Tech Booms: From Dot-Com To AI*

Dan Cao†
Georgetown University

November 2025

Abstract

How long will rapid AI progress—driven by Transformer-based models such as ChatGPT—continue, and how quickly will it translate into economy-wide productivity gains? I develop a two-sector model to assess what stock valuations and real investment activities reveal about these questions and use it to analyze both the 1990s Dot-Com and the ongoing AI booms. The model features co-integrated productivity processes with spillovers from a high-tech sector to the rest of the economy and endogenous capital accumulation. It predicts that when capital supply in the high-tech sector is inelastic and spillovers are slow, high-tech valuations peak as technological progress decelerates—mirroring the Dot-Com boom, and possibly the current AI boom. But with sufficiently fast spillovers, asset prices keep rising even as high-tech progress slows, because the broader economy expands and reinvests into the high tech sector.

*I am grateful to Harrison Hong for insightful early discussions, and to Behzad Diba, Xiang Ding, Mark Huggett, Jean-Paul L'Huillier, Pascual Restrepo, Estaban Rossi-Hansberg, John Rust, Stephen Terry, and various seminar participants for their helpful comments and suggestions. Yichen Wei provided excellent research assistance.

†Department of Economics, Georgetown University. Email:dan.cao@georgetown.edu

"History never repeats itself, but it does often rhyme."[1]

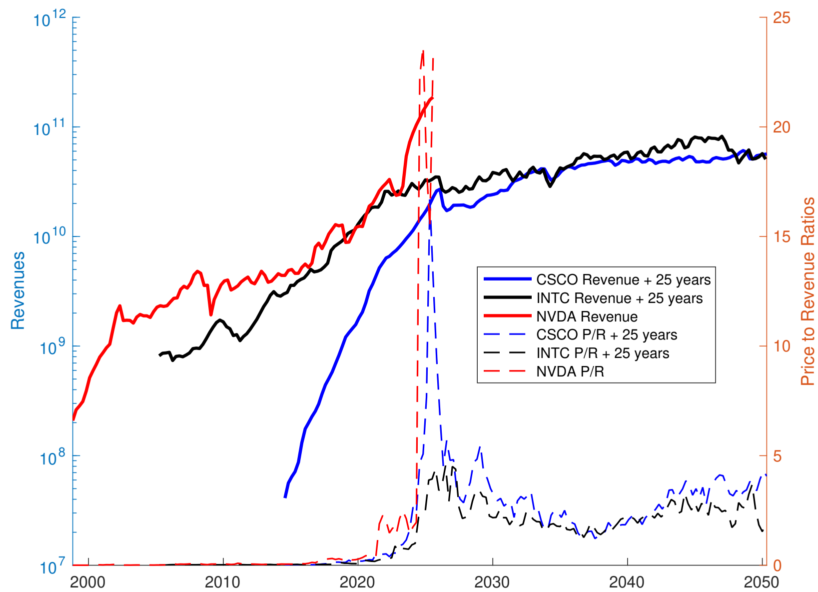
1 Introduction

The introduction of the Transformer model by Google (Vaswani et al., 2023) and ChatGPT by OpenAI in 2022 along with subsequent developments in artificial intelligence (AI), have generated significant excitement about AI while also sparking a major international debate on its potential impact on societies and economies worldwide. The list of major questions about AI is long but in this paper, I focus on two sets of questions: Can the rapid progress in AI be sustained and bring about Artificial General Intelligence (AGI) or even Artificial Super-Intelligence (ASI) in the next few years, or will it slow down soon and we will enter another AI winter like the one in the 1970s? Does the rapid progress in AI in particular and high tech sectors more generally lead to significant economy-wide productivity improvement and if so how long does it take?

I explore these questions by drawing from the experiences of the Internet boom in the 1990s. I argue that equity price and real investment activities in the relevant high-tech sectors could provide useful insights into these questions. Similar to AI, the Internet is a general purpose technology that has revolutionized the way we live and work since the 1990s. The Internet boom was fueled by two major transformations: the massive adoption of personal computers and the spread of the World Wide Web, in the same way that ongoing AI boom being propelled by advances in high performance computing (including Graphics Processing Units, GPUs) and deep neural networks. The upper panel in Figure 1 highlights the similarities by plotting the revenues of the hardware providers in the two booms, Cisco and Intel for the Dot-Com boom and Nvidia for the current AI boom, with the former shifted forward by 25 years (2015 for 1990 and 2025 for 2000). The companies experience rapid revenue growth from the beginning of the booms. The figure also shows the quarterly ratios of market value to revenue of these companies. For the Dot-Com companies, the ratios kept rising during the booms until its burst in 2000. The ratios started to decline at the end of the boom when revenue growth started to slow down. Up until the second quarter of 2025, the AI boom has remained strong, reflected in Nvidia's robust revenue growth and a high market value-to-revenue ratio.

Similarly, the lower panel in Figure 1 shows measures of the capital stock (Property, Plant, and Equipment PPE) and investment (Capital Expenditure, or CAPEX, over PPE) for the same companies. Cisco and Intel both accumulated their capital stock rapidly from

[1] The quote is commonly attributed to Mark Twain, but there is no definitive evidence that he actually said or wrote it.



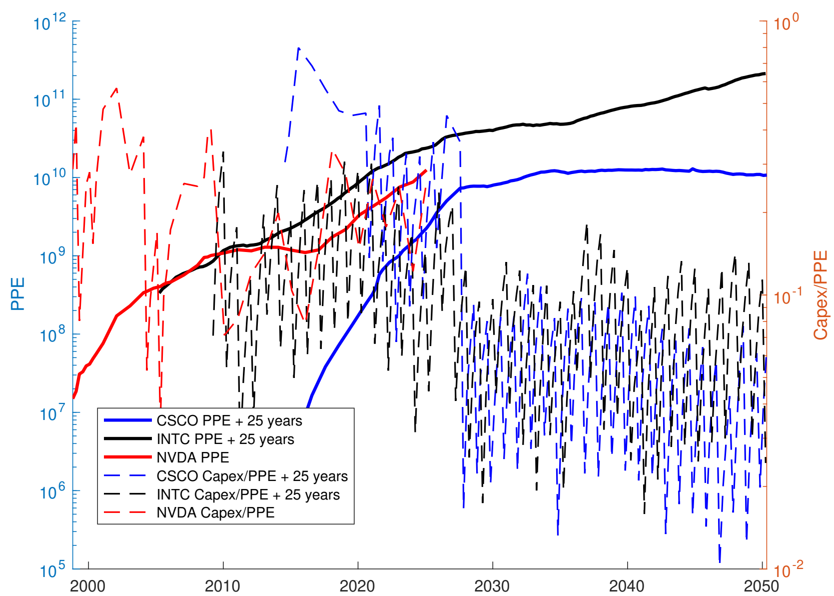


Figure 1: From Dot-Com to AI: Revenue and Market Valuation (upper panel) and Capital Expenditure (lower panel)

1990s to 2000 by undertaking large capital expenditures. However, the capital accumulation slows down significantly after 2000. Nvidia's capital accumulation has accelerated since 2015 and continues to grow rapidly.

In this paper, I build a model to interpret these facts. The model feature two sectors, a high-tech sector A, and the rest of the economy, Sector B. Sector A has a higher productivity growth rate but Sector B benefits from the productivity spillover from Sector A. The capital stock in Sector A is endogenous and can be accumulated using the final good from both sectors, subject to a capital adjustment cost. The economy is populated by two types of agents, optimists and pessimists. They share the same preferences and endowments of the final good and labor, but they might differ in their beliefs about the process of productivity growth in Sector A. The optimists expect that high productivity growth periods (tech booms) last longer and arrive more frequently than the pessimists do. The agents can trade in shares of the capital stock in Sector A and additional financial contracts (collateralized bonds).

I begin the analysis of the model with the special case in which the agents are identical, including their beliefs. On the balanced growth path, I find that the price to dividend ratio of Sector A equity is increasing in the Sector A productivity growth rate if the supply elasticity of Sector A's capital stock is less than the inter-temporal elasticity of substitution of the agents. Outside of the balance growth path, I assume that the productivity growth rate in Sector A follows a two-state Markov chain with a high value state (tech boom) and low value state (normal time). If the spillover is relatively slow, similar to the slow diffusion of influential technological innovations documented in Kalyani et al. (2021), the price-to-dividend ratio increases when the productivity growth rate switches to the high value state from the low value state. As the growth rate remains high, the ratio keeps increasing. The ratio drops when the growth rate transitions back to the low value state, and keeps decreasing towards the pre-boom level, as the growth rate remains low.

However, the dynamics of the ratio differ significantly when the spillover is relative fast. The price to dividend ratio actually peaks after the growth rate transitions back to the low value state. The dividend keeps decreasing but the share price keeps increasing beyond the switching time. This is because the agents' income from Sector B grows quickly due to productivity spillover from Sector A so they have resources to buy Sector A shares and accumulate capital in this sector. On the flip side, if the elasticity of capital increases significantly, the ratio could start to go down before the switching time. When the elasticity increases, capital accumulation speeds up but dividend keeps increasing due to high productivity growth, while asset price stagnates.

Next, I allow the agents to differ in their beliefs. The pessimists have the same beliefs as in the homogeneous beliefs version of the model. However, optimists are more optimistic and believe that the high state arrives more frequently and lasts longer than the pessimists expect. This version of the model leads higher share price and even higher price-to-dividend ratio in the high-tech sector than in the homogenous beliefs version. The latter is due to lower dividend and faster capital accumulation in Sector A. When the optimists can borrow from the pessimists to invest into Sector A capital, the effects are amplified. However, the drops in the share price and price-to-dividend ratio are larger when the state transitions to the low value one.

Going back to the original questions stated at the beginning of the paper, what can the Dot-Com experience tell us about the future of the current AI boom? The dynamics of asset price and capital accumulation described above (when the productivity spillover is slow) resembles the ones observed during the Dot-Com boom in the 1990s. In the current AI boom, amid the real and rapid progress in AI, AI related stocks have been experiencing significant appreciation and have remained at elevated levels. If the model is also applicable to the current AI boom, it suggests that the rapid advancements in AI are expected to continue for the foreseeable future. This expectation is consistent with the growth potentials of AI presented by industry experts. For generative AI and LLMs, while people raise concerned about the progress in LLMs might hit a wall and slows down (partly due to data limitation in the pre-training phrase as analyzed in Villalobos et al. (2024)), there is room for significant improvement in post-training and reasoning using more computing resources (the upper panel in Figure 2 from Nvidia's corporate presentation and Akyurek et al. (2024)). Beyond generative AI, Figure 2 (lower panel) shows that there is more progress to be made in agentic AI and physical AI. According to the model, if the rapid progress continues along these margins, both asset price and capital accumulation rate could rise beyond current levels.

