



Toulouse
School of
Economics

FIT IN Initiative

Research synthesis brief

October 2024

A vibrant night scene of a busy street market. The street is filled with a diverse crowd of people, including men, women, and children. In the background, there are numerous shops and stalls, some with illuminated signs. A prominent red sign for 'AC RESTAURANT AASTHA' is visible on the right. Other signs include 'BEAUTY SALON' and 'ईडेन व्युटीपार्लर'. The street is lined with streetlights, and the overall atmosphere is one of a bustling, active marketplace.

How can data sharing help improve access to credit?

Matthieu Bouvard & Seth Garz

Economics for the Common Good

How can data sharing help improve access to credit?

The growing digitization of consumer transactions, payment transfers, government records, search behavior, and social communication is generating compelling new data sources that are expected to influence credit underwriting by lending institutions and borrowing behavior by consumers. Many analysts suggest these new data sources can help to expand access to credit among consumers and small businesses in low- and middle-income countries (LMICs), in particular, and advocate for initiatives like Open Data and Open Banking to enable newer entrants to share access to data (Jenik et al., 2024). Other analysts note the potential for organizations like Big Tech companies with comparative advantages in access to new proprietary data sources to quickly outcompete incumbent lenders and threaten the effectiveness of regulatory regimes designed to ensure competition, consumer protection, and financial stability (Frost et al., 2019). Whether perceived as an opportunity or a threat, new data sources are likely to continue to disrupt lending. This policy brief offers an up-to-date review of the emerging academic literature on the economics of data sharing and access to credit. It aims to provide practical intuition and a common vocabulary for decision-makers engaged in the high-stakes work of designing new data-use technologies and data-sharing regulations. Key lessons include the importance of considering the incentives of organizations to invest in new data collection if they are compelled to share those data; how payment and credit underwriting services can be complements or substitutes depending on the structure of the payment and lending markets; and how empowering consumers to share data does not necessarily improve lenders' access to information if consumers value privacy differently.

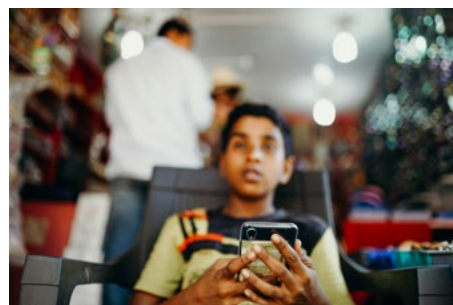
Our analysis is motivated by several related questions:

1. How should we categorize different types of data that may affect the allocation of credit?
2. To what extent do new sources of data, including financial data, improve the profitability and/or reduce the risk of underwriting?
3. To what extent do these new sources of data expand access to credit among those otherwise excluded?
4. How should we understand key policy questions related to data sharing and data protection?

Which types of data affect the allocation of credit?

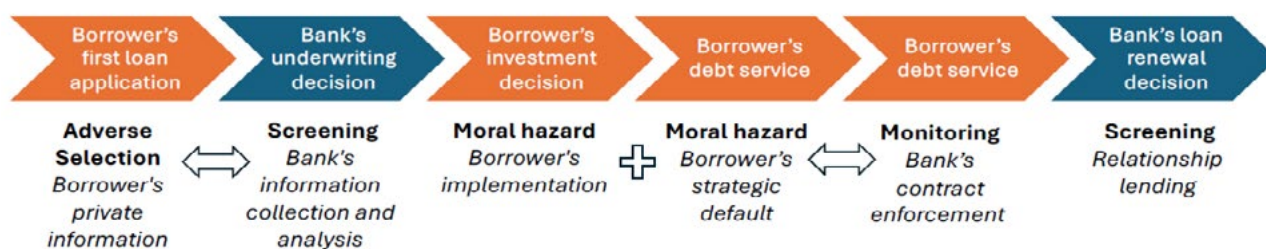
Information that helps predict the credit risk of a loan applicant is a key component of the decision to provide credit. Because credit risk is not just an innate attribute of the loan applicant but also depends on characteristics of the credit contract (e.g., loan size, repayment schedule or interest rate), information affects multiple dimensions of credit provision. In this context, data can be a critical source of information but its importance should be evaluated relative to other sources.

