

Modern Dynamic Asset Pricing Models

Teaching Notes 3.

Uncertainty, Learning and Asset Pricing

Pietro Veronesi

*University of Chicago
Booth School of Business
CEPR, NBER*

Economic Uncertainty, Information Quality, and Learning

- Questions: Can one or more of these puzzles be addressed by lack of information about some key parameter of the model?

- More precisely
 - What is the relationship between the *quality* of information and asset returns?

 - If information is noisy, is there a risk premium?

 - Does information quality affect the relationship between the risk premium and investors' risk aversion?

 - How does the precision of signal affect stock market volatility?

 - Can we infer how good investors' information is from the behavior of stock market returns?

- We now consider a simple modification of the above setting to answer this question (Reference: Veronesi (2000, JF))

A Model of Economic Bayesian Uncertainty

- Dividends grow according to the following process:

$$\frac{dD_t}{D_t} = \theta_t dt + \sigma_D dB_t^D$$

- Investors do not observe the drift θ_t , but they know its law of motion.
- **Assumption 1:** θ_t follows a Markovian process defined on a finite set $\Theta = \{\theta_i\}_{i=1}^n$.
- That is, for every $\theta_i, \theta_j \in \Theta$ there exists λ_{ij} such that in the infinitesimal interval Δ we have:

$$\Pr(\theta_{t+\Delta} = \theta_j | \theta_t = \theta_i) = \lambda_{ij} \Delta$$

- **Remark:** Although I assume that θ_t can take only a finite number of values, the assumption leaves unspecified the *number* of states.
- We can approximate any continuous-time, continuous-state stationary Markov process by choosing a sufficiently fine grid $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ on the real line and by carefully choosing the transition probabilities λ_{ij} .

A Model of Economic Bayesian Uncertainty

- Examples:

1. **Business Cycle:** Use only two states $\theta_1 < \theta_2$. See Veronesi (1999), David and Veronesi (2000) and Locarno and Massa (2000), Cagetti, Hansen, Sargent and Williams (2000).

2. **Pure Jump Process:**

$$d\theta_t = (J_t - \theta_t) dQ_t^{p(\theta_t)}$$

– where $dQ_t^{p(\theta_t)}$ denotes a Poisson process with intensity $p(\theta_t)$ and J_t is a random variable with any density $f(\theta)$.

3. **Pure Drift Uncertainty:** $p(\theta) = 0$ in the previous case.

4. **Mean Reverting Processes:**

$$d\theta_t = k(\bar{\theta} - \theta_t) dt + \sigma_\theta dB_t^\theta$$

– where $dB_t^\theta dB_t^D = 0$.

5. **Mean Reverting Process with Jumps:**

$$d\theta_t = k(\bar{\theta} - \theta_t) dt + \sigma_\theta dB_t^\theta + (J_t - \theta_t) dQ_t^{p(\theta_t)}$$

– This is a very complex filtering problem, used in Veronesi (2003).

Economic Uncertainty and Information Quality

- Investors observe a noisy signal:

$$de_t = \theta_t dt + \sigma_e dB_t^e$$

- where B_t^e is a standard Brownian motion independent of B_t^D .
- This form of the signal is the continuous time analog of the standard “signal equals fundamentals plus noise”, i.e. $e_t = \theta_t + \varepsilon_t$ with ε_t normally distributed, in a discrete time model (see e.g. Detemple (1986)).

$$h_e = 1/\sigma_e = \textit{precision of the external signal}$$

- Similarly, $h_D = 1/\sigma_D = \textit{precision of dividend signal}$

Economic Uncertainty, Information Quality, and Learning

- Denote investors' information set at time t by \mathcal{F}_t , and let

$$\pi_t^i = \text{Prob}(\theta_t = \theta_i | \mathcal{F}_t) \quad (1)$$

- Define $\boldsymbol{\pi}_t = (\pi_t^1, \dots, \pi_t^n)$. This distribution summarizes investors' overall information at time t .
- The expected drift rate at time t $m_t^\theta \equiv E(\theta | \mathcal{F}_t) = \sum_{i=1}^n \pi_t^i \theta_i$
- The evolution of π_t^i is given by
- **Lemma 1:** (a) For all $i = 1, \dots, n$:

$$d\pi_t^i = [\boldsymbol{\pi}_t \boldsymbol{\Lambda}]_i dt + \pi_t^i (\theta_i - m_t^\theta) \left(h_D d\tilde{B}_t^D + h_e d\tilde{B}_t^e \right) \quad (2)$$

- where $\boldsymbol{\Lambda}$ is such that $[\boldsymbol{\Lambda}]_{ij} = \lambda_{ij}$ for $i \neq j$ and $[\boldsymbol{\Lambda}]_{ii} = -\sum_{j \neq i} \lambda_{ij}$.
- In this equation

$$d\tilde{B}_t^D = h_D \left(\frac{dD_t}{D_t} - m_t^\theta dt \right); \quad d\tilde{B}_t^e = h_e (de_t - m_t^\theta dt)$$

Economic Uncertainty and Information Quality

- To understand the effect of information quality on asset prices, it is convenient to look at a specific example.
- Assume θ follows a pure jump process.

$$d\theta_t = (J_t - \theta_t)dQ_t^p \quad \text{with} \quad J_t \sim F = (f_1, \dots, f_n) \quad (\text{stationary distribution})$$

- Then if $\theta(t) = \theta_\ell$:

$$d\pi_i = [p(f_i - \pi_i) + k\pi_i(\theta_i - m_\theta)(\theta_\ell - m_\theta)] dt + \sqrt{k}\pi_i(\theta_i - m_\theta)dB_t$$

- where $m_\theta \equiv E(\theta|\mathcal{F}(t))$, $k = h_D^2 + h_e^2$, and $dB_t = \frac{1}{\sqrt{k}} (h_D dB_t^D + h_e dB_t^e)$
 1. More precise signals \implies posterior π_t more concentrated around the true state θ_ℓ ;
 2. Less precise signals \implies posterior π closer to stationary distribution $F = (f_1, \dots, f_n)$;

Economic Uncertainty and Stock Prices

- The price function is:

$$\frac{P_t}{D_t} = \sum_{i=1}^n \pi_t^i C(\theta_i)$$

- where $C(\theta)$ is defined as follows. Define the constant

$$H = \sum_{i=1}^n \frac{f_i}{\phi + p + (\gamma - 1)\theta_i + \frac{1}{2}\gamma(1 - \gamma)\sigma_D^2} \quad (3)$$

then

$$C(\theta) = \frac{1}{(\phi + p + (\gamma - 1)\theta + \frac{1}{2}\gamma(1 - \gamma)\sigma_D^2) (1 - pH)} \quad (4)$$

- $C(\theta)$ is monotone and convex
- $C(\theta)$ is *decreasing* if and only if $\gamma > 1$.

Economic Uncertainty and the Stock Price

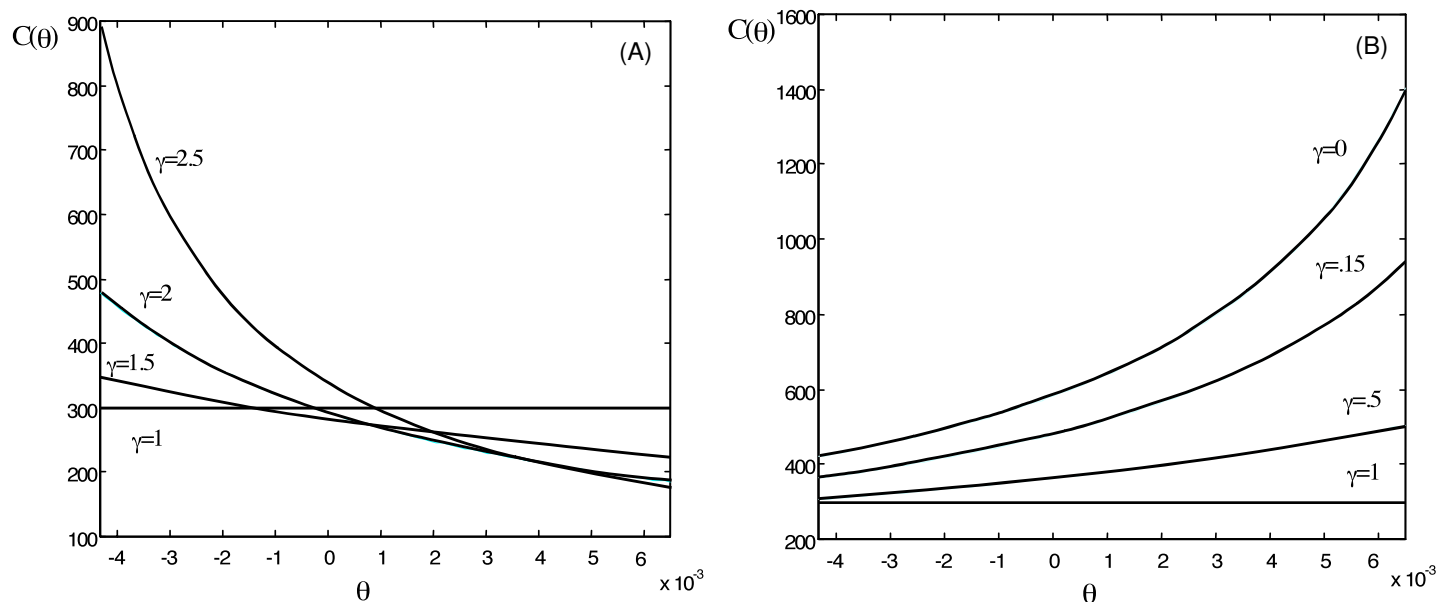


Figure 1. The function $C(\theta)$. (A) plots the function $C(\theta)$ for various values of investors' coefficient of risk aversion $\gamma \geq 1$. This function represents investors' marginal valuation of the stock as a multiple of the current dividend when they condition on the true drift of the dividends process being θ . (B) plots the same function for values of $\gamma \leq 1$.

- Why $C(\theta)$ is decreasing for $\gamma > 1$?
 - Because of low Elasticity of Intertemporal Substitution (EIS)
 - * Low EIS \implies desire of consumption smoothing
 - * \implies Higher $\theta \implies$ higher future consumption \implies lower savings today \implies lower prices (and higher r)

Economic Uncertainty and Stock Prices

- The previous result has an intriguing implication:
 - \implies A mean preserving spread on $\pi \implies$ an increase in P/D .
 - \implies Higher uncertainty increases the P/D ratio.
- Intuition:
 - Higher uncertainty increases the probability of high θ and low θ
 - Compounding effect: 1/2 - 1/2 chance of growing at 10% or 0% is more valuable than sure probability of growing at 5%.
- (Note that the discount adjusts appropriately according to GE model)
- Pastor and Veronesi (2003, JF) document this uncertainty effect on individual stocks, when uncertainty is proxied by firm age.
- Pastor and Veronesi (2006, JFE) uses this intuition to “explain” the tech bubble in the late 1990s.

Economic Uncertainty and Stock Returns

- Let's denote the total excess returns by

$$dR = \frac{dS_t + D_t dt}{S_t} - r_t dt \quad (5)$$

- **Proposition:** The equilibrium excess returns follow the process:

$$dR = \mu_R dt + (\sigma_D + h_D V_\theta(\boldsymbol{\pi}_t)) d\tilde{B}_t^D + V_\theta(\boldsymbol{\pi}_t) h_e d\tilde{B}_t^e \quad (6)$$

where

$$\mu_R = \gamma (\sigma_D^2 + V_\theta(\boldsymbol{\pi}_t)) \quad (7)$$

$$V_\theta(\boldsymbol{\pi}_t) = \frac{\sum_{i=1}^n \pi_t^i C_i (\theta_i - m_t^\theta)}{\sum_{i=1}^n \pi_t^i C_i} = \sum_{i=1}^n \bar{\pi}_t^i \theta_i - \sum_{i=1}^n \bar{\pi}_t^i \theta_i = \bar{m}_t^\theta - m_t^\theta \quad (8)$$

- where

$$\bar{\pi}_t^i = \frac{\pi_t^i C_i}{\sum_{j=1}^n \pi_t^j C_j} \quad (9)$$

Economic Uncertainty and Stock Returns

- The function $V_\theta(\boldsymbol{\pi}_t)$ characterizes both expected returns μ_R and volatility.
- $V_\theta(\boldsymbol{\pi}_t)$ is a measure of both level of uncertainty about the true growth rate θ as well as of the impact of this uncertainty on the investors' own valuations of the asset.
 - For example, when $\gamma = 1$ we have that $C_i = C_j$ for all i and j . Hence, $V_\theta(\boldsymbol{\pi}) = 0$: Investors may be uncertain, but they do not care (because they are myopic).
- It is possible to characterize $V_\theta(\boldsymbol{\pi}_t)$ and we obtain the following
 - **Proposition:**
 1. If $\gamma > 1$, then higher uncertainty decreases the risk premium. That is, a mean preserving spread on investors' beliefs $\boldsymbol{\pi}_t$ decreases μ_R .
 2. If either $m_t^\theta > \sigma_D^2 + \theta_1$ or $\pi_t^1 < \bar{\pi}_1^*$ where $\bar{\pi}_1^*$ is a given constant (quite high), the expected excess return μ_R decreases with γ for γ sufficiently high. Hence, μ_R is bounded above.
 3. If $m_t^\theta > \sigma_D^2 + \theta_1$, there is $\bar{\gamma}$ such that $\mu_R < 0$ for $\gamma > \bar{\gamma}$. Moreover, a mean-preserving spread on $\boldsymbol{\pi}$ decreases $\bar{\gamma}$.

Economic Uncertainty and Stock Returns: Intuition

- Part 1 shows that there is no premium for uncertainty. Actually, quite the opposite holds.

- Intuition?

- Recall

$$P_t = D_t \times \sum_{i=1}^n \pi_t^i C(\theta_i)$$

* $D_t \downarrow \implies P_t \downarrow$ because of the first term.

* $D_t \downarrow \implies E_t[\theta] \downarrow \implies$ consumption smoothing $\implies P_t/D_t \uparrow$ because of the second term.

- Higher uncertainty, second effect stronger.
- Since the premium is given by

$$\mu_R = \gamma \text{Cov}_t \left(dR_t, \frac{dC_t}{C_t} \right)$$

- the stronger the second effect, the lower is the covariance between dR_t and $dC_t/C_t = dD_t/D_t$.
- \implies The equity premium is lower for higher economic uncertainty.

Economic Uncertainty, Risk Aversion and Stock Returns

- Similarly, higher risk aversion \implies lower EIS \implies second effect stronger
 - \implies Equity premium is bounded above.

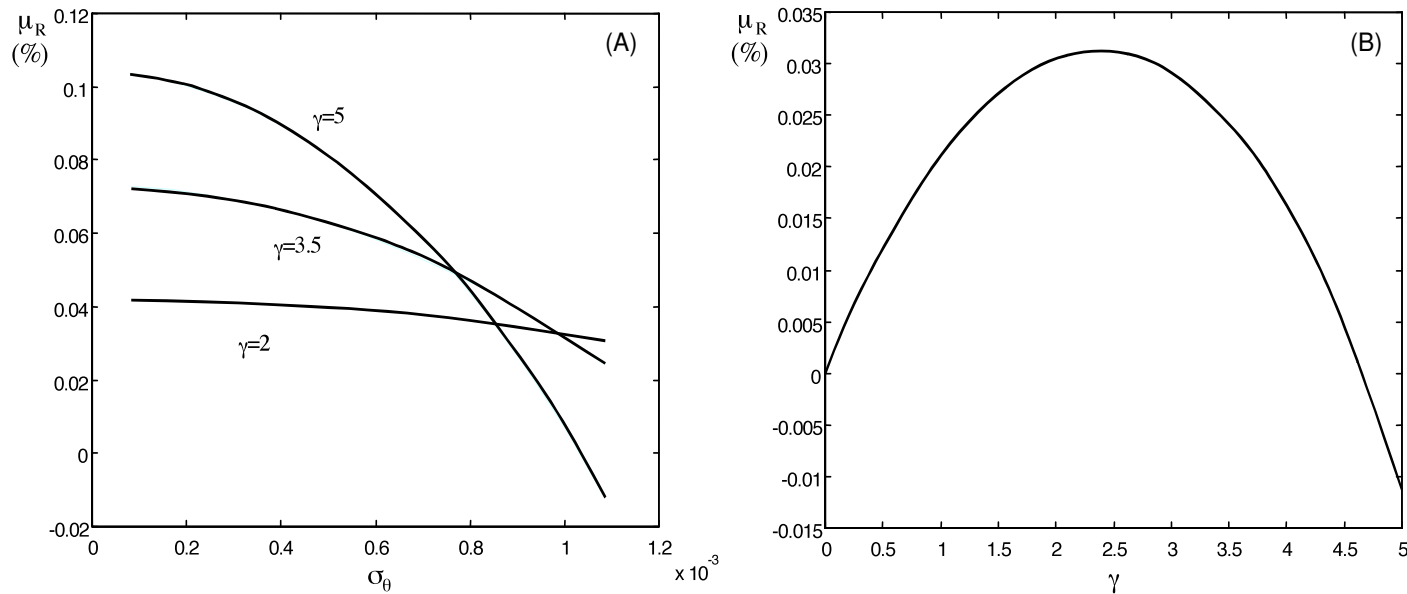


Figure 2. Expected returns and investors' uncertainty. (A) plots the conditional risk premium μ_R against the standard deviation of investors' beliefs $\sigma_\theta = \sqrt{\sum_{i=1}^n \pi_i (\theta_i - m_\theta)}$, which proxies for investors' uncertainty, for various coefficients of risk aversion. (B) plots the conditional risk premium μ_R against the coefficient of risk aversion γ for a level of $\sigma_\theta = 0.11\%$.

Economic Uncertainty and Return Volatility

- Return Volatility is given by

$$\sigma_R^2(\boldsymbol{\pi}_t) = \sigma_D^2 + V_\theta(\boldsymbol{\pi}_t) [2 + (h_e^2 + h_D^2)V_\theta(\boldsymbol{\pi}_t)] \quad (10)$$

- The following proposition then holds:

- **Proposition:**

1. σ_R is a U-shaped function of γ with $\sigma_R = \sigma_D$ for $\gamma = 1$.
2. a mean-preserving spread on $\boldsymbol{\pi}_t$ increases σ_R if $\sigma_R > \sigma_D$. The effect is ambiguous if $\sigma_R < \sigma_D$.
3. Under some conditions, if $h_e > h_D$ then $\sigma_R > \sigma_D$ for a coefficient of risk aversion sufficiently high.

Equity Premium and Return Volatility

- From the results about risk premium and return volatility, it is clear that the relationship between return volatility and expected returns is ambiguous and depends on the degree of investors' uncertainty.

- This statement can be made precise by noticing that we can write:

$$\mu_R(\boldsymbol{\pi}_t) = \gamma\sigma_R^2(\boldsymbol{\pi}_t) - \gamma V_\theta(\boldsymbol{\pi}_t) [1 + (h_e^2 + h_D^2)V_\theta(\boldsymbol{\pi}_t)] \quad (11)$$

- The second term in equation (11) can be positive or negative depending on the magnitude of $V_\theta(\boldsymbol{\pi}_t)$.
- Specifically, for log-utility or when signals are very precise, $V_\theta(\boldsymbol{\pi}_t)$ is approximately zero and hence a linear positive relationship results.
- In contrast, when $\gamma > 1$ and signals are not precise, the second term in equation (11) is positive for $-1/(h_e^2 + h_D^2) < V_\theta(\boldsymbol{\pi}_t) < 0$ and it is negative for $V_\theta(\boldsymbol{\pi}_t) < -1/(h_e^2 + h_D^2)$.
- Since the magnitude of V_θ changes over time due to investors' fluctuating level of uncertainty, equation (11) implies that there is no precise relationship between expected excess returns and conditional volatility.

Equity Premium and Return Volatility

- Higher uncertainty / volatility is related to lower expected return.

