Lagrange Optimisation, Adjustment Costs and Tobin's Q

Investment, Finance and Asset Prices ECON5068

Thomas Walsh

Adam Smith Business School

Lecture Overview

- Toolbox: Optimisation by Lagrange Multiplier
- Adjustment Cost Model
- Tobin's Q

Lagrangians

Lagrangian Function and Lagrange Multiplier

- A useful formulation for **constrained optimization** is the concept of **Lagrangian** function, named after mathematician Joseph-Louis Lagrange
- With **Lagrange multipliers**, the Lagrangian incorporates all constraints into a single function.
- Any constrained optimization becomes **unconstrained** (= easy to solve).
- The multipliers have intuitive **economic interpretation** as a kind of exchange rate.

Constrained Optimization - Primal Problem

$$\max_{\mathbf{x}} f(\mathbf{x})$$
$$\text{subject to: } g_i(\mathbf{x}) = c_i \quad \text{for } i = 1, \ldots, m$$

- $f(\mathbf{x})$ is the objective function.
- $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$
- Constraints usually written as $g_i(\mathbf{x}) = 0$.

$$\max_{x,y} \left\{ F(x,y) = x^2 e^y + 3xy^2 \right\} \quad s.t. \quad x^2 + y^2 = 5$$

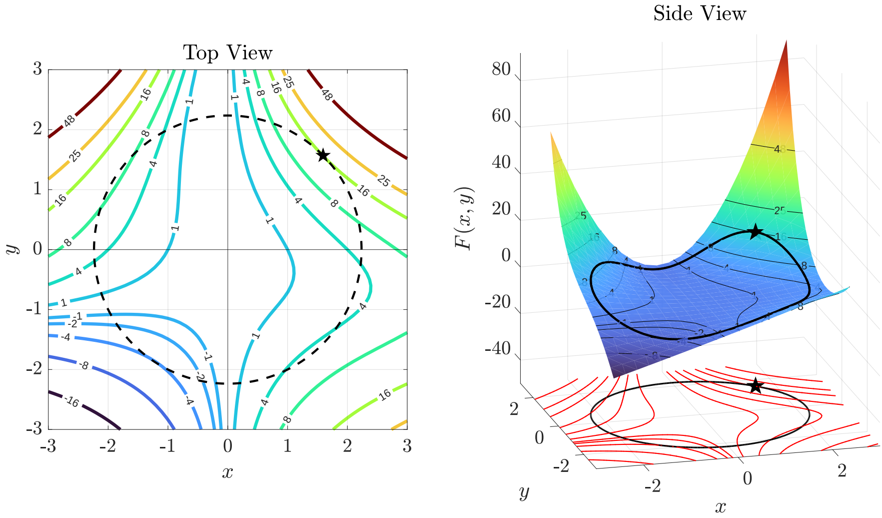


Figure 1: Top and Side Views of Optimised $F(x,y)$ $s.t.$ $g(x,y) = c$

Lagrange Function - Dual Problem

- Lagrangian function, $\mathcal{L}$:

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \sum_i \lambda_i (g_i(\mathbf{x}) - c_i)$$

- $\lambda = (\lambda_1, \ldots, \lambda_m)'$ are **Lagrange multipliers**.
- Equivalent dual maximization:

$$\max_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda)$$

which is an unconstrained maximization problem.[1]

[1] strictly, written fully as: $\min_{\lambda_i} \max_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda)$

Lagrangian Example

Original optimisation:

$$\max_{x,y} \left\{ F(x,y) = x^2 e^y + 3xy^2 \right\} \quad s.t. \quad x^2 + y^2 = 5$$

becomes the Lagrangian of x, y and new variable, λ:

$$\mathcal{L}(x, y, \lambda) = x^2 e^y + 3xy^2 - \lambda(x^2 + y^2 - 5)$$

Lagrange Function

- First-order conditions (FOCs):

$$\frac{\partial \mathcal{L}(x, \lambda)}{\partial x_i} = 0, \quad i = 1, \ldots, n$$

$$\frac{\partial \mathcal{L}(x, \lambda)}{\partial \lambda_i} = 0, \quad i = 1, \ldots, m$$

- If f is concave and g_i convex and differentiable, conditions are **sufficient for a maximum**.

