Deloitte.

Economic impact of real-time payments

RESEARCH REPORT

JUNE 2019

Important notice from Deloitte

This final report (the "Final Report") has been prepared by Deloitte LLP ("Deloitte") for Vocalink Limited, a Mastercard company, ("Vocalink") in accordance with the Framework Agreement dated 8 August 2018 and the Work Order dated 9 August 2018 (together "the Contract") and on the basis of the scope and limitations set out below.

The Final Report has been prepared solely for the purposes of providing a framework for assessing the economic impacts of instant payment schemes, as set out in the Contract. It should not be used for any other purpose or in any other context, and in no way constitutes a replacement for a detailed business case for developing and/or introducing a real-time payments scheme, and Deloitte accepts no responsibility for its use in either regard.

The Final Report is provided exclusively for Vocalink's use under the terms of the Contract. No party other than Vocalink is entitled to rely on the Final Report for any purpose whatsoever and Deloitte accepts no responsibility or liability or duty of care to any party other than Vocalink in respect of the Final Report or any of its contents.

As set out in the Contract, the scope of our work has been limited by the time, information, and explanations made available to us. The information contained in the Final Report has been obtained from Vocalink and third party sources that are clearly referenced in the appropriate sections of the Final Report. Deloitte has neither sought to corroborate this information nor to review its overall reasonableness. Further, any results from the analysis contained in the Final Report are reliant on the information available at the time of writing the Final Report and should not be relied upon in subsequent periods.

This document also includes certain statements, estimates, and projections with respect to anticipated future projections or impacts. Such statements, estimates, and projections reflect various assumptions concerning anticipated results and are subject to significant business, economic, and competitive uncertainties and contingencies, many of which are or may be beyond the control of Vocalink. Accordingly, there can be no assurance that such statements, estimates, and projections will be realised. The actual results may vary from those projected, and those variations may be material. Whilst we have commented on such statements, estimates and projections and their implications, we accept no responsibility for their accuracy or completeness.

We have conducted scenario analysis based on available data and estimated projections. The results produced by our scenarios under different assumptions are dependent upon the information with which we have been provided. Our scenarios are intended only to provide an illustrative analysis of the implications of real-time payments schemes. Actual results are likely to be different from those projected by the scenarios due to unforeseen events and accordingly we can give no assurance as to whether, or how closely, the actual results ultimately achieved will correspond to the outcomes projected in the scenarios.

Any decision to invest, conduct business, enter or exit the markets considered in the Final Report should be made solely on independent advice and no information in the Final Report should be relied upon in any way by any third party. This Final Report and its contents do not constitute financial or other professional advice, and specific advice should be sought about your specific circumstances. In particular, the Final Report does not constitute a recommendation or endorsement by Deloitte to invest or participate in, exit, or otherwise use any of the markets or companies referred to in it. To the fullest extent possible, both Deloitte and Vocalink disclaim any liability arising out of the use (or non-use) of the Final Report and its contents, including any action or decision taken as a result of such use (or non-use).

Contents

Executive summary

Real-time payments enable people, businesses, and governments to make payments more quickly with funds immediately available for use.

Real-time payments have been adopted by a number of countries, but the opportunity now exists to extend the economic benefits through wider adoption.

This report is designed to help policymakers understand the potential impact of wider adoption of the 'payments mix', the benefits that will matter most for their country, and the scale at which those benefits might be realised.

It identifies a range of advantages in using real-time payments, from reducing the size of the shadow economy (and thereby raising tax receipts), to improving the efficiency of the financial system, and helping to tackle financial crime. Underpinned by existing research and new econometric analysis, this report shows that real-time payments can:

Deliver a shift to a more formalised economy

- Case studies in countries such as Thailand suggest real-time payments can help displace cash.
- Real-time payments can therefore reduce the scale of the shadow economy and boost tax collection/receipts and help to reduce financial crime.

Reduce uncertainty and increase working capital

- 63% of businesses surveyed in this study maintain a cash contingency to cover the time it takes to receive payments.
- 50% would be willing to pay a fee to receive payments immediately.
- Real-time payments can reduce business uncertainty from payment delays and therefore boost working capital.

Improve the efficiency of the financial system

- Slow speed of payments being processed results in money being locked in the system and not available for consumers and businesses to use (a kind of 'float').
- The scale of the float at any one time can be substantial. For example, the daily value for cheques alone in Germany, an economy with relatively low usage of cheques as a payment instrument, was USD 654m in 2016. Real-time payments can be used to free up that float.

Enhance financial inclusion

- Real-time payments can make formal financial services more accessible and attractive to consumers who currently use only cash or mobile money services.
- Real-time payments can also help consumers with budgeting.

Reduce the cost of payment systems

- Real-time payments cost about the same as non-instant electronic payments overall, at about USD 1.95 for 10 transactions per capita.
- This is significantly less than cheques, in particular, at USD 2.79.
- Note, that these are ongoing costs and do not reflect migration costs in the transition, which are less salient for the economic impacts and best considered in a specific country business case.

Long-term advantages

The full impact of adopting real-time payments will only become clear with time (and within a given country) as corporates, start-ups, and policymakers deploy innovative services on top of the modern payments infrastructure.

Real-time payments can enable future innovations, creating a platform for the next wave of fintech pioneers. This could include the development of new products and services, which use the data generated by digital payments (subject to the appropriate protections).

The track record of real-time payments is long enough to establish some core facts about the effects of the service, which can:

- Displace a range of other payment instruments, with non-real-time payments, cheques, and other P2P payments likely to be displaced first.
- Increase efficiencies in the payments system, particularly where they displace high-cost instruments such as cheques. Real-time payments can also release money locked up in the financial system and boost competition.
- Open up financial institutions to include more people and more transactions. Over time, consumers will benefit from gaining access to other financial services, while governments will benefit from the ability to distribute benefits and collect taxes more quickly and accurately.

The mix of benefits in any specific country will depend on how real-time payments are implemented – as well as the specific economic conditions in, and demographic characteristics of, the country concerned. This study should provide a valuable starting point for those considering taking the next step in that journey.

Country scenarios

The scale of change delivered by real-time payments is illustrated by comparing three hypothetical economies. The annual impacts after five years are as follows:

Country 1: High-income
- Reduction in costs of up to USD 87m
- Reduction in float value of up to USD 15bn
- An improvement in tax receipts of up to USD 117m

Country 2: Middle-income
- Reduction in costs of up to USD 464m
- Reduction in float value of up to USD 13bn
- An improvement in tax receipts of up to USD 22m

Country 3: Low-income
- Increase in costs of up to USD 1bn (which might be unwound over time and reflect challenges in accounting for all the costs associated with cash)
- Reduction in float value of up to USD 7bn
- An improvement in tax receipts of up to USD 71m

1
Introduction

Real-time payments allow people, businesses, and governments to make payments more quickly with funds immediately available for use by the recipient. Compared with some legacy alternatives, which can take days to reach a recipient, real-time payments offer a faster and more predictable means of payment, potentially creating a range of economic and social impacts that affect consumers, businesses, and governments. Vocalink commissioned Deloitte to analyse the extent of the expected impacts and how they might vary depending on the circumstances (e.g. the country) in which real-time payments are introduced.

There is a broad consensus that real-time payments schemes represent best practice for future payment systems (The Bank for International Settlements, 2016; European Central Bank, 2018; Faster Payments Task Force, 2017; Deloitte, 2015; EY, 2018; HSBC, 2018).

However, while the existing academic literature has considered the effects that new payments options can have on the 'payment mix' (Bolt et al., 2008; Hataiseree & Banchuen, 2010; Trütsch, 2014), the relatively recent introduction of real-time payments schemes in many countries means there is more limited quantitative evidence on the effects of real-time payments in particular on other payment instruments.

Equally, while there are studies on some of the potential economic implications of changes in the 'payment mix' (particularly a shift from cash or cheques to various forms of electronic payments), there is no study that draws the link from the introduction of real-time payments through the impact on the 'payment mix' to the most salient economic and social outcomes.

Finally, while the impact of introducing real-time payments has been studied on a country-by-country basis (e.g. Green et al., 2014), there is no general analysis of the impacts of real-time payments and how they are likely to vary by country based on the full evidence base now available.

This study aims to fill the gap in the literature, quantifying how real-time payments should be expected to affect the mix of payment types and the wider social and economic implications of this shift in the 'payment mix'. It first develops econometric estimates of the impact of real-time payments take-up on the usage of existing payment instruments, based on historical data from countries that have introduced real-time payments thus far. This analysis includes controlling for the economic and structural characteristics of the different economies considered. Using the resulting estimates for take-up and the expected impacts on the existing 'payment mix', the study quantifies, where possible, the economic implications of these 'payment mix' shifts and outlines a number of important qualitative impacts based on the existing evidence, further statistical analysis, and new expert survey evidence.

The aim of this study is to help policymakers understand the likely impacts of introducing real-time payments for a given set of macroeconomic and structural circumstances. It aims to provide an indicative guide to the economic and social impacts of introducing real-time payments based on the experience of countries that have already done so. This can provide an initial guide for those considering its introduction prior to the development of a more detailed business case based on the specific implementation planned.

It is important to note that this study is not:

- a study of the business case for banks or other commercial stakeholders. The question of which of these benefits might be monetised (and by whom) is outside of the scope of this study; or
- a replacement for a detailed study of the likely impacts of a specific introduction of real-time payments, with a specific set of supporting services, and in a specific macroeconomic context.

The goal is to provide a robust starting point for those developing a country-specific understanding of the likely economic and social impacts of the introduction of real-time payments. Policymakers can then work with commercial stakeholders and consider what further analysis is needed to support a decision to go ahead.

The structure of this report is as follows:

Section 2
Provides an introduction to payment systems, and real-time payments in particular; outlines the global take-up of real-time payments to date, the sample of instant payment schemes considered in this study, and an analysis of the drivers of real-time payments take-up. It also discusses the qualitative features of implementation that may further drive real-time payments take-up.

Section 3
Analyses the impacts of real-time payments take-up on the volumes of transactions for other payment instruments. This analysis utilises an econometric model, drawing on the academic literature and economic theory, to allow for estimation of the 'payment mix' in the years following the introduction of real-time payments, including the changes in payment instrument volumes.

Section 4
Outlines the different social and economic impacts that may be expected from the introduction of real-time payments, with a review of the literature associated with each type of impact and, where quantifiable, a methodology for considering these impacts. This section discusses:

- The efficiency gains that may result from instant payment schemes reducing net payment system costs.
- The value of the money that may be unlocked through improvements in transaction speeds.
- The link from faster and more certain transactions to improved business outcomes.
- The potential for real-time payments to support efforts to reduce the shadow economy and, as a result, increase tax receipts.
- The linkages between improvements in electronic payments technologies through real-time payments and financial inclusion in developed and developing countries.
- The ability of real-time payments to serve as the platform for a range of innovative applications, each with its own resulting social and economic benefits.

Section 5
Presents indicative estimates for a set of hypothetical economies, applying the approach developed in this study, demonstrating how the impacts studied are likely to vary.

Section 6
Considers how the results can be applied and identifies avenues for future research.

2
Real-time payments and global take-up

The Bank for International Settlements (BIS, 2013) defines the overall payment system of countries as the set of payment instruments, banking procedures, and interbank funds transfer systems that enable money to circulate. Each payment is comprised of a number of steps:

1. Authorisation of payment/initiation: To initiate a payment, an authorisation must be made by the payer to pay the payee. This may be initiated by individuals, businesses, banks, government, charities, etc. For some payment instruments, this may also be pre-authorised by the payer, with the payee initiating the payment (such as for direct debits).

2. Clearing: Following authorisation, a number of activities must be undertaken prior to payment settlement that together denote "clearing" activities. These are the "transmitting, reconciling, and, in some cases, confirming transfer orders prior to settlement, potentially including the netting of orders and the establishment of final positions for settlement" (BIS, 2003).[1] Clearing is generally undertaken by a payment scheme operator, which acts as an intermediary between the payer and the payee's financial institutions.

3. Settlement: Settlement refers to the discharging of payment obligation between the payer and payee's financial institutions, with the transfer of funds occurring on a gross basis (individually, by payment) or on a net basis (after offsetting all obligations between the financial institutions). The latter allows for settlement in batches to settle, or transfer, a smaller net payment obligation.

4. Notification: This refers to the notification to the payer and payee's financial institution that a payment has been made, with this notification passed on by the respective financial institutions to the payer and payee.

5. Release of funds: This refers to when funds are made available to the payee.

For some retail payment instruments, these steps may be condensed into one step, such as for cash transactions, combined to varying degrees, or remain clearly delineated (see Figure 1), although payment authorisation/initiation is required for all payments. For some payment instruments, the order of these steps may be different, for example some payment instruments allow for notification to the payer and payee of the transaction immediately following authorisation.

Figure 1 – Illustrative steps to payment for traditional payment instruments

CASH

1. In a cash transaction, all steps of the payment process, from **authorisation** to **release of funds** to the payee, are realised at once.

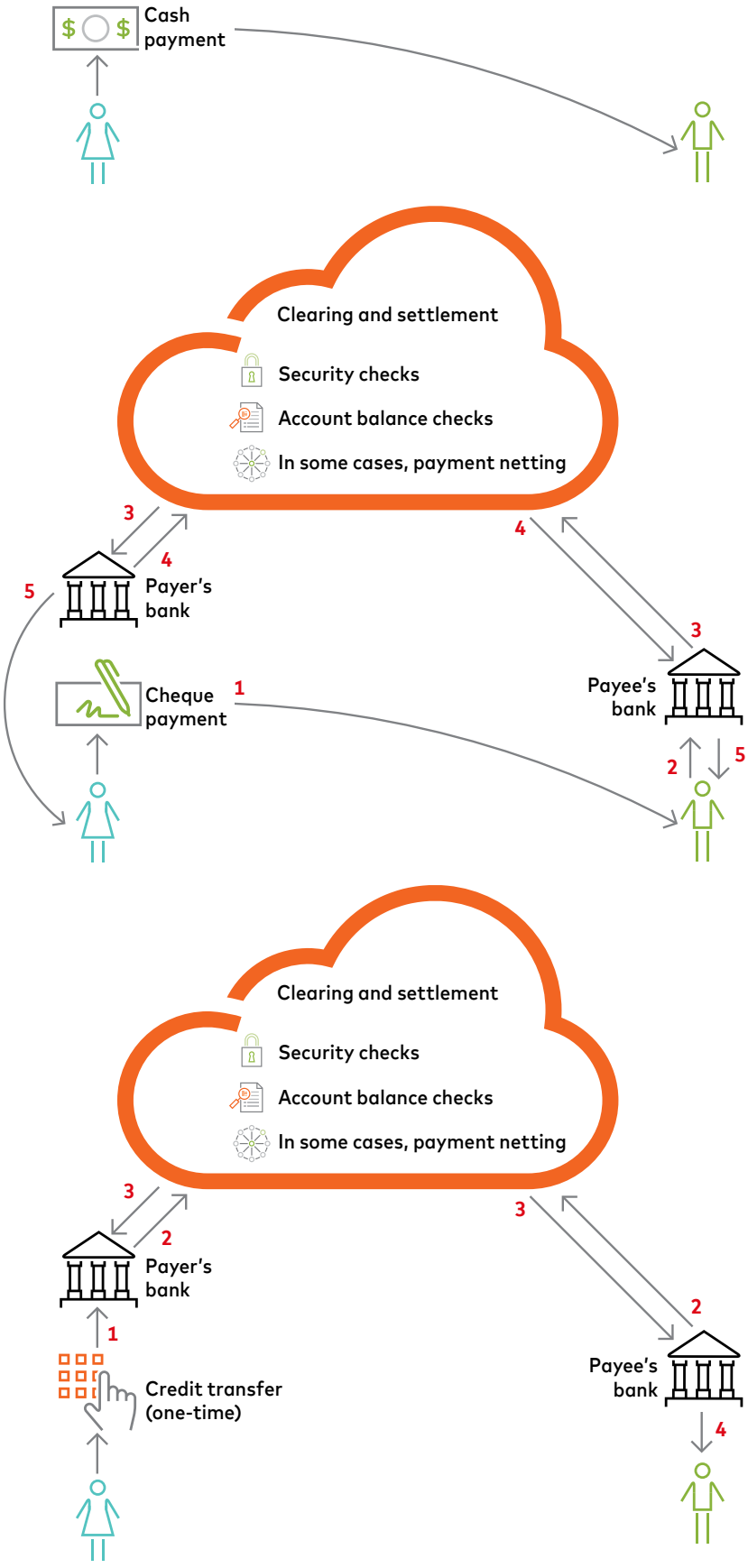
CHEQUE

1. **Authorisation** made by drafting the cheque, with **first notification** of payment by providing it to the payee.
2. Payee deposits into payee's bank.
3. Payee's bank interfaces with the payers bank through the cheque **clearing and settlement** system, generally through a central bank.
4. Banks **settle** via payments through a counterparty, such as the central bank.
5. Payer's bank **debits funds** from payer's account and payee's bank **releases funds** to payee, with banks providing **final notification** of payment to their respective customer.

CREDIT TRANSFER

1. **Authorisation** made by inputting account information into bank's interface. Payer's account may be **debited** at authorisation.
2. Payer's bank interfaces with payee's bank via a payments processor, e.g. Bacs in the UK and FedWire in the US, who act as a counterparty for **clearing** and **settlement**.
3. Banks **settle** via payments through the payments processor.
4. Payee's bank **releases funds** to payee and provides **notification** of payment.

Source: Deloitte illustration



2

In comparison to traditional retail payment methods, a key feature of many real-time payments schemes is that not all steps of the payments process need to be met prior to notification and non-revocable release of funds to the payee.

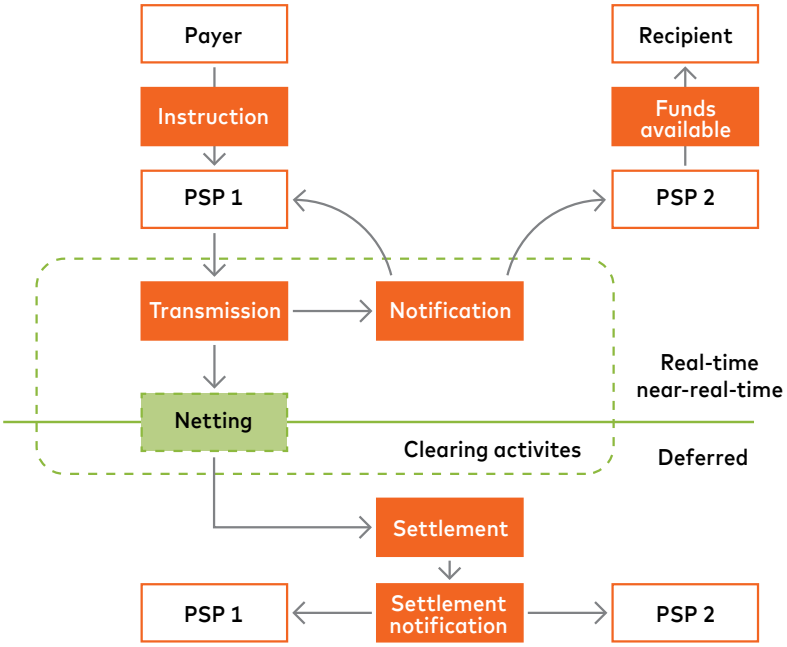
For example, the UK's Faster Payment Service and South Korea's Electronic Banking System operate on a batched, deferred net settlement system but make funds available to the payee immediately following clearing. This allows the schemes to make payments available within a few hours, with most schemes processing and posting payments to payees within a few seconds (BIS, 2016).[2]

Other systems, such as Sweden's Payments in Real-time ("BiR") scheme and Mexico's SPEI scheme, use real-time settlement either on a gross or net basis, with the latter variant netting in a "high number of very short settlement cycles so that settlement can take place in close to real-time" (BIS, 2016).

