Lecture 10. Cross-sectional asset pricing empirics and Expected utility

Lawrence Schmidt

MIT Sloan School of Management

15.470/Fall 2018

Outline

Today: Empirical asset pricing tests and Expected utility

- Finish material on testing the single-factor APT (CAPM) model
- Then, switch to "structural" material & cover choice under uncertainty.
- Campbell Chapter 1 contains an excellent, concise overview of this material.
- Problem set #3 now posted. One more prior to Midterm

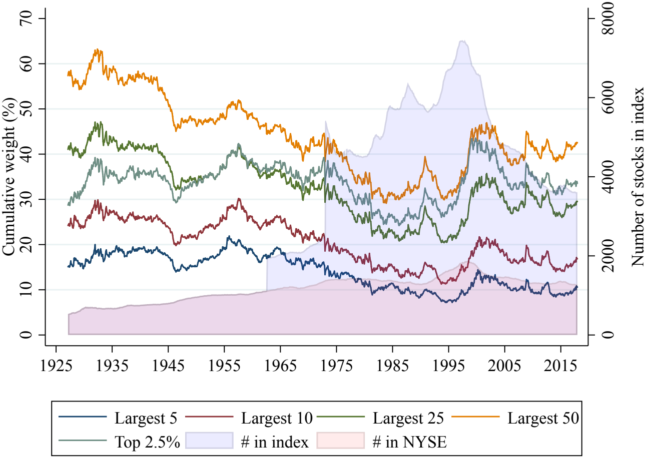
Implications of APT

- With a large number of securities and a small number of risk factors, APT applies (to most – but not all – securities):

$$\bar{r}_i - r_F = \sum_{k=1}^{K} \beta_{ik} \left(\bar{r}_k - r_F \right) = \sum_{k=1}^{K} \beta_{ik} \lambda_k, \ \forall \ i.$$

- β_{ik} is the factor loading of security i on factor k.
- $\bar{r}_k - r_F = \lambda_k$ is risk premium of traded portfolio that mimics factor k.
- The implementation of APT requires the determination of factors and the corresponding factor premiums.
- Note that APT places no restrictions on signs/magnitudes of λ_k, and there could in principle be a large number of factors that affect risk premia
- Main prediction of APT is equivalent to a model of the SDF which is affine in the factors: $\eta = a + Fb$. See Cochrane Chapter 6 for details

In Practice: Why might APT not apply to all securities?



- One reason: market value of public firms is extremely concentrated among a small number of stocks (source: my work with Sung Je Byun)
- Relative supply of these mega-cap stocks is so large that, in the aggregate, *someone* must be exposed to their idiosyncratic risks.

Testing the APT (+ any linear model of the SDF)

- Suppose that I have an empirical proxy for the factor(s) and want to test whether the prediction of the theory holds
- Want to test if expected returns lie on the Securities Market Line (hyperplane w/ > 1 factor) implied by the expected return of the factors
- Two (closely related) common approaches are taken in the literature:
  1. Time-series: Is alpha zero in a regression of excess asset returns on the factors? Only applicable w/ traded factors (or factor-mimicking portfolios)
  2. Cross-sectional: Do estimates of beta explain average returns? Make an empirical counterpart to the SML.
- Note: this material follows Campbell Chapter 3.3

Time-series approach

- Get data (usually monthly) T-bill rates, asset returns, and factor returns
- Run a time-series regression:

$$r_{i,t} - r_{f,t} = \alpha_i + \sum_k \beta_{i,k} (\underbrace{r_{k,t} - r_{F,t}}_{\substack{\text{factor } k \text{ mimicking} \\ \text{portfolio excess return}}}) + \epsilon_{i,t}$$

for $t = 1, \ldots, T$ for an individual tradable asset or a portfolio

- APT predicts that $\alpha_i = 0$
- Test whether $\hat{\alpha}_i = 0$. Can also do this for multiple assets and consider a joint statistical test for all $\{\hat{\alpha}_i\}_{i=1}^N = 0$
- Implicit assumption: $\beta_{i,k}$ is constant over time. Individual stock betas often time-varying and estimated imprecisely $\Rightarrow$ common to use portfolios

Cross-sectional approach

First step:

