

The quest for speed in payments¹

This feature looks at technology in payment systems. It compares the diffusion of real-time gross settlement (RTGS) systems for wholesale payments with that of faster systems for retail payments (fast payments). RTGS systems emerged in the 1980s and were adopted globally within a span of 30 years. Fast payments followed in the early 2000s, offering instant payments on a 24-hour, seven-day basis. So far, the diffusion of fast payments mirrors that of RTGS, and it is primed to take off. Yet even while adoption of fast payments is under way, the next generation of payment systems, such as those based on distributed ledger technology, is under development.

JEL classification: E58, G20, O30.

Mankind has always been in pursuit of speed. In track and field, Jamaican superstar Usain Bolt rules the 100 metre sprint as the fastest man in the world. In the pool, American great Katie Ledecky continues to smash world records. In Formula One, Dutch sensation Max Verstappen is upending the established order and exciting fans with phenomenal speed.

A similar quest is evident in payments. Throughout history, people have worked to accelerate the speed of payments for finance and commerce through the adoption of new technologies, big and small. The use of a simple ledger in the Middle Ages allowed the transfer of credit on the books of a money changer – the precursor of deposit banks (Kohn (1999)). The introduction of the telegraph revolutionised communications and enabled financial institutions to communicate instantly. Electronification and digitalisation in the modern era allowed automation. Few may recall that, before the 1980s, credit card transactions required phone authorisation and imprinting of cards on paper slips.

This quest continues today. Real-time gross settlement (RTGS) systems emerged in the 1980s to speed up wholesale payments and are now the standard around the world. More recently, faster systems for retail payments (fast payments) have emerged. These systems generally allow payees to receive funds within seconds of the payer initiating the payment, anytime and anywhere. A day or more to pay another person used to be acceptable, but in today's fast-paced environment this seems like an eternity. Consumers, who are used to instant communication via e-mail and social media, now expect the same experience when it comes to payments.

¹ The authors thank Claudio Borio, Benjamin Cohen, Ingo Fender, Andreas Schrimpf and Hyun Song Shin for useful comments and suggestions. We are grateful to Codruta Boar for excellent research assistance. The views expressed are those of the authors and do not necessarily reflect those of the BIS.

This article looks at two of the latest leaps forward in speed. Drawing in part on a recent report on fast payments by the Committee on Payments and Market Infrastructures (CPMI),² it compares how RTGS and fast payments have spread around the world. The pattern of diffusion of fast payments is remarkably similar to the move to RTGS for wholesale payments two decades earlier. Like RTGS, fast payments are primed for take-off 15 years after the first implementation. The feature also looks to the future of payments, drawing on two other CPMI reports on distributed ledger technology and digital currencies.

Emergence of fast (retail) payments

Payments are transfers of monetary value from payers to payees, usually in exchange for goods and services or to fulfil contractual obligations. They come in many forms and sizes. *Wholesale payments* are high-priority and typically large-value transfers that are made between financial institutions for their own accounts or on behalf of their customers. Wholesale payments are usually settled via dedicated interbank settlement systems. In contrast, *retail payments* are lower-value transactions between individuals, businesses and governments in such forms as cash, cheques, credit transfers, and debit and credit card transactions.

An important distinction between wholesale and retail payment systems has traditionally been the speed of settlement finality.³ It typically took a day or more for a payee to receive funds using a traditional retail payment system, and for some systems payments were revocable within a certain period, adding an element of uncertainty. Hence, time-sensitive payments (even lower-value ones) were directed via the interbank payment system because of its ability to credit and debit accounts with real-time finality.

The speed of retail payments is now immediate in some countries thanks to improvements in information and communication technologies, including the ubiquity of smartphones and the internet. Fast payments provide retail funds transfer “in which the transmission of the payment message and the availability of ‘final’ funds to the payee occur in real-time or near real-time on as near to a 24/7 basis” (CPMI (2016b)). Further, this feature focuses on *open systems*, where end users can use any number of intermediaries, such as payment service providers (PSPs) and banks, to access the payment system.⁴ (See the box for how fast payments work and examples of fast payment systems.)

² The CPMI is a BIS-based committee of senior central bank officials that promotes the safety and efficiency of payment, clearing, settlement and related arrangements (www.bis.org/cpmi/).

³ *Settlement finality* is defined as the point when the irrevocable and unconditional transfer of an asset occurs.

⁴ *Closed systems* provide payment services to only their customers, and credits and debits occur on their own books. Closed systems often have limitations on the coverage of users within a market or jurisdiction, which is a key element to the successful adoption of new payment services.

How do fast payment systems work?