However, alternative versions of the model suggest that the dynamics of the current AI boom could differ considerable from the 1990s Dot-Com boom. First, if the speed of productivity spillover from the AI sector to the rest of the economy is much faster than the it was for the Information and Communication Technology sector in the 1990s, asset price and the rate of capital accumulation could continue to increase even after a slow down in AI productivity gains. Bick et al. (2024) provide early evidence that the overall adoption rate of generative AI has been faster that the adoption rates for personal computers and the internet. Therefore, this may represent a possible scenario. Second, capital supply elasticity in the current AI boom is much lower than it was during the Dot-Com boom, owing to GPU manufacturing challenges and Nvidia's proprietary control of the GPU software ecosystem, which makes entry difficult. This situation could change

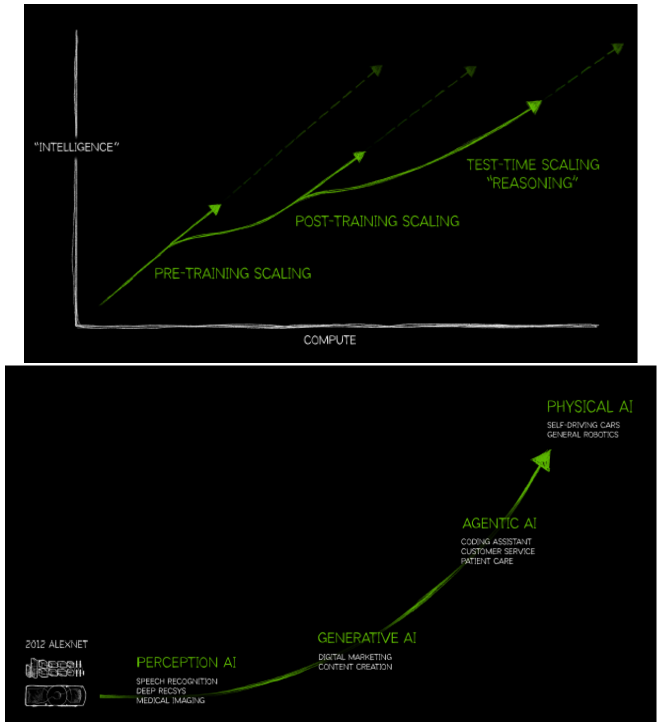


Figure 2: AI Growth Potentials: Generative AI (upper panel) and other AI areas (lower panel) (NVIDIA, 2025)

with further innovation and investment in the sector — for example, through custom AI chips designed by hyperscalers (Google, Meta, Amazon, or even OpenAI) or GPUs developed by Chinese firms such as Huawei and Cambricon. According to the model, once capital supply elasticity increases significantly, asset prices could collapse even if rapid progress in AI continues, while capital accumulation accelerates.

The paper is related to several strands of literature. First, it is part of the effort trying to make sense of the ongoing AI boom and how impactful and world-changing it could be. Not surprisingly, this literature has been expanding very rapidly. Brynjolfsson et al. (2024) propose a comprehensive and impressive agenda for this literature. Second, the paper is related to the literature on asset price or capital accumulation dynamics surrounding major macroeconomic trends including the 1990s Internet boom (Shiller (2000), Pastor and Veronesi (2006), Barbarino and Jovanovic (2007), Cao and L'Huillier (2018)), housing boom and leverage cycles in 2000s (Kaplan et al. (2020), Cao et al. (2023b)). I contributes to this literature by explicitly modeling how the sector of interest interact with the rest of the economy and highlight the importance of the elasticity of capital supply and the speed of spillover. Lastly, at the technical level, the two-sector structure in dynamic general equilibrium also appears in Eberly and Wang (2009) and Cochrane et al. (2007). I solve the model using the novel algorithm and general GDSGE implementation from Cao et al.

(2023a). The extension with belief heterogeneity and leverage is similar to the model in Fostel and Geanakoplos (2008), Geanakoplos (2010), Simsek (2013), and Cao (2018). Hong and Stein (2015) show that belief heterogeneity and short-sale constraints can generate highly nonlinear asset price dynamics including crashes. But these papers only studies endowment economies without endogenous capital accumulation in the present paper.

The rest of the paper is organized as followed. Section 2 presents the general model of a two-sector economy. Section 3 defines Markov equilibrium including the balanced growth path and puts forth a solution method. Section 4 and Section 5 analyzes the dynamics of tech booms under homogenous beliefs and heterogeneous beliefs respectively. Section 6 concludes.

2 A Two-Sector Economy

Consider a single consumption (final) good economy in infinite horizon with infinitely-lived agents. Time runs from $t = 0$ to ∞. In each time, there are S possible exogenous states (or equivalently exogenous shocks)

$$s \in \mathcal{S} = \{1, 2, \ldots, S\}.$$

The states determine the aggregate productivity in the economy.

The evolution of the economy is captured by the realizations of the shocks over time, $s^t = (s_0, s_1, \ldots, s_t)$, which denotes the history of realizations of shocks up to time t. We assume that the shocks follow a Markov process with the transition probabilities $\pi(s, s')$.

Two Sectors: There are two sectors, A and B, in the economy. Both sectors general the final good which could be used for consumption and investment (the price of the final good is normalized to 1). The first sector is the high tech sector with exogenous productivity A_t and endogenous aggregate capital stock K_t. The output in this sector is produced using a standard constant-return-to-scale production function:

$$F(K_t, A_t L_t),$$

where L_t is the labor input and A_t is labor productivity. The productivity follows a random walk:

$$\log A_{t+1} = \log A_t + g_{t+1}. \quad (1)$$

The sector's capital stock accumulation is subject to convex adjustment cost

$$K_t = (1-\delta) K_{t-1} + \Phi(I_t, K_{t-1}),$$

where I_t is investment using the final good.

The remaining sector, Sector B, features exogenous endowment B_t. The dynamics of the endowment depends on A_t through spillover

$$\log B_{t+1} = \log B_t + \rho \log \frac{A_t}{B_t}, \tag{2}$$

where $\rho \in (0,1)$ is the spillover speed.

During period t, a representative final-good producer in sector A produces final good by combining capital K_t, rented at the rate d_t, and labor L_t, at wage rate w_t, using the production above to maximize profit:

$$\max_{K_t, L_t} F(K_t, A_t L_t) - d_t K_t - w_t L_t. \tag{3}$$

At the end of period t, a representative capital-good producers in Sector A produces new units of capital for $t+1$ production by combining old units of capital K_t, purchased at market price q_t and new investment I_t, using the capital adjustment cost function above. The new units of capital is sold at market price q_t^n. The capital good producer chooses inputs to maximize profit:

$$\max_{K_t, I_t} q_t^n ((1-\delta) K_{t-1} + \Phi(I_t, K_{t-1})) - I_t - q_t K_{t-1} \tag{4}$$

I assume that the production function and adjustment cost functions have standard properties.

Assumption 1. *F and Φ have constant-returns to scale and are strictly increasing and weakly concave in their inputs.*

Two Agent Types: There are two types of agents, optimists and pessimists, in this economy with a continuum of measure 1 of identical agents in each type

$$h \in \mathcal{H} = \{O, P\}.$$

In contrast to the standard rational expectation literature, I assume that the agents do not have the perfect estimate of the transition matrix π. The agents might differ in the beliefs on evolution of the exogenous state of the economy. Each of them has their own estimate (belief) of the transition matrix, π^h.

Agents maximize their inter-temporal expected utility with the per period utility CRRA function

$$U(c) = \frac{c^{1-\sigma} - 1}{1-\sigma}$$

At any time t, the agents can buy shares of the new capital stock at prices q_t^n and sell the shares of old capital stock to the capital good producer at prices q_t. They also earn dividend (rent) d_t from the old shares, before selling, to the final-good producers and supply labor inelastically at wage rate w_t. Besides the shares, the agents can trade bonds in zero net supply at price p_t. The bonds allow seller (borrowers) to borrow from the buyers (lenders). The sellers of the bonds needs to pledge capital shares as collateral to make sure that the return from their shares exceeds the promised pays-off from the bonds.

Agent h takes the sequences of prices $\{q_t^n, q_t, p_t, d_t, w_t\}$ as given and chooses consumption c_t^h and capital holding k_{t+1}^h to solve

$$\max_{\{c_t^h, k_t^h, \phi_t^h\}} \mathbf{E}_0^h \left[\sum_{t=0}^{\infty} \beta^t U\left(c_t^h\right) \right] \tag{5}$$

subject to the budget constraint

$$c_t^h + q_t^n k_t^h + p_t \phi_t^h \leq e^h B_t + w_t l^h + (q_t + d_t)\, k_{t-1}^h + \phi_{t-1}^h, \tag{6}$$

the short-sale constraint on capital

$$k_t^h \geq 0, \tag{7}$$

and the leverage/collateral constraint

$$\phi_t^h + (1-m) k_t^h \left(q_{t+1}(s^{t+1}) + d(s_{t+1}) \right) \geq 0, \tag{8}$$

for all histories s^{t+1} that succeeds the history s^t at time t. As in Cao (2018), $m \in [0,1]$ is the fraction of capital the borrowers can borrow against. $m = 1$ implies that there is no borrowing or lending in equilibrium and agents can only use their own available resources to invest in the shares of Sector A capital.

The aggregate labor supply is given by $L = \sum_h l^h$ and we assume that $\sum_h e^h = 1$. In this environment, I define an equilibrium as follows.

Definition 1. A *sequential competitive equilibrium* for an economy with initial asset holdings

$$\left\{k_0^h, \phi_0^h\right\}_{h\in\mathcal{H}}$$

and initial shock s_0 is a collection of consumption, real asset, and bond holdings and prices in each history s^t,

$$\left(\left\{c_t^h\left(s^t\right), k_{t+1}^h\left(s^t\right), \phi_{t+1}^h\left(s^t\right)\right\}_{h\in\mathcal{H}}, q_t^n\left(s^t\right), q_t\left(s^t\right), p_t\left(s^t\right), d_t\left(s^t\right), w_t\left(s^t\right)\right)$$

satisfying the following conditions:

i) The markets for real asset, bond, and final good in each period clear:

$$\sum_{h\in\mathcal{H}} k_t^h\left(s^t\right) = K_t \tag{9a}$$

$$\sum_{h\in\mathcal{H}} \phi_t^h\left(s^t\right) = 0 \tag{9b}$$

$$\sum_{h\in\mathcal{H}} c_t^h\left(s^t\right) + I_t(s^t) = B_t + F(K_t, A_t L). \tag{9c}$$

ii) For each agent h, $\left\{c_t^h\left(s^t\right), k_t^h\left(s^t\right), \phi_t^h\left(s^t\right)\right\}$ solves the individual maximization problem subject to the budget constraint (6), and the no-short sale and collateral constraints, (7) and (8).

iii) The aggregate capital, investment (and labor) $\{K_t(s^t), I_t(s^t)\}$ solves the profit maximization problems of the representative final good and capital producer, (3) and (4).

3 Markov Equilibrium

The productivity dynamics (1) and (2) imply that the model is non-stationary. To obtain stationarity and solve the model numerically, I use scale variables by A_t.

3.1 Scaled Variables

We use the notation $\hat{x}_t = \frac{x_t}{A_t}$ for any variable x(except for the Lagrange multipliers discussed below). For example, $\hat{B}_t = \frac{B_t}{A_t}$ is the productivity of Sector B relatively to Sector A. The spillover dynamics (2) implies

$$\log \hat{B}_{t+1} = (1-\rho)\log \hat{B}_t - g_{t+1},$$

which is a stationary process.

