Generative Adversarial Network (GAN) model in Asset Pricing

Wang Lingjie

Overview

- Motivation
- Literature review
- Research objective
- Methodology
- Evaluation metrics
- Data
- Results
- Discussion
- Conclusion

Motivation

Asset pricing

One important goal in asset pricing

- How do we explain the variation in asset returns?

No-arbitrage pricing theory

- Assets are priced through the stochastic discount factor (SDF, also known as pricing kernel)

SDF estimation

However, how do we estimate the SDF?

1. Choose the right factors (data features)
2. Select the right portfolio (combination of individual assets)
3. Estimate the functional form

Literature review

Factor models: benchmark model

- CAPM[a], Fama & French (1993) three-factor and Carhart (1997) four-factor models
- $\Rightarrow$ SDF may be a linear combination of selected factors

[a] Sharpe (1964), Lintner (1965) and Mossin (1966) proposed Capital Asset Pricing Model in the 1960s independently

Expanding the search for factors

- Feng et al. (2020) describes the hundreds of new factors proposed in the literature as a “factor zoo”
- Freyberger et al. (2017) and Feng et al. (2020) found that only a few factors proposed in literature are statistically significant in explaining the asset returns
- $\Rightarrow$ SDF may be a linear combination of **potentially infinite factors**

Literature review (cont.)

Non-linear, non-parametric SDF estimation

- Gu et al. (2020) compared different machine learning models and found neural network model has the highest out-of-sample R^2 in explaining asset returns
- $\Rightarrow$ SDF may be a **general function** of potentially infinite factors

Machine learning models with economic theory

- Chen et al. (2021) Generative Adversarial Network (GAN) model, Gu et al. (2021) Autoencoder network model impose no-arbitrage condition to improve neural network performance
- Chen et al. (2021) further include macroeconomic data in the GAN model to estimate the SDF
- $\Rightarrow$ SDF may be a general function of potentially infinite factors. **Including macroeconomic data and imposing economic theory can further improve empirical SDF estimation**

Research objective

Goal

- Examine the external validity of Chen et al. (2021)'s GAN model using the United Kingdom (U.K.) London Stock Exchange (LSE) data
- Why the UK market? Avoid "home bias": Karolyi (2012) found only 20% of the papers in top finance journals investigated countries outside of the United States

Results

1. GAN model outperformed benchmark four-factor model in terms of the Sharpe ratio
2. Macroeconomic data improves the SDF estimation
3. Extended FF factors are the most important covariates in SDF estimation
4. Extended FF factors are nearly linear in the SDF. However, there are interaction effects between the factors

Methodology

No-arbitrage asset pricing model

- No-arbitrage asset pricing model assumes the existence of an asset independent, time-dependent SDF M_{t+1} such that

$$E_t(M_{t+1} R^e_{t+1,i}) = 0 \quad (1)$$

- SDF can be expressed as an affine transformation of the tangency portfolio F_{t+1} such that $M_{t+1} = a - bF_{t+1}$
- The tangency portfolio is constructed as a weighted portfolio of all assets, we assume the SDF weight ω is a function of firm characteristics data $I_{t,i}$ and macroeconomic data I_t such that $F_{t+1} = \omega(I_t, I_{t,i})^T R^e_{t+1}$
- Without loss of generality, consider $(a, b) = (1, 1)$, then the no-arbitrage asset pricing model can be expressed as a function of the SDF weight ω such that

$$E_t\left[\left(1 - \omega(I_t, I_{t,i})^T R^e_{t+1}\right) R^e_{t+1,i}\right] = 0 \quad (2)$$

Methodology

No-arbitrage pricing loss

- However, if no-arbitrage pricing model alone is insufficient to explain the variations in asset returns, we can define an alternative conditional moments function $g(I_t, I_{t,i})$ such that

$$E_t(M_{t+1}R^e_{t+1,i}g(I_t, I_{t,i})) = 0 \tag{3}$$

which motivates the **pricing loss** used in the GAN model

$$L(\omega, g|I_t, I_{t,i}) = \frac{1}{N}\sum_{i=1}^{N}\left\{E\left[\left(1-\sum_{j=1}^{N}\omega(I_t, I_{t,j})R^e_{t+1,j}\right)R^e_{t+1,j}g(I_t, I_{t,i})\right]\right\}^2 \tag{4}$$