A defining characteristic of a fast payment system is the ability to complete a payment almost immediately and at any time. To achieve this outcome, all fast payment systems require immediate clearing^① between the payment service providers (PSPs) of the payer and payee. Funds settlements between the PSPs, however, do not necessarily need to occur immediately for each and every payment order. Payee funds availability and inter-PSP settlement can be either coupled (ie real-time settlement) or decoupled (ie deferred settlement).

In *real-time settlement*, payee funds availability and inter-PSP settlements are coupled, with inter-PSP settlements occurring in real time. In other words, the debiting and crediting of funds from the payer to the payee occur at the same time as associated debiting and crediting of the PSP in the fast payment system. In this model, credit risks between participating PSPs do not arise, but participating PSPs continuously require sufficient liquidity to support real-time settlements of fast payments. Therefore, a system is required to address the possible need for liquidity provision to the participant PSPs in the system, the adequacy of the settlement system's operating hours and associated liquidity facilities. Countries that use this model include Mexico and Sweden.

In *deferred settlement*, payee funds availability and inter-PSP settlements are decoupled, with inter-PSP settlements being deferred with batch settlement. That is, while payer and payee accounts are debited and credited in real time or near real time, the associated settlements between the PSPs are batched and executed at pre-specified times. In this model, credit risk inherently arises for PSPs, as the payee's PSP advances the funds to the payee before inter-PSP settlement takes place. A variety of tools can mitigate this risk, including prefunding of positions, a maximum limit on the net debit or credit position that can be established between PSPs, and collateralisation of debit positions. Countries that use this model include India and the United Kingdom.

Examples of fast payment systems

Mexico – The Sistema de Pagos Electrónicos Interbancarios (SPEI) is the Bank of Mexico's main payment system, providing both wholesale and retail payment services. SPEI was launched in 2004 and provided near real-time retail payments. As of November 2015, the service offers 24/7 availability. Funds are available to the payee in less than 15 seconds for mobile payments and less than 60 seconds for other online payments. Currently, 109 institutions (66 banks and 43 non-banks) participate in SPEI as direct members to provide their customers with fast payment services.

Sweden – BiR/Swish, introduced in 2012, is a real-time settlement system for mobile payments in Sweden. Being a privately owned special purpose institution that conducts settlement in commercial bank money, which in turn is fully backed by funding in central bank money, the system allows real-time settlement of fast payments even during times when other settlement facilities (eg the central bank real-time gross settlement system) are closed. The typical time between payment initiation and availability of final funds to the payee for a successful fast payment transaction is one to two seconds. More than half of the country's population uses the Swish mobile app to make fast payments.

India – The Immediate Payment Service (IMPS) went live as a new instant mobile payment system in 2010. The system allows mobile phone subscribers and internet-connected devices to send and receive payments. Payees typically receive funds in less than 30 seconds. The service provides access to fast payments through 190 PSPs. In December 2016, IMPS processed 60.5 million transactions, which represented a 50% increase from the previous month – the largest monthly increase to date – likely driven by the Indian banknote demonetisation directive of November 2016 and the subsequent push from the government to get digital payments adopted nationwide.

United Kingdom – The Faster Payments Service (FPS) is a deferred net settlement system for credit transactions in the form of single, immediate payments, forward-dated payment, or standing orders for households and corporates. The service, which was launched in 2008, allows a payer to initiate a payment simply using the payee's mobile phone number. Funds are typically available to the payee within seconds of the payer initiating the payment transfer. FPS has 10 direct participants, who open up their customer channels to FPS. In December 2016, the service processed 125 million payments totalling £103 billion.

^① CPMI (2016b) defines *clearing* as the process of transmitting, reconciling and, in some cases, confirming transactions prior to settlement, potentially including the netting of transactions and the establishment of final positions for settlement.

Sources: CPMI (2016b); publicly available information.

Fast payment systems in selected countries

Table 1

Year	Country	System	Year	Country	System
2001	Korea	Electronic Banking System	2013	Turkey	BKM Express
2003	Chinese Taipei	ATM, FXML and FEDI systems	2014	Denmark	Nets Real-time 24x7
	Iceland	CBI Retail Netting System (JK)		Italy	Jiffy – Cash in a flash
2006	Malaysia	Instant Transfer		Singapore	Fast And Secure Transfer
	South Africa	Real-Time Clearing	2015	Mexico	SPEI
2007	Korea	CD/ATM System		Switzerland	Twint
	2008	Chile	Transferencias en línea	2017p	Australia
United Kingdom		Faster Payments Service	2017/18p	Saudi Arabia	Future Ready ACH
2010	China	Internet Banking Payment System	2018p	Hong Kong SAR	To be determined
	India	Immediate Payment Service		Japan	Zengin System
2011	Costa Rica	Transferencias de Fondos a Terceros del Sinpe	2019p	Hungary	Instant Payments
	2012	Ecuador		Pago Directo	Netherlands
Poland		Express ELIXIR			
Sweden		BiR/Swish			

p = proposed implementation.

