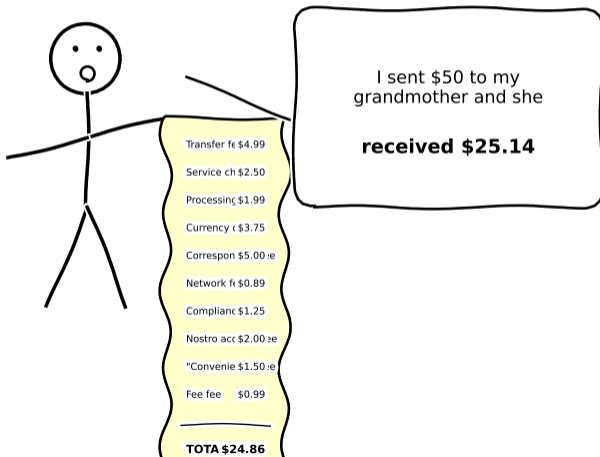


Lesson 1.2: The Economics of Financial Intermediation

Module 1: The Cost Problem

Prof. Dr. Joerg Osterrieder

International Money Transfer: The Experience



After this lesson you will be able to:

- 1 **Apply** transaction cost economics to explain why financial intermediaries exist
- 2 **Decompose** interchange fees into their component parts
- 3 **Trace** a cross-border payment through the correspondent banking system
- 4 **Calculate** the total cost of a remittance including Foreign Exchange (FX) spreads
- 5 **Evaluate** the impact of interchange regulation on different stakeholders
- 6 **Compare** two-sided market dynamics in payment networks

Bloom's taxonomy levels: **Apply** (1–4), **Analyze** (5), **Evaluate** (6).

Lesson 1.1 – Where costs arise:

- Identified the actors (issuer, acquirer, network)
- Mapped the four-party payment flow
- Catalogued visible and hidden fee line items

Lesson 1.2 – Why costs exist:

- Transaction cost economics (Coase, Williamson)
- Interchange as a two-sided market equilibrium
- Correspondent banking and FX cost structure
- Regulatory interventions and their effects

*Now that we know **where** costs arise, we ask **why** they exist.*

This lesson builds the theoretical foundation for understanding payment costs before we examine FinTech disruption in Lessons 1.3–1.4.

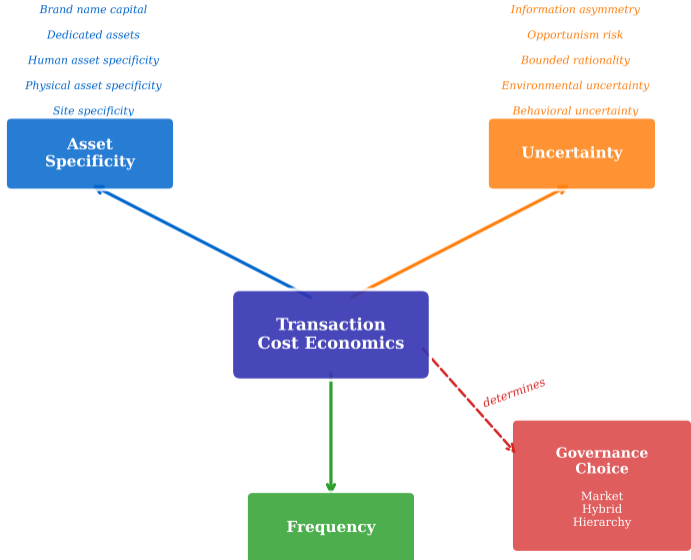
What Is Transaction Cost Economics (TCE)?

Definition: Transaction Cost Economics (TCE) is a framework that explains why economic actors organize transactions through firms and intermediaries rather than pure market exchange.

- **Ronald Coase (1937):** Firms exist because market transactions are costly — there are costs to using the price mechanism (search, negotiation, enforcement)
- **Oliver Williamson (1975, 1985):** Extended Coase's insight by identifying three key dimensions that determine governance choice:
 - *Asset specificity* — how specialized are the investments?
 - *Uncertainty* — how unpredictable is the environment?
 - *Frequency* — how often does the transaction recur?
- When all three are high, hierarchical governance (intermediaries) dominates markets

Coase won the Nobel Prize in Economics in 1991; Williamson in 2009. Both for work on transaction cost theory.

