On the AI Bubble

Presenter: Tomohiro Hirano (Research Fellow, CIGS), tomohih@gmail.com

- Tomohiro Hirano, Keiichi Kishi, and Alexis Akira Toda 2026 "General-Purpose Technologies and Stock Market Bubbles" arXiv:2501.08215
- Tomohiro Hirano, and Shunsuke Hori 2026 "Overvalued Stock Prices and the AI Bubble"

Motivations

- An asset price bubble is a situation in which the asset price exceeds its fundamental value defined by the present value of dividends.
- Looking back at history, major economic crises have invariably been preceded by the emergence and collapse of speculative bubbles (See Kindleberger's book 2000).
- Therefore, there is a substantial interest among policymakers, academics, and the general public in understanding why and how asset price bubbles emerge in the first place.

Motivations (conti.)

- In contrast, in modern macro-finance theory, there is a common belief that "bubbles cannot possibly occur, and asset prices should reflect the fundamentals."
- Indeed, it is well known that there is a fundamental difficulty in generating asset bubbles in dividend-paying assets (real assets, Lucas trees), such as land, stocks, and housing.
- This is known as Santos-Woodford Bubble Impossibility Theorem (Santos and Woodford 1997 ECMA). Their Theorem 3.3 and Corollary 3.4. state

when dividends from assets comprise a non-negligible share of the aggregate endowment

---no matter how small the share is---bubbles are impossible.

Note 1: S-W prove Bubble Impossibility Theorem in deterministic settings.

Note 2: We present a model that shows the inevitable emergence of stock bubbles in a setting with aggregate uncertainty, i.e., despite the expectation that stock bubbles are expected to collapse.

Main messages in today's presentation

1. Firstly, the advent of General Purpose Technologies (GPTs such as electricity, IT, and AI) has a positive spreading effect on the entire economy. During this process, stock price bubbles inevitably emerge, despite the expectation that they may collapse at some point in the future. Once the positive spreading effects begin to diminish, the bubbles collapse.

2. Secondly, the diffusion process of GPTs generates unbalanced growth, with different production factors growing at different rates. As a result, the macroeconomy temporarily sets off on a different dynamic path from a balanced growth path.

3. Thirdly, stock bubbles are characterized by being attached to assets with high growth rates rather than to assets with lower productivity growth, such as land.

4. Simultaneous bubbles in stocks and land: As a result of the advent of GPTs, at the same time when a bubble occurs in stocks with high productivity growth rates, a bubble also arises in land with lower productivity growth rates.

Definition of asset price bubbles (deterministic case)

- $P_t = \frac{D_{t+1}}{1+r_t} + \frac{P_{t+1}}{1+r_t}$ (1)

Iterating (1) forward yields

- $P_t = \underbrace{\sum_{s=1}^{N} \left(\frac{D_{t+j+1}}{\prod_{j=0}^{s-1} 1+r_{t+j}} \right)}_{future\ discounted\ dividends} + \underbrace{\frac{P_{t+N}}{\prod_{j=0}^{N-1} 1+r_{t+j}}}_{future\ discounted\ prices}$ (2)

Letting $N \to \infty$, the fundamental value of the asset is defined as

- $V_t = \underbrace{\sum_{s=1}^{\infty} \left(\frac{D_{t+j+1}}{\prod_{j=0}^{s-1} 1+r_{t+j}} \right)}_{future\ discounted\ dividends}$ (3)

An asset price bubble is defined as

- $B_t = \lim_{N\to\infty} \frac{P_{t+N}}{\prod_{j=0}^{N-1} 1+r_{t+j}} \geq 0$ (4)

$P_t = V_t$ if and only if $B_t = 0$.

$$B_t = \lim_{N\to\infty} \frac{P_{t+N}}{\prod_{j=0}^{N-1} 1+r_{t+j}} \geq 0 \quad (4)$$

Three remarks are in order.

1. Economic meaning of B_t. B_t captures a purely speculative aspect backed by nothing, i.e., one buy assets today for the purpose of resale in the future, rather than for the purpose of receiving dividends.

2. When $B_t = 0$, asset prices are determined only by factors that are backed in equilibrium, namely future dividends

3. Santos-Woodford "Bubble Impossibility Theorem" claims that there are fundamental difficulties in generating $B_t > 0$ for assets with $D_t > 0$ in a general equilibrium model.