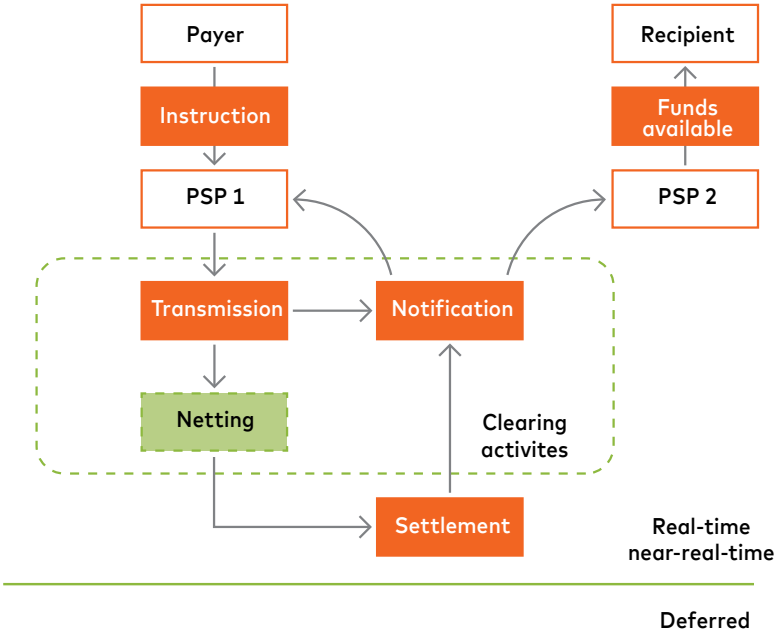
Alternative variations or add-ons may also exist within these systems, such as a pre-funding model whereby banks deposit retail balances that allow for a simulated real-time gross settlement ("RTGS") system (Vocalink, 2016). Figure 2 demonstrates sample models of two instant payment schemes.

Figure 2 – Sample models of instant payment schemes

Model 1 – Deferred settlement



Model 2 – Real-time settlement



Source: BIS, 2016

RTGS systems for high-value payments, such as CHAPS in the UK and FedWire in the US, also allow for same day or real-time payments that are non-revocable (BIS, 2016; Federal Reserve, 2014). However, the key advantage of instant payment schemes is that they make real-time, non-revocable payments accessible to a wider range of consumers and businesses, particularly for lower-value payments. This is due to the near or full 24-hour, 7-day (24/7) availability of the system and the lower payment system costs per transaction compared to more costly RTGS systems. Excluding initial costs of investments, advances in technology, such as more widely available computers and mobile devices for initiating and receiving payments, and cumulative investments in payments infrastructure to date mean that more modern instant payment schemes are less costly for users, payment service providers ("PSPs"), and payment infrastructure operators (BIS, 2016).

Real-time payments schemes present clear benefits to users in terms of a system that can be used flexibly by individuals, allowing for: (a) person-to-person ("P2P") transactions, such as payments between friends and family; and (b) person-to-business ("P2B") transactions, such as from consumers to sole traders or to utilities and landlords. In general, these would otherwise have to be made by costlier payment systems such as cheques or legacy electronic payment systems (see Sections 2 and 4.1 for further discussion).

In addition, the real-time and irrevocable nature of the payment scheme ensures that businesses can make payments flexibly and reliably closer to payment deadlines, easing potential constraints on working capital and the need to hold cash contingencies. Together, these features may also offer the opportunity for innovative third-party applications and use cases, such as on-demand payroll services and faster public transfers to individuals with instant payment schemes serving as the infrastructure on which these applications are built. The benefits of instant payment schemes are explored in Section 4.

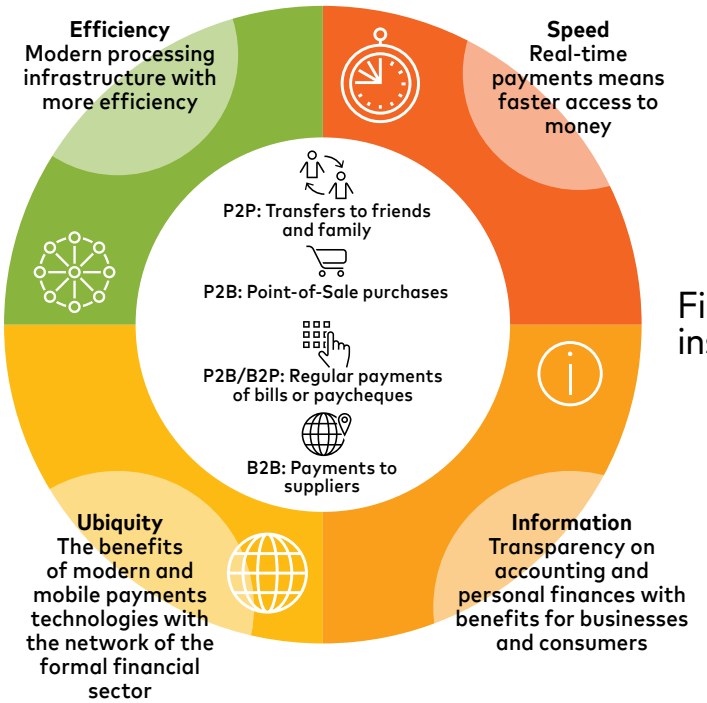


Figure 3 – Benefits of the instant payment schemes

2

2.1
Instant payment schemes in operation today

BIS defines instant payment systems as those "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-hour and seven-day (24/7) basis as possible." BIS also notes that "the characteristics of [instant] payments may also vary by jurisdiction" and that borderline cases may exist, with some systems not initially considered to be fully "real-time payments schemes" developing to meet the criteria (BIS, 2016).

BIS therefore defines the following countries among its Committee on Payments and Market Infrastructures ("CPMI") to have instant payment schemes:

Table 1 – Instant payment schemes in CPMI countries

Country	Implementation	Year commenced[3]	Notes
Australia	New Payments Platform ("NPP")	2018	NPP was launched in February 2018 (New Payments Platform, 2018).
China	Internet Banking Payment System ("IBPS")	2010	
Hong Kong	Faster Payment System ("HK FPS")	2018	HK FPS was launched in September 2018 (Hong Kong Monetary Authority, 2018).
India	Immediate Payment Service ("IMPS")	2010	
Italy	Jiffy	2014	
Republic of Korea	Electronic Banking System ("EBS") / CD/ATM System	2001 / 2007	*"The CD/ATM System has provided near-real-time payments since 1988 with operations on a near-to-24/7 basis…since 2007"* (BIS, 2016).
Mexico	SPEI	2015	*"The SPEI began conducting near-real-time payments in 2004 with operations on a 21/7 basis for mobile payments since March 2015 and on a 24/7 basis since November 2015"* (BIS, 2016).
Selected countries in the European Union	SEPA Credit Transfer instant ("SCT Inst")	2017	SCT Inst is a voluntary real-time payments system for PSPs in the European Union and selected European countries. It initially launched in 2017 with participation by PSPs in four countries: Austria, Estonia, Latvia, and Spain. Since introduction, PSPs in an additional 13 countries have joined the scheme, including from the United Kingdom and Sweden which already operate separate instant payment schemes.
Singapore	Fast and Secure Transfers ("FAST")	2014	
South Africa	Real-Time Clearing ("RTC")	2006	
Sweden	Payments in Real-time ("BiR")	2012	
Switzerland	Twint	2017	In its 2016 report, BIS notes Twint as being launched in 2015; however, SIX Group, the central financial infrastructure provider in Switzerland, reports that Twint was launched in April 2017.
Turkey	BKM Express	2013	
United Kingdom	Faster Payments Service ("UK FPS")	2008	
United States	Real-Time Payments ("RTP")	2017	

Source: BIS, 2016; The Clearing House, 2018; New Payments Platform, 2018; Hong Kong Monetary Authority, 2018; European Payments Council, 2017; European Payments Council, 2018.

A number of additional CPMI countries have payment schemes that are borderline, for example by providing real-time or near-real-time payments but only during working hours. These systems include:

Table 2 – Payment schemes that provide features similar to instant payment schemes

Country	Implementation	Year commenced	Notes
Brazil	System of Funds Transfer ("SITRAF")	2002	Brazil's SITRAF provides near-real-time payments with transaction speed at less than one minute. However, it does not have 24/7 or near-24/7 availability, operating only during working hours.
Japan	Zengin	1993	Japan's Zengin payment system was introduced in 1973 and has provided near-real-time payments since 1993, but this is only available from 08:30 to 15:30. From November 2018, the system will offer 24/7 availability.

Source: BIS, 2016; BIS, 2011; Edgar Dunn, 2018.

Finally, the late 2010s have seen a number of non-CPMI countries introduce or begin plans to introduce instant payment schemes. These include Thailand's PromptPay introduced in 2017, Kenya's PesaLink introduced in 2017, and Canada's Real-Time Rail planned for 2019 (Vocalink, Accessed 24 September 2018; Kenya Integrated Payments Services, 2017; Payments Canada, 2018).

2

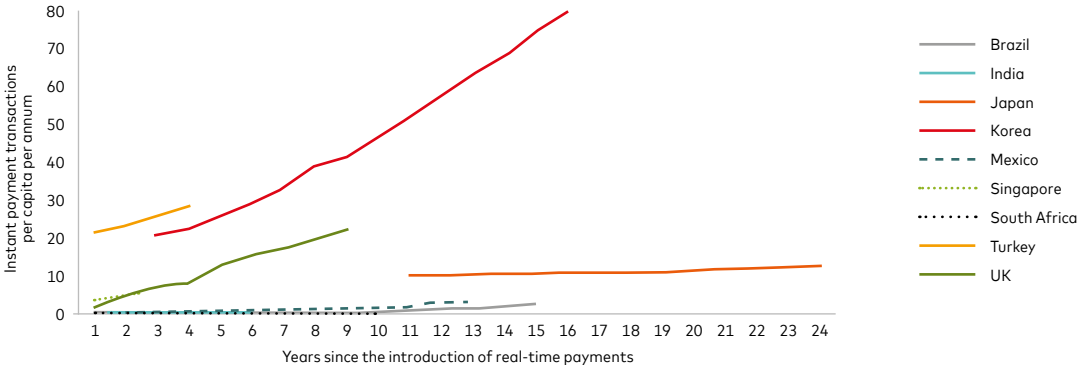
2.2
Estimating the take-up of real-time payments

For the purposes of the econometric analysis in this study, real-time payments schemes are defined according to the characteristics identified by BIS (2016), with some allowances regarding 24/7 availability of the scheme to obtain a sufficient sample size in analysis. In total, nine real-time payments or real-time payments-like schemes are identified for CPMI countries – those in Brazil, India, Japan, Korea, Mexico, Singapore, South Africa, Turkey, and the UK. Table 3 notes the name and data source for each of the schemes and Figure 4 illustrates the take-up of real-time payments in the nine countries.

Table 3 – List of instant payment schemes introduced as at 2016 and with data available for analysis

Country	Real-time payments scheme	Data source
Brazil	SITRAF	Banco Central do Brasil
India	IMPS	Reserve Bank of India
Japan	Zengin	BIS
Korea	EBS, CD/ATM	Bank of Korea
Mexico	SPEI	Banco de Mexico
Singapore	FAST	BIS
South Africa	RTC	Payment Association of South Africa
Turkey	BKM Express	BIS
United Kingdom	FPS	BIS

Figure 4 – Take-up of real-time payments services selected for analysis[4]



Note: Data starts in 2003, hence gaps at the start of some series. Source: Deloitte analysis.

Those historic trends suggest that the take-up of real-time payments follows an adoption curve over time as more consumers and businesses are able to access and start to use the scheme, with high levels of penetration for supportive technologies (e.g. mobile, broadband) and features enabling greater user access also playing important roles. Turkey and the Republic of Korea, for example, have substantially higher usage, and in the case of Republic of Korea, growth rates for real-time payments relative to other countries in the analysis. In contrast, the Zengin service, which was introduced in 1993 and has progressively transitioned to include real-time capabilities (BIS, 2016), still has material limitations (e.g. not available 24/7) versus real-time payments systems that may have constrained adoption.

Using an Ordinary Least Squares ("OLS") method to analyse the impact of economic, structural, and temporal drivers, it is estimated that instant payment transactions per capita increase by 0.59 every year following implementation. Further, the number of broadband subscriptions in a country is estimated to have a particularly positive impact on the take-up of real-time payments, suggesting that high broadband penetration may be a key driver of take-up on a macro level, as is the case in Korea and Japan.

The take-up of real-time payments is driven by a combination of organic growth, broadband penetration, and the displacement of other payment instruments. Abstracting from the impact of continued macroeconomic growth and the associated increase in overall transactions in the economy, take-up may plateau over time once these schemes mature and are fully embedded in the 'payment mix'. The estimate of organic growth of 0.59 transactions per capita reflects the growth rate in relatively new instant payment schemes. Given that most instant payment schemes are relatively new, it is not possible to fully ascertain how take-up may evolve in the long-run. Nevertheless, the estimates suggest that real-time payments are generally an attractive service for consumers and businesses, with organic growth likely driven by end-user expectation and demand for a more modern and faster payment system (BIS, 2016).

Figure 4 also suggests that take-up of real-time payments in the sample countries can be categorised into two groups, roughly according to their income levels. Higher-income countries such as the UK appear to have had a high take-up of real-time payments in the years following introduction. Lower-income countries such as Brazil and South Africa appear to have had a lower take-up of real-time payments. Due to the small sample size of countries with instant payment schemes, the econometrically predicted take-up for countries is likely to provide a middle-range prediction between high take-up and low take-up, which is a poor guide to the expected impact of introduction in a given country, even after controlling for the characteristics discussed above.

As such, in estimating indicative impacts in Section 5, "high" take-up and "low" take-up scenarios are developed to provide an indicative range of the expected impacts from take-up of real-time payments in the areas discussed in Section 4. These scenarios are more or less likely to be realised based on whether features maximising real-time payments take-up are implemented, as described below.

2

2.3
Key features for maximising real-time payments take-up

Real-time payments take-up may be further promoted based on the specific characteristics of the real-time payments scheme's implementation. In addition, investments in the financial and technological supporting services and landscape can stimulate additional take-up and help to maximise the schemes' potential benefits. Important factors include (BIS, 2016; Faster Payments Task Force, 2016):

Coverage and openness of the schemes
Schemes introduced with wider coverage will maximise adoption and usage by both businesses and consumers, and therefore the potential use cases for the schemes. This reflects the network effects associated with payment schemes more generally, whereby the presence of more participants makes the schemes in question more valuable to each participant.

There are a range of potential limitations on the coverage or openness of real-time payments, including:

- non-participating financial institutions;
- closed systems run by banks, or mobile network operators in the case of mobile money systems (see Section 4.5.1); and
- the restriction of real-time payments to a narrow set of transactions, e.g. P2P but not P2B or B2B.

Any such restrictions will act as a deterrent to wider adoption or limit potential use cases, with the scheme itself being less valuable to each participant due to reduced network coverage. The US's Faster Payments Task Force ("FPTF") noted the need for a collective action and collaboration to enhance the benefits resulting from real-time payments (Faster Payments Task Force, 2016).

Geographical limitations may also act as a limitation on coverage and a brake on innovation. To date, most implemented instant payment schemes have been national systems and limited to domestic payments, with only recent initiatives to push for regional integration such as in Europe's SEPA Credit Transfer instant ("SCT Inst") scheme (BIS, 2016). National systems limit the opportunities for developments in cross-border payments by individuals, such as for remittances or transfers to individuals' bank accounts in other countries, and by businesses, such as to international suppliers or from international clients.

Access channels and ease of use
Many instant payment schemes have demonstrated the importance of access through a range of potential devices and generally making usage easy to drive adoption. As a new technology system, barriers to, or frictions in, usage that make the scheme inconvenient or hard to use may act as a drag on adoption.

In a survey conducted by Deloitte as part of this study, a respondent from a consumer-facing business in a country where a real-time payments scheme had been introduced noted that *"bad UX [user experience design] of…bank websites and people struggling with the knowledge of how to"* make electronic payments can pose significant challenges to adoption and usage. Cheaper communications and information technology devices, such as computers, mobile devices, and particularly smartphones, are essential to ensuring a real-time payments scheme is adopted and used, but innovations to improve access and ease-of-use for the schemes themselves also promote take-up.

Examples from both developed and developing countries demonstrate the importance of access channels and ease-of-use:

- In the UK, although real-time payments had been introduced in 2008 and P2P payments using the scheme were technically possible via mobile banking, a 2013 Vocalink study found that only 20% of consumers had made mobile payments and that 39% of consumers did not believe mobile banking methods were good enough (Vocalink, 2013). Following the introduction of Paym in 2014, which reduced the information required to make P2P payments from the recipient's account number/sort-code to their mobile number, payments using the new service saw year-on-year growth from 2015 to 2016 of 259.22%, from approximately 775,000 to over 2 million (Paym, 2016). Consumers highlighted the improvements in ease-of-use for P2P mobile payments, with 27% highlighting the ease of using a phone number rather than account information, 24% highlighting the ease of receiving payments, and 20% suggesting that Paym meant "one less job to do" (Paym, 2016).
- Historically, Thailand has been heavily reliant on cash, with 30% of the population completely unbanked compared to mobile penetration at 137% in 2014 (McKinsey, 2015).[5] In 2017, Thailand launched its National e-Payment Master Plan, which aimed to reduce cash usage and modernise its payment infrastructure. As part of this initiative, the Bank of Thailand initiated the PromptPay service to drive growth in real-time e-payments, making it easier for individuals and businesses to use e-payment services by taking advantage of high mobile penetration rates. Similar setups have also been adopted by other instant payment schemes, such as in the instant payment scheme in Singapore (BIS, 2016).

Similar to the UK's Paym, the service allows the linking of bank accounts to proxy IDs such as phone numbers, national ID numbers, company registration numbers, or email addresses, eliminating the need to share bank account information with payers (Vocalink, accessed 25 September 2018; MobileTopUp, 2018). This has allowed for a significant growth in e-payments in Thailand. From 2016 to 2017, the first year of introduction, growth in the volume of mobile payments was 110%.[6] Additionally, smaller businesses who have implemented e-payment options have increased sales by 17% while larger businesses have increased sales by 22% (ASEAN Today, 2018).

Real-time payments can therefore offer a number of potential social and economic benefits to countries, but it is essential to consider the characteristics of implementation. Schemes must ensure the widest possible coverage and access, maximising scale in order to ensure payment systems are able to benefit from the efficiency, speed, availability, and informational benefits of these types of payments.

Technological and market context

Take-up is also likely to be higher in economies where the wider context enables use of real-time payments. Technological factors might include:

- Mobile penetration.
- Broadband penetration.
- Digital services, particularly mobile banking.