Sources: CPMI (2016b); FIS (2015); national central banks.

Fast payment systems began to emerge in the early 2000s. The first system to satisfy both the fast and continuous service availability requirements under the CPMI (2016b) definition was the Korean Electronic Banking System, which went live in 2001 (Table 1). Two other fast payment systems – in Chinese Taipei and Iceland – were implemented in 2003, and Malaysia and South Africa followed three years later. The two most populous countries in the world joined in 2010. Among the major advanced economies, the first to adopt a fast payment system was the United Kingdom in 2008, followed by Italy in 2014. Japan is planning to make its Zengin System, whose end-to-end speed is already real-time, available 24/7 in 2018.

Technology adoption and diffusion

Adoption is the decision to acquire and use a technology, and involves a weighing of costs and benefits. This is seldom straightforward; a significant part of the costs is incurred upfront, while the benefits tend to accrue over time. Moreover, it might be cheaper to adopt tomorrow rather than today, as the cost of technology tends to decrease over time. Matters are further complicated by the fact that the benefits may depend on the number of other adopters of the technology.

The process by which a new technology spreads is known as *diffusion*. According to Rogers (2003), diffusion is how a new technology (innovation) spreads over time and space among members of a social system through various channels. In other words, diffusion is the cumulative result of a series of adoption decisions, which are often implemented under uncertainty and with limited information. If diffusion passes

a certain threshold, or critical mass, then the technology is likely to take off and be widely adopted. Failure of a technology to take off is often due to either the inertia of an existing technology or the emergence of a superior one.

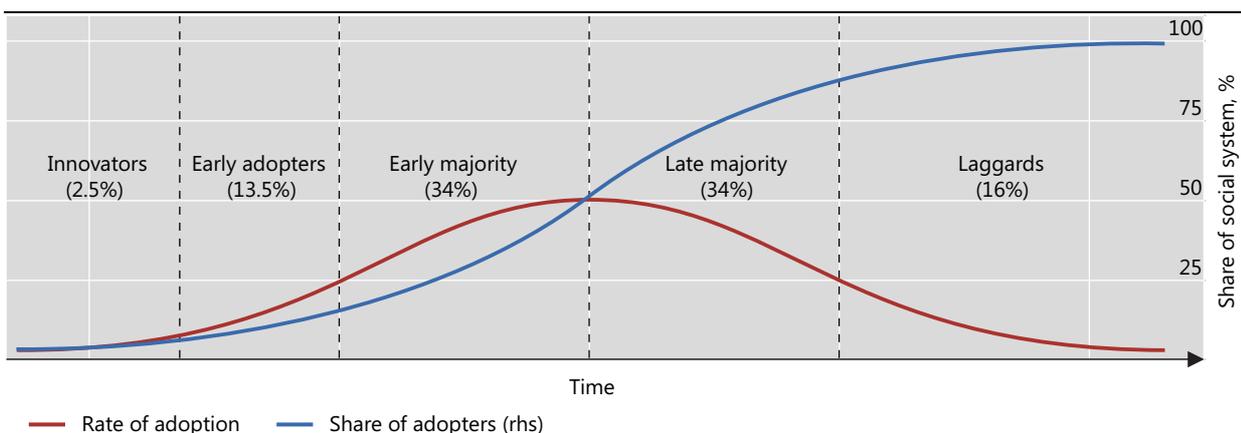
Empirical studies of diffusion suggest that the rate of adoption follows a predictable pattern over time (eg Griliches (1957)). The rate is generally slow at first and starts to accelerate if the technology gains traction within the social system. Rapid adoption continues until a substantial share of agents have shifted to the new technology. At this point, the rate of adoption levels off and eventually falls. That is, the rate of adoption tends to follow a bell curve over time and the share of adopters is a sigmoidal, or S-shaped, curve.

Rogers (2003) provides a simple framework to think about diffusion that builds on this stylised fact (Graph 1). It classifies members of the social system into five categories that reflect their “innovativeness”, or predisposition to adopt a new technology, based on when they adopt relative to the median adopter. The first two categories are called *innovators* and *early adopters*, respectively. Innovators are enterprising and willing to take risk, and early adopters are often key opinion leaders. Early adopters are critical to whether or not the innovation spreads. In Rogers’ framework, the adoption time of the last early adopter corresponds to the first inflection point of the bell curve, ie where the adoption of the technology takes off. The remaining categories are *early majority*, *late majority* and *laggards*.