Methodology

Feed-forward neural network model

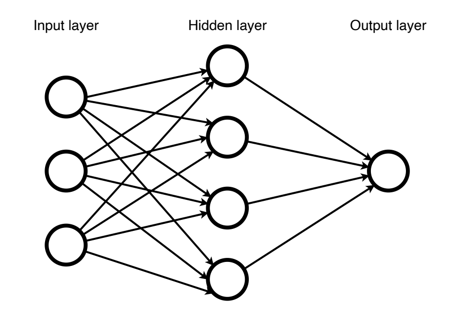

Figure 1: Feed-forward neural network

- Hidden unit: $Z_m = h(\alpha_{0m} + \alpha_m^T \mathbf{X}),\ m = 1, \cdots, M$
- Hidden layer: $\mathbf{Z} = (Z_1, Z_2, \cdots, Z_m)$
- Output unit: $y = \beta_0 + \beta^T \mathbf{Z}$

Methodology

GAN model

- Zero-sum game between two neural networks
  1. Discriminator: estimating $\omega(I_t, I_{t,i})$
  2. Generator: estimating $g(I_t, I_{t,i})$
- Objective: Nash-equilibrium
  - Best discriminator that estimates the SDF
  - Best generator that constructs portfolio and factors that no-arbitrage pricing theory least able to explain
  - Evaluation through no-arbitrage pricing loss function

$$\min_{\omega} \max_{g} L(\omega, g | I_t, I_{t,i})$$

Methodology

GAN model training

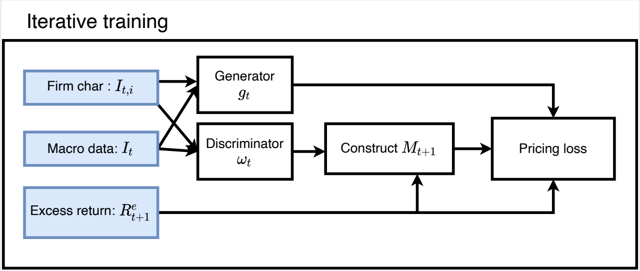

Figure 2: GAN model structure

- Discriminator (ω) estimates SDF weight
- Generator (g) estimates the conditional factors
- Finding the Nash Equilibrium through $\min_\omega \max_g L(\omega, g | I_t, I_{t,i})$

Methodology

Benchmark four-factor model

- Four-factor model
  1. R^e_{mt} (market risk): excess return of market portfolio
  2. SMB (small minus big): outperformance of small market capitalisation companies relative to large market capitalisation companies
  3. HML (high minus low): the outperformance of high book-to-market value companies versus low book-to-market value companies
  4. UMD (momentum): accounts for the speed of price change
- Estimated through OLS
  - $E(\alpha_i) = 0, E(\epsilon_{t,i}) = 0$

$$R^e_{t+1,i} = \alpha_i + \beta_i R^e_{mt} + s_i SMB_t + h_i HML_t + \omega_i UMD_t + \epsilon_{t,i} \quad (5)$$

Evaluation metrics

Sharpe ratio

- Proposed by Sharpe (1966) to evaluate performance of risk-adjusted assets relative to a risk-free asset
- Balanced excess return $(R_i - R^f)$ and the risk of return, measured by the standard deviation of the returns σ_i
- **Higher Sharpe ratio indicates a better performing portfolio**

$$S(R_i) = \frac{E_t[R_i - R^f]}{\sigma_i} \tag{6}$$

Data

United Kingdom London Stock Exchange (1998-2017)

- January 1998 to December 2017 LSE monthly stock prices
- Risk-free rate and four factors: Gregory et al. (2011)
  - R^e_{mt}, SMB, HML, UMD
- Firm characteristic data
  - 2 firm short and mid momentum data (Yahoo! Finance)
  - 14 firm income statement data (Finage LTD)
- Macroeconomic data: Coulombe et al. (2021)
  - 112 macroeconomic indicators
- Total data consist of 132 covariates, 942 stocks with maximum 20 years returns