Market factors might include:

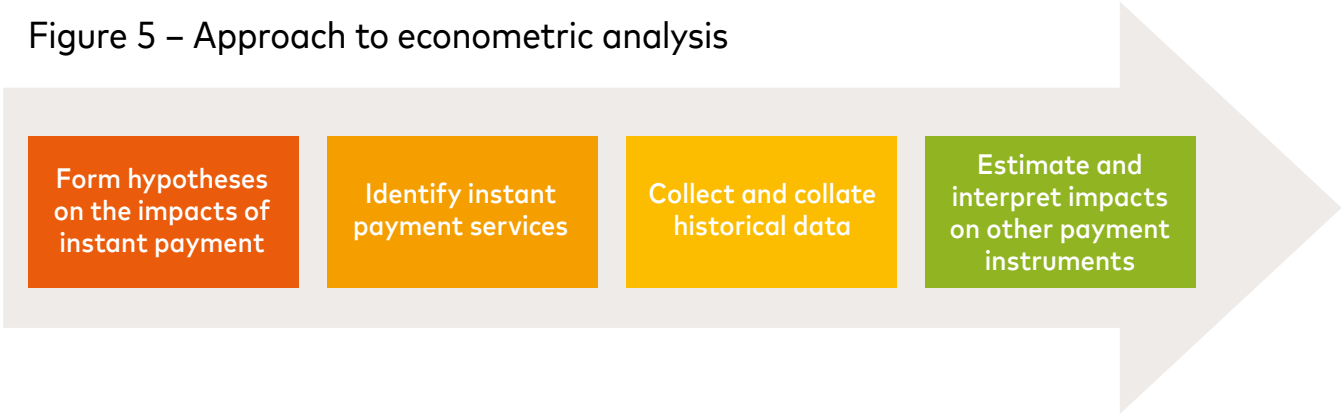
- Competition in affected B2C sectors (e.g. skilled trades; retail with a POS solution).
- Quality of other payment instruments, (e.g. time taken to process cheques, limitations on use of cash).
- Connectivity costs, e.g. data rates. While the volume of data required per person for real-time payments is low, this is likely to affect the use of mobile banking and wider e-commerce.

3
‘Payment mix’ impact

In order to analyse the economic and social impacts of real-time payments, its relative benefits need to be considered versus the likely alternative. Real-time payments may be expected to act as a substitute payment instrument displacing other payment instruments. As such, a key step to considering the impacts of real-time payments is estimating the impact of take-up on the transaction volumes of existing payment instruments and therefore the overall ‘payment mix’.

This study employs an econometric approach to estimating this impact, as illustrated in Figure 5 below. This analysis is conducted at the country-level and aims to identify the impact that an increase in the per capita use of instant payment has on other payment types, while controlling for other economic and demographic factors that affect the use of different methods. By doing this, it allows for estimation of how take-up of instant payment displaces existing payment instruments and how this can have benefits.

Figure 5 – Approach to econometric analysis



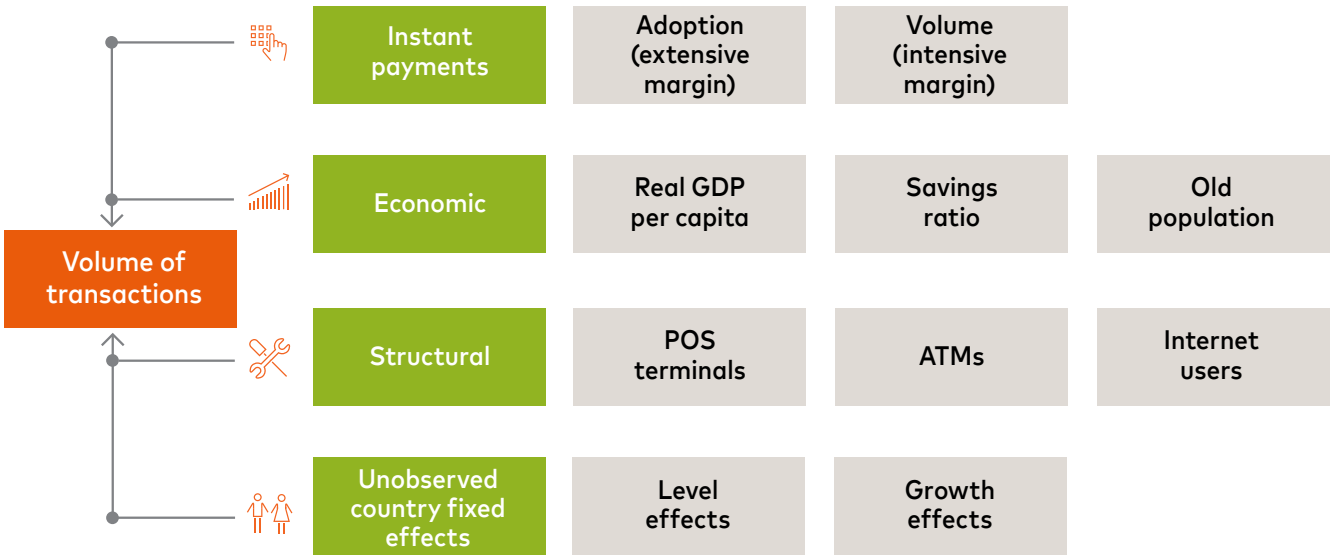
3

3.1
Overview of econometric approach

Section 2.2 outlines the nine CPMI countries with real-time payments schemes, or boundary schemes, for which data is available. A dataset for these countries and additional countries that have not introduced real-time payments schemes was constructed using historical data from BIS, the World Bank, Euromonitor, and various central banks.[7] This dataset includes historical, annual data on instant payment transaction volumes in countries that have adopted real-time payments, transaction volumes for existing payment instruments, and economic and structural variables. In summary, the econometric analysis utilises data covering 20 countries over 26 years (1991 – 2016), with real-time payments volume data available for countries that had introduced instant payment schemes as at 2016.

The estimation of the impacts of real-time payments on the existing payment instruments is undertaken at the per capita level to control for population effects. A Two Stage Least Squares ("2SLS") approach is used to account for various economic and structural variables, unobserved country-specific characteristics, as well as the possibility that the take-up of real-time payments is itself driven by countries' economic and structural characteristics (see Figure 6). Further details on the econometric methodology are provided in Section A Econometrics in the Technical Appendix.

Figure 6 – Drivers of transaction volumes



Source: Deloitte analysis

3.2
Outline of impacts on existing payment instruments

The payment instruments considered in the analysis include cash, cheques, non-instant credit transfers, and non-instant direct debits.[8] For each of these payment instruments, the analysis considers:

- the features of the relevant payment instrument;
- similarities with and differences to real-time payments;
- a resulting hypothesis for the impact of real-time payments take-up on transaction volumes; and
- the econometric results, including directional impact and statistical significance.[9]

3.2.1
Non-instant payment credit transfers

Definition

Similar to real-time payments, a non-instant payment credit transfer is a payment order or a sequence of payment orders to transfer funds to a payee. This can be completed as part of a single immediate payment ("SIP"), a repeating, standing order payment ("SOP") or as forward dated payment ("FDP"), with payments debited at the beginning of the process cycle and credited once funds are settled. Payment instructions are made by a payer to his or her financial institution with details of the payee. Funds are generally made available to the payee after the completion of the payment process steps described in Section 2 (BIS, 2003). Non-instant payment credit transfers can be paper or non-paper initiated, with the former typically requiring the payer to visit a retail bank branch to execute the payment and the latter involving either online or mobile banking.

Credit transfers can be processed through one or multiple wholesale payment systems, some of which may offer the ability for real-time payments but are only economical for higher-value payments due to higher costs (Green et al., 2014).[10] In the UK, the CHAPS system for same-day high-value payments and the Bacs system for lower-value payments with three-day processing cycles can both be used for credit transfers. In the US, the FedACH offers same or next business day payments on batched basis for low-value payments, while the FedWire RTGS system for high-value payments offers "immediate, final, and irrevocable" payments at higher costs with 21/5 availability (Federal Reserve, 2014; Federal Reserve Financial Services, 2017).

Link to real-time payments

Credit transfers as a retail payment instrument more generally include real-time payments schemes. As such, real-time payments are similar to non-instant payment credit transfers in many use cases and functionalities, particularly when used through online and mobile banking. However, the key difference lies in the payment process to enable real-time or near-real-time, low- to medium-value "instant" payments from the perspective of the payer and payee, compared to non-instant payment credit transfers which are either slower or non-economical for low-value payments.

Given these similarities and differences, real-time payments can be expected to displace lower-value credit transfers made through systems such as Bacs in the UK and FedACH in the US, particularly given the similar functionality of the payment instruments and the large improvements in speed and availability of real-time payments.[11] However, it could be less likely to displace payments that would have used RTGS systems such as CHAPS in the UK and FedWire in the US if there are limits on transaction values for instant payment schemes.

Econometric results
The econometric analysis supports this hypothesis, with an additional instant payment per capita associated with a significant and less than proportionate decrease in non-instant payment credit transfers per capita. The analysis estimates that for every ten additional instant payment transactions per capita, the number of non-instant payment credit transfers per capita decreases by approximately four on an aggregate level. While the model does not analyse an individual payer's choices on a micro-level, or distinguish between high-value and low-value payments, the results align with the potential for real-time payments to substitute for low-value payments on a one-for-one basis as well as for single, high-value payments with multiple lower-value real-time payments.[12]

3.2.2 Cheques

Definition
A cheque is a written order requiring the payer's financial institution to pay a specified sum on demand from the payer's financial account to a specified payee, with this instruction typically deposited by the payee's financial institution and the transaction settled through a clearing and settlement system (BIS, 2003). Cheques are generally used to make medium- or high-value payments without having to resort to large amounts of cash, although in some countries, such as the US, they have been commonly used for low-value payments as well (Federal Reserve, 2002).

As such, cheques may be deemed a secure payment method as the responsibility of executing payment and ensuring security lie within the financial system's clearing processes after cheques have been deposited. In addition, cheques can be used to settle debts and withdraw money from banks (BIS, 2003). The heavy use of cheques in certain countries, such as the US, may be attributed to network effects from legacy postage systems (demonstrated by cheque usage being less common in many countries with historically less-developed postage systems), inertia in consumer and business preferences among payment methods, and economic reasons such as the monetary cost of interbank credit transfers relative to cheques (Leibbrandt, 2010).

Link to real-time payments
Similar to cheques, real-time payments offer the ability to make secure low- and medium-value payments. However, a key difference between cheques and real-time payments is the time taken from payment authorisation and initiation to settlement and posting. For example, cheques require two working days to be processed in Brazil, while this process could take up to six days in the UK (BIS, 2011).[13] As such, cheques are a time-inefficient payment instrument. The implications of this inefficiency, such as the payment float, are discussed in further detail in Section 4.2.

In contrast, an area where cheques may confer more flexibility than real-time payments schemes in some countries is transaction value limits. In some schemes, real-time payments may not be used to make high-value payments due to limits set by respective schemes or by individual banks, whereas there are effectively no limits on the transaction values for cheques. For example, in the UK, HSBC sets a daily limit of £250,000 for business FPS transactions, while the limit for FAST transactions in Singapore is S$50,000 (Faster Payments, accessed 24 September 2018; MoneySENSE, 2018). Based on the difference in processing time, real-time payments may be expected to partially substitute for cheques to the extent that such transactions are within the value limits (and other factors such as inertia).

Econometric results
In line with this hypothesis, the econometric analysis finds that additional instant payment transactions per capita are associated with a significant decrease in the volume of cheque transactions per capita. An increase of ten instant payment transactions per capita, controlling for economic, structural, and country-specific drivers, is associated with a decrease of c.6 cheque payments per capita.[14]

On an aggregate level, the analysis suggests that real-time payments displace cheque transactions less than proportionately. For example, multiple medium-value instant payment transactions for charity donations may displace a single high-value equivalent cheque transaction. Instant payment features such as real-time posting and notification, as well as the greater ease of making payments efficiently and securely, are likely drivers of the displacement of low- to medium-value transactions. Other features of real-time payments relative to cheques (e.g. their lower system cost, covered in Section 4.1) will not directly affect consumer behaviour (unless consumers are charged for those additional costs) but could drive financial institutions to promote real-time payments more than cheques.

3.2.3 Cash

Definition
Cash is ubiquitous and remains the most common payment instrument in many countries, despite the growth in electronic payments. The use of cash may be driven by a lack of other options (particularly in less developed economies): behaviours such as the use of cash as a physical instrument to facilitate budgeting, or a preference for its anonymous and instant nature (Iazzolino & Wasike, 2015). However, cash payments incur costs such as storage and security costs and the inconvenience of withdrawing cash from ATMs (Kalckreuth et al., 2014). Cash is typically used as a point-of-sale ("POS") solution for low- and medium-value transactions, in particular P2P and P2B payments. As the transaction value increases, the large number of physical cash notes required may become inconvenient to transport and secure.

Link to real-time payments
Similar to cash, real-time payments offer a real-time or near-real-time method for making payments. However, the majority of instant payment schemes have not yet incorporated a POS solution. This means the displacement of retail cash transactions is likely to be limited. The substitution of cash (or cheque) payments to independent traders may be more likely.

Real-time payments are more likely to substitute for cash P2P transactions and this is normally the principal initial use case. However, the data on cash transactions obtained for the econometric analysis only covers P2B transactions due to difficulties in recording P2P cash transactions.[15] This means that the analysis may not capture the main impact on cash transactions.

The adoption of real-time payments requires penetration among payers and payees (BIS, 2016) as both parties need to have bank accounts capable of real-time payments for a transaction to take place. In turn, this requires, and therefore encourages, a level of financial inclusion (see Section 4.5), which is not required for cash transactions.

Econometric results
The estimated impact of real-time payments as per the econometric analysis was found to be negative but insignificant. This is likely to reflect a combination of the lack of robust and comprehensive data on P2P cash transactions and the general lack of a POS solution. In order to account for the potential that this reflects a shortfall in the data, we consider scenarios for a moderate impact on cash transactions in the later results.

3.2.4 Direct debit

Definition
A direct debit is a pre-authorised payment instruction by a payer, allowing the payee to set a payment amount and initiate payments. Once this arrangement is set up, funds are deducted automatically from the payer's bank account on a set date, with the payee allowed to set the value of the funds to be deducted (BIS, 2003).

This electronic automisation of payments makes direct debits favoured for paying bills such as utilities and credit cards compared to cheques or credit transfers more generally, as payers are able to ensure that payments are made in a timely manner (BACS, 2017). In contrast, some forms of credit transfers may not offer some or all of the same availabilities. For example, credit transfers made as a SIP or FDP require the payer to authorise and initiate each payment, increasing the possibility of missed payments due to human error. SOPs, in contrast, do not offer the flexibility to adjust payment values for bill types that may not be consistent (e.g. credit card bills). However, direct debits also offer less control to payers as opposed to payees, therefore limiting control over budgets in some cases.

Link to real-time payments
As real-time payments schemes have generally been implemented as a form of credit transfer, their introduction may be expected to have minimal impact on the usage of direct debits without additional technological solutions to address the limitations of credit transfers discussed above. In some cases, these limitations are already being addressed in some real-time payments schemes. In 2019, the UK's FPS plans to introduce the "Request to Pay" feature, offering payees the ability to minimise human error by bill payers while improving the control and flexibility for payers through options to pay a requested amount in full or in part or to respond with a request for more time (Faster Payments, accessed 26 September 2018).

While technically possible, instant payment schemes have only been used for direct debit payments in limited cases thus far. China's IBPS is among the few real-time payments systems that allows for real-time direct debit payments processing.

Econometric results
Without technical solutions to address the limitations of credit transfers for paying bill type obligations, the differences in functionalities and use cases between real-time payments and direct debits suggest that instant payment schemes are unlikely to displace direct debits. The econometric analysis supports this hypothesis, with an additional instant payment transaction per capita associated with a negative but insignificant decrease in direct debit transactions per capita.

3.2.5 Summary

Overall, for the instant payment schemes included in the sample analysed, econometric results demonstrate that there is a statistically significant displacement of cheque transactions and credit transfers from the take-up of real-time payments. There is less statistical evidence of an impact on direct debits or cash, although this is likely due to either the lack of suitable data on transactions in the case of cash, or the lack of technological solutions specifically targeting use cases for substitutions in the sample countries in the cases of both cash's POS uses and direct debits.

However, this does not mean that real-time payments may not substitute these transactions. If sufficient technological solutions are implemented to ease access and usability in relevant use cases, as discussed above and in Sections 2.3 and 4, real-time payments may still be expected to displace these payment instruments.

Table 4 – Summary of econometric analysis

Payment method	Econometric result
Cash	Statistically insignificant – possibly due to data issues
Cheque	Negative and less than proportionate displacement
Non-instant payment credit transfer	Negative and less than proportionate displacement
Direct debit	Statistically insignificant

4
Social and economic impacts

Real-time payments have a number of advantages over alternative payment instruments in a range of use cases. The most relevant benefits (at an aggregate level) vary based on the alternative: real-time payments are cheaper and faster than cheques; faster than non-instant electronic payments; and more transparent than cash. The benefits of real-time payments and, in turn, the motivations for introducing them, should therefore vary based on the economy in question.

This section of the report considers a range of potential benefits and how they might affect different kinds of economies. It then provides a quantification of the scale of the impact where possible. This includes considering a range of potential outcomes, such as:

1. The efficiency gains that result from instant payment schemes reducing net payment system costs.
2. The value of the money unlocked through improvements in transaction speeds.
3. The link from faster and more certain transactions to improved business outcomes.
4. Support for efforts to reduce the size of the shadow economy.
5. The links between improvements in electronic payments through real-time payments and financial inclusion in developed and developing countries.

Real-time payments can also serve as the platform or underlying infrastructure for a range of value-added services and innovative use cases, developed either by the real-time payments provider or third parties. Those services each have their own social and economic benefits in addition to the direct impacts of real-time payments schemes themselves. Examples, and an assessment of the resulting additional economic impacts, are provided throughout this section.

Given the relative novelty of real-time payments schemes in many economies, these case studies should be considered examples of the innovations and use cases possible, with more likely to arise over time. While the additional impacts of these use cases are highly context-dependent (and therefore not quantified here), they represent equally legitimate motivations for the introduction of real-time payments at a social level.

4

4.1
Reducing the cost of payment systems

The costs of different payment methods are important to policymakers, with high costs not associated with commensurate benefits to consumers reflecting a pure inefficiency. In the last two decades, there has been a large shift in payment methods from paper-based methods to electronic payments (Hayashi and Keeton, 2012). Several studies have been conducted by central banks and academics to understand what these shifts have meant for costs of the different retail payment instruments.

This literature on the cost of payment systems focuses on what is called the "social cost" to the economy.[16] These are the system costs incurred by all sectors and exclude transfers (such as fees or interest) between them. Examples of such costs include the cost of the production of money to the central bank, time lost in obtaining cash or making a payment, and the cost of processing a cash or cheque transaction at a retail bank (Krüger and Seitz, 2014). Together, they represent the net payment system costs to the economy arising from the use of resources such as capital and labour.

Net payment system costs, and in particular net payment system costs per transaction, differ across payment methods given the varying significance of fixed and variable elements of costs. On a per transaction basis, they can allow for estimation of the impact of changes in the 'payment mix' after controlling for the importance of fixed and variable elements. It is thereby possible to estimate the impact of real-time payments on overall system costs.

4.1.1
Literature on the cost of payment systems

Estimates of net payment system costs, both in aggregate and on a per transaction basis, depend on the specific characteristics of the economy concerned and the scale of use of each payment method (Hayashi and Keeton, 2012; Schmiedel et al., 2012; World Bank, 2016; Kruger and Seitz, 2014). This is largely driven by differences in countries' intensity of use of each payment instrument, which affect the degree to which economies of scale have been realised.