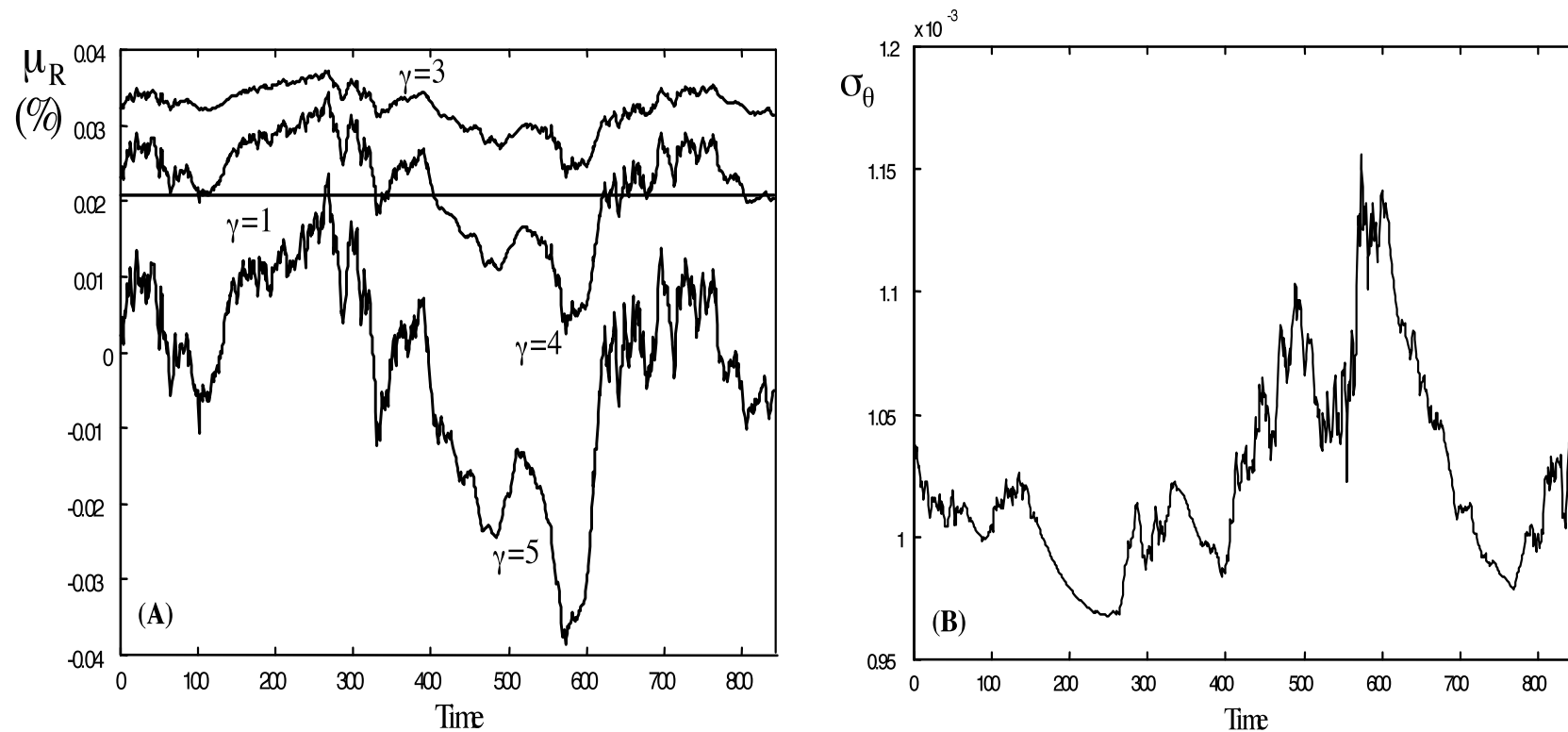


Figure 3. Expected returns and investors' uncertainty over time. (A) plots the conditional risk premium μ_R over time as the result of one Monte Carlo simulation of dividends and posterior distributions. μ_R is computed for the coefficients of risk aversion $\gamma = 1, 3, 4, 5$. (B) plots the standard deviation of investors' beliefs σ_θ across time.

Conclusion

- It is not obvious how economic uncertainty affects stock prices and stock returns.
- The result that P/D ratio increases with uncertainty is rather general.
- The result that expected return decreases with uncertainty crucially depends on $EIS < 1$.
 - There used to be an agreement among macroeconomists that $EIS < 1$.
 - Recent research and new estimates seem to suggest that $EIS > 1$ for stock holders.
 - In this case, expected return and uncertainty would be positively related, as intuition has it.

- What are the ways out?
 - Recent research:
 1. Habit persistence preferences
 2. Risk for the long run
 3. Preferences for robustness

- How far can we go?
- What about the cross-section?

External Habit: Campbell and Cochrane (1999).

- Constantinides (1990), Detemple and Zapatero (1991), Campbell and Cochrane (1999) and others use the following representation of habit preferences.

- The representative agent maximizes

$$E \left[\int_0^{\infty} u(C_t, X_t, t) dt \right], \quad (12)$$

- where the instantaneous utility function is give by

$$u(C_t, X_t, t) = \begin{cases} e^{-\rho t} \frac{(C_t - X_t)^{1-\gamma}}{1-\gamma} & \text{if } \gamma > 1 \\ e^{-\rho t} \log(C_t - X_t) & \text{if } \gamma = 1 \end{cases} \quad (13)$$

- X_t is a habit level.
- Campbell and Cochrane (1999) consider X_t as an *external* habit, rather than internal.
 - Catching up with the Joneses.

A Problem

- The model above is not homogeneous, and thus it is hard to work with.
- One additional problem is that in endowment economies, for standard specification of X_t we cannot guarantee

$$C_t > X_t$$

- For instance, assume that consumption (endowment) follows the geometric Brownian motion

$$\frac{dC_t}{C_t} = \mu_c dt + \sigma_c dB_t$$

- Assume also X_t is just a weighted average of past consumption:

$$X_t = X_0 e^{-\alpha t} + \alpha \int_0^t e^{-\alpha(\tau-t)} C_\tau d\tau$$

- Ito's Lemma implies

$$dX = \alpha (C_t - X_t) dt$$

The Surplus Consumption Ratio Dynamics

- Consider now the following quantity

$$S_t = \frac{C_t - X_t}{C_t} \quad (14)$$

- Campbell and Cochrane (1999) call S_t the *Surplus Consumption Ratio*.
- Clearly, we must have $S_t \in [0, 1]$
- However, if we apply Ito's Lemma to S_t we find

$$dS = k (\bar{S} - S_t) dt + \lambda(S_t) dB_t \quad (15)$$

- where

$$k = \mu_c - \alpha - \sigma_c^2; \quad \bar{S} = (\mu_c - \sigma_c^2) / k; \quad \lambda(S_t) = (1 - S_t)$$

- Note that S_t is:
 - Mean reverting: This is a consequence of habit formation and the fact that X_t is slow moving.
 - Perfectly correlated with innovations to consumption growth, given by dB_t .
 - The volatility of surplus is time varying.

The Surplus Consumption Ratio Dynamics

- Note also that S_t is bounded above:
 - when S_t reaches 1, the diffusion disappears, and the drift is negative.
 - Thus, S_t is dragged down.
- However, nothing stops S_t from going below zero:
 - When $S_t = 0$, the diffusion is still positive (and in fact large).
 - Although the drift is also positive under the sensible assumption that $\mu_c > \sigma_c^2$, there is a non-zero probability that $S_t \leq 0$.
- This event of course is inconsistent with the preference specification.

Campbell and Cochrane Solution

- Campbell and Cochrane (1999) had a great intuition:
 - Specify the mean reverting dynamics for *log* surplus $s_t = \log(S_t)$
 - Specify $\lambda(s_t)$ in a way to ensure $S_t = \exp(s_t) \in [0, 1]$.
- In addition, they specified $\lambda(s_t)$ to obtain specific properties of the interest rate process r_t (e.g. constant!)
- Unfortunately, their specification does not yield closed form solutions for prices.
- I therefore follow Santos and Veronesi (2010), which generalizes the setting in Menzly, Santos, Veronesi (2004) to the power utility case.
- Consider first the stochastic discount factor implied by the habit model

$$\pi_t = e^{-\rho t} \frac{\partial u(C_t, X_t)}{\partial C_t} = e^{-\rho t} (C_t - X_t)^{-\gamma} = e^{-\rho t} C_t^{-\gamma} S_t^{-\gamma}$$

- The surplus consumption ratio acts as a “preference shock”, as it changes the curvature of the utility function: γS_t^{-1} .

Santos and Veronesi (2010) model

- Since S_t is naturally mean reverting, Campbell and Cochrane (1999) consider the particular monotonic transformation $s_t = \log(S_t)$ and model s_t as mean reverting.
- Santos and Veronesi (2010) use a different monotonic transformation, namely

$$G_t = S_t^{-\gamma} \tag{16}$$

- Assume then that G_t is mean reverting

$$dG_t = k (\bar{G} - G_t) dt - \alpha (G_t - \lambda) \sigma_c dB_t \tag{17}$$

- Note the following:

1. G_t is mean reverting, like S_t .
2. G_t is *negatively* perfectly correlated with innovations to consumption dB_t .
3. G_t is bounded below by $\lambda > 1$. That is, we restrict $C_t > X_t$ at all times.

- These are the same properties of Campbell and Cochrane (1999).

Interest Rate in SV model

- Since $\pi_t = e^{-\rho t} C_t^{-\gamma} G_t$, the SDF is given by

$$\frac{d\pi_t}{\pi_t} = -r_t^f dt - \sigma_\pi dB_t,$$

- The risk free rate is given by

$$r_t^f = \rho + \gamma\mu_c - \frac{1}{2}\gamma(\gamma+1)\sigma_c^2 + k(1 - \overline{G}S^\gamma) - \gamma\alpha(1 - \lambda S^\gamma)\sigma_c^2 \quad (18)$$

- **Comments:**

1. The first three terms in r_t are standard.
2. The fourth term $k(1 - \overline{G}S^\gamma)$ represents the intertemporal substitution effect

Low $S_t \rightarrow$ high expected S_τ in future \rightarrow Borrow to increase $C_t \rightarrow r_t$ high

3. The last term $-\gamma\alpha(1 - \lambda S^\gamma)$ represents an additional precautionary savings term:

Low $S_t \rightarrow$ higher probability $C_\tau = X_\tau$ in the future \rightarrow Save more today $\rightarrow r_t$ low

4. Campbell and Cochrane (1999) choose parameters so that these two effects cancel each other \rightarrow constant r_t

The Market Price of Risk in SV model

- The volatility of the stochastic discount factor is given by

$$\sigma_{\pi} = [\gamma + \alpha(1 - \lambda S_t^{\gamma})] \sigma_c. \quad (19)$$

- σ_{π} now depends on S_t^{γ}

Low S_t \rightarrow higher curvature of the utility function $\gamma S_t^{-1} \rightarrow$
 \rightarrow Higher aversion to risk \rightarrow Higher price of risk

Stock Prices in SV model

- Coming to the stock price of a consumption claim, we have

$$P_t = E_t \left[\int_t^\infty \left(\frac{\pi_\tau}{\pi_t} \right) C_\tau d\tau \right] \quad (20)$$

- Substituting, we obtain

$$P_t = C_t^\gamma S^\gamma E_t \left[\int_t^\infty e^{-\rho(\tau-t)} C_\tau^{1-\gamma} G_\tau d\tau \right] \quad (21)$$

- The solution is in closed form

$$P_t = C_t (b_1 + b_2 S_t^\gamma)$$

- where

$$b_1 = \frac{1}{\alpha_1}; \quad b_2 = \frac{k\bar{G} + \alpha(1-\gamma)\lambda\sigma_c^2}{\alpha_1\alpha_2}$$

with

$$\alpha_1 = \rho - (1-\gamma)\mu_c + \frac{1}{2}(1-\gamma)\gamma\sigma_c^2 + k + \alpha(1-\gamma)\sigma_c^2$$

$$\alpha_2 = \rho - (1-\gamma)\mu_c + \frac{1}{2}(1-\gamma)\gamma\sigma_c^2$$

Properties of Stock Prices in SV model

- The implications for

$$P_t/C_t = b_1 + b_2 S_t^\gamma$$

is straightforward:

- a higher surplus consumption ratio S_t translates in lower risk preference, and thus a higher price.
- Intertemporal smoothing hits here too.
 - From the form of b_1 and b_2 , a high consumption growth μ_c translates into a lower P/C ratio, as we saw with learning.
 - Therefore, learning about μ_c , for instance, will generate the same problem it did for the standard power utility case.
 - * Higher uncertainty \implies lower equity premium

Return Volatility and Equity Premium

- What about the volatility and the equity premium?
- By using Ito's Lemma, we have

$$E_t [dR_t] = (\gamma + \alpha (1 - \lambda S_t^\gamma)) \sigma_R(S_t) \sigma_c$$

$$\sigma_R(S_t) = \left[1 + \frac{b_2 S_t^\gamma (1 - \lambda S_t^\gamma) \alpha}{b_1 + b_2 S_t^\gamma} \right] \sigma_c.$$

- How does this model performs?
- The following are some statistics of the market portfolio:

Basic Moments to Explain

Table I
Basic moments

Panel A: Summary statistics for the market portfolio

$E(R^M)$	$\text{vol}(R^M)$	r^f	$\text{vol}(r^f)$
7.71%	16.25%	1.44%	3.08%

Panel B: Predictability regressions

	Panel B-1: Sample 1948-2001				Panel B-2: Sample 1948-1995			
Horizon	4	8	12	16	4	8	12	16
$\ln\left(\frac{D}{P}\right)$.13	.2	.26	.35	.28	.48	.63	.78
t-stat.	(2.13)	(1.65)	(1.34)	(1.29)	(4.04)	(4.00)	(4.49)	(5.41)
R^2	.09	.10	.11	.14	.19	.32	.43	.54

A Calibration

- A simple calibration of the economy (not much parameter search here) is as follows:

Table III
Model parameters used in the simulation

Panel A: Consumption and preference parameters

μ_c	σ_c	γ	ρ	γ/\bar{S}	$\min\{\gamma/S_t\}$	α	k
.02	.015	1.5	.072	48	27.75	77	.13

Model Implications in Simulations

Table IV

Basic moments in simulated data

Panel A: Summary statistics for the aggregate portfolio

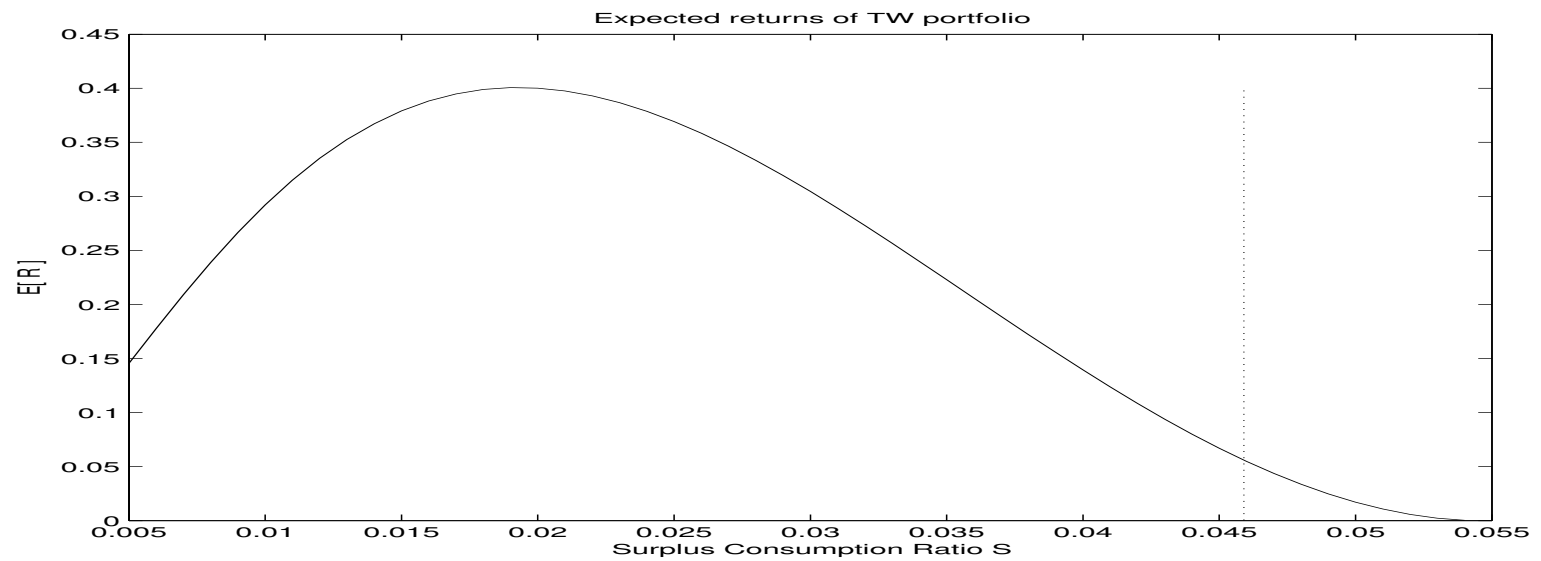
$E(R^M)$	$\text{vol}(R^M)$	r^f	$\text{vol}(r^f)$
9.96%	24.15%	.91%	5.41%

Panel B: Predictability regressions

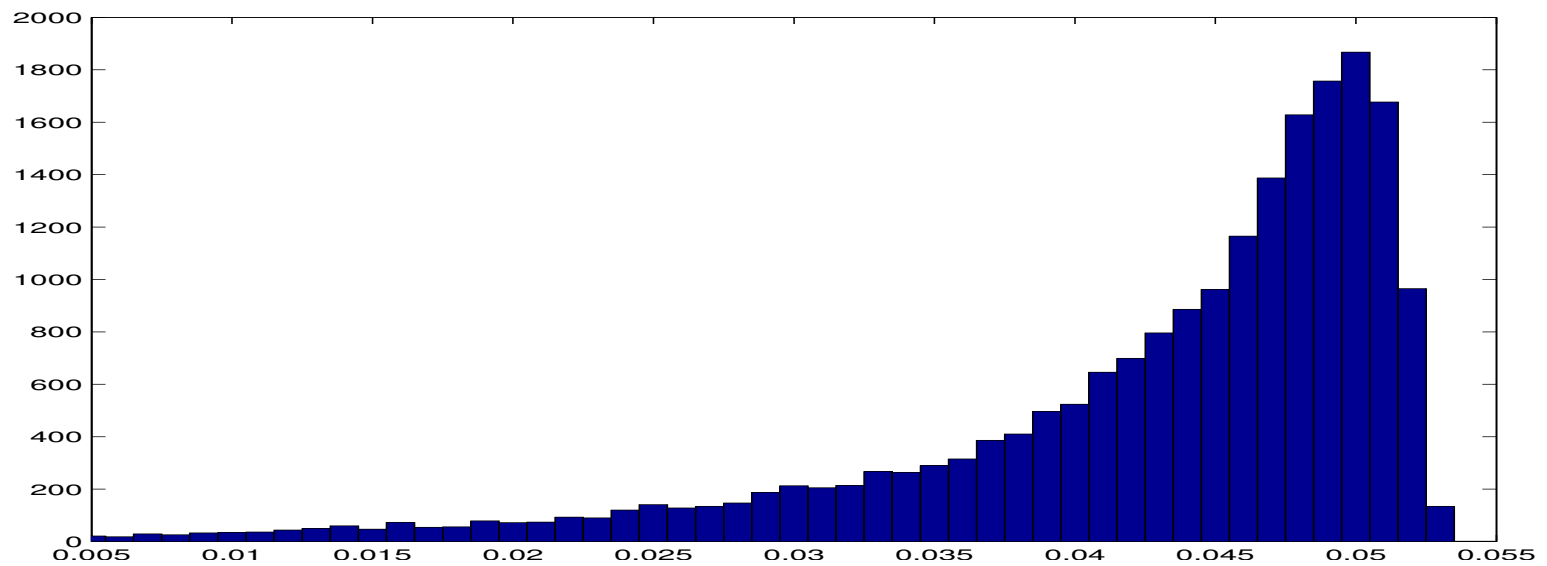
Horizon	4	8	12	16
$\ln\left(\frac{D}{P}\right)$.73	.86	.88	.85
R^2	.25	.30	.29	.27

- The model does well, although the volatility of interest rates is a little too high
- The following figures shows the source of the effects

