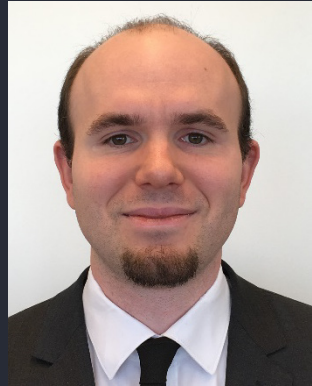# Session 1 Data driven transformation

Jon Frost

Senior Economist, Innovation
and the Digital Economy unit

Bank of International
Settlements

Public policy toward data-driven finance

Jon Frost*, Senior Economist, Innovation and the Digital Economy, BIS

Bank of Finland Conference on Data-driven financial stability, Helsinki, 2 December 2019

*The views expressed here are mine and not necessarily those of the Bank for International Settlements
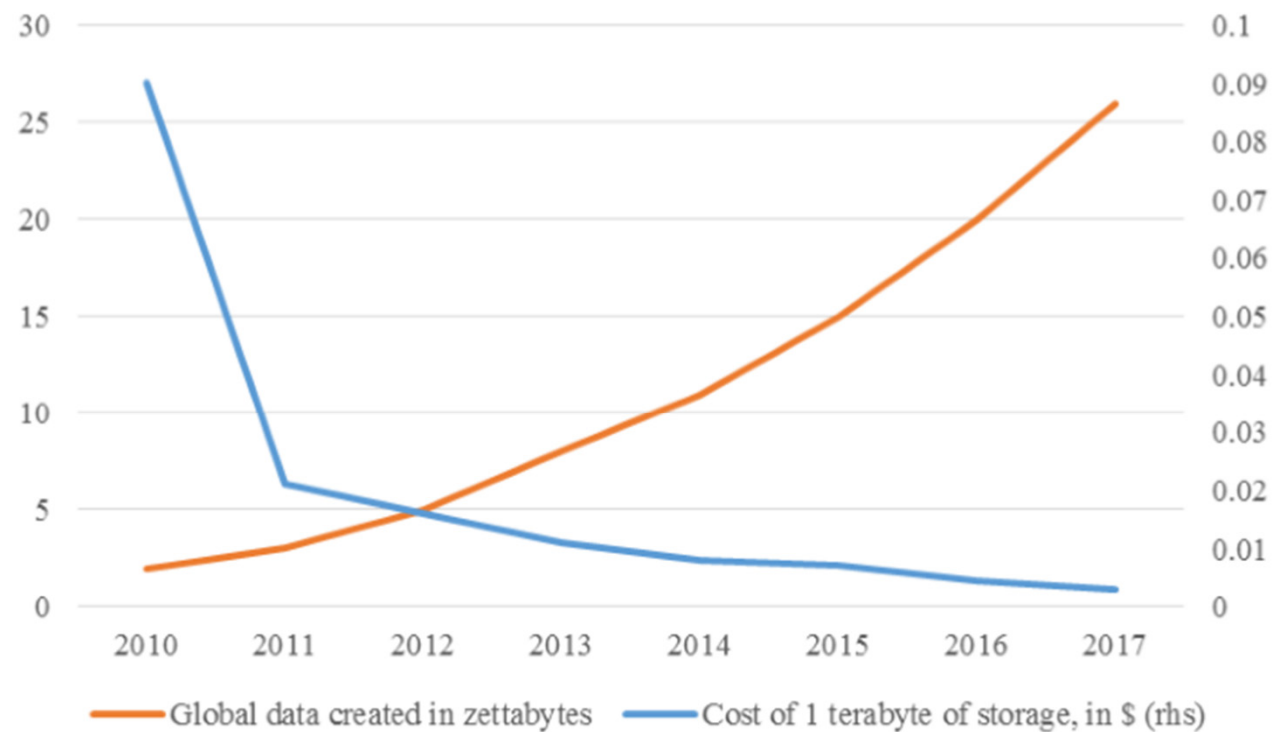
Data in finance are not new

But today, finance has a lot more data

Costs of storage and global data availability, 2009-2017

Source: Reinsel, Gantz and Rydning (2017); Klein (2017). One zettabyte is equal to one billion terabytes.
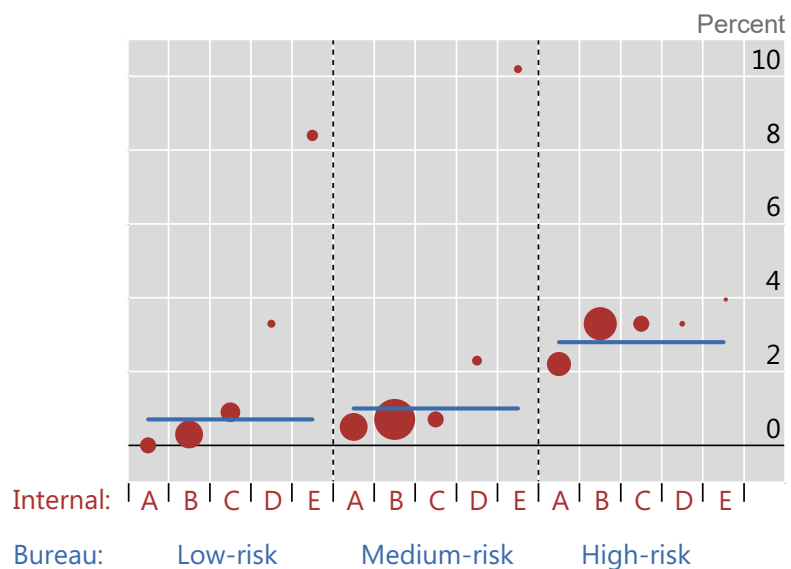
Your digital footprint is valuable

Variable	Description	Information content
Device Type	Main examples: Desktop, Tablet, Mobile.	**Income**
Operating System	Main examples: Windows, iOS, Android.	e.g. Bertrand and Kamenica (2018): iOS best predictor for
Email Provider	Main examples: Gmail, Yahoo, T-Online.	being in Top-Quartile by income
Channel	Channel through which customer has arrived at homepage of the firm. Main examples: paid click vs organic search; affiliate such as price comparison site; direct entering of URL	**Character** e.g. Rook (1987) and Wells et al. (2011): personality traits and impulse shopping
Check-Out Time	Time of day of purchase (morning, afternoon, evening, night)	
Do not track setting	Customer does not allow tracking of device and operating system information, and channel.	
Email Error	Email address contains an error in the first trial (Note: Clients can only order if they register with a correct email address).	
Name in Email	First or last name of customer is part of email address.	**Reputation**
Number in Email	Email address contains number.	e.g. Belenzon, Chatterji, and Daley (2017) and Stern and
Is Lower Case	First name, last name, street, or city are written in lower case.	Guzman (2016): Eponymous Entrepreneurs Effect

Source: Berg et al. (2018)

Alternative data can support credit assessments

Credit Assessment for SMEs and Big Data Analytics

Loss rate[1]

Receiver operating characteristics (ROC) curve[2]