Similarly, the agents' budget constraint (6) becomes

$$\hat{c}_t^h + q_t^n \hat{k}_t^h + p_t \hat{\phi}_t^h \leq e^h \hat{B}_t + \hat{w}_t l^h + (q_t + d_t) \frac{A_{t-1}}{A_t} \hat{k}_{t-1}^h + \frac{A_{t-1}}{A_t} \hat{\phi}_{t-1}^h \tag{10}$$

and the leverage constraints (8) become

$$\hat{\phi}_t^h + (1-m) k_t^h \left(q_{t+1}(s^{t+1}) + d(s_{t+1}) \right) \geq 0.$$

Let λ_t^h and $\mu_t^h(s_{t+1})$ denote the Lagrange multipliers on agents h's no short-selling and collateral constraints. The first-order-conditions of the agents' for capital and financial asset holdings are:

$$\begin{aligned} q_t^n U_h'\left(c_t^h\right) = & \beta \mathbb{E}_t \left[(q_{t+1} + d_{t+1}) U_h'(c_{t+1}^h) \right] \\ & + \lambda_t^h + \sum_{s_{t+1}} \mu_t^h(s_{t+1}) \left(q_{t+1}(s^{t+1}) + d(s_{t+1}) \right) \end{aligned}$$

and

$$p_t U_h'\left(c_t^h\right) = \beta \mathbb{E}_t \left[U_h'(c_{t+1}^h) \Psi_{t+1}(s_{t+1}) \right] + \sum_{s_{t+1}} \mu_t^h(s_{t+1}).$$

Under CRRA utilities, these first order conditions become

$$\begin{aligned} q_t^n \left(\hat{c}_t^h\right)^{-\sigma} = & \beta \mathbb{E}_t^h \left[\left(\frac{A_{t+1}}{A_t} \right)^{-\sigma} (q_{t+1} + d_{t+1}) (\hat{c}_{t+1}^h)^{-\sigma} \right] \\ & + \hat{\lambda}_t^h + \sum_{s_{t+1}} \hat{\mu}_t^h(s_{t+1}) \min_{s^{t+1}|s^t} \left(q_{t+1}(s^{t+1}) + d(s_{t+1}) \right) \end{aligned} \tag{11}$$

and

$$p_t \left(\hat{c}_t^h\right)^{-\sigma} = \beta \mathbb{E}_t^h \left[\left(\frac{A_{t+1}}{A_t} \right)^{-\sigma} (\hat{c}_{t+1}^h)^{-\sigma} \right] + \sum_{s_{t+1}} \hat{\mu}_t^h(s_{t+1}), \tag{12}$$

where

$$\hat{\lambda}_t^h = \lambda_t^h A_t^{-\sigma} \text{ and } \hat{\mu}_t^h = \mu_t^h A_t^{-\sigma}.$$

The first-order conditions for the capital-good producers are

$$\begin{aligned} 0 = & q_t^n \left(1 - \delta + \Phi_2\left(I_t, K_{t-1}\right)\right) - q_t \\ 0 = & q_t^n \Phi_1\left(I_t, K_{t-1}\right) - 1. \end{aligned}$$

which could be rewritten using the normalized variables:

$$0 = q_t^n \left(1 - \delta + \Phi_2 \left(\hat{I}_t, \hat{K}_{t-1} \frac{A_{t-1}}{A_t}\right)\right) - q_t \tag{13a}$$

$$0 = q_t^n \Phi_1 \left(\hat{I}_t, \hat{K}_{t-1} \frac{A_{t-1}}{A_t}\right) - 1. \tag{13b}$$

The corresponding evolution of the capital stock in the tech sector with normalized variables is:

$$\hat{K}_t = (1-\delta) \hat{K}_{t-1} \frac{A_{t-1}}{A_t} + \Phi \left(\hat{I}_t, \hat{K}_{t-1} \frac{A_{t-1}}{A_t}\right).$$

Similarly, the first order conditions for the final good producer are

$$\begin{aligned} 0 =& F_1 \left(K_{t-1}, A_t L_t\right) - d_t \\ 0 =& A_t F_2 \left(K_{t-1}, A_t L_t\right) - w_t, \end{aligned}$$

with the normalized versions:

$$d_t = F_1 \left(\hat{K}_{t-1} \frac{A_{t-1}}{A_t}, L_t\right) \tag{14a}$$

$$\hat{w}_t = F_2 \left(\hat{K}_{t-1} \frac{A_{t-1}}{A_t}, L_t\right) \tag{14b}$$

3.2 Balanced Growth Path

First, I characterize the balanced growth paths (BGPs) with constant productivity growth rate $\bar{G} = \exp(\bar{g})$:

$$\log A_{t+1} = \log A_t + \bar{g}.$$

On balanced growth paths, normalized variables are constant: $x_t \equiv \bar{x}$. Let $\bar{IK}$ denote the BGP value of the investment to capital stock ratio:

$$\bar{IK} = \frac{I_t}{K_t}.$$

The capital accumulation equation becomes

$$\bar{G} = (1-\delta) + \Phi\left(\bar{G}\bar{IK}, 1\right), \tag{15}$$

which uniquely pins down $\bar{IK}$ as a function of the deterministic growth rate $\bar{G}$.

Because Φ has constant returns-to-scale: $\Phi_i \left(\hat{I}, \hat{K}\frac{1}{\bar{G}}\right) = \Phi_i \left(\bar{G}\bar{IK}, 1\right)$ for $i \in \{1, 2\}$.

Therefore,

$$
\begin{aligned}
0 =& \bar{q}^n \left(1-\delta+\Phi_2\left(\bar{G}\bar{I}\bar{K},1\right)\right)-\bar{q} \\
0 =& \bar{q}^n \Phi_1\left(\bar{G}\bar{I}\bar{K},1\right)-1.
\end{aligned}
$$

The second equation pins down the BGP value of new shares $\bar{q}^n$. Combining with the first equation pins down the value of old shares $\bar{q}$.

Without uncertainty, agents have identical beliefs and hence there is no trading. Thus the multipliers in the Euler equations (11) and (12) are zeros. (11) becomes

$$
\bar{q}^n = \beta \bar{G}^{-\sigma}\left(\bar{q}+\bar{d}\right). \tag{16}
$$

This equation gives us the BGP value of the dividend-to-price ratio: $\frac{\bar{d}}{\bar{q}}$. Lastly, the normalized capital stock is pinned down by the marginal product of capital:

$$
F_1\left(\frac{\bar{K}}{\bar{G}},L\right)=\bar{d}.
$$

Proposition 1. *Assume that*

$$
\underbrace{\max_{x>0}\frac{\Theta'(x)}{-x\Theta''(x)}}_{\text{elasticity of capital supply}} \leq \underbrace{\frac{1}{\sigma}}_{\text{intertemporal elasticity of substitution}}, \tag{17}
$$

where $\Theta(x)\equiv\Phi\left(x,1\right)$. *The BGP price-to-dividend ratio is strictly increasing in* $\bar{G}$.

Proof. Appendix B □

When the adjustment cost function is iso-elastic, as assume in Jermann (1998) and Guvenen (2009):

$$
\Theta\left(\frac{I_{t+1}}{K_{t+1}}\right)=a_1\left(\frac{I_{t+1}}{K_{t+1}}\right)^{1-1/\xi}+a_2, \tag{18}
$$

where a_1 and a_2 are chosen so that the steady-state level of capital is independent of ξ:

$$
\begin{aligned}
a_1 &= \frac{\xi}{\xi-1}\delta^{\frac{1}{\xi}} \\
a_2 &= \frac{\delta}{1-\xi}.
\end{aligned} \tag{19}
$$

Inequality (17) becomes

$$\sigma \leq \frac{1}{\xi}.$$

When $\xi \to 0$, the supply of capital is essentially fixed. When $\xi \to \infty$, the supply of capital is fully flexible. In this limit

$$\Theta\left(\frac{I}{K}\right) = \frac{I}{K}$$

and (17) is violated. In this case, the BGP price-to-dividend ratio is strictly decreasing in $\bar{G}$. Indeed, in this case, Euler equation (16) becomes

$$\bar{G}^{\sigma} = \beta \frac{\bar{q}^{n}}{\bar{q}}\left(1 + \frac{\bar{d}}{\bar{q}}\right)$$

where $\bar{q}^{n} = 1$ and $\bar{q} = 1 - \delta$, which implies that $\frac{\bar{d}}{\bar{q}}$ is increasing in $\bar{G}$, and hence, $\frac{\bar{q}}{\bar{d}}$ is decreasing in $\bar{G}$.

3.3 Markov Equilibrium

For the discussions below, I refer to the last two terms in the right hand side of the budget constraint (6)

$$w_t^h = \frac{(q_t + d_t)\, k_t^h + \phi_t^h}{(q_t + d_t)\, K_t}$$

as *financial wealth*, as opposed to *non-financial wealth* $B_t e^h + w_t l^h$, since the later is not tradable nor pledgeable. Because of the collateral constraint and the short-sale constraint, $w_t^h \geq 0$ for all h, t and s^t.

Define the *financial wealth share*, or wealth share for short, of each agent type as

$$\omega_t^h = \frac{(q_t + d_t)\, k_t^h + \phi_t^h}{(q_t + d_t)\, K_t}. \tag{20}$$

Let $\omega_t\left(s^t\right) = \left(\omega_t^O\left(s^t\right), \omega_t^P\left(s^t\right)\right)$ denote the *wealth distribution*. Then in equilibrium $\omega_t\left(s^t\right)$ always lies in the one-dimensional simplex

$$\Omega = \left\{\left(\omega_t^h\right)_{h \in \mathcal{H}} : \omega_t^h \geq 0 \text{ and } \omega_t^O + \omega_t^P = 1\right\}.$$

ω_t^h's are non-negative because the collateral constraint (8) requires the value of each agent's asset holding to exceed the liabilities from their past borrowings. The sum of ω_t^h equals 1 because of the real asset market clearing and bond market clearing conditions.

Using the definition of wealth distribution, we define a Markov equilibrium as follows.

Definition 2. A *Markov equilibrium* is a SCE in which the prices of real and financial assets - q_t, q_t^n, p_t - and the normalized allocation of consumption, real asset and bond, $\hat{c}_t^h, \hat{k}_t^h, \hat{\phi}_t^h$ are (history independent) functions of the exogenous shock s_t and the endogenous financial wealth distribution $\omega_t\left(s^t\right)$.

This Markov equilibrium definition features the endogenous state variable ω_t that depends on equilibrium prices and allocation. This type of endogenous state variable has a long tradition in macroeconomics and finance (Duffie et al. (1994), Brunnermeier and Sannikov (2014), and many others). More importantly, as noted in Cao et al. (2023a), its law of motion is implicit, not explicit. Indeed, ω_{t+1} depends on q_{t+1}which itself depends on ω_{t+1}:

$$\omega_{t+1}^h = \frac{\left(q_{t+1} + d_{t+1}\right)\hat{k}_t^h + \phi_t^h}{\left(q_{t+1} + d_{t+1}\right)\hat{K}_t} = \frac{\hat{k}_t^h}{\hat{K}_t} + \frac{\hat{\phi}_t^h}{\left(q_{t+1}(s_{t+1}, \omega_{t+1}) + d_{t+1}(s_{t+1})\right)\hat{K}_t}.$$

Therefore, it is not explicitly defined even if the current policy variables $\hat{k}_t^h, \phi_t^h, \hat{K}_t$ are known. To tackle this challenge, I use the simultaneous policy function and transition function iterations proposed in Cao et al. (2023a) and impose the consistency equations above during the time iteration. The algorithm and GDSGE implementation are presented in Appendix C.

4 Tech Boom Dynamics: Homogenous Beliefs

In this section, I assume that the agents are identical in beliefs and their labor supply in the high-tech sector and endowment in the low-tech sector are proportional to their initial holding of capital. The following proposition shows that the equilibrium allocation in this case is the same as the solution to the social planner's problem.