In addition, information plays a role at different stages of the lending relationship as illustrated in the figure below. At the underwriting stage, the key friction is asymmetric information. In particular, a lender should be worried that a borrower who accepts an offer could not get a better offer from a competitor, which may indicate this borrower is a bad risk. Information gathering and analysis or "screening" limits the lender's exposure to this adverse selection. But access to information once the loan is disbursed is also useful because it allows the lender to better "monitor" the borrower, that



is, to deploy its enforcement resources to limit the borrower's ability to divert funds or to default strategically (moral hazard). Finally, information collected throughout a lending relationship affects subsequent lending decisions, a pattern known as "relationship lending." We will discuss how access to information matters at these different points.

Information plays a role at different stages of the lending relationship



Information versus data

Information is a broader concept than data, and information economics provides a useful distinction between soft and hard information (Liberti and Petersen, 2019). The key property of hard information is to be verifiable and, therefore, transferable. Data in the form of payment records, business invoices, credit histories, etc. and summary indicators that can be derived from data, such as credit scores, are classic examples of hard information. In contrast to data, information is soft when it is difficult to transfer to a third party, often because proper use of this information is context-dependent in a way that is hard to quantify or formalize. This makes it difficult to separate soft information from the person who collects it. A classic example is the subjective judgment of a loan officer based on knowledge and expertise accumulated through interactions with consumers and businesses. A long-standing literature in banking highlights the role of soft information in the provision of credit. Fisman et al. (2017) provide evidence that cultural proximity between a loan officer and loan applicants in India improves the efficiency of credit allocation with increased access to credit, lower capital requirements, and higher repayment rates. In contrast, other literature suggests loan officer biases can constrain lending to qualified female borrowers or to immigrants (Alesina et al., 2013; Dobbie et al., 2021). The extent to which different types of data are substitutes or complements for soft information remains an important open question. While the potential gains from using data are clear in the case of gender discrimination in lending, gains also exist in the case where cultural proximity is beneficial to underwriting if using data is more cost-effective to produce information. For example, evidence in Petersen and Rajan (2002) for the U.S. or Mian (2006) for Pakistan suggests that the availability of hard information can substitute for the expertise of a local loan officer and improve banks' ability to extend credit far from their geographical base.

Information is a broader concept than data, and information economics provides a useful distinction between soft and hard information

Financial data

Payment data can provide detailed high-frequency hard information about loan applicants who may have never taken out formal loans

Credit histories have traditionally been the basis for evaluating credit risk. Credit data is often collected into credit registries and made available by private or public entities, or aggregated into credit scores. One drawback of this source, particularly in developing countries, is that it only provides information on consumers or businesses with previous interactions with a formal channel of credit. According to the 2021 Global Findex database, for example, only 23% of adults in developing countries had borrowed from a formal financial institution (Demirgüç, Kunt et al., 2022). Payment data can overcome this limitation by providing detailed high-frequency hard information about loan applicants who may have never taken out formal loans. The proliferation of digital financial

services such as mobile money in LMICs suggests such payment data are increasingly available even for inexperienced borrowers. Rishabh (2024) provides evidence in India that payment data complements sources of both hard and soft information traditionally used for loan underwriting. Specifically, payment data improves the prediction of loan delinquency relative to both credit bureau data and to what the authors identify as indicators of banks' soft information. In addition, a less appreciated benefit of payment data is that continuous access to payment information significantly improves lenders' ability to predict delinquency after loan disbursement. In other words, high-frequency payment data can be valuable not just for the underwriting stage of the lending cycle, but also for the loan monitoring stage. While payment data may improve loan underwriting and monitoring by lenders, it is not immediately obvious these efficiency gains will benefit otherwise disadvantaged potential borrowers. Fortunately, this link between payment data and improved access to credit is supported by Ghosh et al. (2023), who show, also in the context of India, that access to digital payment by small businesses is associated with an increased likelihood of obtaining a loan and with lower interest rates. Similarly, Ouyang (2023) uses Alipay data in China to provide evidence of a causal link between the adoption of cashless payments by poorer consumers and access to credit.