Internal: A B C D E | A B C D E | A B C D E

Bureau: Low-risk   Medium-risk   High-risk

Model:
- I - Logistic, only bureau score
- II - Logistic, bureau score and borrower characteristics
- III - Machine learning, only Mercado Libre credit score
- Random
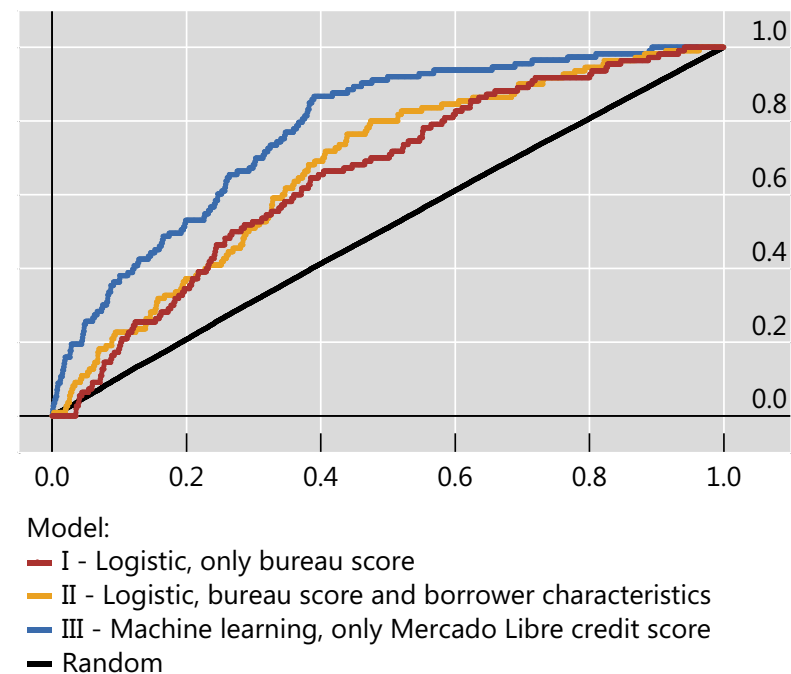
Source: Frost et al. (2019)

Fintech credit has grown fast, from a small base

Global Volume of New Fintech and Big Tech Credit

Credit (lhs): █ Big tech █ Other fintech   Rhs: ● Total fintech credit as a share of total stock of credit[1,2]
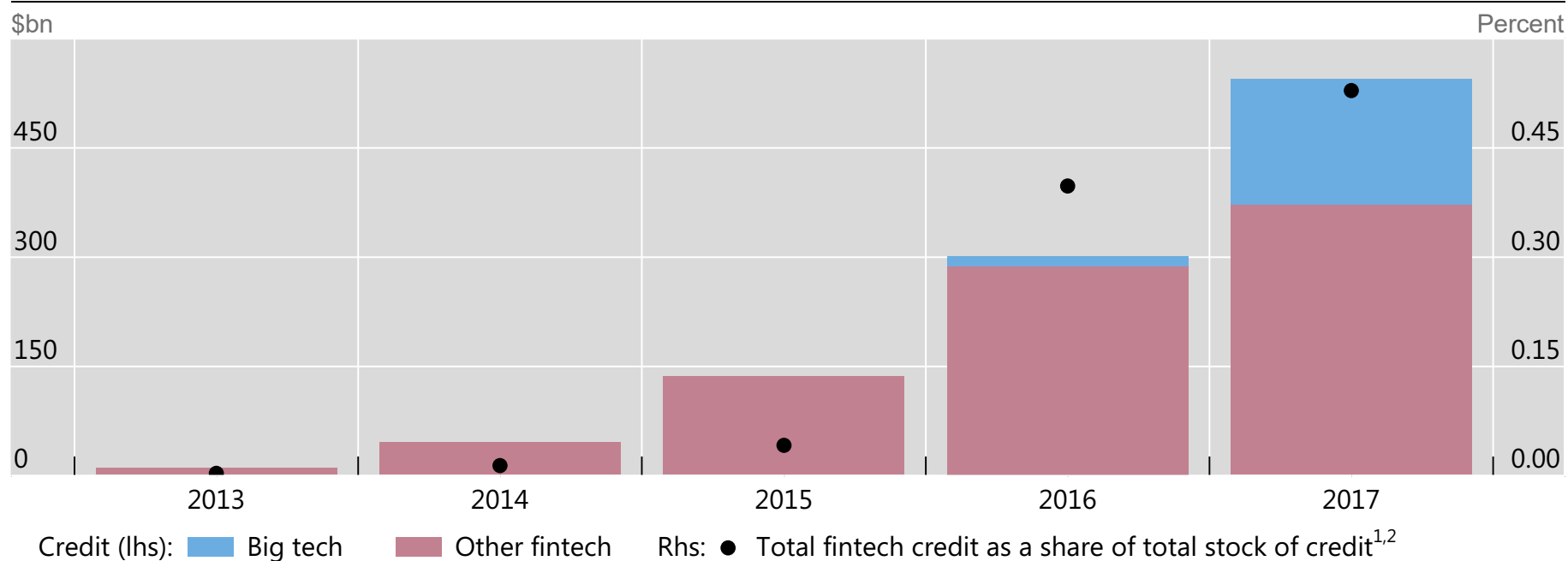
Figures includes estimates.

Source: Frost et al. (2019)

The DNA of big techs

- Big techs' business model rests on enabling direct interactions among a large number of users (e-commerce platforms, social media, search)

- An essential by-product of their business is the large stock of user data which are utilised as input to offer a range of services that exploit natural network effects, generating further user activity

- **D**ata analytics, **n**etwork externalities and interwoven **a**ctivities ("**DNA**") constitute the key features of big techs' business models. These three elements reinforce each other.
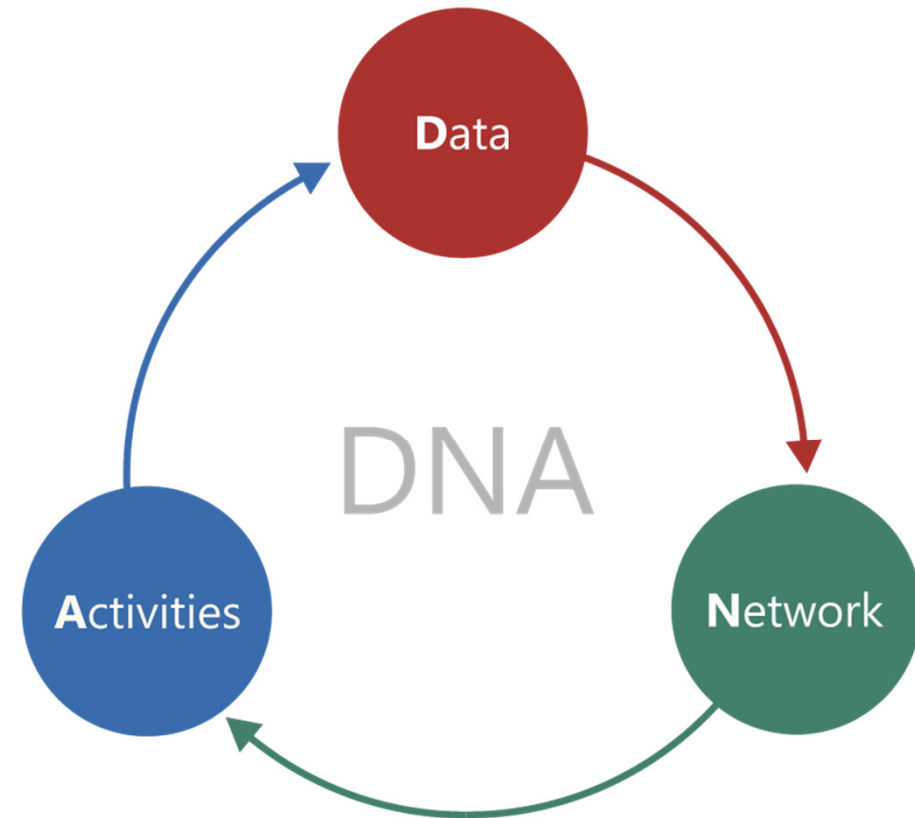
Source: BIS (2019)