- Get data (usually monthly) T-bill rates, asset returns, and factor returns
- Estimate betas for each asset on factors by running a time-series regression:

$$r_{i,t} - r_{F,t} = \alpha_i + \sum_k \beta_{i,k} F_{k,t} + \epsilon_{i,t}$$

and save $\{\alpha_i, \beta_i\}_{i=1}^N$

Second step:

- Use your data from the first step to run a cross-sectional regression:

$$\frac{1}{T}\sum_{t=1}^{T}(r_{i,t} - rR_{f,t}) \xrightarrow{p} E[r_i - r_F] = \lambda_0 + \sum_{k=1}^{K} \lambda_k \hat{\beta}_{i,k} + u_i$$

for $i = 1, \ldots, N$

- APT predicts $\lambda_0 = 0$ and $\lambda_k = E[f_{k,t}]$ if factors are tradable excess returns.
- Geometrically, you are estimating the SML with a linear regression
- Works w/ non-traded factors too! Difference: intercept λ_0 isn't restricted.
- Perfect model + enough data $\Rightarrow u_i = 0$ for all i, $R^2 = 100\%$

Testing the APT: Practical considerations

- Goal: characterize all major sources of systematic risk and individual stocks/portfolios' exposures to them. Data $\Rightarrow$ risk prices λ_k
- Step 1: find data on the factors. Sources used in practice:
  - Returns (endogenous): Factor/principal component analysis (PCA).
  - Fundamentals: Macroeconomic shocks (consumption, GDP, etc.)
- Step 2: estimate β and λ. Harder than it sounds! Potential challenges:
  - Validity of the factor model: Are residuals actually uncorrelated or are there missing factors?
  - Stability of factor structure and factor premiums: β and λ may vary over time and be hard to estimate
    - Exposures to macro factors often particularly hard to estimate: often only available at low frequencies and there may be lead/lag relationships
    - Realized returns may not accurately reflect ex-ante expected returns. Beware of data mining
  - Interpretation: factors may work empirically but not be interpretable

The Cross-Sectional Asset Pricing Literature

- Eugene Fama and Kenneth French started a massive literature which tries to identify sources of priced systematic risk in the data
- Their approach (still the standard):
  1. Use observable characteristics of individual stocks (e.g., accounting variables) to form portfolios
  2. Test the ability of a given factor model to explain variation in expected returns across portfolios in the data
- If model can't explain the data, call it an **asset pricing anomaly**
- Because 1 factor explains majority of comovements across stocks + reasons that you'll learn about later in the course, the literature began by testing the Capital Asset Pricing Model (CAPM)
- CAPM is equivalent to the APT using the value-weighted portfolio of all traded assets as a single pricing factor.

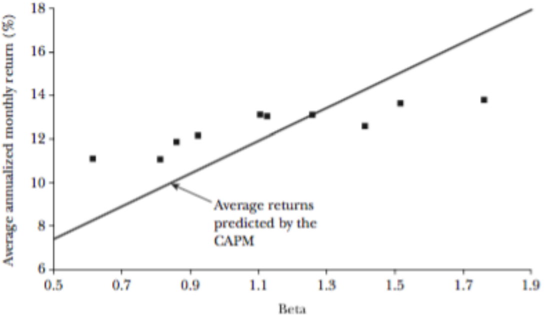
Asset pricing patterns unexplained by the CAPM

- Fama and French (1992) showed that the CAPM could not explain expected returns of portfolios sorted on several very simple firm characteristics:
  1. **Beta**: Using data from prior to time t estimated β, then formed portfolios sorting on these estimates
  2. **Size**: A small cap stock is a stock with a low market capitalization. F-F compared returns of small cap vs large cap stocks
  3. **Book to Market Ratio**: is the ratio of its book value to its market value. Equivalently, it is the ratio of the book value per share to the stock price per share.
     - ★ Growth stocks have low book-to-market ratios, tend to have high expected growth rates in earnings, sales, cash flow, or book value.
     - ★ Value stocks have high book-to-market ratios, thought to have lower growth prospects
- Hundreds of other anomalies have been discovered (see Campbell book), one of the most notable and robust being **momentum**: sorting on returns over the prior year

Historical average returns: Sorting on Beta

Figure 2

Average Annualized Monthly Return versus Beta for Value Weight Portfolios Formed on Prior Beta, 1928–2003