Proposition 2. *Assume that agents has the same beliefs, $\pi^P = \pi^O$, and the initial financial propositions are proportional to their non-financial endowments*

$$\frac{k_{-1}^O}{k_{-1}^P} = \frac{e^O}{e^P} = \frac{l^O}{l^P},$$

and $\phi_{-1}^O = \phi_{-1}^P = 0$. Starting from initial capital stock K_0, there exists a unique equilibrium in *which the aggregate variables* ($K_t, I_t, c_t = c_t^O + c_t^P$) *are given by the planner's problem*

$$\max_{\{c_t, I_t, K_t\}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma}}{1-\sigma} \right]$$

s.t.

$$c_t + I_t \leq B_t + F\left(K_t, A_t L\right)$$

and

$$K_{t+1} = (1-\delta)K_t + \Phi\left(I_t, K_t\right).$$

Proof. Appendix B □

To illustrate equilibrium properties, I assume that s follows a two-state Markov process with low growth or high growth rate in the tech sector:

$$g \in \{g_L, g_H\},$$

with the transition matrix

$$\pi^P = \pi^O = \begin{bmatrix} 1 - p_{LH} & p_{LH} \\ 1 - p_{HH} & p_{HH} \end{bmatrix}$$

I pick $g_L = 0.01$ and $g_H = 0.04$, in line with the estimates of potential growth increase due to AI and other automation technologies in Chui et al. (2023).

The total labor supply in the high tech sector is L_A and divided between the optimists and pessimists

$$l^O = 0.1 L_A \text{ and } l^P = 0.9 L_A.$$

Total endowment in sector B is 1 with $e^O = 0.1$ and $e^P = 0.9$. In the steady-state the total output produced in the high tech sector relative to the other sector is

$$\frac{A}{B} \bar{K}^{\alpha} L_A^{1-\alpha}.$$

I choose L_A so that Sector A's output is around 10% of the total output.

Slow Spillover I set the spillover factor ρ in (2) at

$$\rho = 0.025,$$

i.e., it takes around 50 years for the productivity gain in Sector A to diffuse fully to Sector B, consistent with the estimate of the diffusion speed of influential technological innovations in Kalyani et al. (2021). I show below that this level of spillover generates empirically relevant asset price dynamics. In general, the spillover speed plays an important role in the dynamics of asset prices.

In this section, I assume that the agents have exactly the same beliefs:

$$p^O_{LH} = p^P_{LH} = 0.05$$
$$p^O_{HH} = p^P_{HH} = 0.80.$$

This transition matrix implies that tech booms ($g = g_H$) lasts on average for 5 years and happens every 10 years.

The policy functions for Markov equilibrium are presented in Supplemental Appendix C. In this section, I explore how the equilibrium variables respond to a sequence of growth shocks g_t. I assume that g_t stays at g_L for a long time ($t = 1$ to 100). At $t = 101$, it switches to the high state (tech booms) and remains there until $t = 115$ before switching back to the low state. The upper panel in Figure 3 shows the equilibrium dynamics of the price-to-dividend ratio, q_t/d_t, over time. The ratio jumps up immediately when productivity increases but evolve non-monotonically over time. Notably, the ratio peaks at $t = 139$, more than 10 periods after productivity growth comes down to g_L.

To better understand the dynamics, the lower panel of Figure 3 plots q_t and d_t separately. Dividend d_t moves more slowly than share price q_t. This is because d_t is determined by the normalized capital stock (14a) which adjust slowly (middle panel) due to the capital adjustment cost. At $t = 101$, q_t jumps up while d_t remains close to the pre-switch level. This leads to a jump in the price-to-dividend ratio. Afterwards, the ratio keeps increasing as q_t increases faster than d_t does. Increasing q_t reflects the persistently higher level of relative productivity in the high-tech sector after the switch, which also implies high levels of investment (bottom panel for I_t/K_t). Equation (13b) shows that a higher $\hat{I}_t$ corresponds to a higher marginal cost of a new units of capital, and hence higher q^n_t and, by Equation (13a), a higher q_t. Both d_t and q_t peak at $t = 115$ start to decline as soon as the growth rate switches from g_H to g_L at $t = 116$. Similar to the dynamics before the switch, q_t declines faster than d_t does and, hence their ratio goes down over time.

The long-run levels of $\hat{K}_t = K_t/A_t$ when $g_t \equiv g_L$ or g_H are approximately 0.16 and 0.01 respectively. The middle panel in Figure 3 shows that starting from the high initial level around 0.16 at $t = 101$, $\hat{K}_t$ decreases towards the lower long-run level 0.01 because K_t keeps increasing but not as fast as A_t goes (at the rate g_H). After g_t switches to g_L at

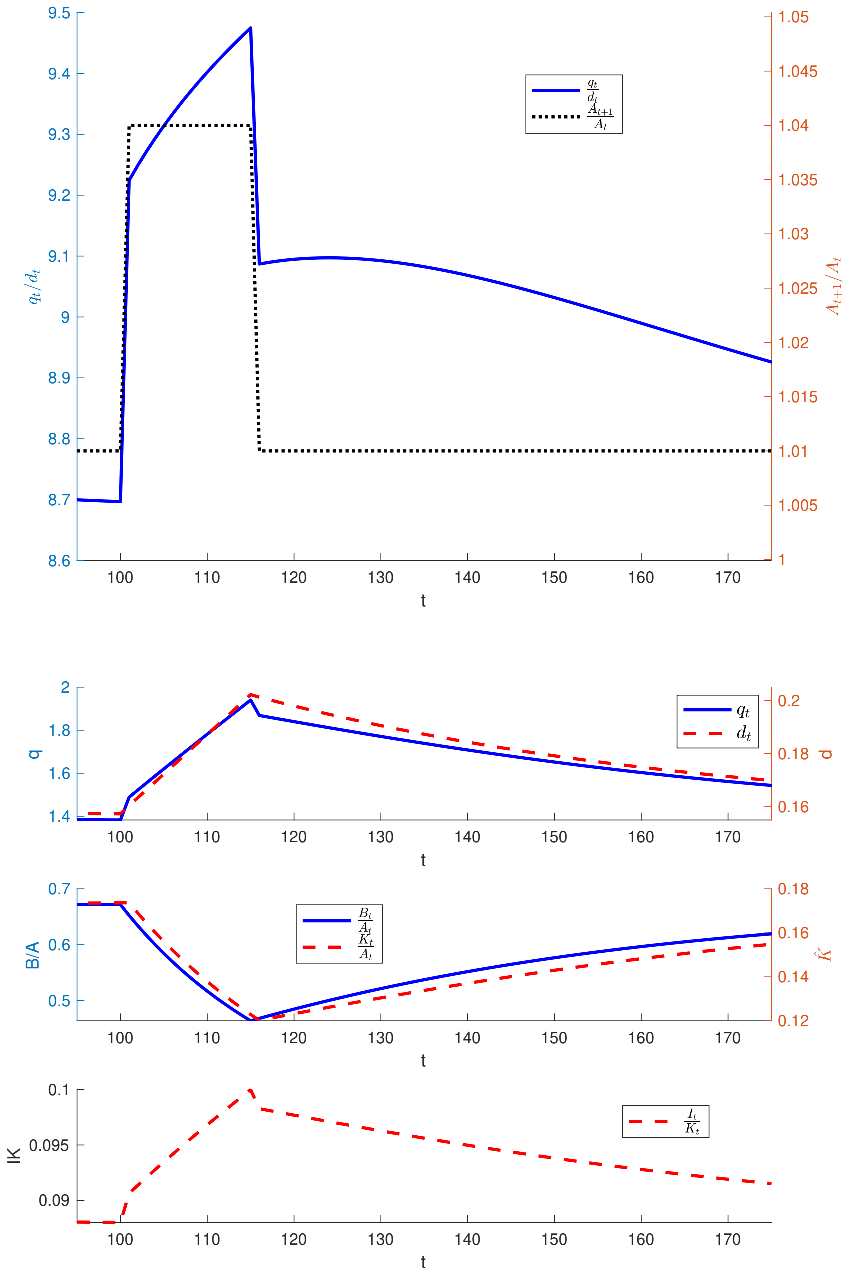


Figure 3: Price-to-Dividend Ratio (upper panel) and Price, Dividend, Capital and Investment (lower panels)

$t = 116$, K_t keeps expanding at rates higher than g_L. Therefore $\hat{K}_t$ increases over time towards its long run level around 0.16. Figure 9 in Appendix C shows that the price to dividend ratio is decreasing in $\hat{K}_t$. Therefore q_t/d_t (and q_t) increases over time while $g_t = g_H$.[2]

These dynamics are similar to the dynamics of PPE and Capex shown in Figure 1. Despite the decline in $\hat{K}_t$, K_t increases rapidly over time due to the high investment rate, I_t/K_t. In this sense, there is no over-investment as in the popular narrative during the DotCom and the current AI booms. If anything, one could argue that K_t is "too low" compared to A_t and the economy experiences underinvestment into Sector A. Indeed, even after g_t switches to g_L at $t = 116$, K_t keeps going up albeit at lower rate. Overall, the dynamics of investment and capital are fully optimal as they result from the planner's optimal control problem.

Fast Spillover To examine the importance of productivity spillover, I solve and simulate the model for a higher value of $\rho = 0.10$, instead of the benchmark value 0.025. Figure 4 (upper panel) shows that the price to dividend ratio peaks much later, around 5-10 years after g_t switches from g_L to g_H. The lower panel shows that this is because asset price q_t keeps appreciating even as growth rate slows down at $t = 116$, while capital accumulates and d_t declines at similar speeds as when $\rho = 0.025$. The scaled budget constraint (10) suggests that q_t keeps increasing from $t = 116$ to 121 because the agents' endowment from Sector B grows faster thanks to the rapid spillover (as shown in the middle panel) and the agents can use some of the endowment to accumulate capital shares in Sector A, driving up q_t^n and q_t. Faster endowment growth allows agents to increase both consumption and investment faster. The upper panel of Figure 5 shows that consumption growth fasters with faster spillover from $t = 116$ to $t = 121$. Faster consumption growth requires that rate of return to investing in capital shares is higher to be consistent with the representative consumer's Euler equation:

$$1 = \beta \mathbb{E}_t \left[\frac{u'(c_{t+1})}{u'(c_t)} \frac{q_{t+1} + d_{t+1}}{q_t^n} \right]$$

The middle panel of of Figure 5 shows that this is indeed the case $(q_{t+1} + d_{t+1})/q_t^n$ is higher from $t = 116$ until $t = 121$ when $\rho = 0.10$ than when $\rho = 0.025$. The bottom panel shows that this is due to both a higher appreciation rate of q_t and a higher dividend rate.

[2] Asset price rises gradually despite the standard constant return to scale investment cost due to shocks to the growth rate of productivity instead of the level shocks in Barbarino and Jovanovic (2007) (who rely on a non-standard investment function to produce the dynamics with a gradual rise in asset prices).

These dynamics are inconsistent with the empirical patterns documented in Figure 1 for the Dot-Com boom. Therefore, the benchmark model with slow spillover is more empirically relevant for this episode. However, the spillover could be much faster for current AI boom. Bick et al. (2024) provide early evidence that the overall adoption rate of generative AI has been faster than the adoption rates for personal computers and the internet. If this also applies to productivity spillover, asset price and the rate of capital accumulation could continue to increase even after a slow down in AI productivity gains as shown in Figure 4.