Technology diffusion in payments is potentially different from other areas, as the adoption decision may involve more than profit maximisation. Payment systems often exhibit significant economies of scale and scope, so that they tend to be owned by either the central bank or an industry association/consortium. A central bank is likely to consider monetary and financial stability issues. And decision-making by an industry association or consortium is often complicated and time-consuming: it may require a catalyst or strong outside impetus.

Rogers’ diffusion of innovation model

Graph 1



Source: Rogers (2003).

The diffusion of real-time gross settlement

Prior to the 1980s, wholesale payments typically accumulated over the business day and were settled by netting obligations on the central bank books either at the end of the day or the next morning. This method – known as deferred net settlement (DNS) – significantly reduces the amount of money that needs to change hands but also gives rise to potential settlement risk.⁵ If a bank with a net amount due defaults, payments involving that bank may need to be unwound. This implies new net obligations for all other banks. Conceivably, some other bank – expecting incoming funds from the failed bank – may then not be able to meet its new net obligation and may thus fail, potentially setting off a cascade of failures.⁶

With wholesale payment values increasing in the 1980s, a number of central banks became more conscious of the settlement risk involved in large-value funds transfer and its potential financial stability implications. Driven by advances in information and communication technology, central banks shifted to RTGS, where payments are settled immediately one by one.⁷ As payments are final and irrevocable, settlement risk is eliminated (eg Borio and Van den Bergh (1993)). In 1985, three central banks had RTGS systems; by 1990, the number was eight – including the Deutsche Bundesbank, the Bank of Japan, the Swiss National Bank and the Federal Reserve.

The adoption of RTGS took off in the mid-1990s in part because it was made a prerequisite for joining the Economic and Monetary Union of the European Union. This led to a flurry of new systems or upgrades to existing ones by eventual euro area members as well as prospective ones (Graph 2, left-hand panel).

Another factor contributing to the take-off was the guidance provided in a number of BIS committee reports that analysed the risks and benefits of netting and RTGS systems.⁸ These reports later developed into a set of principles for the design of wholesale payment systems.⁹ In turn, these Core Principles for Systemically Important Payment Systems became part of the toolkit of the Financial Sector Assessment Programs (FSAPs) and technical assistance programmes conducted by the International Monetary Fund and The World Bank Group. This supranational impetus – as well as lower implementation costs due to the emergence of several off-

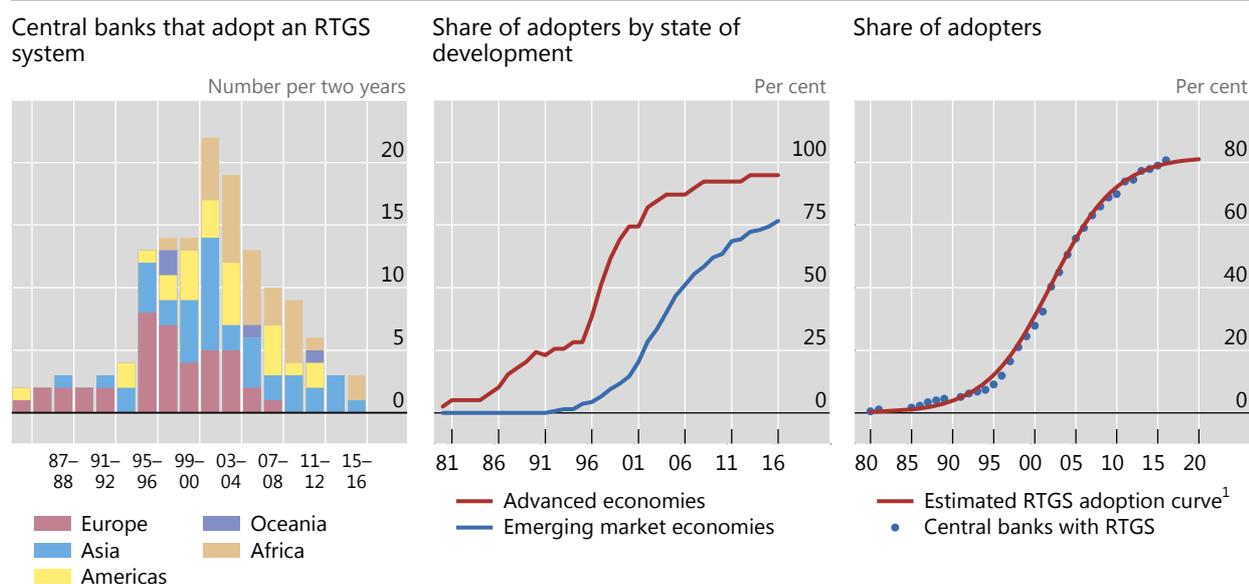
⁵ *Settlement risk* is the general term used to designate the risk that settlement in a funds or securities transfer system will not take place as expected. This risk may comprise both credit and liquidity risk. See CPMI (2016a).