- Qualititatively, things look good.
- Quantitatively, line is too flat
- Source: Fama and French (2004)

Historical average returns: Sorting on Size (market cap)

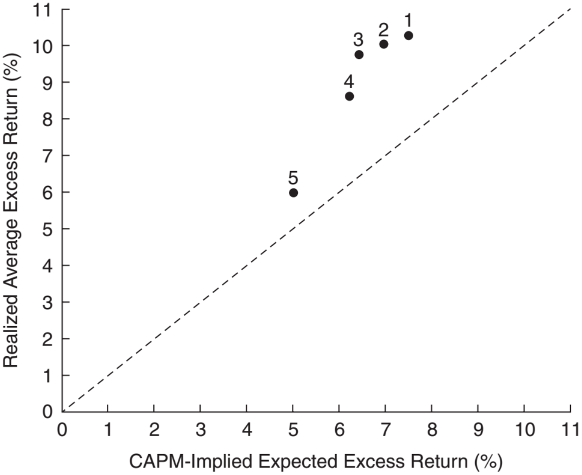


Figure 3.5. *The CAPM and the Size Effect*

- Slope is too steep relative to what CAPM would predict

Size effect was concentrated in January

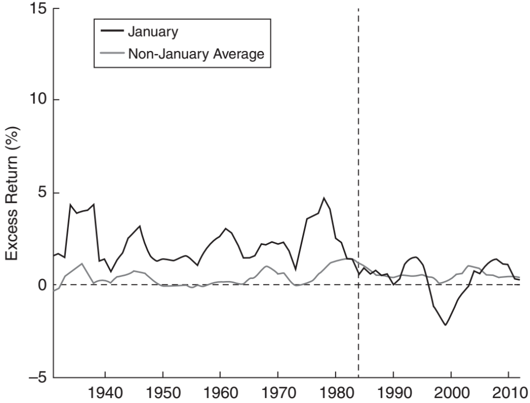


Figure 3.9. *Five-Year Moving Average Excess Returns to Small-Cap Stocks, January vs. Other Months*

- Rationale: capital losses on stocks are tax-deductible, investors tend to sell at end of year to take deduction. Hits smaller, less-liquid stocks harder
- Vertical line is when original paper documenting the phenomenon was published. Magnitudes smaller now!

Size effect was concentrated in January

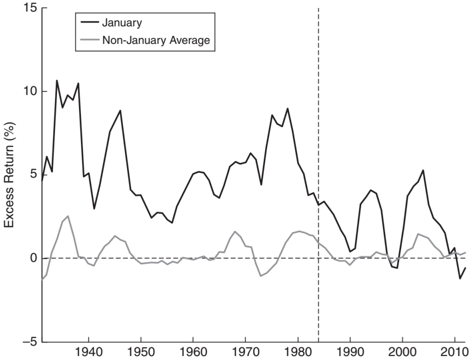


Figure 3.10. *Five-Year Moving Average Excess Returns to Micro-Cap Stocks, January vs. Other Months*

- Magnitudes even bigger for really small (micro-cap) stocks, but also decay
- One reaction to smaller effects later in sample: investors are trading to eliminate them. McLean & Pontiff (2016) show that volumes & volatility increase following publication: this is anomaly elimination

Historical average returns: Sorting on Book-to-market ratio

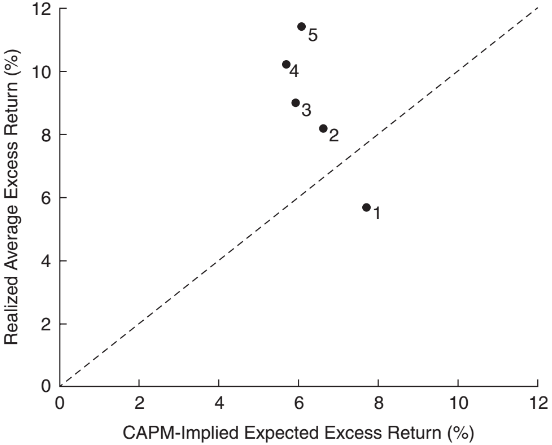
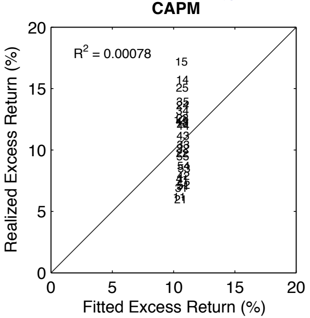