Three objectives: stability, efficiency and privacy. A trilogy or another trilemma?
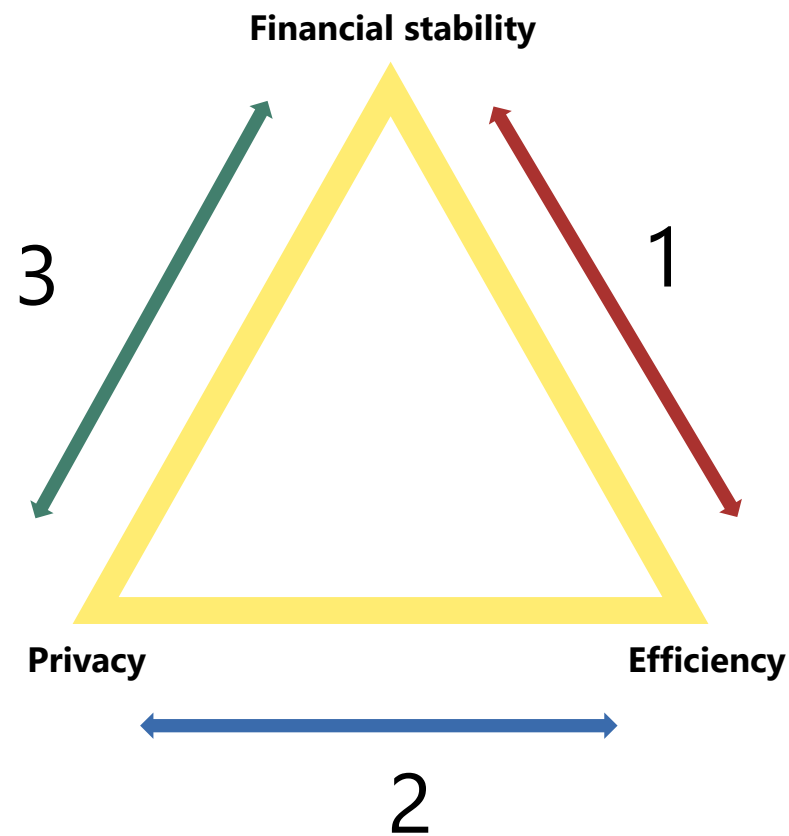
1. "Traditional" stability-competition tradeoff

. Access to data for private providers vs anonymity (eg better/worse access to credit; misuse of data)

3. Access to data for regulatory goals vs anonymity (eg AML/CFT, supervisory data)

Source: Petralia et al. (2019)

Willingness of users to share data for better offers is lower with higher income
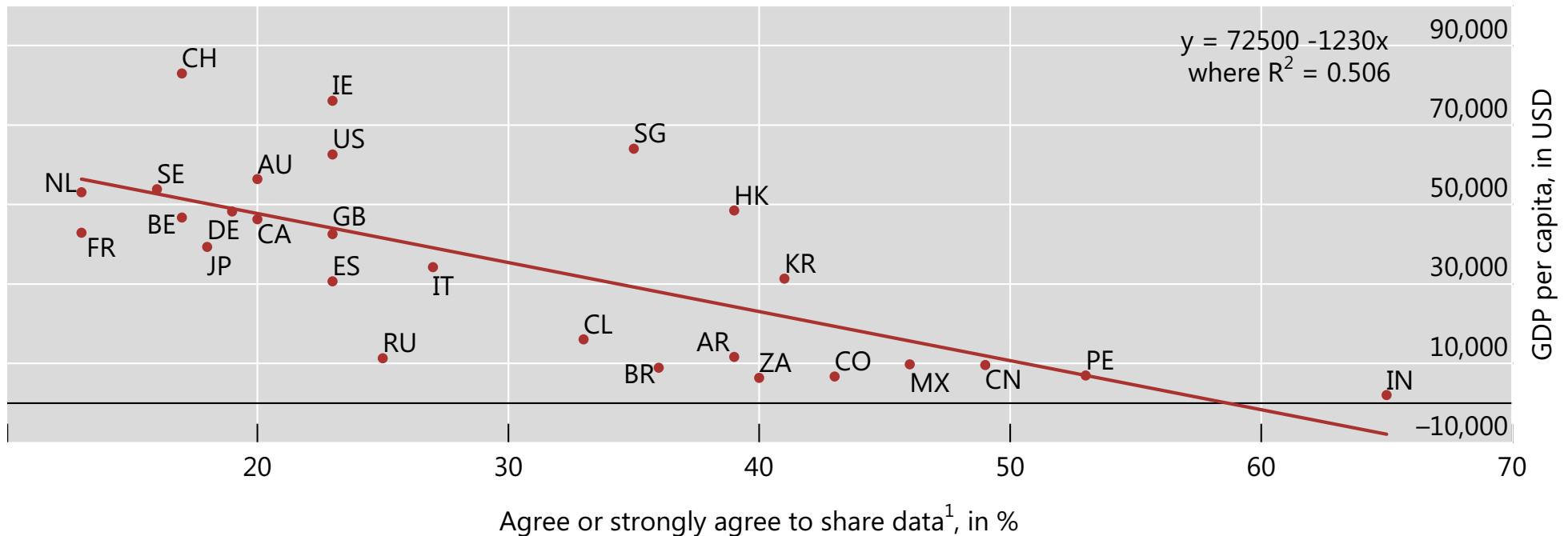
$y = 72500 - 1230x$
where $R^2 = 0.506$

GDP per capita, in USD

Agree or strongly agree to share data[1], in %

[1] The question asked in the survey reads "I would be comfortable with my main bank securely sharing my financial data with other organisations if it meant that I received better offers from other financial intermediaries"; for Belgium the figure is calculated over Belgium and Luxembourg.

Sources: EY (2019), "Global FinTech Adoption Index 2019"; IMF, *World Economic Outlook*.