Where does the Lagrangian come from?

$$\max \mathbf{F}(\mathbf{x}, \mathbf{y}) \ s.t. \ \mathbf{g}(\mathbf{x}, \mathbf{y}) = \mathbf{c} \tag{1}$$

➔ **Must stay on constraint**, budget is fixed (total differential $dg = 0$):

$$dg = g_x dx + g_y dy = 0 \Rightarrow dy = -\frac{g_x}{g_y} dx \tag{2}$$

➔ **How does F change** along g-contour, for small steps (dx, dy):

$$dF = F_x dx + F_y dy \tag{3}$$

$$= (F_x - \frac{F_y g_x}{g_y}) dx \tag{4}$$

➔ **At optimum**, $\frac{dF}{dx} = 0$, ratios F_i/g_i are a constant: λ

$$(F_x/g_x - F_y/g_y) = 0 \tag{5}$$

Where does the Lagrangian come from?

F and g have been transformed into a new system

System defined by 2 new optimality ratios and 1 level constraint

$$F_x = \lambda g_x \tag{6}$$

$$F_y = \lambda g_y \tag{7}$$

$$g(x, y) = c \tag{8}$$

The function $\mathcal{L}(x, y, \lambda)$ will give exactly [FOCs $\mathcal{L}_x, \mathcal{L}y, \mathcal{L}_\lambda = 0$] we need:

$$\begin{aligned}\mathcal{L}(x, y, \lambda) &= F(x, y) - \lambda(g(x, y) - c)\\ &= F(x, y) + \lambda(c - g(x, y))\end{aligned}$$

The Lagrange Multiplier has an Economic Interpretation

Interpretation

The Lagrange multiplier can be interpreted as the rate of change in the maximal value of the objective function as the constraint is relaxed

$$\lambda_i^* = \frac{\partial F(x^*)}{\partial c_i} = \frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial c_i}$$

- **Shadow price**: converts one unit to another (e.g.: £$ to utility)

$$\partial F(x^*) = \lambda_i^* \partial c_i$$

➤ combine total differentials with FOCs

we will go over this on next slide, but don't worry the practice questions will guide you through examples

Shadow Price

At the optimum choices given c:

TD the constraint wrt c

Sub FOCs

Direct Attack w/ envelope condition[2]: $\frac{d\mathcal{L}}{dc} = \frac{\partial \mathcal{L}}{\partial c}$

[2] See Practice Questions

General Approach

- You need to maximize an (objective) function:

$$\max_x f(x)$$

subject to some constraints

$$g(x) = 0$$

- Step 1: Write down the Lagrange function that converts this to an unconstrained maximization problem by penalizing any constraint violations:

$$\mathcal{L}(x, \lambda) = f(x) - \lambda_i g_i(x)$$

- Step 2: First order conditions:

$$\frac{\partial \mathcal{L}}{\partial x_i} = 0, \quad \frac{\partial \mathcal{L}}{\partial \lambda_i} = 0$$