Williamson's Transaction Cost Economics Framework



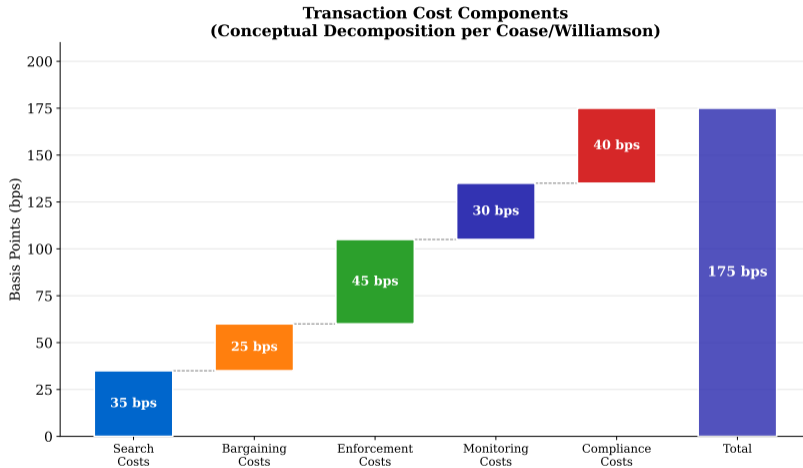
Why do banks, card networks, and clearinghouses exist?

Dimension	Payments	Lending	Securities
Asset specificity	Network infrastructure, fraud systems	Credit scoring models, customer data	Trading platforms, custody systems
Uncertainty	Counterparty fraud, settlement risk	Default risk, interest rate risk	Price volatility, liquidity risk
Frequency	Millions of transactions per day	Periodic (monthly payments)	Continuous (real-time trading)

All three dimensions are **high** in finance → intermediaries dominate.

TCE predicts that disintermediation only succeeds when technology reduces one or more of these dimensions.

Transaction Cost Components



Each component represents a real cost that intermediaries absorb. Reducing any component creates a FinTech opportunity.

What Is Interchange?

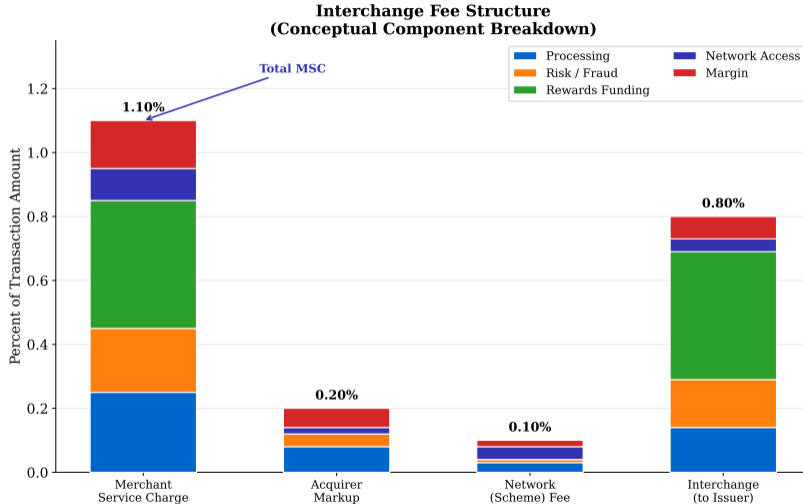
Definition: An **interchange fee** is the fee paid by the merchant's bank (acquirer) to the cardholder's bank (issuer) each time a card payment is processed.

- Set by the card network (e.g., Visa, Mastercard), not by individual banks
- Typically expressed as a percentage of transaction value plus a fixed amount
- Represents the **largest single component** of the Merchant Service Charge (MSC)
- Funds issuer costs: fraud prevention, interest-free period, rewards programs
- **Key economic insight:** interchange is a *balancing mechanism* in a two-sided market

Flow: Merchant $\xrightarrow{\text{MSC}}$ Acquirer $\xrightarrow{\text{Interchange}}$ Issuer

Interchange is invisible to consumers but ultimately affects prices. It is the most debated fee in payment economics.

