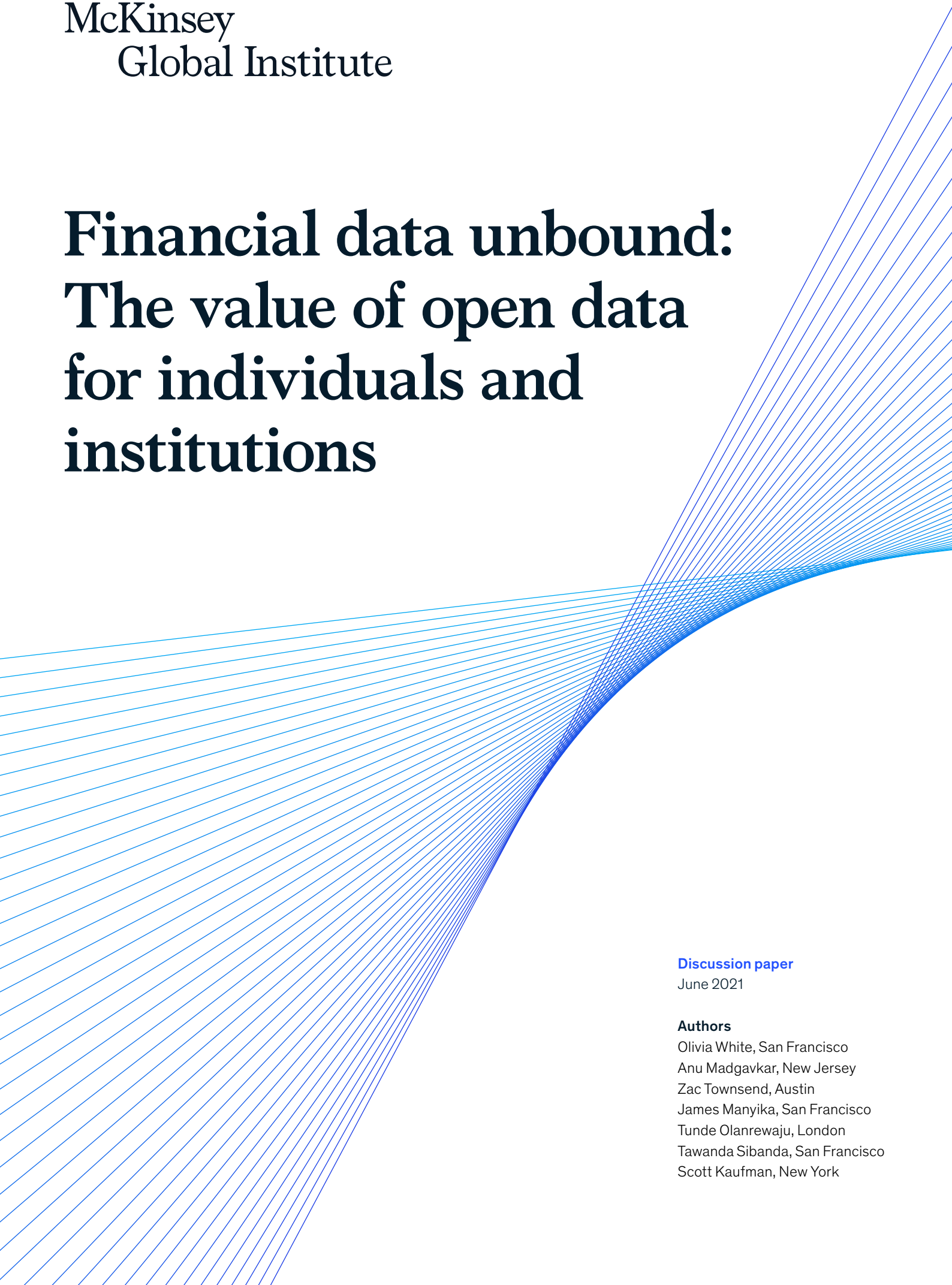
McKinsey Global Institute

Financial data unbound: The value of open data for individuals and institutions



Discussion paper
June 2021

Authors
Olivia White, San Francisco
Anu Madgavkar, New Jersey
Zac Townsend, Austin
James Manyika, San Francisco
Tunde Olanrewaju, London
Tawanda Sibanda, San Francisco
Scott Kaufman, New York

McKinsey Global Institute

Since its founding in 1990, the McKinsey Global Institute (MGI) has sought to develop a deeper understanding of the evolving global economy. As the business and economics research arm of McKinsey & Company, MGI aims to help leaders in the commercial, public, and social sectors understand trends and forces shaping the global economy.

MGI research combines the disciplines of economics and management, employing the analytical tools of economics with the insights of business leaders. Our "micro-to-macro" methodology examines microeconomic industry trends to better understand the broad macroeconomic forces affecting business strategy and public policy. MGI's in-depth reports have covered more than 20 countries and 30 industries. Current research focuses on six themes: productivity and growth, natural resources, labor markets, the evolution of global financial markets, the economic impact of technology and innovation, and urbanization. Recent reports have assessed the digital economy, the impact of AI and automation on employment, physical climate risk, global health, income inequality, the productivity puzzle, the economic benefits of tackling gender inequality, a new era of global competition, Chinese innovation, and digital and financial globalization.

MGI is led by three McKinsey & Company senior partners: co-chairs James Manyika and Sven Smit and director Jonathan Woetzel. Michael Chui, Mekala Krishnan, Susan Lund, Anu Madgavkar, Jan Mischke, Jaana Remes, Jeongmin Seong, and Tilman Tacke are MGI partners. Project teams are led by the MGI partners and include consultants from McKinsey offices around the world. These teams draw on McKinsey's global network of partners and industry and management experts.

The MGI Council is made up of McKinsey leaders and includes Hemant Ahlawat, Michael Birshan, Andrés Cadena, Sandrine Devillard, André Dua, Kweilin Ellingrud, Katy George, Rajat Gupta, Eric Hazan, Solveigh Hieronimus, Acha Leke, Clarisse Magnin-Mallez, Jurica Novak, Gary Pinkus, Hamid Samandari, Sha Sha, Oliver Tonby, and Eckart Windhagen. The Council members help shape the research agenda, lead high-impact research, and share the findings with decision makers around the world. In addition, leading economists, including Nobel laureates, advise MGI research.

This report contributes to MGI's mission to help business and policy leaders understand the forces transforming the global economy and prepare for the next wave of growth. As with all MGI research and reports, this work is independent and reflects our own views. This report was not commissioned or paid for by any business, government, or other institution, and it is not intended to promote the interests of McKinsey's clients. For further information about MGI and to download reports, please visit www.mckinsey.com/mgi.

Preface

This discussion paper looks at the potential value that could be created—and the key issues that need to be addressed—by adoption of open data for finance. It comes at a time of lively debate among financial institutions and policy makers in several countries about creating data-sharing financial ecosystems that can improve efficiencies for customers and institutions alike.

The research is part of an ongoing series by the McKinsey Global Institute examining aspects of digital financial infrastructure. It follows a report in April 2019, *Digital identification: A key to inclusive growth*, which examined how digital identification could be a new frontier in value creation, and an article published in January 2021, *COVID-19: Making the case for robust digital financial infrastructure*, which highlighted how the pandemic was a tough, real-life stress test for government disbursement schemes, with success hinging both on availability of digital IDs and on data tethered to those IDs.

The research was conducted by Olivia White, a McKinsey senior partner in San Francisco, and Anu Madgavkar, a McKinsey Global Institute partner based in New Jersey, together with Zac Townsend, an associate partner based in Austin; James Manyika, a senior partner based in San Francisco; Tunde Olanrewaju, a senior partner based in London; Tawanda Sibanda, a partner based in San Francisco; and Scott Kaufman, a consultant based in New York. The working team comprised Maria Jesus Ramirez, Maclaine Fields, Nawel Gabouge, Krzysztof Kwiatkowski, Tomasz Mataczynski, and Alexis Yang.

Our research was informed by extensive consultations with our research collaboration partners, Flourish Ventures. We owe them thanks for their expertise and network, with a particular debt of gratitude to Kabir Kumar, Tilman Ehrbeck, Tina Moran, and Chuck Muckenfuss.

Several individuals and organizations provided valuable input and guidance into the research: Nandan Nilekani, co-founder and chairman of Infosys and founding chairman of the Unique Identification Authority of India; Pia Tayag and David Symington at the office of the UN Secretary General's Special Advocate for Inclusive Finance for Development; Jason Lamb at the Bill and Melinda Gates Foundation; Kimberley Bella and Bob Hedges at Visa; Jo Ann Barefoot at Barefoot Innovation; Jennifer Tescher and Da Murphy at the Financial Health Network; Aaron Klein at the Brookings Institution; Melissa Koide at FinRegLab; and Nick Cook at the Financial Conduct Authority.

Several McKinsey colleagues provided expert advice: Lindsay Anan, Chandana Asif, Oliver Bevan, Jim Boehm, Adam Broitman, Renzo Comolli, Klaus Dallerup, Raj Dash, Olivier Denecker, Malcolm Gomes, Arvind Govindarajan, Sid Gupta, Solveigh Hieronimus, Nikhil Kaithwal, Akshay Kapoor, Somesh Khanna, Andreas Kramer, Dmitry Krivin, Alexis Krivkovich, Alok Kshirsagar, Pankaj Kumar, Glen Kushta, John Larson, Ryan Luby, Susan Lund, Deepa Mahajan, Adrian Murphy, Marie-Claude Nadeau, Fritz Nauck, Kirtiman Pathak, Greta Radaelli, Kausik Rajgopal, Bryan Richardson, Kayvaun Rowshankish, Hiro Sayama, Eric Schweikert, Joydeep Sengupta, Marcus Sieberer, Shwaitang Singh, Owen Sperling, Dan Stephens, Renny Thomas, Chetan Venkatesh, Ahalya Vijayasrinivasan, Daniel Wallance, Rob Wavra, Dan Williams, Eckart Windhagen, and Kuba Zielinski.

The paper was edited and produced by Peter Gumbel, MGI editorial director, together with operations manager Vasudha Gupta and senior graphic designers Marisa Carder and Richard Johnson. Cathy Gui and Rebeca Robboy, MGI directors of external communications, together with Matt Cooke, communications director for McKinsey's global banking practice, helped disseminate and publicize the research. Lauren Meling, digital editor, ensured that this paper was digitally available to all.

This discussion paper contributes to MGI's mission to help business and policy leaders understand the forces transforming the global economy. As with all MGI research, this research is independent and has not been commissioned or sponsored in any way by business, government, or other institution. We welcome your comments at MGI@mckinsey.com

James Manyika
Director and Co-chair, McKinsey Global Institute
Senior Partner, McKinsey & Company
San Francisco

Sven Smit
Director and Co-chair, McKinsey Global Institute
Senior Partner, McKinsey & Company
Amsterdam

Jonathan Woetzel
Director, McKinsey Global Institute
Senior Partner, McKinsey & Company
Shanghai

June 2021

Contents

In brief

Financial data unbound: The value of open data for individuals and institutions

Open data for finance is the ability to share financial data across financial institutions with limited effort or manipulation through a digital ecosystem. At a time when many countries are looking to strengthen their digital financial infrastructure, this paper examines the main mechanisms through which financial data sharing can create value, and for which market participants, and sizes the potential lift to economic growth that a well-functioning system could bring about. To help identify similarities and differences across a range of economies, we focus on four regions: the European Union, India, the United Kingdom, and the United States. Key findings include the following:

Open-data ecosystems facilitate frictionless interactions between financial institutions and consumers—individuals as well as micro-, small, and medium-size businesses—generating value for both sides. We find seven mechanisms that drive value creation across the financial services life cycle. Three of these directly benefit consumers: increased access to financial services, which in turn can boost credit; greater user convenience; and improved product options. The four others directly benefit institutions providing financial services: increased operational efficiency; better fraud protection; improved workforce allocation, for example using open data to identify high-risk customers who can become the focus of collections efforts; and reduced friction in data intermediation.

The boost to the economy from broad adoption of open-data ecosystems could be as high as 1.5 percent of GDP in 2030 in the European Union, the United Kingdom, and the United States, and as much as 4 to 5 percent in India. Emerging economies stand to benefit more than advanced ones because they tend to have lower levels of financial inclusion and less financial depth. All market participants benefit, although to varying degrees depending on region. In India, for example, consumers would stand to gain the most from the newfound ability to access appropriate financial services with relative ease. In advanced economies, institutions would stand to gain a relatively larger share of potential economic value from being able to carry out existing processes more efficiently, in a more targeted way, and with less fraud.

Capturing the full value of open data requires both a level of data standardization and a breadth of data sharing that are not yet enabled in many economies. Our research suggests that, in the developed economies we examined, only 10 to 20 percent of the potential value from open financial data is currently accessible. In India, we estimate that roughly 60 percent is accessible. In the European Union and the United Kingdom, standardization levels are high, but the breadth of data shared is more limited. By contrast, in the United States, a wide range of data is shared, but standardization is limited; private financial data aggregators there broker data flows between providers and users, with limited consumer control. In India, the data ecosystem has moderately high openness as well as breadth of sharing, using the nation's IndiaStack ecosystem, which includes layers for identity, authentication, payments, paperless data exchange, and user consent.