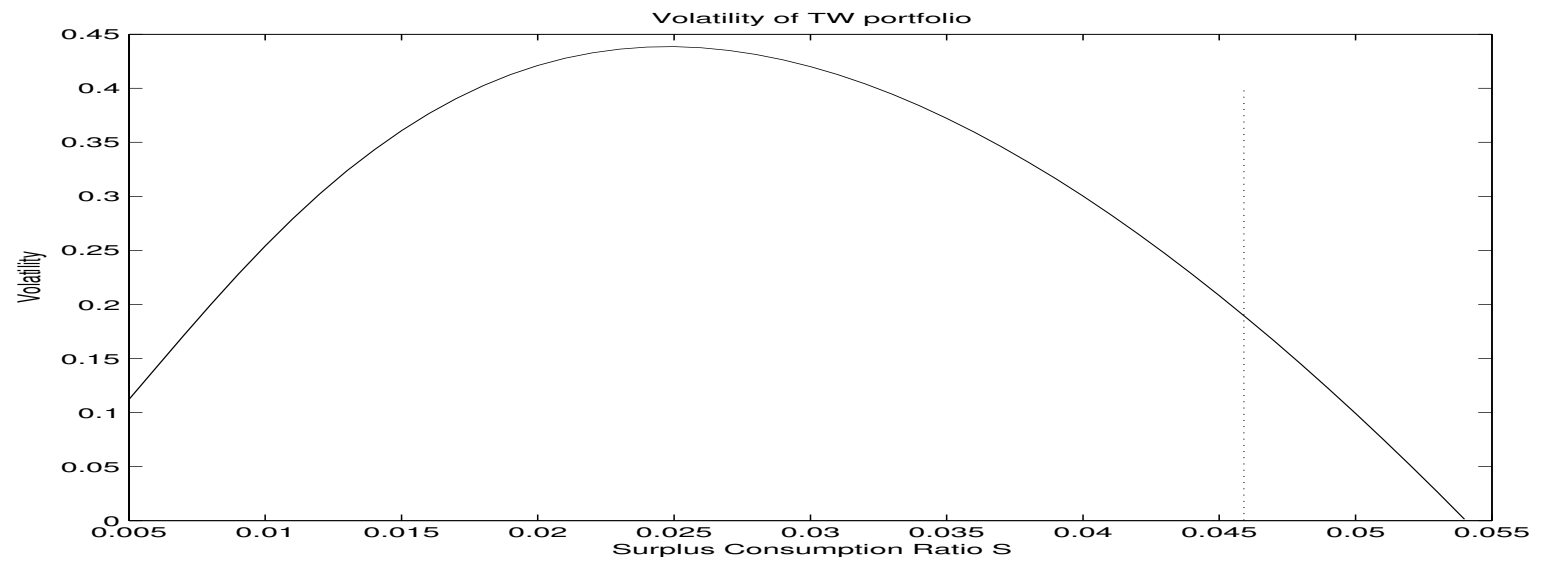
Expected Return



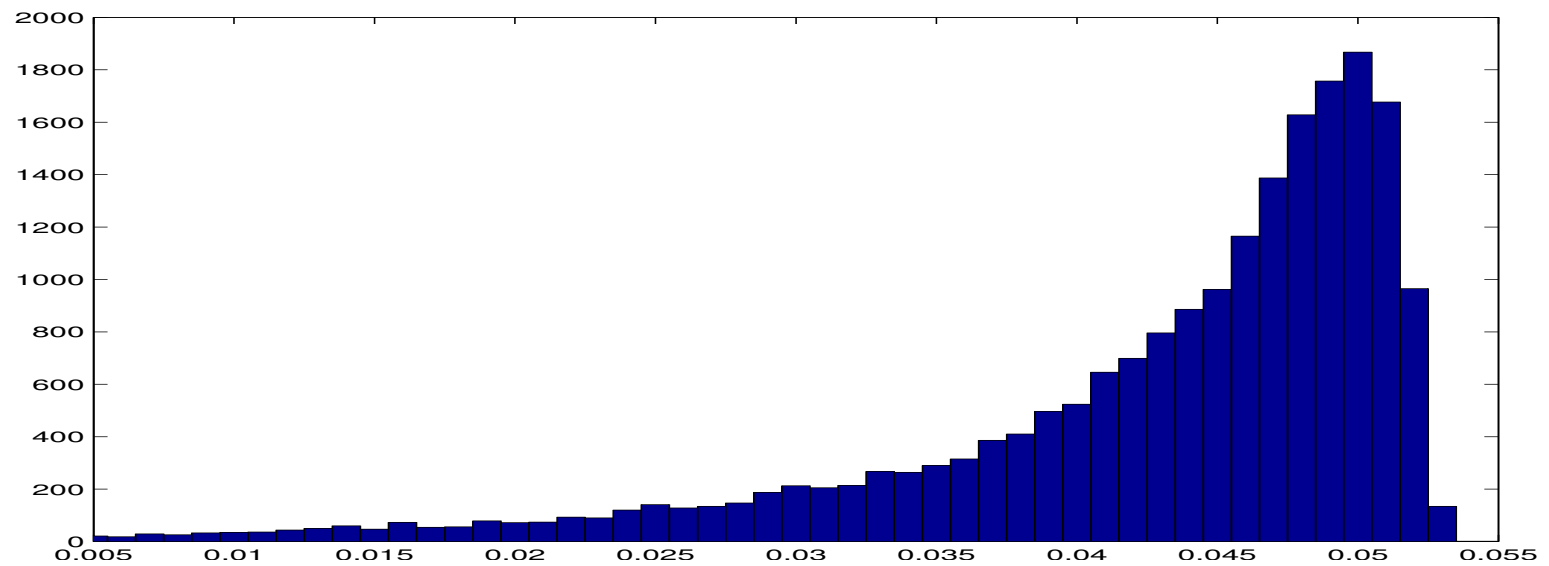
Distribution of S_t



Conditional Volatility



Distribution of S_t



Habit Preferences, Uncertainty and the 1990s “Bubble”

- Pastor and Veronesi (2006) use a similar setting with uncertainty about average profitability to rationalize the high valuations in the late 1990s



Why Was Uncertainty about Average Profitability High in the Late 1990s?

- Technological revolution; “new era”
- Stock price reaction to earnings surprises was stronger in 1996-1999 than in 1992-1995 (Ahmed, Schneible, and Stevens, 2003)
- Firms went public earlier in their life-cycles (Schultz and Zaman, 2001)
- Tech stock profitability was highly volatile
- Tech stock prices were highly volatile
- Anecdotal evidence

“...the projections of revenue growth were, by and large, wild guesses.”

Investment Dealers Digest, 23 October 2000.

“Internet firms’ highly unpredictable growth rates make historical information less useful.”

TIAA-CREF Investment Forum, March 2001.

“...being wrong isn’t very costly, and being right has a high payoff... With Amazon, we believe the payoff for being right is high.”

Bill Miller, portfolio manager of the Legg Mason Value Trust, in Barron’s, 15 Nov 1999.

Habit Preferences in PV

- PV use a different transformation of surplus consumption ratio

$$S_t = e^{s_t}$$

$$s_t = a_0 + a_1 y_t + a_2 y_t^2$$

$$dy_t = k_y (\bar{y} - y_t) dt + \sigma_y dW_{0,t}$$

- Choosing a_i appropriately (in particular, $a_2 < 0$) $\implies s_t < 0 \implies S_t \in [0, 1]$.
- The SDF is then given by

$$\pi_t = e^{-\eta t - \gamma(s_t + c_t)},$$

- with dynamics

$$\frac{d\pi_t}{\pi_t} = -r_t dt - \sigma_{\pi,t} dW_{0,t}$$

$$\sigma_{\pi,t} = \gamma (\sigma_\varepsilon + (a_1 + 2a_2 y_t) \sigma_y)$$

– when $y_t \uparrow \implies \sigma_{\pi,t} \downarrow \implies$ equity premium \downarrow

Model: Profitability

- Firm profitability is measured as the accounting return on equity (ROE),

$$\rho_t = \frac{\text{Earnings}_t}{\text{Book Equity}_t} = \frac{Y_t}{B_t}$$

- Mean reversion in firm profitability:

$$d\rho_t^i = \phi^i \underbrace{(\bar{\rho}_t + \bar{\psi}_t^i - \rho_t^i)}_{\bar{\rho}_t^i} dt + \sigma_{i,0} \underbrace{dW_{0,t}}_{\text{systematic}} + \sigma_{i,i} \underbrace{dW_{i,t}}_{\text{idiosyncratic}}$$

– $\bar{\rho}_t$... Average aggregate profitability

$$d\bar{\rho}_t = k_L (\bar{\rho}_L - \bar{\rho}_t) dt + \sigma_{L,0} dW_{0,t} + \sigma_{L,L} dW_{L,t}$$

– $\bar{\psi}_t^i$... Average firm-specific excess profitability

$$d\bar{\psi}_t^i = -k_\psi \bar{\psi}_t^i dt$$

Model: Dividends

- Dividends are proportional to book equity:

$$D_t = c B_t, \quad c \geq 0$$

- Clean surplus relation (assuming no new equity issues/withdrawals):

$$dB_t = (Y_t - D_t) dt = (\rho_t - c) B_t dt$$

Note: Since $\frac{dB_t}{B_t} = (\rho_t - c) dt$,

uncertainty about average ρ_t = uncertainty about the average growth rate of B_t

Model: Market Value

- The abnormal earnings model of Ohlson (1990, 1995):

$$M_t = B_t + \text{present value of future abnormal earnings}$$

- Competition $\implies M_T = B_T$ at some future time T
 - T is random, exponentially distributed with density $h(T; p)$
 - At any point in time, there is probability p that T arrives in the next instant

- Market value of equity:

$$M_t = E_t \left[\int_t^\infty \left(\int_t^T \frac{\pi_s}{\pi_t} D_s ds + \frac{\pi_T}{\pi_t} B_T \right) h(T; p) dT \right],$$

where π_t is given earlier

Valuation Formula

- **Proposition 1.** Suppose $\bar{\psi}_t^i$ is known.

$$\frac{M_t}{B_t} = G\left(\bar{\psi}_t^i; y_t, \bar{\rho}_t, \rho_t^i\right) = (c + p) \int_0^\infty Z\left(y_t, \bar{\rho}_t, \rho_t^i, \bar{\psi}_t^i, s\right) ds$$

- When $\bar{\psi}_t^i$ is unknown, the law of iterated expectations yields

$$\frac{M_t}{B_t} = E\left[G\left(\bar{\psi}_t^i; y_t, \bar{\rho}_t, \rho_t^i\right)\right] = \int G\left(\bar{\psi}_t^i; y_t, \bar{\rho}_t, \rho_t^i\right) f_t\left(\bar{\psi}_t^i\right) d\bar{\psi}_t^i$$

Note: Since G is convex in $\bar{\psi}_t^i$, greater dispersion in $f_t\left(\bar{\psi}_t^i\right)$ increases M/B

- **Proposition 2.** Suppose that $\bar{\psi}_t^i$ is unknown, and that $f_t\left(\bar{\psi}_t^i\right) = N\left(\hat{\psi}_t^i, \hat{\sigma}_{i,t}^2\right)$.

$$\frac{M_t}{B_t} = (c + p) \int_0^\infty Z\left(y_t, \bar{\rho}_t, \rho_t^i, \hat{\psi}_t^i, s\right) e^{\frac{1}{2}Q_4(s)^2 \hat{\sigma}_{i,t}^2} ds$$

- Note: M/B increases if

- (i) expected profitability increases ($\hat{\psi}_t^i \uparrow, \bar{\rho}_t \uparrow, \rho_t^i \uparrow$)
- (ii) the discount rate decreases ($y_t \uparrow \Rightarrow$ equity premium \downarrow)
- (iii) uncertainty about $\bar{\psi}_t^i$ increases ($\hat{\sigma}_{i,t} \uparrow$)

Calibration

- Two sectors:

- “New economy” (Nasdaq): described above
- “Old economy” (NYSE/Amex): pays dividends $D_t^O = c^O B_t^O$ forever

Old economy’s market value is $M_t^O = E_t \left[\int_t^\infty \frac{\pi_s}{\pi_t} D_s^O ds \right]$, we derive $M_t^O / B_t^O = \Phi(\bar{\rho}_t, y_t)$

- The profitability parameters are estimated from the data
- The SDF parameters are calibrated to match the old economy’s average return, volatility, M/B, and the level of interest rate to their empirical counterparts

Table 1

Old Economy Profitability				New Economy Profitability			Individ. Firm Profitability		
k_L	$\bar{\rho}_L$	σ_{LL}	$\sigma_{L,0}$	ϕ^N	$\sigma_{0,N}$	$\sigma_{N,N}$	ϕ^i	$\sigma_{i,0}$	$\sigma_{i,i}$
0.3574	12.17%	1.47%	1.31%	0.3551	2.93%	4.88%	0.3891	6.65%	8.07%
Stochastic Discount Factor									
η	γ	k_y	\bar{y}	σ_y	a_0	a_1	a_2	μ_ε	σ_ε
0.0471	3.9474	0.0367	-0.08%	25.30%	-2.8780	0.3084	-0.0413	2%	1%
Means of Fitted Quantities				Std Deviations of Fitted Quantities					
$E[M/B]$	$E[\mu_{R,t}^{mkt}]$	$E[\sigma_{R,t}^{mkt}]$	$E[r_{f,t}]$	$\sigma[M/B]$	$\sigma[\mu_{R,t}^{mkt}]$	$\sigma[\sigma_{R,t}^{mkt}]$	$\sigma[r_{f,t}]$	c^O	k_ψ
1.77	5.06%	14.47%	6.25%	0.6477	1.72%	2.24%	1.55%	5.67%	0.0139

Table 2. Nasdaq's Valuation on March 10, 2000 Assuming Zero Uncertainty

$\rho_t^N = 9.96\%$ per year, $c = 1.35\%$ per year, $E(T) = 20$ years.

Excess ROE	Equity Premium (% per year)							
$\hat{\psi}^N$ (% per year)	1	2	3	4	5	6	7	8
Panel A: Model-implied M/B with zero uncertainty (Actual M/B: 8.55)								
0	3.33	3.02	2.63	2.23	1.84	1.47	1.12	0.76
1	4.15	3.70	3.17	2.64	2.14	1.68	1.25	0.83
2	5.27	4.62	3.89	3.19	2.53	1.95	1.41	0.90
3	6.83	5.89	4.87	3.92	3.05	2.29	1.62	1.00
4	9.06	7.68	6.23	4.92	3.75	2.74	1.88	1.11
5	12.28	10.22	8.15	6.31	4.71	3.36	2.23	1.26
6	17.02	13.92	10.90	8.28	6.04	4.19	2.69	1.45
7	24.09	19.38	14.91	11.12	7.93	5.36	3.32	1.69
Panel B: Implied return volatility with zero uncertainty (Actual volatility: 41.5% in March 2000, 47% in 2000)								
0	18.09	20.17	21.76	22.93	23.76	24.18	24.10	23.04
1	18.69	20.93	22.65	23.92	24.83	25.31	25.22	24.05
2	19.31	21.71	23.57	24.97	25.97	26.52	26.46	25.18
3	19.93	22.50	24.52	26.05	27.18	27.83	27.81	26.45
4	20.54	23.30	25.47	27.16	28.44	29.21	29.27	27.85
5	21.14	24.07	26.42	28.27	29.72	30.66	30.85	29.42
6	21.71	24.82	27.34	29.37	31.01	32.15	32.52	31.15
7	22.25	25.53	28.23	30.44	32.28	33.65	34.27	33.05

Table 3. Nasdaq's Valuation on March 10, 2000 Assuming Uncertainty of 3% Per Year

$\rho_t^N = 9.96\%$ per year, $c = 1.35\%$ per year, $E(T) = 20$ years.

Excess ROE	Equity Premium (% per year)							
$\hat{\psi}^N$ (% per year)	1	2	3	4	5	6	7	8
Panel A: Model-implied M/B with 3% uncertainty (Actual M/B: 8.55)								
0	4.70	4.09	3.43	2.81	2.23	1.72	1.26	0.82
1	6.16	5.27	4.33	3.47	2.69	2.02	1.44	0.90
2	8.29	6.95	5.59	4.39	3.33	2.43	1.67	1.00
3	11.44	9.40	7.41	5.69	4.21	2.99	1.98	1.13
4	16.17	13.03	10.07	7.57	5.46	3.76	2.40	1.30
5	23.39	18.53	14.05	10.35	7.29	4.87	2.99	1.52
6	34.59	26.96	20.10	14.53	9.99	6.49	3.83	1.82
7	52.23	40.10	29.44	20.91	14.08	8.89	5.04	2.24
Panel B: Implied return volatility with 3% uncertainty (Actual volatility: 41.5% in March 2000, 47% in 2000)								
0	33.66	33.93	33.97	33.76	33.22	32.25	30.62	27.60
1	35.63	35.96	36.06	35.89	35.37	34.36	32.60	29.23
2	37.59	38.02	38.21	38.12	37.66	36.66	34.81	31.11
3	39.51	40.05	40.37	40.41	40.05	39.13	37.26	33.25
4	41.33	42.03	42.50	42.70	42.51	41.73	39.93	35.70
5	43.05	43.91	44.56	44.94	44.96	44.40	42.77	38.46
6	44.65	45.67	46.50	47.09	47.35	47.07	45.73	41.52
7	46.11	47.29	48.31	49.10	49.63	49.67	48.72	44.85

Table 4. Matching Nasdaq's Valuation on March 10, 2000

$\rho_t^N = 9.96\%$ per year, $c = 1.35\%$ per year, $E(T) = 20$ years.

Excess ROE $\hat{\psi}^N$ (% per year)	Equity Premium (% per year)							
	1	2	3	4	5	6	7	8
Panel A. Uncertainty needed to match the observed M/B								
0	4.39	4.71	5.06	5.43	5.81	6.22	6.67	7.27
1	3.81	4.17	4.59	5.01	5.44	5.89	6.38	7.03
2	3.08	3.54	4.04	4.53	5.03	5.54	6.08	6.77
3	2.08	2.73	3.38	3.98	4.57	5.15	5.75	6.50
4	0.00	1.45	2.51	3.32	4.04	4.71	5.39	6.22
5	0.00	0.00	0.97	2.43	3.40	4.22	5.00	5.91
6	0.00	0.00	0.00	0.78	2.56	3.63	4.56	5.58
7	0.00	0.00	0.00	0.00	1.18	2.90	4.06	5.23
Panel B. Return volatility under implied uncertainty (Actual volatility: 41.5% in March 2000, 47% in 2000)								
0	64.69	73.70	85.81	100.56	119.11	142.51	173.80	223.37
1	51.35	60.15	71.80	85.98	103.93	126.70	157.49	206.85
2	38.94	47.54	58.69	72.23	89.41	111.45	141.54	190.50
3	27.79	36.12	46.66	59.43	75.73	96.84	126.12	174.41
4	20.54	26.53	36.07	47.81	63.03	83.03	111.22	158.63
5	21.14	24.07	27.70	37.78	51.53	70.16	96.98	143.20
6	21.71	24.82	27.34	30.14	41.66	58.44	83.59	128.22
7	22.25	25.53	28.23	30.44	34.09	48.24	71.18	113.79

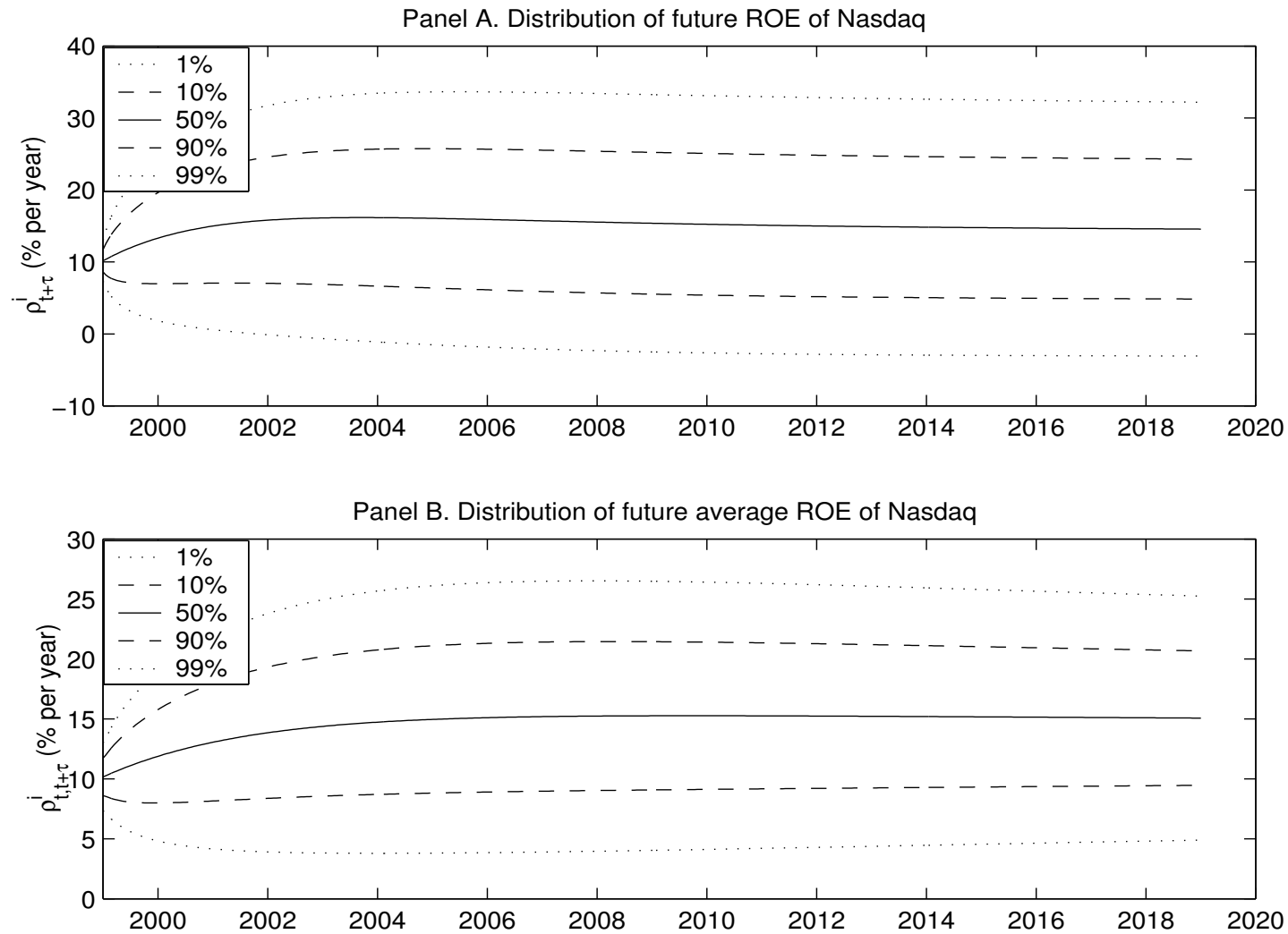
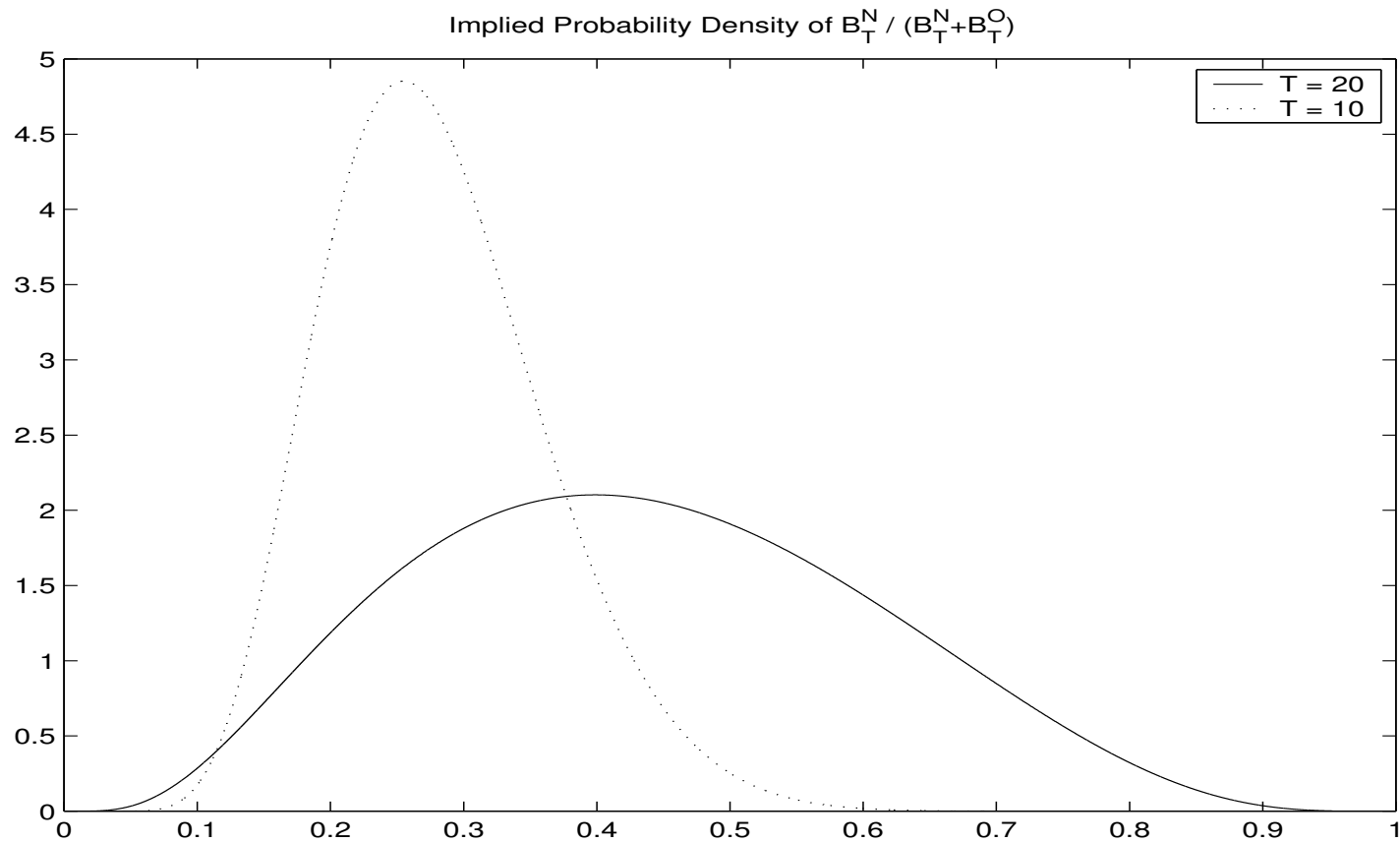


Figure 1. Model-predicted distributions of future profitability and average future profitability for Nasdaq.