Alternative data

The growing digitization of commerce, government records, search behavior, and social communication is generating new "alternative" data sources that are expected to influence credit underwriting by lending institutions and borrowing behavior by consumers. This catch-all phrase - "alternative data" - covers non-banking data from demographics (e.g., education, residence) to commercial transaction-related information (e.g., customer reviews for businesses, text analysis of listings on P2P platforms) to the broadest notion of digital footprint (e.g., social network, search history, operating system used). Many fintech entrants (e.g., online lenders, P2P lending platforms) have explicitly advertised the ability of their business models to take advantage of alternative data to compete with traditional banks (Jagtiani and Lemieux, 2019). The entry of Big Tech platforms into the credit market may reflect their superior access to proprietary data that improves ex-ante screening and ex-post monitoring. For example, whereas payments and credit-card companies facilitating merchant payments may only have access to microdata on the amount transferred between parties, the general sector of a merchant's business, and a few other data fields, e-commerce platforms know specific products purchased, prior search history, etc. In principle, search engines, could infer the likelihood of economically meaningful events - such as divorce, new children, or changes in employment - from an individual's search history without even observing an individual's transactions or payments. Combined with the lack or ambiguity of regulation of digital platforms in some countries, these examples illustrate the potential for Big Tech to profoundly disrupt the traditional banking business model (Frost et al., 2019). Disrupting the traditional banking model may not be bad news for the vast underbanked population in LMICs. Advocates suggest alternative data could broaden credit access to applicants with a limited financial history and increase financial inclusion. While uses of alternative data are still relatively new in LMICs, in particular, there is already some select evidence of impact from the US, China, and Germany. Di Maggio et al. (2022) show that a U.S. fintech business model based on alternative data, notably education and job history, outperforms the model used by the Consumer Financial Protection Bureau and is particularly beneficial to borrowers with low credit scores and short credit histories. Hau et al. (2019) show that the use of Alibaba e-commerce transaction data led Ant Financial to target credit at borrowers with low credit scores. Berg et al. (2020) provide direct evidence that the digital footprint of consumers on an e-commerce platform provides information that complements credit bureau information to help predict delinquencies. On the other



hand, Berg et al. (2022) highlight that models based on alternative data do not seem to show a significant competitive advantage over banks' models, raising the question of the complementarity or substitutability of alternative data with more traditional sources of information. In particular, the authors argue there is no consistent evidence that fintechs outperform banks in terms of realized default. Notably, there is a clear need to expand the evidence on the use of alternative data for credit in LMICs.

A holistic approach to data sharing

A combination of regulations and technological infrastructures (e.g., APIs, access to mobile devices, internet connectivity) is typically required to facilitate the exchange of data. As policymakers and providers advance efforts to improve the sharing of financial and alternative data for lending purposes, it will be important to think holistically about the potential effects of data sharing. By impacting the competitive landscape and the distribution of credit across different segments, data sharing can have complicated implications for consumer welfare that demand further investigation.

As policymakers and providers advance efforts to improve the sharing of financial and alternative data for lending purposes, it will be important to think holistically about the potential effects of data sharing

Competitive effects

To the extent that access to data improves the evaluation of credit risk, data sharing should make the market for credit more competitive. In particular, by leveling access to information, data sharing can alleviate adverse selection concerns for lenders and lead them to compete more aggressively. This holds for incumbents but the entry of new players such as fintechs could also contribute to more intense competition. For example, Babina et al. (2024) show that open banking regulations in the U.K. that mandate banks to share customer information with third parties (with customer consent) led to SMEs getting improved access to credit, from fintech lenders, in particular. Payment data sharing could also have an impact on competition in the payment market (Parlour et al., 2022). With no data sharing, the competitive advantage provided by access to payment data could lead payment providers to underprice their services. Absent this motivation (say, if data is shared with no compensation), the price of payment services could go up. While we often think of price increases as harming consumers, this price increase need not be harmful if consumers equally benefit from changes enabled by data sharing in the downstream credit market. In other words, it is important to consider the pricing and regulation of payments and credit together, particularly in markets where credit is offered by payment service providers.

Distributional effects

As well as impacting the competitive advantage of providers, data sharing will also likely effect the distribution of credit among consumers. This raises the concern that data sharing may not benefit all consumers in the credit market and, under some circumstances, could expand inequalities in credit access. In Babina et al. (2024), for example, the positive effect of open banking on the credit offering for small and medium-sized enterprises (SMEs) is concentrated among borrowers with existing access to credit. One explanation is that data sharing is more beneficial to those who are in a position to share richer financial data, i.e. those who already have access to the financial system. Babina et al. (2024) conclude there is little evidence of an effect of open banking on financial inclusion in the U.K. case. Additionally, there is broad concern that the training datasets and machine learning models enabled by these data may amplify structural biases. For example, Fuster et al. (2022) show that the introduction of machine learning models in the U.S. mortgage market is less likely to benefit Black and Hispanic borrowers. In other words, it will be imperative to monitor the changing distribution of credit and to evaluate trade-offs that may arise between more efficient allocation of credit and more equal distribution of credit. There is a danger that more data and more powerful algorithms will only be more efficient at excluding the underbanked.