⁶ Humphreys (1986) found that the failure of a major participant in the New York-based Clearing House Inter-Bank Payments System (CHIPS) could, given the rules at the time, trigger a system-wide crisis. In 2001, CHIPS moved to a so-called “hybrid” settlement system which combines features of DNS and RTGS.

⁷ The Federal Reserve’s Fedwire Funds Service® is the world’s oldest RTGS system. Its origins go back to 1918, when the Federal Reserve created a network of wire communications among the individual Federal Reserve Banks.

⁸ Group of Experts on Payment Systems (1989), Committee on Interbank Netting Schemes (1990) and CPSS (1997).

⁹ The Core Principles for Systemically Important Payment Systems were later replaced by the principles for financial market infrastructures (CPSS (2001) and CPSS-IOSCO (2012), respectively).



¹ Using a generalised logistic model.

Sources: Bech and Hobijn (2007); Rogers (2003); national central banks; The World Bank Group; authors' calculations.

the-shelf RTGS solutions – led to the adoption of RTGS systems by emerging market economies from the late 1990s (Graph 2, centre panel).¹⁰

By 2000, 49 out of 176 central banks worldwide had RTGS systems; and by 2005, all advanced economy central banks had adopted RTGS with the exception of Canada.¹¹ At the end of 2016 (or some 35 years after “first” implementation), there are only a handful of late adopters (Graph 4). Consistent with models of technology diffusion, the rate of RTGS adoption by central banks followed a bell curve, and consequently the share of adopters takes the form of an S-curve (Graph 2, left- and right-hand panels).

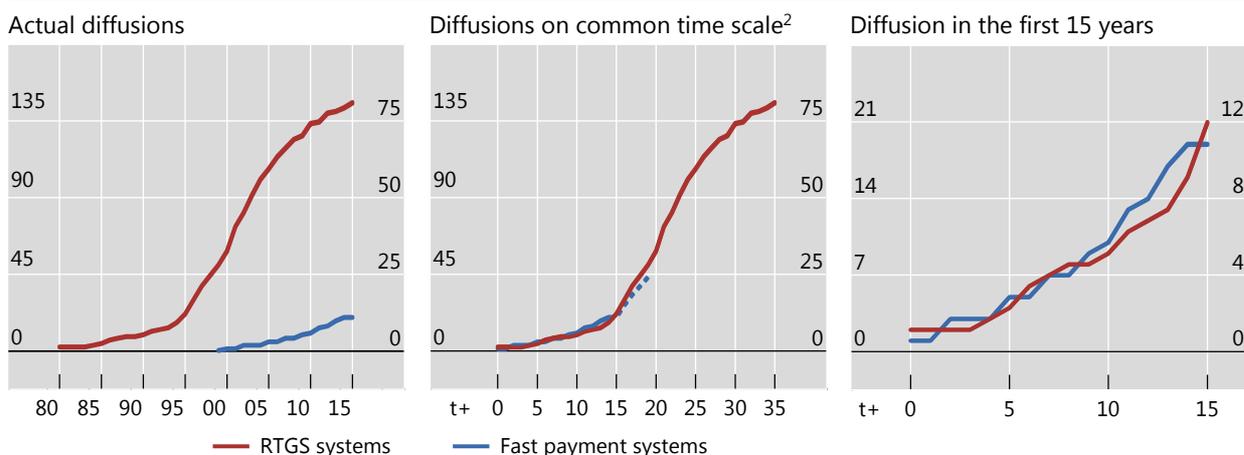
The diffusion of fast payments

Comparing the diffusions of fast payments and RTGS reveals some interesting similarities and differences. Despite taking place about 20 years apart, for example, the diffusion of fast payments is so far surprisingly similar to that of RTGS (Graph 3, left-hand panel).¹² In fact, when put on the same time scale, the two diffusion curves are virtually identical (Graph 3, centre and right-hand panels). This also holds when projecting diffusion five years ahead based on systems under consideration or

¹⁰ Lester et al (2008) provide a theoretical model of RTGS adoption. They argue that deferred net settlement is less costly due to the higher information and communication technology costs associated with an RTGS environment. If these costs fall, then the settlement risk reduction achieved in an RTGS system may be enough to justify a switch from DNS.

