

Empirical Asset Pricing: Introduction

Ralph S.J. Koijen Stijn Van Nieuwerburgh*

March 17, 2024

*Koijen: University of Chicago, Booth School of Business, NBER, and CEPR. Van Nieuwerburgh: Columbia Business School, CEPR, and NBER. If you find typos, or have any comments or suggestions, then please let us know via ralph.koijen@chicagobooth.edu or svnieuwe@gsb.columbia.edu.

1. Basic structure of the notes

- High-level summary of theoretical frameworks to interpret empirical facts.
- Per asset class, we will discuss:
 1. Key empirical facts in terms of prices (unconditional and conditional risk premia) and asset ownership.
 2. Interpret the facts using the theoretical frameworks.
 3. Facts and theories linking financial markets and the real economy.
 4. Active areas of research and some potentially interesting directions for future research.
- The notes cover the following asset classes:
 1. Equities (weeks 1-5).
 - Predictability and the term structure of risk (week 1)
 - Cross-section and the factor zoo (week 2)
 - Intermediary-based asset pricing (week 3)
 - Production-based asset pricing (week 4)
 - Asset pricing via demand systems (week 5)
 2. Mutual Funds and Hedge Funds (week 6).
 3. Options and volatility (week 7).
 4. Government bonds (week 8).
 5. Corporate bonds and CDS (week 9).
 6. Currencies and international finance (week 10).
 7. Commodities (week 11).
 8. Real estate (week 12).

2. High-level summary of theoretical frameworks

- To organize and interpret the empirical facts, various theoretical frameworks will be used.
 1. **Factor models** for the cross-section of expected returns.
 2. **Stochastic discount factor** (SDF) *models* based on “a” marginal investor.
 3. **Structural equilibrium models**: Endowment and production economies.
 4. Equilibrium models: **Asset pricing using demand systems**.
- By moving from the first to the third framework, we impose more structure and hence there are more testable predictions.
- Every equilibrium asset pricing model implies a demand system. Oftentimes, the implications are not very interesting (e.g., the representative-agent model) or the implications have not been tested using data on asset holdings. The fourth framework tries to match holdings data and asset pricing data jointly.
- We start with a high-level overview of these frameworks. Classic asset pricing textbooks provide a more comprehensive treatment of the first and third frameworks.

2.1. Factor Models

- Cochrane's *Asset Pricing* book contains a detailed treatment of factor models, including their estimation.
- The absence of arbitrage opportunities implies the existence of a positive stochastic discount factor (SDF), M_{t+1} , which is unique if markets are complete.
- Given how competitive financial markets are, we typically assume that there are no arbitrage opportunities (after accounting properly for trading costs).
- For any valid SDF M_{t+1} , we have

$$E_t[M_{t+1}R_{t+1}^e] = 0,$$

for any excess return, $R_{t+1}^e = R_{t+1} - R_t^f$, where R_{t+1} denotes the return and R_t^f the one-period risk-free rate.

- Straightforward rewriting implies

$$E_t[R_{t+1}^e] = \beta_t \lambda_t,$$

where $\beta_t = Cov_t[R_{t+1}^e, M_{t+1}] / Var_t[M_{t+1}]$ and $\lambda_t = -Var_t[M_{t+1}] R_t^f$.

- **Implication:** The absence of arbitrage opportunities implies a single-factor model for expected returns, where the single factor is the stochastic discount factor.

- We do not observe M_t . It is common practice to use linear factor models to approximate the SDF:

$$M_{t+1} = a + b'F_{t+1},$$

with interpretable factors (e.g., a business cycle factor like consumption growth or investment growth, an interest rate factor) that plausibly correspond to states of high marginal utility for households, binding constraints of institutions,

- The above derivation shows, however, that any successful K -factor model implies a 1-factor model.
- Ideally, we find a low-factor representation of the SDF with a clear economic interpretation of the factors.
- Two standard ways to motivate factor models:

1. APT:

- The APT starts with a statistical description of returns. In particular, suppose that there exists a K -factor model in returns

$$R_{i,t} = a_i + \beta_i'F_t + e_{i,t},$$

where $F_t \in \mathbb{R}^K$, $E[e_{i,t} | F_t] = 0$, and $Var(e_t)$ is a diagonal matrix (this can be relaxed to allow for some dependence that vanishes in large, diversified portfolios).

- The absence of arbitrage implies:

$$E[R_{i,t+1}] = \lambda_0 + \beta_i'\lambda,$$

where $\beta, \lambda \in \mathbb{R}^K$. λ_0 is the zero-beta rate, which equals

the risk-free rate when it exists.