...and views differ within society, as well

Willingness to share data is higher for men than women[1]

Female:
- ◇ Asia Pacific
- ◇ Americas
- ◆ Africa
- ◇ Europe

Male:
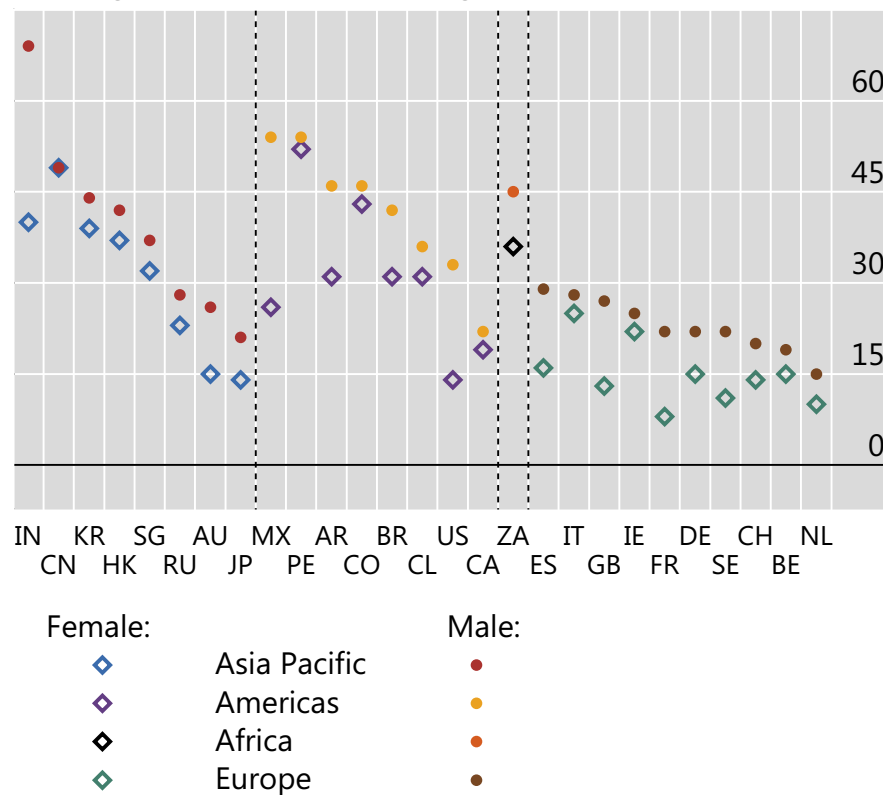- ● Asia Pacific
- ● Americas
- ● Africa
- ● Europe

[1] The question asked in the survey reads "I would be comfortable with my main bank securely sharing my financial data with other organisations if it meant that I received better offers from other financial intermediaries"; for Belgium the figure is calculated over Belgium and Luxembourg.
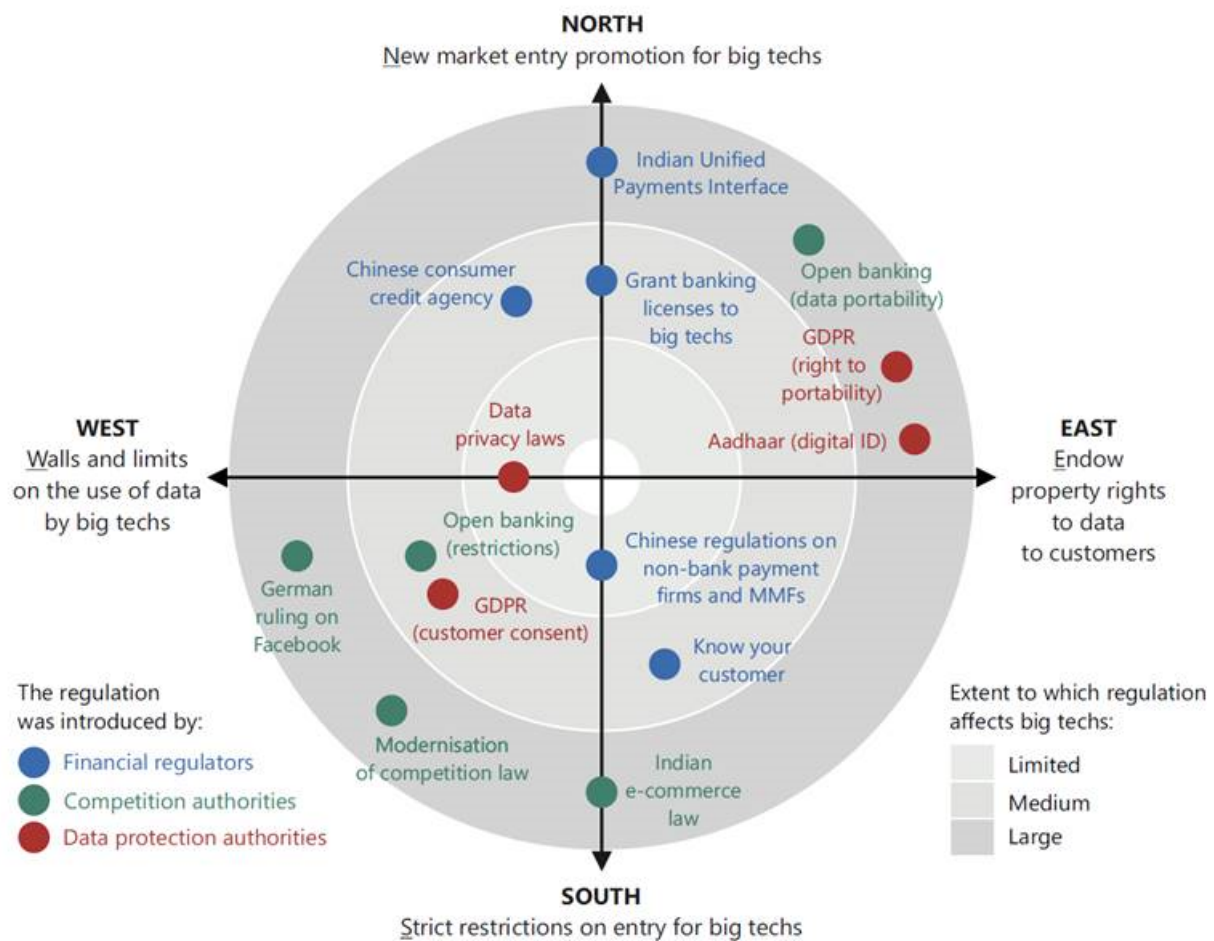Sources: EY (2019), "Global FinTech Adoption Index 2019".

BIS

10

A new regulatory compass

NORTH
New market entry promotion for big techs

Indian Unified Payments Interface

Open banking (data portability)

Chinese consumer credit agency

Grant banking licenses to big techs

GDPR (right to portability)

WEST
Walls and limits on the use of data by big techs

Data privacy laws

Aadhaar (digital ID)

EAST
Endow property rights to data to customers

Open banking (restrictions)

Chinese regulations on non-bank payment firms and MMFs

German ruling on Facebook

GDPR (customer consent)

Know your customer

The regulation was introduced by:
- Financial regulators
- Competition authorities
- Data protection authorities

Modernisation of competition law

Indian e-commerce law

Extent to which regulation affects big techs:
- Limited
- Medium
- Large

SOUTH
Strict restrictions on entry for big techs

Source: BIS (2019)

Digital ID as a foundation?