Data sharing may not benefit all consumers in the credit market and, under some circumstances, could expand inequalities in credit access

Privacy concerns affect data-sharing incentives

A straightforward effect of privacy concerns is to limit the effect of an improved ability to share data. To the extent that customers ultimately control the amount of data they share, a lack of trust in financial intermediaries, particularly in new entrants such as fintechs, can undo the positive competitive effects mentioned above (Tang, 2019). Survey evidence in Armantier et al. (2024) shows systematic differences in willingness to share data across genders and age groups with implications for the distributional effects of data sharing. A more indirect effect of enabling data sharing is to create new costs for maintaining one's privacy. Even when consumers can choose whether or not to share data, the fact that sharing data is possible renders data retention suspicious. In a credit market context, for example, a lender may struggle to discern whether a consumer is not sharing data due to privacy concerns or to hide negative information (He et al., 2023). This can harm consumers with high privacy concerns. If they still choose to retain data despite the stigma, they might face worse credit terms. If they choose to share data, they bear the corresponding cost of losing privacy contrary to their preferences.



A straightforward effect of privacy concerns is to limit the effect of an improved ability to share data

Finally, a third effect relates to what economists refer to as the “externalities” created by consumers’ decisions to share data. The concern here is that data sharing by a group of consumers helps uncover patterns that provide information about other (similar) consumers. In other words, data sharing creates a costly loss of privacy for those who place a higher value on privacy and are viewed similarly by machine algorithms (Acemoglu et al., 2022). As a stylized example, if lending algorithms are trained on data from other borrowers with similar last names to yours from similar postal codes, they may not need your personal transaction information to characterize your credit rating as good or bad.

Policy implications

Financial inclusion

Because of the limited experience with formal credit in LMICs, access to payment data will be critical to expand lending for the underbanked. Adoption of digital payment by those who otherwise use cash will only increase the population of consumers who can benefit from the use of payments data for underwriting. To ensure competition among lenders that use payments data for lending, policymakers may seek to encourage or even mandate data sharing. However, policymakers in LMICs with large unbanked populations must carefully balance how mandating data sharing can discourage payment providers from investing in efforts to acquire new clients who would otherwise be financially excluded. This is because payment providers may not enjoy the full value of payment data when compelled to share it with other potential lenders. Conversely, enabling data sharing may increase consumer (and merchant) adoption of digital payments if consumers can realize improved access to credit through the digitization of their transactions (Ghosh et al., 2023). Ultimately, the design of the policy should depend on the market penetration of digital payment and it may need to be differently applied to data from entrepreneurs, merchants, and other firms versus consumers.



The design of the policy should depend on the market penetration of digital payment and it may need to be differently applied to data from entrepreneurs, merchants, and other firms versus consumers

Open banking, open finance and open data

Open banking – or, more broadly, open finance – has emerged as the regulatory framework of choice for the sharing of financial data. While implementation differs across jurisdictions, the unifying principle is a mandate for financial institutions to give qualified third parties access to a consumer’s financial data (mostly payment data at this stage) conditional on this consumer’s consent. The rationale for regulation is based on the idea discussed earlier that exclusive access to consumers’ financial data gives a competitive advantage in the credit market and is therefore a barrier to entry. Jones and Tonetti (2020) show that in such context, firms have insufficient incentives to share data even when a market for data exists and that giving consumers ownership over their data (consistent with the open-banking approach) can enhance welfare. One insight from the cross-country study in Babina et al. (2024) is that the technical aspects of such regulations matter. Without technical requirements on the implementation – for instance, about how APIs should be designed – it will be more difficult for institutions to coordinate on a standard and there will be more room for these institutions to effectively limit data sharing. To the extent that alternative data also impacts access to credit, broader regulations on the sharing of non-financial data (such as the EU’s Data Act) should also be relevant, particularly in a context where large social media or e-commerce platforms become active in the credit market.