Stochastic Capital Supply Elasticity The benchmark assumes that the elasticity of capital is relatively low so that (17) is satisfied and the price to dividend ratio tends to be increasing in the expected growth rate of the high-tech productivity. Low capital supply elasticity is consistent with the supply constraints the whole AI sector has been facing since the release of ChatGPT in 2022. Nvidia has been a dominant supplier of the GPU chips used to train and deploy AI models. On the Q3 FY2025 earning call in November 2024, Nvidia CFO stated that the demand for Blackwell, Nvidia's latest GPU chip design, was expected to exceed supply for several quarters in fiscal 2026. The excess demand translates into significant pricing power and profit margin of over 70% for Nvidia since 2022. GPU supply is difficult to expand rapidly because 1. the manufacturing process and supply chain for these chips are extremely complex (for example, the GPUs are made on the most advanced semiconductor nodes which pack billions of nanometer-sized transistors on these chips; at these tiny scales, even a defect at the atomic level can ruin an entire chip); and 2. Nvidia's GPUs are programmed using CUDA (Compute Unified Device Architecture) which is a proprietary software platform and prevents other manufacturers from supplying GPUs using the platform. Millions of lines of AI and high performance computing (HPC) code have been written in CUDA creating a vast software ecosystem that makes GPUs useful for AI and HPC, lock in customers, and make the entry into GPU market difficult. Similarly, demand for Cisco's networking gear outstripped supply at the peak of the Dot-Com booms in 1999-2000 with some high-end core routers had months-long backlogs because component suppliers (albeit these constraints were much less severe than the structural bottlenecks we see in the GPU world today). Cisco also enjoyed a strong lock-in effect in the 1990s through its internetwork operating system and proprietary protocols for its routers and switches, workforce training, and certifications.

However, since 2022, there has been significant expansion of supply of AI chips. Hyperscalers, primary developers of foundational large language models, such as Google, Meta, and even OpenAI, develop their own custom AI chips, such as Google's TPUs,

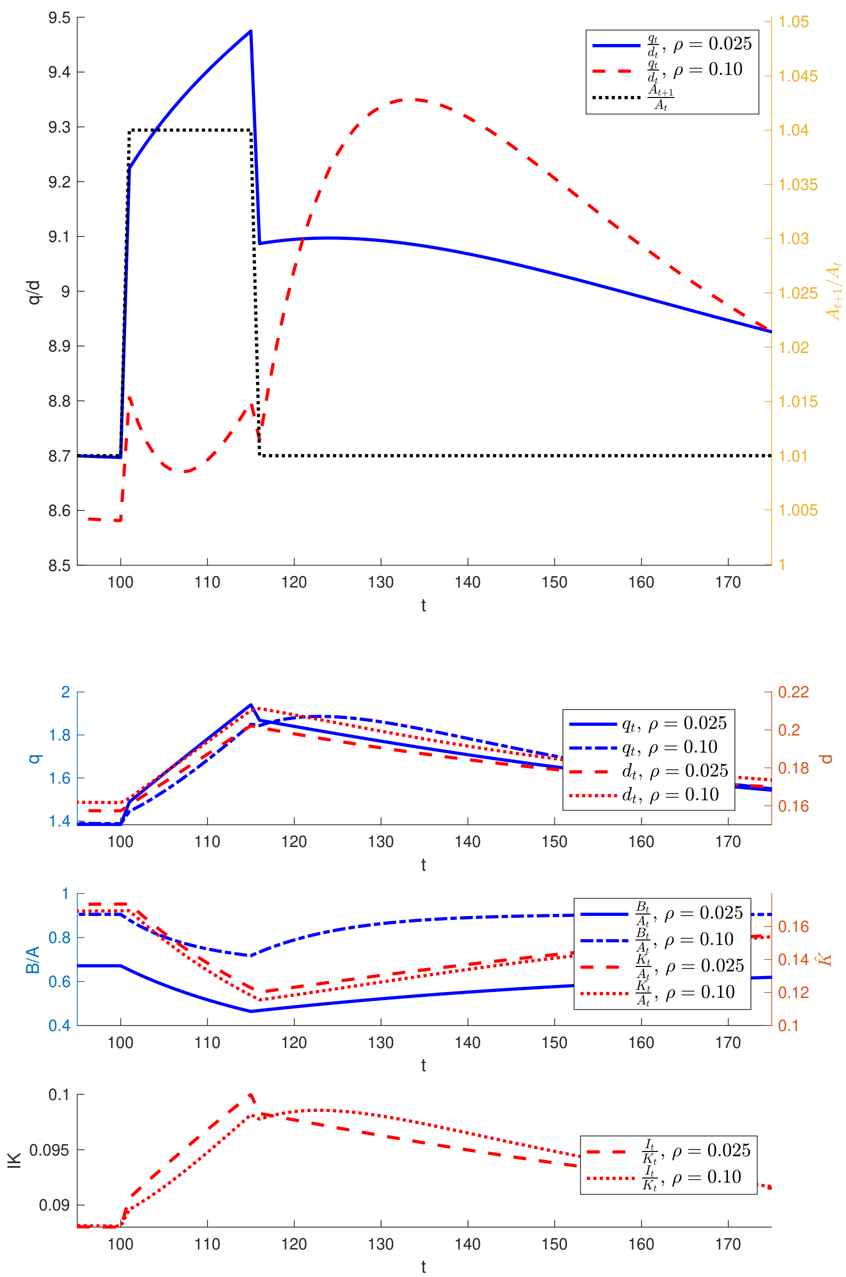


Figure 4: Price-to-Dividend Ratio (upper panel) and Price, Dividend, Capital and Investment (lower panel)

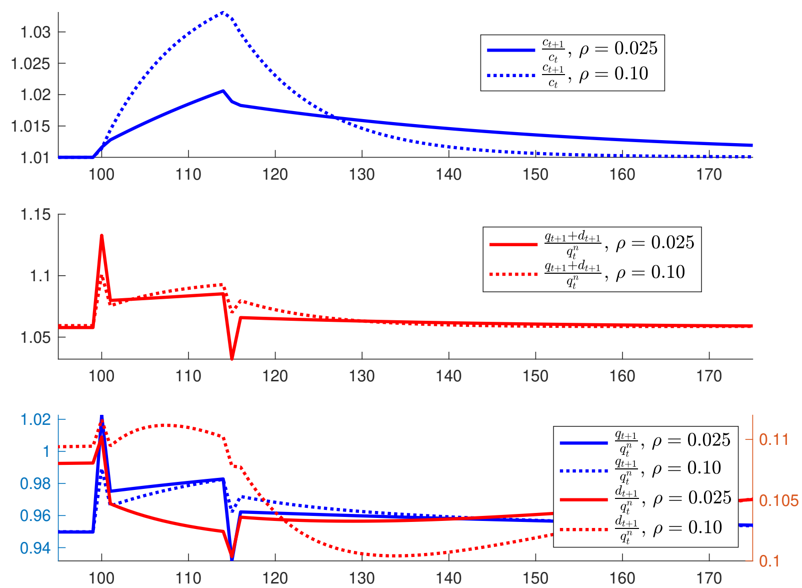


Figure 5: Consumption growth (upper panel) and Stock return (lower panel)

tailored for training and inferencing with these models. Amazon have developed their own AI chips (Trainium for training and Inferentia for inferencing) and have made them available to customers through AWS cloud. AMD and Intel have also started to produce AI GPUs. In China, many manufacturers like Huawei and Cambricon have started to developed and offered AI chips to the Chinese market thanks to the restrictions to Nvidia's GPUs imposed by the U.S. government.

To capture these dynamics, I extend the model to allow for stochastic transition of capital supply elasticity

$$\Theta_t\left(\frac{I_{t+1}}{K_{t+1}}\right) = a_{1,t}\left(\frac{I_{t+1}}{K_{t+1}}\right)^{1-1/\xi_t} + a_{2,t}, \tag{21}$$

where $a_{1,t}$ and $a_{2,t}$ are functions of ξ_t given in (19). The supply elasticity increases when the high tech productivity enters the high growth state. Formally, I assume that the elasticity follows a two-state Markov chain, conditional on Sector A's productivity growth, g_t:

$$\xi_t \in \{\xi_L, \xi_H\}$$

with

$$\Pr\left(\xi_{t+1} = \xi_H | \xi_t, g_t\right) = \begin{cases} 0.25 & \text{if } \xi_t = \xi_L, g_t = g_H \\ 0 & \text{otherwise.} \end{cases}$$

Figure 6 shows the simulation from this model. As in Figure 3, I assume that g_t stays at g_L from $t = 1$ to 100 and at $t = 101$, it switches to the high state (tech booms) and remains there until $t = 115$ before switching back to the low state. However, at $t = 110$, the elasticity ξ_t switches from ξ_L to ξ_H and remains there until the end of the simulation period. The upper panel in Figure 6 shows the equilibrium dynamics of the price-to-dividend ratio, q_t/d_t, over time for the benchmark model with fixed elasticity (solid lines) and for this model with stochastic elasticity (dashed lines). Unlike in the benchmark model, the ratio starts to decreases at $t = 110$ before g_t switches back to g_L at $t = 115$. The bottom panel in Figure 6 helps explain the dynamics. Investment and capital stock expands more rapidly with higher elasticity and dividend d_t keeps going up while capital price q_t remains flat from $t = 110$ to 115.

5 Tech Boom Dynamics: Heterogenous Beliefs

Major tech booms are often accompanied by significant disagreement on the magnitude and the economy-wide impact of these booms. For example, in the ongoing AI boom, leading AI researchers disagree on whether GPT LLMs like ChatGPT are promising paths towards AGI. Geoffrey Hinton, who won the 2024 Nobel Prize for his contribution in AI, argues that LLMs already have significant cognitive abilities and can be improved relatively quickly to achieve AGI. While, Yan Lecun, another foremost AI pioneer, asserted that a significant gap exists in achieving human-level AI, and current LLMs are not even at cat-level intelligence, indicating a need for substantial advancements in AI systems' understanding of the physical world and planning abilities. On the economy-wide impact, Acemoglu (2024) estimates that new advances in AI will lead to only 0.66% increase in total factor productivity for the U.S. over 10 years. Goldman Sachs (2023) is much more optimistic with 4% increase in GDP over the same time span.

In this section, I examine how belief heterogeneity affects the dynamics of the booms. To model belief heterogeneity, I assume that the agents to differ in their subjective transition matrix:

$$p^P_{LH} = 0.05 < p^O_{LH} = 0.10$$
$$p^P_{HH} = 0.80 < p^O_{HH} = 0.95.$$

The pessimists' transition matrix is the same as in the previous section. But the optimists expect that tech booms last 20 years on average (instead of the pessimists' 5 years average) and takes on average 10 years to emerge (compared to the pessimists' 20 years average).