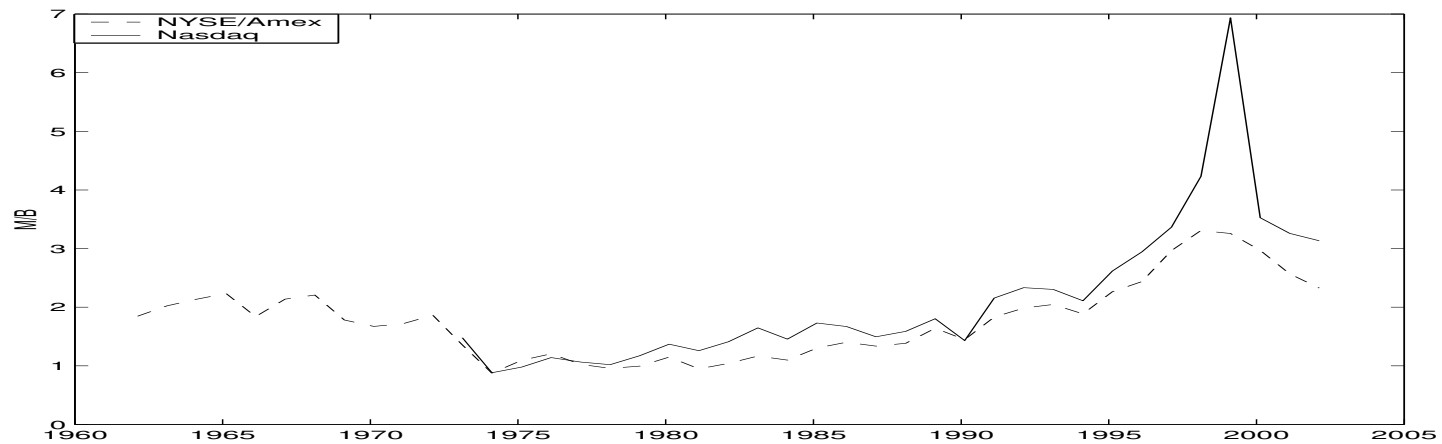
Model-predicted distribution of the future ratio of Nasdaq book value to NYSE/Amex/Nasdaq book value.

T	Percentile								
	1	5	10	25	50	75	90	95	99
10	0.12	0.16	0.18	0.22	0.27	0.33	0.39	0.43	0.50
20	0.11	0.17	0.22	0.31	0.43	0.56	0.67	0.73	0.82

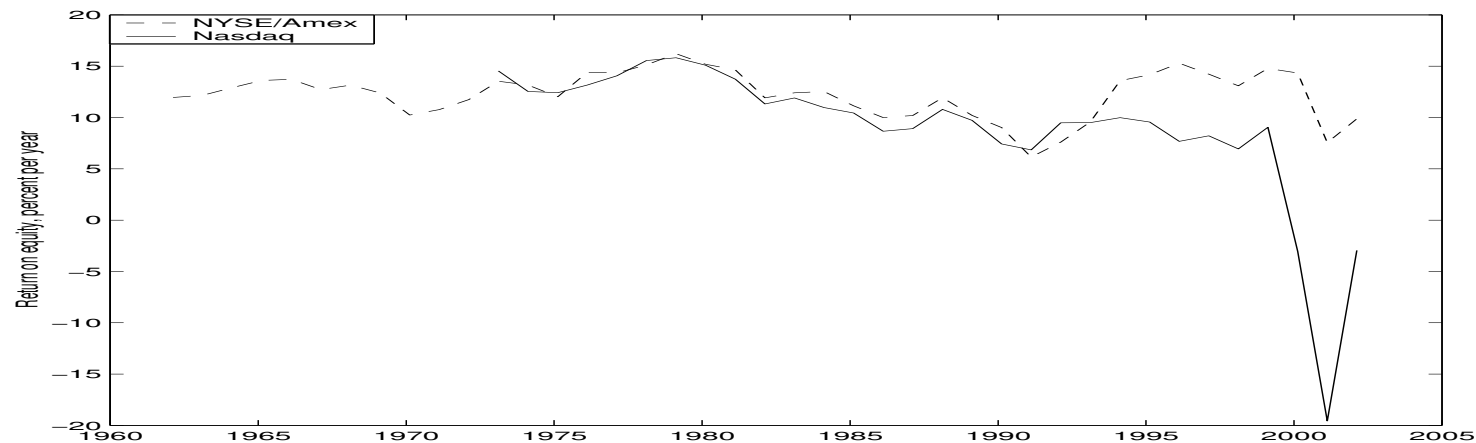
Why Did the "Bubble" Burst?

- There is little doubt of what caused tech stock prices to drop in 2000.
 - Nasdaq's profitability plummeted in 2000.

Nasdaq and NYSE/Amex M/B ratios



Nasdaq and NYSE/Amex profitability



Why Did the “Bubble” Burst?

- Is this large drop consistent with our model?
 - Yes
 - * A high uncertainty about long term profitability implies large revisions when there are large unexpected events.
 - * Our model implies a similar drop in M/B in 2000, and an even larger drop in 2001.

- Return volatility did not move much after March 2000.
 - This is also consistent with our model: Uncertainty remained high even after March 2000.

Technological Revolutions and Asset Prices

- The 1990s tech revolution and tech “bubble” was just the last example of a pattern repeated several times in history.

“Technological revolutions and financial bubbles seem to go hand in hand.”

*“Every previous technological revolution has created a speculative bubble...
With each wave of technology, share prices soared and later fell...”*

(The Economist, September 21, 2000)

- Stock prices tend to exhibit bubble-like patterns during technological revolutions
 - Prices rise and then fall, especially for innovative firms
 - Return volatility is high, especially for innovative firms
- Examples:
 - the early 1980s (biotechnology, PC)
 - the early 1960s (electronics)
 - the 1920s (electricity, automobiles)
 - the early 1900s (radio)

Repeated Irrational Exuberance?

- The bubble-like stock price behavior is commonly attributed to irrationality (e.g., Shiller, 2000, Perez, 2002, popular press)
 - Investors get too excited about the new technology
- We propose a rational explanation
 - Time-varying nature of uncertainty about the new technology

Pastor and Veronesi (2009, AER)

- New technologies have **high uncertainty** about average future productivity
 - This uncertainty makes returns highly volatile
- Initially, this uncertainty is mostly **idiosyncratic**
 - Because the new technology is initially developed on a small scale
 - The idiosyncratic uncertainty increases stock prices (PV 2003, 2006)
- In *technological revolutions*, new technologies are widely adopted
- For those technologies that are eventually adopted by the whole economy, the uncertainty gradually changes from idiosyncratic to **systematic**
 - As a result, discount rates rise and stock prices fall
- The “bubble” in prices is observable *ex post* but unpredictable *ex ante*
 - **Ex post selection bias**: We know *ex post* that a technological revolution took place, but investors did not know that *ex ante*

Outline of the Model

- We develop a general equilibrium model with a representative agent
- Two sectors: the “new economy” and the “old economy”
 - **Old economy:** *Large-scale* production using *old* technology
 - * Affects the representative agent’s wealth
 - **New economy:** *Small-scale* production using *new* technology
 - * Does not affect the representative agent’s wealth
- The representative agent (the social planner)
 1. Sets up the new economy to “experiment” with the new technology
 2. Learns about the average productivity of the new technology
 3. Decides whether/when to adopt the new technology on a large scale
- If the technology is adopted, we call this a **technological revolution**

Preferences and Technology

- Representative agent has utility from final wealth, $u(W_T) = \frac{W_T^{1-\gamma}}{1-\gamma}$, with $\gamma > 1$
- The agent is endowed with capital B_0 at time $t = 0$
- Capital produces output $Y_t = \rho_t B_t$, and follows $dB_t = Y_t dt = \rho_t B_t dt$
- Market clearing: $W_T = B_T$

- Productivity ρ_t follows a mean-reverting process:

$$d\rho_t = \phi(\bar{\rho} - \rho_t) dt + \sigma dZ_{0,t}, \quad (\text{under old technology})$$

$$d\rho_t = \phi(\bar{\rho} + \psi - \rho_t) dt + \sigma dZ_{0,t}, \quad (\text{under new technology})$$

- The “productivity gain” ψ is *unobservable*
 - When the new technology arrives at time t^* , ψ is drawn as normal:

$$\psi \sim N(0, \hat{\sigma}_{t^*}^2)$$

- After time t^* , the agent learns about ψ in a Bayesian fashion

Learning in the New Economy

- The agent learns about ψ by observing productivity in the new economy
- Capital used in the new economy, B_t^N , is infinitely smaller than B_t
 - \Rightarrow the new technology affects W_T only if adopted by the old economy

- Capital in the new economy evolves as

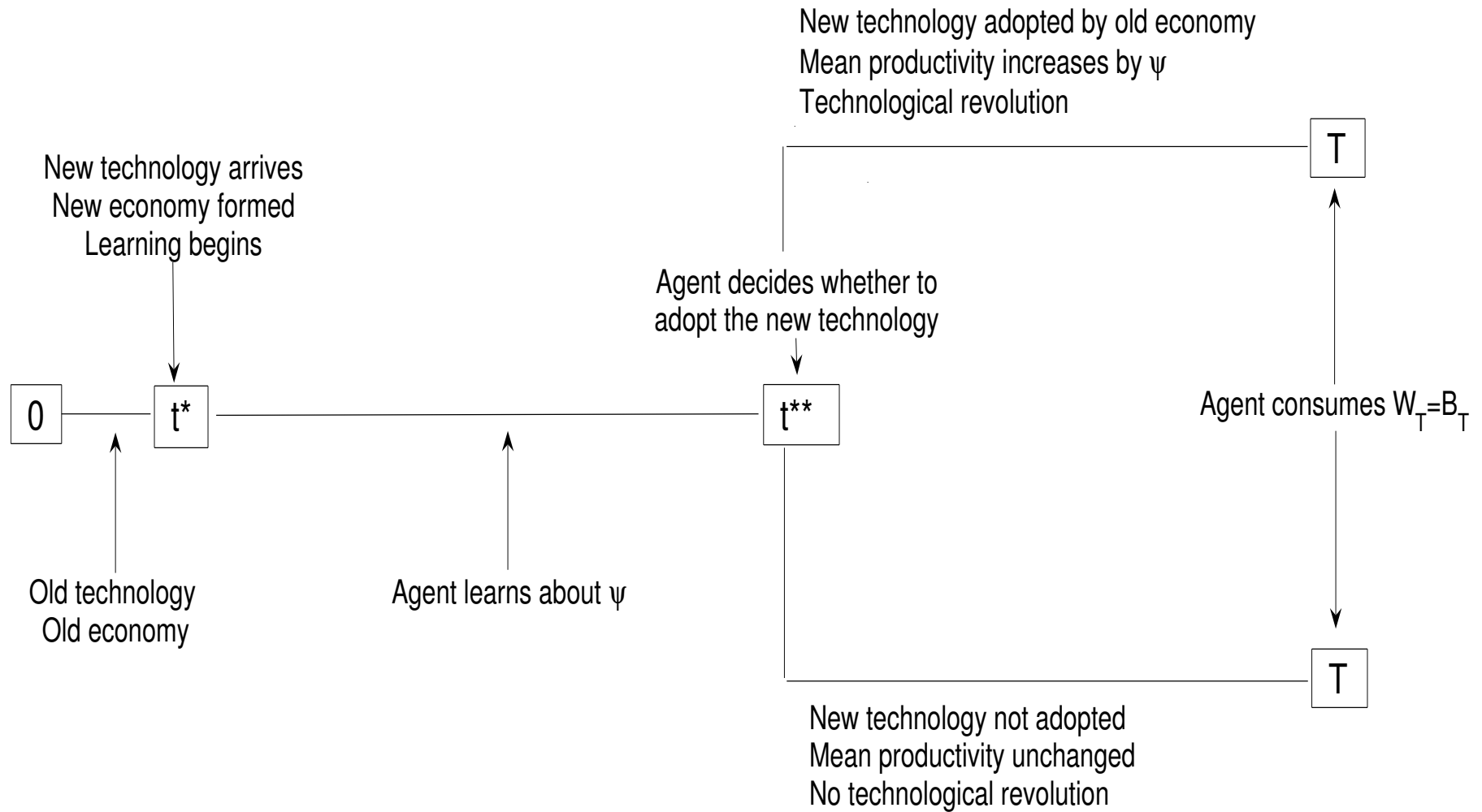
$$dB_t^N = \rho_t^N B_t^N dt$$

$$d\rho_t^N = \phi (\bar{\rho} + \psi - \rho_t^N) dt + \sigma_{N,0} dZ_{0,t} + \sigma_{N,1} dZ_{1,t}$$

Learning: Given the prior $\psi | \mathcal{F}_{t^*} \sim N(0, \hat{\sigma}_{t^*}^2)$, the posterior of ψ is also normal, $\psi | \mathcal{F}_t \sim N(\hat{\psi}_t, \hat{\sigma}_t^2)$, where

$$d\hat{\psi}_t = \hat{\sigma}_t^2 \frac{\phi}{\sigma_{N,1}} d\tilde{Z}_{1,t}$$

$$\hat{\sigma}_t^2 = \frac{1}{\hat{\sigma}_{t^*}^{-2} + \left(\frac{\phi}{\sigma_{N,1}}\right)^2 (t - t^*)}$$



- t^{**} is chosen to maximize utility, but first we take it as given, for simplicity

Technology Adoption

- The agent chooses if/when to adopt the new technology to maximize utility
- The adoption incurs a proportional conversion cost $\kappa \geq 0$ and it is irreversible

Proposition 1: It is never optimal to adopt the new technology at time t^* .

- The prior at time t^* is $\psi \sim N(0, \hat{\sigma}_{t^*}^2)$

Proposition 2: The new technology is adopted at time $t^{**} > t^*$ iff

$$\hat{\psi}_{t^{**}} > \underline{\psi} = -\frac{\log(1 - \kappa)}{A_2(\tau^{**})} + \frac{1}{2}(\gamma - 1) A_2(\tau^{**}) \hat{\sigma}_{t^{**}}^2 ,$$

- Adopt if the new technology is perceived as sufficiently productive

Proposition 3: It is optimal to begin experimenting with new technology at t^* .

- Experimenting provides a valuable option for free

Stock Prices and the Changing Nature of Uncertainty

- Market values of stocks in the old and new economies:

$$M_t = E_t \left[\frac{\pi_T B_T}{\pi_t} \right] \quad \text{and} \quad M_t^N = E_t \left[\frac{\pi_T B_T^N}{\pi_t} \right]$$

- State price density: $\pi_t = E_t [W_T^{-\gamma}] / \lambda$

Propositions 4 - 6: Closed-form formulas for π_t , M_t/B_t and M_t^N/B_t^N

- Stochastic discount factor:

$$\frac{d\pi_t}{\pi_t} = -\gamma A_1(\tau) \sigma d\tilde{Z}_{0,t} - S_{\pi,t} \hat{\sigma}_t^2 \frac{\phi}{\sigma_{N,1}} d\tilde{Z}_{1,t}$$

- In a technological revolution, adoption probability increases from ≈ 0 to 1, so $S_{\pi,t} \uparrow$ and the nature of $\hat{\sigma}_t$ changes from idiosyncratic to systematic

The Dynamics of Prices During a Technological Revolution

- In a technological revolution, $\hat{\psi}_t$ increases from $\hat{\psi}_{t^*} = 0$ to $\hat{\psi}_{t^{**}} > \underline{\psi} > 0$
- The increase in $\hat{\psi}_t$ has two opposing effects on prices:
 - *Cash flow effect:* Expected dividend $\uparrow \Rightarrow M/B \uparrow$
 - *Discount rate effect:* Systematic risk $\uparrow \Rightarrow M/B \downarrow$
- For the new economy, the cash flow effect tends to prevail initially (close to t^*), but the discount rate effect prevails in the end (close to t^{**})
 - \Rightarrow “bubble” in the new economy

Simulations

- We simulate 50,000 samples of shocks in our economy
 - Plot average paths of M/B and volatility across simulations
 - See how these paths differ depending on whether the new technology was eventually adopted (revolution) or not (no revolution)
- ⇒ Tackle the ex post selection bias

Table 1: Parameters used in Simulations.

$\bar{\rho}_L$	$\widehat{\psi}_{t^*}$	$\widehat{\sigma}_{t^*}$	ϕ	σ_0	$\sigma_{N,0}$	$\sigma_{N,1}$	κ	$t^{**} - t^*$	T	γ
0.1217	0	0.04	0.3551	0.07	0.07	0.07	0.1	8	30	4

Figure 3. Average M/B and Volatility in Simulations.