Beyond open-data enablement, countries would need to develop supportive digital infrastructure and frameworks to safeguard consumers. For data ecosystems to flourish, significant questions about how to ensure user consent and data security need to be addressed. In addition, robust digital financial infrastructure, including digital ID, and product innovation are essential. While the practical and policy implications are challenging to navigate, our research shows that the innovation such data ecosystems could enable would be a spur to economic recovery and broader-based prosperity in the postpandemic era.

Financial data unbound: The value of open data for individuals and institutions

1. How open financial data creates value

Financial data are created or used throughout the life cycle of financial services. They accompany every step of the consumer journey, from understanding available lending, insurance, payments, and wealth management products and services, to purchasing, using, and finally exiting them. Similarly, financial data are created or used in every action taken by financial institutions as they engage with customers: product and service design, marketing, decision making, onboarding, servicing, monitoring, making termination decisions, and closing accounts. Each interaction between customers and financial institutions provides a test case for how effectively data are shared and the value created for both sides. Financial data sharing is currently limited in many parts of the financial services value chain, resulting in avoidable friction, cost, and reduced financial access.

That could change: technological, regulatory, and competitive forces are moving markets toward easier and safer financial data sharing around the world.[1] Over the past few years, the United Kingdom formed the Open Banking Implementation Entity, a multibank platform to facilitate open application programming interfaces (APIs) and data-sharing standards, while the European Union moved ahead with new electronic payments regulations, the second payment services directive.[2] In the United States, an initiative of the Consumer Financial Protection Bureau aims to facilitate a consumer-authorized data-sharing market, while the Financial Data Exchange consortium attempts to promote common, interoperable standards for secure access to financial data.[3] In Asia, open API frameworks have been initiated in Hong Kong, and India is fast scaling its interoperable unified payment stack and financial account aggregator mechanism.[4] New consumer protection laws have been framed in Australia.[5] In Latin America, Brazil's central bank has drafted guidelines around data access, technical standards, and consumer consent that will require financial institutions to give open-data access to nonbank third parties, and other countries in the region may follow closely behind.[6] In Africa, a not-for-profit industry group in Nigeria—the Open Technology Foundation—has been set up to develop open data, and market players are also taking the lead in South Africa.[7] At the same time, many African nations have adopted data protection regulation, using the European Union's General Data Protection Regulation (GDPR) as an example framework.

1 See Laura Brodsky and Liz Oaks, "Data sharing and open banking," September 2017, McKinsey.com; Laura Brodsky, Chris Ip, and Tobias Lundberg, "Open banking's next wave: Perspectives from three fintech CEOs," August 2018, McKinsey.com; and Max Chuard, *What's next for Open Banking*, World Economic Forum, April 2021.

2 *Annual report,* Open Banking Implementation Entity, openbanking.org.uk; Napala Pratini, *An introduction to UK open banking*, May 2019, fin.plaid.com; Alessio Botta, Nunzio Digiacomo, Reinhard Höll, and Liz Oakes, "PSD2: Taking advantage of open-banking disruption," January 2018, McKinsey.com; and *Everything you need to know about PSD2,* BBVA, October 17, 2019.

3 *CFPB outlines principles for consumer-authorized financial data sharing and aggregation*, Consumer Financial Protection Bureau, October 18, 2017; *Financial Data Exchange (FDX) adds 18 new members*, Financial Data Exchange, May 4, 2021.

4 *Open application programming interface for the banking sector,* Hong Kong Monetary Authority, hkma.gov.hk; Sampath Putrevu, "IndiaStack introduces account aggregator initiative to drive financial inclusion in a fair and equitable manner," YourStory, February 23, 2020, yourstory.com.

5 Andrada Coos, *Australian government kicks of privacy act review*, Endpoint Protector, November 18, 2020.

6 *Brazil open banking model: First steps*, Banco Central do Brasil, August 6, 2019.

7 Topsy Kola-Oyeneyin, Mayowa Kuyora, and Tunde Olanrewaju, "Harnessing Nigeria's fintech potential," September 23, 2020, McKinsey.com.

Yet even as many countries are evolving stronger digital financial infrastructure and data-sharing mechanisms, COVID-19 has exposed the significant limitations and gaps in their reach, a theme we explored in earlier research.[8]

This new research focuses on the potential economic value of open financial data, a key component of strong digital financial ecosystems. We identify seven broad mechanisms through which financial data sharing can create economic value. Together, these seven mechanisms span the financial services life cycle. Three directly benefit individual and micro-, small, and medium-size enterprise (MSME) customers. These are increased access to financial services, greater user convenience, and improved product options. The remaining four mechanisms directly benefit financial institutions: increased operational efficiency, better fraud protection, improved workforce allocation, and reduced friction in data intermediation.

To size the overall economic value that open financial data can create, we quantify the potential of 24 use cases in banking and payments that are grouped under these seven mechanisms. We then scale up from these use cases to develop a broader view of the macroeconomic gains (see Box 1, "Our research methodology").

Each use case translates into some economic gain for consumers, institutions, and the economy. Exhibit 1 shows the use cases grouped by mechanism and where they operate in the financial services life cycle.

[8] Olivia White, Anu Madgavkar, Tawanda Sibanda, Zac Townsend, and Maria Jesus Ramirez, "COVID-19: Making the case for robust digital financial infrastructure," January 2021, McKinsey.com.

Exhibit 1

Open financial data creates value for individuals, MSMEs, and financial institutions across the financial services life cycle.

Use cases benefiting individuals and MSMEs[1]

Consumer actions	Increased access to financial services	Greater user convenience	Improved product options
Understand			6. Access to competitive mortgages, traditionally facilitated by brokers, for individuals
Purchase	1. Newly allowing access to retail credit through alternative credit underwriting for individuals 2. Newly allowing access to retail credit through alternative credit underwriting for MSMEs	4. Simplified application and onboarding process for MSMEs	
Use	3. Retail debt consolidation at reduced interest rates for individuals		7. Increased deposit yields through easier account switching for individuals 8. Increased deposit yields through easier account switching for MSMEs 9. Improved customer segmentation to reduce switching costs for individuals 10. Improved customer segmentation to reduce switching costs for MSMEs
Exit		5. Automated data portability between accounts for individuals	

Use cases benefiting financial institutions

Provider actions	Increased operational efficiency	Better fraud protection	Improved workforce allocation	Reduced friction in data intermediation
Design and market	11. Marketing efficiency through data-driven targeting	20. Fraud reduction through timely and comprehensive data	21. Access to market and customer data for product design	23. Direct access to lead-generation data traditionally brokered by 3rd-party providers
Decision and onboard	12. Automated underwriting of standard mortgages 13. Data availability to drive faster mortgage closure 14. Automated KYC[2] for individuals 15. Automated KYC[2] for MSMEs			24. Direct access to mortgage data traditionally brokered by 3rd-party providers
Service and monitor	16. Streamlined data entry into CRM[3] systems 17. Predictive data-driven digital/IVR[4] call center operations 18. Data-driven reductions in the cost of credit recovery		22. Data-driven reorientation of collections teams toward higher-risk borrowers	
Terminate and close	19. Automated notification of events that should trigger account closure			

1. Micro-, small, and medium-size enterprises, defined by the International Finance Corporation as enterprises with sales and/or assets less than $15 million and/or with fewer than 300 employees.
2. Know your customer data.
3. Customer relationship management.
4. Interactive voice response.

Note: Use cases are drawn from the customer life cycle in banking and payments; while not comprehensive, the 24 use cases listed here represent the most significant opportunities to create economic value through open financial data in the banking and payments value chain.

Source: McKinsey Global Institute analysis

Box 1.

Our research methodology

This research builds on previous MGI work on digital ID and the importance of digital financial infrastructure in the context of COVID-19. Our research and understanding of the potential of open financial data draws on work by our research collaborators, Flourish Ventures, a venture company that is part of the Omidyar Group.

We quantify the economic value potential of open data for finance, its attribution to different types of market participants, and how capturing the potential value relates to the shape of the data ecosystem in standardization and openness.

We define open data for finance (a term we use interchangeably with open financial data and financial data sharing) as the ability of market participants to share financial data in a manner that requires limited effort or manipulation once a data ecosystem is established. Participants range from regulators, banks, nonbank financial institutions, and nonfinancial institutions to consumers themselves—both small and medium-size enterprises and individuals—who variously act as creators, holders, and aggregators of financial data in the data-sharing ecosystem. Financial data, for the purpose of our research, is defined as any data that can be used or created by a financial services provider (for example, account balances) or a consumer (for example, name or birthdate) during a financial transaction.

Our definition of open data for finance is not tied to any specific enabling regulation but is rather a description of an outcome in the ease of data sharing that could be enabled by regulation, market forces, or some combination of the two. The level of enablement of data sharing is characterized by two critical axes: the level of data standardization and the breadth of data sharing experienced in the ecosystem by its participants.

We studied four economic regions: the European Union, India, the United Kingdom, and the United States. While the advanced economies share many similarities, including mature financial sectors, they nonetheless have some meaningful differences in their approaches to open data for finance. India provides an example of an emerging economy with a less mature financial sector, but which has taken significant steps to develop a deliberate approach to data sharing.

To quantify the potential economic impact of open data for finance in the four regions, we focused on consumer banking and payments, identifying 24 use cases, or data-sharing applications that span the customer life cycle. We grouped the use cases into seven mechanisms of value creation. The first three center on consumers: increased access to financial services, greater user convenience, and improved product options. The other four accrue mainly to financial institutions: increased operational efficiency, better fraud protection, improved workforce allocation, and reduced friction in data intermediation. The use cases we identified are not comprehensive but represent the major opportunities we see based on consumer needs, productivity and efficiency gaps, and levels of innovation in the banking and payments industry. In each case, we applied our understanding of the financial services value chain to determine if the financial institution or the consumer is the likely direct beneficiary of the value created, although the value captured would be subject to additional factors that vary across markets, such as industry structure, levels of competition, and regulation.

We first quantified the potential gains of data sharing through a set of micromodels (or use case-specific models) that estimated the economic impact of scaling the 24 specific use cases, each on a stand-alone basis, by 2030 in our focus economies. To inform our approach for the micromodels, we scanned literature for case examples from various financial ecosystems and institutions, and conducted expert interviews about potential efficiency gains and volume growth.

We relied on a variety of public and government sources of data for our micromodel estimates. Public sources included the International Monetary Fund, the Organisation for Economic Co-operation and Development, the American Bankers Association, Eurofinas, and the World Bank. National government sources included the Consumer Financial Protection Bureau, the Federal Deposit Insurance Corporation, the Federal Trade Commission, and the Small Business Administration in the United States; the Bank of England and the Financial Conduct Authority in the United Kingdom; the European Banking Authority and the Bundesbank in Europe; and the Reserve Bank of India and the Unique Identification Authority of India.

We then leveraged McKinsey Global Institute's Global Growth Model (an econometric model spanning more than 100 countries) to translate the potential micro-impacts of each use case to their GDP impact in 2030, taking dynamic macroeconomic feedback loops into account. The GDP impact was modeled using levers such as cost saving, the value of time saved, increase in physical capital, reduced cost of fraud, and higher labor productivity.

The potential economic gain from financial data sharing that we size for each economic region is sensitive to factors like the current structure of the economy, levels of financial product access, fraud, service costs, and access to digital infrastructure including the internet and various forms of digital ID. In determining the potential economic value, our model assumes a high level of open data enablement and adoption of use cases by 2030. In practice, the economy's trajectory and eventual level of data-sharing capabilities and practices would influence how much of the potential is captured. Accordingly, we compare the current level of data sharing enabled in each of our focus regions with the potential economic opportunity. The aim is to stimulate a discussion about what is needed to help capture all the value at stake.

While this research focuses on four economic regions, our approach to assessing the potential economic value of open data for finance, and the path to capturing it, is relevant more broadly, although the value for other economies would depend on industry characteristics in each country.

The topic of open data is inextricably linked with data privacy, data protection, and user consent, as well as with digital ID and authentication systems. We touch on all these aspects in our research. Other important questions related to open financial data go beyond the scope of our research, however. These include the shape of the data value chain, the roles and economics of its players, how data infrastructure can be designed and managed, and corporate governance for data.

For consumers, open data can improve access to financial services, user convenience, and product and service options

Consumers see value in the products and services that data sharing can enable. A McKinsey survey of 3,000 individual consumers and MSMEs in the United Kingdom showed that the willingness to share data doubles when customers find an appealing product or service enabled by it or understand the value it might bring them, for example an application that helps track and improve credit scores or a marketplace through which individuals can easily switch between different savings accounts based on interest rates.[9]

This represents an opportunity for increased revenue and growth for financial services providers, and both startups and more traditional players are responding with products to meet and shape demand. Key potential benefits from open financial data for consumers include the following:

Increased access to financial services. Data sharing enables customers to buy and use financial services to which they might not otherwise have access. Where limited data from traditional documentary sources may disqualify consumers from accessing loans, for example, open financial data can help assess the creditworthiness of borrowers by sourcing rent, phone, utility, and other bills. Individuals and SMEs with thin files or no formal records can thereby gain access to formal credit, often for the first time. Improved access to alternative data and customer data residing with other credit providers can also enable individual and SME borrowers to access loans that consolidate their debt across multiple institutions and credit lines, at lower average interest rates.