Figure 3.6. *The CAPM and the Value Effect*

- Betas and expected returns not even *qualitatively* related in direction predicted by the CAPM (though this feature can change depending on sample)

Things look even worse if we sort on Size AND Value



CAPM$_\alpha$					t-value				
−5.14	1.84	3.84	7.65	8.08	−1.38	0.61	1.45	2.90	2.82
−3.99	1.51	4.58	5.87	6.79	−1.62	0.79	2.12	2.64	2.90
−2.23	2.45	3.36	5.33	6.31	−1.34	1.52	1.97	2.56	2.55
−0.55	0.90	3.68	4.65	3.98	−0.36	0.66	2.41	2.48	1.96
−0.46	0.40	2.22	1.85	1.88	−0.38	0.44	1.99	1.26	0.98

Rows: sort on size from small to big

Columns sort on book-to-market from low (growth) to high (value)

Sample: 1954-2003. Source: Jagannathan and Wang (2007)

Historical average returns: Prior returns (Momentum)

Table I

Momentum Portfolio Returns

This table reports the monthly returns for momentum portfolios formed based on past six-month returns and held for six months. P1 is the equal-weighted portfolio of 10 percent of the stocks with the highest returns over the previous six months, P2 is the equal-weighted portfolio of the 10 percent of the stocks with the next highest returns, and so on. The "All stocks" sample includes all stocks traded on the NYSE, AMEX, or Nasdaq excluding stocks priced less than $5 at the beginning of the holding period and stocks in the smallest market cap decile (NYSE size decile cutoff). The "Small Cap" and "Large Cap" subsamples comprise stocks in the "All Stocks" sample that are smaller and larger than the median market cap NYSE stock respectively. "EWI" is the returns on the equal-weighted index of stocks in each sample.

	All Stocks			Small Cap			Large Cap		
	1965–1998	1965–1989	1990–1998	1965–1998	1965–1989	1990–1998	1965–1998	1965–1989	1990–1998
P1 (Past winners)	1.65	1.63	1.69	1.70	1.69	1.73	1.56	1.52	1.66
P2	1.39	1.41	1.32	1.45	1.50	1.33	1.25	1.24	1.27
P3	1.28	1.30	1.21	1.37	1.42	1.23	1.12	1.10	1.19
P4	1.19	1.21	1.13	1.26	1.34	1.05	1.10	1.07	1.20
P5	1.17	1.18	1.12	1.26	1.33	1.06	1.05	1.00	1.19
P6	1.13	1.15	1.09	1.19	1.26	1.01	1.09	1.05	1.20
P7	1.11	1.12	1.09	1.14	1.20	0.99	1.09	1.04	1.23
P8	1.05	1.05	1.03	1.09	1.17	0.89	1.04	1.00	1.17
P9	0.90	0.94	0.77	0.84	0.95	0.54	1.00	0.96	1.09
P10 (Past losers)	0.42	0.46	0.30	0.28	0.35	0.08	0.70	0.68	0.78
P1–P10	1.23	1.17	1.39	1.42	1.34	1.65	0.86	0.85	0.88
t statistic	6.46	4.96	4.71	7.41	5.60	5.74	4.34	3.55	2.59
EWI	1.09	1.10	1.04	1.13	1.19	0.98	1.03	1.00	1.12

Source: Jegadeesh and Titman (2001)

Historical CAPM alphas: Prior returns (Momentum)

	CAPM Alpha
P1	0.46
	(3.03)
P2	0.29
	(2.86)
P3	0.21
	(2.53)
P4	0.15
	(1.92)
P5	0.13
	(1.70)
P6	0.10
	(1.22)
P7	0.07
	(0.75)
P8	−0.02
	(−0.19)
P9	−0.21
	(−1.69)
P10	−0.79
	(−4.59)
P1–P10	1.24
	(6.50)

T-stats are in parentheses. Source: Jegadeesh and Titman (2001)