Conclusion

- Data are valuable, but data are also (market) power
- Policymakers need to consider the complex trade-offs between policy objectives – stability, efficiency and privacy
- Standards complicated by differing preferences to data privacy across and within societies
- Need for domestic collaboration (with competition and data protection authorities) and international dialogue
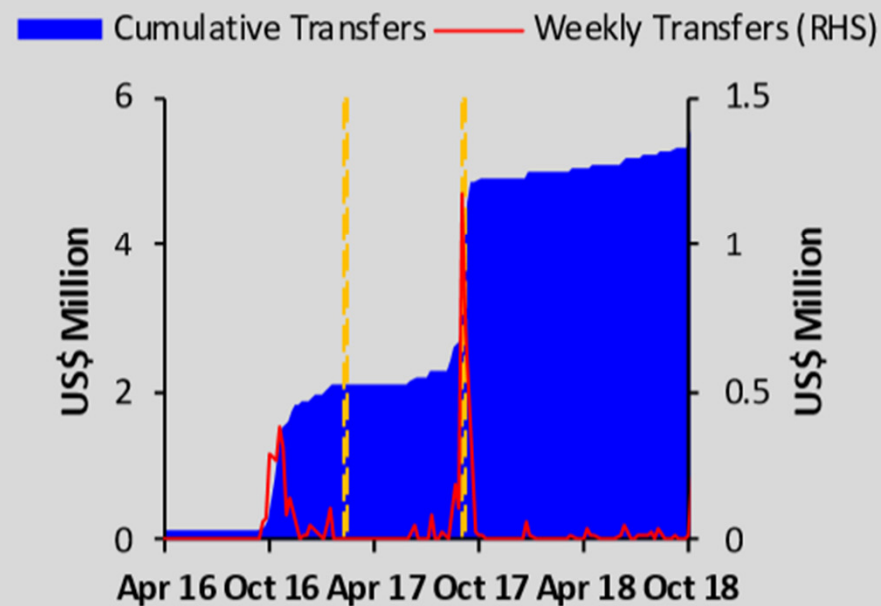- Digital ID as a foundation?

#questions?

BIS

Data science skills can help to monitor opaque markets
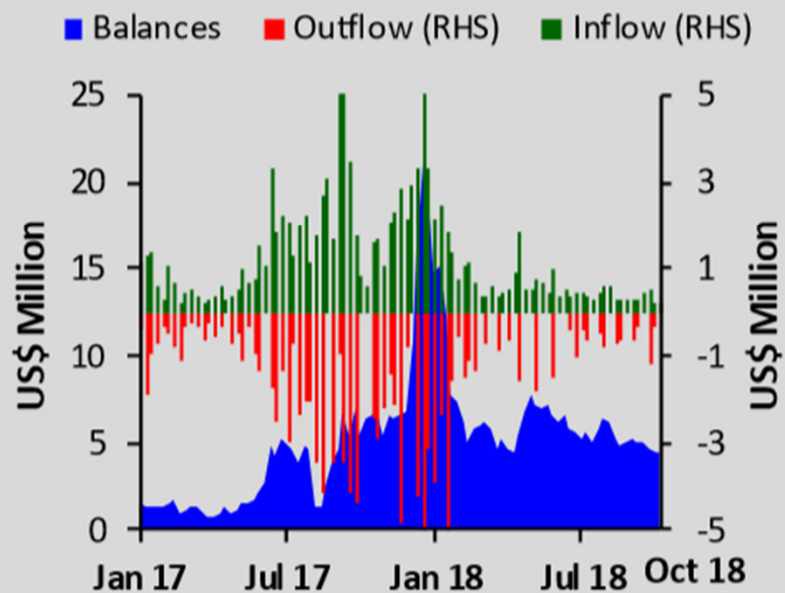
Source: MAS (2018); Bitcoin blockchain

Big techs have a rising share in overall market capitalisation and higher profitability

Share of big techs[1] in overall stock index[2] and profit margins

Lhs: — Big techs' share in world market capitalisation   Profit margin (rhs): — MSCI All Country Global Index   — Big techs[3]
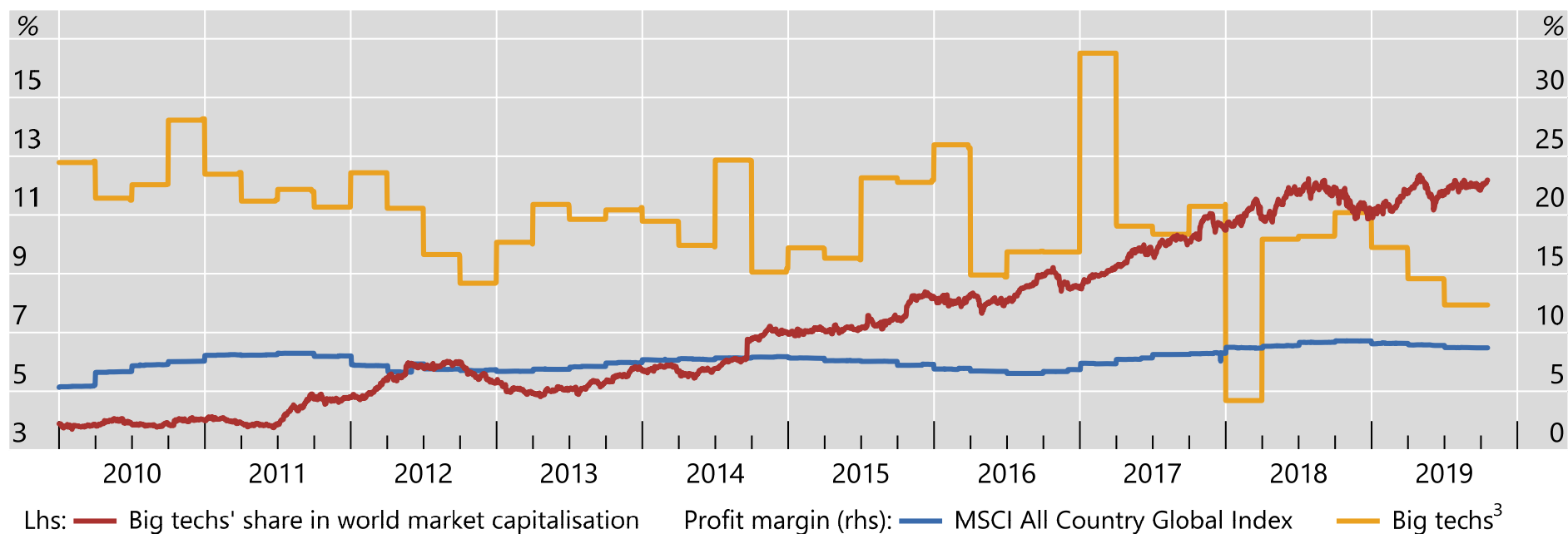
[1] The sample includes Alibaba, Amazon, Apple, Baidu, eBay, Facebook, Google, Kakao Corp, Line, Microsoft, NTT Docomo, Rakuten, Samsung and Tencent.
[2] MSCI All Country Global Index.   [3] Average profit margin.

Source: Bloomberg; Pereira et al. (2019).