Lagrange Function: Transversality Conditions

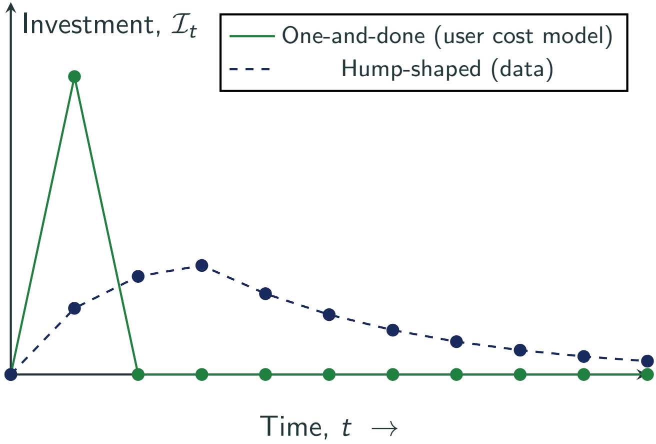
- In infinite horizon problems, transversality conditions are used to prevent divergence as $t \to \infty$.
- e.g.: I could transfer more and more of my wealth to the infinitely far future, and consume $c_t \to \infty$
- Discounting future payoffs so

$$\lim_{T \to \infty} \beta^T \pi_T = 0 \quad if \quad \beta \in [0, 1)$$

- All models in this course satisfy these conditions.

The “Adjustment Cost” or “Tobin” Model

Motivation: Investment Dynamics of a firm over time



- Firm investment response to changes in economic conditions in UC Model and more realistic path like in data

Adjustment Cost Model

- **Costs** arise when the capital stock is **adjusted quickly**.
- **Expansion** (or reversal) of capital is **painful**
- Examples: **installation, training, learning, shutdowns.**
  - installations (/removals) have specific requirements (skills, other machines)
  - works must be trained or get experience using new capital
  - replaced machine cannot produce while it is being removed
- We focus on **convex** adjustment costs
  - $(\mathcal{AC}'(x) > 0, \mathcal{AC}''(x) > 0)$
  - **more smaller changes** favoured over one very large installation
  - **humps** vs one-and-done spikes in $\mathcal{I}_t$ in data
- Introduces **Tobin's Q**.

Adjustment Costs: Internal vs External

- **Internal**: direct costs of changing capital stocks. e.g:
  - **installation**
  - **training** workers to operate new machines
  - **temporary shutdowns** or other disruption
  - **overtime** or slack
  - **management burden:** integrating new projects, restructuring departments
  - **supply chains** must be coordinated

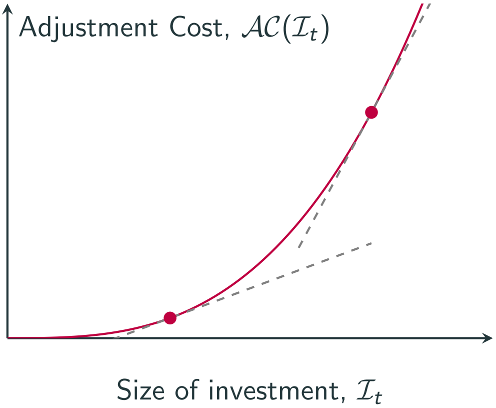
- (**External**: capital prices fluctuate)

- We focus on **internal** adjustment costs.

Adjustment Cost Model - Assumptions

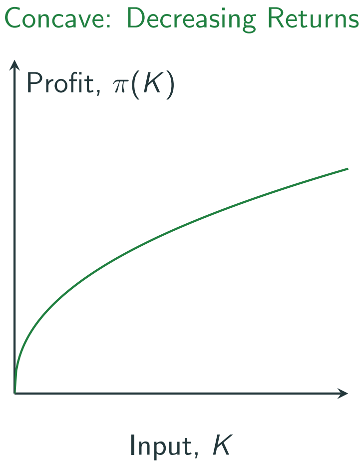
- **Infinite time horizon**, $T = \infty$, firm lives forever, so **no entry/exit**
  - think as unknown very far away end point
  - Firm treats every day as "business as usual"
  - Doesn't worry about exit / end conditions on an average day
  - discounting = the very fare future has tiny extra value
- The firm **maximises its value** = present value of dividends
- Constant interest rate, r
- **Convex, increasing adjustment cost**, $\mathcal{AC}(\mathcal{I}_t)$