For example, cash transactions have large physical infrastructure costs, such as building and operating an ATM or bank branch network, and relatively low direct, variable transaction costs. But the cost of handling transactions increases with larger values of transactions due to additional costs such as security (Schmiedel et al., 2012). This therefore suggests a low variable cost with volume and a high variable cost with value. If cash is used very frequently for low-value transactions in a country, net system cost per transaction, or "unit social cost" in the literature, would be relatively low, although total net system costs would be high.

In comparison, electronic payments have relatively lower fixed infrastructure costs and higher fixed operating costs. Further, variable costs are low and not impacted by volume and value of transactions. However, if a country's frequency of usage is relatively low, this can result in relatively higher net system costs per transaction than cash due to the emphasis on fixed costs.

As a result of these differences, the relative ranking of payment methods' net payment system costs per transaction can vary across countries depending on the intensity and scale of use of each payment method. Several central banks and academics have conducted studies of the costs of retail payment methods in their respective economies. The literature notes the presence of economies of scale in all payment instruments to varying degrees, and therefore the importance of the "volume" variable cost element to net system costs per transaction is the common theme.

A study conducted by the ECB provides among the most comprehensive comparisons to date of the net payment system costs of different payment instruments across countries, rather than within a single country as in most studies, using data from 13 national central banks in the EU and extrapolating to the rest of the EU (Schmiedel et al., 2012). This study shows that, given its relatively high intensity of use in many EU countries, cash transactions have on average the lowest net payment system cost per transaction.

However, detailed studies on particular European countries with low cash usage suggest this is not always the case. In Sweden, net payment system costs per transaction for cash were higher than those for card payments in both 2002 and 2009 (Guibourg and Segendorf, 2007; Segendorf and Jansson, 2012). In 2009, net payment system cost per transaction for cash in Sweden was EUR 0.78 in comparison to EUR 0.42 for debit cards in 2009 (Segendorf and Jansson, 2012). The same study showed that cash comprised 34% of all transactions and 41% of total payment system costs, whereas debit cards comprised 43% of all transactions but only 28% of total payment system costs.[17] Given the economies of scale in the cost of cash and the downward trend in cash usage in Sweden, this suggests that costs per transaction for cash will increase further unless there is a commensurate reduction in the cost of the cash-supporting fixed infrastructure network. Sweden's experience may represent a better guide to how costs are likely to vary between payment instruments in the future, as the electronic share in transactions rises globally, than studies of cash-intensive economies today.

It also seems likely that the costs of cash, particularly the time spent in processing it and the resulting opportunity costs, are the most likely to be underestimated. These costs are harder to estimate than the formal prices paid for electronic payments and less likely to be monitored.

4.1.2 Real-time payments and payment system costs

In a 2014 study of the potential benefits of introducing real-time payments in the US, the Federal Reserve found that the UK's Faster Payment Systems cost GBP 0.14 per transaction in 2008, the first year following introduction, and decreased with increased transaction volumes to GBP 0.02 per transaction in 2014 (Greene et al., 2014). These estimates included the installation costs of development and maintaining the required infrastructure, although it did not include all resource costs as in Figure 7.

Nonetheless, this suggests that net payment system costs per real-time payments transaction may be lower than that for other payment methods, and in particular non-electronic payment methods. Therefore, the introduction of real-time payments would likely reduce total net payment system costs. The potential savings from substituted volumes can be estimated by comparing net payment system costs per transaction for real-time payments and for the mix of other payment methods substituted.

4

Figure 7 – Fixed and variable cost elements of payment systems[18]

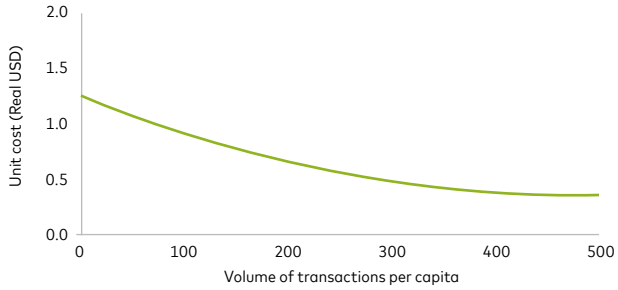
Type of costs	Cost category	Cost element	Variability of cost element	
			Fixed	Variable
Resource	Time	Travel time		✓
		Waiting		✓
		Transaction time	✓	✓
		Reconciliation		✓
		Operations	✓	✓
	Logistics	Travel costs		✓
		Communication	✓	✓
	Infrastructure	Maintenance	✓	
		Depreciation	✓	
	Damage	Error		✓
		Theft		✓
		Fraud		✓
		Insurance	✓	
	Production	Production costs of paper forms		✓
		Production costs of cheques		✓
		Production costs of cards		✓
	Cost of funds	Holding costs		✓
Transfer	Service costs	Periodic fees	✓	✓
		Per transaction fees	✓	✓
		Fees for services outsourced	✓	
	Cost of funds	Float costs		✓

Source: World Bank, 2016.

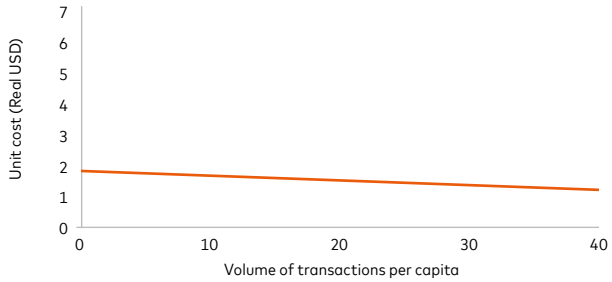
Given the importance of fixed and variable cost elements to estimating net payment system costs, it is essential to consider how these would adjust when estimating the impact of changes in the 'payment mix' (Krüger and Seitz, 2014; Hayashi and Keeton, 2012).[19] Similar to the approach taken by the ECB, a best-fit relationship is estimated between net payment system costs per transaction and volumes per capita for each payment method (Schmiedel et al., 2012). These curves allow for approximation of net payment system costs per transaction for the existing payment method and the real-time payments method before and after the change in 'payment mix', controlling for average fixed and variable costs by payment method.[20] Figure 8 illustrates the estimated relationships between net payment system costs and volumes per capita for cash, cheques, and credit transfers.[21]

Figure 8– Unit payment system costs vs volume per capita for different payment methods[22]

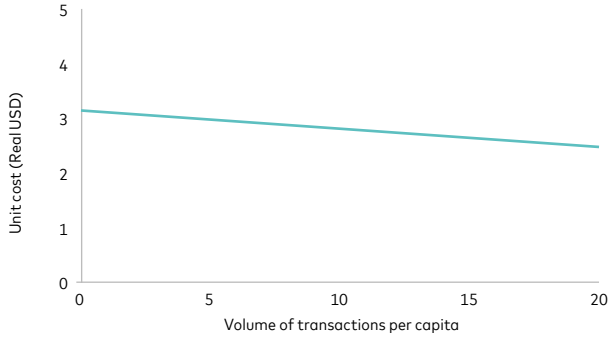
Cash



Credit transfers



Cheques



Source: Deloitte analysis based on data from the World Bank, FRED, and the studies in Appendix B1 Payment system costs.

Using the curves estimated above, the new unit payment system costs for each payment type are calculated based on the reduced number of transactions, after considering the substitution impact that takes place due to the introduction of real-time payments. The impact on net payment system costs for each payment type is then calculated using the new estimate of unit payment system costs and the estimated (reduced) number of transactions.

There is limited relevant data on the cost of real-time payments, particularly in prospect for economies that introduce real-time payments in future. The safest assumption outside of a detailed business case is that the costs are similar to non-instant credit transfers. There are some factors which increase costs in a real-time system (e.g. 24/7 availability) but others that reduce them (e.g. improved data). Therefore, changes in total net payment system costs are driven by the difference between credit transfer costs and non-electronic payment costs, particularly cheques. Refer to Appendix B1 for further details.

Using the methodology described above, the unit payment system costs at a level of ten transactions per capita for each of the payment instruments would be:

- Cash: USD 1.21
- Cheques: USD 2.79
- Credit transfers: USD 1.95

The cost of credit transfers will fall relative to cash, in particular, to the extent volumes increase further. Note, however, that these are ongoing costs and do not reflect migration costs in the transition, which are less salient for the economic impacts and best considered in a specific country business case.

4

4.2
Improving efficiency in the financial system

When payments are made using non-cash payment instruments, there may be a delay between when payments are debited from the payer and credited to the payee.[23] This delay has important implications for both financial institutions and their customers, creating an economic inefficiency.

Economic inefficiencies as a result of payment delays may include agents adapting payment behaviour and timing to accommodate the payment system: for example, a payer may make an early payment in order to account for processing time. Payment delays may also mean that either the payee or payer's financial institution may have additional funds in their corresponding bank deposit accounts, artificially inflating the level of bank reserves as money is being held in the financial system. Finally, payment system delays may result in working capital being locked in the financial system and unavailable to consumers and businesses, limiting short-term consumption or investment or requiring that this be financed at a cost.

Real-time payments schemes can play an important role in addressing these inefficiencies. Real-time or near-real-time payments can reduce the value of payments locked in the system and ensure that economic agents to do not engage in inefficient payment behaviour. Money would be unlocked and made available to consumers and businesses, with wider macroeconomic benefits in terms of reduced risk of inefficient market outcomes or failures.

4.2.1
Literature review on inefficiencies in the payment system

Inefficiencies in the payment system, sometimes called the "payment float", have been an important area of research by central banks and macroeconomists. Studies have focused primarily on the role payment float plays in bank recovery of payment system costs, with literature considering the differences in bank internal costs and the direct and indirect pricing of different payment instruments. Others have highlighted the economic costs resulting from inefficiencies in the payment system.

Greene et al., (2014) divide float costs and gains into three types:

- the interest costs to the consumers between the time funds are credited to the receiver's account and the settlement time;
- interest gains to the financial institutions from holding the funds until settlement time; and
- the costs arising from a delay in the time when a payment has been processed.

In many jurisdictions, the interest costs and gains have been restricted prior to the introduction of real-time payments. However, historically, by not getting interest on transaction float, consumers have implicitly paid for bank payment instrument costs, with banks using interest on the float from delaying availability of funds credited to accounts and debiting accounts prior to bill-payment value dates to recoup payment system costs (Bolt et al., 2008). Given that interest returns from payment floats would be on very short-

term deposits (e.g. one to three days), lost interest costs are unlikely to be substantial for any individual payer or payee. In aggregate, however, payment floats could mean significant additional funds on banks' balances, with interest gains and losses more substantial to financial institutions that trade in large volumes or values.

Any interest cost of the payment represents a highly opaque cost to consumers, which may have implications for aggregate system costs as consumers favour payment instruments with higher effort costs than monetary costs. For example, paper-based payment instruments are associated with higher system costs but lower user monetary costs, with most of the user cost coming from effort such as withdrawing cash from an ATM or drafting and mailing a cheque (see Section 4.1), for example, and Humphrey et al., (2006) argue that more direct pricing of payment instruments could increase shifts to electronic payments given cheaper transaction costs.

Beyond the interest cost, the delay itself will have an economic impact. Borzekowski and Kiser (2006) argue that implicit payment system costs, such as availability, convenience, and speediness, can have significant social implications. As a potential financial inefficiency, payment floats may act as a drag on economic activity. Money locked in the system due to the float is not available to be spent. Mile and Tang (2005) describe further direct costs to bank customers, arguing that in addition to lost interest, consumers face a cost from uncertainty around payment posting due to the gap between the timing of debiting and crediting of payments. This reduces their ability to recognise and respond to payment failures and increases the potential for fees from late or incomplete payments. Lacker (1997) also makes an argument, restated by McAndrews and Roberds (2000), that costly attempts to manipulate payment clearing and settlement times create inefficiencies in the system with zero value to society as a whole.

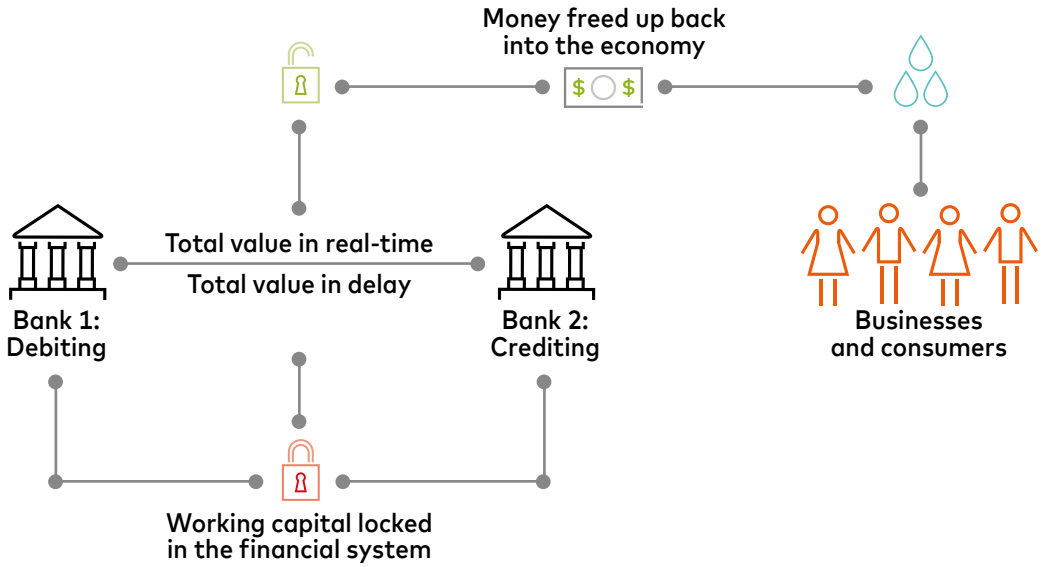
In recognition of these inefficiencies, central banks and policymakers have put forward measures for reducing the float. For example, in an attempt to reduce float, the Federal Reserve has since the 1980s been installing and advocating for new technology to speed the cheque clearing process. This has allowed paying banks to have their cheques scanned and converted into electronic representations (called cheque imaging), eliminating the need for paper cheque processing and allowing for faster processing of cheque payments (Federal Reserve Bank of New York, 2007). In the UK, the Office of Fair Trading requested the elimination of float by mandate for the purpose of improving information flows and helping to avoid late fees for consumers (Office of Fair Trading, 2005; Greene et al., 2014).

4

4.2.2
Real-time payments and the payment float

When a non-instant electronic payment is made, the delay in payment processing results in capital being locked in the financial system and unavailable to either the payer or payee in the transaction. This process is shown in the illustrative figure below.

Figure 9 – Illustrative diagram of the impact of real-time payments on financial system inefficiency



Source: Deloitte

As discussed in Section 4.2.1, money locked in the financial system may potentially limit short-term economic activity and lead to economic inefficiencies. For example, if a payment from a consumer to a business is delayed, that might affect the business' ability to finance working capital (see Section 4.3 for further discussion). On the consumer side, delays in paycheques paid by employers to employees can constrain consumer spending power in the short-term, potentially leading to additional costs in the form of short-term financing products. While the impact of any one delayed payment may be minor, in aggregate this may represent a material economic drag.

If payments are received in real-time or near-real-time due to real-time payments, the total float value may be reduced, with less money locked in the system and an increase in the velocity of money. Even if regulations limit the revenue banks may earn on interest or short-term investments, unlocked money can represent a boost to short-term spending and aggregate economic activity and help to reduce inefficiencies in user choice of payment instruments.

The impact of reducing financial system inefficiencies can be calculated by estimating the total daily value of money locked in the system for each payment method, and therefore money freed up, given the substituted float value, due to the take-up of real-time payments. Using the reduced float value, a further quantification of benefits can be estimated by applying a weighted average delay and using discount rates for individual countries generally assumed to account for the overall time value of money. This is complicated in practice, however, by the considerable variation in the time taken for transactions of different kinds to be processed.

As a thought experiment to consider how the scale of the inefficiency with money unavailable to ultimate end-users might be quantified, we can assume a one-day mismatch between debiting and crediting times. Data from BIS suggests that the value of daily cheque transactions in Germany was USD 654m in 2016 (low given the scale of the economy concerned). If we use Germany's daily discount rate as a reasonable proxy for the time value of money, the time value associated with that float would be USD 74,000 a day.[24] The reduction in the value of this payment float on cheques can then be calculated by applying econometric estimates of the volume of cheque payments displaced from take-up of real-time payments, representing the money that is unlocked for economic use. More detail on this estimation framework is provided in Appendix B2.

CASE STUDY ON PAYROLL: SUPPORTING THE GIG ECONOMY

Canada, the US, and India

Payment platforms are central to what is now called the gig economy – where a payment for every service replaces the traditional, long-term labour contract. With the majority of gig workers being paid by direct deposits, any delay in receiving their money on time may also be a cost to employers as workers gravitate to platforms that allow them to be paid more conveniently (PYMNTS, 2018). In a recent study on gig platforms, 84% of surveyed employees reported they would do more gig work if they were paid faster (PYMNTS, 2018).

ZayZoon, which operates in Canada and the US, is aiming to address this issue. Combined with Push Payments' real-time payment services, ZayZoon offers an on-demand payroll service which can reduce financial insecurity among gig economy workers (Finextra, 2018).

Real-time payments capability also supports customer demand for services in the gig economy. For example, Ola – an Indian taxi aggregator– use the Universal Payments Interface ("UPI") API in its app to allow real-time payments for booking journeys. Ola has also partnered with other payment companies in the market such as Mobikwik and PhonePe in order to boost usage and increase its share of passenger volume in the domestic cab hailing company (The Economic Times, 2018).

4

4.3
Increasing businesses' ability to finance working capital

In addition to the lost economic activity from money locked in the financial system, slower payments may reduce the availability of money for businesses to use as working capital. Studies point to the need for businesses to be able to finance working capital to support employment growth and reach optimum production levels. Delays and uncertainty in payments may limit the ability to finance working capital internally, with many businesses needing to hold additional cash contingencies on balances to cover working capital requirements. These cash contingencies may be associated with opportunity costs in terms of investment, or be associated with direct costs from financing.

Real-time payments schemes may therefore ease these restrictions through faster processing of payments from consumers or other businesses. This may have knock-on impacts in terms of either increasing the capacity of companies to invest directly or through reducing their external financing costs. In addition, businesses of all sizes may be able to improve cash management by reducing inefficiencies related to the inability to see real-time balances, reducing costs and further improving their ability to invest or avoid unnecessary financing costs.

4.3.1
Literature review on the importance of financing working capital

Literature considering the theory of the firm has increasingly sought to consider the role of working capital in firm growth and development through the capital investment and labour expenditure decision. Generally, working capital is needed by businesses to cover costs of operations before revenue is received. Recent studies have noted that economic models excluding the dynamics of working capital financing in firm decisions may understate the predicted effects of financial constraints on production efficiency, firm profit, and growth over time (Dao and Liu, 2017; Chan, 2010).

Chan (2010) highlights the importance of financing working capital through internal credit on the firm's investment decision and firms' ability to reach optimum production levels, as financially constrained firms tend to decrease investment in order to focus on current production, causing a trade-off with future production. Small- and medium-sized enterprises ("SMEs") in particular are noted to be more adversely impacted by financial and other obstacles (Beck et al., 2007), with financial constraints having the largest negative impact on firm growth. Other studies emphasise the need for firms to finance a portion of working capital in advance of production completion using external financing, particularly in developing countries (Dao and Liu, 2017).