¹¹ Canada opted for a hybrid system that is considered equivalent to RTGS in terms of settlement risk.

¹² This comparison does not include the US Fedwire Funds Service® that dates back to 1918 (see footnote 7).



¹ Left-hand scale: number of central bank adopters; right-hand scale: adopters as a percentage of total number of central banks. ² The dashed line corresponds to projected implementation based on public project announcements and assumptions that central banks currently considering a fast payment system would implement one within the next five years.

Sources: Bech and Hobijn (2007); CPMI (2016b); FIS (2015); national central banks; The World Bank Group; authors' calculations.

development. The implementation of fast payments is now at a similar stage to the implementation of RTGS in 1995.

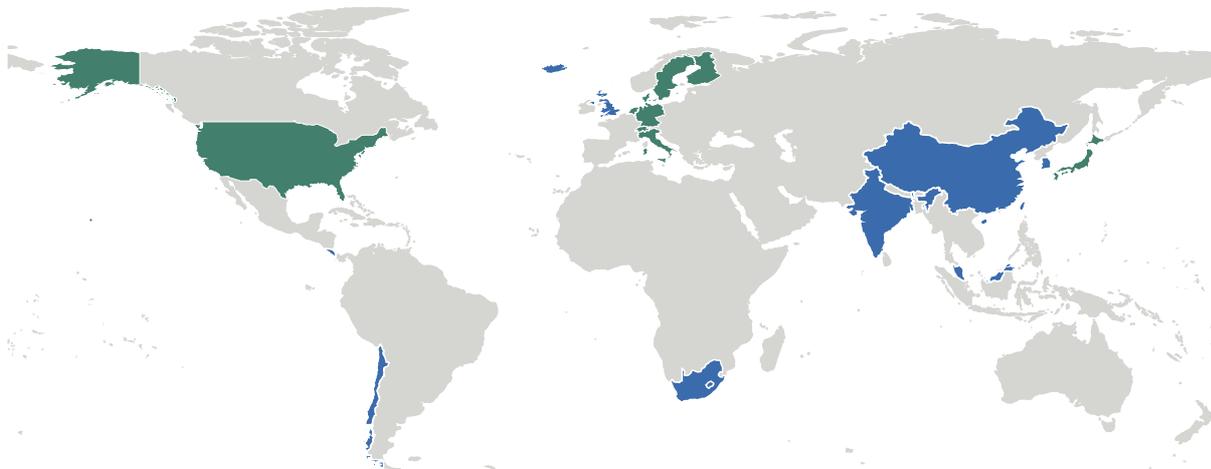
This similarity is in itself somewhat surprising, however. It runs contrary to the perception that the general pace of technology adoption is speeding up (McGrath (2013)). One possible explanation is that adoption of a fast payment system is typically not solely an individual decision. Rather, it tends to be a decision that requires coordination and collective decision-making. This may be slowing diffusion in some cases, even if the technology itself has spread faster. In fact, a common challenge in many countries is to overcome potential coordination issues between different stakeholders.¹³

In contrast to the diffusion of RTGS, fast payment innovators and early adopters are not predominantly advanced economies, and indeed there is no discernible difference in the diffusion of fast payments between emerging market and advanced economies. A possible explanation as to why emerging market economies may be adopting fast payments at a similar rate to advanced economies is the lack of existing electronic retail payment infrastructures. This means that the net benefit of adoption is likely to be higher and the decision-making process may be easier in the absence of well established infrastructures.

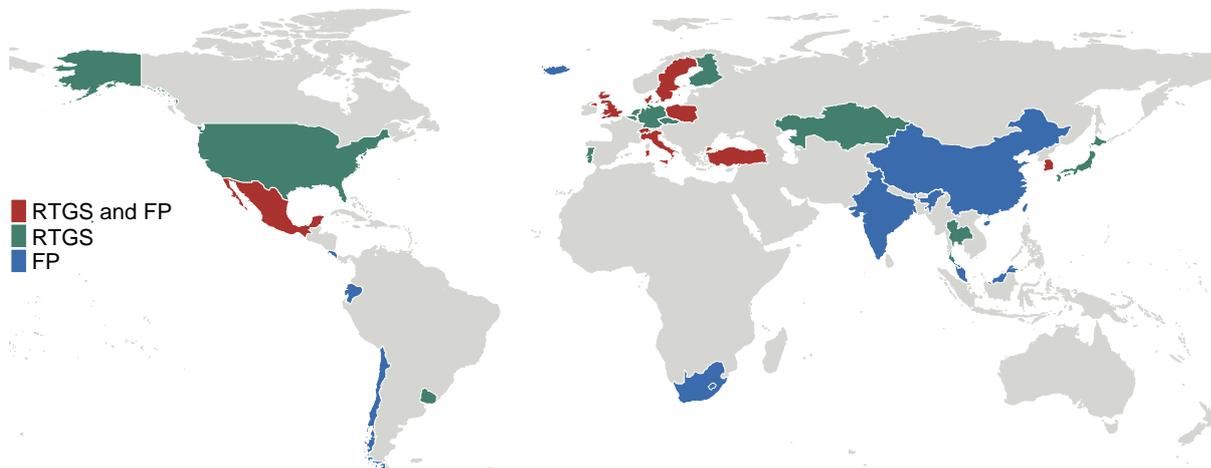
As of the end of 2016 (or 15 years after implementation of the first fast payment system), there were 20 such systems in operation meeting the CPMI (2016b) definition. The countries with these systems cover more than 40% of the world's population – twice that of RTGS systems at the same point in the diffusion of RTGS – and roughly 30% of global GDP – roughly half of where RTGS was.