As we will see, the dynamics of the price-dividend ratio changes markedly between $B_t > 0$ and $B_t = 0$.

→ possible to build a bridge to the bubble detection literature

Related literature: Due to S-W Theorem, the literature has almost exclusively focused on "pure bubbles", i.e., assets that pay no dividend like money.

Common criticisms

1. unrealistic: pure bubble assets other than money/cryptocurrencies are hard to find in reality

2. difficult to apply the theory for empirical or quantitative analysis

3. impossible to build a bridge to bubble detection literature using P-D ratio

4. suffer from equilibrium indeterminacy, i.e., a continuum of bubble equilibria→ select only one of them→ model prediction is fragile (e.g., Jordi Gali 2014 AER paper)

5. pure bubbles are likely to arise when financial conditions get tighter, which contradicts stylized facts.

6. impossible to explain how the bubble starts and collapses

→1-6 simply show that pure bubble models are fundamentally difficult for applications.

Our simple model overcomes all of these criticisms.

A simple model of stock market bubbles that are expected to collapse

- Consider an OLG model with workers (size N) and entrepreneurs (size N').
- There are two sectors, investment and consumption good sectors.
- Each worker has a unit of labor endowment (only when young), supplies it inelastically in the labor market and earns wage w_t.
- Each entrepreneur has an entrepreneurial ability of starting up a business using a unit of his/her labor (only when young), while workers do not.
- Occupational choice: Each entrepreneur decides whether to start up a business or work in the labor market.

Model (conti.)

- When an entrepreneur starts up a business, his/her labor input at date t produces A_t units of technology (AI service) at date t.

- Technology (AI service) created at date t can be sold at the price of P_t in terms of consumption goods at date t.

→ P_t can be interpreted as the share price of a company providing AI services.

- Hence, an entrepreneur receives the initial wealth of $P_t A_t$ by initial public offering (IPO).

- Technology (AI service) produced at date t will be available for production from the next period $t+1$ onwards.

- Technology (AI service) depreciates by δ every period.

- The maximization problem of each young person is given by

$$Max_{c_{t+1},k_{t+1}} \quad u_t = E_t[c_{t+1}] \quad (5)$$

$$s.t \quad P_t k_{t+1} = e_t \quad (6)$$

$$c_{t+1} = (D_{t+1} + P_{t+1})k_{t+1} \quad (7)$$

where $e_t = w_t$ or $P_t A_t$, or a combination of them if labor is divisible. k_{t+1} is the amount of technology holding (AI related stocks) at date t, and P_t is the stock price at date t, and D_t is dividend technology (AI service) generates at date t.

- In AI services, D_t corresponds to subscription fees, AI usage fees, etc. The AI company profits from that.

- From the maximization problem,

$$u_t = E_t[c_{t+1}] = E_t\left[\frac{D_{t+1}+P_{t+1}}{P_t}\right] e_t \equiv R_t e_t \quad (8)$$

R_t captures how society trades off resources between dates t and $t+1$. Because hypothetically if a social planner took resources from a young agent at date t, (s)he would require R_t units of consumption goods to maintain the same expected utility.

Consumption-good sector

- The production function of the consumption-good sector

$$Y_t = \left(\alpha(Z_t K_t)^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(H_t M_t)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \quad (9)$$

where M_t is labor employed in the consumption-good sector. Z_t and H_t are productivities of technology (AI services) and labor, respectively.

- The wage rate is

$$w_t = \left(\alpha(Z_t K_t)^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(H_t M_t)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{\sigma-1}} (1-\alpha)(H_t M_t)^{\frac{-1}{\sigma}} H_t \quad (10)$$

- Dividend technology (AI service) yields is

$$D_t = \left(\alpha(Z_t K_t)^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(H_t L_t)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{\sigma-1}} \alpha(Z_t K_t)^{\frac{-1}{\sigma}} Z_t \quad (11)$$

Assumption: we focus on the case of $\sigma < 1$. Technology (AI service) and labor complement each other. $\sigma < 1$ is empirically plausible.

- **The evolution of stock of technology (AI services)**

$K_{t+1} = I_t + (1-\delta)K_t$, with $I_t = A_t \eta_t N'$, (12)

where I_t is net investment at date t and $\eta_t \in [0,1]$ is the proportion of entrepreneurs who start up a business at date t and will be determined endogenously in equilibrium.