Interchange Fee Structure



The issuer receives the largest share because it bears the most risk (fraud, credit losses, float).

What Is a Two-Sided Market?

Definition: A **two-sided market** (or platform market) is one where an intermediary serves two distinct user groups, and the value to each group depends on participation by the other.

Card network example:

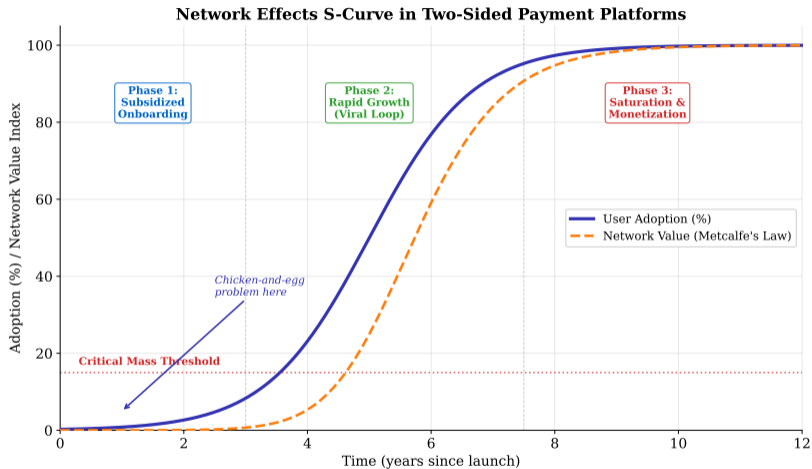
- **Side 1:** Cardholders (consumers)
- **Side 2:** Merchants (acceptors)
- More cardholders → merchants want to accept
- More merchants → consumers want cards

Pricing implication:

- Platform subsidizes the price-sensitive side
- Interchange shifts costs to merchants (less price-elastic)
- Consumers receive rewards funded by merchant fees
- Rochet & Tirole (2003): optimal interchange balances both sides

Jean Tirole won the 2014 Nobel Prize partly for his work on two-sided markets and platform economics.

Network Effects in Payment Platforms



The “chicken-and-egg” problem is the central challenge for new payment networks: both sides must be onboarded simultaneously.

What Is Correspondent Banking?

Definition: Correspondent banking is an arrangement where one bank (the correspondent) holds deposits and provides services on behalf of another bank (the respondent), enabling cross-border payments.

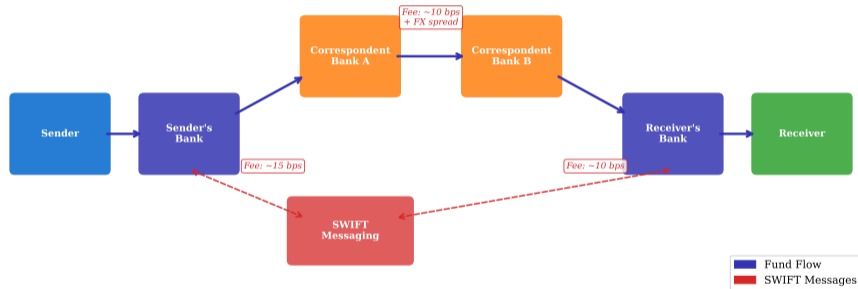
- Most banks do not have direct relationships with banks in every country
- Instead, they maintain accounts at large global banks that act as intermediaries
- A payment may pass through 1–4 correspondent banks in a chain
- Each link adds fees, time, and potential points of failure
- The Society for Worldwide Interbank Financial Telecommunication (SWIFT) provides the messaging layer but does **not** move funds

Key problem: The more links in the chain, the higher the cost and the longer the settlement time.

Correspondent banking relationships have been declining since 2011 (“de-risking”), making some corridors more expensive.

Cross-Border Payment Flow (Correspondent Banking Model)

Fund Flow (solid) vs. Message Flow (dashed)



A typical cross-border payment takes 2–5 business days. Each intermediary adds latency and cost.

What Are Nostro and Vostro Accounts?