The economic benefits of this improved access can be significant. For example, an Experian study showed that including utility data allowed 20 percent of "thin-file" credit customers with scant documentation to support their credit application to become "thick-file" customers, raising the thick-file segment of total credit applicants from 55 percent to 64 percent.[10] In a related use case, FICO research shows that, on average, each additional data type beyond traditional sources (for example, data publicly shared on social media or utility data) adds 5 percent more predictive power to credit underwriting. By combining traditional and alternative data, FICO was able to create a predictive underwriting model for a personal loan originator that was significantly more accurate than a model using only traditional data.[11] Scaling such gains to an economy-wide level, we find that increased access to credit using alternative data could raise the economy's credit-to-GDP ratio by 20 basis points in the United States and the European Union. In India, this would be as much as 130 basis points—the equivalent of about $80 billion to $90 billion in GDP by 2030.

Greater user convenience. Data sharing saves time for customers in their interactions with financial services providers, most importantly during product purchase and exit. MSMEs can provide documentation faster during customer onboarding, for example. Open access to data on available mortgage products, with applications automatically prefilled, allows consumers to apply for loans without needing to use mortgage brokers. This not only eases the process but enables customers to benefit from the best rates. In the United Kingdom, which introduced its Open Banking system in 2018, startups use open-banking data to enable quick and easy mortgage applications to all participating mortgage providers for free, unlike traditional mortgage brokers who charge arrangement fees. These startups also notify consumers when it might make sense for them to refinance and then facilitate the process using available data.

[9] "Financial services unchained: The ongoing rise of open banking," forthcoming on McKinsey.com.
[10] *Let there be light: The impact of positive energy-utility reporting on consumers,* Experian.
[11] *FICO blog*, "Using alternative data in credit risk modelling," August 29, 2017.

Improved product options. Open financial data can broaden and improve the range of product options available to customers, saving them money. This is especially relevant during the understanding and use phases of the financial services life cycle, when customers are selecting products for the first time or switching to new products. For example, an open-data ecosystem makes it easier to switch accounts from one institution to another, helping retail and MSME customers achieve the best yield. Some cash management startups notify users when the rates they receive are less advantageous than the best ones on the market and allow for quick transfer of funds. This may help consumers narrow the difference between the yield they realize and the best yield available. For example, in the United Kingdom in 2020, this represented a difference of 48 basis points.[12]

Four major mechanisms of open-data systems benefit financial institutions

Fintech companies and nontraditional financial services players have taken the lead in financial data sharing in many countries, seeing it as a way to achieve a breakthrough in serving customer needs with new financial products and services offerings. Meanwhile, incumbent players are looking to respond to the open banking trend. Both types of providers could see economic benefit from data sharing. Traditional players in particular could reduce costs by being able to streamline and automate various operational processes that currently need data drawn from multiple, disjointed sources.[13] These benefits flow mainly to financial institutions, although they may also be passed on to consumers through more competitive pricing and help improve customer experience. Four such mechanisms can create economic value across the life cycle:

Increased operational efficiency. Since most data are still found in physical documents or disparate digitized sources, open financial data could cut costs by providing verified data digitally and make it easier to adopt automation technologies, with the associated efficiency boost. All of this can also improve experience for customers by promoting faster and more transparent interactions with providers.

For instance, it is possible to entirely replace manual processes with automated know-your-customer (KYC) processes for retail and MSME customers at much lower cost. In mortgage underwriting, sharing borrowers' data, which for now are siloed and manually aggregated, allows standard mortgages to go through automated underwriting, reducing operating costs and speeding up the lending timeline. Customer profiles built on open data provided by other institutions can increase use of predictive analytics and AI in digital and interactive voice response–based call center operations. When customers switch financial services providers, open data reduces the time spent between account dormancy and closure by enabling the automation of notifications to dependent institutions.

Financial data sharing also helps avoid multiple manual data handoffs that lead to errors, rework, and less efficient outcomes. An open-data ecosystem acts as a "source of truth," with data entries and adjustments occurring at only one time, producing better quality and cleaner data. This significantly reduces the costs associated with remediating bad customer relationship management data, currently estimated at 20 percent of a typical financial institution's income.[14] On the customer acquisition side, an open-data ecosystem can help financial institutions engage in more targeted, data-driven marketing, leading to improvements in conversion rates and more efficient spending on marketing. Credit costs could be reduced by identifying more targeted credit recovery strategies using open financial data.

[12] Derin Clark, *Average savings rates at their lowest levels on record*, Moneyfacts.co.uk, August 17, 2020; George Nixon, *Savings wipe: Just two easy-access accounts now beat inflation and those with big banks could be losing £69 a year*, Thisismoney.co.uk, November 19, 2020.

[13] Open data could potentially improve credit risk assessment and lower credit risk costs and accuracy of risk-weighted asset assessment. We have not sought to quantify this in our research as it is highly dependent on portfolio composition and pricing strategy.

[14] *Only 50% of respondents believe their CRM/ERP data is clean*, Experian, edq.com.

One example of the time and cost savings that are possible through data sharing comes from India. There the use of the national digital identification system, Aadhaar, for KYC verification for retail consumers reportedly reduced costs for financial institutions from about $5 per customer to approximately $0.70.[15] Another example comes from Estonia, where a study found that X-Road, the internet-based data exchange layer of the country, serviced queries across a variety of applications of open data yielding a total of 2.8 million hours in time savings in 2014, comparable to 3,225 full-time-equivalent employees working for one whole year.[16]

Better fraud prediction. This is relevant across the full financial services life cycle and can lead to a significant reduction in costs for institutions, as well as improved customer experience. The Association of Certified Fraud Examiners estimates total fraud (including but also extending beyond financial services) at more than $4.5 trillion annually, or the equivalent of about 5 percent of global corporate revenue.[17] Fraud in financial services takes multiple guises, including synthetic and traditional ID fraud, payments fraud, and credit application fraud. Real-time access to a full set of customer data can support advanced techniques to identify and reduce costs related to these and other types of fraud. Sharing fraud-specific information and other kinds of data provides more evidence and clues with which to flag suspicious activity. For example, synthetic IDs can be spotted by leveraging data trails of people from dozens of different data systems, both physical and digital, and scoring them for depth and consistency using machine learning tools. Low scores would imply a higher probability of synthetic ID fraud.[18] Real-time data sharing would help institutions build out their predictive modeling of fraud and catch cases earlier.

In one example in the United Kingdom, Cifas, a not-for-profit fraud database and fraud prevention organization that facilitates data sharing among its members, says those members reported more than 350,000 cases of fraud in 2019, preventing fraud totaling £1.5 billion. Cifas members say the fraud database identifies and prevents 91 percent of external fraud.[19]

Improved workforce allocation. This is particularly important during the period of product design and marketing and during servicing and monitoring. Companies can use open data to better allocate and target their workforce, assigning staff to the highest-value activities. For example, companies can draw on external data sources to help collections staff better focus their calls on high-risk customers, reduce the time spent monitoring low-risk customers, and ultimately recover more debt. During the pandemic, one major bank in the United Kingdom that used open-banking transaction data rather than traditional credit bureau data to better understand its customers' credit risk reduced underwriting losses by 40 percent. Banks using such models can better allocate their collections analysts toward truly high-risk customers rather than customers inaccurately perceived as being high risk, for example those with limited credit histories. Similarly, pulling reliable data on customers, including that housed in external sources, such as social data, increases the productivity of product researchers and designers by reducing the time they spend in sourcing data from vendors.

Reduced friction in data intermediation. This value-creation mechanism is most relevant to financial institutions before they have direct knowledge of a prospective customer, such as in lead generation or loan origination, and so look to acquire and use data from third-party providers. The missing details could range from basic identification to behavioral information. Siloed information can lead to data intermediation at several stages of the customer journey, and the process can be cumbersome and costly. Open-data systems enable direct access to data by using APIs for intermediation, which reduces friction. Data sharing reduces or eliminates the costs financial institutions incur in sourcing data from third-party data providers and aggregators for the purposes of lead generation and customer targeting as well

[15] "Use of Aadhaar for KYC authentication will cut costs," *Hindu Business Line*, January 20, 2018.
[16] Kristjan Vassil, *World Development Report 2016 background paper: Estonian e-government ecosystem: Foundation, applications, outcomes*, World Bank, June 2015.
[17] *Report to the nations: 2020 global study on occupational fraud and abuse*, Association of Certified Fraud Examiners, 2020.
[18] For further details see Bryan Richardson and Derek Waldron, "Fighting back against synthetic identity fraud," McKinsey.com, January 2, 2019.
[19] *Annual report and financial statements for the year ended 2019*, Cifas, December 2020, cifas.org.uk.

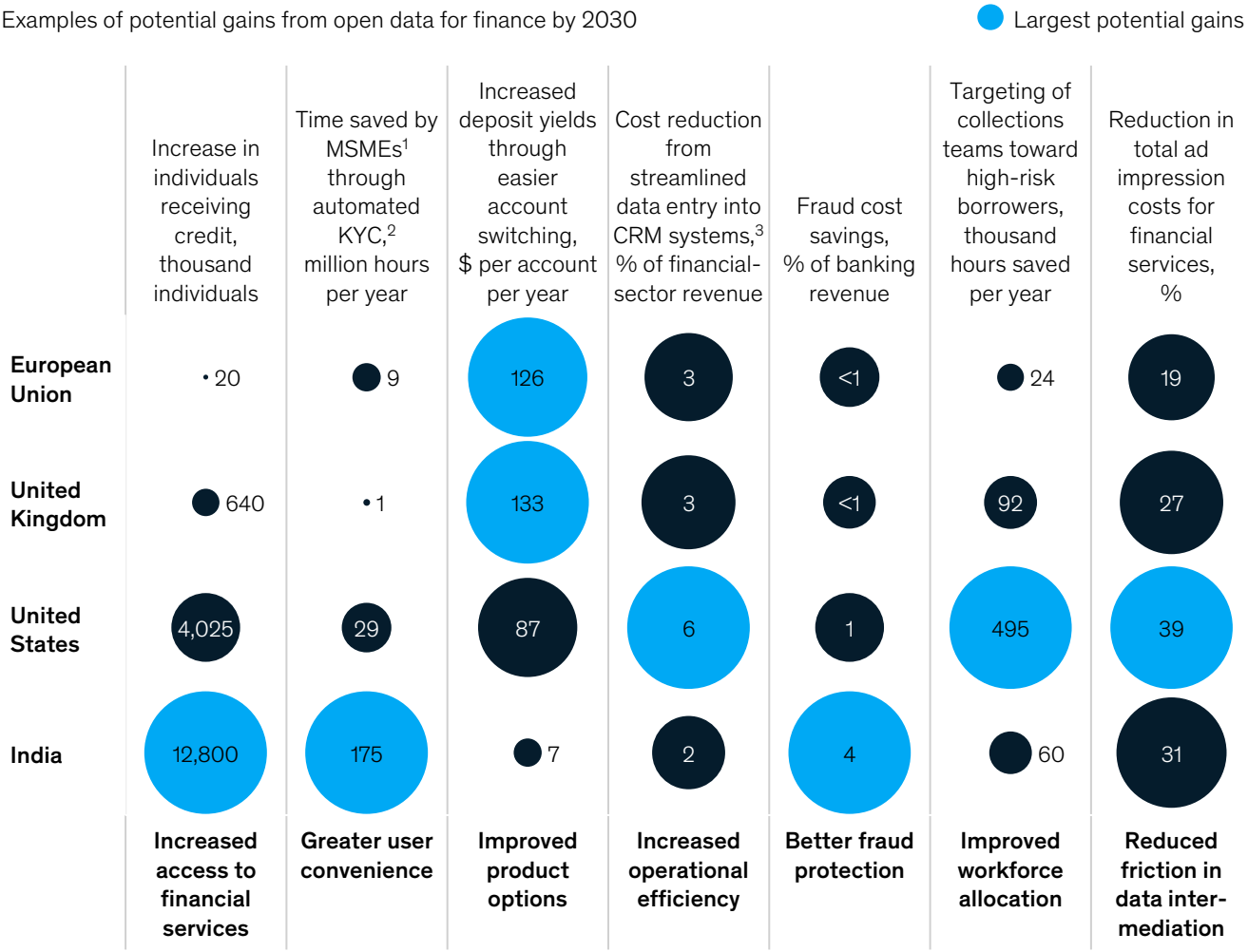
as mortgage underwriting. In the United States, for example, where nearly half of all mortgage providers rely on third-party data for mortgage origination, such data can cost as much as $80 per mortgage application. The data provided typically entails consolidated customer credit data, KYC data, and property valuation data. With open data for finance, much of these data are becoming more publicly available. For example, Zillow, an online real estate marketing company in Seattle, provides a "Zestimate" using public information to provide a property's approximate appraisal value. Combining public data sources using open-data APIs could significantly reduce—and in some cases eliminate entirely—the cost of third-party mortgage origination.

Exhibit 2 demonstrates the potential economic gains from an example use case linked to each of the seven mechanisms for creating value.

Exhibit 2

Open financial data ecosystems can scale to create significant potential economic gains.

Examples of potential gains from open data for finance by 2030



1. Micro-, small, and medium-size enterprises, defined by the International Finance Corporation as enterprises with sales and/or assets less than $15 million and/or with fewer than 300 employees.
2. Know your customer data.
3. Customer relationship management.