The "tech premium" in labour markets is rising

Wage differential between the technology and the financial sector increases

US wages per full-time equivalent employee, in US dollars
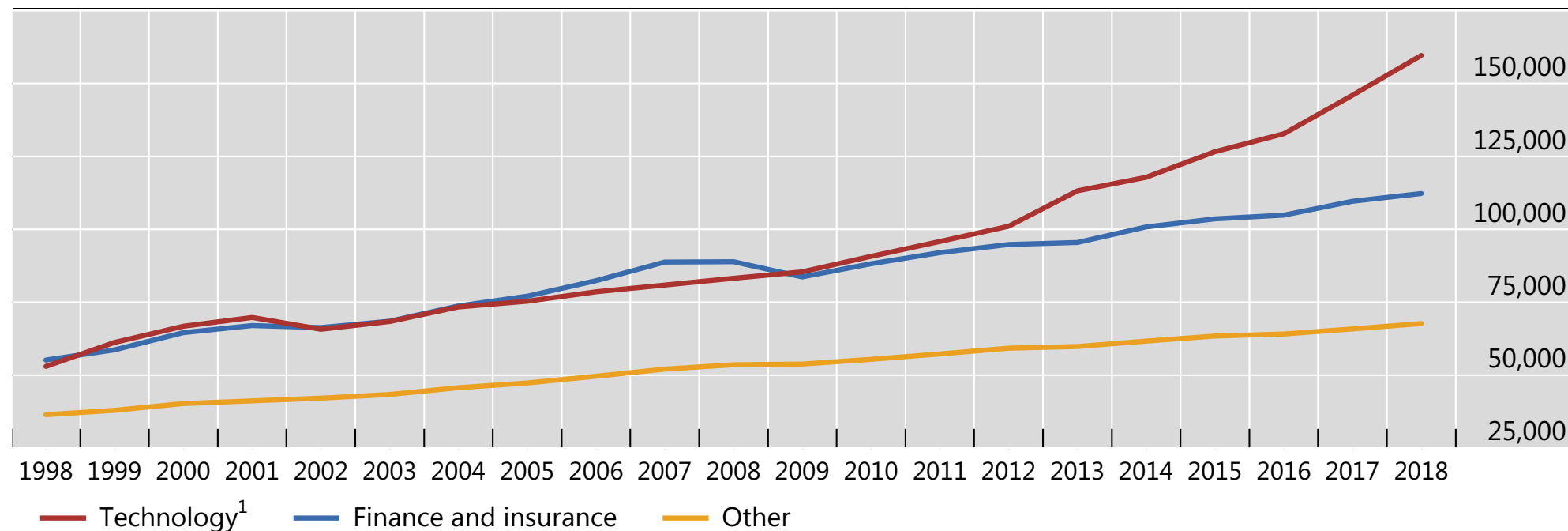
Source: Pereira et al. (2019)

Fintech as a source of SME finance

FinTech credit volumes as a proportion of bank loan volumes to SMEs (2017)

In percent
Graph 9

Legend: US, UK | CA, AU, NL, DE, SG | BR, AR, RU

Note: US total bank loan volume data is based on 2016 figures.

Source: FSB (2019)

Equity crowdfunding remains small, and governance of ICOs is problematic
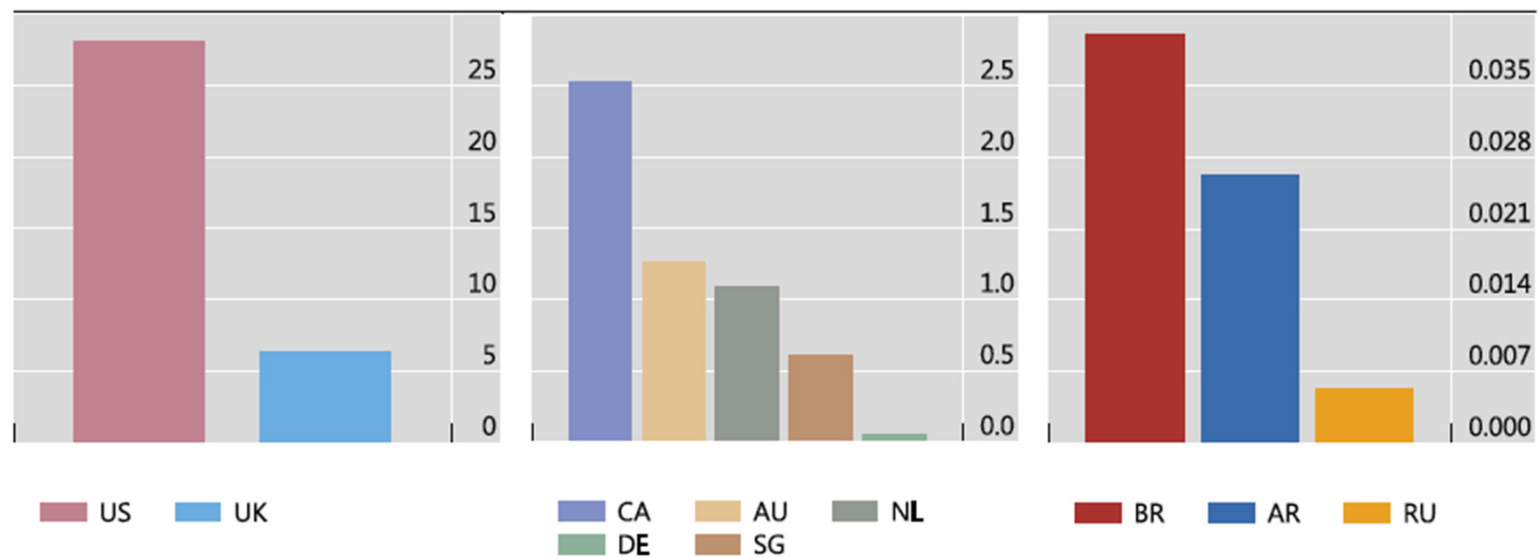
Equity Crowdfunding and ICO Volumes

Equity crowdfunding by jurisdiction[1,2]
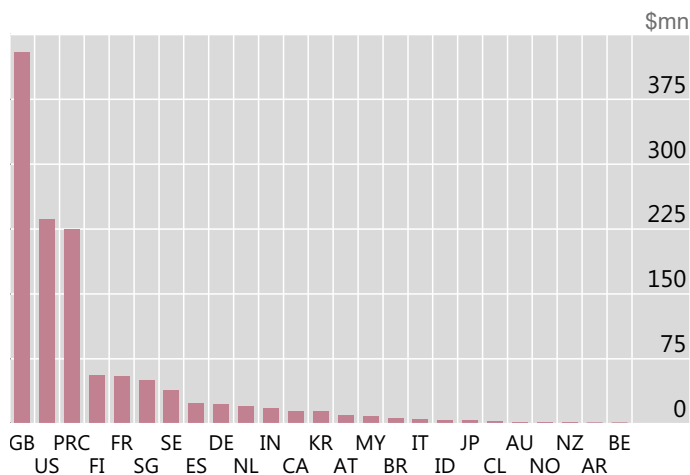
Initial coin offerings
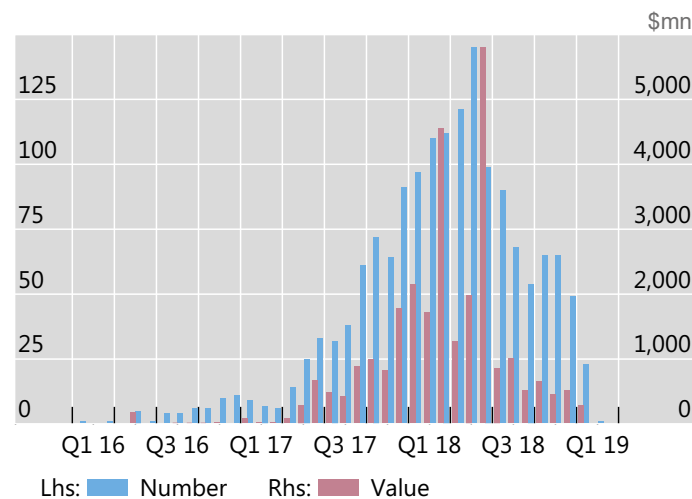
[1] Top 25 countries; 2017 data.