Broader regulations on the sharing of non-financial data (such as the EU’s Data Act) should also be relevant



Privacy protection

As discussed earlier, privacy concerns can limit data sharing and its potential benefits, as well as creating costs for consumers when data is shared. Evidence in Babina et al. (2024) suggests that open-banking policies are less effective when trust in fintechs is lower. On the other hand, respondents in Armantier et al. (2024) report a higher willingness to share transaction data when privacy regulations are stricter, suggesting a role for policy. Consistent with this finding, Doerr et al. (2023) provide evidence that the implementation of

stricter consumer privacy laws in California in 2020 led to a higher volume of loan applications to fintechs. Rather than being driven by concerns about externalities, the rationale for this regulation is to provide a credible way for firms to commit not to misuse consumer data and to implement appropriate security. Rules on privacy protection inherently restrict firms’ ability to use data (or make it more costly) so they need to strike a balance between bolstering consumer confidence and letting firms exploit richer data (Rishabh, 2024).

Privacy protection rules need to strike a trade-off between bolstering consumer confidence and letting firms exploit the potential of accessing richer data

KEY POLICY INSIGHTS

Financial inclusion

- Expanding access to payment data is essential to increase lending opportunities for the underbanked.
- Policymakers in LMICs must carefully consider how to encourage data sharing.
 - Payment providers may be reluctant to invest in acquiring new clients if compelled to share data with rivals.
 - Mandated data sharing may increase adoption of digital payments if consumers gain access to credit (Ghosh et al., 2023).
 - Policy design should depend on market penetration of digital payments and may need to be differentially applied.

Regulation of open finance

- Open banking policies usually require financial institutions to grant third-party access to consumer data.
- Effective implementation requires addressing technical standards to facilitate seamless data sharing across platforms.
- Broader regulations on sharing of non-financial data (such as the EU's Data Act) are needed, particularly for large social media or e-commerce platforms.

Privacy protection

- Open-banking policies are less effective when trust in fintechs is lower (Babina et al., 2024).
- Stricter privacy rules may increase loan applications (Doerr et al., 2023) and willingness to share transaction data (Armantier et al., 2024).
- Privacy rules must navigate the balance between bolstering consumer confidence and letting firms exploit richer data (Rishabh, 2024).

About the author

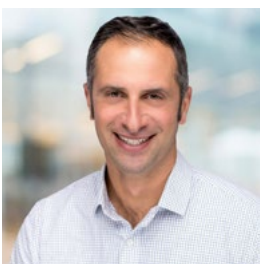


Matthieu Bouvard

Toulouse School of Economics

Matthieu's research focuses on financial intermediation. He investigates how information technologies affect the design and delivery of financial products and the implications for risk and stability in financial systems. His current research studies the entry of e-commerce platforms into the payment and the credit market and the role of mobile money agents in developing countries. His paper on blockchains published in *The Review of Financial Studies* has received the Swiss Finance Institute Outstanding paper award in 2017.

matthieu.bouvard@tse-fr.eu



Seth Garz

Bill & Melinda Gates Foundation

Seth Garz is Deputy Director for Data & Evidence with the Women's Economic Empowerment strategy at the Bill & Melinda Gates Foundation where he also manages the research portfolio for the Inclusive Financial Systems strategy. Seth serves as a consultative member of the advisory board of the FIT IN Initiative and is an Affiliate Professor in the Economics Department at the University of Washington.