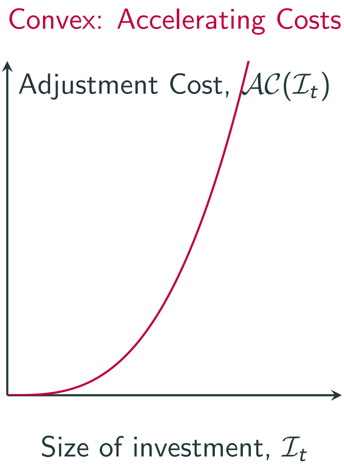
Convex Costs



- **Increasing Marginal Cost** of installation (slope) in investment
- Ikea furniture: building the next wardrobe is harder than the last (**tiredness accumulates**)
- $\mathcal{AC}(1) = 1, \mathcal{AC}(2) = 4, \mathcal{AC}(3) = 9, ...$

Reminder: Concave vs. Convex shapes





- Tip: con**C**ave looks a bit like a C, con**V**ex looks like a V?

Firm's Problem

- **Value**: The Value of the firm at time t is given by:

$$V_t = \mathbb{E}_t\left[\sum_{i=0}^{\infty}\left(\frac{1}{1+r}\right)^i D_{t+i}\right]$$

- **Expectations**: We use the operator $\mathbb{E}_t$ since *future* dividends are random variables, but we can forecast given information *today*

- **LOM**: The firm aims to maximise this Value, given that current investments $\mathcal{I}_t$ become productive with a 1-period lag. Capital therefore follows this law of motion:

$$K_{t+1} = (1-\delta)K_t + \mathcal{I}_t \quad \forall t$$

Adjustment Costs

Adjustment Every unit of investment incurs quadratic adjustment cost $\mathcal{AC}_t$, representing lost revenues of disruption, compatibility issues etc.

$$\mathcal{AC}_t = \frac{\phi}{2}(\mathcal{I}_t)^2$$

- Note: When $\phi = 0$, we have no adjustment costs.
- This AC:
  - lost revenues scale convexly with investment in levels.
  - $\frac{\kappa}{2}(\frac{I}{K})^2$ scales with **investment rate**
  - $\frac{\mu}{2}(\frac{I}{K})^2 K$ scales with **investment rate, invariant to firm size**
- see practice problems for this last cost function

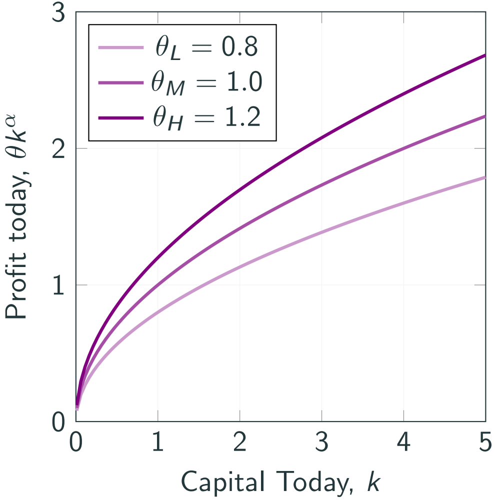
Dividends can be defined as profits net of investment expenditures and adjustment costs:

$$D_t = \pi(\theta_t, K_t) - \mathcal{I}_t - \mathcal{AC}_t$$

Here we have assumed that the price of capital is one.

- $\pi(\theta_t, K_t)$: **profit function**,e.g $\theta_t K_t^\alpha$
- θ_t: **productivity shock**
- Modeled as **stochastic** process, future is **uncertain**, but we can form **conditional expectations**, and know we will be **optimising**

Effect of θ in $\pi(\theta, k) = \theta k^\alpha$



- Scales profits: shrinks/magnifies profit function by factor θ

Lagrangian Formulation

the complete firm's problem as follows. Objective function:

$$\max_{\{I_t\}_0^\infty} \ \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\pi(\theta_t, K_t) - \mathcal{I}_t - \frac{\phi}{2} \mathcal{I}_t^2 \right]$$

subject to:

$$K_{t+1} = (1-\delta) K_t + \mathcal{I}_t \quad \forall t$$

- where $\beta = \frac{1}{1+r}$ is the discount factor; $r = \frac{1-\beta}{\beta}$is the discount rate
- If we specify one time-preference parameter, it implies the other.