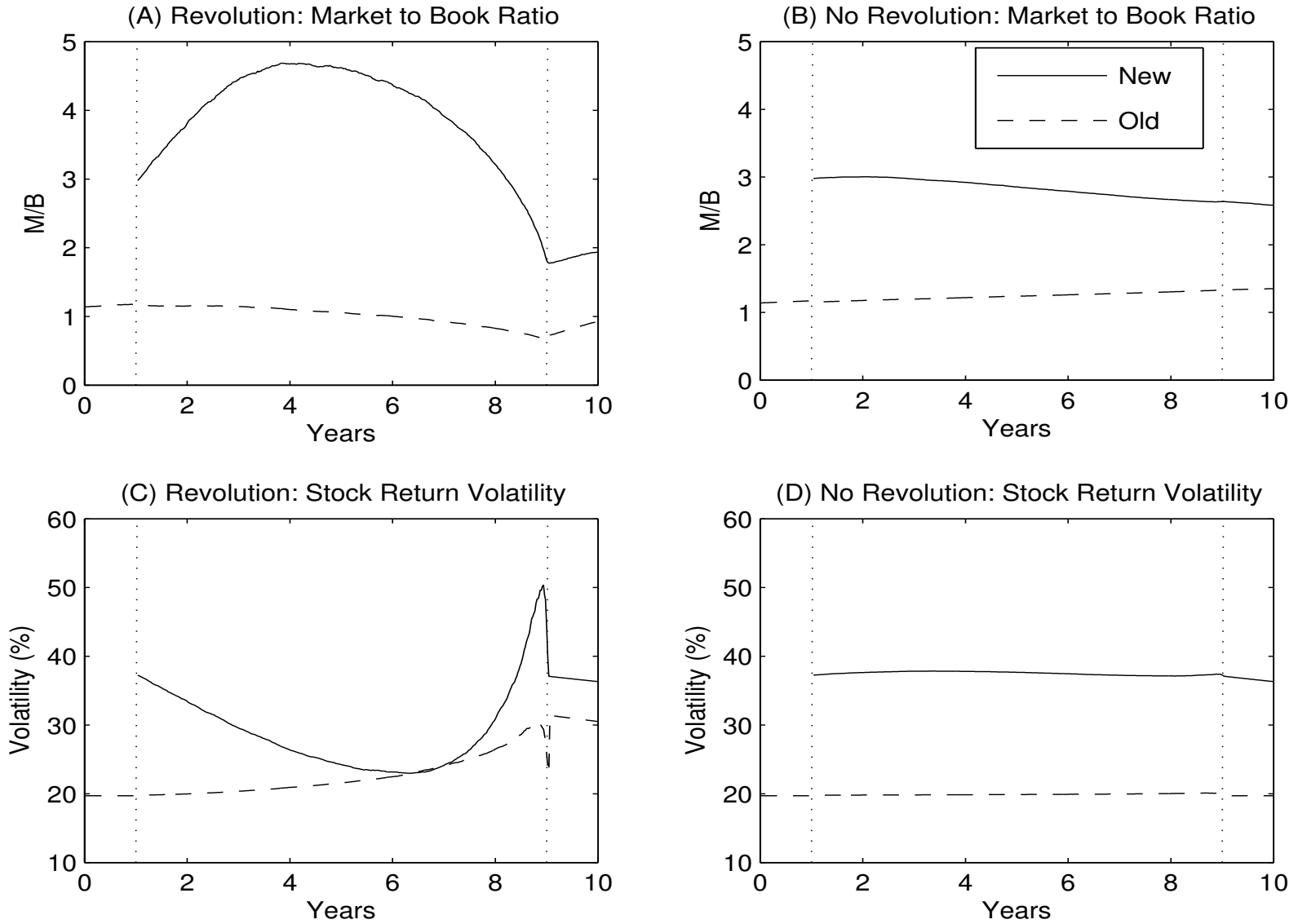


Figure 4. Beta and Average Stock Return in Simulations.

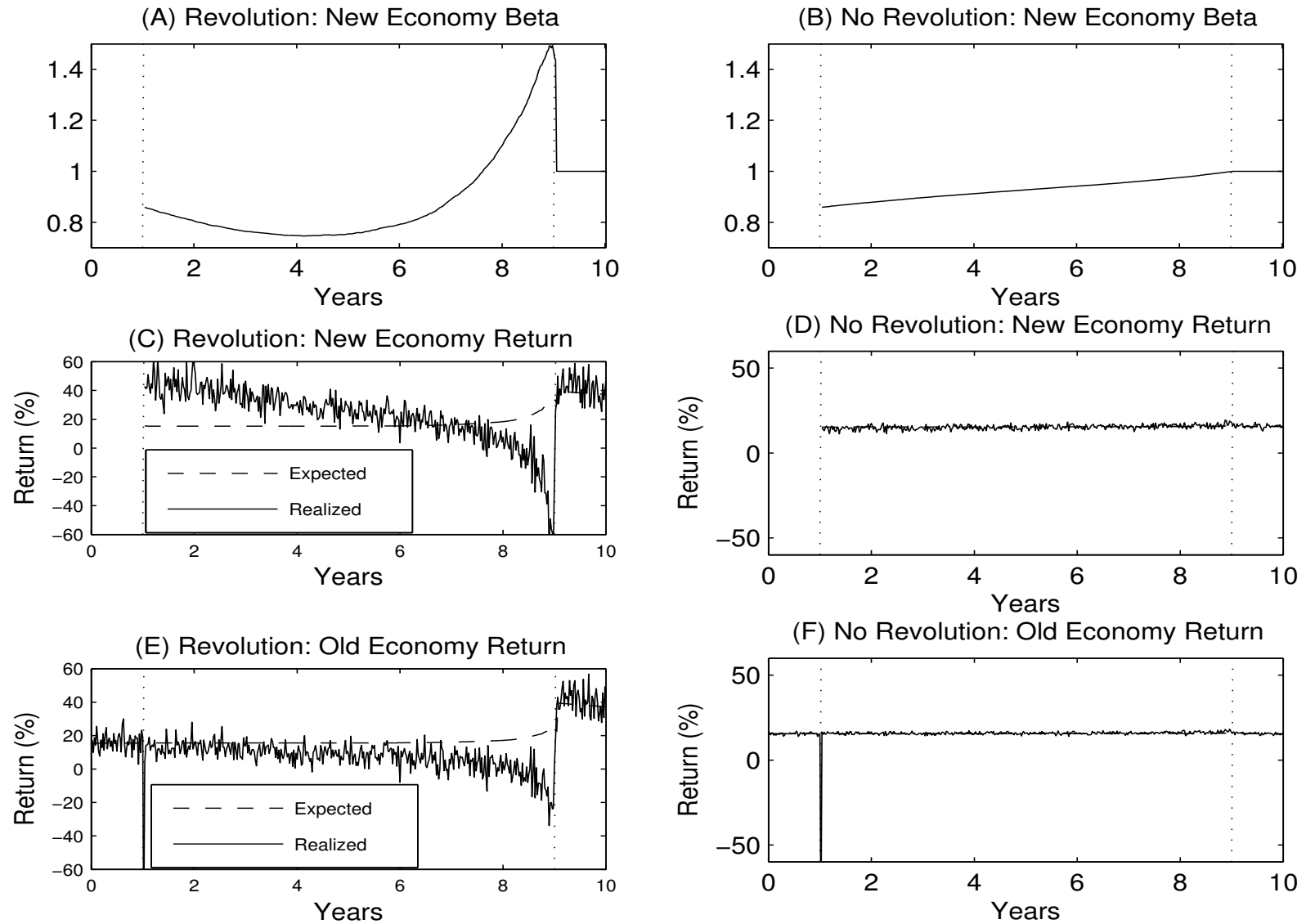


Figure 6. Optimal Adoption Time.

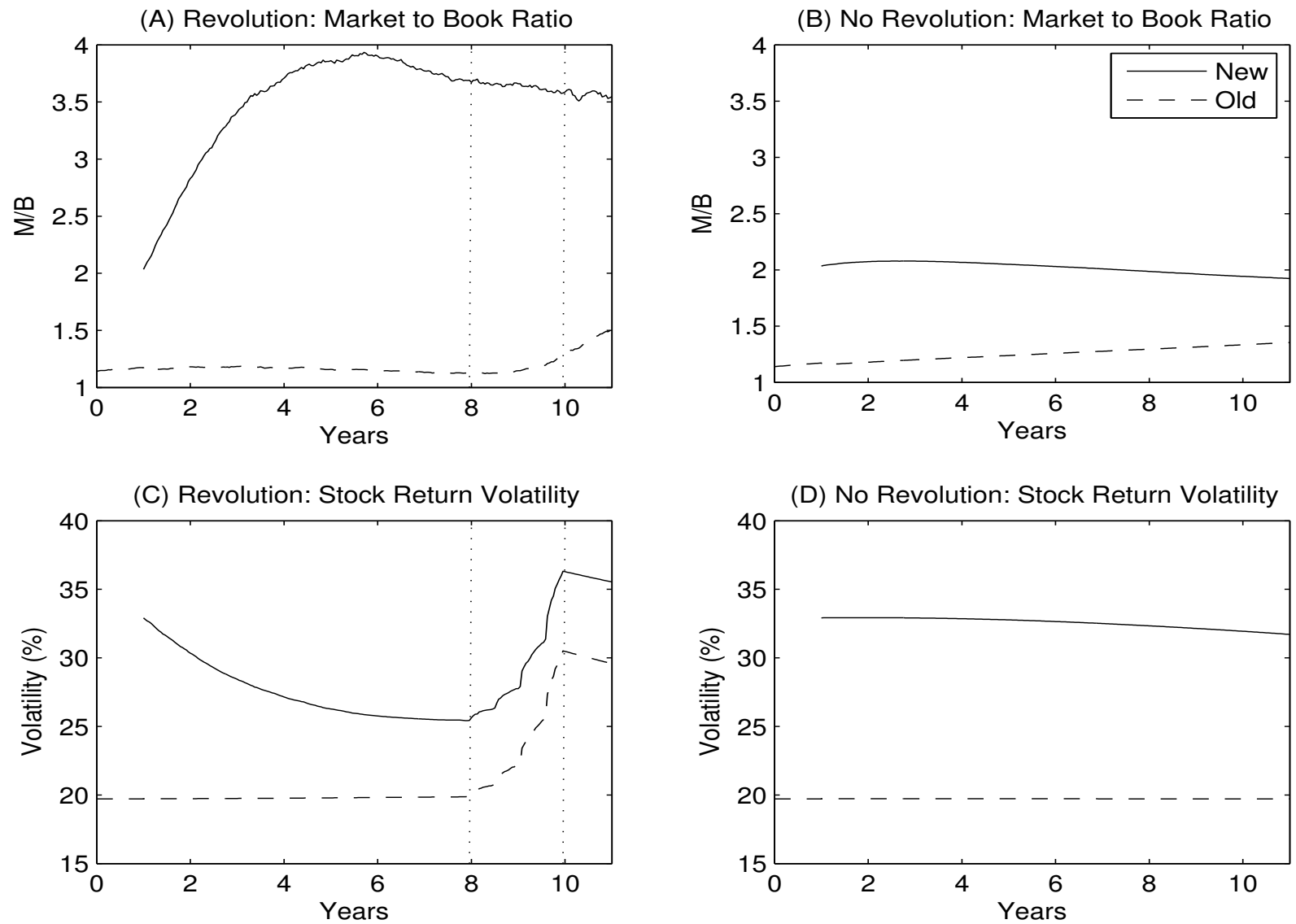


Figure 5. Sensitivity Analysis.

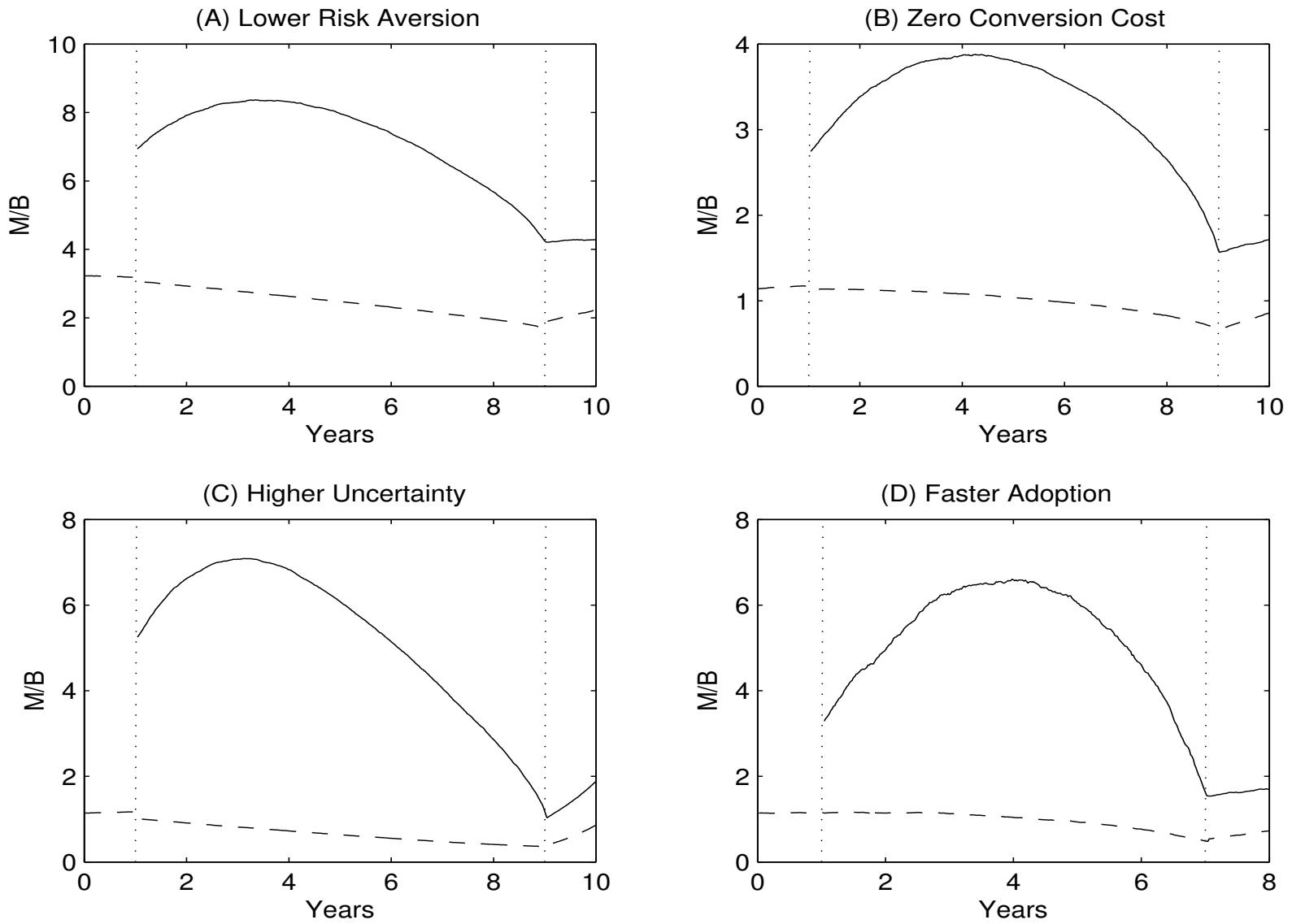


Figure 7. Internet Revolution: Theory.

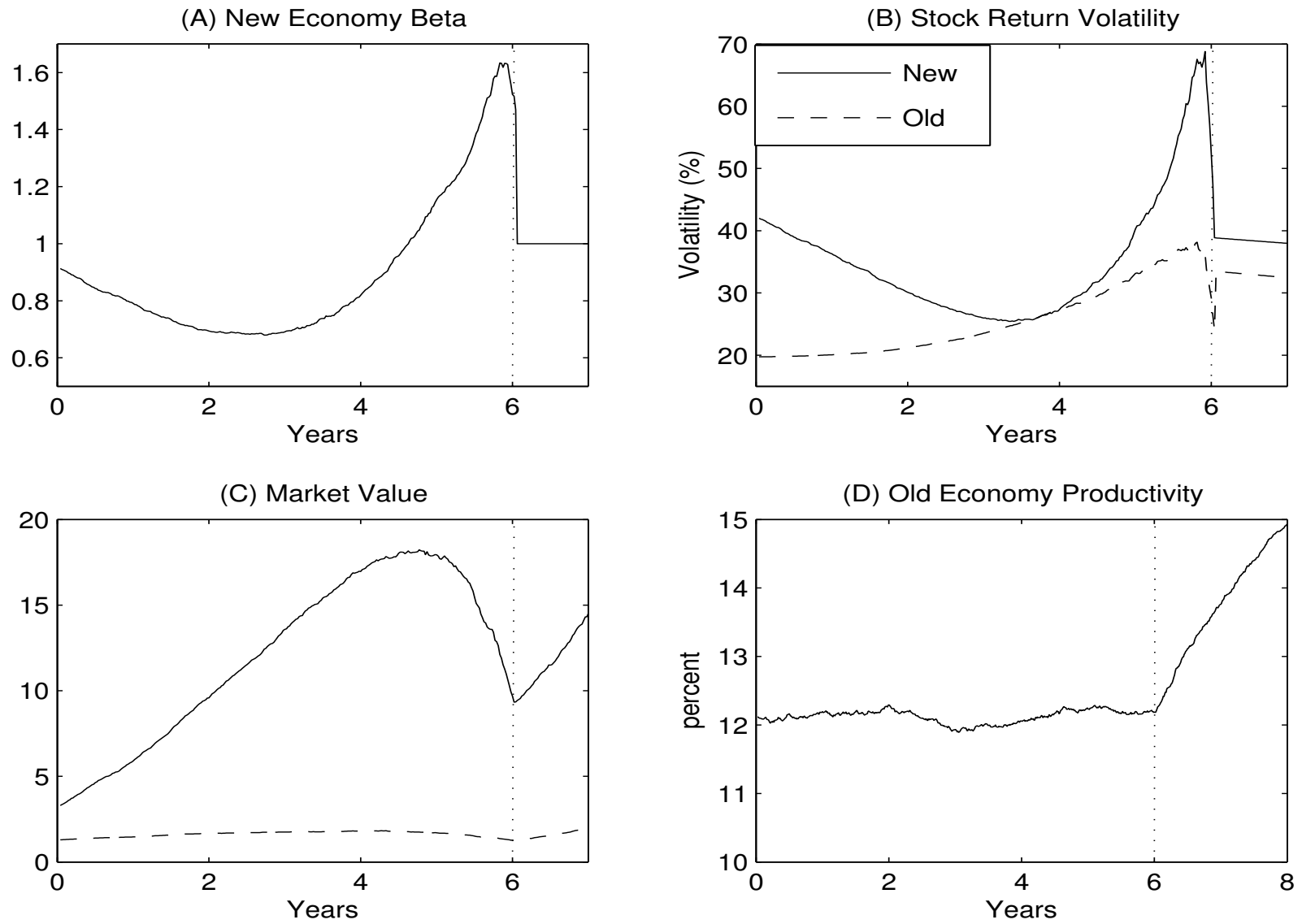
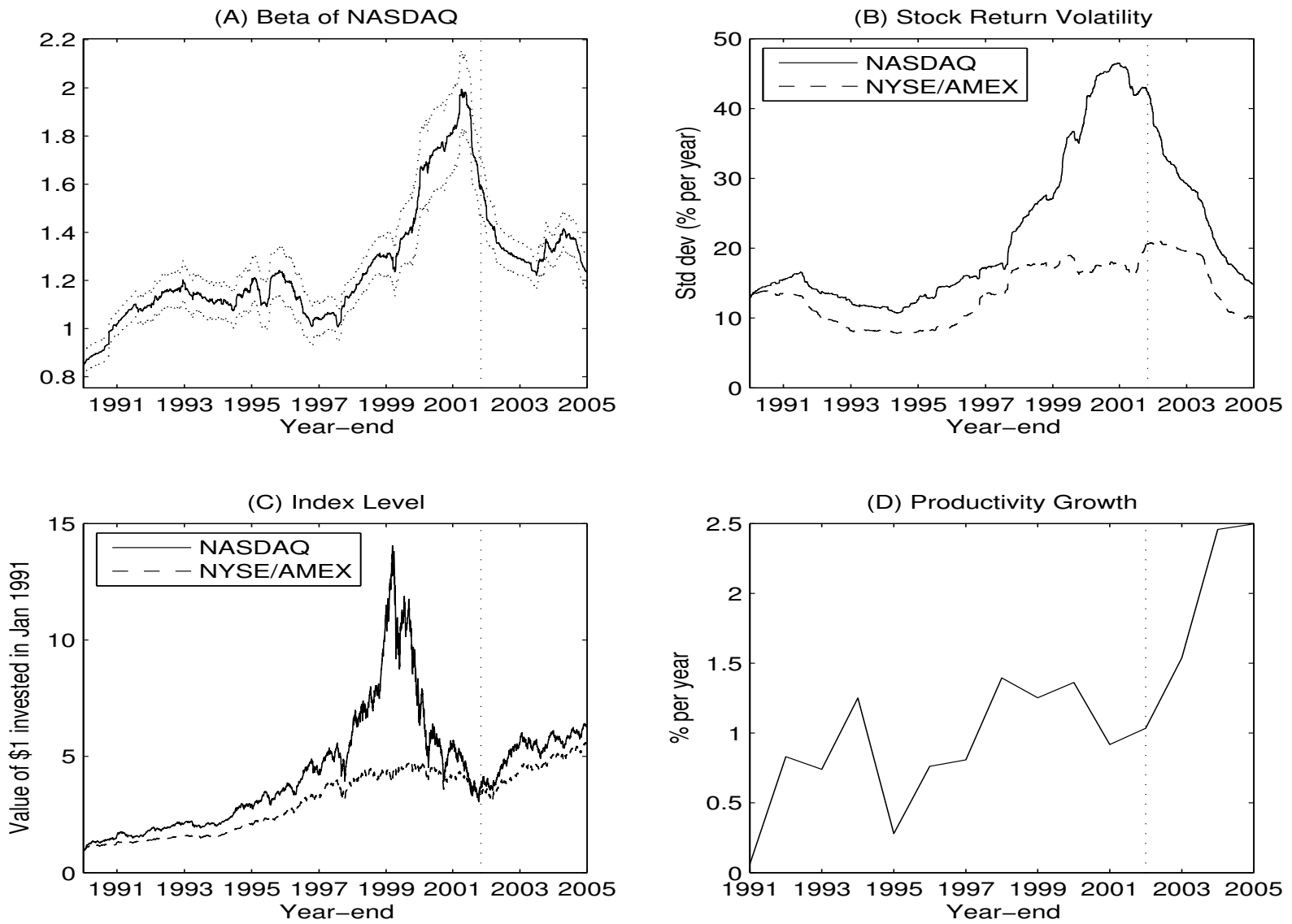


Figure 8. Internet Revolution: Data.