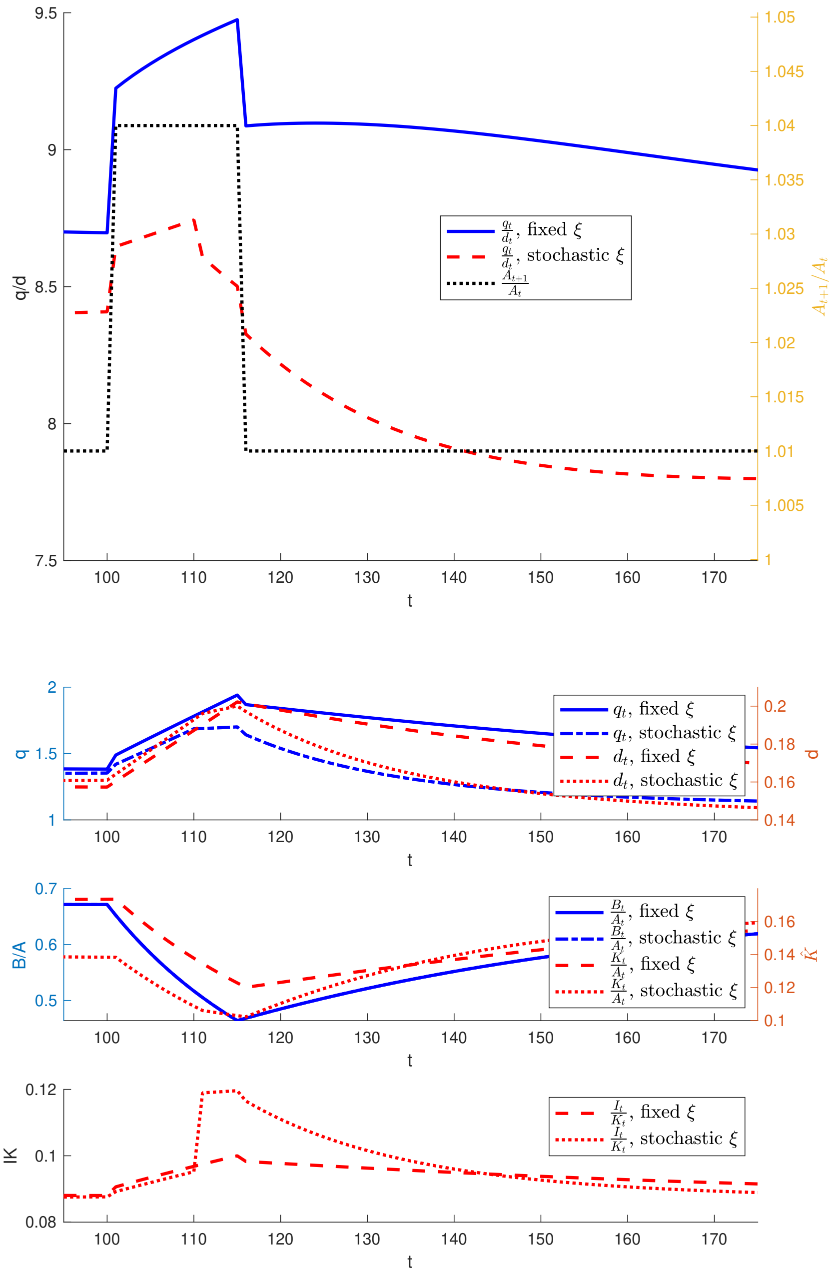


Figure 6: Price-to-Dividend Ratio (upper panel) and Price, Dividend, Capital and Investment (lower panel)

First, I consider the special case in which the optimists cannot borrow from the pessimists to invest in high-tech equity ($m = 1$). This is consistent with the fact that during the ongoing AI boom, many big tech firms reinvest the revenues from their AI related products into new AI businesses. The agents' budget constraints (6) become

$$c_t^h + q_t^n k_t^h \leq e^h B_t + w_t l^h + (q_t + d_t)\, k_{t-1}^h. \tag{22}$$

Without borrowing, the optimists can still accumulate high-tech equity from their labor earning $w_t l^O$ and dividends and existing equity holding plus dividend: $(q_t + d_t)k_{t-1}^O$. The dashed red line in Figure 7 shows the dynamics of price-to-dividend ratio in this model. The price-to-dividend ratio is higher under heterogenous beliefs than in the previous model with homogenous beliefs. But the magnitude of the relative changes given the same sequence of productivity shocks is similar.

Next, I consider the general general model which allows the optimists to borrow from the pessimists to acquire high-tech shares, i.e., $m < 1$. Many of the companies in the Internet boom and the current AI boom raise funding through a variety of financial contracts such as venture capital, partnership, and leverage. For example, Microsoft has invested approximately \$13 billion in OpenAI since their partnership began in 2019. The latter retains some of the upside potentials because the former would lose access to the former AI technologies when its achieves AGI and generates more than \$100 billion dollars in profits. BlackRock (2024)'s new AI partnership raised \$70 billion dollars from bond issuance, on top of \$30 billion dollars in equity. Figure 7 (solid lines) plots the equilibrium dynamics for $m = 0.5$. The upper panel shows that leverage leads to a larger increase in price-to-dividend ratio, compared to the previous model without leverage, when the tech boom arrives at $t = 101$ but also a larger decrease when the tech boom ends at $t = 115$. The bottom panel tells us that this is due to both higher equity price and lower dividend, i.e, faster capital accumulation.

6 Conclusion

In this paper, I propose a theoretical model to study the dynamics of the 1990s Dot-Com boom and the ongoing AI boom. The dynamics of asset prices and capital accumulation in the 1990s could be explained by persistent productivity growth in the high-tech sectors with very slow spillover to the rest of the economy. If the model is applicable to the current AI boom, we should expect the rapid progress in AI to continue until AI valuation reaches its peak. However, the result depends on the supply elasticity of capital in the

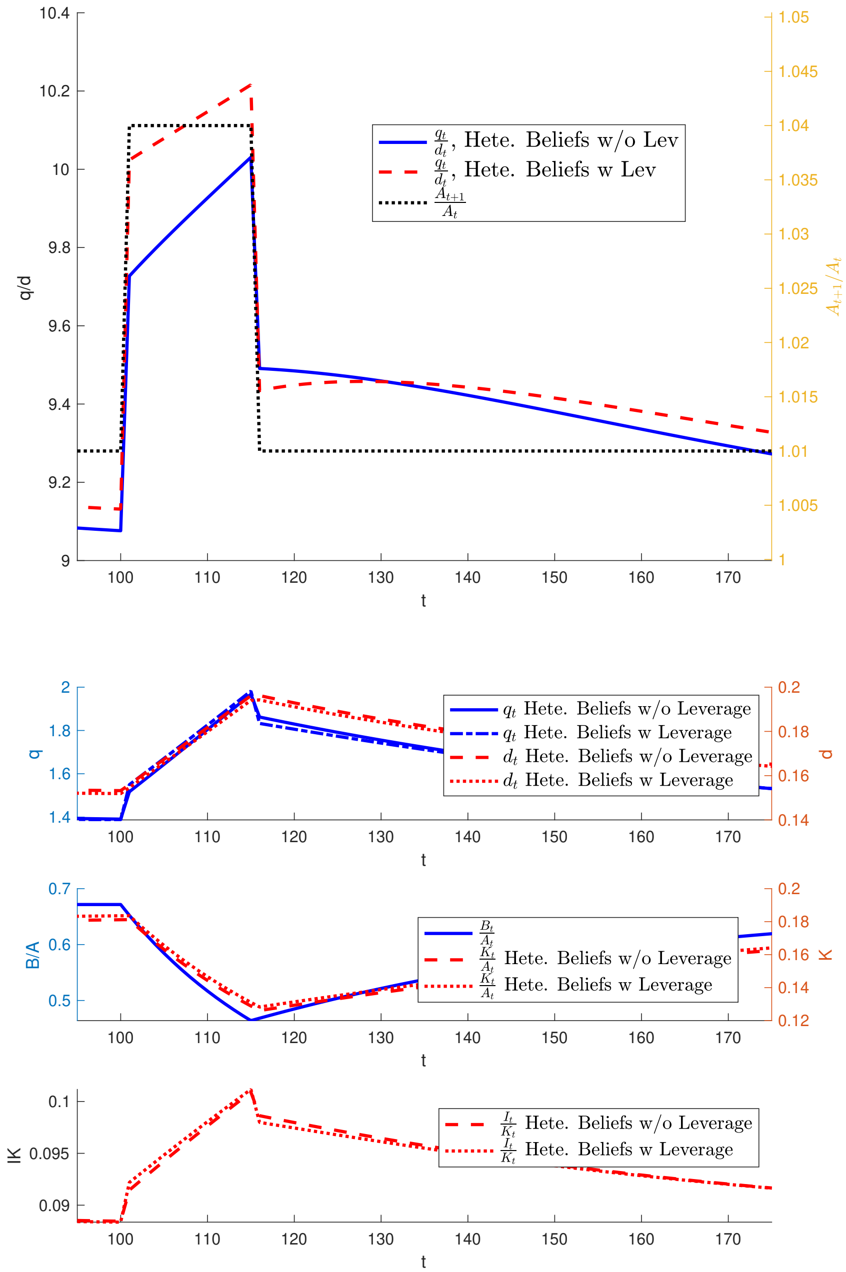


Figure 7: Price-to-Dividend Ratio (upper panel) and Price, Dividend, Capital and Investment (lower panel)

high-tech sectors and the speed of spillover.

References

Acemoglu, D. (2024). The simple macroeconomics of ai. *Economic Policy 39*(120).

Akyurek, E., M. Damani, L. Qiu, H. Guo, Y. Kim, and J. Andreas (2024). The surprising effectiveness of test-time training for abstract reasoning.

Barbarino, A. and B. Jovanovic (2007). Shakeouts and market crashes. *International Economic Review 48*(2), 385–420.

Bick, A., A. Blandin, and D. J. Deming (2024, September). The rapid adoption of generative ai. Working Paper 32966, National Bureau of Economic Research.

BlackRock (2024). Blackrock, global infrastructure partners, microsoft, and mgx launch new ai partnership to invest in data centers and supporting power infrastructure. Accessed: 2025-02-09.

Brunnermeier, M. and Y. Sannikov (2014). A macroeconomic model with a financial sector. *American Economic Reviews 104*(2), 379–421.

Brynjolfsson, E., A. Korinek, and A. Agrawal (2024). The economics of transformative ai: A research agenda. Technical report. UVA Working Paper.

Cao, D. (2018). Speculation and financial wealth distribution under belief heterogeneity. *The Economic Journal*.

Cao, D. and J.-P. L'Huillier (2018). Technological revolutions and the three great slumps: A medium-run analysis. *Journal of Monetary Economics 96*, 93–108.

Cao, D., W. Luo, and G. Nie (2023a). Global dsge models. *Review of Economic Dynamics 51*, 199–225.

Cao, D., W. Luo, and G. Nie (2023b). Uncovering the effects of the zero lower bound with an endogenous financial wedge. *American Economic Journal: Macroeconomics 15*(1), 135–172.

Chui, M., E. Hazan, R. Roberts, A. Singla, K. Smaje, A. Sukharevsky, L. Yee, and R. Zemmel (2023). *The Economic Potential of Generative AI*. New York: McKinsey and Company.

Cochrane, J. H., F. A. Longstaff, and P. Santa-Clara (2007, 11). Two trees. *The Review of Financial Studies 21*(1), 347–385.

Duffie, D., J. Geanakoplos, A. Mas-Colell, and A. McLennan (1994). Stationary markov equilibria. *Econometrica 62*(4), 745–781.

Eberly, J. and N. Wang (2009, May). Capital reallocation and growth. *American Economic Review 99*(2), 560–566.

Fostel, A. and J. Geanakoplos (2008). Leverage cycles and the anxious economy. *American Economic Reviews 98*(4), 1211–1244.

Geanakoplos, J. (2010). Leverage cycles. In K. R. Daron Acemoglu and M. Woodford (Eds.), *NBER Macroeconomics Annual 2009*, Volume 24, pp. 1–65. University of Chicago Press.

Goldman Sachs (2023). Ai may start to boost us gdp in 2027. Accessed: 2025-01-15.

Guvenen, F. (2009). A parsimonious macroeconomic model for asset pricing. *Econometrica 77*(6), 1711–1750.

Hong, H. and J. C. Stein (2015, 06). Differences of opinion, short-sales constraints, and market crashes. *The Review of Financial Studies 16*(2), 487–525.

Jermann, U. J. (1998). Asset pricing in production economies. *Journal of Monetary Economics 41*(2), 257–275.