Data

Data splitting

- Training, validation and out-of-sample testing period
  - Training: 1998 - 2009 (12 years)
  - Validation: 2010 - 2013 (4 years)
  - Out-of-sample: 2014 - 2017 (4 years)
- Two data splits to maximise data availability
  - Remove stocks without complete characteristic data in a particular month. However, Finage database only contains limited income statement data for a subset of stocks

1. Factor data
   - Only consist of factors and macroeconomic data
   - 942 stocks with 116 covariates across maximum 20 years
2. Fundamental data
   - Consist of all covariates, including firm income statement data
   - 242 stocks with 132 covariates across maximum 20 years

Results

Sharpe ratio

Table 1: Sharpe ratio results

Model	Data used	Included Macro data	Train	Valid	Test
Factor	factor	N	0.29	0.62	0.42
GAN	factor	N	0.65	0.49	0.33
GAN	factor	Y	1.88	1.76	1.99
GAN	fundamental	N	0.70	0.35	0.49
GAN	fundamental	Y	1.18	0.39	0.69

- GAN models achieve higher out-of-sample Sharpe ratio than factor model, except for the GAN with only factor data
- Including macroeconomic data improves the Sharpe ratio

Variable Importance

- Sensitivity$(x_j) = \frac{1}{C}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\frac{\partial\omega(I_t, I_{t,i})}{\partial x_j}\right|$

Most important data features

1. Extended FF factors
2. Firm specific past return data
3. Macroeconomic indexes affecting consumers, business and general stock markets

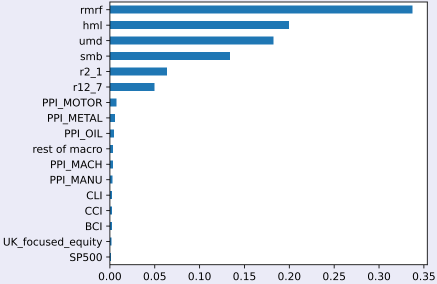

Figure 3: Variable importance

SDF structure

GAN SDF weight structure

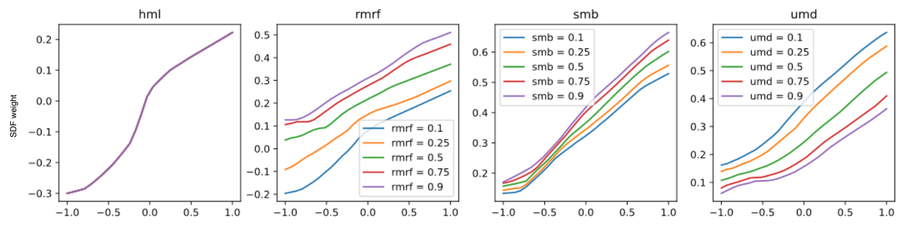

Figure 4: GAN SDF weight as a function of factors

- First image: near linear relationship between `hml` (high minus low) to SDF weight
- Other images: interaction effect between `hml` and `rmrf` (market risk), `smb` (small minus big) and `umd` (momentum)
- Observation holds for other factors as well, similar result as Chen et al. (2021) US SDF estimation

Discussion

Limitation

- Considerable gap between Chen et al. (2021) dataset and this paper
  - Chen et al. (2021): 50 years of historical data with over 10,000 stocks
  - This paper: 20 years of 942 stocks
- However, even the paid data subscription *Finage LTD* (2022) was not able to furnish complete and rich data comparable to the US
- Nonetheless, results in this paper are still relevant in assessing the performance of the GAN model when subject to data limitations

Conclusion

Research objective

Examine the external validity of Chen et al. (2021)'s GAN model using the UK LSE 1998 - 2017 data

Results

1. GAN model outperformed benchmark four-factor model in terms of Sharpe ratio
2. Macroeconomic data improves the SDF estimation
3. Extended FF factors are the most important covariates in SDF estimation
4. Extended FF factors are nearly linear in the SDF. However, there are interaction effects between factors