- The key observation is that idiosyncratic risk is not priced and only systematic risk factors matter for expected returns.
- However, the APT does not tell you what the factors are, nor what the signs and magnitudes of the risk prices are.
 - ⇒ It may well be the case that the risk price of a given factor is zero while the factor explains a significant fraction of realized returns.

2. ICAPM / Equilibrium models:

- Equilibrium asset pricing models imply a SDF, and therefore specify the relevant pricing factors.
- The ICAPM suggests that we identify innovations to factors that capture the distribution of future returns, summarizing future investment opportunities.
- The ICAPM often implies testable restrictions that the factors should predict future returns and return variances, or restrictions on the signs (and sometimes magnitudes) of risk prices.
- See [Maio and Santa-Clara \(2012\)](#) for a discussion and tests of the restrictions on multi-factor models implied by the ICAPM.
- We will see various examples of this in different asset classes.

- Why are factor models useful?

1. A reduced-form, low-dimensional description of expected returns may guide the [design of equilibrium models](#).
 - The reduced-form factor literature may be able to identify the key factors that price the cross-section of expected returns. This may provide new moment conditions for theoretical models.
 - For instance, the Fama and French 3-factor model ([Fama and French, 1992](#)) had a large impact on the theoretical asset pricing literature where researchers tried to explain the value premium. Three examples are [Zhang \(2005\)](#), [Lettau and Wachter \(2007\)](#), and [Garleanu, Panageas, and Yu \(2012\)](#).
 - Analogously, in the currency literature, [Lustig, Rousanov, and Verdelhan, 2011](#) summarize the main dimensions of risk premia that is guiding the design of international finance models.
2. If a factor model holds perfectly, [optimal portfolio choice reduces to choosing an optimal portfolio over the factors](#). This dramatically reduces the number of parameters that need to be estimated, which could dramatically improve the **stability** of the optimal portfolio.

Consider the standard mean-variance solution

$$w = \frac{1}{\gamma} \Sigma^{-1} \mu.$$

If we plug in sample estimates of average returns and the covariance matrix, the solution tends to be unstable and perform poorly out of sample.

If a factor model holds perfectly, we can replace

$$\mu = \beta\lambda,$$

where we can estimate β accurately and we can estimate λ potentially over longer samples. Moreover, if we have a statistical factor model of returns, we need to estimate fewer parameters for the covariance matrix Σ .

3. **Performance measurement** of mutual funds and hedge funds.

If you find non-zero alphas relative to the factor model, you want to deviate from holding a portfolio based on the factors only.

Both points above explain the huge interest from the asset management industry in factor models / factor investing.

4. It may help help to compute the **cost of capital for firms** in case of equities and corporate bonds.

5. However, in practice, estimating factor betas and risk premia remains a difficult task (**Fama and French, 1997**).

- Limitations of factor models

1. Factor models are silent about the underlying **economics** behind the factors. For instance, the fact that we see co-movement for certain anomalies such as momentum does not mean that there is a rational, risk-based explanation. A factor structure in irrational beliefs (for instance, extrapolation as in [Barberis, Greenwood, Shleifer, and Yin, 2015](#)) can also generate co-movement and can give rise to a factor model. See [Kozak, Nagel, and Santosh \(2020\)](#) for a detailed discussion on the interpretation of factor models. All that is needed is that riskless arbitrage opportunities do not exist and there exists a single-factor model. This is a very weak condition.

The value anomaly is a prime example where researchers (still) disagree whether the value premium is compensation for risk ([Campbell, Giglio, Polk, and Turley, 2018](#) or [Koijen, Lustig, and Van Nieuwerburgh, 2017](#)) or due to expectation errors (e.g., [La Porta, Lakonishok, Shleifer, and Vishny 1997](#)).

2. Factor models cannot be used to think about policy questions or counterfactuals. This requires a fully-specified equilibrium model.

Of course, both points simply reflect the fact that factor models are reduced-form ways of looking at data.

- Some broader concerns with the current factor model literature

1. The number of factors has exploded. We now have over 400 factors that (supposedly) drive return differences in the cross-section of expected returns. Once we add (non-linear) combinations of these factors, we have thousands. This provides little guidance for the design of equilibrium models. Dimension-reducing techniques are necessary to guide theory. Much more on this in week 2.
2. The current generation of factor models appears rather “static.”
 - The nature of shocks changes all the time (e.g., the credit crisis in the U.S., the earthquake in Japan, the Greek crisis/Euro crisis, geopolitical risk in Russia or the Middle East, COVID-19 crisis, ...).
 - It would be surprising if changes in the nature of shocks would map into stable factor structures that apply uniformly across asset classes and across countries.
 - One view is that some shocks are structural, giving rise to the traditional, stable factors. Also, some factors (e.g., value) may generally capture cross-sectional dispersion in expected returns, regardless of its source.
 - However, the fragility of factor models may be due to the changing nature of shocks or demand curves.
 - To understand whether “changing risks” are important for asset pricing, we may need different types of models or data.
 - In this context, it may be interesting to explore *text data*. So-called **topic models** are designed to capture

exactly the changing nature of factors (Blei, Ng, and Jordan, 2003).