Note: Estimated potential value assumes data standardization and breadth of sharing as well as robust data privacy and consent frameworks, widespread access to digital financial infrastructure including digital ID, and regulations enabling strong product innovation. Bubbles scaled by column.

Source: Experian; Glassdoor; IFC; IMF; OECD; Refinitiv KYC survey; World Bank (Global Findex); national sources and databases (see methodology); Zillow; McKinsey Global Institute analysis

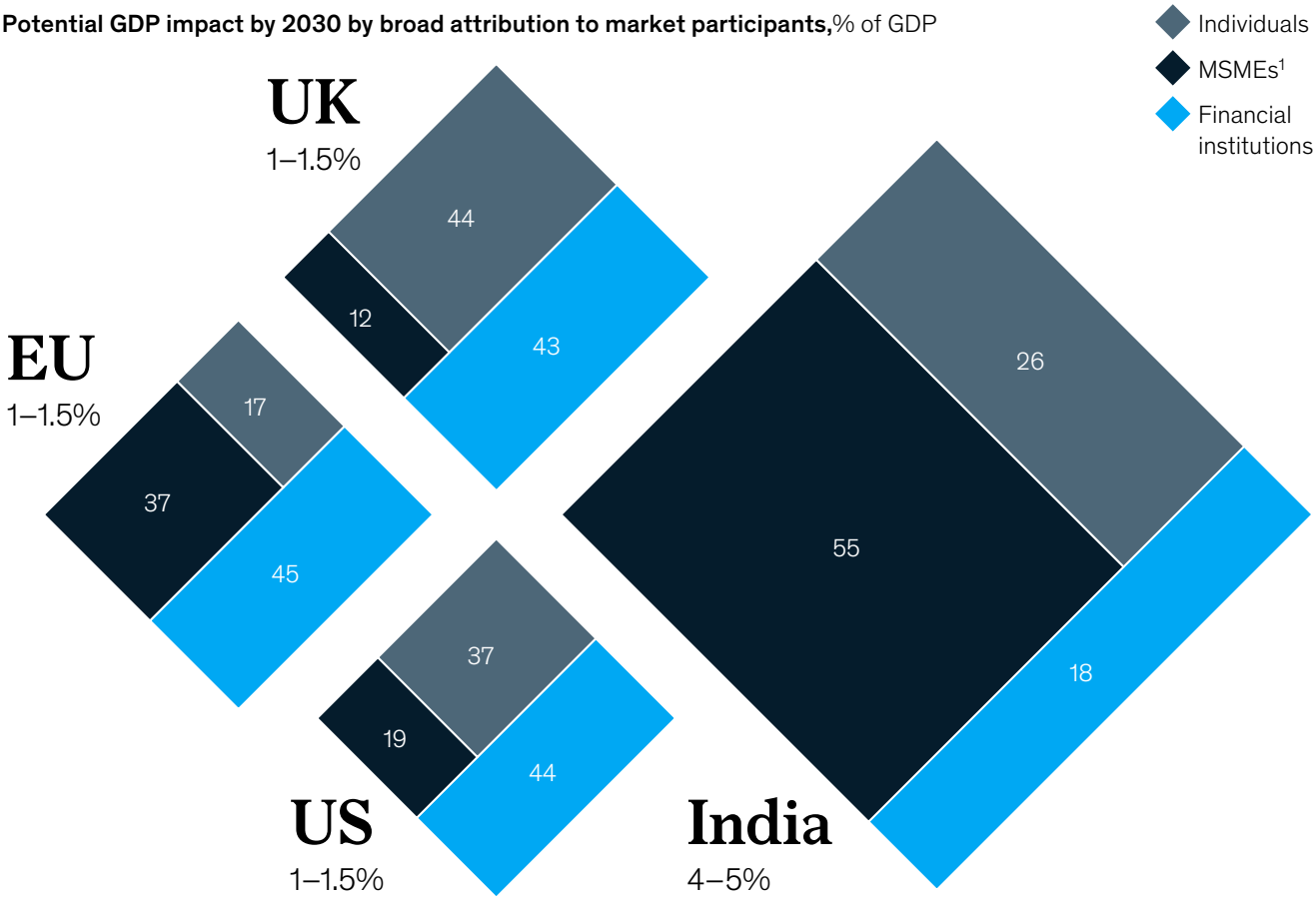
2. Sizing the value at stake from open financial data

Aggregating the potential economic impact across our 24 use cases, we find significant value at stake overall and for all market participants (Exhibit 3). The total potential GDP impact from open financial data in 2030 is highest for India, at 4 to 5 percent of GDP, while we estimate the impact for the European Union, the United Kingdom, and the United States to be between 1 and 1.5 percent of GDP. The differences are the result of several factors, typically structural features of economies.

Exhibit 3

The potential GDP impact of open financial data and the share accruing to different market participants varies by region.

Potential GDP impact by 2030 by broad attribution to market participants, % of GDP



1. Micro-, small, and medium-size enterprises, defined by the International Finance Corporation as enterprises with sales and/or assets less than $15 million and/or with fewer than 300 employees.

Note: The GDP impact for each economy in 2030 is estimated using MGI's Global Growth Model (GGM), a dynamic general equilibrium economic model spanning 100 countries; GGM estimates the economy-wide potential GDP impact by 2030 of 24 banking and payments use cases, each of which individually generates economic value. The attribution of potential economic value across market participants is estimated based on the potential impact of the 24 use cases on a standalone basis, without taking into account dynamic macroeconomic feedback loops that determine GDP impact. Figures may not sum to 100% because of rounding.

Source: McKinsey Global Institute analysis

In the European Union, the United Kingdom, and the United States, the share of economic value from decisioning and onboarding is lower than that in India, while the share from servicing and monitoring is higher, ranging between 35 percent and 45 percent of the total. Many more small-ticket individuals and MSMEs already have credit access in advanced economies, reducing the share of value in upstream customer acquisition and moving it to the account servicing and monitoring stage. In the United States, the value at stake in relation to decisioning and onboarding is higher than in the European Union because of the size of the credit gap for the country's many MSMEs.

The main difference between the European Union and the United Kingdom is in the breakdown of potential value associated with the servicing and monitoring portion of the life cycle. In the United Kingdom, we estimate nearly half of this value could flow from improved product options, driven by the potential for MSMEs to increase their deposit yields through easier account switching. In the European Union, where the average MSME is more than half as small and total MSME savings are lower, the relative potential value coming from such account switching is smaller.

Notably, emerging economies tend to have lower levels of financial access and less financial depth, which means the lift in value creation they could achieve with open data is large. India has significant unmet need for retail and MSME credit.[20] It thus has higher economic growth potential for every unit of physical capital added as open financial data improves credit access. MSMEs also benefit from time saved in opening accounts; this would translate into higher GDP, assuming the time saved was deployed in market-based economic activities rather than, say, leisure. In India, the time saving accruing to MSMEs could make them the largest beneficiary segment, with more than 60 percent of the potential economic value across our quantified use cases accruing to them.

In both the United Kingdom and the United States, by comparison, while we estimate that individuals would capture the largest share of value, financial institutions have sizable value at stake. In the European Union, financial institutions would gain the largest share of value, or up to nearly 45 percent of the total economic value across our quantified use cases. In tangible terms, this could translate into economic value of $90 billion to $100 billion for financial institutions, $80 billion to $90 billion for individuals, and between $40 billion and $50 billion for MSMEs in 2030 in the United States, for example.

Comparing across the advanced economies, the relative benefit to individual consumers compared to MSMEs is meaningfully smaller in the European Union than in the United Kingdom and the United States. This reflects industry practice in the European Union, in which credit decisions for individuals are made largely on the basis of delinquencies rather than credit information, limiting the role that data sharing can play in increasing credit access.

The ways in which open-data ecosystems create value across the financial services life cycle vary by economy

The European Union, the United Kingdom, the United States, and India share some similarities in the way open financial data creates value. Exhibit 4 shows the share of potential economic value for each region (measured on a standalone basis for each use case) mapped to each major step in the financial services life cycle and for each of the seven mechanisms for value creation. In all four regions, individuals and MSMEs stand to gain most from the way open financial data can transform both product design and decisioning and onboarding, boosting ease of access to financial services. For institutions, the potential value is relatively larger during servicing and monitoring as well as in account termination. In these phases, institutions are better able to undertake existing processes more efficiently, in a more targeted way, and with less fraud.

Yet there are also some significant differences across economies relating to the seven mechanisms for value creation we outlined in the previous section.

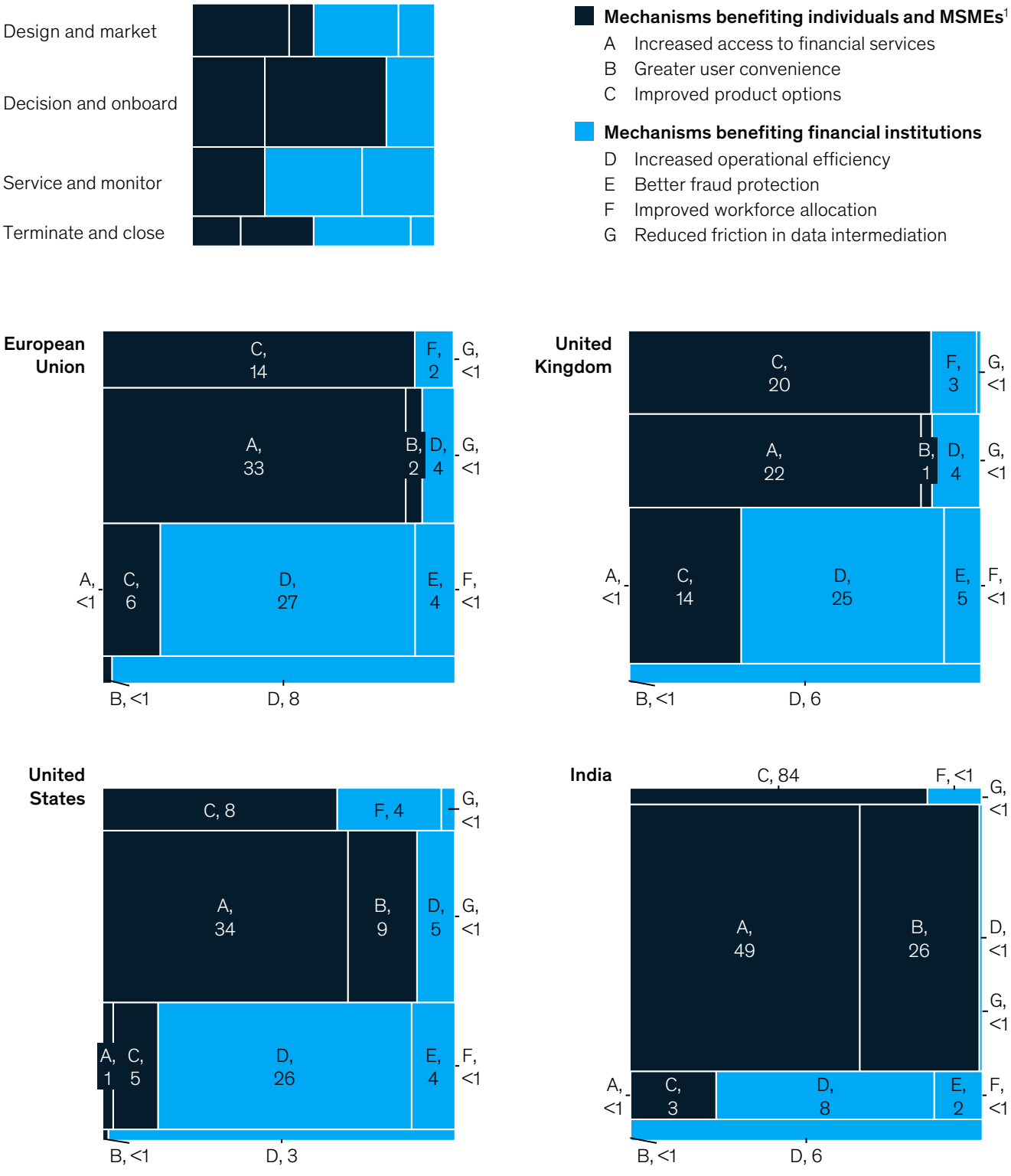
India stands out, with much greater share of value, about 75 percent of total economic value, coming from the decisioning and onboarding component of the life cycle, particularly linked to increased access to financial services. This is due to the large potential of opening access to credit to currently excluded individuals and MSMEs. Greater user convenience also contributes significantly to the potential value at stake in India; this comes from the potential for simplified onboarding processes for new customers, particularly easing the burden for MSMEs of having to produce paper documentation, saving meaningful amounts of time.

[20] *MSME finance gap: Assessment of the shortfalls and opportunities in financing micro, small, and medium enterprises in emerging markets*, International Finance Corporation, 2017.

Exhibit 4

The mix of potential economic value of open financial data varies across economies.

Attribution of potential economic value by mechanism of value creation and stage in the customer life cycle,
% of economic value for the region



1. Micro-, small, and medium-size enterprises, defined by the International Finance Corporation as enterprises with sales and/or assets less than $15 million and/or with fewer than 300 employees.