- **The labor market clearing condition**

$N + (1-\eta_t)N' = M_t$. (13)

The proportion $1-\eta_t$ of entrepreneurs work in the labor market. Hence, the left and right-hand sides are aggregate labor supply and demand, respectively.

- **The consumption-goods market clearing condition**: $C_t = Y_t$. (14)

C_t is aggregate consumption of young and old generation at date t.

- $\boldsymbol{GDP_t} = Y_t + P_t A_t \eta_t N' = C_t + P_t I_t$ (15)

Asset market equilibrium

- All savings finance the existing and new stocks.

$$P_t(I_t + (1-\delta)K_t) = w_t N + w_t(1-\eta_t)N' + P_t A_t \eta_t N' \quad (16)$$

The right-and left-hand sides are aggregate savings and the total value of stocks at date t.

- Solving for P_t yields

$$P_t = \frac{w_t(N+(1-\eta_t)N')}{(1-\delta)K_t}. \quad (17)$$

Hence, the initial wealth entrepreneurs receive by IPOs is

$$P_t A_t = \frac{w_t(N+(1-\eta_t)N')A_t}{(1-\delta)K_t}. \quad (18)$$

- Comparing $P_t A_t$ to w_t, entrepreneurs decide whether to start up a business or work in the labor market. That is,

$$P_t A_t = \frac{w_t(N+(1-\eta_t)N')A_t}{(1-\delta)K_t} > or \ = or \ < w_t \quad (19)$$

→ $\eta_t \in [0,1]$ will be endogenously determined.

→ If the solution is interior, $\eta_t = \frac{A_t(N+N')-(1-\delta)K_t}{A_t N'}$. (20)

- As in endogenous growth theory, there are positive spillover effects of technology stocks (AI services) on the productivity of entrepreneurs starting up their businesses.

$A_t = K_t^{\phi_t}$ (21)

→ ϕ_t captures the strength of the positive spillover effects.

- We also assume.

$Z_t = K_t^{\epsilon_t}$ and $H_t = K_t^{\mu_t}$ (22)

→ Increases in K_t (stock of AI services) generates the positive spreading spillover effects on overall productivities (of technology and labor).

Introducing aggregate uncertainty

- With probability π, $\phi_t = \phi^H \geq 1, \epsilon_t = \epsilon^H, \mu_t = \mu^H$ (23)

We assume $1 + \epsilon^H - \mu^H > 0$. → generates unbalanced growth where different production factors grow at different rates.

- With probability $1 - \pi$, $\phi_t = \phi^L < 1, \epsilon_t = \epsilon^L, \mu_t = \mu^L$. (24)

$\phi^L < 1$ → the positive spillover effect begins to diminish→converges to balanced growth

We assume that state with $\phi_t = \phi^L$ is an absorbing state.

What is ϕ_t in AI?

Through data network effect, we have positive loops.

1. More Users: As more people use a product, they generate more data.
2. More Data: This data is fed back into the machine learning models.
3. Better Models: The AI learns from the new patterns, edge cases, and user feedback, becoming more accurate and efficient.
4. Better Product: A smarter product attracts even more users, and the cycle repeats.

- So long as probability π persists, $\phi_t = \phi^H \geq 1$. The AI data network effect works exponentially or super-exponentially, generating economic growth.
- Once probability $1 - \pi$ arises, $\phi_t = \phi^L < 1$. Those diffusion effects begin to diminish.

→ Growth is expected to come to a halt.

e.g., increased computational cost (GPU, energy), data saturation in terms of quantity and quality.

- The no-arbitrage equation conditional on state being UG (unbalanced growth) at date *t*,

$$R_t^{UG} = \frac{\pi D_{t+1}^{UG}}{P_t^{UG}} + \frac{(1-\pi) D_{t+1}^{BG}}{P_t^{UG}} + \frac{\pi P_{t+1}^{UG}}{P_t^{UG}} + \frac{(1-\pi) P_{t+1}^{BG}}{P_t^{UG}} \quad (25)$$