[2] GB = Great Britain, US = United States, PRC = People's Republic of China, FI = Finland, FR = France, SG = Singapore, SE = Sweden, ES = Spain, DE = Germany, NL = Netherlands, IN = India, CA = Canada, KR = Republic of Korea, AT = Austria, MY = Malaysia, BR = Brazil, IT = Italy, ID = Indonesia, JP = Japan, CL = Chile, AU = Australia, NO = Norway, NZ = New Zealand, AR = Argentina, BE = Belgium.

Source: Cornelli et al. (2019)

Keeping things in perspective

Global FinTech activity relative to the global financial system

Total assets of the global financial system, 2017

■ Crypto-assets (at peak, Jan. 2018)
■ FinTech credit (incl. BigTech), 2017

■ =USD 1 trillion

The value of financial assets is represented by the area of the coloured field.

Source: Frost (2019)

Welfare implications of innovations: payments and credit markets
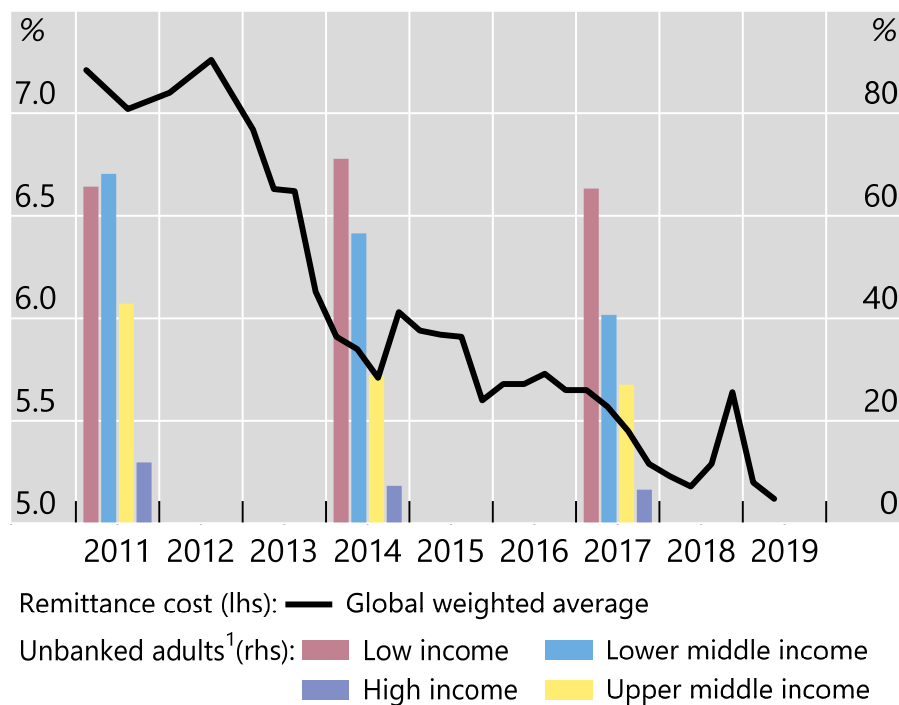
Payment innovations reduce remittances costs

Remittance cost (lhs): ▬ Global weighted average

Unbanked adults[1](rhs): ▮ Low income   ▮ Lower middle income
                     ▮ High income   ▮ Upper middle income

Credit innovations may benefit differently ethnic groups
(cumulative share of borrowers)

▬ Asian   ▬ White Hispanic   - - - 50%
▬ White non-Hispanic   ▬ Black
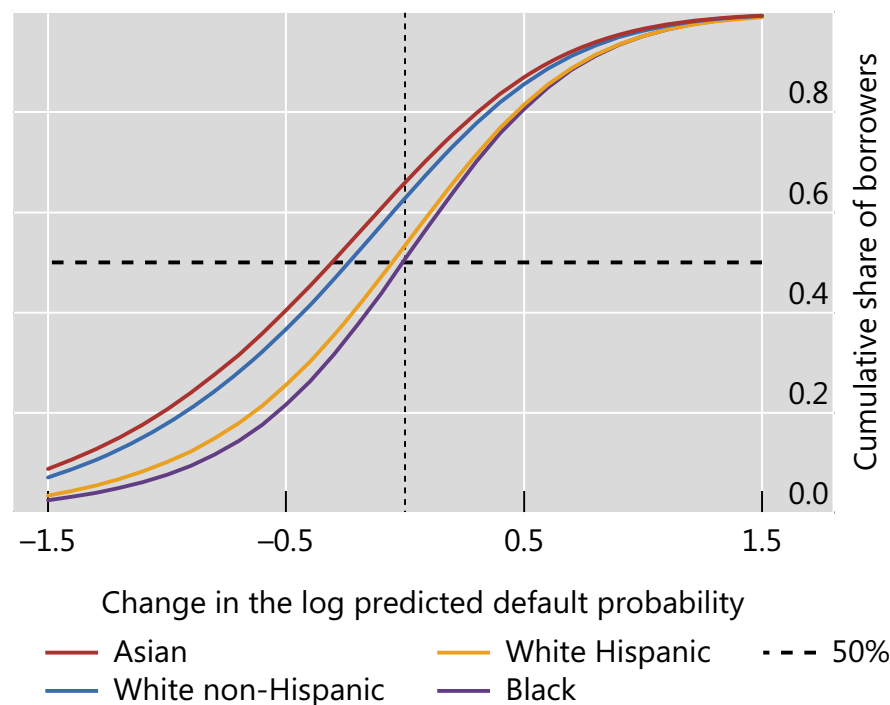
[1] Adults without an account; for each income group, weighted average by population for a selected set of countries. [2] On the horizontal axis is reported the change in the log predicted default probability as lenders move from traditional predictive technology (a "Logit" classifier) to machine learning technology (a "Random Forest" classifier). On the vertical axis is reported the cumulative share of borrowers from each ethnic group that experience a given level of change.

Sources: World Bank, *Global Findex;* World Bank, *Remittance Prices Worldwide,* remittanceprices.worldbank.org; Fuster et al (2018).

References

Berg, T., V. Burg, A. Gombović and M. Puri (2018). "On the rise of fintechs – credit scoring using digital footprints", NBER Working Paper no 24551.

BIS (2019). Annual Economic Report. Chapter III: "Big tech in finance: opportunities and risks".

Cornelli, G., V. Davidson, J. Frost, L. Gambacorta, K. Oishi (2019), "SME Finance in Asia: Recent Innovations in Fintech Credit, Trade Finance and Beyond", ADBI Working Paper 1027.

Financial Stability Board (2019). "Evaluation of the effects of financial regulatory reforms on small and medium-sized enterprise (SME) financing: Consultation report", June.

Frost, J. (forthcoming). "The Economic Forces Driving FinTech Adoption across Countries", in King, M. and Nesbitt, R. (eds.), *The Technological Revolution in Banking*, Toronto: University of Toronto Press.