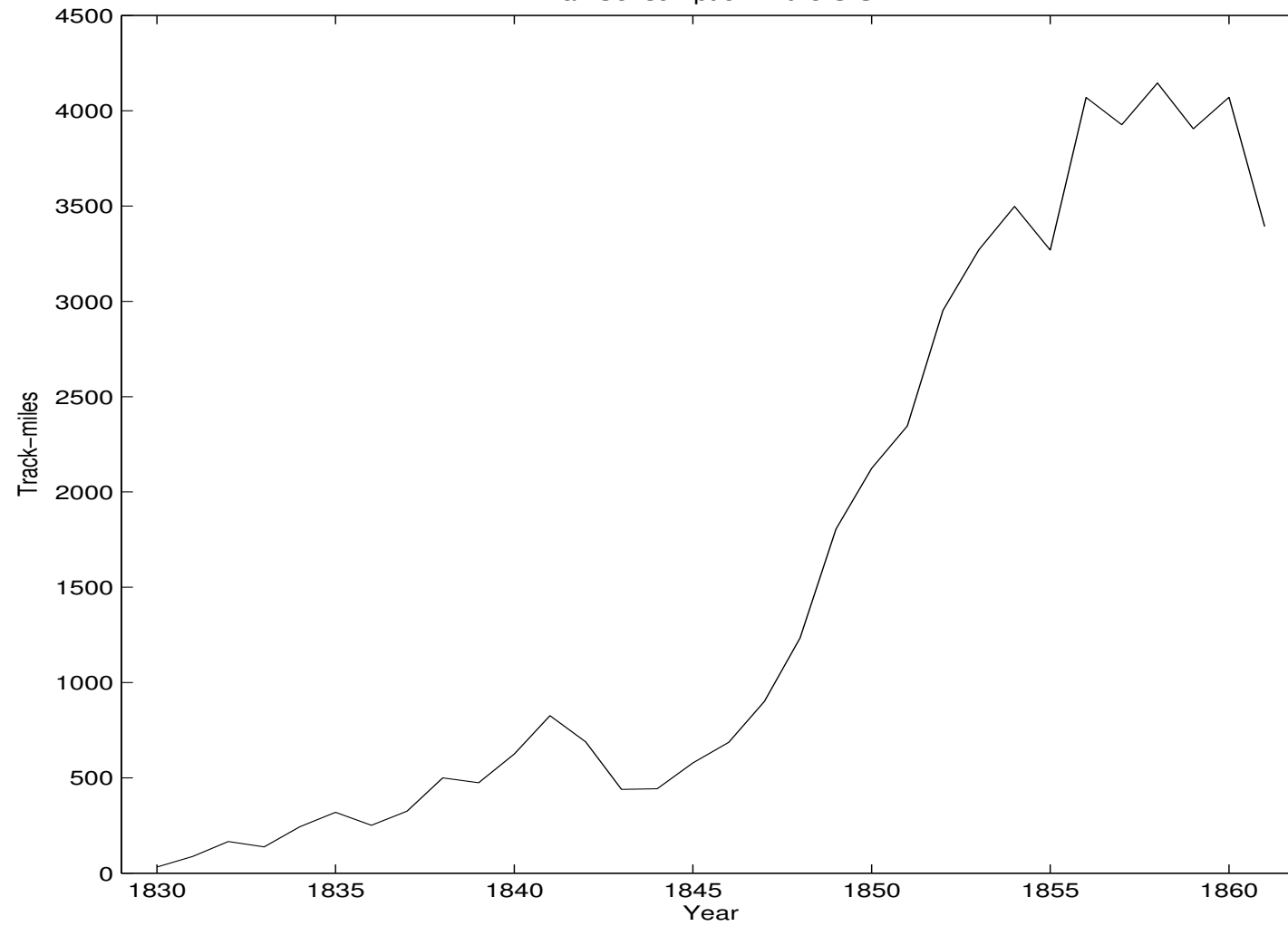
American Railroads Before the Civil War

- Early milestones:
 - 1825: First steam locomotive run (John Stevens)
 - 1828: First RR construction begins (Baltimore & Ohio)
 - 1830: First scheduled steam train run (Charleston)
- It was far from obvious in the 1830s-40s that RRs would later come to dominate the transportation industry
 - Competition with other modes of transportation: wagons, steamboats, canals
 - Waterways were cheaper, wagons more flexible

“Far from being viewed as essential to economic development, the first RRs were widely regarded as having only limited commercial application. Extreme skeptics argued that RRs were too crude to insure regular service, that the sparks thrown off by belching engines would set fire to buildings and fields, and that speeds of 20 to 30 miles per hour could be “fatal to wagons, road and loading, as well as to human life.” More sober critics questioned the ability of RRs to provide low cost transportation. [Some] placed “a RR between a good turnpike and a canal” in transportation efficiency.” (Fogel, 1964)

Figure 9.

Rail Consumption in the U.S.



Railroad Expansion

- Large-scale adoption of RR technology appears to have taken place by 1860
 - 1856: Leap in RR diffusion
 - * Two milestone RRs completed
 - Illinois Central, the longest RR in the world (705 miles)
 - Sacramento Valley, the first RR in California
 - * First RR bridge across Mississippi, heralding westward expansion

“By 1860... the RR had emerged not only as the preferred form of transportation but also as the chief weapon of commercial rivalry.” (Klein, 1994)

- Do stock prices agree with this assessment?

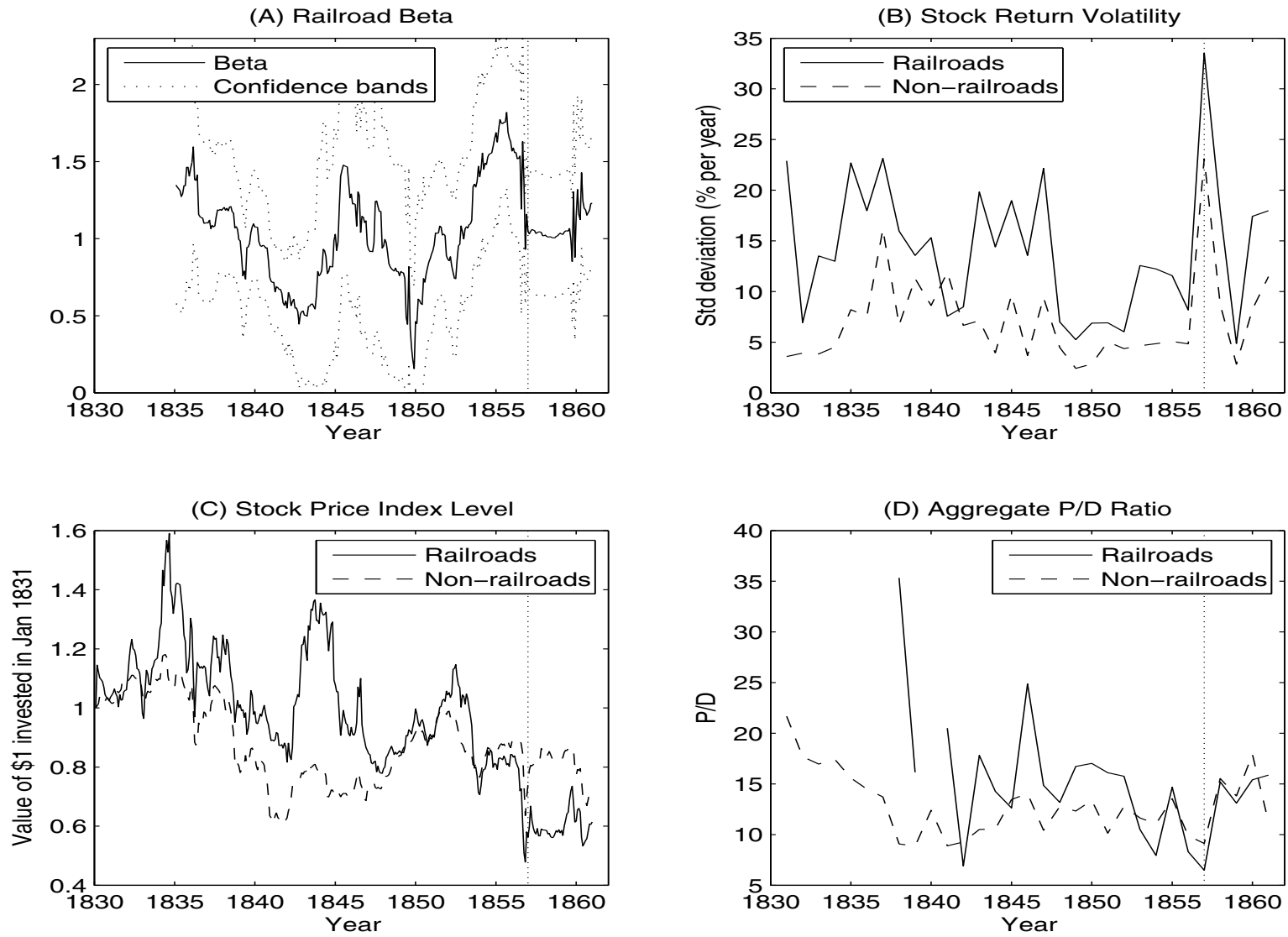
Railroad Stock Prices

- We examine RR stock prices in the early days of the RR (1830–1861)
- Nearly all RRs organized as corporations funded by private investors
 - More than half of the \$300m+ RR investment in 1850 was stock-financed
- Data compiled by Goetzmann, Ibbotson, and Peng (2001)
 - Monthly individual stock prices for 671 NYSE stocks in 1815 to 1925
 - Annual dividends for a subset of stocks in 1825 to 1870
- We focus on common stocks (exclude 85 preferred stocks and 29 scrips)
- We delete apparent data errors (40 of 15,276 prices; 0.26% of observations)
- We fill in price gaps no more than three months long by linear interpolation
 - Before 1848, uninterrupted price sequences for RR stocks are rare
- We identify RRs by name (284 stocks)

Table 2: Railroads Appearing in our Price Index.

Year	Railroad
1831	Camden & Amboy; Canajoharie & Catskill; Harlem; Ithaca & Oswego
1832	Boston & Providence
1833	Boston & Worcester; Brooklyn & Jamaica
1835	Hudson & Berkshire; Long Island
1839	Auburn & Syracuse
1841	Auburn & Rochester
1844	Housatonic
1847	Hudson River; Macon & West
1848	Hartford & New Haven; New York & Erie
1849	Erie
1850	Albany & Schenectady; Baltimore & Ohio; Michigan Central; New York & Harlem
1851	Chemung
1852	Michigan & Southern
1853	Cincinnati, Hamilton & Dayton; Cleveland, Columbus & Cincinnati; Cleveland & Pittsburg; Cleveland & Toledo; Galena & Chicago; Illinois Central; Little Miami
1854	Chicago & Rock Island
1855	Michigan Southern & Northern Indiana
1856	Eighth Avenue; Lacrosse & Milwaukee; Macon & Western
1857	Chicago, Burlington & Quincy; Delaware, Lackawanna & Western; Indianapolis & Cincinnati
1858	Brooklyn City; Buffalo & State Line; Cleveland, Painesville & Ashtabula

Figure 10. The Railroad Revolution: Data.



Another Filtering Result

- We start by providing a result about learning when an underlying (vector) process has a drift that is not observable and jumps from one state to another according to some Markov process.
- This is a slight generalization to the vector case of Theorem 9.1, page 333 in Liptser and Shiriyayev (1977).
- Let \mathbf{X}_t be a N dimensional Ito process described by

$$d\mathbf{X}_t = \boldsymbol{\mu}_t dt + \boldsymbol{\Sigma} d\mathbf{B}_t$$

- where \mathbf{B}_t is a M dimensional Brownian motion.
- Assume that the N dimensional vector $\boldsymbol{\mu}_t$ follows a n state continuous time Markov chain, with states

$$\boldsymbol{\mu}_t \in \{\boldsymbol{\mu}^1, \boldsymbol{\mu}^2, \dots, \boldsymbol{\mu}^n\}$$

Another Filtering Result

- Denote the infinitesimal generator (i.e. transition matrix)

$$\mathbf{\Lambda} = \begin{pmatrix} -\sum_{j \neq 1} \lambda_{1j} & \lambda_{12} & \lambda_{13} & \dots & \lambda_{1n} \\ \lambda_{21} & -\sum_{j \neq 2} \lambda_{2j} & \lambda_{23} & \dots & \lambda_{2n} \\ \lambda_{31} & \lambda_{32} & -\sum_{j \neq 3} \lambda_{3j} & \dots & \lambda_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{n1} & \lambda_{n2} & \lambda_{n3} & \dots & -\sum_{j \neq n} \lambda_{nj} \end{pmatrix}$$

- Notice that for all $i = 1, \dots, n$, $\lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}$ (see e.g. Karlin and Taylor (1975), p.151).
- Both $\boldsymbol{\mu}$ and Σ can be functions of \mathbf{X} .
- We then have the following set of results:

Another Filtering Result

- (A) For given prior distribution $(\widehat{\pi}^1, \dots, \widehat{\pi}^n)$ on $(\boldsymbol{\mu}^1, \dots, \boldsymbol{\mu}^n)$, under some technical conditions (see Liptser and Shirayayev (1977), Ch. 9), the posterior probability

$$\pi_t^i = \Pr(\boldsymbol{\mu}_t = \boldsymbol{\mu}^i \mid \mathcal{F}_t)$$

- satisfies the system of stochastic differential equations:

$$d\pi_t^i = \sum_{j=1}^n \lambda_{ji} \pi_t^j dt + \pi_t^i (\boldsymbol{\mu}^i - \mathbf{m}_t^\mu)' \left(\boldsymbol{\Sigma} \boldsymbol{\Sigma}' \right)^{-\frac{1}{2}} d\widetilde{\mathbf{B}}_t \quad (\text{A1})$$

- under the condition $\pi_0^i = \widehat{\pi}^i$ where

$$\mathbf{m}_t^\mu = \sum_{j=1}^n \pi_t^j \boldsymbol{\mu}^j$$

$$d\widetilde{\mathbf{B}}_t = \left(\boldsymbol{\Sigma} \boldsymbol{\Sigma}' \right)^{-\frac{1}{2}} (d\mathbf{X}_t - \mathbf{m}_t^\mu dt)$$

- (B) If for all $i = 1, \dots, n$, $\pi_0^i > 0$ and $\lambda_{ij} > 0$ for all ij , then for every finite t

$$\text{Prob}(\pi_t^i > 0) = 1$$

Another Filtering Result

- **Remark 1:** It is immediate to check that we always have

$$\sum_{i=1}^n \pi_t^i \equiv 1$$

- In fact, define $K_t = \sum_{i=1}^n \pi_t^i$, from Ito's lemma

$$\begin{aligned} dK_t &= \sum_{i=1}^n d\pi_t^i = \sum_{i=1}^n \sum_{j=1}^n \lambda_{ji} \pi_t^j dt + \sum_{i=1}^n \pi_t^i (\boldsymbol{\mu}^i - \mathbf{m}_t^\mu)' (\boldsymbol{\Sigma} \boldsymbol{\Sigma}')^{-\frac{1}{2}} d\tilde{\mathbf{B}}_t \\ &= \sum_{j=1}^n \pi_t^j \sum_{i=1}^n \lambda_{ji} dt + (\mathbf{m}_t^\mu - \mathbf{m}_t^\mu)' (\boldsymbol{\Sigma} \boldsymbol{\Sigma}')^{-\frac{1}{2}} d\tilde{\mathbf{B}}_t \\ &= 0 \end{aligned}$$

- because $\sum_{i=1}^n \lambda_{ji} = \sum_{i \neq j} \lambda_{ji} + \lambda_{jj} = \sum_{i \neq j} \lambda_{ji} + \left(-\sum_{i \neq j} \lambda_{ji}\right) = 0$.
- Hence, if the initial condition $\hat{\boldsymbol{\pi}} = (\hat{\pi}^1, \dots, \hat{\pi}^n)$ is such that $\sum_{i=1}^n \hat{\pi}^i = 1$, the equality is preserved over time.
- **Remark 2:** Because of learning, the effect of the n Poisson processes implicit in the continuous time Markov chain $\boldsymbol{\Lambda}$ are “smoothed out.”
- In other words, they give rise to n state variables (the posteriors π_t^i) but they are all driven by M Brownian motions.

Some Additional Useful Results (Veronesi 2001)

- Consider now the case in which:

$$\frac{dD}{D} = \theta_t dt + \sigma_D dB_t^D; \quad de = \theta_t dt + \sigma_e dB_t^e$$

- **Remark 3:** From diffusion process for π_t , we then obtain a formula for:

$$\pi_{t,\tau}^i = \Pr(\theta_\tau = \theta_i | \mathcal{F}_t)$$

- This stems from the linear structure of the probabilities. Since they multiply a matrix $\mathbf{\Lambda}$, we have the following implications.
- Let's rewrite the process for π_t in vector form:

$$d\pi_t' = \mathbf{\Lambda}' \pi_t' dt + \pi_t \odot (\Theta - m_t^\theta \mathbf{1}_n) \mathbf{h} d\tilde{\mathbf{B}}_t \quad (22)$$

- where π_t is a $1 \times N$ vector, $\Theta = [\theta_1, \theta_2, \dots, \theta_n]$, and $\mathbf{h} = (h_D, h_e)$.

Some Additional Useful Results (Veronesi 2001)

- Let

$$\tilde{\pi}_t(u) = E[\pi_u | \mathcal{F}_t]$$

- We can write (22) in integral form

$$\pi'_u = \pi'_t + \int_t^u \Lambda' \pi'_s ds + \int_t^u \pi_s \odot (\Theta - m_s^\theta \mathbf{1}_n) \mathbf{h} d\tilde{\mathbf{B}}_s$$

- The diffusion term is bounded. Hence, the stochastic integral is a martingale.
- Take expectations on both sides and using the fact that the stochastic integral has zero expectations, we then have

$$\tilde{\pi}'_t(u) = \pi'_t + \int_t^u \Lambda' \tilde{\pi}'_t(s) ds$$

- This can be rewritten as

$$d\tilde{\pi}'_t(s) = \Lambda' \tilde{\pi}'_t(s) ds$$

- (recall that t is fixed!)

Some Additional Useful Results (Veronesi 2001)

- The solution to this system of ordinary differential equations with initial condition $\tilde{\boldsymbol{\pi}}_t(t) = \boldsymbol{\pi}_t$ is

$$\tilde{\pi}_t^i(u) = E[\pi_u^i | \mathcal{F}_t] = \sum_{k=1}^N \pi_t^k \sum_{j=1}^N w_{jk}^{-1} w_{ij} e^{\omega_j(u-t)}$$

- where ω_j are the eigenvalues of $\boldsymbol{\Lambda}'$, w_{ij} are associated eigenvectors, collected in the matrix \mathbf{W} , and $w_{ij}^{-1} = [\mathbf{W}^{-1}]_{ij}$.
- In other words, define the diagonal matrix

$$\mathbf{E}(\tau) = \text{diag}(e^{\omega_j \tau})$$

- we can define the transition matrix between t and $t + \tau$ as

$$\boldsymbol{\Lambda}(\tau) = \mathbf{W}^{-1} \mathbf{E}(\tau) \mathbf{W}$$

- It is easy to show that

$$\boldsymbol{\Lambda} = \lim_{\tau \rightarrow 0} \frac{1}{\tau} (\boldsymbol{\Lambda}(\tau) - \mathbf{I})$$

Some Additional Useful Results (Veronesi 2001)

- This implies that we can compute the expectations of all the linear functions that depend $\boldsymbol{\pi}_t$.

$$\begin{aligned}
 E \left[\sum_{i=1}^n \alpha_i(\tau) \pi_{\tau}^i | \mathcal{F}_t \right] &= \sum_{i=1}^n \alpha_i(\tau) E [\pi_{\tau}^i | \mathcal{F}_t] \\
 &= \sum_{i=1}^n \alpha_i(\tau) \sum_{k=1}^N \pi_t^k \sum_{j=1}^N w_{jk}^{-1} w_{ij} e^{\omega_j(u-t)} \\
 &= \sum_{k=1}^N \pi_t^k \sum_{i=1}^n \alpha_i(\tau) w_{ij} \sum_{j=1}^N w_{jk}^{-1} e^{\omega_j(u-t)} \\
 &= \boldsymbol{\pi}_t \boldsymbol{\Lambda}(\tau) \boldsymbol{\alpha}(\tau)
 \end{aligned}$$

- Finally, as a special case, we have

$$\pi_{t,\tau}^i = \Pr(\theta_{\tau} = \theta_i | \mathcal{F}_t) = E [1_{\{\theta_{\tau} = \theta_i\}} | \mathcal{F}_t] = [\boldsymbol{\pi}_t \boldsymbol{\Lambda}(\tau)]_i$$

Some Additional Useful Results (Veronesi 2001)

- Identify the state price density as the marginal utility of the representative agent:

$$\chi_t = e^{-\phi t} u_c(c_t)$$

- We can use this to obtain the prices for stocks and bonds.
- In order to do this, we need the following Lemma (see Veronesi (2001)):
- **Lemma 2:** Let β be any constant and define $n_t^i = c_t^\beta \pi_t^i$. Let also

$$\bar{\Lambda}_\beta = \Lambda + \text{diag}(\theta_1^\beta, \dots, \theta_n^\beta)$$

- with $\theta_i^\beta = \beta\theta_i + \frac{1}{2}\beta(\beta-1)\sigma_D^2$. Then for $u > t$ we have

$$E[n_u^i | \mathcal{F}_t] = \sum_{k=1}^N n_t^k \sum_{j=1}^N w(\beta)_{jk}^{-1} w_{ij}(\beta) e^{\omega_j(\beta)(u-t)}$$

- where $\omega_j(\beta)$ are the eigenvalues of $\bar{\Lambda}'_\beta$ and $w_{ij}(\beta)$ are associated eigenvectors and $w(\beta)_{ij}^{-1} = [W^{-1}]_{ij}$.
- Clearly, the result obtained above for the posterior π is a special case of this one, where $\beta = 0$.