¹³ See CPMI (2016b, pp 27–8).

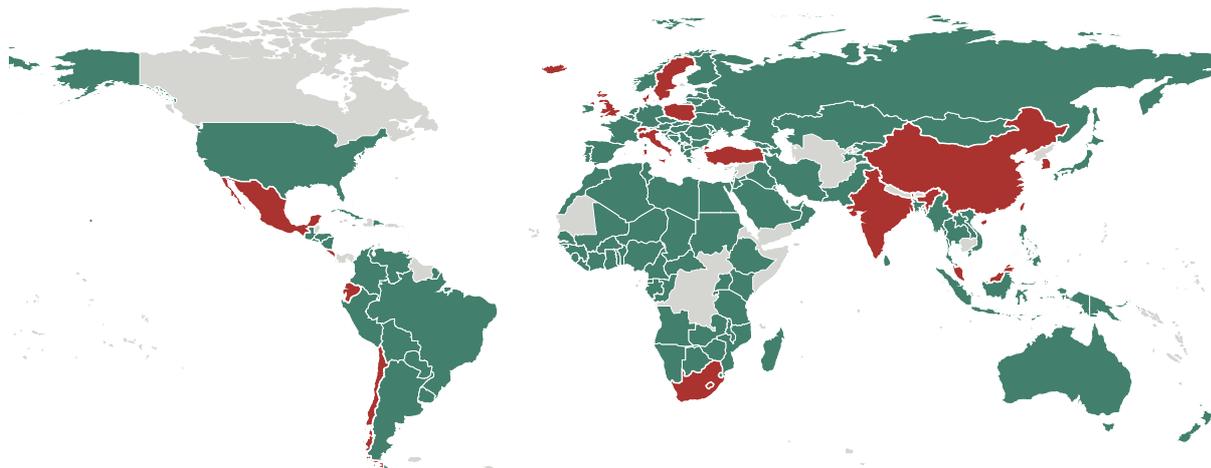
Ten years after implementation



Fifteen years after implementation



Thirty-five years after implementation



¹ For RTGS, 10, 15 and 35 years after implementation reflects 1991, 1996 and 2016, respectively. For fast payments, 10 and 15 years after implementation reflects 2011 and 2016, respectively.

Sources: Bech and Hobijn (2007); CPMI (2016b); FIS (2015); national central banks; The World Bank Group.

Continuing evolution of the payments landscape

As countries progress in implementing fast payments, work is under way to develop technologies and features to enhance the speed of certain types of payment. Unlike RTGS and fast payments, these efforts are not primarily driven by central banks or commercial banks but rather by entrepreneurs, technology firms and venture capitalists. Over the last couple of years, significant investment has gone into financial technology (fintech). A sizeable share has been directed towards payments-related projects – targeting all parts of the payment-processing chain, from user interfaces to clearing and settlement.

One particular technology that is receiving significant attention is *distributed ledger technology* (DLT) – commonly referred to as *blockchain* technology. DLT promises to streamline payment, clearing and settlement processes by, for example, reducing the number of intermediaries and eliminating the need for reconciliation among those that remain.¹⁴ It allows participants in a payment system (or other arrangement) to jointly manage and update a synchronised, distributed ledger. This contrasts sharply with existing payment systems, where a single authority manages a central ledger. In DLT-based payment systems, participants can submit, validate and record transfers on the distributed ledger with little or no need for special intermediaries.