Definitions:

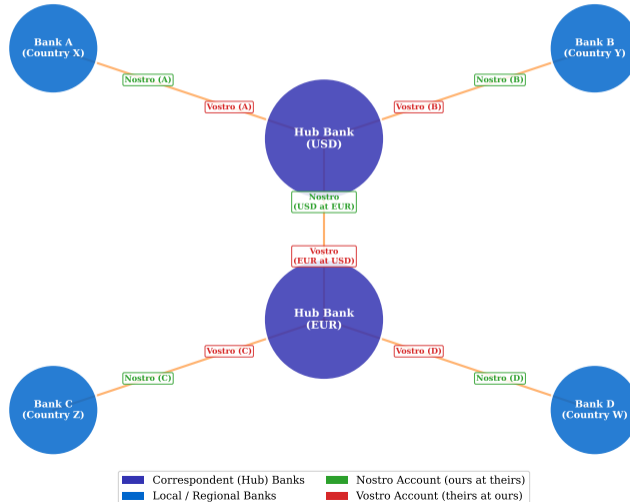
- **Nostro account** (“ours at theirs”): An account that Bank A holds *at* Bank B, denominated in Bank B’s local currency
- **Vostro account** (“theirs at ours”): The same account from Bank B’s perspective — Bank A’s money held *at* Bank B
- They are the **same account** seen from two different viewpoints

Why they matter for costs:

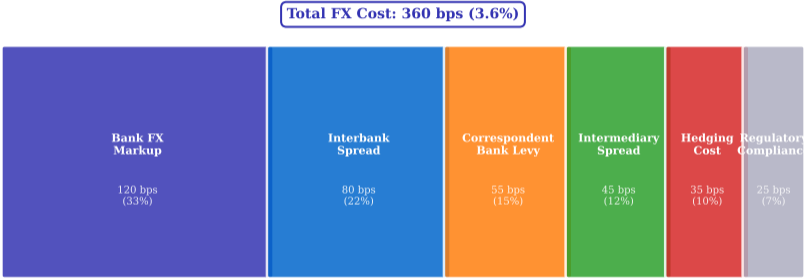
- ① Banks must pre-fund nostro accounts → **trapped liquidity**
- ② Maintaining accounts in many currencies → **capital costs**
- ③ Reconciliation between nostro and vostro → **operational overhead**
- ④ Failed reconciliation → **investigation costs**

Global nostro account balances are estimated at over \$27 trillion in pre-funded liquidity (McKinsey, conceptual estimate).

Correspondent Banking Network (Nostro / Vostro Account Relationships)



FX Spread Decomposition (Conceptual Cost Breakdown for Cross-Border Payment)



The mid-market rate is the “true” exchange rate. Every deviation from it is a cost to the sender or receiver.

Calculating Total Remittance Cost

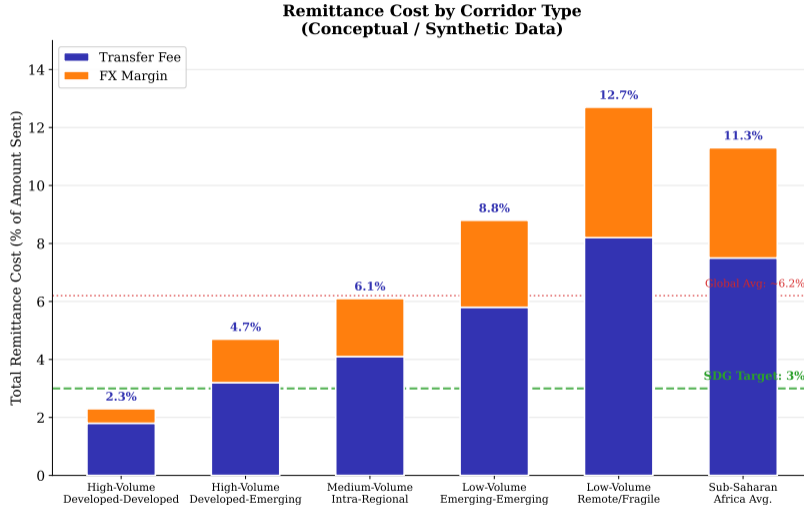
Example: Sending \$200 from Country A to Country B.