● BG denotes balanced growth

- Solving this for P_t^{UG} and then solving it forward yields

$$P_t^{UG} = \frac{\pi D_{t+1}^{UG}}{R_t^{UG}} + \frac{\pi^2 D_{t+2}^{UG}}{R_t^{UG} R_{t+1}^{UG}} + \frac{\pi^3 D_{t+3}^{UG}}{R_t^{UG} R_{t+1}^{UG} R_{t+2}^{UG}} + \cdots$$
$$+ (1-\pi)\left[\frac{D_{t+1}^{BG} + P_{t+1}^{BG}}{R_t^{UG}} + \frac{\pi\left(D_{t+1}^{BG} + P_{t+1}^{BG}\right)}{R_t^{UG} R_{t+1}^{UG}} + \frac{\pi^2\left(D_{t+1}^{BG} + P_{t+1}^{BG}\right)}{R_t^{UG} R_{t+1}^{UG} R_{t+2}^{UG}} + \cdots\right] + \frac{\pi^n P_{t+n}^{UG}}{R_t^{UG} R_{t+1}^{UG} R_{t+2}^{UG} \cdots R_{t+n-1}^{UG}} \quad (26)$$

- The fundamental value at date *t* is equal to

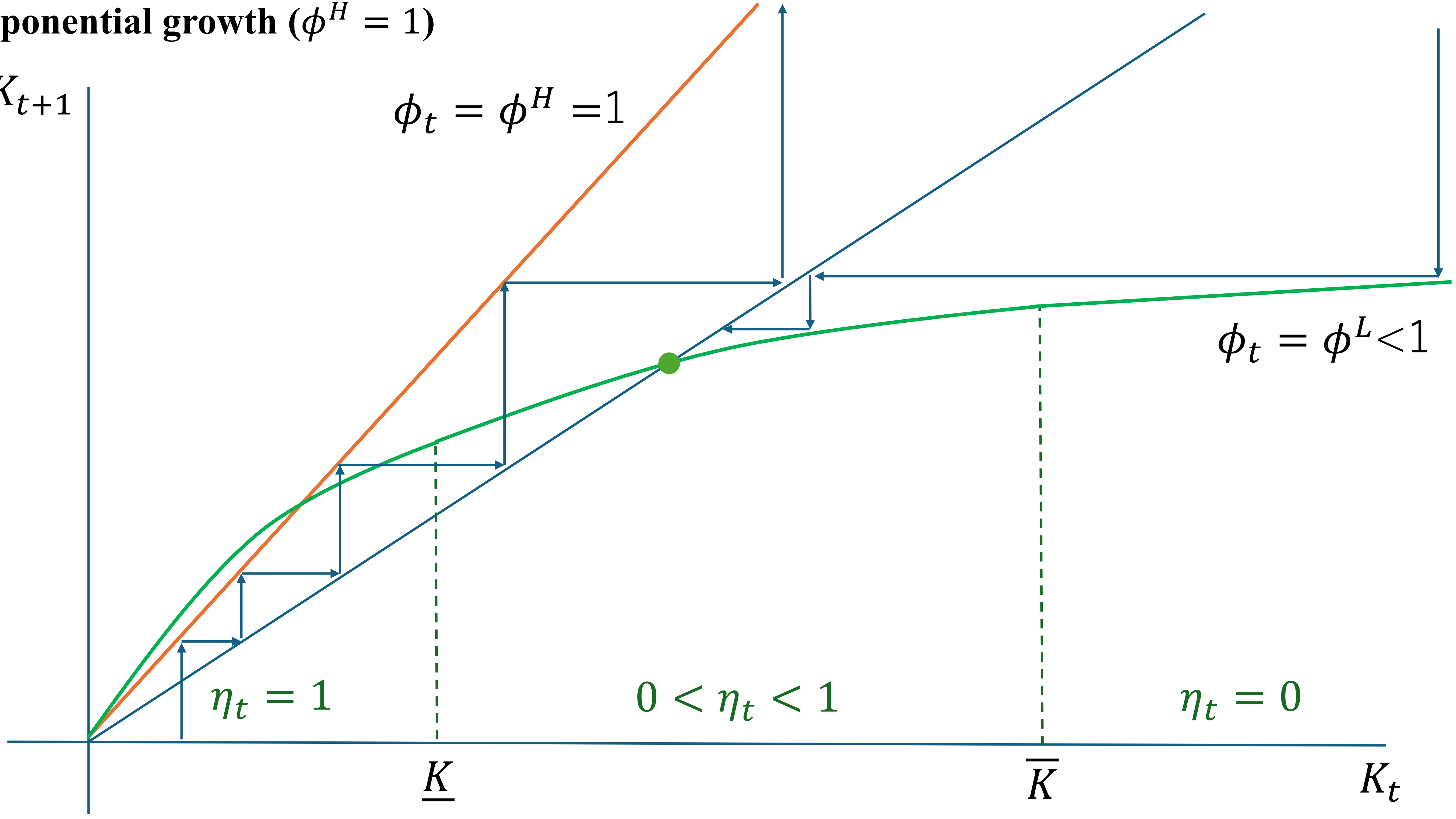
$$V_t^{UG} = \frac{\pi D_{t+1}^{UG}}{R_t^{UG}} + \frac{\pi^2 D_{t+2}^{UG}}{R_t^{UG} R_{t+1}^{UG}} + \frac{\pi^3 D_{t+3}^{UG}}{R_t^{UG} R_{t+1}^{UG} R_{t+2}^{UG}} + \cdots$$
$$+ (1-\pi)\left[\frac{D_{t+1}^{BG} + P_{t+1}^{BG}}{R_t^{UG}} + \frac{\pi\left(D_{t+1}^{BG} + P_{t+1}^{BG}\right)}{R_t^{UG} R_{t+1}^{UG}} + \frac{\pi^2\left(D_{t+1}^{BG} + P_{t+1}^{BG}\right)}{R_t^{UG} R_{t+1}^{UG} R_{t+2}^{UG}} + \cdots\right] \quad (27)$$