Kalyani, A., N. Bloom, M. Carvalho, T. A. Hassan, J. Lerner, and A. Tahoun (2021, July). The diffusion of new technologies. Working Paper 28999, National Bureau of Economic Research.

Kaplan, G., K. Mitman, and G. L. Violante (2020). The housing boom and bust: Model meets evidence. *Journal of Political Economy 128*(9), 3285–3345.

NVIDIA (2025, March). GTC 2025 Keynote Presentation. https://s201.q4cdn.com/141608511/files/doc_downloads/2025/03/GTC2025_Keynote.pdf. Presented by Jensen Huang at NVIDIA GTC 2025, San Jose, CA.

Pastor, L. and P. Veronesi (2006). Was there a nasdaq bubble in the late 1990s? *Journal of Financial Economics 81*(1), 61–100.

Shiller, R. J. (2000). *Irrational Exuberance* (1st ed.). Princeton, NJ: Princeton University Press.

Simsek, A. (2013). Belief disagreements and collateral constraints. *Econometrica 81*(1), 1–53.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin (2023). Attention is all you need.

Villalobos, P., A. Ho, J. Sevilla, T. Besiroglu, L. Heim, and M. Hobbhahn (2024). Will we run out of data? limits of llm scaling based on human-generated data.

Appendix

A Additional Empirical Results

To expand the set of DotCom and AI firms from Intel, Cisco, and Nvidia, I began with the firm-level data from WRDS (Compustat and CRSP) and applied multiple selection criteria. First, I excluded industries entirely unrelated to the Internet or AI and retained only those in technology-related sectors. Next, we reviewed business descriptions to remove firms operating at the periphery, such as traditional communication or peripheral hardware producers. Finally, I kept only firms that made direct contributions to the core technologies of the Internet or artificial intelligence. I ended up with 35 DotCom companies in the 1990s including Cisco and Intel and 26 AI companies in 2020s including Nvidia. Figure 8 present similar metrics as in Figure 1 for these companies, relative to companies in the S&P 500 index. The top panel shows that the aggregate price to revenue ratios peaked for DotCom companies in 2000 and has been rising rapidly for AI companies since 2020 and exceed the level for SP500 companies. The bottom panel shows that the investment rate (CAPEX/PPE) was higher for DotCom companies than for SP500 companies until 2000 and has been higher for AI companies since 2010.

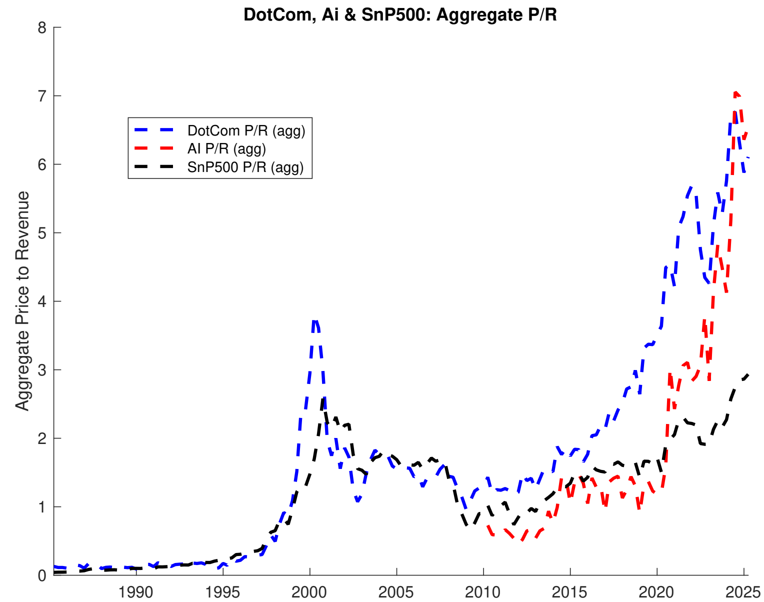
B Proofs

Proof of Proposition 1.

$$\bar{q}^n = \frac{1}{\Phi_1\left(\bar{G}\bar{I}\bar{K}, 1\right)}$$

and

$$\bar{q} = \frac{1}{\left(1 - \delta + \Phi_2\left(\bar{G}\bar{I}\bar{K}, 1\right)\right)\Phi_1\left(\bar{G}\bar{I}\bar{K}, 1\right)}$$



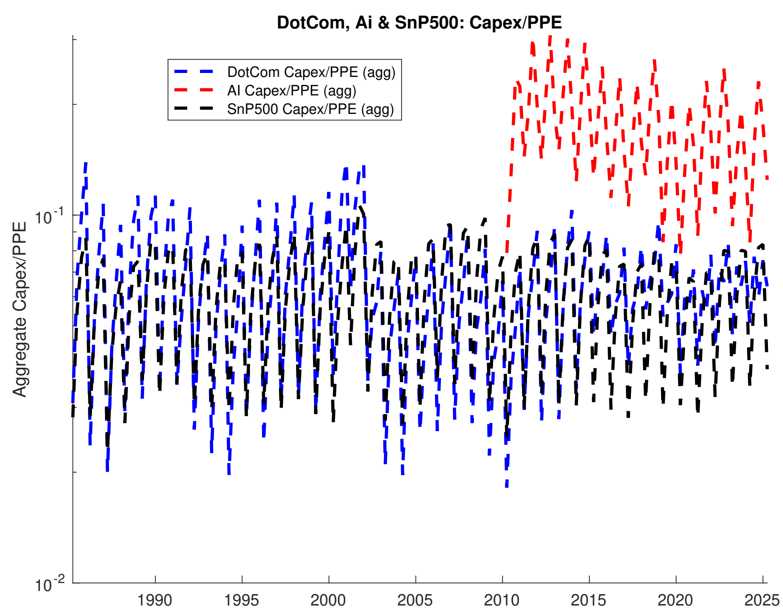


Figure 8: From Dot-Com to AI: Revenue and Market Valuation (upper panel) and Capital Expenditure (lower panel)

and

$$\frac{\bar{q}^n}{\bar{q}}\bar{G}^\sigma = \beta\left(1+\frac{\bar{d}}{\bar{q}}\right) = \frac{\bar{G}^\sigma}{1-\delta+\Phi_2\left(\bar{G}\bar{IK},1\right)}.$$

where

$$\bar{G} = (1-\delta) + \Phi\left(\bar{G}\bar{IK},1\right)$$

Differentiate both sides with respect to $\bar{G}$, we obtain:

$$1 = \Phi_1\left(\bar{G}\bar{IK},1\right)\left(\bar{IK}+\bar{G}\frac{d\bar{IK}}{d\bar{G}}\right).$$

Therefore

$$\begin{aligned}
&\frac{d}{d\bar{G}}\frac{\bar{G}^\sigma}{1-\delta+\Phi_2\left(\bar{G}\bar{IK},1\right)} \\
&= \frac{\sigma\bar{G}^{\sigma-1}\left(1-\delta+\Phi_2\left(\bar{G}\bar{IK},1\right)\right)-\bar{G}^\sigma\Phi_{21}\left(\bar{G}\bar{IK},1\right)\left(\bar{IK}+\bar{G}\frac{d\bar{IK}}{d\bar{G}}\right)}{\left(1-\delta+\Phi_2\left(\bar{G}\bar{IK},1\right)\right)^2} \\
&= \frac{\bar{G}^{\sigma-1}}{\left(1-\delta+\Phi_2\left(\bar{G}\bar{IK},1\right)\right)^2}\left(\sigma\left(1-\delta+\Phi_2\left(\bar{G}\bar{IK},1\right)\right)-\frac{\bar{G}\Phi_{21}\left(\bar{G}\bar{IK},1\right)}{\Phi_1\left(\bar{G}\bar{IK},1\right)}\right) \\
&= \frac{\bar{G}^{\sigma-1}}{\left(1-\delta+\Phi_2\left(\bar{G}\bar{IK},1\right)\right)^2}\left(\sigma\left(1-\delta+\Theta\left(\bar{G}\bar{IK}\right)-\bar{G}\bar{IK}\Theta'\left(\bar{G}\bar{IK}\right)\right)-\frac{\bar{G}\Phi_{21}\left(\bar{G}\bar{IK},1\right)}{\Phi_1\left(\bar{G}\bar{IK},1\right)}\right) \\
&= \frac{\bar{G}^{\sigma}}{\left(1-\delta+\Phi_2\left(\bar{G}\bar{IK},1\right)\right)^2}\left(\sigma\left(1-\bar{IK}\Theta'\left(\bar{G}\bar{IK}\right)\right)-\frac{\Phi_{21}\left(\bar{G}\bar{IK},1\right)}{\Phi_1\left(\bar{G}\bar{IK},1\right)}\right) \qquad (23)
\end{aligned}$$

Notice that

$$\Phi(I,K) = K\Theta\left(\frac{I}{K}\right)$$

which implies

$$\Phi_1(I,K) = \Theta'\left(\frac{I}{K}\right)$$

and

$$\Phi_2(I,K) = \Theta\left(\frac{I}{K}\right)-\frac{I}{K}\Theta'\left(\frac{I}{K}\right)$$

and

$$\Phi_{21}(I,K) = -\frac{I}{K}\frac{1}{K}\Theta''\left(\frac{I}{K}\right).$$

Plugging these back into (23), we arrive at

$$\begin{aligned}
&\frac{d}{d\bar{G}}\frac{\bar{G}^{\sigma}}{1-\delta+\Phi_2\left(\bar{G}\bar{I}\bar{K},1\right)}\\
&=\frac{\bar{G}^{\sigma}}{\left(1-\delta+\Phi_2\left(\bar{G}\bar{I}\bar{K},1\right)\right)^2}\left(\sigma\left(1-\bar{I}\bar{K}\Theta'\left(\bar{G}\bar{I}\bar{K}\right)\right)-\frac{-\bar{G}\bar{I}\bar{K}\Theta''\left(\bar{G}\bar{I}\bar{K}\right)}{\Theta'\left(\bar{G}\bar{I}\bar{K},1\right)}\right)\\
&<\frac{\bar{G}^{\sigma}}{\left(1-\delta+\Phi_2\left(\bar{G}\bar{I}\bar{K},1\right)\right)^2}\left(\sigma-\frac{-\bar{G}\bar{I}\bar{K}\Theta''\left(\bar{G}\bar{I}\bar{K}\right)}{\Theta'\left(\bar{G}\bar{I}\bar{K},1\right)}\right)\leq 0.
\end{aligned}$$

□

Proof of Proposition 2. For equilibrium existence: The planner's problem is a globally concave maximization problem, which ensures the existence of a unique optimal contingent consumption and investment plan$\{c_t, I_t, K_t\}$. Given this sequence, we can derive the corresponding prices $\{q_t^n, q_t, d_t, w_t\}$ from the producers' first-order conditions. We then allocate the aggregate consumption and capital to the agents $\{c_t^h, k_t^h\}_{h\in\{O,P\}}$ in proportion to their initial capital holdings. This allocation, along with the derived prices, satisfies all the conditions for a sequential competitive equilibrium as defined in Definition 1.

For equilibrium uniqueness: Given any sequential competitive equilibrium, the agents' utility maximization problems are globally concave, and therefore each admits a unique solution. Since the initial conditions of the two agents are proportional, their optimal consumption and capital paths must also be proportional, preserving the same ratio throughout time. The resulting aggregate sequences for consumption, investment, and capital satisfy the first-order conditions of the planner's problem. Thus, they coincide with the planner's optimal allocation. Because the planner's problem admits a unique solution, it follows that the sequential competitive equilibrium is also unique. □

C Numerical Algorithm

To solve for the Markov equilibria of this model, I use the time-iteration algorithm developed in Cao et al. (2023a). In the model with financial assets in Subsection 5, I solve for the policy and transition functions simultaneously during the time-iterations (STPFI).