Note: The composition of potential economic value for each economy is estimated based on the potential impact of 24 banking and payments use cases in terms of time savings, cost savings, enhanced physical capital, reduced cost of data intermediation, fraud cost reduction, and higher labor productivity. These estimates of economic value represent potential impact on a standalone basis; they do not take into account dynamic macroeconomic feedback loops that determine GDP impact. Figures may not sum to 100% because of rounding.

Source: American Bankers Association; CreditKarma; Eurofinas; Experian; FICO; Glassdoor; GPFI; IFC; IMF; LendingTree; national sources and databases (see methodology); OECD; Refinitiv KYC survey; World Bank; Zillow; McKinsey Global Institute analysis

3. The need for data standardization and breadth of sharing

The value creation mechanisms we outlined require varying levels of data standardization and breadth of data sharing for their full potential to be captured.

A combination of high standardization and broad data sharing ensures accessibility to data for all interested parties, across a wide range of participants and use cases. This incentivizes adoption, distribution, and participation in the data ecosystem. The value of open data that is captured and the ecosystem participants to whom that value would accrue depend to a large extent on these two axes of standardization and openness.

Data standardization refers to the extent to which standardized mechanisms exist for sharing data and the associated cost of access. In some use cases, data sharing occurs only through ad-hoc means. For example, consumers wanting to receive automated access to competitive mortgages need to provide the same specific mortgage application data to multiple providers during a mortgage search process. To operate at scale, other data-sharing use cases require data to be sourced easily through standardized APIs at minimal cost. For example, for customers to be able to switch accounts easily to increase yield on their deposits, consumer data needs to seamlessly move between providers so that accounts can be closed and opened fast and automatically. An example of an ecosystem with high data standardization is the United Kingdom, where large banks are required to share transaction-level consumer financial data, at no cost, with licensed third-party service providers (including other banks), all via highly standardized and regulated "Open Banking" APIs.[21]

Breadth refers to how broadly data are shared and the mechanisms in place that drive the data sharing. Some use cases work when individuals can request specific data to be shared on an ad-hoc basis. For example, consumers can benefit from faster mortgage closure when they are able to grant their prospective lender on-off access to the required data. To operate at scale, other use cases require data sharing over time across a wide range of types of financial data, albeit with consumer consent. For example, financial institutions require continuous access to a range of consumer data to improve and personalize products. An example of an ecosystem with broad data sharing is India, where banks must share all consumer data including personal nonfinancial and financial data at the request of consumers via private but highly standardized APIs developed on a publicly built technology ecosystem known as IndiaStack.[22] Consumers can choose to share their data with digital nonbank lenders via an app to secure loans.

From our research, we see that consumers (both individuals and SMEs) require moderate levels ofstandardization and breadth of data sharing to reap benefits of open data coming from increased access to financial services, greater convenience for users, and improved product options (Exhibit 5). Institutions, by contrast, can only access full benefits when the levels of standardization and breadth are higher.

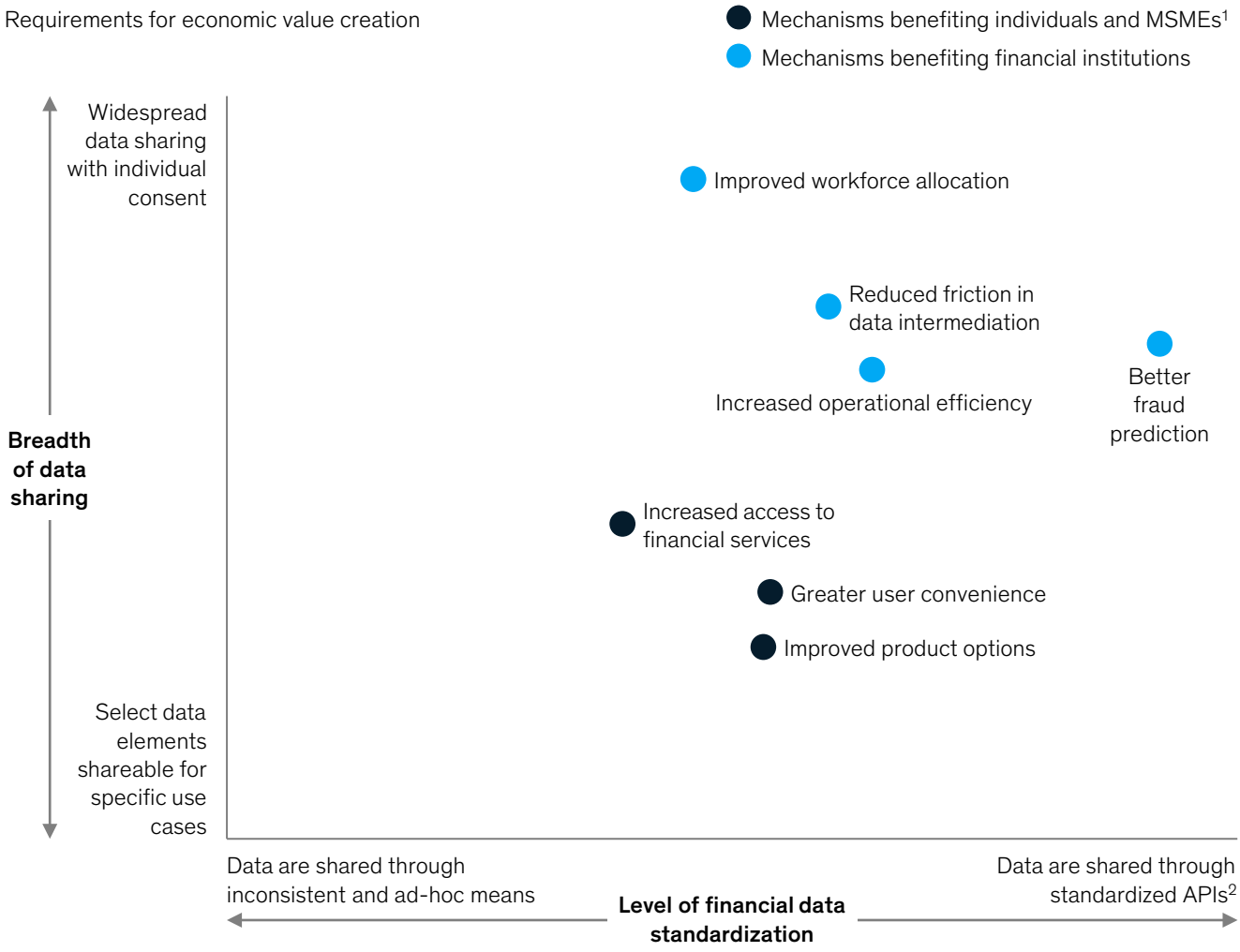
Significantly greater breadth of data shared could enable full value capture through improved operational efficiency, improved workforce allocation, and lower friction in data intermediation, while greater data standardization could incrementally yield the full value of better fraud prediction.

[21] See openbanking.org.uk.
[22] For details about IndiaStack, see *Digital India: Technology to transform a connected nation*, McKinsey Global Institute, March 2019.

Exhibit 5

Capturing the full value of open financial data requires a high degree of data standardization and breadth of sharing.

Requirements for economic value creation



1. Micro-, small, and medium-size enterprises, defined by the International Finance Corporation as enterprises with sales and/or assets less than $15 million and/or with fewer than 300 employees.
2. Application programming interfaces; used here to include similar mechanisms of data sharing.

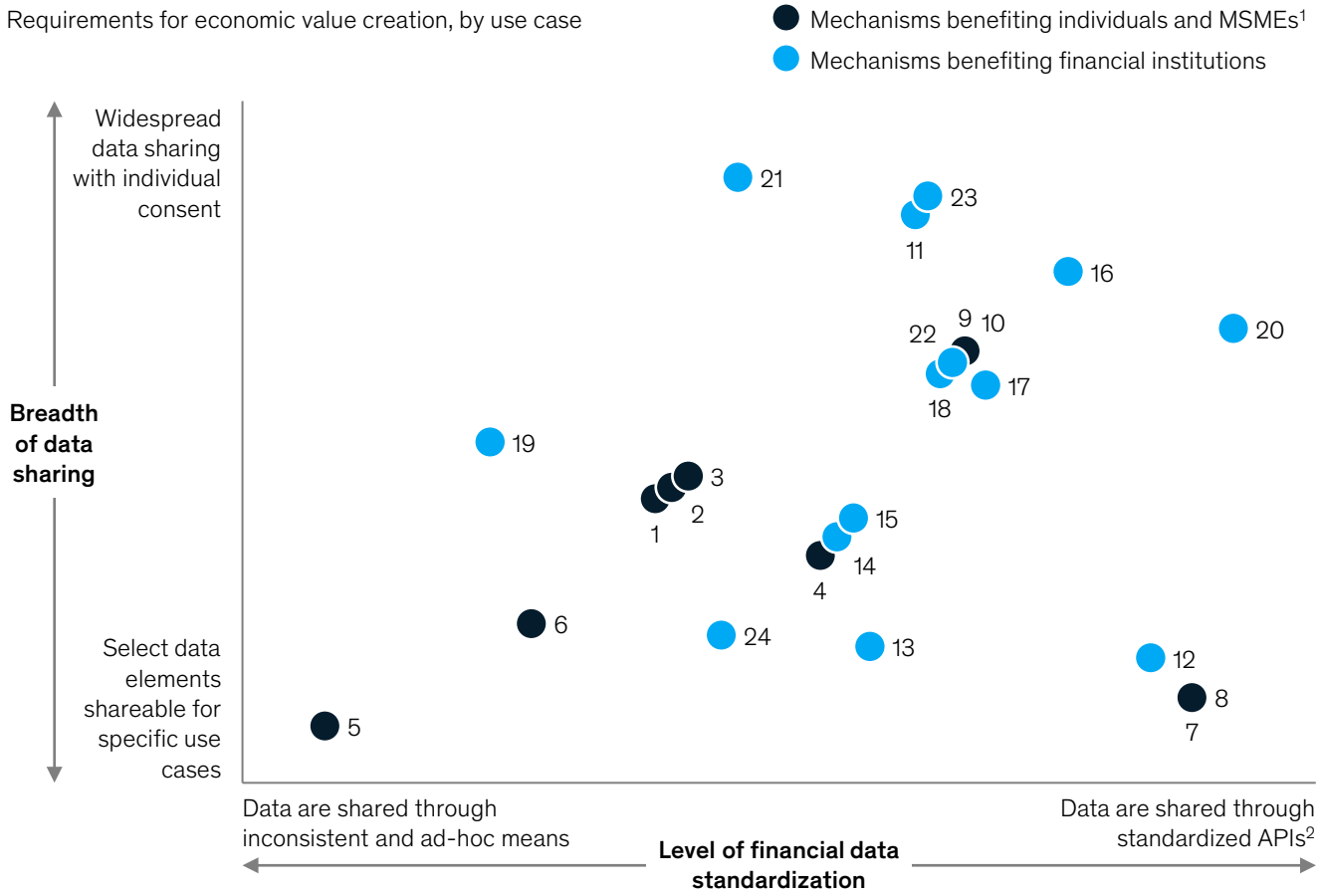
Note: The location of each bubble on the grid represents the average level of data standardization and breadth of sharing necessary for value capture for use cases within that mechanism; bubbles are not scaled to the amount of potential economic value they could create.

Source: McKinsey Global Institute analysis

Exhibit 5 (continued)

Capturing the full value of open financial data requires a high degree of data standardization and breadth of sharing (continued).

Requirements for economic value creation, by use case



Increased access to financial services

- 1. Newly allowing access to retail credit through alternative credit underwriting (individuals)
- 2. Newly allowing access to retail credit through alternative credit underwriting (MSMEs)
- 3. Debt consolidation at reduced interest rates (individuals)

Greater user convenience

- 4. Simplified application and onboarding process for small businesses
- 5. Automated data portability between accounts

Improved product options

- 6. Access to competitive mortgages traditionally facilitated by brokers
- 7. Increased deposit yields through easier account switching (individuals)
- 8. Increased deposit yields through easier account switching (MSMEs)
- 9. Improved customer segmentation to reduce switching costs (individuals)
- 10. Improved customer segmentation to reduce switching costs (MSMEs)

Increased operational efficiency

- 11. Marketing efficiency through data-driven targeting
- 12. Automated underwriting of standard mortgages
- 13. Data availability to drive faster mortgage closure
- 14. Automated KYC[3] (individuals)
- 15. Automated KYC[3] (MSMEs)
- 16. Streamlined data entry into CRM[4] systems
- 17. Predictive data-driven digital/IVR[5] call center operations
- 18. Data-driven reductions in the cost of credit recovery
- 19. Automated notification of events that should trigger account closure

Better fraud protection

- 20. Fraud reduction through timely and comprehensive data

Improved workforce allocation

- 21. Access to market and customer data for product design
- 22. Data-driven reorientation of collections teams toward higher-risk borrowers

Less friction in data intermediation

- 23. Direct access to lead-generation data traditionally brokered by 3rd-party providers
- 24. Direct access to mortgage data traditionally brokered by 3rd-party providers

1. Micro-, small, and medium-size enterprises, defined by the International Finance Corporation as enterprises with sales and/or assets less than $15 million and/or with fewer than 300 employees.
2. Application programming interfaces; used here to include similar mechanisms of data sharing.
3. Know your customer data.
4. Customer relationship management.
5. Interactive voice response.