- An asset bubble is defined as

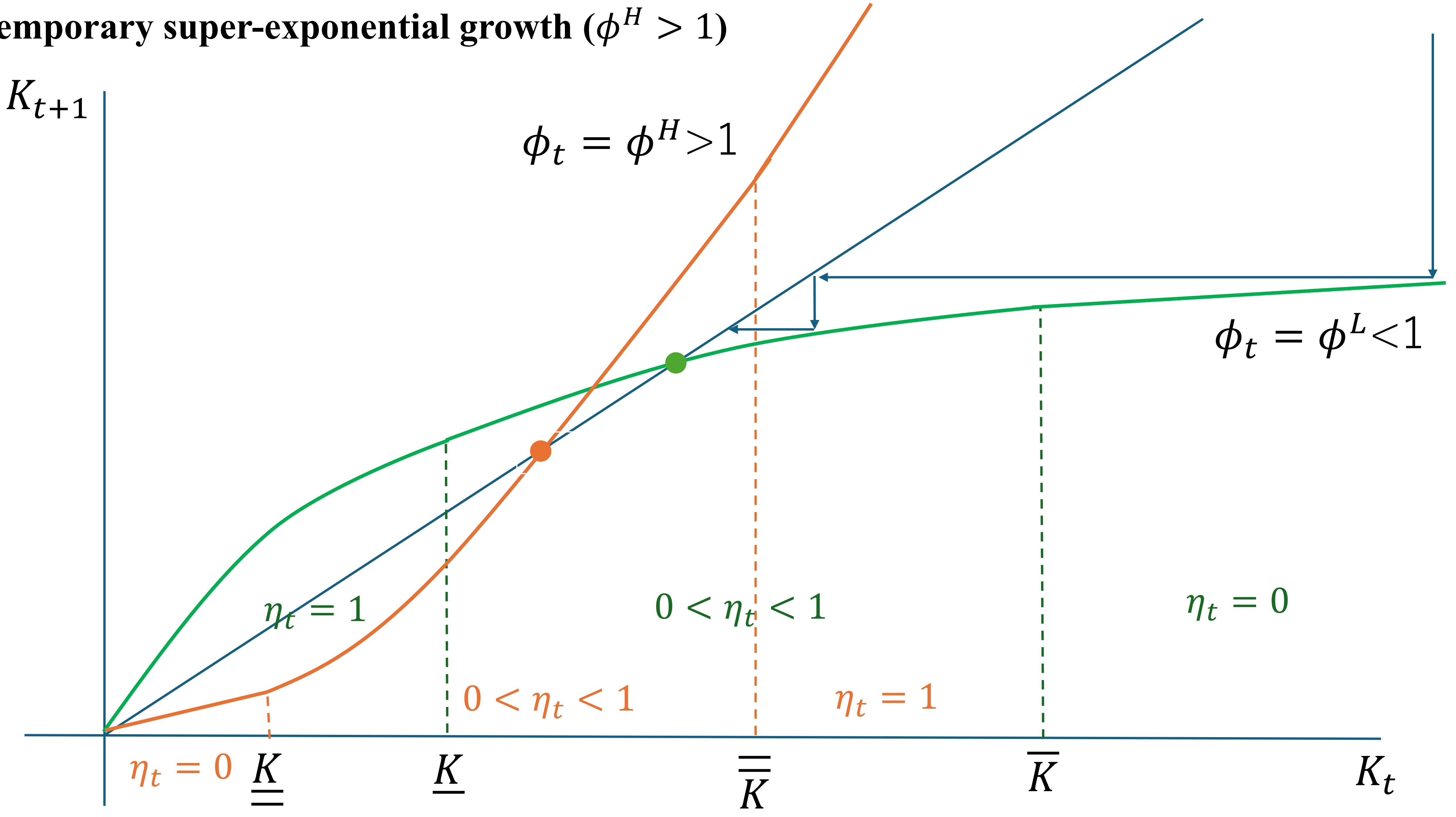
$$B_t \equiv \lim_{n\to\infty} \frac{\pi^n P_{t+n}^{UG}}{R_t^{UG} R_{t+1}^{UG} R_{t+2}^{UG} \cdots R_{t+n-1}^{UG}} \geq 0 \quad (28)$$

