Module 2 Summary: The Access Problem

Prof. Dr. Joerg Osterrieder

Digital Finance — BSc Course

Module 2: Prerequisites

Theme: Access (financial inclusion) **From prior modules:**

- **M1L1:** the five-party payment model and who bears the MDR
- **M1L2:** two-sided market economics of intermediaries
- **M1L4:** unit economics (LTV/CAC) for fintech viability

External knowledge assumed:

- Basic probability (independent events, conditional probability)
- Elementary descriptive statistics (mean, standard deviation, percentile)
- Pandas basics for reading and filtering tabular data

Will be introduced this module: We introduce financial exclusion data, credit-scoring models, robo-advice mechanics, and algorithmic-fairness metrics (demographic parity, equal opportunity, Chouldechova).

Prerequisites are advisory; lessons remain self-contained where feasible. Forward references inside lessons flag any concept used before its canonical introduction.

Key Concepts by Lesson

L1: Financial Exclusion

- 1.4 billion adults lack formal identity documents
- Mobile money (M-Pesa) as leapfrog technology
- KYC requirements as barrier to inclusion

L2: Credit Scoring & Alternative Data

- Traditional scoring excludes "credit invisibles"
- Alternative data: mobile usage, utility payments, social graphs
- Trade-off: inclusion vs. privacy and bias risk

L3: Democratizing Investment

- Fractional shares lower minimum investment to $1
- Robo-advisors: algorithm-driven portfolio management
- Gamification risks: overtrading, herd behavior

L4: Algorithmic Fairness

- Disparate impact vs. disparate treatment
- Fairness metrics: demographic parity, equalized odds
- Impossibility theorem: cannot satisfy all fairness criteria simultaneously

---

Module 2 answers: Who is excluded from finance, why, and how can technology close the gap without creating new harms?

Key Formulas and Definitions

Financial Exclusion

Unbanked: No account at a formal financial institution. **Underbanked:** Has an account but relies on alternative financial services (payday lenders, check cashers).

Credit Scoring

$$\text{Default Probability} = f(\text{payment history, utilization, length, mix, inquiries})$$

Alternative data extends this: $f(\ldots, \text{mobile top-ups, rent payments, utility bills})$

Fairness Metrics

$$\text{Demographic Parity:} \quad P(\hat{Y} = 1 \mid A = 0) = P(\hat{Y} = 1 \mid A = 1)$$

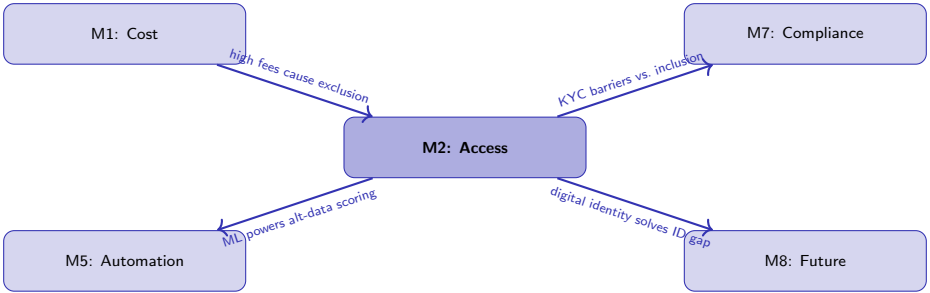
$$\text{Equalized Odds:} \quad P(\hat{Y} = 1 \mid Y = y, A = 0) = P(\hat{Y} = 1 \mid Y = y, A = 1)\ \forall y$$

Impossibility Theorem (Chouldechova, 2017)

When base rates differ across groups ($P(Y = 1|A = 0) \neq P(Y = 1|A = 1)$), it is **mathematically impossible** to simultaneously achieve equal false positive rates, equal false negative rates, and equal predictive values.

Fairness in lending is a mathematical constraint problem, not just an ethical aspiration.

Connections to Other Modules



- **Cost → Access (M1):** High MDR and account fees are direct drivers of financial exclusion
- **Access → Automation (M5):** ML models enable alternative credit scoring but introduce bias-variance trade-offs
- **Access → Compliance (M7):** KYC/AML requirements designed to prevent crime also exclude legitimate users without documents
- **Access → Future (M8):** Self-sovereign identity (SSI) could give the unbanked verifiable credentials without centralized gatekeepers

Access is the human consequence of cost, technology, and regulation decisions made across every other module.

Module 2: Integrative Questions

Two questions that need more than one lesson to answer:

1. A neobank claims its model serves the unbanked; using M2L2 (credit scoring) + M2L4 (algorithmic fairness), what evidence would falsify the claim?
2. Combine M2L1 (financial exclusion data) with M2L3 (democratising investment) to design an inclusion metric that is not gameable by liquidity-mining incentives.

Use these as study prompts before the module exam; each integrates concepts that span lessons.

Module 2: Worked Multi-Lesson Example

Worked example traces a single low-income borrower across M2L1 (currently unbanked), M2L2 (thin-file credit score), M2L3 (fractional ETF access through a neobank), and M2L4 (fairness-audit on the credit model's allocation).
Pedagogical pattern: the example is intentionally end-to-end. Solve it lesson-by-lesson, then step back and identify the lesson whose assumption was the binding constraint.

The exam-style version of this example appears in extttv4/exam_prep/exam_bank.tex for module 2.

Module 2: Forward-Pointing Dependency Map

Concepts from Module 2 that later modules will use:

- **M5L3:** ML fairness limits flow back into M2L4 as the operating constraint for credit-scoring models
- **M7L3:** EU AI Act Annex III lists credit scoring as high-risk; M7L3 governs the M2L2 model in practice
- **M8L1:** Digital-identity primitives from M8L1 are the prerequisite for portable thin-file credit signals

Forward-pointing dependencies; concepts not in this map are local to Module 2.