seth.garz@gatesfoundation.org

References

- Acemoglu, D., Makhdoumi, A., Malekian, A. and Ozdaglar, A., 2022. [Too much data: Prices and inefficiencies in data markets](#). *American Economic Journal: Microeconomics*, 14(4), 218-256.
- Armantier, O., Doerr, S., Frost, J., Fuster, A. and Shue, K., 2024. [Nothing to hide? Gender and age differences in willingness to share data](#). SSRN Working Paper, <https://ssrn.com/abstract=4808467>.
- Alesina, A. F., Lotti, F., and Mistrulli, P. E., 2013. [Do women pay more for credit? Evidence from Italy](#). *Journal of the European Economic Association*, 11(1), 45-66.
- Babina, T., Bahaj, S.A., Buchak, G., De Marco, F., Foulis, A.K., Gornall, W., Mazzola, F. and Yu, T., 2024. [Customer data access and fintech entry: Early evidence from open banking](#). NBER Working Paper 32089.
- Berg, T., Burg, V., Gombović, A. and Puri, M., 2020. [On the rise of fintechs: Credit scoring using digital footprints](#). *The Review of Financial Studies*, 33(7), 2845-2897.
- Berg, T., Fuster, A. and Puri, M., 2022. [Fintech lending](#). *Annual Review of Financial Economics*, 14(1), 187-207.
- Demirgüç-Kunt, A., Klapper, L., Singer, D., and Ansar, S., 2022. [The Global Findex Database 2021: Financial inclusion, digital payments, and resilience in the age of COVID-19](#). World Bank Publications.
- Di Maggio, M., Ratnadiwakara, D. and Carmichael, D., 2022. [Invisible primes: Fintech lending with alternative data](#). NBER Working Paper 29840.
- Dobbie, W., Liberman, A., Paravisini, D., and Pathania, V., 2021. [Measuring bias in consumer lending](#). *The Review of Economic Studies*, 88(6), 2799-2832.
- Doerr, S., Gambacorta, L., Guiso, L. and Sanchez del Villar, M., 2023. [Privacy regulation and fintech lending](#). SSRN Working Paper, <https://ssrn.com/abstract=4353798>.
- Fisman, R., Paravisini, D. and Vig, V., 2017. [Cultural proximity and loan outcomes](#). *American Economic Review*, 107(2), 457-492.
- Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T. and Walther, A., 2022. [Predictably unequal? The effects of machine learning on credit markets](#). *The Journal of Finance*, 77(1), 5-47.
- Frost, J., Gambacorta, L., Huang, Y., Shin, H. S., and Zbinden, P., 2019. [BigTech and the changing structure of financial intermediation](#). *Economic Policy*, 34(100), 761-799.
- Hau, H., Huang, Y., Shan, H. and Sheng, Z., 2019. [How FinTech enters China's credit market](#). *AEA Papers and Proceedings*, 109, 60-64.
- He, Z., Huang, J. and Zhou, J., 2023. [Open banking: Credit market competition when borrowers own the data](#). *Journal of Financial Economics*, 147(2), pp.449-474.
- Ghosh, P., Vallee, B. and Zeng, Y., 2023, March. [FinTech lending and cashless payments](#). *Journal of Finance*, forthcoming.
- Jagtiani, J. and Lemieux, C., 2019. [The roles of alternative data and machine learning in fintech lending: evidence from the LendingClub consumer platform](#). *Financial Management*, 48(4), 1009-1029.
- Jenik, I., Mazer, R. and Fernandez Vidal, M., 2024. [The Building Blocks Supporting Open Finance](#). CGAP Working Paper.
- Jones, C.I. and Tonetti, C., 2020. [Nonrivalry and the Economics of Data](#). *American Economic Review*, 110(9), 2819-2858.
- Liberti, J.M. and Petersen, M.A., 2019. [Information: Hard and soft](#). *Review of Corporate Finance Studies*, 8(1), 1-41.
- Mian, A., 2006. [Distance constraints: The limits of foreign lending in poor economies](#). *The Journal of Finance*, 61(3), 1465-1505.
- Ouyang, S., 2023. [Cashless payment and financial inclusion](#). SSRN Working Paper, <https://ssrn.com/abstract=3948925>.
- Parlour, C.A., Rajan, U. and Zhu, H., 2022. [When fintech competes for payment flows](#). *The Review of Financial Studies*, 35(11), pp.4985-5024.
- Petersen, M.A. and Rajan, R.G., 2002. [Does distance still matter? The information revolution in small business lending](#). *The Journal of Finance*, 57(6), 2533-2570.
- Rishabh, K., 2024. [Beyond the Bureau: Interoperable Payment Data for Loan Screening and Monitoring](#). FIT IN Working Paper.
- Tang, H., 2019. [The value of privacy: Evidence from online borrowers](#). Working Paper

About the FIT IN Initiative

In November 2020, the Toulouse School of Economics launched the Financial Inclusion Through Interoperability Initiative to catalyze new research to constructively influence the design and regulation of interoperable digital financial services systems in low- and middle-income countries.

The main objective of this four-year research initiative is to better understand the implications of alternative competition and regulatory policies and ultimately inform policies to expand the scope, improve the quality and reduce the cost of digital payment systems for impoverished users.

The FIT IN Initiative receives support from the Bill & Melinda Gates Foundation's Financial Services for the Poor program.

For more information: www.tse-fr.eu/groups/FIT-IN-Initiative / ftininitiative@tse-fr.eu



FIT IN Initiative

Photos: itsehsanh, jack-sparro, framesbypc, cottonbro, TSE

Toulouse School of Economics

1, Esplanade de l'Université

31080 Toulouse Cedex 06

FRANCE

Tel: +33 (0)5 67 73 27 68

www.tse-fr.eu/groups/FIT-IN-Initiative