Lagrangian Formulation

$$\mathcal{L} = \max_{\{I_t\}_0^\infty} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\pi(\theta_t, K_t) - \mathcal{I}_t - \frac{\phi}{2}\mathcal{I}_t^2 - q_t(K_{t+1} - (1-\delta)K_t - \mathcal{I}_t) \right] \tag{9}$$

For each period t = 0,1,2,..., we have:

- Flow operating profits
- costs of investing
- law of motion constraint
- lagrange multiplier, q_t

Initial condition: Assume firm starts with some known capital and productivity (K_0, θ_0).

First Order Conditions

$$\frac{\partial L}{\partial \mathcal{I}_t} = 0 \Rightarrow q_t = 1 + \phi \mathcal{I}_t \tag{10}$$

$$\frac{\partial L}{\partial K_{t+1}} = 0 \Rightarrow q_t = \beta \mathbb{E}_t \big[\pi_K(\theta_{t+1}, K_{t+1}) + q_{t+1}(1-\delta) \big] \tag{11}$$

$$\frac{\partial L}{\partial q_t} = 0 \Rightarrow K_{t+1} = (1-\delta) K_t + \mathcal{I}_t \tag{12}$$

Investment Rule

Marginal Cost of Investment = Marginal Benefit of Investment

$$1 + \phi \mathcal{I}_t = \beta \mathbb{E}_t \Big(\pi_K(\theta_{t+1}, K_{t+1}) + q_{t+1}(1-\delta) \Big)$$

$$\textit{price} + \textit{marginal AC} = \beta \mathbb{E}_t (\textit{ MPK} + \textit{shadow value of capital})$$

- LHS: **marginal cost** of an additional unit of capital, the price of capital ($p_k = 1$) plus the marginal adjustment cost ($\phi \mathcal{I}_t$).
- RHS: **expected discounted value** of marginal profitability and **value of non-depreciated capital**.
- **Shadow v market prices:** The firm prices capital at q compared to the market for captial goods $p_k = 1$

Recursive Form

We can keep expanding FOC(K_{t+1}) by **recursive subsitution**, we can move time forward one period, and substitute on the RHS[3]

$$q_t = \beta \mathbb{E}_t \left[\pi_K(\theta_{t+1}, K_{t+1}) + q_{t+1}(1-\delta) \right];$$

$$\text{(Law of iterated expectations:} E_t(E_{t+k}[X]) = E_t(X)) \tag{13}$$

$$= \beta \mathbb{E}_t \left[MPK_{t+1} + \beta(1-\delta) MPK_{t+2} + \beta(1-\delta)^2 q_{t+2} \right]$$

$$= \beta \mathbb{E}_t \left[MPK_{t+1} + \beta(1-\delta) MPK_{t+2} + \beta^2(1-\delta)^2 MPK_{t+3} + ... \right] \tag{14}$$

[3] "My best guess today of what my best guess will be tomorrow ...**must already be my best guess today**". This is saying we only have information up to time t for **all** future forecasting

Full Sequence

$$q_t = \beta \mathbb{E}_t \left[MPK_{t+1} + \beta(1-\delta) MPK_{t+2} + \beta^2 (1-\delta)^2 MPK_{t+3} + ... \right]$$

Let h be the number of steps into the future from today:

$$q_t = \beta \mathbb{E}_t \sum_{h=0}^{\infty} \beta^h (1-\delta)^h MPK_{(t+1)+h} \quad (15)$$

Interpretation: The firm values capital according to the marginal increase in profits generated for the rest of its useful lifetime