Orders of magnitude are large

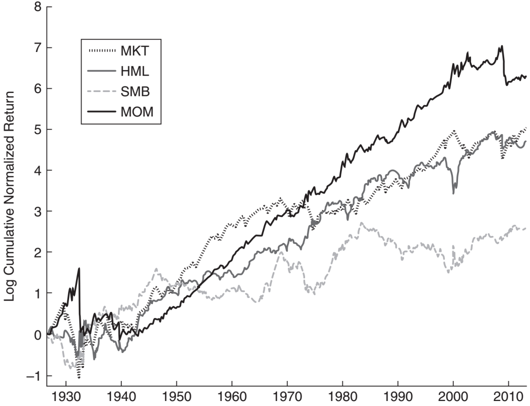


Figure 3.7. *Fama-French-Carhart Log Cumulative Normalized Factor Returns, 1926–2013*

Some anomaly portfolios earn expected returns of the same order of magnitude, or even slightly larger, than the equity premium (MKT $\equiv E[r_m - r_f]$) (!!)
Source: Campbell Textbook

Interpretations of CAPM failures

- These anomaly portfolios also feature stocks that strongly co-move $\Rightarrow$ there are additional sources of systematic risk in the data
- APT $\Rightarrow$ any systematic factor may be priced. Work with multi-factor representations. Most popular empirical specifications:
  - Fama-French 3 factor model: add Value and Size portfolios as additional pricing factors
  - Carhart 4 factor model: Fama French + momentum
  - Fama-French now have a 5 factor model
  - Lu Zhang + coauthors: factor models sorting on investment
  - Rise of the machines! Examples: Giglio and Xiu (2017); Kelly, Pruitt, & Siu (2018), among many others. This is a rapidly growing area
- Another reaction: much of this is a statistical fluke. Data snooping has led to a "zoo" of factors, many of which are spurious.
- Limits to arbitrage: many "overpriced" stocks are a small fraction of the market and hard to short-sell.

One reaction to the "factor zoo"

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Predicting anomaly performance with politics, the weather, global warming, sunspots, and the stars

Robert Novy-Marx

Simon Graduate School of Business, University of Rochester, 500 Joseph C. Wilson Blvd., Box 270100, Rochester, NY 14627, United States

ARTICLE INFO

Article history:
Received 18 September 2012
Received in revised form
2 August 2013
Accepted 5 August 2013
Available online 11 February 2014

JEL classification:

ABSTRACT

Predictive regressions find that the party of the US president, the weather in Manhattan, global warming, the El Niño phenomenon, sunspots, and the conjunctions of the planets all have significant power predicting the performance of popular anomalies. The interpretation of these results has important implications for the asset pricing literature.

- Along these lines, if you want a laugh, check out `http://www.tylervigen.com/spurious-correlations` An Example!

A similar reaction

… and the Cross-Section of Expected Returns

Campbell R. Harvey
Duke University, National Bureau of Economic Research

Yan Liu
Texas A&M University

Heqing Zhu
The University of Oklahoma

Hundreds of papers and factors attempt to explain the cross-section of expected returns. Given this extensive data mining, it does not make sense to use the usual criteria for establishing significance. Which hurdle should be used for current research? Our paper introduces a new multiple testing framework and provides historical cutoffs from the first empirical tests in 1967 to today. A new factor needs to clear a much higher hurdle, with a t-statistic greater than 3.0. We argue that most claimed research findings in financial economics are likely false. (*JEL* C12, C52, G12)

Where we're going next

- Roughly speaking, this concludes the "reduced-form" part of the course
- Next, we will start working from the bottom up to build a micro-founded model
- One important lesson to keep in mind: the data are highly informative about models and investors' expectations. Ignore data at your own risk!
- Most successful researchers in asset pricing work with both theory and data because they complement one another.

Preferences

Introduction

- We now start to explore the connection between asset prices and the economic activities of the agents in the economy.
- We begin by analyzing how agents utilize the financial market to best meet their economic needs.
- As discussed before, an agent's economic need is captured by her preferences
- A preference is a binary relation $\succsim$ on consumption set C, satisfying:
  1. Completeness,
  2. Transitivity.
- It will be common to assume certain additional properties for preferences

Preferences

Axioms of Preferences

Axiom (Continuity)

$\forall\ c \in C$, the sets $\{a \in C : a \succsim c\}$ and $\{b \in C : b \precsim c\}$ are closed.

Continuity means two close consumption plans should also be "close" in their rankings. (Counterexample: lexicographic preferences)

Axiom (Insatiability)

$a \succ b$ if $a > b$.