These studies have also noted that firms need to finance working capital, and this affects their ability to create employment, with better access to finance allowing firms to hire more labour and increase overall employment growth (Dao and Liu, 2017). The authors argue that this effect is particularly large for smaller firms, more labour-intensive production firms, and for some sectors with a generally greater need to finance working capital. Ayyagari et al., (2016) make a similar argument, showing that increased supply of credit results in higher employment growth, especially for SMEs in developing countries.

In terms of working capital and cash management more generally, real-time information on payments and balances may also play an important role. Real-time information may not only support cash management (FinExtra, 2017), but could also support access to finance for small and large businesses. Nienhus et. al., (2013) note how real-time exchange of payment status information on payable invoices, available on request to creditors and lenders, could allow for SMEs to receive real-time financing on accounts receivable. Mile and Tang (2005) also highlight the impact of reduced uncertainty resulting from real-time information on payments, with recipients able to recognise and respond earlier to payment failures or delays. This would mean that businesses can avoid inefficiencies resulting from the inability to see real-time balances, e.g. the need to hold additional contingencies on balances that may have opportunity costs.

4.3.2
Real-time payments and cash contingency

Businesses normally maintain a cash contingency to cover the expected time for payment crediting and thus ensure that they continue to have the cash required in the meantime. By reducing the time associated with payment processing, real-time payments could ease cash contingency requirements, allowing for greater internal financing of working capital and increased economic activity.

As part of this study, Deloitte conducted a crowd-sourced expert survey of 68 corporate treasury professionals and SME owners across a range of business segments and sectors, both in countries that have introduced real-time payments schemes and those that have not. The survey investigated the challenges and costs associated with processing payments made by cheque or electronic fund transfer, particularly the value of cash contingencies held due to delays in payments.

The survey found that handling and processing cheques was particularly challenging, and that significant benefits could be gained through digitisation of payments. Respondents highlighted that cheques could:

- Create delays in payment, with a significant amount of *"time to issue and to process the money into the company's account"*, according to a respondent from a consultancy in Brazil.
- Require the business to *"deploy people to collect, deposit, and track the processing of the payment. This is the least preferable as a mode of payment as in many cases, it could also give a huge false alarm of sales"*, according to a respondent from a private, skilled-trades business in India.
- Increase the likelihood of "*late paying and... error[s] in payment"*, according to a respondent from a bank in the US.
- Lead to potential lost payments, as cheques *"need to be stored until they can be cashed"* and *"can get lost in the cashing process"*, according to a respondent from a contractor business in Canada.

34%

of respondents think bank clearing has a significant or very significant impact on the time it takes for cheque payments to be credited to their business' accounts

4

Delays in payment were considered to be a significant source of business frustration and cost. A respondent from a construction company in Australia highlighted that late payments were *"time consuming to follow up"* while another construction company in Brazil noted that delayed payments could require them to renegotiate payment terms with creditors. While some businesses noted a strong preference for electronic payments and direct transfers, these were also noted to have issues. A telecommunications company in Australia, which introduced a real-time payments scheme in 2018, noted that the lack of *"instant transfers and a fee for same-day transfers"* presented challenges to collection.

40%

of respondents think having to wait to receive cheques from customers significantly or very significantly impacts payment times for their business

35%

of respondents think the time it takes for consumers to write and send a cheque significantly or very significantly impacts payment times for their business

However, even in countries where real-time payments schemes had been introduced, survey responses suggested the potential for greater adoption. One respondent from a small shuttle services business in South Africa, a country with a real-time payments scheme since 2006, noted that the time taken to receive electronic fund transfers posed challenges to usage, suggesting that greater facilitation, access, and ease-of-use for small businesses may be needed to drive adoption and address challenges with electronic payments.

Systems alignment was also highlighted as an area where electronic payments offered advantages. A respondent in the wholesale and retail products industry noted that electronic fund transfers such as real-time payments *"are faster, more accurate and easier to track. This is crucial as it is important to show the financial reports showing the most up to date information."* However, others argued that other services were needed alongside instant payment schemes to realise the potential benefits. One respondent noted that *"aligning customer bills to payments is always dependent on customers using a reference number that can be automatically aligned by our systems, else it becomes a manual process"*.

30%

of respondents think needing to align incoming electronic payments with outstanding customer bills has a significant or very significant impact on their business' operations

This suggests that even if a payment is made in real-time or near-real-time, the lack of a tracking number could cause delays in following payment due to the need for systems alignment. Other services building on standard instant payment schemes could help to address this issue. In the UK, the planned "Request to Pay" feature, a new method of initiating real-time payments whereby a request for payment is sent by a payee, could allow for easier tracking of when a particular outstanding payment is made while offering additional benefits to payers.

To face these challenges, the survey found that businesses tend to hold a cash contingency for outgoing payments, suggesting potential opportunity costs due to delays associated with receiving payments. Asked to elaborate on the challenges and impacts they face in receiving payments, 43% of respondents thought the risks from the time it takes for electronic payments to be credited to their businesses' accounts after a transfer is initiated were significant or very significant, with 34% of respondents highlighting the same for cheque payments. Due to these inefficiencies, 63% of all respondents reported maintaining a cash contingency to cover the expected time for payment crediting.

To assess the impact real-time payments could have on businesses' finances, respondents were asked to provide the size of the cash contingency their business held to cover the time it takes for payments to be credited to their business' accounts. Normalising these as a percentage of sales, businesses held a cash contingency equivalent to approximately 4.6% of annual sales, implying that a significant portion of businesses' revenue was held as a reserve and could not be used for internal financing of working capital. This figure was higher for middle and lower-income countries, where businesses reported holding cash contingencies equivalent to 6.8% of sales. By using real-time payments, businesses could therefore unlock a significant portion of revenues for expenditure on labour and investment in capital, contributing to further firm growth and development.

The survey also highlighted the impact real-time payments can have in correcting for information inefficiencies when processing payments. In circumstances when cash can be delayed or direct debits cancelled, certainty of receipt of funds becomes important for businesses. The survey asked respondents' views on the willingness to pay for improved information and speed of payments processing, which is associated with real-time payments. 31% said they would pay a set fee per payment or a share of the average payment value to ensure they receive all payments immediately. It is important to note, however, that this reduction in uncertainty for business is accompanied by an increased need to manage risks for banks (albeit one they may be better able to accommodate than other businesses, particularly SMEs).

63%

of respondents report that their business maintains a cash contingency to cover the time it takes to receive payments

31%

of respondents report that their business would be willing to pay some type of fee to ensure they receive all payments immediately

CASE STUDY ON CONSUMER CASH MANAGEMENT:
PROVIDING INCREASED CONTROL OVER PERSONAL FINANCES

United Kingdom

The convenience of making payments instantly and at any time of the day can drive usage by giving people greater control over finances. However, real-time payments are less common as a means to pay bills, where direct debits are more common (giving the seller the ability to vary the amount charged). However, the forthcoming introduction of "Request to Pay" will extend real-time payments into paying for bills and give users more control over how such payments are made.

This service allows people to pay their pre-existing bills at a time of their choosing, paying in part or requesting more time. Customers will have more information and certainty about upcoming individual payments, helping them avoid late payment charges. Businesses might also benefit: Accenture estimates that businesses may be able to save up to 8% on accounts receivable costs through a reduction in the average cost of the invoice-to-pay process (Faster Payments, 2017).

Additional interactions between real-time payments schemes and Open Banking initiatives also offer opportunities for more control over personal finances. For example, in their report on the latest stress testing exercise of the financial industry, the Bank of England proposes that Faster Payment Service combined with the new Open Banking regulations could make it easier for consumers to automatically net off payments across accounts at different banks and therefore avoid overdraft fees (Bank of England, 2017). This could help boost customers' cash resilience through better cash management on accounts and reduce financial risks from becoming overdrawn.

CASE STUDY ON DUTY FEE PAYMENTS:
VALUE-ADD CAPABILITIES TO STREAMLINE EXISTING PROCESSES

Poland

In Poland, Express Elixir, one of the first instant payment schemes in Europe, allows shipping companies to make duty fee payments to the local port authorities in real-time, allowing for a streamlined and efficient turnaround for unloading and loading without the customary delays (FIS, 2017).

The use of Express Elixir to streamline duty fee payments provides an example of the potential for real-time payments to boost economic mobility and activity, with more efficient turnaround for shipping operations. The system has also allowed businesses to increase efficiency in costs incurred by allowing a quicker flow in information and resources between the two sides of any transaction, contributing to greater trust between market participants and encouraging competitiveness (Sadlakowski, 2017).

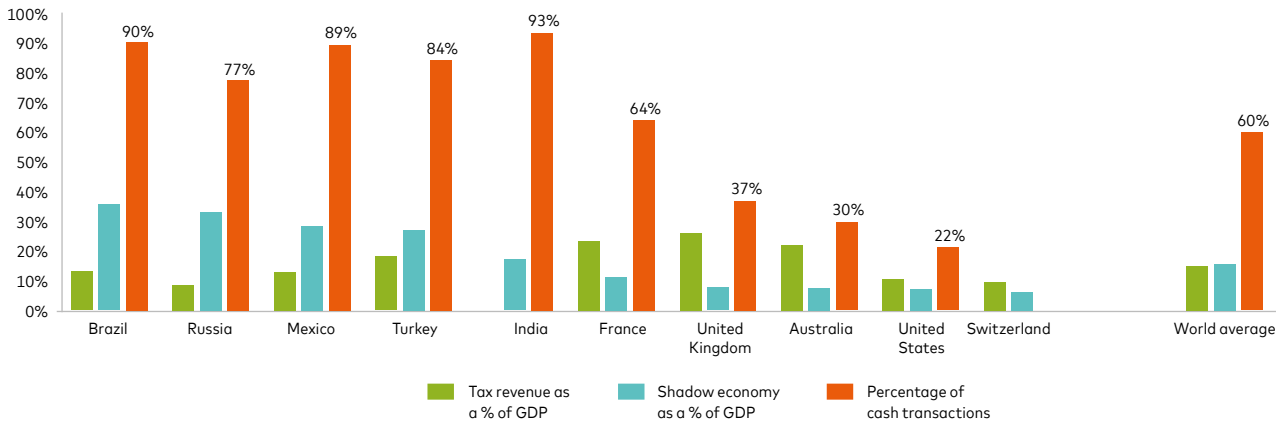
4.4
Supporting a shift to a more formalised economy

The term "shadow economy" refers to work that is either illicit or undeclared, and therefore is untaxed. Some of the adverse consequences of shadow economies include a reduced tax base, distortions in market competition, the degradation of social and economic institutions, and lower economic growth (EY, 2017).

Several studies have discussed the relationship between the usage of different payment types and the extent of the shadow economy. In general, these studies suggest that more use of cash is associated with more tax evasion. Electronic payments are associated with less tax evasion because of the enhanced transparency they provide on informal, although not necessarily illicit, economic activity (with transactions tracked as a matter of course). This section discusses the impact the introduction of real-time payments could have on government tax receipts through a decrease in the usage of cash and therefore a reduction in the size of the shadow economy.

The size of the shadow economy varies around the world. Figure 10 shows the size of the shadow economy for different countries as a percentage of GDP in 2015, as estimated by Medina and Schneider (2018), as well as their tax revenue as a percentage of GDP according to World Bank data. The size of the shadow economy appears to be correlated with levels of economic development: for example in 2015, in richer countries such as Switzerland, France, and Australia, the estimated size of the shadow economy ranged from 7-12%, while in low- and middle-income countries this ranged from 18-35% (Medina and Schneider, 2018). This gives a sense of the potential additional tax revenues resulting from formalisation of the shadow economy.

Figure 10 – Tax revenue and shadow economy sizes as a percentage of GDP, 2015[25]



Source: World Bank; Shadow economy estimates from Medina and Schneider (2018); Cash shares of total transactions based on data from Euromonitor.
Note that Euromonitor data only includes P2B cash transactions. Cash data for Switzerland is unavailable.

4

4.4.1
Literature on the links between payment methods and the shadow economy

The existing literature on the relationship between the payment types and the shadow economy has focused mainly on European countries. The literature suggests a link between cash usage generally and the size of the shadow economy, likely due to the ease with which cash enables unrecorded transactions. A 2017 study by Schneider focusing on European countries used two different methods to investigate the importance of cash usage on the shadow economy. Using a simulation method, Schneider found that when the share of cash payments decreases by 10%, the shadow economy decreases by 2%. Further, this method was used to show that if cash were no longer available at all, the shadow economy would decrease by 20%. Alternatively, using the MIMIC estimation method led to an estimate that a 1% reduction in the cash share of the economy leads to a 0.075% reduction in the size of the shadow economy (Schneider, 2017).

Other studies have used different indicators of reduction in cash to investigate the impact of reducing cash on the size of the shadow economy. Many of these studies have focused on measurement through the take-up of electronic payment methods such as cards. For example:

- One study, based on an analysis of Central and Eastern European economies, found that an increase of 1% in the total value of card payments at physical terminals as a percentage of GDP led to a decrease of 0.037% of GDP in the shadow economy (EY, 2017).
- Another study done in Greece using 15 years of annual data estimated that for every 1% increase in the use of card payments, there is a 0.24% increase in tax revenue (Foundation for Economic & Industrial Research, 2015).
- Based on analysis covering 25 countries in the EU, a 2016 study estimated that 10 additional card transactions per capita per year reduces the VAT gap[26] by between 0.08% and 0.2% of GDP (Immordino & Russo, 2016).

In developing countries, studies have considered the impact of more prevalent forms of electronic payments as an alternative to cash. Existing literature has considered, for example, the impact of mobile payments take-up on cash, by far the most prevalent payment method in many developing countries. A 2012 study by the Bankable Frontier Associates found that even after the introduction of M-Pesa, 99% of all retail transactions in Kenya in their sample were carried out in cash.[27] The study showed that the 1% of transactions that were carried out using mobile money tended to represent the high-value transactions. While the average transaction size of a mobile money transaction was USD 75, more than 90% of all transactions in the study were less than USD 10 and two thirds of the transactions in the study were found to be below USD 3 (Collins et al., 2012).

Several studies have explored reasons for cash being "sticky" in the context of Africa and the introduction of mobile payments, often with a focus on Kenya and the introduction of M-Pesa. Besides financial reasons for cash not being completely substituted by mobile payments, such as fees associated with mobile payments, many studies have explored behavioural factors in relation to cash being widely used. These include factors such as cash allowing an individual to mentally account and budget better, cash being anonymous, and transactions made in cash being instantaneous and easier to keep track of (Iazzolino and Wasike, 2015). Real-time payments may address some of these reasons to continue using cash.

4.4.2
Real-time payments and tax receipts

While studies investigating the impact of real-time payments on the shadow economy and tax receipts are limited, studies investigating the impact of electronic payments more generally may be indicative of the expected impact the introduction of real-time payments could have. Studies looking at electronic payments have found an association between take-up at the expense of cash and the shadow economy (negative) and tax receipts (positive), suggesting that adoption of real-time payments could reduce the size of the shadow economy and increase tax receipts. This may be through direct substitution of cash transactions to real-time payments, or through indirect impacts in the form of substitution of cash for electronic payment instruments more generally – for example, if the introduction and take-up of real-time payments encourages use of more formal banking and electronic systems more widely. Note that the analysis does not account for the latter potential indirect impacts of real-time payments.

In considering the impact of similar payment methods such as mobile money, studies have to date found a more limited impact on cash usage. As discussed in Iazzolino and Wasike (2015), this is likely given the potential fees associated with this service, which increase monetary marginal costs of each transaction for price-sensitive consumers and traders who are engaging in informal but not illicit activities. This limitation may be addressed if real-time payments schemes were to be introduced with an operating model similar to countries such as the United Kingdom, with costless transactions for P2P and P2B transactions driving potentially greater take-up than previous electronic payment alternatives. However, this would also depend on additional factors in implementation and investments made in access and ease-of-use for the scheme. For example:

Availability and access
In the low- to medium-income countries in Figure 10, Mexico, Turkey, and Brazil have all introduced real-time or near-real-time payments schemes as at 2015 but with some limitations on access – the Brazilian SITRAF system does not offer 24/7 availability and Mexico's SPEI scheme only introduced 24/7 availability in late 2015 (BIS, 2016). In contrast, Thailand's PromptPay scheme introduced features that make it easy to access from launch, such as 24/7 availability and usage of multiple different identifiers to register for the system and enabling mobile payments, leading some to suggest that Thailand could make rapid progress in reducing the cash share in spending (ASEAN Today, 2018).[28]

POS facilitation and ease-of-use
Currently, cash usage remains predominant in POS payments in many countries due to its low infrastructure prerequisites. To support digitisation of POS payments, implementation would need to consider integration of POS-facilitating solutions such as Quick Response ("QR") codes, near-field communication ("NFC"), or other technologies. Combined with mobile or broadband penetration, and potentially other incentives such as receipt lotteries, these could allow for better recording of sales in P2B situations where other electronic payments take-up has been low, such as transactions between sole traders and consumers.

CASE STUDY ON DIGITISATION:
OPEN-ACCESS TO SUPPORT CUSTOMER ADOPTION AND INTEGRATION

India

India's rapid uptake and enhancement of electronic payments has encouraged customer adoption, and more broadly promoted economic growth (Lipis Advisors et al., 2017). A primary driver and an example of the importance of innovation and transformation in the customer journey is the development of India's UPI in 2016, introduced on top of its instant payment scheme IMPS. The interface works by instantly transferring funds between two bank accounts on a mobile platform.

UPI ensures open access to support innovation and drive customer adoption by allowing payments to be directly processed in external business applications. According to the 2018 Faster Payment Innovation Index, daily IMPS transaction volumes have risen from about two million transactions per day in 2017 to about 2.8 million a day in 2018, primarily driven by the additional overlay services offered on top of IMPS by UPI (FIS, 2018).

In addition, the introduction of UPI has led to a number of new and innovative payment solutions, combining both the reach in the infrastructure and conversion through innovative functionalities (FIS, 2018). For example, Google has announced the launch of the Tez chat-like interface application built specifically for India on its UPI platform. It allows users to send and receive payments directly into bank accounts of individuals and merchants (FIS, 2018).

Investments supporting access and ease-of-use of these schemes could therefore help drive lower cash usage in countries where cash remains predominant, reducing the size of the shadow economy. Both low-value and medium-value P2P cash transactions could be substituted by the introduction of real-time payments, while integration of POS ease-of-use features as above could support substitution from cash in P2B settings, particularly for the areas that are not illicit but may not be recorded for a number of reasons. In countries still heavily reliant on cash to pay sole traders, this would not only increase the amount of declared economic activity in an economy, which would then increase the tax base, but could also lead to economic growth (by reducing transaction costs and thereby enabling more transactions).

As discussed in Section 3.2.3, adequately estimating the impact of introducing real-time payments, particularly with a POS solution, on the volume of cash transactions in the P2P and P2B space, and as such quantifying the potential impact on the shadow economy and tax receipts, is not possible within the scope of this study. This is due to a number of reasons, in particular:

- The size of the shadow economy countries can only be approximately estimated, most recently based on the method developed by Medina and Schneider (2018).
- Data on cash transactions of interest, such as those that are P2P or with sole and smaller traders in the informal economy, is difficult, if not impossible, to measure. Additionally, econometric estimates may be insignificant due to the lack of POS solutions in many of the sample countries during the period of coverage, reducing the scale of the impact.[29]

- Finally, data on successful implementations of real-time payments in developing countries with high cash usage and shadow economies, and with the appropriate mechanisms for usage to support take-up of real-time payments, remain limited and limits econometric estimation of substitution effects.