Note: The location of each bubble on the grid represents the average level of data standardization and breadth of data sharing necessary for value capture for use cases within that particular mechanism; bubbles are not scaled to the amount of potential economic value they could create.

Source: McKinsey Global Institute analysis

Economies vary in their current levels of data standardization and breadth of data shared

Our research finds notable differences in the manner, speed, and extent of open-data deployment by country. Each economic region we studied has different levels of open financial data enablement in place along the axes of data standardization and breadth of data shared, influencing the pattern of data flows between market participants. The type of data-sharing ecosystem used in an economy depends on multiple factors including local market conditions, the robustness of existing digital financial infrastructure, and regulation, including consumer protection laws and mechanisms.

In the four economic regions on which we focus, we find significant differences in the degree of standardization and breadth of data shared (Exhibit 6).

In the European Union and the United Kingdom, for example, we find a high level of standardization combined with relatively less breadth of data shared. In such data ecosystems, a limited subset of financial data is available via highly standardized and regulated APIs. The data are accessible and usable for third-party service providers such as financial technology firms and other banks. For example, in the United Kingdom, APIs are mandatory for most account and direct debit data, such as address, beneficiary, and payment timestamp. Access to data outside of the subset is more restricted and at the discretion of each individual financial service provider.

In the European Union, the second payment services directive stipulates that data specific to payments be shared via highly standardized APIs. However, separate data privacy regulations restrict the sharing, harvesting, and use of other financial and nonfinancial data, including, for instance, screen scraping, a prevalent form of data gathering in the United States.[23]

In the United States, by comparison, our research finds that a broad range of data is shared but that data standardization is more limited. Financial data aggregators broker data flow between providers and users, with limited consumer control. This is primarily because of a lack of strong federal regulation regarding data privacy and a private-market approach to data sharing, in which institutions share data when competitively advantageous, for example to meet consumer demand or to monetize data. This entirely private-market approach has made private data aggregators the de facto standard setters for how data are shared, leading to more limited standardization in data sharing and relative opacity in the cost of data.

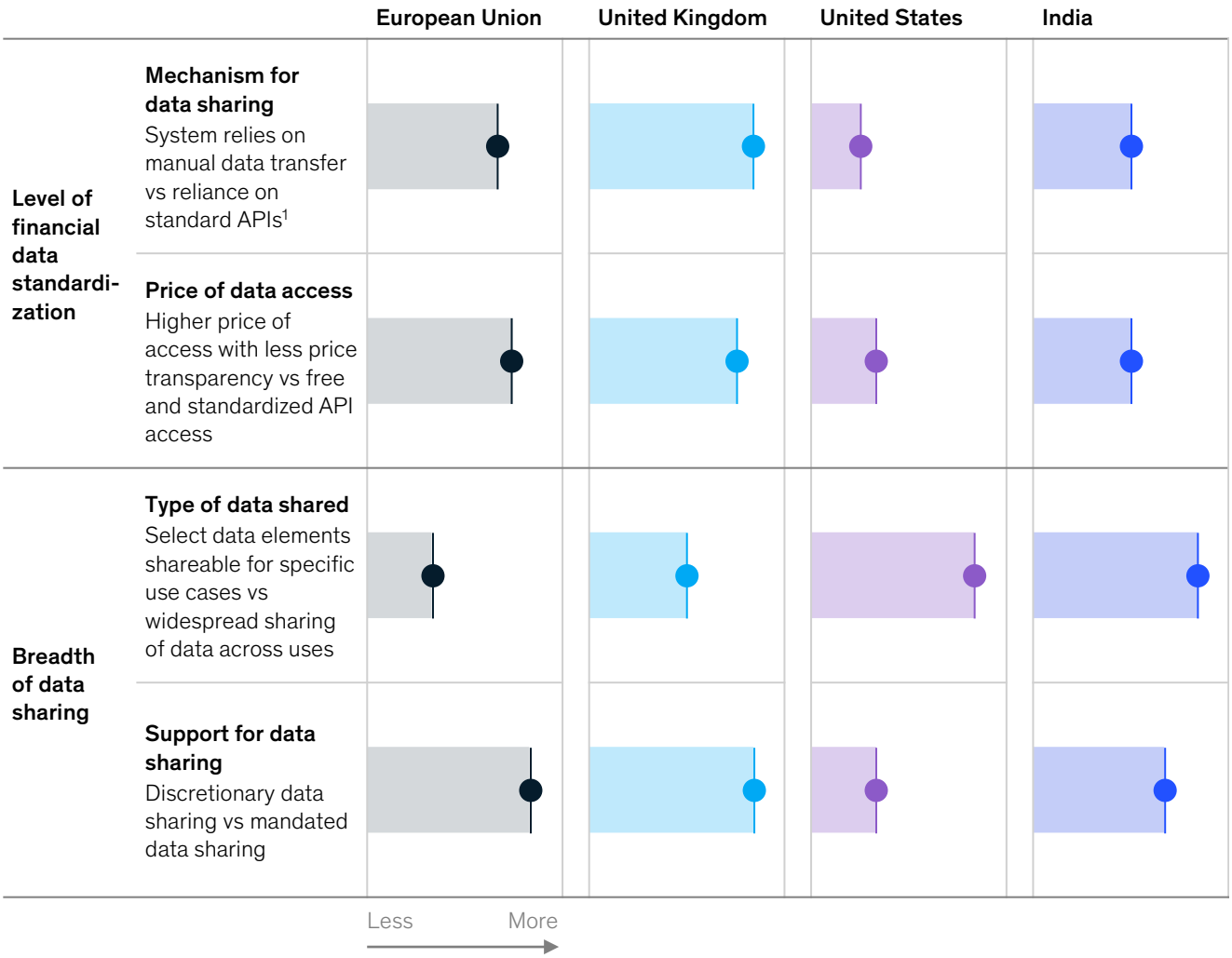
In India, a relatively broad range of data is shared and there is some degree of standardization. While all data are not shared via public APIs, as in the case in the United Kingdom and the European Union, private APIs in India are built by licensed data aggregators on IndiaStack.[24] This public technology stack includes layers for identity, authentication, payments, paperless data exchange, and user consent and has a relatively high level of openness and standardization. Data aggregators broker data sharing across institutions using this standard, and consumers can view the data and share directly with payment systems players.

[23] *Payments services directive (PSD2): Regulatory technical standards enabling consumers to benefit from safer and more innovative electronic payments*, European Commission, November 27, 2017.
[24] *Digital India: Technology to transform a connected nation*, McKinsey Global Institute, March 2019.

Exhibit 6

Economies have different levels of financial data sharing enablement in place.

Increasing enablement by impacting factor



1. Application programming interfaces; used here to include similar mechanisms of data sharing.

Note: Average scores for each economy along each dimension are set based on expert interviews, literature review, and a review of existing policies and practices on different aspects of data standardization and breadth of sharing.

Source: McKinsey Global Institute analysis

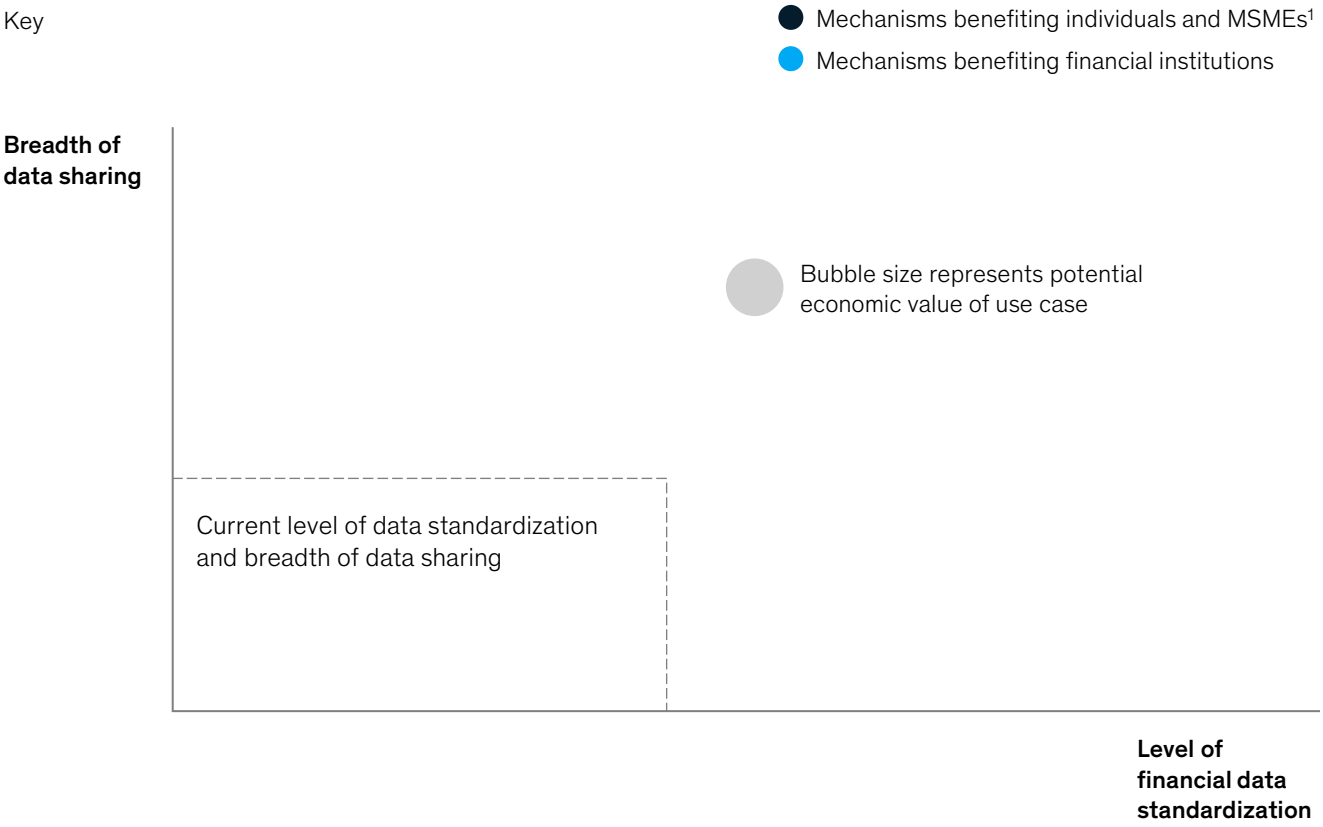
Only low to moderate levels of economic value are accessible today, given current levels of data standardization and breadth of data sharing

A country's level of data standardization and breadth of data sharing sets the potential value from open financial data that it might access today (Exhibit 7). In the European Union, the United Kingdom, and the United States, current data ecosystems leave much of the potential value at stake inaccessible. Both the United States and the European Union are, for now, in a position to capture only a small fraction of the potential value from open financial data—less than 10 percent, in our estimate. In the United States, the constraint is lack of standardization, while in the European Union it is limited breadth of data sharing. In the United Kingdom, somewhat more value is currently accessible—we estimate between 30 and 40 percent—but nonetheless limited by the breadth of data sharing. To capture more value from open financial data, these regions can consider raising both standardization and breadth of sharing to expand the realm of the possible, whether through regulatory or market forces.

India is better poised to capture value today. Its open data environment positions it to access between 60 and 70 percent of the potential value that open financial data could offer, provided other enablers are in place, as we describe in the next section.

Exhibit 7

Current levels of data standardization and breadth of data sharing would enable economies to create low to moderate levels of economic value.



Increased access to financial services

- 1. Newly allowing access to retail credit through alternative credit underwriting (individuals)
- 2. Newly allowing access to retail credit through alternative credit underwriting (MSMEs)
- 3. Debt consolidation at reduced interest rates (individuals)

Greater user convenience

- 4. Simplified application and onboarding process for small businesses
- 5. Automated data portability between accounts

Improved product options

- 6. Access to competitive mortgages traditionally facilitated by brokers
- 7. Increased deposit yields through easier account switching (individuals)
- 8. Increased deposit yields through easier account switching (MSMEs)
- 9. Improved customer segmentation to reduce switching costs (individuals)
- 10. Improved customer segmentation to reduce switching costs (MSMEs)

Increased operational efficiency

- 11. Marketing efficiency through data-driven targeting
- 12. Automated underwriting of standard mortgages
- 13. Data availability to drive faster mortgage closure
- 14. Automated KYC[2] (individuals)
- 15. Automated KYC[2] (MSMEs)
- 16. Streamlined data entry into CRM[3] systems
- 17. Predictive data-driven digital/IVR[4] call center operations
- 18. Data-driven reductions in the cost of credit recovery
- 19. Automated notification of events that should trigger account closure

Better fraud protection

- 20. Fraud reduction through timely and comprehensive data

Improved workforce allocation

- 21. Access to market and customer data for product design
- 22. Data-driven reorientation of collections teams toward higher-risk borrowers

Less friction in data intermediation

- 23. Direct access to lead-generation data traditionally brokered by 3rd-party providers
- 24. Direct access to mortgage data traditionally brokered by 3rd-party providers

1. Micro-, small, and medium-size enterprises, defined by the International Finance Corporation as enterprises with sales and/or assets less than $15 million and/or with fewer than 300 employees.
2. Know your customer data.
3. Customer relationship management.
4. Interactive voice response.