Questions

Thank you

Annex

SDF as an affine transformation of the tangency portfolio

- $E(M_{t+1}R^e_{t+1,i}) = 0 = E(M_{t+1})E(R^e_{t+1,i}) + Cov(M_{t+1}R^e_{t+1,i})$
- $Cov(M_{t+1}R^e_{t+1,i}) = \rho(M_{t+1}, R^e_{t+1,i})\sigma(M_{t+1})\sigma(R^e_{t+1,i})$
- $\Rightarrow E(M_{t+1})E(R^e_{t+1,i}) = -\rho(M_{t+1}, R^e_{t+1,i})\sigma(M_{t+1})\sigma(R^e_{t+1,i})$
- $\Rightarrow E(R^e_{t+1,i}) = -\rho(M_{t+1}, R^e_{t+1,i})\frac{\sigma(M_{t+1})}{E(M_{t+1})}\sigma(R^e_{t+1,i})$
- Tangency portfolio: a portfolio that maximises mean-variance efficiency
- $\because$ max $\frac{E(F_{t+1})}{\sigma(F_{t+1})}$ when $\rho(M_{t+1}, F_{t+1}) = -1$
- $\therefore M_{t+1} = a - bF_{t+1}$

Empirical no-arbitrage pricing loss

- The **empirical pricing loss** takes into consideration the varying length of ith firm by weighting the number of observation T_i

$$\frac{1}{N}\sum_{i=1}^{N}\frac{T_i}{T}\left[\frac{1}{T_i}\sum_{t\in T_i}\left(1-\sum_{j=1}^{N}\omega(I_t, I_{t,j})R^e_{t+1,j}\right)R^e_{t+1,i}g(I_t, I_{t,i})\right]^2 \quad (7)$$

Long Short-Term Memory (LSTM)

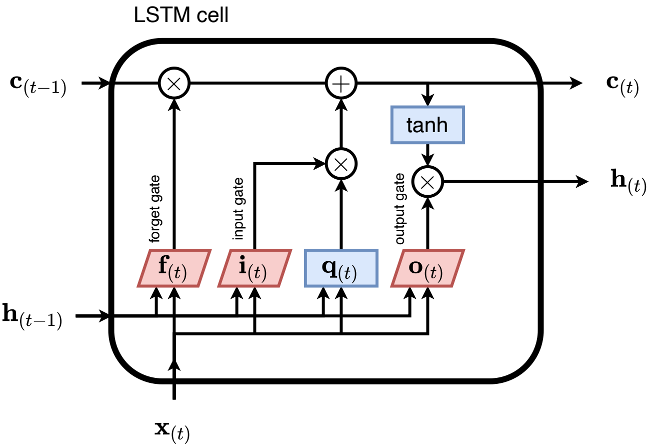

Figure 5: Long short-term memory cell

Long Short-Term Memory (LSTM)

LSTM details

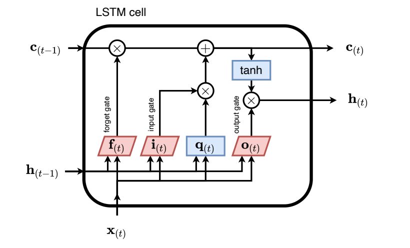

Figure 6: Long short-term memory cell

- input gate: $\mathbf{i}_{(t)} = \sigma(\mathbf{W}_{xi}^T \mathbf{x}_{(t)} + \mathbf{W}_{hi}^T \mathbf{h}_{(t-1)} + \mathbf{b}_i)$
- forget gate: $\mathbf{f}_{(t)} = \sigma(\mathbf{W}_{xf}^T \mathbf{x}_{(t)} + \mathbf{W}_{hf}^T \mathbf{h}_{(t-1)} + \mathbf{b}_f)$
- output gate: $\mathbf{o}_{(t)} = \sigma(\mathbf{W}_{xo}^T \mathbf{x}_{(t)} + \mathbf{W}_{ho}^T \mathbf{h}_{(t-1)} + \mathbf{b}_o)$
- temp long-term state: $\mathbf{q}_{(t)} = tanh(\mathbf{W}_{xg}^T \mathbf{x}_{(t)} + \mathbf{W}_{hg}^T \mathbf{h}_{(t-1)} + \mathbf{b}_g)$
- long-term memory: $\mathbf{c}_{(t)} = \mathbf{f}_{(t)} \otimes \mathbf{c}_{(t-1)} + \mathbf{i}_{(t)} \otimes \mathbf{q}_{(t)}$
- output (short-term memory): $\mathbf{y}_{(t)} = \mathbf{h}_{(t)} = \mathbf{o}_{(t)} \otimes tanh(\mathbf{c}_{(t)})$