Nevertheless, an estimate of the impact on tax receipts can be developed using scenarios for the reduction in the cash share of the economy and coefficients in the literature on the impact of this on the shadow economy. Using the coefficient from Schneider (2017), if the introduction of real-time payments led to a decrease of 10 percentage points in the cash share of the economy, the shadow economy can be expected to decrease by 0.75 percentage points of GDP. Figures for France in 2015 (see Figure 10) suggest that the shadow economy would therefore decrease from 12% of GDP to 11.25% of GDP based on these estimates. Given that in 2015 France's GDP was USD 2.47tn and tax as a percentage of GDP was 23%, this suggests that in this scenario tax receipts could increase by USD 4.26bn.

4.5 Supporting financial inclusion

Financial inclusion is the expansion of financial products and services to unbanked or underserved segments of the population. This includes both formal and informal financial services such as retail banking and mobile financial services, respectively.

Financial inclusion is generally thought to support economic growth through a number of mechanisms, particularly in developing economies (Beck et al., 2007; Ang, 2010; Sahay et al., 2015; Demirgüç-Kunt et al., 2017). Increased financial inclusion in developed countries will still be important for underserved populations (Consumer Focus, 2010), even if the impact on aggregate economic growth is more limited. This chapter discusses the impacts of real-time payments on financial inclusion by using case studies to discuss and contrast the impacts on developing and developed countries.

4.5.1 Developing countries

The impact of real-time payments in developing economies is likely to vary depending on the wider financial services context and the services with which instant payment schemes can compete and/or connect. In this section we consider three examples:

- Real-time payments with a user-focused implementation, with Thailand and PromptPay as a case study.
- The relationship of real-time payments to mobile money, with Kenya and M-Pesa as a live case study.
- Real-time payments as a complement to open banking, with analysis focusing on South Africa as a case study.

A focus on user access: Thailand

In December 2015, the Thai government launched an e-payment initiative, with an aim to create a cashless society. This was an extremely ambitious target given that 97% of Thailand's retail transactions were handled in cash in 2010 (Bank of Thailand, 2017).

As part of this initiative, Thailand launched PromptPay, a real-time payments scheme, in January 2017 to allow consumers to make real-time payments and transfers with a mobile phone using a number of different identifiers for individuals and businesses to sign up, such as national ID numbers, email addresses, mobile numbers, and corporate registration numbers. In addition to pushing for government use of the scheme for transfers to individuals (see case study below), the ease-of-access allowed businesses to easily adopt PromptPay to lower the cost of transactions, and some retailers have been pushed by consumers to adopt PromptPay POS solutions to increase convenience. These features have driven high PromptPay adoption rates, with 36.2 million users, or more than half the population, signed up for PromptPay, many of whom were not previously included in the financial system (Vocalink, 2018a). Of these, 23.5 million users have signed up using ID cards, further increasing transparency in financial transaction for segments of the informal economy.

This has allowed PromptPay to offer important economic benefits. By allowing for quicker and lower cost interbank transactions, PromptPay encourages greater adoption by consumers and allows inclusion of lower-income individuals in economic activity. This can result in a higher velocity of money circulation in the economy, as more people of different income groups use the service (National e-Payment, 2017; Burapongbandhit, 2017). Additionally, through lower transfer services fees, Thailand's PromptPay enhances business competitiveness within the country by increasing opportunities for the business sector to expand trade channels and online payments (Bank of Thailand, 2016).

With mobile penetration expected to increase, additional innovative features being introduced on top of PromptPay, e.g. links to e-Wallets, may in the future increase the payment capability of consumers without bank accounts further (Vocalink, 2018b). This allows real-time payments to be not merely a conduit for formal financial system accounts, but for all banking-type products, potentially driving greater financial inclusion.

Real-time payments and mobile money: Kenya

M-Pesa, introduced in Kenya in 2007, is a mobile money scheme that allows money to be transferred using mobile devices in real-time. It was originally designed to be a system that allowed microfinance loans to be repaid using mobile phones and reduced the costs associated with making cash transactions, with usage expanding to cover a wider variety of traditional formal banking services (The Economist, 2015). M-Pesa uses mobile networks as an intermediary to make cash payments. Users can credit or debit their accounts using licensed agents and can transfer the money stored in their M-Pesa accounts to others.

In effect, mobile money services simulate a payment scheme similar in nature to real-time payments schemes from the perspective of an end-user. Payments are processed in a method similar to intra-bank fund transfers in the formal financial sector, with transfers between mobile money accounts not necessitating an actual transfer of funds and operating instead as adjustments to internal balances for different accounts under the same operator. As the mobile networks for Safaricom, M-Pesa's operator, cover a large majority of the population (with a market share of 67%), this in effect leverages network effects to create a ubiquitous payment system network (Bloomberg, 2018).

CASE STUDY ON PUBLIC TRANSFERS: EFFICIENCIES WITH WIDER SOCIAL IMPLICATIONS

Thailand and South Africa

In a report in 2012, the World Bank suggested that digitising government transfers to individuals could help to make welfare systems more impactful, by:

- *"lowering the cost of disbursing and receiving payments;*
- *increasing individuals' risk management capacity;*
- *increasing the privacy of payments;*
- *increasing control over the funds received;*
- *increasing the security of payments and reducing the incidence of crimes associated with them;*
- *increasing the transparency of payments;*
- *increasing the speed of payments; and*
- *providing a first entry point into the formal financial system."*

More prompt payment of benefits thanks to real-time payments, in particular, can reduce recourse to emergency support (e.g. food banks); improve the ability of claimants to manage their spending; and enhance the flexibility of welfare provision to respond to changing claimant circumstances (in turn improving the potential for claimants to take temporary work, without a shortfall when payments restart).

In addition to other benefits from adoption, the ease-of-access features underlying PromptPay have allowed the Thai government to make progress towards achieving many of the above impacts by digitising government transfer systems and disbursing social benefits to individuals more efficiently while providing faster accessibility for claimants. According to the Bank of Thailand, the Thai government has increasingly transitioned government payments to PromptPay, starting with child support payments and progressing to other welfare payments and tax refunds (Bank of Thailand, 2016). The use of PromptPay for government payments has been a large driver of formal banking adoption and inclusion for many low-income people, helping them to transition into the formal banking system and away from cash (ASEAN Today, 2018).

With a different technical underpinning, but a similar result, the South African Social Services Agency ("SASSA") began distributing social benefits payments to underprivileged, disabled, and retired communities digitally in 2012 (SASSA & MasterCard, 2016). This has been paired with the use of biometrics to validate the identity of recipients, allowing people to receive benefits on their SASSA MasterCard debit card without the need to travel long distances to collect their money or worry about their safety. The addition of biometric technology for cardholder authentication has enabled wider public sector efficiency, accuracy, and transparency, lowering the cost of delivering payments further by reducing fraud in addition to improving operational efficiency (SASSA & Mastercard, 2016). After requiring re-registration for social benefits delivery, the South African government reported that 850,000 social grants were cancelled due to illegal and duplicate grant collections discovered by the system, equating to a total savings of R3.4 billion in 2013/2014 (Carnejo & Madia, 2015).

The use of instant payment schemes may not be limited to government welfare transfers. The Thai central bank is currently developing an e-Donation system in collaboration with financial institutions, NITMX (the central payments infrastructure operator), and the Thai Revenue Department, whereby donations made by specific QR codes linked to the PromptPay system allow for more ease in individual donations and for information to be shared automatically with the Revenue Department for tax deduction purposes (Bank of Thailand, 2018).

The distribution of welfare payments electronically more generally can create added value for merchants through greater consumption. Estimates suggest that retailers who accept the SASSA MasterCard debit cards have benefited from a 30% uplift in sales in the first year following introduction of the SASSA MasterCard scheme due to higher foot traffic and sales, as many of these retail locations are now convenient for consumers to withdraw cash in addition to purchasing products (Carnejo & Madia, 2015). Similarly, under its National E-Payment Master Plan, the Thai government and central bank are looking to leverage the government-driven adoption of PromptPay to support greater use in commerce, particularly through features such as QR codes for PromptPay (KPMG, 2018; Bank of Thailand, 2018).

Some mobile money accounts do also allow for transfers with different mobile money networks. This operates on a system similar to deferred settlement, and consumers can choose to cash-out at any point (similar to withdrawing money from a bank account) by visiting one of the licensed agents that buy mobile money and sell cash in a method similar to cashback. Mobile money services are also moving towards offering many more services that are available in the formal banking sector. For example, similar to formal banking services that are able to use information on microbusinesses and consumer finances to offer loans, M-Pesa's high penetration allows it to offer micro-loans using the M-Shwari product (Cook & McKay, 2015).

Due to the high penetration of mobile networks in developing countries, the introduction of mobile money services has led to a higher degree of financial inclusion. After the introduction of mobile money services, financial inclusion grew from 27% in 2006 to over 75% in 2016 in Kenya, with the number of M-Pesa accounts in Kenya overtaking the number of bank accounts from 2010 to 2015 (Cook, 2017). Given its similarity to instant payment schemes, the popularity of M-Pesa in Kenya therefore demonstrates the impact real-time payments schemes may have on financial inclusion. However, it also suggests that in countries where there is an existing large-scale adoption of mobile money accounts, real-time payments may not directly improve financial inclusion.

Nonetheless, the introduction of real-time payments schemes could offer other benefits as a driver of greater formal banking innovation and, as a result, greater competition for mobile money services, to the benefit of consumers and greater banking access. In a study on the impacts of M-Pesa in Kenya, Mbiti and Weil (2011) found that existing competitors, such as Western Union and MoneyGram, decreased prices in response to competition from the mobile money scheme. Higher quality competition for money transfer services could increase competitive pressure further, leading to further reductions in the quality-adjusted price (Cook & McKay, 2017); currently, Safaricom's M-Pesa charges at least KSH 11 (USD 0.10) for payments to other M-Pesa users, rising up to KSH 97 (USD 0.96) for payments of KSH 10,0001 (USD 99.06; Safaricom, accessed 2 October 2018). In a review of the impacts of the introduction of PesaLink, an instant payment scheme, on the banking sector in Kenya, Cook and McKay (2017) suggest the new scheme has enabled and spurred innovation in the banking sector and increased its focus on low-income consumers. This has encouraged consumers to become part of the formal financial system and accelerated the long term development of formal financial services.[30]

Formal banking services may also fill gaps in the services offered by mobile money operators. For example, while M-Pesa has limited daily transfer amounts to approximately USD 1,400, banks have set a much higher limit of USD 10,000 for real-time payments (Quartz Africa, 2017). While the M-Pesa limit would likely be satisfactory for most low-value, consumer uses in a developing country, PesaLink's higher limit provides an option for small businesses to use real-time payments for medium-value transactions with much more convenience, potentially reducing the size of the informal economy.

Real-time payments and Open Banking: South Africa

Some of the common problems faced by small informal businesses that limit their ability to develop and scale relate to their receiving and making payments in cash – for example, the need for greater security and proximity to suppliers and customers and the lack of electronic records. To the extent that real-time payments increase the propensity to use the banking system, and associated payment instruments, it can help mitigate those problems.

In 2018, Lipis Advisors studied the potential for real-time payments schemes to contribute to addressing these issues and reduce the population of unbanked small businesses in South Africa through links with open banking. Open banking allows third parties to build applications around financial institutions and services that would allow access to data that was previously not available. This allows users greater access to on-demand account information, on-demand payment initiation, and other services such as savings plans, etc., through third-party applications that may be more suited to ease access than those developed through the direct banking institution.

By linking with real-time payments, Open Banking can provide even greater benefits to users. For example, third-party applications, which would be more likely to be adopted due to ease-of-use, can then be used to initiate and receive real-time payments, helping to reduce the delay in receiving payments and providing a guarantee of funds that helps reduce the "trust gap" between users in the informal sector and formal banking (Lipis Advisors, 2018). By providing more incentives to migrate to digitised payments and formal banking (through third-party access channels), small businesses can benefit from better liquidity and cash management through increased flexibility when making and receiving payments and reduced security risk for entrepreneurs who may be hesitant to carry large amounts of cash, giving them greater opportunity to develop and scale their businesses.

4.5.2
Developed countries

Developed countries generally have a much smaller proportion of unbanked individuals compared to developing countries.[31] However, the high rates of financial penetration and inclusion in developed economies does not imply that unbanked populations do not exist in these countries, with important social consequences. Penetration for formal financial services and infrastructure is often lower in certain regions and demographics. In 2017, the share of the population with no formal or informal financial accounts in the US was 7%, with a similar figure of 6-7% of the population unbanked in some developed EU countries.[32] Figure 11 shows the share of unbanked for a set of developed countries as well as the OECD average.

Evidence from the UK suggests that financial exclusion may be related to financial capability and the ability to make budgeting decisions. In a publication prior to the introduction of FPS, the lack of balance transparency for many account-based transactions using traditional electronic payment services was found to be linked to a hesitation to move away from the visibility and physical aspects of cash and towards the use of bank accounts (Consumer Focus, 2010). These concerns may therefore be addressed by real-time payments that allow for real-time transfers and balance updates, and that consumers are required to trust the counterparties to their transactions less in comparison to non-instant electronic payments because they will know more quickly if a payment has not been made. The introduction of real-time payments may therefore play a role in enabling marginal consumers to move away from cash by addressing their concerns. This can in turn draw them more into using formal financial institutions more broadly and thereby encourage financial inclusion.

Figure 11 – Share of unbanked in an economy, 2017

Source: Deloitte calculations based on World Bank data.

CASE STUDY ON MICRO-INSURANCE:
INNOVATIVE PRODUCTS WITH A CONSUMER-CENTRIC FOCUS

China, the US, and the UK

In a 2017 study in the US, Vertafore found that a large majority of young people believed having no insurance was risky but were still less insured than older generations, with particularly high rates of underinsurance for personal items and renters insurance (Vertafore, 2017). Underinsurance was particularly found to be linked to personal finance constraints. Constrained consumers appreciated the flexibility of on-demand services.

To address this gap, technology companies and insurers have begun to offer micro-insurance services aimed at helping younger people receive on-demand coverage (Business Insider, 2016). Key to the provision of micro-insurance services is the ability to provide coverage flexibly, reliably, and quickly. Current iterations have provided real-time claims processing through AI technologies, such as Lemonade in the US, (PR Newswire, 2017) and location-based, micro-term coverage using accurate smartphone GPS, such as Trov in the UK (The Telegraph, 2017).

Real-time payments are likely to be an important factor in improving the reliability and speed of micro insurance services. Real-time, irrevocable payments can provide assurance to the insurer of payment of the premium, while, combined with real-time claims processing through AI, real-time payments technologies can guarantee immediate claims payment. For example, Lemonade was able to process and initiate a payment through bank wire on a claim within three seconds of submission, using AI and without needing any paperwork. However, as real-time payments infrastructure was not available in the US at the time, this payment would have taken an additional day or more to reach the beneficiary (PR Newswire, 2017). With real-time payments, the beneficiary would have received the payment immediately.

China in particular has seen increasing popularity for micro-insurance services. Zhong An, launched in 2013, became the first digital micro-insurance service relying on real-time payments infrastructure. It offers a number of micro-insurance products, such as e-commerce return shipping insurance for less than 50 cents and flight-delay insurance, which, when a claim was made, provides vouchers instantly for redemption in the airport (Ledger Insights, 2018; Financial Times, 2016). Since its launch, Zhong An has underwritten over 630 million insurance policies and serviced 150 million clients in its first year of operation (Majesco, 2018).

5
Country impact scenarios

To demonstrate the potential social and economic impacts of introducing an instant payment scheme in different types of countries, this section uses the methods discussed in Sections 2, 3 and 4 to develop estimates for three hypothetical countries. These countries are specified to have a range of different macroeconomic characteristics, with different levels of usage of the relevant payment instruments. For each of the hypothetical countries, estimates are presented for the annual impact at five years after the introduction of real-time payments, allowing time for market penetration in the country.

The three hypothetical countries have been developed to illustrate indicative impacts for the introduction of real-time payments and how they are likely to vary between countries, with GDP normalised to USD 1.5tn for each of the three countries for ease of comparison. The impacts are then estimated for the hypothetical countries, taking GDP and other macroeconomic characteristics as constant (to control for variation) and assuming zero inflation. The hypothetical countries are:

Country 1
A high-income country, with over 50% of transactions in cash but some usage of credit transfers, direct debits, cards, and cheques.

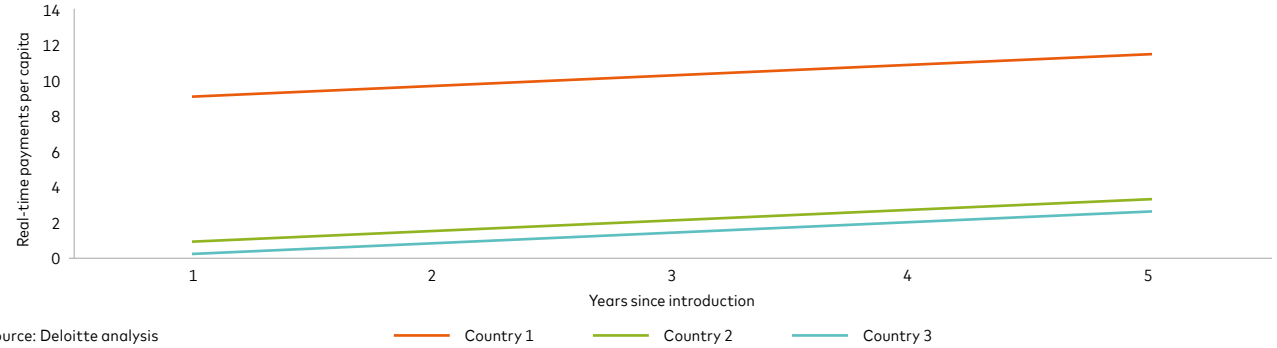
Country 2
A middle-income country, with some usage of cards and credit transfers but minimal usage of cheques and direct debits and a predominant (90%) usage of cash.

Country 3
A lower-income country, with low-to-no usage of non-cash payment instruments.

The macroeconomic and financial sector characteristics of the countries are specified in more detail in Appendix B3: Indicative impacts for hypothetical countries. The econometric results presented in Section 2 suggest that a high-income country with greater internet penetration and formal financial system integration (i.e. Country 1) is likely to see a higher take-up of real-time payments.

To account for how the real-time payments scheme might be implemented, and to reflect the diversity in take-up rates among countries that have implemented real-time payments thus far, estimates are also calculated for "low" and "high" take-up scenarios.[33] This gives a realistic range for the impacts depending on the characteristics of the real-time payments scheme introduced and whether it is supported with investments in potential use cases.

Figure 12 – Predicted take-up of real-time payments over five years following scheme introduction as estimated using the econometric model outlined in Section 2.2



Source: Deloitte analysis

5

5.1
Net payment system cost savings

Savings in net payment system costs are driven by three factors: the degree to which the take-up of real-time payments displace existing payment instruments, the cost of the payment instruments, and population. Real-time payments generate substantial cost savings by primarily displacing the usage of cheques, which constitute the most costly payment instrument (see Section 4.1). Population then becomes a key driver as small per capita displacements are multiplied out.

Estimates suggest that while cash impacts on a per capita basis may be large, as modelled in the analysis, impacts on net payment system costs may be low due to the significant economies of scale and high starting proportion of existing payments. For developing countries with low to no cheque usage and significant cash usage, the high levels of cash transactions means that real-time payments schemes may actually suggest a net increase in payment system costs in the short-term. However, as cash use falls further, the dilution of fixed costs means that costs are more likely to fall overall, while the other impacts become more important.

