Week 12: Synthesis, Ethics & Future Directions

FIN306 · Final taught week · Coursework 2 due Week 13

Learning Objectives

- Synthesise 12 weeks of FinTech concepts into coherent framework
- Evaluate ethical implications of financial technology innovations
- Assess AI governance challenges in finance (bias, transparency, accountability)
- Analyse tensions between innovation and regulation
- Identify emerging trends shaping financial services evolution
- Prepare strategic responses to FinTech disruption
- Finalise **FIN306 Coursework 2** (Week 13 submission; confirm date and time on Blackboard)
- Develop professional skills and career pathways in FinTech

Agenda

Part I: Course Synthesis: Integrating themes across 12 weeks
Part II: Ethics & Governance: AI bias, privacy, accountability
Part III: Future Directions: Emerging trends, strategic implications
Part IV: **FIN306 Coursework 2** (Week 13) and careers

Part I: The Spine of FIN306, Statistical Science

FIN306 in one sentence

> ***Data science is statistical science****: the disciplined study of variation and uncertainty in financial data.*

Three challenges from Week 1, applied every week after (Gelman, Hill, and Vehtari 2020):

- **Sample to Population**: do my data represent the market I claim?
- **Treatment to Control**: would the result hold without the signal?
- **Measurement to Construct**: am I measuring what I think I am?

Bias, variance, and the virtue of complexity

- **Bias-variance trade-off** (Week 1): no model is free; we choose where to pay
- More parameters can **help** out-of-sample if we regularise (Kelly, Malamud, and Zhou 2024)
- This reframes "more data" and "more features" from danger into opportunity
- Direct implication for **Scaffold B**: tree ensembles vs OLS on JKP UK (Gu, Kelly, and Xiu 2020)

Why FinTech exists: the cost puzzle

- The unit cost of US financial intermediation has been **roughly constant for 130 years** (Philippon 2016)
- Technology has reshaped retail, telecoms, and manufacturing; **finance has resisted**
- That stubborn cost is the **opportunity** FinTech firms chase
- It is also the **first place to be sceptical**: easy gains have been claimed many times before

Data quality: where models go wrong before they start (Week 2)

- **Survivorship bias**: today's index hides yesterday's failures
- **Look-ahead bias**: information you would not have had in real time
- **Selection bias**: who shows up in your data, and who is missing
- **Measurement to construct gap**: VIX is *implied volatility of an option set*, not "fear"

  *Your **CW1 risk register** was the first formal practice of this discipline. **CW2 Section B** should sound the same.*

Part II: What FIN306 actually built

Predicting return vs predicting risk (Weeks 3 and 4)

- **Returns** (ARIMA, Week 3): tiny R^2, near-zero predictability, and that is honest
- **Risk** (GARCH, Week 4): meaningful R^2, because volatility clusters
- The lesson is not "models fail"; it is **what they can and cannot do**
- VIX example: a good measurement, a contested construct

Robo-advisors and portfolio reality (Week 5)

- Mean-variance optimisation (Markowitz 1952) is fragile under estimation error
- Robo-advisors **widen access** to portfolio construction (Reher and Sokolinski 2024)
- Live questions: governance, suitability, fee transparency
- The Bloomberg lab made the estimation problem concrete

Alternative finance and credit scoring (Week 6)

- Marketplace lending data is **selection-biased by design** (who applies, who is funded)
- Digital footprint variables (device type, email provider) **add about 1.3pp AUC** over standard credit scores (Berg et al. 2020)
- That is real, and also **opaque to the borrower**
- Inclusion and fairness are both genuine concerns; they are not the same concern

Cryptocurrency and rare-event detection (Week 8)

- Crypto microstructure: thin books, regime shifts, fat tails
- Fraud detection as **rare-event classification**: imbalanced classes, walk-forward CV, cost-sensitive thresholds
- Class imbalance changes which metric matters: AUC, precision-recall, expected cost
- Direct anchor for **Scaffold A** (Elliptic Bitcoin, around 10 percent illicit)

Factor investing and the replication crisis (Week 9)

- Fama-French is a starting point, not an end state
- Harvey, Liu and Zhu: the **factor zoo**; the right t-hurdle is **above 3** (Harvey, Liu, and Zhu 2016)
- Jensen, Kelly and Pedersen: a careful global re-test (Jensen, Kelly, and Pedersen 2024)
- Direct anchor for **Scaffold B** (JKP UK monthly factors with tree models and SHAP)

Backtesting, validation, sequence learning (Weeks 10 and 11)

- Walk-forward CV with embargoes; in-sample is **not** evidence
- **PBO** from CSCV (Bailey et al. 2015); **DSR** (Bailey and Prado 2014): the gap measures the cost of search
- Sequence learning showed why **temporal structure** matters in finance
- For **Scaffold C** (volatility), evaluate forecasts with Mincer-Zarnowitz, not Sharpe

Part III: Ethics, regulation, judgement

Algorithmic accountability in lending

- Digital footprint scoring (Berg et al. 2020) is **opaque to the borrower** by construction
- Mortgage technology can be **predictably unequal** across protected groups (Fuster et al. 2022)
- The accountability question: when a model denies credit, **who is on the hook**?
  - Data scientist? Product owner? Vendor? Risk function? Regulator?
- "Compliance with the rules" and "treats customers fairly" are not the same test

Transparency vs interpretability

- **SHAP** (used in Scaffold B) gives feature attributions; it does **not** give causal explanations
- Quoting model output is not the same as **explaining a decision**
- The CW2 report tests this skill: tell the marker what your model **cannot** justify
- Fairness has multiple, mathematically incompatible definitions; the choice is normative