Interpretation

- The multiplier q_t gives us **shadow price of capital**
- The shadow price describes how much the **value of the firm** will rise if we were to have an **additional unit of capital**.
- The advantage of this model is that we have also defined the value of capital or the value of firm

Adjustment Cost Model: Remarks

- Since the price of a new capital good is equal to one, the optimal investment rule says to **keep investing in capital until the marginal value of this action given by q_t equals its cost.**

- q_t is called **Marginal Q** or **Tobin's Q**, named after the economist **James Tobin** (1918-2002) winner of the 1981 Nobel Prize

Q-theory a.k.a. Tobin's Q

Tobin's Q

From the first order condition in eq. (6), we have:

- $q_t = 1 + \phi \mathcal{I}_t$

in terms of investments:

- $I_t = \frac{1}{\phi}(q_t - 1)$

Tobin Model Investment Rule:

Investment positive iff $q_t > 1$

- ϕ controls sensitivity of investment to changes in q
- investment should ONLY be a function of q, ϕ and other parameters

Tobin's Q Model: The Pros

✔ **Intuitive rule**: The investment rule clearly shows that investment depends on future expected profitability. Since capital is durable and capital boosts production and profits this makes sense.

✔ **Sufficient statistic:** q_t or marginal Q is what we call in statistics a sufficient statistic for investment

- That is, knowing Q is sufficient to understand all relevant information related to the investment decision

✔ **More realistic Hump-shaped dynamics**: no more one-and-done, slow decay (some investment over many periods), sometimes hump-shaped

Responses after a shock to Earnings

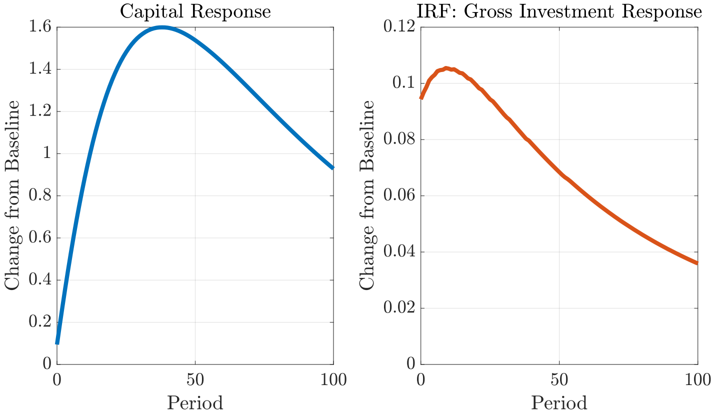


Figure 2: Responses of Capital and Investment to a Shock to Revenue

- Very sensitive to model parameters; so probably can't generate hump with realistic values

Tobin's Q Model: The Cons

- ✖ **Always-investing problem**: firms respond continuously to changes in the environment.

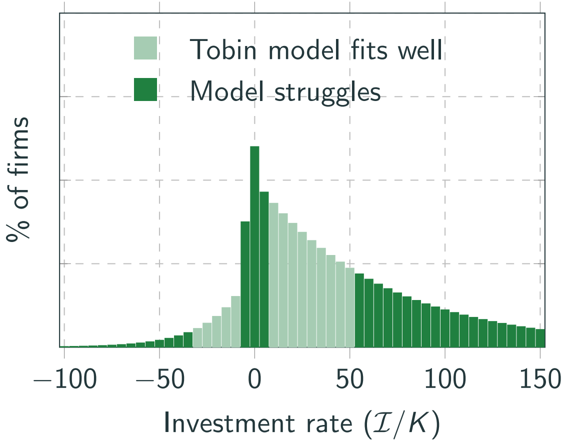
⇒ Investment predicted to always be small and continuous in Tobin's world, its never 0

- ✖ **Lumps and Bumps** Unfortunately, this is not true in empirical data where investment is lumpy

⇒ Firms often go many periods with no significant adjusment before beginning the installation cycle

- ✖ **Zeroes**: predicts too little inaction (≈ 0s),
- ✖ **Spikes**: the model underestimates extreme investment events in the tails of the distribution (not enough mega-installs, e.g. $>$100%)