Discriminator network structure

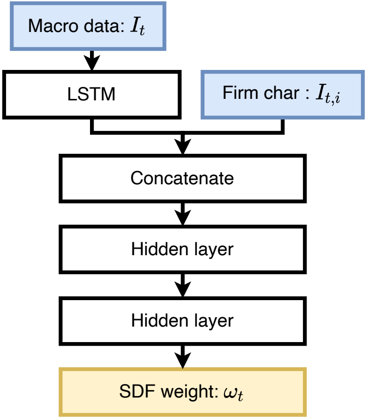

Figure 7: Discriminator structure

Generator network structure

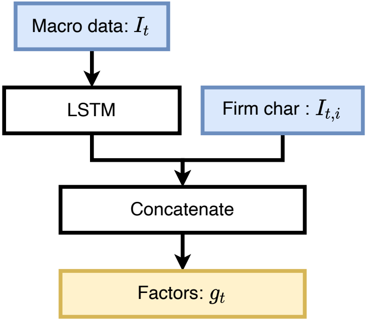

Figure 8: Generator structure

SDF structure

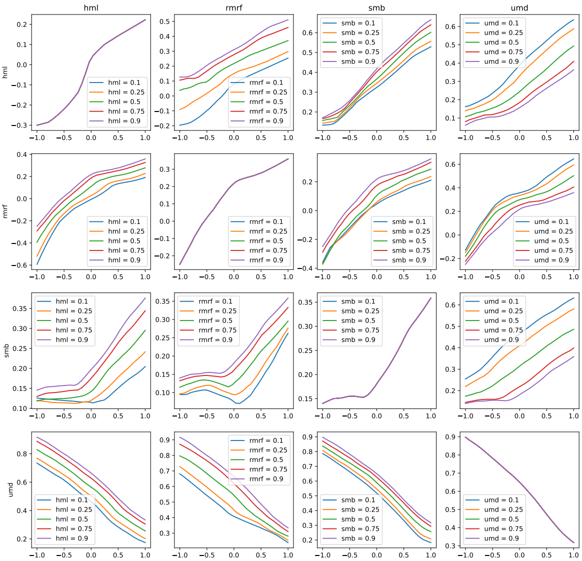

Figure 9: GAN SDF weight as a function of factors

References I

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, *52*, 57–82. https://doi.org/10.2307/2329556

Chen, L., Pelger, M., & Zhu, J. (2021). Deep learning in asset pricing. *Research Methods & Methodology in Accounting eJournal*. https://doi.org/10.2139/ssrn.3350138

Coulombe, P. G., Marcellino, M., & Stevanović, D. (2021). Can machine learning catch the COVID-19 recession? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3796421

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*, 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Feng, G., Giglio, S., & Xiu, D. (2020). Taming the factor zoo: A test of new factors. *The Journal of Finance*, *75*(3), 1327–1370. https://doi.org/10.1111/jofi.12883

Finage LTD. Finage | real-time stock APIs and websocket. (2022). https://finage.co.uk/

References II

Freyberger, J., Neuhierl, A., & Weber, M. (2017). Dissecting characteristics nonparametrically. *CESifo: Macro*. https://doi.org/10.3386/w23227

Gregory, A., Tharyan, R., & Christidis, A. (2011). Constructing and testing alternative versions of the fama-french and carhart models in the UK. *European Finance eJournal*.

Gu, S., Kelly, B. T., & Xiu, D. (2020). Empirical asset pricing via machine learning. *Review of Financial Studies*, *33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Gu, S., Kelly, B., & Xiu, D. (2021). Autoencoder asset pricing models. *Journal of Econometrics*, *222*(1), 429–450. https://doi.org/https://doi.org/10.1016/j.jeconom.2020.07.009

Karolyi, G. (2012). Home bias, an academic puzzle. *Emerging Markets: Finance eJournal*. https://doi.org/10.2139/ssrn.2153206

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, *47*(1), 13–37. https://doi.org/10.2307/1924119

References III

Mossin, J. (1966). Equilibrium in a capital asset market. *The Econometric Society*, *34*, 768–783. https://doi.org/10.2307/1910098

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, *19*(3), 425–442.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, *39*(1), 119–138. http://www.jstor.org/stable/2351741