebra (we consider terms of $O(\Delta t)$ but ignore higher order terms):

$$J''(x) - \frac{2\mu}{\sigma^2}J'(x) - \frac{2r}{\sigma^2}J(x) + \frac{2\pi(x)}{\sigma^2} = 0. \qquad (13)$$

- The solution to this differential equation is of the form:

$$J(x) = A_1 e^{\lambda_1 x} + A_2 e^{\lambda_2 x} + J_0(x)$$

where λ_1 and λ_2 are the roots of:

$$\lambda^2 - \frac{2\mu}{\sigma^2}\lambda - \frac{2r\mu}{\sigma} = 0$$

and $J_0(x)$ denotes the expected discounted flow payoff ignoring all barriers and controls (see Dixit for a proof that $J_0(x)$ satisfies (13)).

Finding the Bands of the Optimal Policy

- To determine d, u, D, U, A_1 and A_2 we need six "independent" equations
- Value matching

  The value of the program should not change when adjustment takes place. This condition is equivalent to imposing that, when the agent adjusts, she must be indifferent between adjusting and not adjusting. Hence:

$$J(U) - J(u) = K^- + k^-(u - U),$$

  and, analogously:

$$J(D) - J(d) = K^+ + k^+(D - d).$$

Finding the Bands of the Optimal Policy

- Smooth pasting

  The rate at which the value function changes should be the same before and after you adjust (smooth pasting) refers to the behavior of the value function being "similar" at both pairs of trigger and target points) .

- This leads to conditions on the first derivatives of $J(x)$, one of which we derive formally in what follows:

$$J(U) - B(U - u) = \max_x \{J(u + x) - B(x)\},$$

where x denotes the change in z that takes place when adjusting from u to U ($\Delta z < 0$ in this case).

Finding the Bands of the Optimal Policy

- The FOC for this problem is:

$$J'(u+x) = B'(x),$$

with the optimal x equal to $U - u$. It follows that:

$$J'(U) = -k^{-}.$$

- Dixit provides similar derivations for $J'(u) = -k^{-}$, $J'(D) = J'(d) = k^{+}$, thereby providing four smooth pasting conditions, which combined with two value matching conditions can be used to fully specify the solution $J(x)$. He also provides a graphical interpretation of these conditions.

8.5. Invariant and Ergodic Distribution

Aggregate and Idiosyncratic Shocks

- We consider a large number (continuum) of agents that follow the same *Ss* policies, but may face different shocks
- Shocks faced by agents have a common and an agent-specific component.
- The former are referred to as aggregate shocks, the latter as idiosyncratic shocks.
- Idiosyncratic shocks are independent across agents, and independent from the aggregate shock

Aggregate and Idiosyncratic Shocks

- If the frictionless variable follows a BM we assume that the aggregate (common) shock follows a $BM(\mu_A\,,\ \sigma_A^2)$ and idiosyncratic shocks are independent across agents (and independent from the aggregate shocks) and follow a $BM(0\,,\ \sigma_I^2)$.
- Similarly, in discrete time, if the frictionless variable follows a random walk, then we assume that the aggregate shock follows a random walk with drift μ and variance of innovations σ_A^2, while idiosyncratic shocks are independent across agents and independent from the aggregate shock, and follow a random walk with zero drift and variance of innovations σ_I^2.
- In both cases the relative importance of aggregate and idiosyncratic shocks is captured by σ_A/σ_I.

Invariant and Ergodic Density

- An invariant density (or invariant distribution) is such that, if this density describes the x-section of agents' state variables at time t it also describes it at $t + dt$ (if working in continuous time) or $t + 1$ (if working in discrete time).
- Next we find the invariant density for the policies described above.
- In the presence of aggregate shocks (i.e., if $\sigma_A > 0$), there does not exist an invariant distribution. What is well defined in this case, though, is an ergodic distribution, which can be interpreted as the weighted average over all possible distributions you might observe, with weights reflecting the likelihood of observing a given distribution.

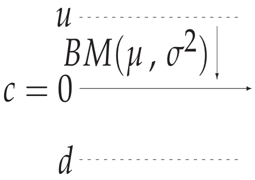
Invariant and Ergodic Density

- For example, working in discrete time, denote by G_t the operator that transforms the density of deviations at time $t-1$ into the density of deviations at t. This operator depends on v_t^A, σ_I and the parameters that characterize the *Ss* policy. Then we say there exists an ergodic density if:

$$f_E(x) = \lim_{T\to\infty} \frac{1}{T} \mathrm{E} \sum_{t=1}^{T} G_t G_{t-1} \cdots G_1 f_0$$

  exists and does not depend on the initial density f_0, where the expectations operator averages over the aggregate shocks.
- Next we derive the invariant density when there are no aggregate shocks ($\sigma_A = 0$). It is easy to argue that this is also the ergodic density, as long as we consider the variance of the corresponding process to be $\sigma_T^2 \equiv \sigma_A^2 + \sigma_I^2$.

Informal Derivation



Informal Derivation

- In what follows we assume $U = D = c$. Also, without loss of generality we assume $c = 0$ (see panel b in the Figure on page 127).
- An invariant density $f(x)$ must satisfy (we are using the alternative discretization of a BM here):

$$f_{t+\Delta t}(z) = \frac{1}{2}\left[f_t(z + \mu\Delta t - \sigma\sqrt{\Delta t}) + f_t(z + \mu\Delta t + \sigma\sqrt{\Delta t})\right],$$

  where the first term on the r.h.s. represents those agents who were one unit above z and the second term those who were one unit below z, in both cases at time t.
- Taking a Taylor expansion (strictly: Ito's Lemma) leads to:

$$\alpha f'(z) = f''(z) + O(\Delta t)$$

  with $\alpha = -2\mu/\sigma^2$.