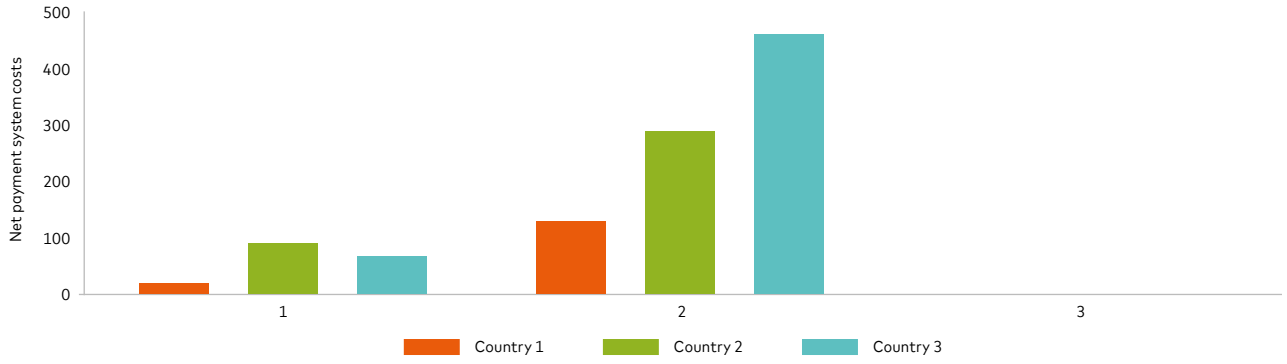
Analysis for each of the hypothetical countries reveals:

Country 1
In Country 1, the hypothetical high-income country, take-up of real-time payments is expected to reduce net payment system costs by between USD 17m and USD 62m, with most of the savings reflecting a reduction in the costs associated with cheques (between USD 262m and USD 703m) partly offset by the costs of real-time payments (between USD 294m and USD 797m). This is despite the relatively low usage of cheques, reflecting the high cost per transaction for this payment instrument.

Country 2
In Country 2, the hypothetical middle-income country, take-up of real-time payments is expected to reduce costs by between USD 132m and USD 464m, again with most of the savings reflecting a reduction in the costs associated with cheques (between USD 413m and USD 1,266m) partly offset by the costs of real-time payments (between USD 441m and USD 1,287m). The greater cost savings relative to Country 1 reflect the larger population of Country 2, with even a small decrease in per capita usage resulting in large savings. In comparison, savings in the cost of cash appear to be low due to the significant economies of scale. This suggests that savings in the net payment system costs of cash may increase steeply as cash use falls further over time.

Country 3
In Country 3, the hypothetical lower-income country, take-up of real-time payments is expected to result in a net increase in net payment system costs. This is because real-time payments are expected to cost between USD 576m and USD 1,690m, driven by the large population and hence large volume of transactions, even at a low per capita usage, with no offsetting savings from reduced cheques usage. Similar to Country 2, the high levels of cash usage and large population mean that savings in the net payment system costs of cash may increase steeply as real-time payments substitute more significantly for cash transactions in the economy over time.

Figure 13 – Indicative annual decrease in net payment system costs at five years after the introduction of real-time payments, USD m



Source: Deloitte analysis. Note that net savings are largely driven by reductions in cheque usage and to a lesser extent cash.
As Country 3 does not have any significant cheque usage, this results in a net positive impact on total payment system costs, not shown here.

5

5.2
Financial system efficiency gains

Efficiency gains, as indicated by reductions in the one-day float value associated with different payment instruments, are driven by the existing average transaction value, the usage rates, and the degree to which real-time payments take-up displaces different payment instruments. In high-income countries, this suggests that the higher displacement of cheques relative to credit transfers results in more significant impacts on overall financial system efficiency.

However, this is not always the case in middle- and lower-income countries. This is likely due to the lower level of credit transfer usage by consumers relative to businesses, resulting in much lower average transaction values for cheques than credit transfers. Nonetheless, higher interest rates likely suggest a higher time value of money for money locked in the system; as such, the higher interest rates in middle- and lower-income countries suggest a greater time value of money gain for these countries.

Analysis for each of the hypothetical countries reveals:

Country 1
In Country 1, the hypothetical high-income country, take-up of real-time payments is expected to lead to a reduction in the daily float value of between USD 5,638m to USD 15,316m, with the large majority of this coming from a reduction in the float associated with cheque transactions (between USD 3,383m and USD 8,550m). This is because while cheques account for only 10% of the daily float for credit transfers and cheques combined, real-time payments are expected to have a larger displacement impact on this payment instrument, completely displacing it in the high take-up scenario. Using the interest rate for Country 1, this suggests a gain in the daily time value of money of up to USD 586,000.

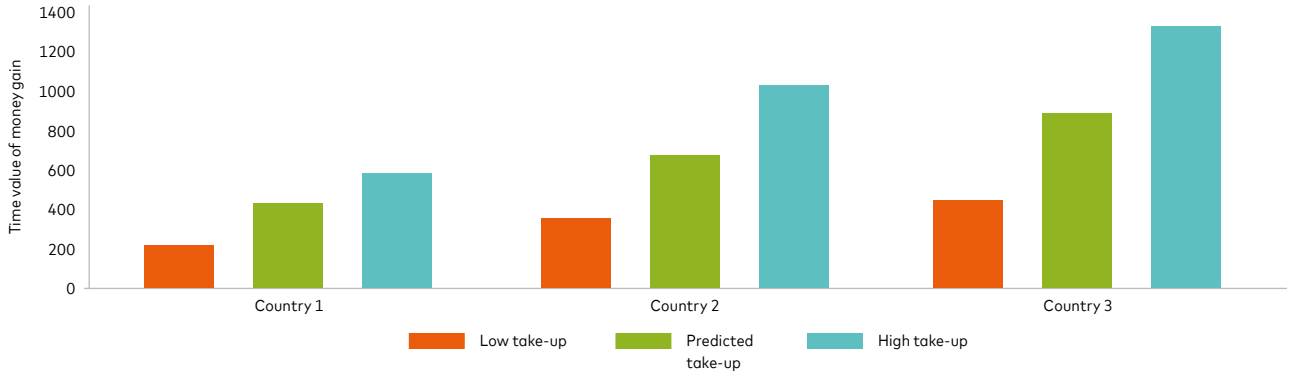
Country 2

In Country 2, the hypothetical middle-income country, take-up of real-time payments is expected to lead to a reduction in the daily float value of between USD 4,455m and USD 13,364m, with the large majority of this from the reduction in float associated with credit transfers (between USD 3,636m and USD 10,909m). This is driven by the higher average transaction values for credit transfers relative to cheques, likely due to the greater use of credit transfers by businesses than consumers for the former. However, Country 2 achieves a larger reduction in proportional terms for cheques, with up to a 79% reduction in the daily float value for the payment instrument. This is again due to the larger expected displacement of cheques by real-time payments. Because of the higher interest rate for Country 2 relative to Country 1, the gain in the daily time value of money is estimated at up to USD 1,018,000.

Country 3

In Country 3, the hypothetical lower-income country, take-up of real-time payments is expected to lead to a reduction in the daily float value of between USD 2,367m and USD 7,101m, with this coming entirely from the reduction in float associated with credit transfers due to the lack of cheque usage in the economy. This is driven by many of the same factors as Country 2. The higher interest rate suggests a more significant gain in the daily time value of money, estimated at up to USD 1,333,000.

Figure 14 – Indicative one-day time value of money gain at five years after the introduction of real-time payments, USD thousands



Source: Deloitte analysis

5

5.3 Tax benefits

As Figure 10 in Section 4.4 shows, high-income countries generally collect a higher percentage of national income in tax, with some exceptions. This suggests that gains in tax revenue are likely to be highest for high-income countries due to their larger take of any formalised portions of GDP.

However, given the large usage of cash in middle- and low-income countries, there is a large potential for revenue gains from VAT for countries able to reduce cash usage further and formalise larger portions of the shadow economy. Additional investments in cash-reducing features of instant payment schemes, such as POS solutions and features to increase accessibility (see Sections 3.2.3, and 4.5), would increase impacts on tax receipts for middle- and low-income countries.

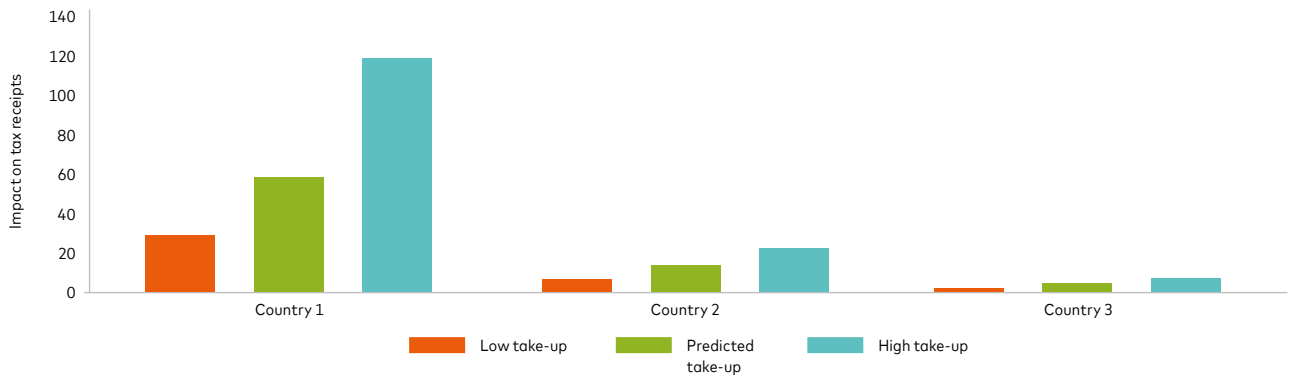
Analysis for each of the hypothetical countries reveals:

Country 1
In Country 1, the hypothetical high-income country, take-up of real-time payments is expected to lead to an impact on tax revenue of between USD 29m and USD 117m. This reflects that real-time payments take-up is estimated to be higher, and therefore reductions in cash usage are estimated to be more significant. In addition, the generally higher rates of standard indirect tax,[34] and thus the higher rates of tax revenue as a percentage of GDP, means that a greater proportion of formalised portions of the shadow economy would be collected in tax.

Country 2 and Country 3
In Country 2, the hypothetical middle-income country, take-up of real-time payments is expected to lead to a more modest impact on tax revenue (between USD 7m and USD 22m). The impact on tax revenue is also expected to be smaller for Country 3, with an estimated impact of between USD 3m and USD 8m.

However, for both of these countries, the high levels of cash usage suggest the potential for greater gains if more investments are made in instant payment schemes to target cash use cases. For example, this may include increasing ease of access and promoting POS usage through integration of technological solutions, as in Thailand's PromptPay scheme.

Figure 15 – Indicative annual impact on tax receipts at five years after the introduction of real-time payments, USD m



Source: Deloitte analysis

5.4
Country summary tables

Table 5 – Indicative annual impact at five years after the introduction of real-time payments, USD – Country 1

	Low take-up	Predicted take-up	High take-up
Payment system costs			
Cost of real-time payments	294,000,000	559,000,000	797,000,000
Decrease in net payment system costs			
Cash	69,000,000	137,000,000	203,000,000
Cheques	262,000,000	545,000,000	703,000,000
Credit transfers	-20,000,000	-35,000,000	-47,000,000
TOTAL	**17,000,000**	**87,000,000**	**62,000,000**
Financial system efficiency			
Cheques			
Total daily float value	8,550,000,000	8,550,000,000	8,550,000,000
Reduction in daily float value	3,383,000,000	6,766,000,000	8,550,000,000
Time value of money gain (one-day)	129,000	259,000	327,000
Credit transfers			
Total daily float value	76,950,000,000	76,950,000,000	76,950,000,000
Reduction in daily float value	2,255,000,000	4,510,000,000	6,766,000,000
Time value of money gain (one-day)	86,000	173,000	259,000
TOTAL			
Total daily float value	85,500,000,000	85,500,000,000	85,500,000,000
Reduction in daily float value	5,638,000,000	11,276,000,000	15,316,000,000
Time value of money gain (one-day)	216,000	432,000	586,000
Tax receipts			
Impact on tax receipts	29,000,000	58,000,000	117,000,000
Impact on tax receipts as a percentage of GDP	0.00%	0.00%	0.01%

Source: Deloitte analysis

5

Table 6 – Indicative annual impact at five years after the introduction of real-time payments, USD – Country 2

	Low take-up	Predicted take-up	High take-up
Payment system costs			
Cost of real-time payments	441,000,000	870,000,000	1,287,000,000
Decrease in net payment system costs			
Cash	9,000,000	18,000,000	27,000,000
Cheques	413,000,000	834,000,000	1,266,000,000
Credit transfers	151,000,000	304,000,000	458,000,000
TOTAL	**132,000,000**	**287,000,000**	**464,000,000**
Financial system efficiency			
Cheques			
Total daily float value	3,115,000,000	3,115,000,000	3,115,000,000
Reduction in daily float value	818,000,000	1,636,000,000	2,455,000,000
Time value of money gain (one-day)	62,000	125,000	187,000
Credit transfers			
Total daily float value	41,538,000,000	41,538,000,000	41,538,000,000
Reduction in daily float value	3,636,000,000	7,273,000,000	10,909,000,000
Time value of money gain (one-day)	277,000	554,000	831,000
TOTAL			
Total daily float value	44,654,000,000	44,654,000,000	44,654,000,000
Reduction in daily float value	4,455,000,000	8,909,000,000	13,364,000,000
Time value of money gain (one-day)	339,000	679,000	1,018,000
Tax receipts			
Impact on tax receipts	7,000,000	15,000,000	22,000,000
Impact on tax receipts as a percentage of GDP	0.00%	0.00%	0.00%

Source: Deloitte analysis

Table 7 – Indicative annual impact at five years after the introduction of real-time payments, USD – Country 3

	Low take-up	Predicted take-up	High take-up
Payment system costs			
Cost of real-time payments	576,000,000	1,139,000,000	1,690,000,000
Decrease in net payment system costs			
Cash	15,000,000	29,000,000	44,000,000
Cheques	-	-	-
Credit transfers	208,000,000	419,000,000	630,000,000
TOTAL	**-352,000,000**	**-691,000,000**	**-1,016,000,000**
Financial system efficiency			
Cheques			
Total daily float value	-	-	-
Reduction in daily float value	-	-	-
Time value of money gain (one-day)	-	-	-
Credit transfers			
Total daily float value	18,269,000,000	18,269,000,000	18,269,000,000
Reduction in daily float value	2,367,000,000	4,734,000,000	7,101,000,000
Time value of money gain (one-day)	444,000	888,000	1,333,000
TOTAL			
Total daily float value	18,269,000,000	18,269,000,000	18,269,000,000
Reduction in daily float value	2,367,000,000	4,734,000,000	7,101,000,000
Time value of money gain (one-day)	444,000	888,000	1,333,000
Tax receipts			
Impact on tax receipts	24,000,000	47,000,000	71,000,000
Impact on tax receipts as a percentage of GDP	0.00%	0.00%	0.00%

Source: Deloitte analysis

6
Conclusion

In most of the markets in which financial institutions have implemented a real-time payment system in its most complete form, with immediate transfer around the clock, it remains relatively nascent.

The full impact will only become clear with time as the track record gets longer and entrepreneurial corporates, start-ups, and policymakers have more time to deploy innovative services on top of the modern real-time payments infrastructure. There is a value in implementing real-time payments in order to increase a country's exposure to future innovations and create a platform for the next wave of use cases developed by fintech pioneers.

However, the track record is long enough to establish some core facts about the impacts of real-time payments:

- First, that real-time payments are likely to displace a range of other payment instruments. The exact set of instruments displaced will depend on what is in place before real-time payments are introduced, but non-real-time payments, cheques, and other P2P payments seem likely to be displaced first.
- Second, that the increased efficiency of real-time payments has a range of material economic and social impacts. Real-time payments are a practical and cost-effective alternative to legacy payment instruments. The straightforward financial savings will be particularly pronounced where real-time payments displace high-cost instruments such as cheques. Real-time payments can also remove an economic inefficiency that impairs household finances and business operations, with money locked up in the financial system instead put to productive uses.
- Finally, modernising payments infrastructure by implementing real-time payments can open up formal financial institutions to include more people, and more transactions. The rapid development of real-time payments in the previously cash-dominated economy in Thailand is a particularly vivid example. Over time, consumers will benefit from gaining access to other valuable financial services, whilst governments will benefit from the ability to distribute benefits and collect taxes more quickly and accurately.

The mix of benefits in any specific country will depend on how the institutions responsible implement real-time payments, and the specific economic conditions in the country concerned. This study is not intended to provide an alternative to such a detailed business case.

However, this study should provide a starting point for those considering taking that next step.

By helping policymakers to understand the likely impact on the 'payment mix', the benefits that will matter most for their country, and the scale at which those benefits might be realised, this study can hopefully support policymakers in considering more specifically the expected impacts of real-time payments in their respective countries.

Technical appendix

This appendix provides further detail on the econometric analysis and additional review of the literature. It also provides further details on methodologies used to quantify economic and social benefits.

A Econometrics

A1 Objective

The objective of the econometric analysis is to analyse:

- the impact of real-time payments on the transaction volume of existing payment instruments; and
- the impact of economic and structural variables on the transaction volume of each payment instrument.

This is done through an econometric methodology that examines how transaction volumes of different payment instruments have varied with the take-up of real-time payments, controlling for economic and structural variables that simultaneously drive transaction volumes. The econometric outputs are then used to evaluate the economic impact of real-time payments.

A2 Background

A2.1 Payment methods

The analysis addresses how the introduction of real-time payments has affected existing payment instruments such as cash, cheque, card, non-instant payment credit transfer and direct debit. Each of these methods has its unique benefits and costs. For example, cash transactions are immediate and convenient although they incur storage and shoe leather costs (Kalckreuth et al., 2014). Cheque transactions facilitate large payments (when large amounts of cash notes may be inconvenient) but typically require a few working days to process.

Real-time payments enable real-time or near-real-time clearing and settlement of transactions, and in doing so yield economic benefits. At the macro level, real-time payments may lead to a reduction in the shadow economy and an increase in tax revenue. At the micro level, real-time payments may yield savings in transactions and social costs. These benefits are discussed in further detail in Section 3.

The introduction of real-time payments add an alternative to an agent's choice set of payment instruments. In particular, real-time payments may substitute for existing payment instruments at the transaction level if the characteristics of real-time payments are deemed to be more attractive than existing payment instruments.

A2.2 Economic and structural variables

The inclusion of economic and structural variables in the econometric analysis is motivated by theoretical and practical considerations. For example, economic theory suggests that higher GDP per capita may reflect higher transaction demand and thus may drive higher transaction volumes. In addition to economic variables that are hypothesised to affect transaction volumes, structural variables are included to reflect practical factors – such as greater convenience and accessibility – that drive take-up of specific payment instruments. For example, a higher number of Automated Teller Machines ("ATMs") facilitates more cash transactions while a higher number of cards issued facilitates more card transactions.

A3 Data

The data required for econometric analysis can be grouped into two categories: variables related to payment instruments and economic and structural variables.