There are three areas where DLT could have a significant impact on the speed of payments. One is the payments associated with the settlement of securities. Today, it typically takes several days after the trade date for the security and the associated payment to change hands. If DLT arrangements could provide real-time or near real-time settlement of the securities and associated funds transfer on the trade date, the cost savings could be significant. It would reduce record-keeping and reconciliation costs, as well as settlement costs, by eg eliminating the use of collateral to guarantee the exchange of securities and cash.

The second area is cross-border payments, which are currently time-consuming. A cross-border payment typically involves the use of a local bank, a foreign bank and one or more correspondent banks. Other intermediaries involved may include financial services or communications companies such as SWIFT or Western Union. By using a distributed ledger, the sender and beneficiary could in principle settle cross-border funds transfers in real time without the need for financial intermediaries. Yet there may be significant barriers to implementing such a solution due to jurisdictional differences in the legal, regulatory and operational frameworks.

The third area is central bank-issued digital currencies. Sveriges Riksbank is studying the issuance of an e-krona as a complement to physical cash (Skingsley (2016)).¹⁵ Several other central banks have also publicly announced internal efforts to study digital currencies for either retail payments or wholesale payments, or both (eg Mersch (2017)). If implemented, the impact would be significant: banks have traditionally played a central role in supporting payments, so that removing them from the centre of this system could reshape banking and, more broadly, the financial markets.

¹⁴ See CPMI (2015) and CPMI (2017).

¹⁵ Although digital currencies in the form of e-money have existed for decades, DLT is slightly different in that it could provide for the possibility of peer-to-peer payments in a decentralised network without the need for a financial intermediary (CPMI (2015)).

Conclusion

Payments are a dynamic, constantly evolving business. As the diffusion of RTGS ends, the implementation of fast payments is primed to take off. In fast payments, emerging market economies are likely to leapfrog advanced ones. Still, efforts are already under way to design the next-generation payment system. Blockchain and other distributed ledger technology holds great promise, but projects are currently only in the proof-of-concept phase. The first wide-scale use of distributed ledgers in payments is likely to be years away, as technological, legal and other hurdles will need to be overcome. Central banks and other authorities will continue to play a critical part in furthering greater efficiency and resilience of payments.¹⁶

¹⁶ To assist in these efforts, the CPMI has developed an analytical framework to assess and analyse DLT-based projects in payment, clearing and settlement activities (CPMI (2017)).

References

Bech, M and B Hobijn (2007): "Technology diffusion within central banking: the case of real-time gross settlement", *International Journal of Central Banking*, September, pp 147–81.

Borio, C and P Van den Bergh (1993): "The nature and management of payment system risk: an international perspective", *BIS Economic Papers*, no 36.

Committee on Interbank Netting Schemes (1990): *Report of the Committee on Interbank Netting Schemes of the central banks of the Group of Ten countries* (Lamfalussy Report), November.

Committee on Payment and Settlement Systems (1997): *Real-time gross settlement systems*, no 22, March.

——— (2001): *Core Principles for Systemically Important Payment Systems*, no 43, January.

Committee on Payment and Settlement Systems and International Organization of Securities Commissions (2012): *Principles for financial market infrastructures*, December.

Committee on Payments and Market Infrastructures (2015): *Digital currencies*, no 137, November.

——— (2016a): *A glossary of terms used in payments and settlement systems*, October.

——— (2016b): *Fast payments – enhancing the speed and availability of retail payments*, no 154, November.

——— (2017): *Distributed ledger technology in payment, clearing and settlement – an analytical framework*, February.

FIS (2015): *Flavours of fast – a trip around the world in immediate payments*, May.

Griliches, Z (1957): "Hybrid corn: an exploration in the economics of technological change", *Econometrica*, vol 25, no 4, pp 501–22.

Group of Experts on Payment Systems (1989): *Report on netting schemes* (Angell Report), February.

Humphrey, D (1986): "Payments finality and risk of settlement failure", in A Saunders and L Whyte (eds), *Technology and the regulation of financial markets: securities, futures, and banking*, Lexington, MA: DC Heath and Company.

Kohn, M (1999): "Early deposit banking", Dartmouth College, *Working Paper* 99–03, February.

Lester, B, S Millard and M Willison (2008): "Optimal settlement rules for payment systems", in S Millard and V Saporta (eds), *The future of payment systems*, London: Routledge.

McGrath, R (2013): "The pace of technology adoption is speeding up", *Harvard Business Review*, November.

Mersch, Y (2017): "Digital base money – an assessment from the European Central Bank's perspective", speech, Helsinki, 16 January.

Rogers, E (2003): *Diffusion of innovations*, fifth edition, New York, Free Press.

Skingsley, C (2016): "Should the Riksbank issue e-krona?", speech at FinTech Stockholm 2016, 16 November.