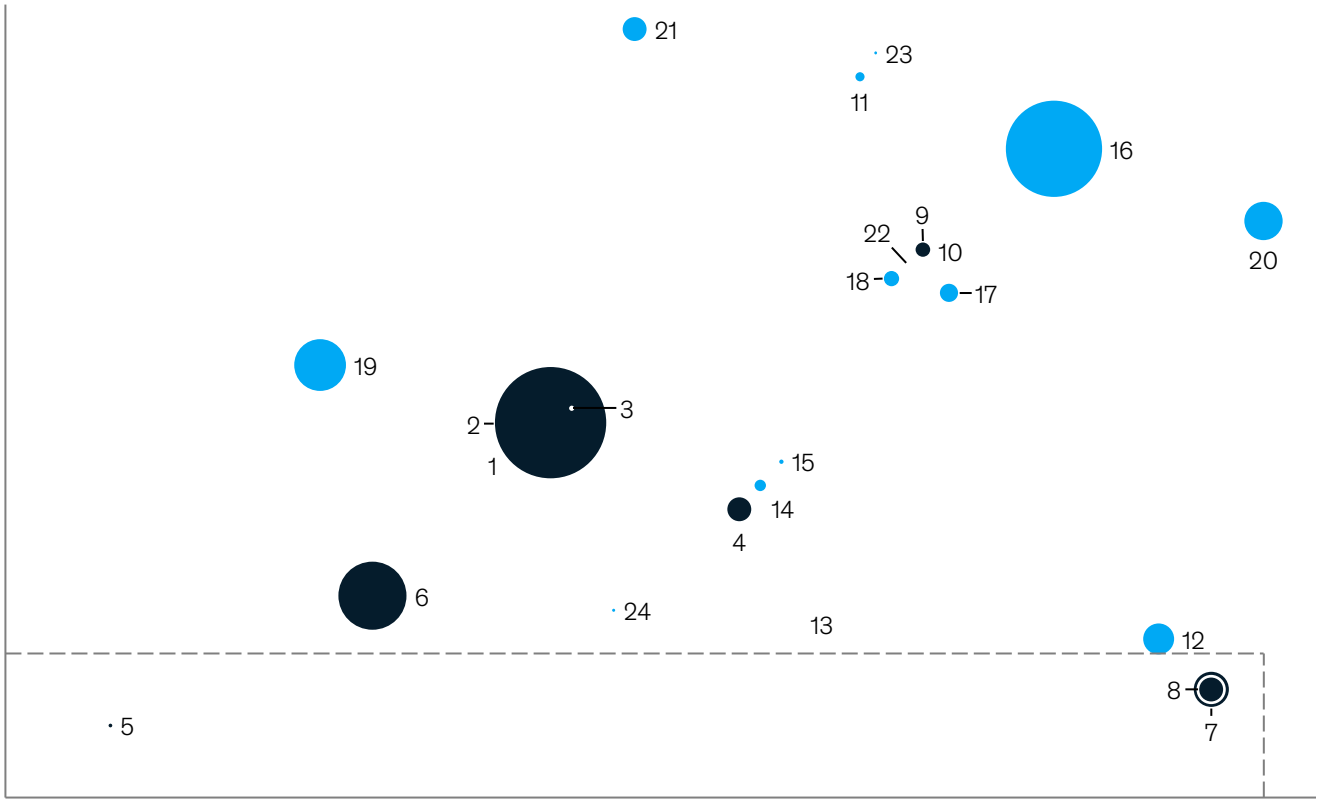
Note: The location of each bubble on the grid represents the average level of data standardization and breadth of data sharing necessary for value capture for that use case; bubbles are scaled to the amount of potential economic value they could create. The dashed-line area represents the current data standardization and breadth of data sharing in each economy; for each use case that falls within this area, the necessary levels of data standardization and breadth of data sharing to capture value are in place, although additional digital infrastructure and product innovation may be needed. Use cases outside the shaded area would require greater enablement of financial data sharing in order for the associated economic value to be realized.

Source: American Bankers Association; CreditKarma; Eurofinas; Experian; FICO; Glassdoor; GPFI; IFC; IMF; LendingTree; national sources and databases (see methodology); OECD; Refinitiv KYC survey; World Bank; Zillow; McKinsey Global Institute analysis

Exhibit 7 (continued)

Current levels of data standardization and breadth of data sharing would enable economies to create low to moderate levels of economic value (continued).

European Union



United Kingdom

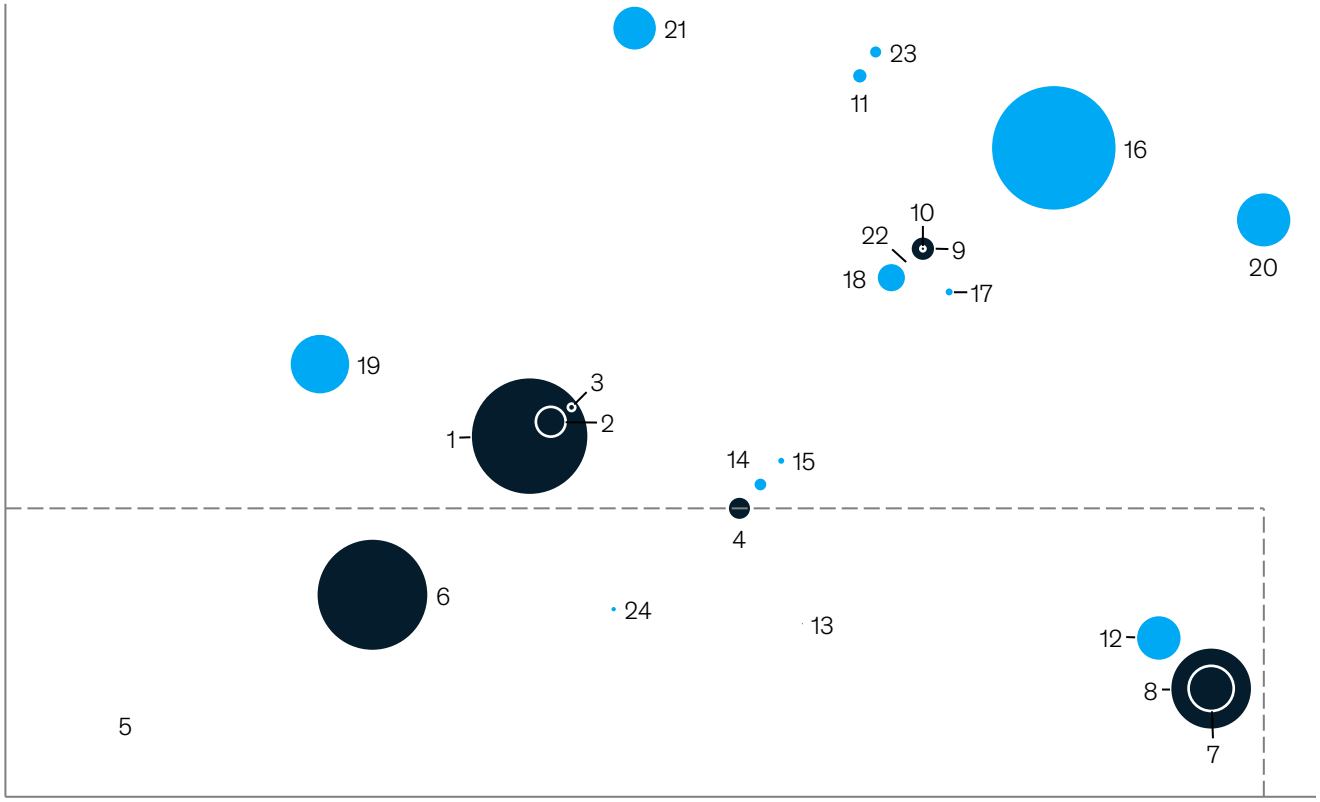
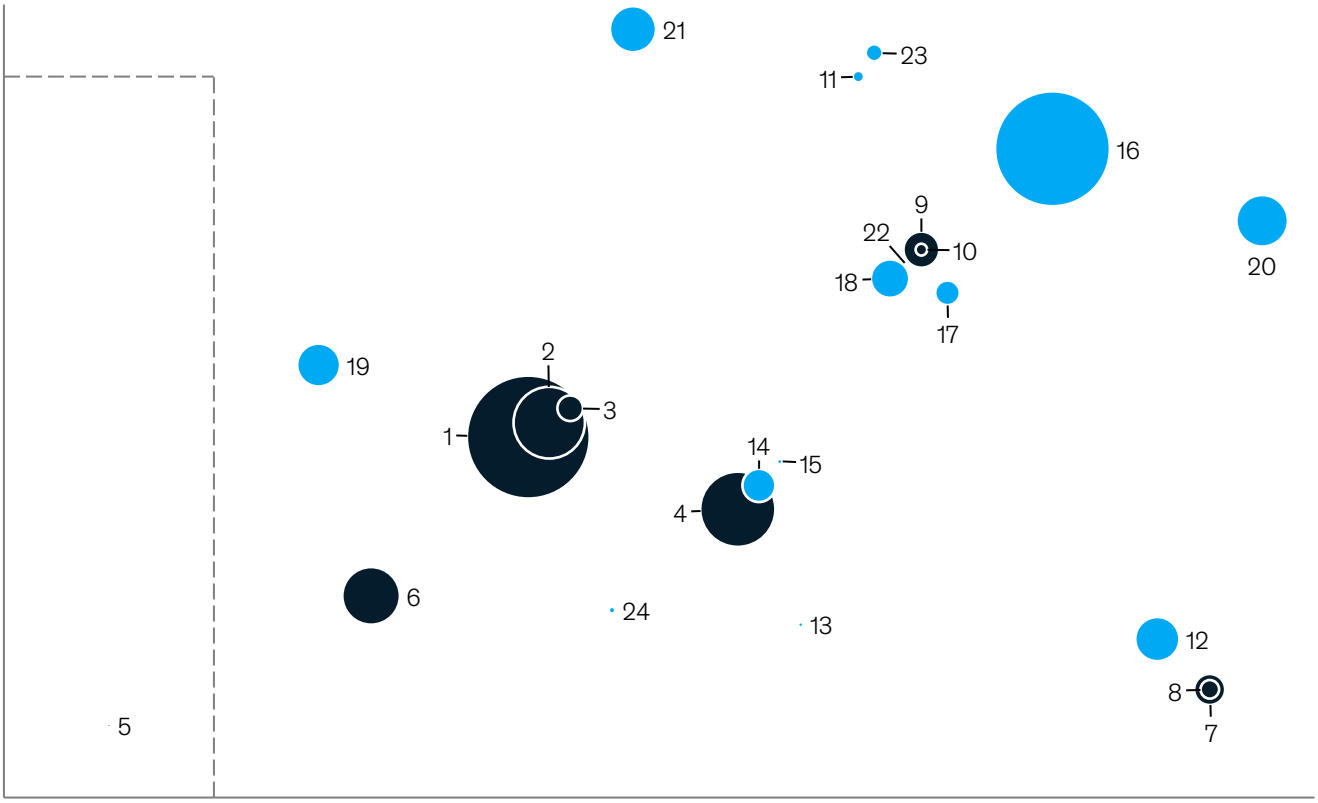


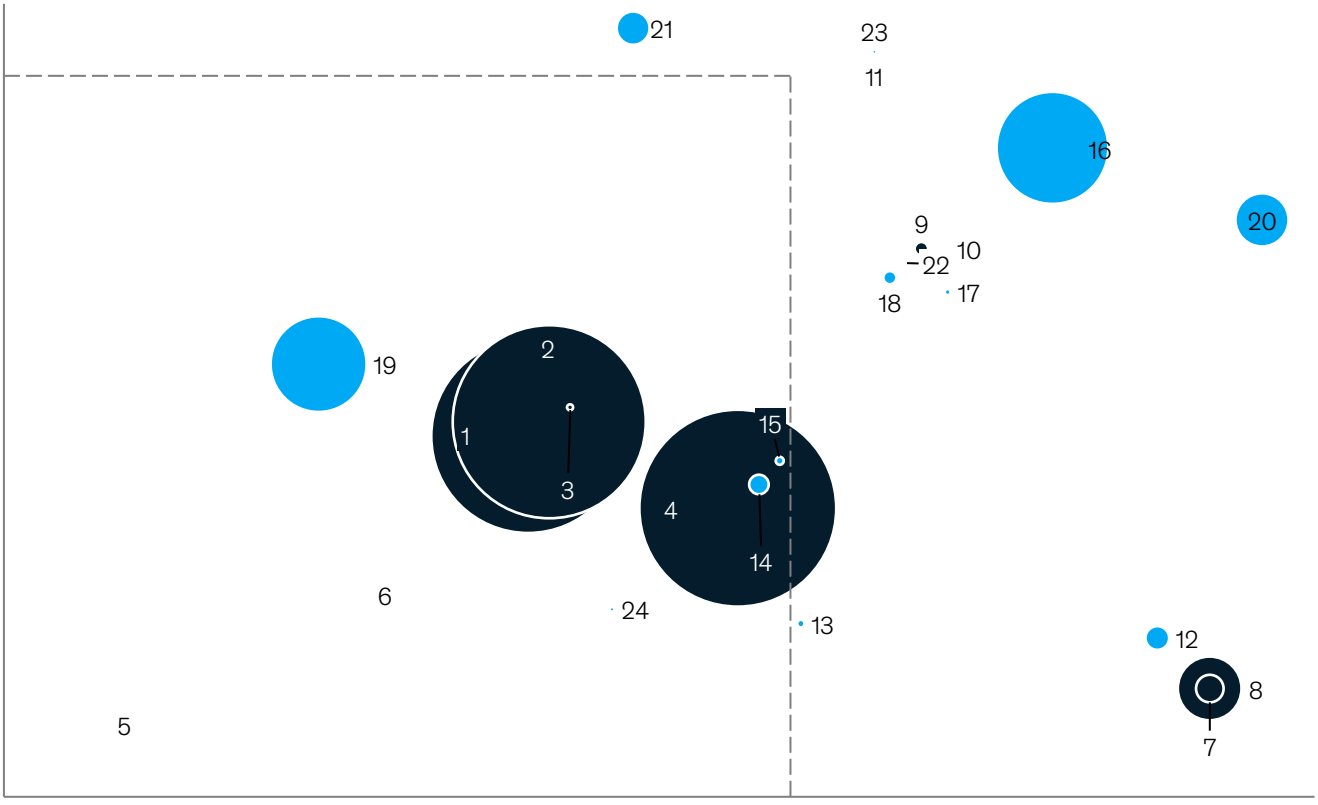
Exhibit 7 (continued)

Current levels of data standardization and breadth of data sharing would enable economies to create low to moderate levels of economic value (continued).