- *Proof*: The proof is basically identical to the previous one. By Ito's lemma, one can rewrite

$$\begin{aligned} dn_t^i &= c_t^\beta d\pi_t^i + \beta c_t^{\beta-1} \pi_t^i dc_t + \frac{1}{2} \beta (\beta - 1) \pi_t^i c_t^{\beta-2} dc_t^2 + \beta c_t^{\beta-1} d\pi_t^i dc_t \\ &= \left\{ [\mathbf{n}_t \boldsymbol{\Lambda}]_i + \beta n_t^i m_t^\theta + \frac{1}{2} \beta (\beta - 1) n_t^i \sigma_D^2 + \beta n_t^i (\theta_i - m_t^\theta) \mathbf{h} (\sigma_D, 0)' \right\} dt \\ &\quad + n_t^i ((\theta_i - m_t^\theta) \mathbf{h} + (\sigma_D, 0)) d\tilde{\mathbf{B}}_t \end{aligned}$$

- where $\mathbf{n}_t = (n_t^1, \dots, n_t^n)$. Since

$$(\theta_i - m_t^\theta) \mathbf{h} (\sigma_D, 0)' = (\theta_i - m_t^\theta) \frac{1}{\sigma_D} \times \sigma_D = \theta_i - m_t^\theta$$

- we have

$$dn_t^i = \left\{ [\mathbf{n}_t \boldsymbol{\Lambda}]_i + \frac{1}{2} \beta (\beta - 1) n_t^i \sigma_D^2 + \beta n_t^i \theta_i \right\} dt + n_t^i ((\theta_i - m_t^\theta) \mathbf{h} + (\sigma_D, 0)) d\tilde{\mathbf{B}}_t$$

- As before, we can now write this in vector form:

$$d\mathbf{n}' = \overline{\boldsymbol{\Lambda}}_\beta' \mathbf{n}' dt + \mathbf{n} \odot \boldsymbol{\Sigma}(\pi) d\tilde{\mathbf{B}}_t \quad (23)$$

- where $\overline{\boldsymbol{\Lambda}}_\beta$ is defined above and $\boldsymbol{\Sigma}(\pi)$ is some bounded $n \times 2$ matrix.
- At this point the proof is the same as before, obtaining the representation in the lemma.

Application: Bond Prices and Interest Rates

- Let's first find the instantaneous interest rate.
- By definition of the state-price density, we have

$$r_t = E \left[-\frac{d\chi_t}{\chi_t} \middle| \mathcal{F}_t \right] = \phi + \gamma m_t^\theta - \frac{1}{2} \gamma (\gamma + 1) \sigma_D^2$$

- Let's look at bond prices, now. This result is due Veronesi and Yared (2000) (based on a previous work by Yared (1999)), although these papers investigate nominal bonds and not only real bonds.
 - It is simple to introduce inflation in this setting: Just assume that the price index follows a lognormal process with unobservable drift rate q_t , where the latter varies over time according to a m-state, continuous time Markov chain model analogous to the one assumed here for θ_t .
 - It is simple to see that all results then follow also for nominal bonds (with the obvious change in interpretation). See e.g. Veronesi and Yared (2000). See also David and Veronesi (2008).

Application: Bond Prices and Interest Rates

- Let $P_t(\tau)$ be price of a bond paying 1 unit of consumption good at time $t + \tau$.
- By definition, we then have

$$\begin{aligned} P_t(\tau) &= \frac{1}{\chi_t} E[\chi_{t+\tau} \times 1 | \mathcal{F}_t] = \frac{1}{e^{-\phi t} u_c(c_t)} E[e^{-\phi(t+\tau)} u_c(c_{t+\tau}) \times 1 | \mathcal{F}_t] \\ &= \frac{e^{-\phi\tau}}{c_t^{-\gamma}} E[c_{t+\tau}^{-\gamma} \times 1 | \mathcal{F}_t] \end{aligned}$$

- The problem with evaluating $E[c_{t+\tau}^{-\gamma} | \mathcal{F}_t]$ is that the drift rate of consumption changes over time according to the linear process $m_t^\theta = \sum_{i=1}^n \pi_t^i \theta_i$.
- However, since we know that $\sum_{i=1}^n \pi_t^i = 1$, we can equivalently write

$$P_t(\tau) = \frac{e^{-\phi\tau}}{c_t^{-\gamma}} E\left[c_{t+\tau}^{-\gamma} \times \sum_{i=1}^n \pi_{t+\tau}^i | \mathcal{F}_t\right] = \frac{e^{-\phi\tau}}{c_t^{-\gamma}} \sum_{i=1}^n E\left[c_{t+\tau}^{-\gamma} \times \pi_{t+\tau}^i | \mathcal{F}_t\right]$$

- We can then use the Lemma with $\beta = -\gamma$ to find immediately

$$P_t(\tau) = \boldsymbol{\pi}_t \mathbf{K}(\tau)$$

Application: Bond Prices and Interest Rates

- where

$$K_i(\tau) = \sum_{k=1}^N \sum_{j=1}^N w_{jk}^{-1} w_{ij} e^{-(\phi - \omega_j)\tau}$$

- and where $w_{ij} = [\mathbf{W}]_{ij}$, $w_{ij}^{-1} = [\mathbf{W}^{-1}]_{ij}$, with \mathbf{W} given by the column eigenvectors of the matrix

$$\bar{\mathbf{\Lambda}} = \mathbf{\Lambda} - \gamma \text{diag}(\theta_1, \dots, \theta_n) - \frac{1}{2} \gamma (\gamma + 1) \sigma_D^2 \mathbf{I}_n$$

- and $(\omega_1, \dots, \omega_n)$ being its eigenvalues.

Application: Stock Prices

- Turning to stock prices, we can use a very similar methodology, with the only detail that we now have to take an integral.
- Using again the asset pricing formula

$$\begin{aligned}
 S_t &= \frac{1}{\chi_t} E \left[\int_0^\infty \chi_{t+\tau} \times D_{t+\tau} d\tau \mid \mathcal{F}_t \right] \\
 &= \frac{1}{e^{-\phi t} c_t^{-\gamma}} E \left[\int_0^\infty e^{-\phi(t+\tau)} c_{t+\tau}^{-\gamma} \times D_{t+\tau} d\tau \mid \mathcal{F}_t \right] \\
 &= \frac{1}{e^{-\phi t} c_t^{-\gamma}} E \left[\int_0^\infty e^{-\phi(t+\tau)} c_{t+\tau}^{1-\gamma} d\tau \mid \mathcal{F}_t \right]
 \end{aligned}$$

- Using the same trick that for every t $\sum_{i=1}^n \pi_t = 1$, we can rewrite

$$\begin{aligned}
 S_t &= \frac{1}{e^{-\phi t} c_t^{-\gamma}} E \left[\int_0^\infty e^{-\phi(t+\tau)} c_{t+\tau}^{1-\gamma} \sum_{i=1}^n \pi_{t+\tau}^i d\tau \mid \mathcal{F}_t \right] \\
 &= \frac{1}{c_t^{-\gamma}} \sum_{i=1}^n E \left[\int_0^\infty e^{-\phi\tau} c_{t+\tau}^{1-\gamma} \pi_{t+\tau}^i d\tau \mid \mathcal{F}_t \right]
 \end{aligned}$$

Application: Stock Prices

- Assuming that all the integrals are finite, we can use Fubini to invert the order of integration, and then Lemma 2, with $\beta = 1 - \gamma$, to find

$$\begin{aligned}
 S_t &= \frac{1}{c_t^{-\gamma}} \sum_{i=1}^n \int_0^{\infty} e^{-\phi\tau} E \left[c_{t+\tau}^{1-\gamma} \pi_{t+\tau}^i | \mathcal{F}_t \right] d\tau \\
 &= \frac{1}{c_t^{-\gamma}} \sum_{i=1}^n \int_0^{\infty} e^{-\phi\tau} \sum_{k=1}^n c_t^{1-\gamma} \pi_t^k \sum_{j=1}^n w_{jk}^{-1} w_{ij} e^{\omega_j \tau} d\tau \\
 &= c_t \sum_{k=1}^n \pi_t^k C_k
 \end{aligned}$$

- where

$$C_k = \sum_{i=1}^n \sum_{j=1}^n w_{jk}^{-1} w_{ij} \int_0^{\infty} e^{-(\phi - \omega_j)\tau} d\tau = \sum_{i=1}^n \sum_{j=1}^n w_{jk}^{-1} w_{ij} \frac{1}{(\phi - \omega_j)}$$

- where now $w_{ij} = [\mathbf{W}]_{ij}$, $w_{ij}^{-1} = [\mathbf{W}^{-1}]_{ij}$, with $(\omega_1, \dots, \omega_n)$ and \mathbf{W} given by the eigenvalues and column eigenvectors, respectively, of the matrix

$$\bar{\mathbf{\Lambda}} = \mathbf{\Lambda} + (1 - \gamma) \text{diag}(\theta_1, \dots, \theta_n) - \frac{1}{2} \gamma (1 - \gamma) \sigma_D^2 \mathbf{I}_n$$

Application: Stock Prices

- Using $D_t = c_t$, we finally obtain that the price dividend ratio is

$$\frac{S_t}{D_t} = \boldsymbol{\pi}_t \cdot \mathbf{C}$$

- It is convenient to obtain a nicer representation for the constants C_k . We find

$$\mathbf{C} = (\phi \mathbf{I}_n - \bar{\boldsymbol{\Lambda}})^{-1} \mathbf{1}_n$$

- To show this, we must show

$$\sum_{j=1}^n w_{jk}^{-1} w_{ij} \frac{1}{\phi - \omega_j} = e_i \left(\phi \mathbf{I} - \bar{\boldsymbol{\Lambda}}' \right)^{-1} e_k$$

- Let $\boldsymbol{\Omega}$ be the diagonal matrix with the eigenvalues ω_j of $\bar{\boldsymbol{\Lambda}}'$ on the principal diagonal.

- Then we know that

$$\left(\phi \mathbf{I} - \bar{\boldsymbol{\Lambda}}' \right)^{-1} = \mathbf{W} (\mathbf{I}\phi - \boldsymbol{\Omega})^{-1} \mathbf{W}^{-1}$$

- Let $\mathbf{D} = (\mathbf{I}\phi - \boldsymbol{\Omega})^{-1}$. Since this is a diagonal matrix, we finally obtain

$$e_i \left(\phi \mathbf{I} - \bar{\boldsymbol{\Lambda}}' \right)^{-1} e_k = \sum_{j=1}^n \sum_{\ell=1}^n w_{ij} \mathbf{D}_{j\ell} w_{\ell k}^{-1} = \sum_{j=1}^n \frac{w_{ij} w_{jk}^{-1}}{\phi - \omega_j}$$

Linearity Generating Models: Gabaix (2009)

- Reverse Engineering.
- Start from the pricing formula

$$\chi_t P_t = E \left[\int_t^\infty \chi_u D_u du \mid \mathbf{X}_t \right]$$

– where $\{D_t\}$ is a cash flow process, and \mathbf{X}_t is a set of state variables.

- The LG Twist:

1. Call

$$\mathbf{Y}_t = \begin{pmatrix} \chi_t D_t \\ \chi_t D_t \mathbf{X}_t \end{pmatrix}$$

2. Assume \mathbf{Y}_t is a process such that

$$E[d\mathbf{Y}_t] = -\boldsymbol{\omega} \mathbf{Y}_t dt$$

$$\implies E[Y_{i,u} \mid \mathbf{Y}_t] = \sum_{k=0}^n Y_{k,t} \sum_{j=0}^n [\mathbf{U}]_{ij} [\mathbf{U}^{-1}]_{jk} e^{-\Lambda_{jj}(u-t)}$$

* where \mathbf{U} and $\boldsymbol{\Lambda}$ are the eigenvectors / eigenvalue matrices of $\boldsymbol{\omega}$.

Linearity Generating Models: Gabaix (2009)

3. Assume that all integration conditions are satisfied (etc etc) and solve:

$$\begin{aligned}
 \chi_t P_t &= E \left[\int_t^\infty \chi_u D_u du \mid X_t \right] = E \left[\int_t^\infty Y_{0,u} du \mid Y_t \right] = \int_t^\infty E [Y_{0,u} \mid Y_t] du \\
 &= \sum_{k=0}^n \sum_{j=0}^n Y_{k,t} [U]_{ij} [U^{-1}]_{jk} \int_t^\infty e^{-\Lambda_{jj}(u-t)} du \\
 &= \sum_{k=0}^n Y_{k,t} b_k
 \end{aligned}$$

where

$$b_k = \sum_{j=0}^n \frac{[U]_{ij} [U^{-1}]_{jk}}{\Lambda_{jj}}$$

4. Finally, substitute $Y_{k,t}$

$$\chi_t P_t = Y_{0,t} b_0 + \sum_{k=1}^n Y_{k,t} b_k = \chi_t D_t b_0 + \sum_{k=1}^n \chi_t D_t X_{k,t} b_k$$

* and obtain

$$P_t = D_t (b_0 + \mathbf{b}' \cdot \mathbf{X}_t)$$

Examples in the Paper

1. Generalized Gordon Model 1: stochastic dividend drift

$$\left. \begin{aligned} \chi_t &= e^{-rt} \\ dD_t/D_t &= (g_* + \gamma_t) dt + \sigma dW_1 \\ d\gamma_t &= -(\phi\gamma_t + \gamma_t^2) dt + \sigma_\gamma(\gamma_t) dW_2 \end{aligned} \right\} \implies \frac{P_t}{D_t} = b_0 + b_1\gamma_t$$

2. Generalized Gordon Model 2: stochastic dividend drift and stochastic risk premium

$$\left. \begin{aligned} d\chi_t/\chi_t &= -r dt - (\pi_* + \hat{\pi}_t) / \sigma dW_1 \\ dD_t/D_t &= (g_* + \hat{g}_t) dt + \sigma dW_1 \\ d\hat{g}_t &= -\phi_g \hat{g}_t dt + \hat{g}_t (\hat{\pi}_t - \hat{g}_t) dt + \sigma_g (\hat{g}_t, \hat{\pi}_t) dW_2 \\ d\hat{\pi}_t &= -\phi_\pi \hat{\pi}_t dt + \hat{\pi}_t (\hat{\pi}_t - \hat{g}_t) dt + \sigma_\pi (\hat{g}_t, \hat{\pi}_t) dW_3 \end{aligned} \right\} \implies \frac{P_t}{D_t} = b_0 + b_1 \hat{g}_t + b_2 \hat{\pi}_t$$

(with $b_2 < 0$)

3. Generalized Gordon Model 3: multiple factors

$$\left. \begin{aligned} \chi_t &= e^{-rt} \\ \frac{dD_t}{D_t} &= \left(g_* + \sum_{i=1}^n X_{i,t} \right) dt + \sigma dW_1 \\ E[dX_{i,t}] / dt &= -\phi_i X_{i,t} - (g_t - g_*) X_{i,t} \end{aligned} \right\} \implies \frac{P_t}{D_t} = b_0 + \mathbf{b}'_1 \cdot \mathbf{X}_t$$

Examples in the Paper

4. Aggregate Model of Menzly, Santos and Veronesi (2004):

$$\left. \begin{aligned} \chi_t &= e^{-\rho t}; && \text{(and not } = e^{-\rho t} Y_t / C_t \text{ !)} \\ dC_t / C_t &= \mu dt + \sigma dW_1 \\ dY_t &= k (\bar{Y} - Y_t) dt - \alpha (Y_t - \lambda) dW_1 \end{aligned} \right\} \implies \frac{P_t}{C_t} = b_0 + b_1 Y_t$$

5. A LG Process where the stock price is convex in the growth rate of dividend

$$\left. \begin{aligned} \chi_t &= e^{-rt} \\ \frac{dD_t}{D_t} &= g_t dt + \sigma dW_1 \\ dg_t &= - (g_t^2 / 2 + \phi g_t) dt + \sqrt{k (G^2 - g_t^2)} dW_2 \end{aligned} \right\} \implies \frac{P_t}{D_t} = b_0 + b_1 g_t + b_2 g_t^2$$

Examples in the Paper

6. A Multifactor Bond Model

$$\left. \begin{aligned} d\chi_t/\chi_t &= -r_t dt + dN_t \\ r_t &= r_* + \sum_{i=1}^n r_{it} \\ E[dr_{it}] + \langle dr_{it}, dM_t/M_t \rangle &= [-\phi_i r_{i,t} + (r_t - r_*) r_t] \end{aligned} \right\} \implies$$

$$\implies B_t(T) = e^{-r_* T} \left(1 - \sum_{i=1}^n \frac{1 - e^{-\phi_i T}}{\phi_i} r_{i,t} \right)$$

7. A one factor bond model with always positive nominal rate.

$$\left. \begin{aligned} \chi_t &= e^{-\int_0^t r_s ds} \\ r_t &= r_* + \hat{r}_t \\ d\hat{r}_t &= -(\phi - \hat{r}_t) \hat{r}_t + dN_t \end{aligned} \right\} \implies B_t(T) = e^{-r_* T} \left(1 - \frac{1 - e^{-\phi T}}{\phi} \hat{r}_t \right)$$

8. A model in the spirit of Brennan and Schwartz, where the factors are the term rate and the perpetuity rate.

...

9. r_t having a time trend.

...

The LG-Twist: in the Drift or in the Covariance?

- Learning model is a special case of Gabaix (2009) where state variables are π_t .
- Differently from Gabaix (2009), state variables **need not** have quadratic terms in their drifts (so, no LG-twist in drift) to obtain linear model. E.g.

$$d\pi_t = \Lambda' \cdot \pi_t dt + \pi_t \odot (\theta - \mathbf{1}_n \pi_t' \cdot \theta) \mathbf{h} d\widetilde{\mathbf{W}}_t$$

- Why is this happening?
 - In Learning Models, there is a natural covariance between shocks to dividends and the “state variables” (i.e. beliefs) π_t .
 - It turns out that this covariance kills the quadratic terms stemming from the application of Ito’s lemma to $\eta_{i,t} = C_t^{-\gamma} \pi_t^i$.
 - Thus $\eta_{i,t} = C^{-\gamma} \pi_t^i$ has a linear drift, which is the defining characteristics of LG models.
- Bottom Line: The existence of the quadratic term in LG models may or may not be there, depending on the specification of some covariance terms.

The LG-Twist: in the Drift or in the Covariance?

- This argument is easy to see also in the simplest model of Gabaix (2009).

- Consider again the Generalized Gordon Formula

$$dD_t/D_t = (g_* + \gamma_t)dt + \sigma dW_{1,t}$$

- Assume now a constant discount rate R and that

$$d\gamma_t = \mu(\gamma) dt + \sigma_1(\gamma_t) dW_{1,t} + \sigma_2(\gamma_t) dW_{2,t}$$

- The pricing function is $P_t = E \left[\int_t^\infty e^{-R(u-t)} D_u du \right]$

- Let

$$V(D_t, \gamma_t, t) = E \left[e^{-R(u-t)} D_u \right]$$

- Then Feynman Kac Theorem states that $V(D_t, \gamma_t, t)$ is the solution to

$$RV = V_t + V_D E[dD] + V_\gamma E[d\gamma] + \frac{1}{2} V_{DD} E[dD^2] + \frac{1}{2} V_{\gamma\gamma} E[d\gamma^2] + V_{\gamma D} E[dD d\gamma]$$

- Conjecture a linear solution

$$V(D, \gamma, t) = D(A(t) + B(t)\gamma_t)$$

with $A(u) = 1$ and $B(u) = 0$

The LG-Twist: in the Drift or in the Covariance?

- Then, some computations lead to

$$R(A(t) + B(t)\gamma_t) = \left(\frac{\partial A(t)}{\partial t} + \frac{\partial B(t)}{\partial t}\gamma_t \right) + A(t)(g_* + \gamma_t) + B(t)\gamma_t g_* + B(t) [\gamma_t^2 + \mu(\gamma) + \sigma_1(\gamma)\sigma_D]$$

- The first row has only linear terms, but the second row has one quadratic term.
- How can we kill it? We only need $\mu(\gamma)$ and $\sigma_1(\gamma)$ such that

$$\gamma_t^2 + \mu(\gamma) + \sigma_1(\gamma)\sigma_D = a + b\gamma_t$$

- This can be achieved by assuming

$$\begin{aligned} \text{Gabaix : } & \mu(\gamma) = -\phi\gamma - \gamma^2 \quad \text{and} \quad \sigma_1(\gamma) = \alpha_1 + \alpha_2\gamma_t \\ \text{Alternative : } & \mu(\gamma) = -\phi\gamma; \quad \text{and} \quad \sigma_1(\gamma)\sigma_D = \alpha_1 + \alpha_2\gamma_t - \gamma_t^2 \end{aligned}$$

- Either way, there is a strong restriction to be made.
- As shown earlier, the latter case naturally arises in Learning Models
- Either solution $\implies P_t = D_t(A + B\gamma_t)$

Overconfidence and Speculative Bubbles

- Scheinkman and Xiong (JPE, 2003) follow Harrison and Kreps (QJE, 1978) to show that biased differences of beliefs may generate high prices.
- What is a bubble in this context?
 - It is a component of the price that is not related to fundamentals but rather on the (speculative) expectation of being able to resell the stock to somebody else for a higher price later.
- There is a stock which with dividend process

$$dD_t = f_t dt + \sigma_D dZ_t^D$$

where

$$df_t = \lambda(\bar{f} - f_t)dt + \sigma_f dZ_t^f$$

- f_t is not observable.
- There are two types of risk neutral investors A and B, which differ only in their beliefs about f_t . In particular, their beliefs are generated by the following signal

$$\begin{aligned} ds_t^A &= f_t dt + \sigma_s dZ_t^A \\ ds_t^B &= f_t dt + \sigma_s dZ_t^B \end{aligned}$$

Overconfidence

- Key assumptions:

1. All agents observe all signals, but A thinks ds_t^A is the right signal, while B thinks ds_t^B is the right signal.
2. In addition, each group think that the informativeness of each own signal is larger than what really is. That is:

- A believes

$$ds_t^A = f_t dt + \sigma_s \phi dZ_t^f + \sigma_s \sqrt{1 - \phi^2} dZ_t^A$$

- B believes

$$ds_t^B = f_t dt + \sigma_s \phi dZ_t^f + \sigma_s \sqrt{1 - \phi^2} dZ_t^B$$

- \implies Implication: A overreact to signal ds_t^A , while B overreact to signal ds_t^B .