Insatiability means more is always preferred to less.

Axiom (Convexity)

$\forall\ a, b, c \in C$ and $\alpha \in (0, 1)$, if $a \succsim b$ and $c \succsim b$, then $\alpha a + (1-\alpha)c \succsim b$.

For continuous preferences, convexity implies that the sets of preferred bundles $\{a \in C : a \succsim c\}$ are convex. Strict convexity replaces $\succsim$ with $\succ$ above

Utility Function

Definition (Utility Function)

A utility function for a preference relation $\succsim$ is a function u from C to R:

$$u: \ C \rightarrow R$$

such that:

$$\forall \ a, \, b \in C, \quad u(a) \geq u(b) \ \text{ if and only if } \ a \succsim b.$$

Theorem (Debreu)

For a preference $\succsim$ defined on a closed, convex consumption set C satisfying the Continuity Axiom, there exist a continuous utility function $u(\cdot)$ on C such that:

$$\forall \ a, b \in C, \quad a \succsim b \ \text{ if and only if } \ u(a) \geq u(b).$$

Given that preferences can be fully defined by a utility function, going forward we will start with the utility function when describing an agent's preference.

If a preference exhibits insatiability, the utility function is strictly increasing.

Expected Utility

Utility and Probability

- A consumption plan c includes consumption in all possible future states.
- The utility function $u(\cdot)$ also depends on probabilities of these states $\Rightarrow$ want to define preferences over consumption lotteries $[c;p]$
- All else constant, we may expect small marginal utility from consumption in a state with very small probability.
- We would like to distinguish the influence on utility by the likelihood of a state, which is exogenous, from the influence by consumption itself.
- In particular, we would like to express the utility function as the expected value of utility over different consumption paths:

$$u(c;p) = \sum_{\omega \in \Omega} p_\omega u_\omega(c_0, c_{1\omega}). \tag{1}$$

This form of utility function is called expected utility.

Expected Utility

Ordinal vs Cardinal preferences

- In general, utility function $u(c;p)$ provides ordinal information (only relative rankings are informative):
  - Given a function $\phi : \mathbb{R} \to \mathbb{R}$ that is strictly increasing ($\phi'(x) > 0$ for all x), alternative utility function $\tilde{u}(c;p) \equiv \phi(u(c;p))$ encodes the same preferences
  - Key concept encoded by utility function is its upper contour set: the set of bundles preferred to c. Units are meaningless.
- Expected utility $u(c;p) = \sum_{\omega \in \Omega} p_\omega u_\omega(c_0, c_{1\omega})$ is no exception
- Can take any monotonic transformation of the *entire expression* for $u(c)$ and preserve same preference ordering
- However, we can only to take affine transformations of the function $u_\omega(c_0, c_{1\omega})$ on the RHS and preserve inequalities across consumption bundles
- von Neumann Morgenstern (vNM) utility function $u_\omega(c_0, c_{1\omega})$ is cardinal. Comparisons like $u_\omega(c_0'; c_{1\omega}') - u_\omega(c_0; c_{1\omega}) > u_\omega(c_0''; c_{1\omega}'') - u_\omega(c_0'; c_{1\omega}')$ are informative.

Expected Utility

Adding uncertainty: Continuity and Independence Axioms

Axiom (Continuity)

For all consumption c, probabilities p_a, p_b, and p_c:

$$[c; p_a] \succsim [c; p_b] \succsim [c; p_c] \Rightarrow \exists \alpha \in (0,1) : [c; (1-\alpha)p_a + \alpha p_c] \sim [c; p_b]$$

Axiom (Independence)

For all consumption c, probabilities p_a, p_b, and p_c and $\alpha \in (0,1)$:

$$[c; p_a] \succsim [c; p_b] \Rightarrow [c; (1-\alpha)p_a + \alpha p_c] \succsim [c; (1-\alpha)p_b + \alpha p_c]$$

- IA in words: if lottery A is preferred to lottery B, then a compound lottery that mixes A and B with C w/ prob α has the same preference ordering.
- It is easy to show that expected utility satisfies the Independence Axiom.

Expected Utility

Theorem (Debreu)

Under the Independence Axiom, the utility function can be expressed in the form of expected utility as given in (1).