United States



India



4. Navigating risk and implementation challenges

Capturing the value from open financial data requires more than sufficient data standardization and broad data sharing: users—both consumers and providers—must trust the system and infrastructure must support financial data sharing. These factors together provide the groundwork for potential economic gains from open financial data, both those which we size and those that future innovation may make possible.

Data-sharing ecosystems require well-founded trust to encourage usage, protect users, and guard against other risks

An open-data ecosystem can function effectively only by achieving a level of well-founded trust among all participants. Without this, market participants—whether individual consumers or businesses—may opt out. Yet, the unauthorized use of personal data for economic gain has been an issue of rising concern for individuals globally. Financial data are particularly sensitive, and the long-term viability of data openness will depend on users' trust in the data-sharing ecosystem and the personal data protections it provides.

Willingness to participate and share data will require perceptions of agency, data protection, and consent. However, building on our earlier work on digital ID, we find that maximum value will be captured only in instances where individuals are knowing participants in data sharing, with clear knowledge of what personal data will be captured and how they will be used, or trust in automatic safeguards in place. Users are also more likely to want to share if they know what they are sharing and why that sharing is valuable to them. Open financial data regimes that are not coupled with consent are likely to erode trust and limit financial access, slowing or stopping positive economic impact.

The mere appearance of trust is not sufficient. For all the potential for value from open financial data, without the proper controls it might also inflict harm. Financial services can be a vector for unequal treatment or discrimination. Open financial data may make those abuses easier or more prevalent. For example, institutions might be able to use open financial data to more accurately prevent customers deemed unprofitable from opening bank accounts, which could systematically increase the unbanked population in certain segments.

Strong consumer financial protections may be necessary to prevent financial abuse whether data openness is driven by regulation or markets. Consumer financial data that is shared can be incorrect or require updating. Even beyond data errors, life circumstances change—someone might join the armed forces, have a conviction overturned, or fall victim to identity theft and need to expunge a series of fraudulent actions from their record. Without an automatic or easy mechanism for correction, problematic data might block an individual or MSME from accessing a needed financial product at a fair price.

Threats to cybersecurity could compromise user data and provider ability to use them as intended. Breaches can occur during transfer of data, or at any institution involved in the open data ecosystem, such as a bank or fintech. For example, when data transfer is achieved via APIs, a hacker who breaches such an API can hijack any apps that use the interface to collect data. As a result, open APIs require strong customer authentication. Breaches at providers themselves can come externally or from inside. For example, newer and less established fintechs in an ecosystem may be less experienced in confronting financial crimes. At the same time, identity and access management is required to guard either employee error or malicious behavior.

Successful data ecosystems tend to have built-in safeguards to ensure privacy and security while also giving users access to their personal data, decision rights over who else has access to that data, and transparency about who has accessed it. Whether created through regulatory regimes or enabled by private-sector actors, thoughtful system design with built-in privacy provisions like data minimization and correctability, well-controlled processes, and robust governance, together with the established rule of law, are essential to controlling risk and creating user trust. Key components are already in place at many institutions, and if leveraged and built upon systematically could help minimize risk and maximize trust.

Robust financial infrastructure is needed to support data sharing

At a basic level, the presence of digital financial infrastructure plays a critical role in helping the data ecosystem to flourish. An open-data ecosystem is a technologically complex endeavor that depends on the extent of a digital payments infrastructure and the presence or absence of key elements of a digital identification system for both individuals and businesses, as well as the degree of IT adoption more broadly.

Financial accounts and digital payment channels, along with digital identification systems with broad population coverage, are critical structural features of financial infrastructure needed to harness the value of open financial data. In related research spanning emerging and advanced economies, we found that when all these structural features—digital payments channels, digital ID, and data tethered to ID—were present in country-level financial infrastructure, COVID-19 disbursement programs could be optimally designed and delivered quickly; when not present, countries had to make trade-offs between the design ambition of their programs and their delivery success in terms of speed, coverage, and fraud levels in rolling them out.[25]

High-assurance digital ID facilitates user control of data, privacy protections, and security for online interactions, and reduces friction in managing online accounts. Open data systems without high-assurance digital IDs could mean consumers face growing complexity and struggle to keep track of their digital footprint or use their data securely and efficiently. At the same time, digital ID can support strong customer authentication, helping to control against cyber-attacks at APIs.

Additionally, for many emerging economies, basic internet access, smartphone penetration, and reliability of electricity constitute pre-requisites for capturing the full economic value of a data sharing ecosystem. Shortfalls in these elements limit the value of open financial data that could be captured. IndiaStack provides one example: it is revolutionary in its aspiration, and has the grounding of digital ID coverage for about 90 percent of the population, but still, India lacks full coverage of smartphone and internet access, particularly in rural areas, which would limit the value realized.[26]

[25] Olivia White, Anu Madgavkar, Tawanda Sibanda, Zac Townsend, and Maria Jesus Ramirez, "COVID-19: Making the case for robust digital financial infrastructure," January 2021, McKinsey.com.
[26] Data from Statista.com.

Trade-offs are required in implementing open financial data systems

Economies looking to gain value from open financial data have some important choices to make about the design of an ecosystem. These include the level of regulation, the extent of data standardization, the degree of control that users have over their own data, and whether the data are free or come at a market price. Some of the choices may involve trade-offs between potential beneficiaries. For example, while high levels of standardization and breadth will benefit financial institutions, they may come at the expense of individual data privacy and control. Where foundational digital infrastructure, such as digital payment channels and ID systems, is still emerging, economies will need to make critical design and implementation decisions on how to structure these in a way that drives consumer access and user adoption.

By a similar token, innovative product design will be needed to support value-creating use cases, but innovators will inevitably seek out and pursue opportunities for profit. If not well managed through appropriate policies, regulation, and infrastructure frameworks, the innovation spurred by open data ecosystems could serve to exacerbate existing inequalities and sources of discrimination or lead to less than optimal outcomes for certain stakeholders. The path forward, while fostering and promoting greater innovation, also needs careful consideration on how to ensure that all participants, and the ecosystem itself, benefits from its fruits, rather than a narrow subset.

All these issues have implications for the cost of implementing a data-sharing ecosystem, its security, the degree of competition that is encouraged, the restrictions put in place, and the customer experience. Whatever choices are made, economies would do well to preserve the extraordinary vibrancy of innovation in financial services in place across the globe today.

5. Looking ahead: the role of innovation in open data unbound

The landscape for open financial data across the globe is evolving rapidly. Exactly where it is heading is uncertain, but one thing is clear: innovation will be essential in unlocking the potential value from open financial data.

An economy can establish a virtuous cycle. Capturing the value from open financial data accessible today—given current levels of data standardization and breadth of data sharing—will require innovation. At the same time, the more value becomes accessible, through increases in standardization, breadth, or both, the more the potential for innovation will grow, likely beyond use cases we can envisage today.

Capturing the value available today will require market participants to develop and scale products and services that address specific use cases, including the 24 major ones we profile in this research. That will entail identifying new business opportunities, designing customer value propositions, and scaling new business models across the financial services value chain.

Private-market participants—both investors and providers—will play a central role. Different types of innovators, ranging from traditional banking incumbents to technology platform-based players and new fintech startups, could all play meaningful roles, focusing on their areas of strength and competitive advantage. The specific types of innovators would vary by market and depend on the structure of the financial data ecosystem.

In the United States, for example, where there is minimal regulatory intervention and more reliance on private-market solutions, a number of fintechs have gained scale. They are focused on laying the foundation of openness by playing an important role in data connection and aggregation. Future innovation that can yield the benefits described in our research will depend either on regulatory standards or on the ability of these players to build high-quality APIs and set de facto data standards in conjunction with a robust fintech environment ready to capitalize on this openness.

In the United Kingdom, on the other hand, where there is a regulatory mandate for Open Banking and successful fintech startups, companies are leveraging openness to innovate. Examples include apps that connect to all bank accounts and allow gig economy workers and MSMEs to quickly file tax returns; credit providers using transaction data to lend money at low rates including to individuals who were previously financially excluded; and streamlined account switching to enable consumers to easily access higher-yielding current accounts. In India, the investment in openness via IndiaStack positions the country to capture significant value. However, the relative difficulty of starting new businesses may hamper private innovation and inhibit how much of this value can be realized.

Innovation today is necessarily limited by current levels of data standardization and breadth of data sharing. Expanding the boundaries of open-data enablement would make new types of use cases possible, fueling greater innovation and greater value capture. For example, the proposed UK Financial Conduct Authority's regulation that would move the market from open banking to open finance, covering investments and mortgages in addition to payments, would open possibilities for many additional use cases for fintechs to explore. In the United States, as the Financial Data Exchange common standard for secure access to financial data expands, so also would the space and appetite for innovation coming from both traditional incumbents and startups.

Our prior research on robust digital financial infrastructure during the COVID-19 crisis suggested that such infrastructure not only provides economies with greater resilience in times of stress but also serves as a potential driver of greater productivity, efficiency, and economic growth. Open data is one example of how countries with a strong digital financial backbone can reap the benefits. If well designed and executed, a data-sharing ecosystem can bring multiple benefits to all participants. In the aftermath of the COVID-19 crisis, data sharing looks set to become an important differentiator for governments, financial systems, and financial institutions globally.

Further reading

The following is a selection of papers and open-data resources that go into greater depth on certain aspects of open data for finance, beyond the works cited in the footnotes of this discussion paper.

The appropriate use of customer data in financial services, World Economic Forum, September 2018.

Kaitlin Asrow, *The role of individuals in the data ecosystem: Current debates and considerations for data protection and data rights in the United States*, Federal Reserve Bank of San Francisco, June 2020.

Kaitlin Asrow and Beth Brockland, *Liability, transparency and consumer control in data sharing*, Center for Financial Services Innovation, September 2017.

Brazilian open banking model: First steps, Banco Central do Brasil, August 2019.

Yan Carrière-Swallow, Vikram Haksar, and Manasa Patnam, *India's approach to Open Banking: Some implications for financial inclusion*, IMF working paper, February 2021.

CFPB symposium: Consumer access to financial records, Consumer Finance Protection Bureau, February 2020.

Competition issues in data-driven consumer and small business financial services, Financial Data and Technology Association, June 2020.

Consumer financial data: Legal and regulatory landscape, Financial Health Network, Flourish, FinRegLab, and Mitchell Sandler, October 2020.

Consumer protection principles: Consumer-authorized financial data sharing and aggregation, Consumer Finance Protection Bureau, October 2017.

Open Banking year one: Insights from the CMA9 and more, Finextra, January 2019.

Open Data Barometer, World Wide Web Foundation, webfoundation.org.

Tracking the state of open government data, Global Open Data Index, index.okfn.org.

The use of cash-flow data in underwriting credit: Market context and policy analysis, FinRegLab, February 2020.

The use of cash-flow data in underwriting credit: Small business spotlight, FinRegLab, September 2019.

Related MGI and McKinsey research

Digital identification: A key to inclusive growth (April 2019)
This report shows how "good" digital ID is a new frontier in value creation for individuals and institutions around the world.

COVID-19: Making the case for robust digital financial infrastructure (January 2021)
This article explores how the pandemic was a tough real-life stress test for government disbursement schemes, highlighting opportunities but also gaps and vulnerabilities.

Digital India: Technology to transform a connected nation (March 2019)
This report looks at the rapid growth in digital adoption in India and focuses on some of the drivers, including the Aadhaar digital ID system and the IndiaStack technology referenced in this discussion paper.

www.mckinsey.com/mgi

Download and listen to MGI podcasts on iTunes or at
www.mckinsey.com/mgi/publications/multimedia/

McKinsey Global Institute
June 2021

Designed by the McKinsey Global Institute

www.mckinsey.com/mgi

@McKinsey_MGI
@McKinseyGlobalInstitute
@McKinseyGlobalInstitute

Subscribe to MGI's podcast, *Forward Thinking*:
mck.co/forwardthinking