Frost, J., L. Gambacorta, Y. Huang, H. Shin, and P. Zbinden (2019). "BigTech and the changing structure of financial intermediation", Economic Policy, forthcoming (also published as BIS Working Paper Series, no. 779).

Fuster, A. P. Goldsmith-Pinkham, T. Ramadorai and A. Walther (2019). "Predictably unequal? The effect of machine learning on credit markets", 2018, mimeo.

Klein, A. (2017). "Hard Drive Cost Per Gigabyte", Backblaze, July

Monetary Authority of Singapore (MAS) (2018). "Financial Stability Report", November.

Pereira L., J Frost and L. Gambacorta (2019). "Welfare implications of digital financial innovation".

Petralia, K., T. Philippon, T. Rice and N. Veron (2019). "Banking Disrupted? Financial Intermediation in an Era of Transformational Technology", Geneva Report No. 22.

Reinsel, D., J. Gantz and J. Rydning (2017). "Data Age 2025: The Evolution of Data to Life-Critical", IDC White Paper, April.

Background
OOOOOO

Artificial Intelligence
OOOO

The limits of AI
OOOOO

To conclude
OO

Data driven financial stability

Robert Macrae

Systemic Risk Centre
London School of Economics

modelsandrisk.org/AI

r.macrae@lse.ac.uk
Data driven financial stability
2 December 2019

From

- Artificial intelligence and Systemic Risk. Jón Daníelsson, Robert Macrae and Andreas Uthemann
- Available at SSRN: https://ssrn.com/abstract=3410948
- VoxEU Artificial intelligence and- the stability of markets https://voxeu.org/article/artificil-intelligence-and-stability-markets

What is machine learning?

- Broad family of highly flexible, expressive models
- Forests of boosted or bagged regression trees, neural networks, SVMs...
- Theory-light
- Overfitting controlled by cross validation

What ML can and cannot do

- AI can master any decision process with a *defined action space* better than any human
  - chess, go, computer games,...
- It can handle decisions to the extent they can be mapped onto a contained local problem
  - driving a car, medical diagnosis, allocation of credit
- If the action space is ill defined (like all human endeavours), not so easy
- AI today is *unable to reason about things it has not seen*
- It can generalise within a local problem but cannot apply experiences from one domain to another
- It does not understand the global problem that contains the local one

Systemic risk

Probability of an unlikely financial crisis causing a severe economic recession

- Systemic risk can not be eliminated
- Happens on the boundaries of silos
- One year out of every 43 for OECD countries
- Does generally not arise from the behaviour or failure of any individual financial institution
- For the United States it is not an event with costs in the USD billions or tens of billions or hundreds of billions
- Instead, several trillions
- Perhaps more than 10% of GDP

The time dimension of risk

Easy to measure risk   Measuring risk almost impossible

Frequency per century				
Daily	10	5	2 or 3	1 or 2
Event				
Client losses	Large bank losses	Large banking failure	Banking crises local systemic	Global systemic crises
Drivers				
Stocks and factors	Technological development	Social / political?	The unforeseen?	The wholly unforeseen…?

Systemic risk is all about the unknown unknowns

- The US stock market goes down by $200 billion in one day and nobody cares
- Potential subprime losses of less than $200 billion and a global crisis happens
- Risk we know we prepare for — *known unknowns*
- Risk we don't know is the dangerous type
- *Unknown unknowns* are most damaging
- Almost axiomatic that the next crisis will happen where the nobody looking

Background
000000

Artificial Intelligence
●000

The limits of AI
00000

To conclude
00

Bob — The Bank of England Bot ... and Gus, and Mel, and Barry

- MLs that are grandmasters in portfolio construction and risk management
- Trained, theory-light, in the entire corpus of historic asset price movements
- They know everything that has ever happened
- They have induced a sensible structure that permits extensive interpolation
- To the extent that the same drivers continue to apply they will do a great job

Risk management, compliance and micropru

- There is no technological reason why AI cannot play a major role in most
  - risk management functions in financial institutions
  - microprudential supervision
- Job of the supervisor and the risk manager will become high level interaction with their respective AIs
- BoB talks to the banks' bots
- Passing data, model, rules, questions and decisions
- Objections are cultural, political, legal
- Technology is mostly here
- The financial case is clear
- Substantial investment but even larger savings

Macro. Case not so clear

- There are limits to AI
- Some can be overcome with technical developments
- Others at present can not
    1. understanding reasoning
    2. procyclicality
    3. unknown unknowns
    4. optimisation against the system

The time dimension of risk

Easy to measure risk
Easy for BoB

Measuring risk almost impossible
Impossible for BoB

Frequency per century				
Daily	10	5	2 or 3	1 or 2
Event				
Client losses	Large bank losses	Large banking failure	Banking crises local systemic	Global systemic crises
Drivers				
Stocks and factors	Technological development	Social / political?	The unforeseen?	The wholly unforeseen...?

Artificial intelligence and systemic risk© 2019

Background
000000

Artificial Intelligence
0000

The limits of AI
●0000

To conclude
00

Trust, causality and reason

- A 1980s AI, EURISKO, played a naval wargame
- It found the best solution was to sink its own slowest ships
- It is impossible to specify all eventualities
- Humans can reason about unseen things, AI cannot
- Current AI without a kill switch will do something catastrophic

Procyclicality

- BoB will favour homogeneous best–of–breed methodologies and standardised processes
- Even stronger than human authorities
- In-breeding and homogeneity make behaviour more procyclical
- Which increases systemic risk

Procyclicality 2

- Financial entities have, in my experience, rather similar objectives
- "Better" solutions mean closer to the optimum
- As a consequence better solutions must also be closer to each other
- Crowded trade... or rather, crowded perception
- When new information arrives all AIs will update their models in a similar way
- They will all want to sell the same things

BoB cannot find unknown–unknowns

- Systemic vulnerabilities tend to happen on the boundaries of areas of responsibilities (the silos)
  - subprime mortgages put into structured credit products with hidden liquidity guarantees
  - crossing multiple jurisdictions, institutional categories and countries
- Where humans and AI are least likely to look
- Current AI can easily be trained on events that have happened: risk
- It can perhaps be trained on events that have been imagined: known-unknowns
- Our system is *endogenously infinitely complex*
- It will always miss unknown-unknowns.

Optimise against the system

- BoB's optimisation is harder than that of a malicious actor because BoB faces an infinitely complex computational problem
- A malicious actor only has to optimise against very small part of that domain
- A human regulator provides a natural defence because they create randomness, nuance and interpretation which varies across individuals and time
- Regulators also use common sense and understand out-of-domain constraints such as the limits to regulator powers
- For AI, such randomised responses, would have to be programmed in, and hence would not be acceptable

Effective macroprudential AI needs to:

1. control across borders
2. control across silos
3. share data across borders and silos
4. randomise responses
5. create rules in a nontransparent way

6. understand causality in in unforeseen cases
7. reason on a global rather than local basis
8. identify threats that have not yet had bad outcomes

The first 5 are unacceptable; the last 3 are beyond current capabilities

Background
oooooo

Artificial Intelligence
oooo

The limits of AI
ooooo

To conclude
o●

So...

- BoB and his friends will become increasingly useful
- Reduce costs for financial institutions
- Highly successful at least 999 days out of 1000
- Need a kill switch
- Increase systemic risk
- Increase procyclicality
- Reduce volatility and fatten tails