In addition, we normalized the Lagrange multipliers by present marginal utilities in order to have fixed bounds for these multipliers over the iterations. More specifically, I rewrite the financial asset Euler equation (12) as

$$1 = \frac{1}{p_t \left(\hat{c}_t^h\right)^{-\sigma}} \beta \mathbb{E}_t \left[\left(\frac{A_{t+1}}{A_t} \right)^{-\sigma} (\hat{c}_{t+1}^h)^{-\sigma} \right] + \tilde{\mu}_t^h,$$

where

$$\tilde{\mu}_t^h = \frac{\mu_t^h A_t^{-\sigma}}{p_t \left(\hat{c}_t^h\right)^{-\sigma}}.$$

This implies

$$0 \leq \tilde{\mu}_t^h \leq 1$$

in equilibrium. Similarly, the real asset Euler equation can be rewritten as

$$\begin{aligned} 1 = & \frac{1}{q_t^n \left(\hat{c}_t^h\right)^{-\sigma}} \beta \mathbb{E}_t \left[\left(\frac{A_{t+1}}{A_t} \right)^{-\sigma} (q_{t+1} + d_{t+1}) \left(\hat{c}_{t+1}^h\right)^{-\sigma} \right] \\ & + \tilde{\lambda}_t^h + \frac{p_t}{q_t^n} \tilde{\mu}_t^h (1-m) \min_{s^{t+1}|s^t} \left(q_{t+1}(s^{t+1}) + d(s_{t+1}) \right) \end{aligned}$$

where

$$\tilde{\lambda}_t^h = \frac{\lambda_t^h A_t^{-\sigma}}{q_t^n \left(\hat{c}_t^h\right)^{-\sigma}}.$$

This implies

$$0 \leq \tilde{\lambda}_t^h \leq 1.$$

The collateral constraint (8) could be re-written as:

$$\zeta_t^h = \frac{\hat{b}_t^h}{(1-m) \min_{s^{t+1}|s^t} \left(q_{t+1}(s^{t+1}) + d(s_{t+1})\right) \hat{K}_t} + \frac{k_t^h}{K_t} \geq 0.$$

In equilibrium

$$\zeta_t^O + \zeta_t^P = 1.$$

Therefore,

$$0 \leq \zeta_t^h \leq 1.$$

Figure 9 and Figure 10 present the policy functions for the model with homogenous beliefs in Section 4. The price to dividend ratio and the investment rate are decreasing strongly in the scaled capital stock K_t/A_t They are increasing in the productivity of Sector B relative to Sector A, B_t/A_t, but not as strongly.

Figure 11 and Figure 12 present the policy functions for the model with heterogenous

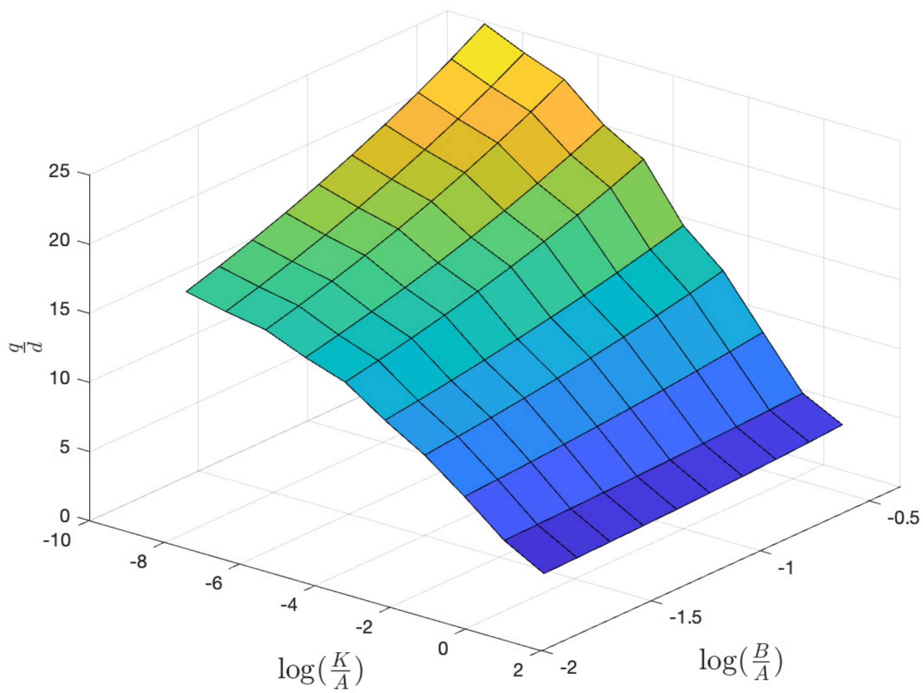


Figure 9: Policy Function: Homogenous Beliefs

beliefs and leverage in Section 5. The share of capital owned by the optimists are increasing in their financial wealth shares. The excess borrowing capacity ζ_t^O is zero when ω_t^O is low, i.e., the optimists' borrowing constraint binds strictly.

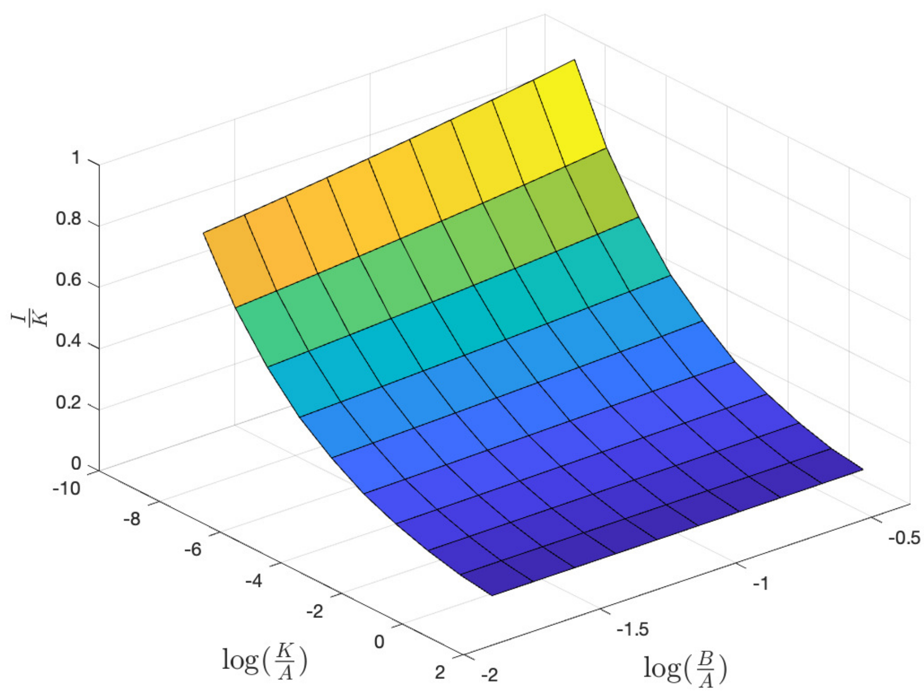


Figure 10: Policy Function: Homogenous Beliefs

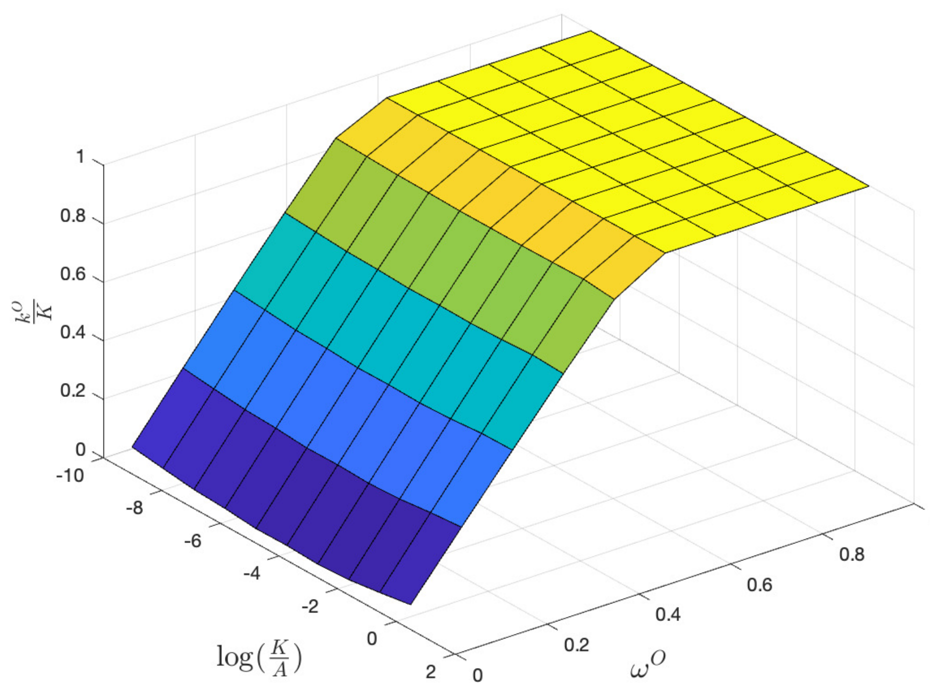


Figure 11: Policy Function: Heterogeneous Beliefs with Leverage

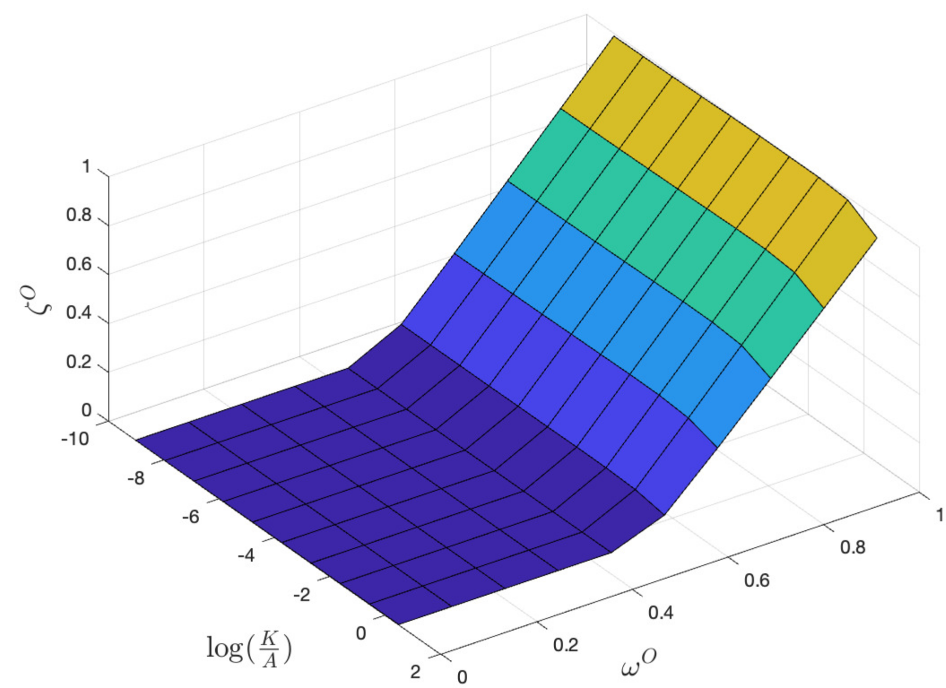


Figure 12: Policy Function: Heterogeneous Beliefs with Leverage