- Text data is high-dimensional, and hence sensitive to small variations in methodology. Most of the work is taking off-the-shelf methods from machine learning. See Gentzkow, Kelly, Taddy (2019 JEL) for an introduction.
 - One concern is that any statistical approach will always have limitations. Using economic priors to discipline the dimension-reduction problem, analogous to VAR models in macro (Del Negro and Schorfheide, 2004) may be a productive way to go.
3. Although it would have been nice to uncover a simple structure, the evidence so far suggests that the research agenda to uncover a simple factor model has made limited progress. There are many different factors for each asset class and the factors across asset classes and countries are not that highly correlated.
 4. This is not because the field did not try. There has been a large research effort in both academia and in the industry.
 5. We may need to explore more (non-return) data and may need to use more theory to understand the beliefs and objectives of the major investors in different asset classes. This is where the work on “the” marginal investor is relevant. Then, we need to impose market clearing and figure out equilibrium asset prices. This is where equilibrium models come in.

Traditional structural asset pricing models with heterogeneous agents (who differ in their preferences, beliefs, trading/production technologies, or endowments) have made

some strides in simultaneously accounting for prices and quantities.

However, with current computational techniques, these methods run into curse of dimensionality problems once more than two or three (types) of agents are considered.

We will discuss recent advances in demand-based equilibrium asset pricing models which can deal with more agents. In both cases, taking seriously the role of institutions, such as financial intermediaries, has gained new prominence in the aftermath of the Great Financial Crisis.

2.2. SDFs Based on “A” Marginal Investor

- Instead of modeling the behavior of all investors, a natural starting point may be to model the behavior of a class of investors.
- For instance, we can model the behavior of households that hold stocks or institutions that trade in various asset classes (e.g., broker-dealers).
- Assume households maximize their value function over consumption and their investment portfolio,

$$\max \sum_{s=0}^{\infty} \beta^s E_t [u(C_{t+s})],$$

for some utility function $u(\cdot)$.

- The first-order condition is given by

$$E_t \left[\beta \frac{u'(C_{t+1})}{u'(C_t)} R_{t+1}^e \right] = 0.$$

- We can check whether this condition holds for subsets of households (e.g., rich, highly-educated households). This is not necessarily the representative agent.
- Importantly, in this literature, we do not impose market clearing.

- To understand equilibrium asset prices, we want to understand the investment decisions of the major investors in a given asset class. E.g., mutual funds for equities, insurance companies for corporate bonds, and banks for mortgage-backed securities.
- We will see examples below where first-order conditions appear to hold, but the investors in question are small players in a given asset class. What do we learn in such case?
- Their first-order condition still provides a valid stochastic discount factor, but these investors are likely to be price-taking agents. Hence, this may be a useful way to understand the priced sources of risk (risk factors) in a particular asset class. This represents significant progress.
- However, in any counterfactual or policy experiment, these investors may not matter much, so that a more holistic view of the asset market is required for normative questions.

2.3. *Equilibrium Models: Endowment and Production Economies*

- The next step up from models that focus on a subgroup of investors (first-order conditions) is to impose market clearing.
- Endowment economies specify preferences, endowments, and beliefs, starting from the classic paper by Lucas (1978).
- Leading asset pricing models in this class:
 1. The habit model (Campbell and Cochrane, 1999).
 2. The long-run risks model (Bansal and Yaron, 2003).
 3. The variable rare disasters model (Gabaix, 2012 and Wachter, 2013).
- To keep things tractable, these models share three common features:
 1. Representative agent: No heterogeneity.
 2. Institutions and intermediation plays no role.
 3. No frictions (e.g. borrowing constraints).
- Assumptions 1+3 imply that markets are complete. In the Lucas model, agents differ in the realizations of their endowments, but they are all ex-ante identical and can trade a full set of state-contingent securities. This enables them to share risk perfectly. Asset prices are the same as if there were a representative agent.
- Of course, there are many interesting extensions of these models that relax assumptions 1-3.
- Most of the models that relax assumptions 1 and 3 are heterogeneous-agent incomplete markets models which feature imperfect risk

sharing. This is a large and active area of research in both macro-economics and finance. Since closed-form solutions are usually not available, numerical solution techniques and powerful computers are in order.