3. There are short sale constraints.

- Why do we need overconfidence and not only differences of opinion?

- The key is to generate *time varying* differences of opinion.
- Otherwise, the resale option value and trading volume are zero.

Differences of Opinion

- We can apply the usual filtering algorithm for each class of agents.
- We have the following two processes

$$d\hat{f}_t^A = \lambda(\bar{f} - f_t^A)dt + \frac{\phi\sigma_s\sigma_f + \gamma}{\sigma_s^2}(ds_t^A - \hat{f}_t^A dt) \\ + \frac{\gamma}{\sigma_s^2}(ds_t^B - \hat{f}_t^A dt) + \frac{\gamma}{\sigma_D^2}(dD - f_t^A dt)$$

$$d\hat{f}_t^B = \lambda(\bar{f} - f_t^B)dt + \frac{\phi\sigma_s\sigma_f + \gamma}{\sigma_s^2}(ds_t^B - \hat{f}_t^B dt) \\ + \frac{\gamma}{\sigma_s^2}(ds_t^A - \hat{f}_t^B dt) + \frac{\gamma}{\sigma_D^2}(dD - f_t^B dt)$$

- where γ is the posterior stationary variance $E[(f^A - \hat{f}^A)^2] = E[(f^B - \hat{f}^B)^2]$. It is the solution to the usual riccati equation under the assumption $t \rightarrow \infty$

$$\gamma = \frac{\sqrt{(\lambda + \phi\sigma_f/\sigma_s)^2 + (1 - \phi^2)(2\sigma_f^2/\sigma_s^2 + \sigma_f^2/\sigma_D^2)} - (\lambda + \phi\sigma_f/\sigma_s)}{\frac{1}{\sigma_D^2} + \frac{1}{\sigma_s^2}}$$

Differences of Opinion

- Define the difference of opinion variable $g_t^A = \hat{f}^A - \hat{f}^B$. It immediately follows that one can write

$$dg_t^A = -\rho g_t^A + \sigma_g dW_t^A$$

where dW_t^A is a Wiener process orthogonal to innovations to \hat{f}^A , and

$$\rho = \sqrt{(\lambda + \phi\sigma_f/\sigma_s)^2 + (1 - \phi^2)\sigma_f^2(2/\sigma_s^2 + 1/\sigma_D^2)}; \quad \sigma_g = \sqrt{2}\phi\sigma_f$$

- Differences of beliefs are time varying, and mean reverting around zero.
- It is convenient to also consider the difference in beliefs $g_t^B = \hat{f}^B - \hat{f}^A$. It immediately follows that one can write

$$dg_t^B = -\rho g_t^B + \sigma_g dW_t^B$$

where dW_t^B is a Wiener process orthogonal to innovations to \hat{f}^B .

Trading

- Consider now the personal valuation of the asset from each group perspective.
- Assume that there is a trading cost c anytime the holder of the asset sell it to somebody else.
- The value of the asset to whoever holds (owner) it is (for $o \in \{A, B\}$):

$$p_t^o = \max_{\tau \geq 0} E^o \left[\int_t^{t+\tau} e^{-r(s-t)} dD_s + e^{-r\tau} (p_{t+\tau}^{\bar{o}} - c) \right]$$

- where $p_{t+\tau}^{\bar{o}}$ is the valuation of the buyer.
- Given the mean reverting process for f_t (under the o filtration) we have

$$\int_t^{t+\tau} e^{-r(s-t)} dD_s = \int_t^{t+\tau} e^{-r(s-t)} [\bar{f} + e^{-\lambda(s-t)} (\hat{f}_t^o - \bar{f})] ds + M_{t+\tau}$$

with $E^o[M_{t+\tau}] = 0$

$$p_t^o = \max_{\tau \geq 0} E^o \left[\int_t^{t+\tau} e^{-r(s-t)} [\bar{f} + e^{-\lambda(s-t)} (\hat{f}_t^o - \bar{f})] ds + e^{-r\tau} (p_{t+\tau}^{\bar{o}} - c) \right]$$

Solution

- Postulate the following form for the solution

$$p_t^o = p(\hat{f}^o, g^o) = \frac{\bar{f}}{r} + \frac{\hat{f}_t^o - \bar{f}}{r + \lambda} + q(g_t^o)$$

- Substitute in the pricing equation for $p_{t+\tau}^o$ and rearrange (intermediate step: rewrite the integral as $\int_t^{t+\tau} = \int_t^\infty - \int_{t+\tau}^\infty$).

$$p_t^o = \frac{\bar{f}}{r} + \frac{\hat{f}_t^o - \bar{f}}{r + \lambda} + \max_{\tau \geq 0} E \left[e^{-r\tau} \left(\frac{g_{t+\tau}^o}{r + \lambda} + q(g_{t+\tau}^o) - c \right) \right]$$

- The last term is the *resale option*, which then satisfies

$$q(g_t^o) = \max_{\tau \geq 0} E \left[e^{-r\tau} \left(\frac{g_{t+\tau}^o}{r + \lambda} + q(g_{t+\tau}^o) - c \right) \right]$$

Solution

- This is a relatively standard American option pricing problem.
- First, the option value should always be greater than or equal to intrinsic value

$$q(x) \geq \frac{x}{r + \lambda} + q(-x) - c$$

where we use $g_t^o = -g_t^{\bar{o}}$.

- Second, in the continuation region (i.e. conditional on not exercising the option), we should have the martingale condition $E[dq] = r q dt$

$$q' \times (-\rho x) + q'' \times \frac{\sigma_g^2}{2} - r q = 0$$

- Finally, at exercise, $q(x) = \frac{x}{r + \lambda} + q(-x) - c$
- Scheinkman and Xiong obtain an explicit solution to the ODE on the continuation region $(-\infty, k^*)$.
 - k^* is the minimum difference of opinion that is necessary to generate a trade.

Solution

- We have

$$q(x) = \begin{cases} \frac{b}{h(-k^*)}h(x) & \text{if } x < k^* \\ \frac{x}{r+\lambda} + h(-x) - c & \text{if } x \geq k^* \end{cases}$$

- where

$$h(x) = \begin{cases} U\left(\frac{r}{2\rho}, \frac{1}{2}, \frac{\rho}{\sigma_g^2}x^2\right) & \text{if } x < 0 \\ \frac{2\pi}{\Gamma(1/2+r/2\rho)\Gamma(1/2)}M\left(\frac{r}{2\rho}, \frac{1}{2}, \frac{\rho}{\sigma_g^2}x^2\right) - U\left(\frac{r}{2\rho}, \frac{1}{2}, \frac{\rho}{\sigma_g^2}x^2\right) & \text{if } x > 0 \end{cases}$$

where U and M are two Kummer functions, and Γ is the Gamma distribution.

- k^* satisfy the equation

$$[k^* + c(r + \lambda)](h'(k^*) + h'(-k^*)) - h(k^*) + h(-k^*) = 0$$

Implications

- Bubble: Value in the mind of the owner is higher than the PV of future dividends because of resale option:

$$p_t^o = \frac{\bar{f}}{r} + \frac{\hat{f}_t^o - \bar{f}}{r + \lambda} + q(g_t^o)$$

- Additional volatility driven by the resale option

$$\sigma(dp) = \frac{1}{r + \lambda} \sigma(df^o) + g' \sigma(dg_t^o)$$

- Trading “volume:” There is trading anytime the difference of beliefs reach k^* .
 - When trading occurs, the new owner beliefs are such that

$$g^{\text{newowner}} = -g^{\text{oldowner}}$$

so the relevant “ g ” is reset to a lower value. Next trading is when new g crosses again k^* .

- If cost $c \rightarrow 0$ then $k^* \rightarrow 0$, and trading becomes more and more frequent.

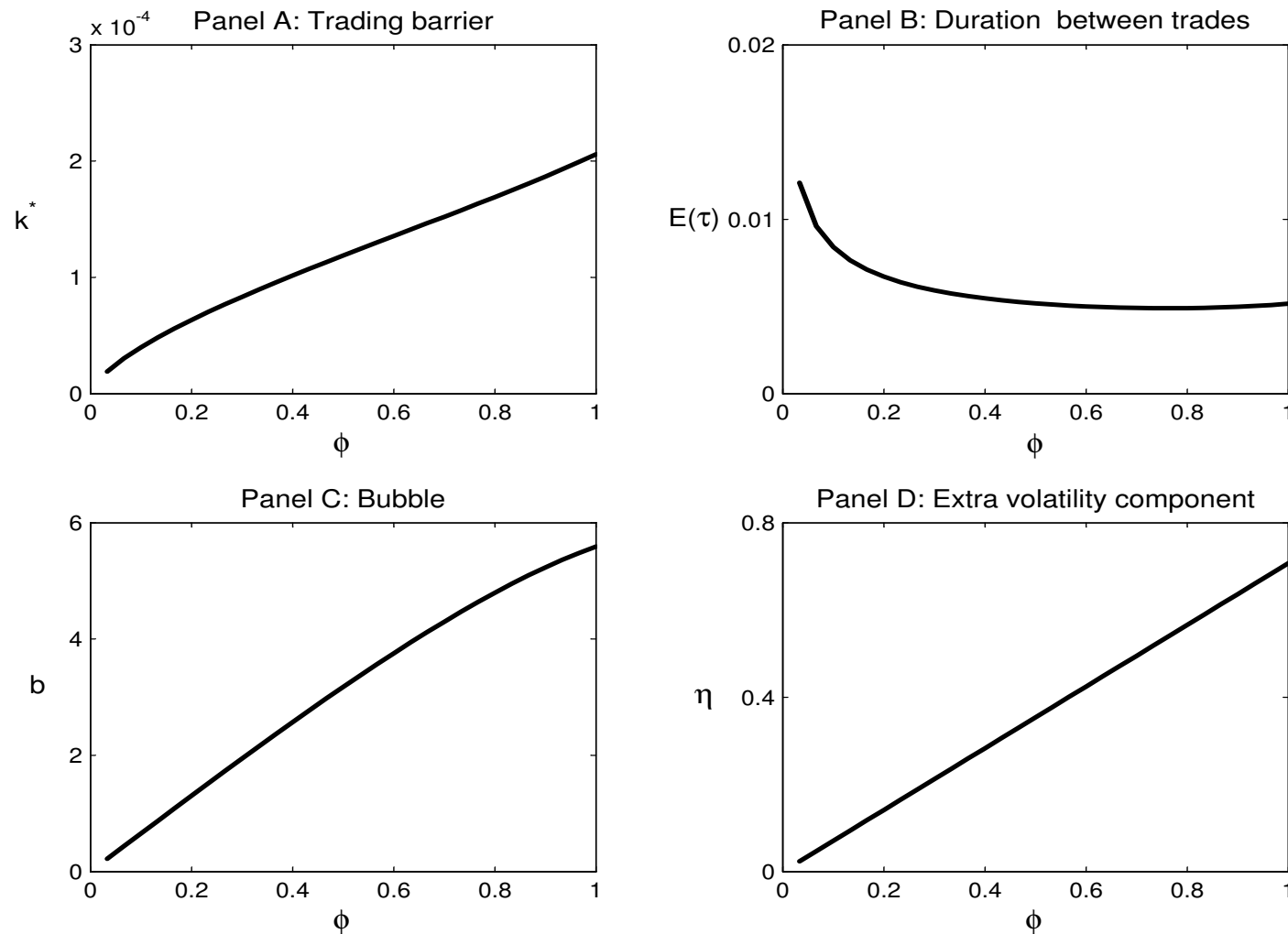


Figure 1: Effect of overconfidence level. Here, $r = 5\%$, $\lambda = 0$, $\theta = 0.1$, $i_s = 2.0$, $i_D = 0$, and $c = 10^{-6}$. The values of the bubble and the extra volatility component are computed at the trading point. The trading barrier, the bubble and the extra volatility component are measured as multiples of $\frac{\sigma_f}{r+\lambda}$, the fundamental volatility of the asset.

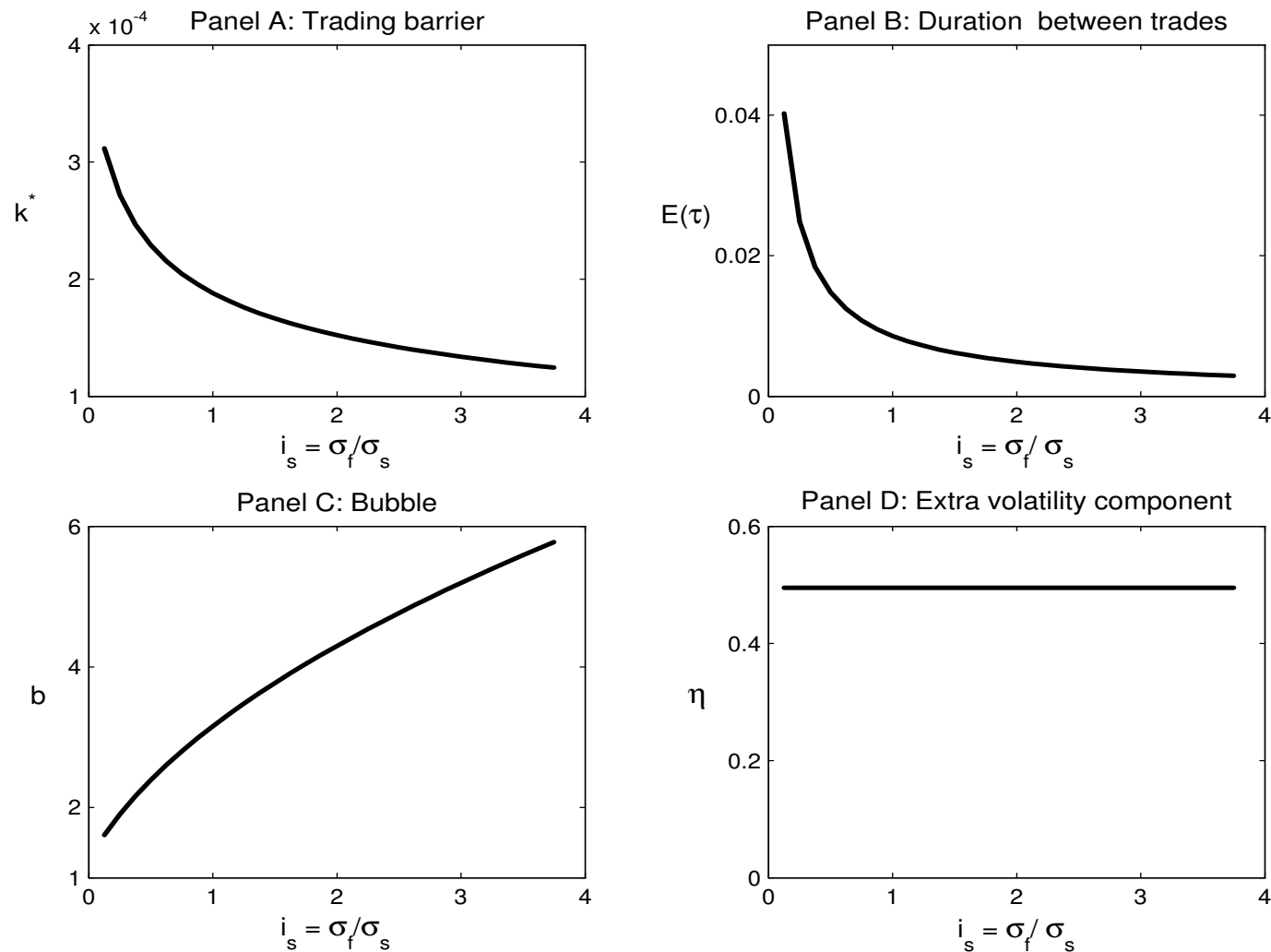


Figure 2: Effect of information in signals. Here, $r = 5\%$, $\lambda = 0$, $\theta = 0.1$, $\phi = 0.7$, $i_D = 0$, and $c = 10^{-6}$. The values of the bubble and the extra volatility component are computed at the trading point. The trading barrier, the bubble and the extra volatility component are measured as multiples of $\frac{\sigma_f}{r+\lambda}$, the fundamental volatility of the asset.

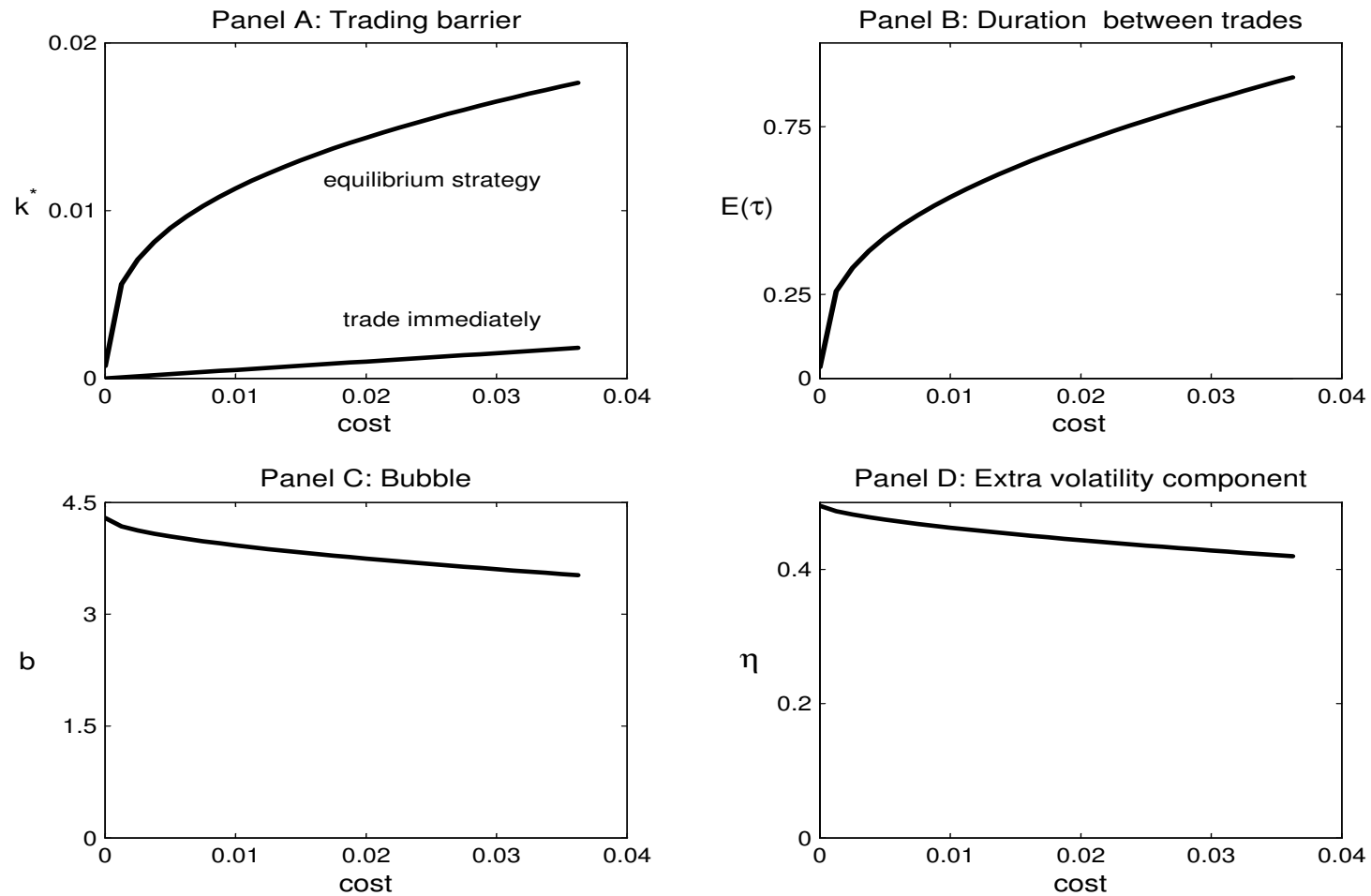


Figure 3: Effect of trading costs. Here, $r = 5\%$, $\phi = 0.7$, $\lambda = 0$, $\theta = 0.1$, $i_s = 2.0$, and $i_D = 0$. The values of the bubble and the extra volatility component are computed at the trading point. The trading barrier, the bubble, the extra volatility component, and trading cost are measured as multiples of $\frac{\sigma_f}{r+\lambda}$, the fundamental volatility of the asset.