Cost Component	Amount	% of \$200
Transfer fee (flat)	\$4.99	2.50%
FX markup (1.5% of \$200)	\$3.00	1.50%
Correspondent bank fee	\$3.50	1.75%
Receiving bank fee	\$1.50	0.75%
Total cost	\$12.99	6.50%
Amount received	\$187.01	93.50%

- The United Nations Sustainable Development Goal (SDG) 10.c target: reduce remittance costs to below 3% by 2030
- Current global average: approximately 6.2% (conceptual benchmark)
- Some corridors exceed 10–15% due to thin volume and regulatory friction

Smaller remittances are hit harder: the same \$4.99 flat fee is 10% of a \$50 transfer but only 0.5% of a \$1,000 transfer.

Remittance Cost by Corridor Type



Corridor cost depends on volume, competition, regulation, and availability of digital channels.

Why Remittance Costs Vary by Corridor

Low-cost corridors (below 3%):

- High volume (economies of scale)
- Multiple competing providers
- Digital channels dominant
- Stable, convertible currencies
- Light regulatory burden

High-cost corridors (above 8%):

- Low volume (no scale)
- Few providers (monopoly rents)
- Cash-only last mile
- Exotic or illiquid currencies
- Heavy compliance requirements (Anti-Money Laundering / Know Your Customer)

Key insight: The people who can least afford fees — low-income migrant workers — pay the highest rates.

De-risking by correspondent banks has made some corridors even more expensive by reducing competition.

EU Interchange Fee Regulation (IFR, 2015):

- Caps on **all** consumer card interchange
- Debit: 0.2% of transaction value
- Credit: 0.3% of transaction value
- Applies across the European Economic Area (EEA)
- Commercial cards exempt

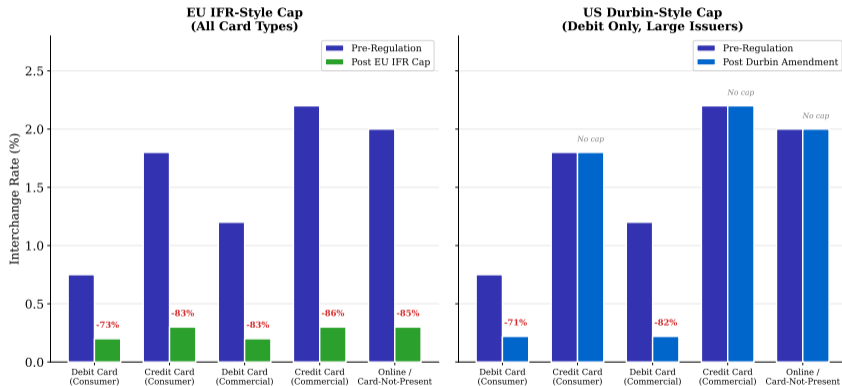
US Durbin Amendment (2010):

- Caps on **debit card** interchange only
- Applies to banks with assets > \$10 billion
- Cap: \$0.21 + 0.05% per transaction
- Credit cards unregulated
- Small banks exempt

Core debate: Do interchange caps help consumers (lower prices) or hurt them (reduced rewards, higher bank fees)?

Australia was the first country to regulate interchange (2003). The Reserve Bank of Australia pioneered cost-based interchange setting.

Impact of Interchange Fee Regulation (Conceptual)



EU caps reduced interchange by 60–80%. The US Durbin Amendment reduced debit interchange by approximately 70% for large issuers.

Who Wins and Who Loses from Interchange Caps?

Stakeholder	Potential Benefit	Potential Cost
Merchants	Lower MSC → savings	May not pass savings to consumers
Consumers	Lower retail prices (in theory)	Reduced card rewards; higher bank fees
Issuers	Clearer cost structure	Revenue loss → offset via other fees
Networks	Regulatory certainty	Volume pressure if cards become less attractive
FinTechs	Level playing field	Less room for disruption if fees already low

Empirical finding: Merchant savings are often not fully passed through to consumer prices (“merchant passthrough puzzle”).

The distributional effects of interchange regulation remain actively debated among economists and regulators.

The Waterbed Effect in Interchange Regulation

Definition: The **waterbed effect** occurs when capping one fee causes other fees to increase to compensate for lost revenue.