Hence, $P_t^{UG} = V_t^{UG}$ if and only if $B_t = 0$.

Exponential growth ($\phi^H = 1$)



Temporary super-exponential growth ($\phi^H > 1$)



Stock price bubbles inevitably emerge if and only if conditions 1 and 2 are simultaneously satisfied, despite the expectation that they may collapse at some point in the future:

- Condition 1: $(1 + \epsilon - \mu)(1 - \sigma) > 0$
- Condition 2: $1 + g_t > 1$

→ Given $\sigma < 1$, $1 + \epsilon^H - \mu^H > 0$, the condition that generates unbalanced growth, is crucial.

→ with $\epsilon^H > 0$ (capital-augmenting technological progress), unbalanced growth occurs.

Note: "Uzawa Steady-State Growth Theorem" claims that to obtain BGP, ϵ^H has to be zero.

Key features of our model

1. So long as probability π persists, the dynamics of the macroeconomy exhibits "unbalanced growth", where different production factors grow at different rates.

→ the dynamics of asset price bubbles can be seen as a temporary deviation from the balanced growth path where asset prices equal the present discounted value of future dividends.

2. During unbalanced growth dynamics, the P-D ratio exhibits an "explosive" increase.

→ capital gains account for the majority of returns on stocks.

→ provides a micro-foundation for bubble detection literature developed by Phillips and Shi.

3. Bubbles are attached to assets with high rates of productivity growth.

Important implications for macro-theory construction

- Any balanced growth model is knife-edge theory. This is well known. The hope was that the results obtained regarding asset pricing implications and macro-dynamics would maintain generality.
- We show that the slightest deviation from knife-edge cases leads to markedly different macro-dynamics and asset pricing implications.

→ Our construction of the macro-finance model, where unbalanced growth dynamics can temporarily occur, provides a new perspective on the methodology of macro-theory construction.

Note: Our methodology is similar to the one employed by Joseph Stiglitz in establishing economics of asymmetric information. Indeed, we learn this methodology from him.

Simultaneous bubbles in the stock and land (real estate) markets

$$Q_t = \left(\alpha(Z_t K_t)^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(H_t M_t)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} + D_t' X, \quad (29)$$

with $D_t' = K_t^{\lambda}$.

- Assumption: $1 + \epsilon > \mu > \lambda$.

→ AI sector grows faster than land sector.

Then, we can prove

- so long as π persists, simultaneous bubbles in stocks and land occur.

→ funds generated in AI sector are flowing into land sector, resulting in the emergence of land bubbles.

→ At the same time when a bubble occurs in stocks with high productivity growth rates, a bubble also arises in land with lower productivity growth rates.

- once the positive spillover effect begins to diminish at probability $1 - \pi$, both stock and land bubbles collapse simultaneously.