- [Common justifications for assumption 2 that institutions and intermediaries play no role:](#)
 1. Intermediaries act in the best interest of households who own the capital that is managed by the intermediaries; they are a *veil*.
 2. If intermediaries did not act in the best interest of households, then households would reallocate their capital.
- The first justification requires that your fund manager, your pension fund, . . . knows your $U'(C)$. This is a strong assumption.
- The second justification has testable implications. For instance, if a mutual fund under-performs, we should see households reallocate capital away from that fund. If regulation changes the asset allocation of insurance companies, then households should reallocate capital toward/away from insurers.
- In reality, it may not be easy to move your capital around (e.g., capital locked up in pension funds and insurance companies). Or households may not be aware of the regulatory changes.
- Lots of interesting work can be done to understand whether these frictions are indeed important for asset pricing. So far, these assumptions have been largely taken for granted.
- Using detailed data on flows and holdings of different institutions, we can test whether these mechanisms are active.

- A recent generation of models explicitly models financial institutions, see for instance [Brunnermeier and Pedersen \(2009\)](#) and [He and Krisnamurthy \(2013\)](#). Much more on this in week 3.
- Note also that in models without financial institutions, many interesting policy questions involving such institutions cannot be answered. An example is optimal macro-prudential policy (Ex. [Elenev, Landvoigt, and Van Nieuwerburgh \(2021\)](#)).
- Endowment economy models have dominated the theoretical asset pricing literature for the last two decades. We will discuss the empirical successes and failures of these models.
- Production-based models peel the onion one layer back and endogenize the endowments. They start from a process for technology and model the production and investment decisions of firms, sometimes alongside the consumption decisions of households. Consumption, labor income, and dividend income are now endogenous, which makes it much more challenging to match the data on prices and quantities. More in week 4.
- In some cases, the real side and the asset pricing side can be “disconnected” through particular assumptions about preferences ([Tallerini, 2000](#)). It is unclear whether we want asset markets to be this disconnected/neutral.
- A recent strand of the literature features production economy models with financial intermediaries, see [Adrien and Boyarchenko \(2015\)](#), [Brunnermeier and Sannikov \(2014\)](#), [He and Krisnamurthy \(2019\)](#), and [Elenev, Landvoigt, and Van Nieuwerburgh \(2021\)](#).

- There is virtually no empirical work testing these new models. An interesting challenge is how to match these models to the real world with heterogeneous institutions that differ in terms of their regulatory frameworks and the liquidity of their liabilities (e.g., banks versus insurance companies).

2.4. *Asset Pricing via Demand Systems*

- The models and frameworks discussed so far focus on asset prices and macro-economic quantities (e.g., consumption).
- While some of these models are tested using micro-data such as household-level consumption data, predictions for asset holdings are rarely confronted with actual portfolio holdings data.
- There are many questions that involve a shift in the demand curve of a group of agents:
 - Do large institutions amplify volatility in bad times? Should they be regulated as SIFIs (OFR 2013)?
 - How do large-scale asset purchases (i.e., QE) affect asset prices through institutional holdings?
 - What if bond mutual funds experience large outflows?
 - What is the impact on asset prices if we change the risk weights of banks or insurance companies?
 - What is the impact of China and Japan buying large quantities of U.S. Treasuries?
- To answer these questions, we need a model with
 1. Rich heterogeneity that matches the holdings of different types of institutions and the household sector.
 2. Correct, i.e., empirically plausible, substitution patterns across investors.
- In short, we need a realistic model of the demand system for financial assets.

- Any equilibrium model implies a demand system, but the predictions tend to be rather stark. In particular, assets are often very close substitutes and demand curves are virtually flat.
- **IF** true, then many policy questions have trivial answers and the impact of changing regulation, flows, et cetera, is minimal.
- Although this implication / assumption is shared by many models, there is little direct evidence of its empirical relevance.
- We will discuss in week 5, however, evidence of downward-sloping demand curves in equity and fixed income markets.
- Also, we will discuss a recent literature on demand systems in asset pricing, launched by [Kojen and Yogo \(2019\)](#). This model revisits mostly forgotten work going back to [Brainard and Tobin \(1968\)](#), [Friedman \(1980\)](#) and [Friedman \(1985\)](#).
- These models start from an empirical specification of demand curves, which, combined with market clearing, leads to equilibrium asset prices.
- The model can be “closed” by modeling how households allocate capital to various institutions.