Proof. The necessity of the Independence Axiom for expected utility is given before. See Debreu (1959) for the proof of sufficiency.

- Expected utility has intuitive appeal: overall utility from different consumption paths is the average utility from each possible consumption path, corresponding to each state, weighted by its probability.
- A formulation of the expected utility theory concerning only random payoffs (without consumption at 0) was developed by von Neumann and Morgenstern (1947).
- Is expected utility a good description of preferences? We can attack this from normative vs. positive perspectives

Additional Simplifications

State Independence

- For simplicity, we consider further simplifications of the expected utility.
- State Independence: Utility over a consumption path depends only on the consumption levels along the path (state), not the path/state itself.
- That is:

$$u(c) = \sum_{\omega \in \Omega} p_\omega u(c_0, c_{1\omega}). \tag{2}$$

- State independence is mostly for parsimony.
- Benefit: preferences apply to *any* consumption lottery (e.g., changes in Ω)!
- State dependence can be a convenient modeling device for certain effects.

Additional Simplifications

Time Additivity

- Consumption at different times can be complements or substitutes:

$$\frac{\partial^2 u(c_0, c_{1\omega})}{\partial c_0 \partial c_{1\omega}} \neq 0.$$

  - Complements if the above derivative is positive.
  - Substitutes if the above derivative is negative.
- For simplicity, we assume no complementarity and substitutability.
- Time Additivity: Utility over a consumption path is the sum of utility over consumption at each date:

$$u(c_0, c_{1\omega}) = u_0(c_0) + u_1(c_{1\omega}). \tag{3}$$

- Often, we may even assume that:

$$u(c_0, c_{1\omega}) = u(c_0) + \rho\, u(c_{1\omega}), \quad \rho > 0.$$

ρ is called the time preference/discount coefficient.

Additional Simplifications

State-Independent and Time-Additive Expected Utility

A state-independent and time-additive, discounted expected utility function then takes the following form:

$$u(c_0) + \rho \sum_{\omega \in \Omega} p_\omega u(c_{1\omega}), \quad \rho > 0. \tag{4}$$

This will often be used as a canonical specification of an agent's utility function.

Going forward, we also assume that $u(\cdot)$ is twice differentiable.

- We will refer to $u'(c)$ as the marginal utility at consumption level c.
- Insatiability implies

  $$u'(\cdot) > 0,$$

  i.e., the agent has strictly increasing utility or positive marginal utility.

Critiques/Extensions: Behavioral Biases and Beyond

- Evidence inconsistent with expected utility
  - Allais paradox: evidence of direct violations of the independence axiom
  - Ellsberg paradox: people dislike ambiguous lotteries
  - Rabin critique: rejecting small mean zero bets implies ridiculous aversion to large bets.
  - . . .
- More flexible utility functions – "behavioral economics/finance"
  - Habit formation
  - Catching up with the Jones
  - First order risk aversion (prospect theory/generalized disappointment aversion)
  - Uncertainty/ambiguity aversion
  - Difference in beliefs, general state-dependent preferences
  - . . .
- Are these concerns really about preferences?

Concavity of Utility Function

Axiom (Convexity of preferences)

$\forall\ a, b, c \in C$ and $\alpha \in (0,\ 1)$, if $a \succsim b$ and $c \succsim b$, then $\alpha a + (1-\alpha)c \succsim b$.

For continuous preferences, convexity implies that the sets of preferred bundles $\{a \in C : a \succsim c\}$ are convex. Strict convexity replaces $\succsim$ with $\succ$ above

Definition

A function $u(\cdot)$ is concave if $\forall\ x$, x' and $\alpha \in [0, 1]$,

$$u(\alpha x + (1-\alpha)x') \geq \alpha u(x) + (1-\alpha)u(x').$$

Theorem (Concavity)

If a preference satisfies the Continuity, Independence and Convexity axioms, and can be represented by a discounted expected utility function of the form:

$$u(c_0) + \rho \sum_{\omega \in \Omega} p_\omega u(c_{1\omega}), \quad \rho > 0,$$

then $u(\cdot)$ is a concave function.

Decreasing Marginal Utility

Theorem

If a concave utility function $u(\cdot)$ *is twice differentiable, then* $u'' \leq 0$.

Negative second order derivative for $u(\cdot)$ means decreasing marginal utility.