Regulating AI in finance: the UK lens

- **EU AI Act** (Regulation 2024/1689): credit scoring and market surveillance fall under **high-risk** obligations
- **UK FCA**: principles-based; supervisory sandboxes; consumer duty
- **Bank of England SS1/23**: model risk management principles for PRA-regulated firms
- "AI vs no AI" is the wrong question; **"governable AI vs ungoverned AI"** is the real one

The scarce skill: judgement under uncertainty

- LLMs can **produce output**; they cannot be **accountable** for it
- FIN306 trains the rarer skill: deciding what matters, weighing evidence, communicating limits
- That is also what employers in FCA-regulated firms hire for
- It is the **same** habit your CW1 risk register and CW2 limitations section reward

Part IV: FIN306 Coursework 2 and careers

FIN306 schedule: this week and next

- **Week 12 (today):** Module synthesis; **CW2 finalisation** (see module schedule)
- **Week 13: Coursework 2 due**. Confirm **date, time, and portal** on **Blackboard** (module outline currently lists **8 May 2026, 23:59**; always verify live)
- **What you submit:** one completed **scaffold notebook** plus a **reflective report** (2,500 words, +10% rule on the brief). Full file types and naming on Blackboard

FIN306 Coursework 2: Applied Data Science with Critical Reflection

- **Weight: 70%** of the module (CW1 was the recorded presentation and risk register)
- **You chose one scaffold in Week 8**; you complete the **TODO** sections and write the **report**. We assess **judgement and reflection**, not coding from scratch
- **Scaffold A, Blockchain fraud (Elliptic):** rare-event classification, walk-forward CV, cost-sensitive decisions; ties to fraud week
- **Scaffold B, Tree-based factor investing (JKP UK):** ensembles, SHAP, OOS evaluation; ties to factor week
- **Scaffold C, Volatility forecasting:** GARCH family, forecast evaluation; ties to volatility week and Bloomberg-style data

How FIN306 Coursework 2 is marked

Criterion	Weight	Examiners look for
Content: analysis depth and insight	**25%**	Honest interpretation; limitations; not a marketing pitch
Application of theory	**20%**	Correct use of the method; validation; why this specification
Knowledge and understanding	**20%**	Links to module themes (bias, OOS discipline, FinTech context)
Evidence of reading	**15%**	Academic and professional sources used properly
Referencing	**10%**	Harvard (or as brief); consistent
Communication	**10%**	Structure, figures, professional tone

Week 13 checklist (before you submit)

- ☐ **Notebook:** every **TODO** completed; **Kernel → Restart & Run All** with no errors
- ☐ **Report:** method rationale, **data quality** choices, results, **limitations**, responsible deployment. Use CW1 feedback where relevant
- ☐ **Word count** within brief (remember what counts toward the limit)
- ☐ **Both files** uploaded as Blackboard requires (e.g. `.ipynb` plus `.docx`; confirm)
- ☐ **References** complete; no placeholder citations
- ☐ Any **AI-use** or integrity statement required by the brief is included

Career Pathways in FinTech and Financial Data Science

Roles:

1. Data Science / Quantitative Analysis - Develop trading algorithms, risk models, fraud detection - Requires: Statistics, ML, programming, financial knowledge - Employers: Asset managers, banks, hedge funds, FinTech startups

2. Product Management - Design financial products/services, translate needs to specifications - Requires: Technical understanding, business acumen, communication - Employers: FinTech companies, banks' digital divisions

3. Regulatory / Compliance - Navigate regulation, design compliance systems, engage regulators - Requires: Legal/regulatory knowledge, risk management, technology understanding - Employers: All financial institutions, consultancies, regulators

4. Research / Consulting - Analyse industry trends, advise firms/governments on strategy - Requires: Analytical skills, industry knowledge, communication - Employers: Central banks, think tanks, consultancies, academia

Final Reflection and Moving Forward

What we've covered (FIN306 arc):

- Foundations: statistical science, returns, bias–variance, honest inference
- Data quality and measurement (Bloomberg and structured labs)
- Time series and volatility modelling
- Robo-advisory and alternative finance
- Machine-learning bridge and **CW2 scaffolds** (Weeks 7–8)
- Cryptocurrency, fraud detection, factor investing, backtesting and validation
- Sequence learning (Week 11), then synthesis, ethics, and futures (this week)
- Ethical frameworks evaluating FinTech impacts

What remains uncertain:

- Will decentralisation succeed or will re-intermediation dominate?
- How will AI transform finance and employment?
- Will regulation enable innovation or stifle it?
- How will society navigate privacy vs surveillance?
- Will FinTech reduce inequality or increase it?

References and Further Reading

Statistical foundations of FIN306

- Gelman, Hill and Vehtari (2020). *Regression and Other Stories*. The three challenges of inference.
- Kelly, Malamud and Zhou (2024). The virtue of complexity in return prediction. Frames bias-variance for ML in finance.

Economics of technology enabled financial services

- Philippon (2016). The FinTech opportunity and the 130-year cost puzzle.
- Reher and Sokolinski (2024). Robo-advisors and access to wealth management.

Methods, replication, and honest reporting

- Gu, Kelly and Xiu (2020). Empirical asset pricing via machine learning.
- Harvey, Liu and Zhu (2016). The factor zoo and the $t > 3$ hurdle.
- Jensen, Kelly and Pedersen (2024). The replication crisis re-examined; source dataset for Scaffold B.
- Bailey et al. (2015). Probability of Backtest Overfitting via CSCV.
- López de Prado (2014). The Deflated Sharpe Ratio.