Investment Distribution



Firm Value and Investment

- Firm value ≈ stock market value (market cap)
- Future expected profits raise firm value
- Tobin model says: Stock market value (expected profits), & investment comove together

Expectation of
future profitable opportunities
$\mathbb{E}_t(MPK_{t+1}, MPK_{t+2}, ...)$

↓

Higher Value, V_t → + Stock Market Capitalisation (valuation £$)

↓

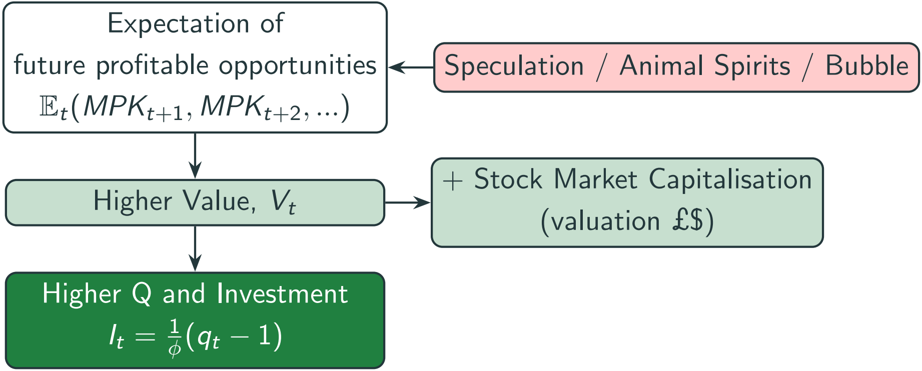
Higher Q and Investment
$I_t = \frac{1}{\phi}(q_t - 1)$

Empirical Confirmation

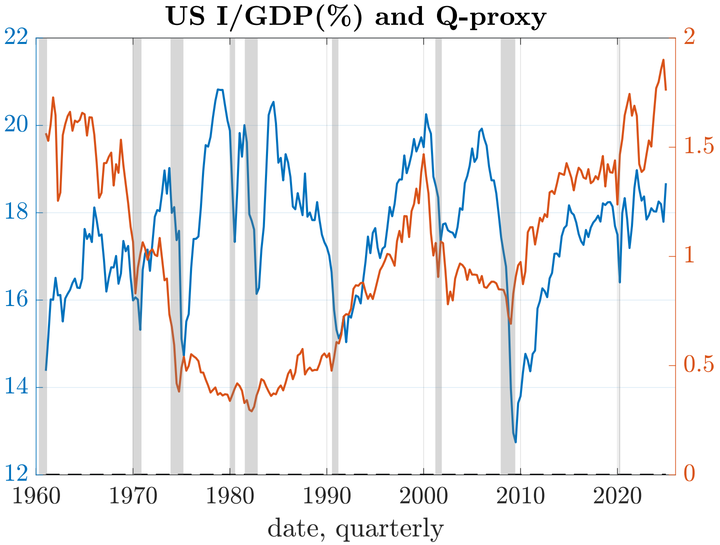
- Example: 1990s tech boom on NASDAQ
  - Stock prices & investment surged (and crashed) together

- Right Now: US Tech Giants (Google, Meta, Amazon, AI firms): surging market caps, large capex
  - ignores bubble dynamics (see: pets.com)
  - Ultra-large players have other strategic reasons for high capex
  - recall the challenges of valuing intangibles (lecture 1)
  - Market is betting (expecting) that large AI investments will pay off

Firm Value and Investment



... in the data



US I/GDP(%) and Q-proxy
22
20
18
16
14
12
2
1.5
1
0.5
0
1960
1970
1980
1990
2000
2010
2020
date, quarterly

Further Reading

- Gregory Chow, *Dynamic Economics: Optimization by the Lagrange Method*, Chapter 1: 1.1–1.3, 1.8