- Interchange caps reduce issuer revenue from merchants
- Issuers respond by:
 - Introducing or increasing annual card fees
 - Reducing or eliminating rewards programs
 - Increasing fees on bank accounts (overdraft, maintenance)
 - Tightening credit availability
- **Net effect on total consumer welfare is ambiguous**
- Regulators must consider the *entire* fee ecosystem, not just interchange in isolation

Analogy: Pushing down on one side of a waterbed simply pushes the other side up.

The waterbed effect is a key argument against simplistic fee caps. Comprehensive regulation must anticipate offsetting fee changes.

Why are cross-border payments 5–10× more expensive than domestic?

- 1 **FX conversion** — the mid-market rate is marked up at every step
- 2 **Correspondent chain** — each bank in the chain takes a fee
- 3 **Compliance screening** — Anti-Money Laundering (AML) / Countering the Financing of Terrorism (CFT) checks at each node
- 4 **Pre-funded liquidity** — nostro balances tie up capital across currencies
- 5 **Time zone mismatch** — settlement windows do not overlap → delays
- 6 **Lack of transparency** — sender often does not know total cost upfront

Total cost estimate for a \$1,000 cross-border B2B payment:

Domestic equivalent: \$0.50–2.00 vs. Cross-border: \$25–50

The G20 Roadmap for Enhancing Cross-Border Payments (2020) identified these as the five key “friction points” to address.

Cost Comparison: Domestic vs. Cross-Border

Cost Driver	Domestic	Cross-Border	Multiplier
Processing fee	\$0.20	\$2.00	10×
FX conversion	—	\$8.00	n/a
Compliance / AML screening	\$0.10	\$5.00	50×
Correspondent fees	—	\$10.00	n/a
Liquidity / funding cost	\$0.05	\$3.00	60×
Failed payment investigation	\$0.02	\$2.00	100×
Total (est.)	\$0.37	\$30.00	~80×

Conceptual estimates for a \$1,000 business-to-business payment.

These cost multipliers explain why cross-border payments are the primary target for FinTech and blockchain disruption.

Reducing transaction costs:

- **Search costs:** Aggregators, comparison platforms
- **Bargaining costs:** Algorithmic pricing, real-time quotes
- **Monitoring costs:** Automated reconciliation, Application Programming Interfaces (APIs)
- **Enforcement costs:** Smart contracts, programmable escrow

Disrupting the cost stack:

- **FX:** Mid-market rate access, peer-to-peer matching
- **Correspondent banking:** Direct connections, blockchain settlement
- **Compliance:** Shared Know Your Customer (KYC) utilities, regtech
- **Liquidity:** Pre-positioned digital assets, on-demand liquidity

Key insight: FinTech does not eliminate intermediation — it *restructures* it with lower transaction costs.

We will analyze specific FinTech business models attacking each cost layer in Lessons 1.3 and 1.4.

- 1 **Transaction cost economics** (Coase, Williamson) explains *why* financial intermediaries exist: search, bargaining, enforcement, and monitoring costs
- 2 **Interchange fees** are a two-sided market balancing mechanism, not a simple monopoly rent
- 3 **Correspondent banking** creates long chains of intermediaries, each adding cost and latency to cross-border payments
- 4 **Nostro/vostro accounts** trap trillions in pre-funded liquidity across currencies
- 5 **FX spreads** contain multiple hidden components beyond the visible exchange rate
- 6 **Interchange regulation** reduces visible fees but may trigger waterbed effects elsewhere
- 7 Cross-border payments cost **50–100× more** than domestic equivalents
- 8 FinTech disruption targets specific **transaction cost components**, not intermediation itself

Understanding why costs exist is essential before evaluating whether a FinTech solution actually reduces them.

This lesson:

- Built the theoretical foundation: TCE explains intermediary existence
- Analyzed interchange as a two-sided market equilibrium
- Traced cross-border payment costs through the correspondent banking chain
- Evaluated regulatory approaches (EU IFR, US Durbin) and the waterbed effect

Next lesson — Lesson 1.3: Revenue Models Under Pressure:

- How do traditional financial institutions *make money* from payments?
- Net Interest Margin (NIM) compression and fee income trends
- Where FinTechs capture revenue that incumbents lose
- The “unbundling” thesis: picking apart the bank revenue stack

With the cost structure clear, we next examine the revenue side — and where the margin opportunity lies for disruptors.