A3.1 Transaction volume

Existing payment methods

The main source of payment-related data is BIS.[35] BIS provides data on the transaction volume of cheques, cards, credit transfers, and direct debits. In the cross-section dimension, data is only available for CPMI member countries.[36] In the time series dimension, data is available from 1991 on an annual basis.[37] Since cash data is not available from BIS, this data is obtained from Euromonitor. However, cash data is only available for 20 CPMI countries from 2003 whereas data for other payment methods is available from 1991.[38]

The dataset of transaction volumes is an unbalanced macro panel consisting of 20 countries from 1991 to 2016. In this dataset, the number of observations across time is not the same for each country because the number of CPMI member countries increased from 10 to 22 in 2005. Therefore, there are more observations across time (1991 – 2016) for the original 10 CPMI member countries compared to the additional 12 CPMI member countries (2005 – 2016).

Table 8 below summarises information on the data on existing payment instruments.

Table 8 – Availability of data on existing payment instruments

Payment instrument	Main source	Countries available	Years available
Cash	Euromonitor	20 CPMI countries	2003 – 2016
Cheque, card, non-instant credit transfer, direct debit	BIS	22 CPMI countries	1991 – 2016

In constructing the dataset on payment instruments, the following issues were considered:

- Comparability: The unique nature of each country's payment system poses a challenge to the comparability of payment instruments across countries. For example, Switzerland implements mobile-only real-time payments (Twint) while the UK implements real-time credit transfers (FPS).
- Granularity: The inclusion of more payment instruments in the econometric analysis may reduce the degrees of freedom. In contrast, an overly high-level set of payment instruments may limit the granularity of insights.
- Coverage: The Euromonitor cash data only includes P2B cash transactions but does not include P2P, B2P, and B2B cash transactions due to data measurement issues. While the volume of B2P and B2B transactions may not be large, the exclusion of P2P transactions may limit the analysis on cash transactions.

The set of payment instruments considered for econometric analysis includes cash, cheques, cards, non-instant credit transfers, direct debits, and real-time payments.

Real-time payments

BIS defines real-time payments according to two key features: [39] [40]

- Speed: Transmission of the payment message and the accessibility of final funds to the payee occur in real-time or near real-time.
- Availability: Near to a 24-hour and 7-day ("24/7") basis as possible.

The CPMI member countries with real-time payments considered for econometric analysis are listed in Table 9 below. The main source underlying this list is BIS.[41] Table 9 is not an exhaustive list as there are other instant payment schemes in non-CPMI member countries, such as Thailand and Australia.[42]

In this dataset, there are 14 CPMI member countries which have implemented real-time payments. However, due to data availability issues (China, Italy, Sweden) and the sample period of the dataset (Switzerland, USA), only 9 instant payment schemes are considered for analysis. In the cases of China, Italy, Sweden, Switzerland, and the US, the country-year observations retained in the dataset are those without instant payment schemes.

Table 9 – CPMI member countries with real-time payments

Country	Instant payment	Note	Source
Brazil	SITRAF	Operates during business hours only	Banco Central do Brasil
China	IBPS	Data unavailable	
India	IMPS		Reserve Bank of India
Italy	Nexi	Data unavailable	
Japan	Zengin	Operates during business hours only	BIS
Korea	EBS, CD/ATM		Bank of Korea
Mexico	SPEI	Operates during business hours only	Banco de Mexico
Singapore	FAST		BIS
South Africa	RTC		Payment Association of South Africa
Sweden	BiR	Data unavailable	
Switzerland	Twint	Year of implementation outside sample	
Turkey	BKM Express	Operates during business hours only	BIS
United Kingdom	FPS		BIS
United States	RTP	Year of implementation outside sample	

The econometric analysis adopts a less restrictive definition of real-time payments to obtain a larger sample of countries with real-time payments. In particular, the services in Brazil, Japan, Mexico, and Turkey are considered real-time payments despite involving real-time settlements only (and not posting) and not being available on a 24/7 basis.[43]

Real-time payments transactions are assumed to be instant credit transfer transactions in the econometric analysis. Therefore in order to estimate the impact of real-time payments on non-instant payment credit transfers, the latter is obtained by subtracting real-time payments from total credit transfers.

A3.2 Economic and structural variables

Data for economic and structural variables is obtained from BIS and World Bank. The variables proposed for inclusion in the econometric analysis are listed in Table 10 below.

Table 10 – Economic and structural variables

Category	Variable	Definition	Source
Economic	Population	Total population	World Bank
	GDP per capita	Real GDP per capita	
	Savings ratio	Savings as a proportion of GDP	
Structural (Payment)	Card issued	Number of cards issued	BIS
	POS terminal	Number of POS terminals	
	ATM	Number of ATMs	
	Bank branches	Number of bank branches	
Structural (Non-payment)	Broadband	Number of fixed broadband subscriptions	World Bank
	Mobile	Number of mobile subscriptions	
	Old population	Proportion of population aged above 65	
	Share of internet users	Proportion of population who used internet	

A3.3 Exclusions

There are two data issues which require the exclusion of some observations:

- Potential zeros: Where zero transaction volumes represent observations where certain payment instruments are unavailable, these are excluded. Specifically, these observations relate to direct debit in Japan, as well as cheque in Netherlands and Russia.[44]
- Real-time payments: Where real-time payments volume data is unavailable, such observations are excluded because the derivation of non-instant payment credit transfer transaction volumes would be inaccurate. In contrast, observations where real-time payments schemes themselves are not in place are not excluded because BIS data on credit transfer transaction volumes would fully reflect non-instant payment credit transfers.

A4 Econometric Model

The types of econometric models in the literature depend on the granularity of data available. For example, Bounie and François (2006) used transaction-level data to estimate a multinomial logit model. This model analysed the determinants (such as transaction value) of the probability of a transaction being paid by cash, cheque, or card as a POS solution. In another example, Schuh and Stavins (2010) used individual level data to estimate Heckman selection models of adoption (extensive margin) and volume (intensive margin) of different payment instruments.

The analysis adopts the econometric specification used by Humphrey et al., (1996), Drehmann et al., (2002) and Amromin et al., (2007). In particular, this specification is formulated at the macro level, such that macro data is used for variables expressed at the per capita level on an annual basis. Therefore, parameters in such model specifications should be interpreted at the appropriate level of granularity.

A4.1 Regression equation

Transaction volumes can be described by the following equation:

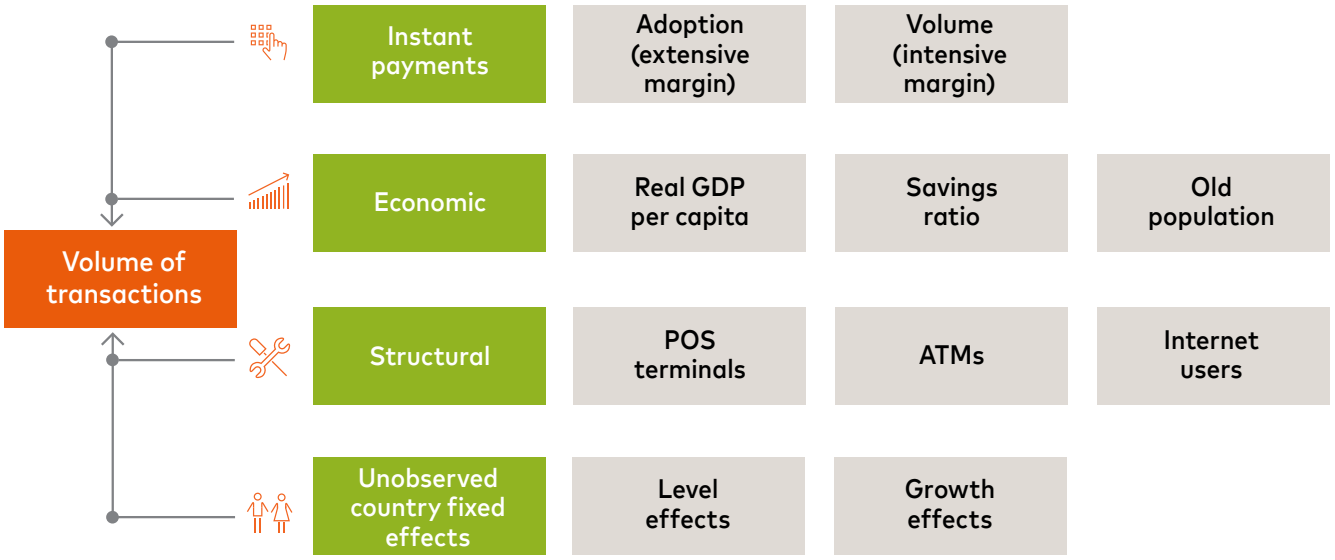
$$y_{m,it} = \alpha_m + \gamma_{m1}\mathbf{1}(IP_{it} \geq 0) + \gamma_{m2}IP_{it} + \boldsymbol{\pi}'_m\boldsymbol{X}_{m,it} + \boldsymbol{\theta}'_m\boldsymbol{D}_{m,i} + \varepsilon_{m,it}, \varepsilon_{m,it} \sim (0, \sigma_m^2) \quad \text{(Equation 1)}$$

$y_{m,it}$ is the volume of transactions using payment instrument m in country *i* in year *t*. $X_{m,it}$ is a set of economic and structural variables which may differ across payment instruments. $\boldsymbol{D}_{m,i}$ is a set of dummies to allow for potential unobserved fixed effects while $\varepsilon_{m,it}$ is an idiosyncratic error with zero mean.

The non-linear specification of real-time payments variables is motivated by a potential two-fold effect of real-time payments.[45] Firstly, the availability of real-time payments would inform consumers and businesses of a new payment instrument, thus motivating a dummy variable $1\,(IP_{it} \geq 0)$ to indicate whether real-time payments were available in country *i* in year *t*.[46] Secondly, the effect of an additional instant payment transaction may be different from that of the initial transaction when real-time payments were introduced, thus motivating the continuous variable IP_{it} to capture the marginal impact of an additional instant payment transaction.

The different factors driving the volume of transactions are illustrated in Figure 16 below.

Figure 16 – Drivers of transaction volume



Source: Deloitte analysis

The interpretation of parameters in Equation 1 are provided in Table 11 below.

Table 11 – Interpretation of parameters

Interpretation of parameters	
$\boldsymbol{\alpha_m}$	Constant
$\boldsymbol{\gamma_{m1}}$	Impact of the availability of real-time payments (level effect)
$\boldsymbol{\gamma_{m2}}$	Impact of an additional real-time payments transaction (marginal effect)
$\boldsymbol{\pi'_m}$	Impact of economic and structural variables
$\boldsymbol{\theta'_m}$	Unobserved fixed effects
$\boldsymbol{\sigma_m^2}$	Variance of idiosyncratic error term

All model parameters are identified by the variation in the explanatory variables: $\mathbf{1}(IP_{it} \geq 0)$, IP_{it}, $\boldsymbol{X}_{m,it}$ and $\boldsymbol{D}_i$. Since Equation 1 is specified at the country-level, the parameters should not be interpreted at the consumer or transaction level. Whereas alternative payment instruments must be substitutes for each other at the transaction level, there may be complementarities between different types of payment instruments at the country level. For example, if the availability of real-time payments makes mobile payments, including those made by card, more convenient in general, then real-time payments and card transactions may be complements.

The parameter of interest in Equation 1 is γ_{m2} which captures the impact of an additional real-time payments transaction per capita. The parameter γ_{m1} is not used for analysis due to concerns that the dummy variable $\mathbf{1}(IP_{it} \geq 0)$ may capture other macroeconomic events and hence γ_{m1} may reflect additional impacts. Therefore, inference in this econometric analysis is based on the intensive margin of the impact of the take-up of real-time payments.

A4.2 Endogeneity issues

The next stage of the analysis selects an estimator to estimate model parameters. The starting point is the Ordinary Least Squares (OLS) estimator. However, the OLS estimator would not be appropriate if some explanatory variables $(\mathbf{1}(IP_{it} \geq 0), IP_{it}, \boldsymbol{X}_{m,it})$ are correlated with the structural error terms $\varepsilon_{m,it}$. This could occur if there are unobservable cultural factors or underlying preferences which drive both real-time payments and existing payment methods. In this case, the OLS estimator would misstate the impact of real-time payments without controlling for such unobservable factors.

Two potential sources of endogeneity are identified: the real-time payments variables and unobserved fixed effects at the country or region level.

Real-time payments variables

The econometric analysis aims to achieve model parsimony by refraining from overfitting the regression equations with explanatory variables, particularly given the modest sample size.[47] Consequently, there may be omitted variables which are correlated with the real-time payments variables and affect existing payment methods. In this case, the real-time payments variables are endogenous, resulting in biased and inconsistent OLS estimates.

Omitted variables may include other payment method volumes in individual payment method equations. For example, the omission of cash and cheque from the credit transfer equation may lead to omitted variable bias if both variables are correlated with real-time payments and affect credit transfers simultaneously.

Unobserved fixed effects

Another potential source of endogeneity is the correlation between unobserved fixed effects and explanatory variables. For example, unobserved and time-invariant cultural factors, such as the legacy usage of cheques in the US, may simultaneously correlate with the take-up of real-time payments and the accessibility of payment instruments (captured by $\boldsymbol{X}_{it}$). Ignoring the presence of unobserved fixed effects would lead to biased and inconsistent estimates.

Dummies are therefore included to address potential endogeneity caused by unobserved regional, country and outlier fixed effects.[49] This formulation facilitates an analogous Fixed Effects estimator.[50] While regional dummies are less granular compared to country dummies, there would be greater degrees of freedom for statistical tests particularly given the modest sample size. In certain equations with regional dummies, dummies are included to account for outliers as shown in Table 12 below.[51] These dummies are formulated based on a qualitative analysis of time series plots of transaction volumes. This method is applied consistently across all payment instruments.

Table 12 – Specification of country, regional and outlier dummies

Payment instrument	Regional	Country	Outlier
Total existing	✓		-
Cash		✓[52]	-
Cheque	✓		United States
Card	✓		Singapore
Credit transfer	✓		-
Direct debit	✓		Germany, Netherlands
Instant payment			Korea, Japan, Turkey

These dummies account for permanent cross-country or cross-region differences in the transaction volumes of different payment instruments. In addition, these dummies are interacted with linear time trends to account for differences in growth rates.

A4.3 Approach

The proposed approach estimates three regression equations to identify and analyse the drivers of transaction volumes:

- Total existing payment instruments (excluding real-time payments)
- Individual payment instruments (excluding real-time payments)
- Real-time payments

The transaction volumes and certain economic and structural variables are expressed in per capita terms to remove any confounding population impacts.

Individual payment instruments

Individual payment instruments volumes can be described for the following payment instruments:

$$y_{m,it} = \{cash_{it}, cheque_{it}, card_{it}, ni\_cred\_trans_{it}, dir\_debit_{it}\} \qquad \text{(Equation 2)}$$

The set of explanatory variables included in each payment instrument equation is outlined in Table 13 below. These variables were selected on the basis of theoretical considerations. For example, the number of ATMs and POS terminals may affect the convenience of cash payments, which in turn may affect the volume of cash transactions. In addition, the statistical significance of coefficients was taken into account when refining the set of explanatory variables. For example, the number of bank branches was excluded from the non-instant payment credit transfer equation as the coefficient on this variable was statistically insignificant in various model specifications.[53]

The purpose of these individual equations is to estimate the impact of real-time payments across different payment instruments. The null hypothesis of interest is $H_0: \gamma_{\mathrm{m2}} = 0$ for payment method m to test whether real-time payments substitute for existing payment instruments.

Table 13 – Explanatory variables of individual payment methods

Variables	Cash	Cheque	Card	NIP-CT	DD
Economic					
Log real GDP per capita	✓	✓	✓	✓	✓
Savings ratio	✓	✓	✓	✓	✓
Structural					
POS terminal per capita	✓		✓		
ATM per capita	✓				
Card issued per capita			✓		
Old population		✓			
Log real GDP per capita Old populations[54]		✓			
Share of internet users				✓	
Country specific trend	✓				
Region specific trend		✓	✓	✓	✓
Outlier specific trend		✓			
Dummies					
Country	✓				
Region		✓	✓	✓	✓
Outliers		✓	✓		✓

Real-time payments

The transaction volume of real-time payments can be described by the following equation:

$$IP_{it} = \alpha_6 + \boldsymbol{\pi}_6' \boldsymbol{X}_{6,it} + \boldsymbol{\theta}_6' \boldsymbol{D}_i + \varepsilon_{6,it}, \varepsilon_{6,it} \sim (0, \sigma_6^2) \quad \text{(Equation 3)}$$

The set of explanatory variables included in Equation 3 are outlined in Table 14 below.

Table 14 – Explanatory variables of real-time payments

Category	Variables
Economic	Log real GDP per capita[55]
Structural	Card issued per capita
	Broadband per capita
	Years since IP implemented
Dummies	Outlier

The purpose of this auxiliary regression is to analyse the impact of economic and structural variables on the take-up of real-time payments. These impacts can be used in counterfactual analyses to predict the volume of real-time payments per capita for countries which do not currently implement real-time payments.

A4.4 Estimation

The inclusion of economic and structural variables as well as dummies to control for unobserved fixed effects may be insufficient to address concerns around the endogeneity of real-time payments variables. The Two Stage Least Squares ("2SLS") estimator is able to obtain consistent estimates by using instrumental variables.

The necessary conditions for consistent estimation by 2SLS are instrument relevance and instrument exogeneity. The first condition requires instruments to be correlated with the endogenous variables. Instrument exogeneity requires instruments to be uncorrelated with the idiosyncratic error terms. However, the 2SLS estimator may perform poorly in small samples or in the presence of weak instruments.[56]

The proposed set of instrumental variables for each payment instrument is outlined in Table 15 below.

Table 15 – Economic and structural variables by payment instrument

Equation	Instruments
Total existing payment instrument	Broadband per capita Card issued per capita Country dummies: Republic of Korea, Japan, Turkey
Cash	Broadband per capita Card issued per capita
Cheque	Broadband per capita Card issued per capita Country dummies: Republic of Korea, Turkey
Card	Broadband per capita Country dummies: Republic of Korea, Japan
Non-instant payment credit transfer	Broadband per capita Country dummies: Republic of Korea, Japan
Direct debit	Broadband per capita Card issued per capita Country dummies: Republic of Korea, Japan, Turkey

The preferred estimator for the regressions of total existing payment instrument volumes and individual payment instrument volumes is the 2SLS estimator with unobserved fixed effects. For the auxiliary regression of real-time payments volume, the OLS estimator with unobserved fixed effects is preferred because this regression is used for predictive purposes only.

The analysis has sought to utilise a consistent set of instruments across equations. Differences in instrument sets arise from statistical tests underpinning the preferred model specifications. The inclusion of outlier dummies in instrument sets increases the correlation between instruments and endogenous variables, which in turn reduces the risk of weak instruments. However, the exogeneity of such additional instruments needs to be tested. For example, cards issued per capita was excluded from the instrument set in the non-instant payment credit transfer equation because test results suggested that this instrument is endogenous.[57]

A5 Model output

The parameter of interest is the coefficient on IP_{it}, which measures the impact of an additional real-time payments transaction on existing payment method volumes. The impacts of economic and structural variables are reported to assess the relative importance of different variables. For brevity, the coefficients on dummies and time trends are not reported.

In addition to estimates of model parameters, results from various statistical tests are reported to assess model diagnostics. These tests are outlined in Table 16 below.

Table 16 – Statistical tests

Statistical tests	
2SLS estimator	
Endogeneity	Durbin-Wu-Hausman test of the endogeneity of real-time payments variables. If the real-time payments variables are not endogenous, the OLS estimator should be used instead of the 2SLS estimator.
Weak instruments	Based on the strength of the correlation between instruments and endogenous variables. If instruments are weakly correlated with endogenous variables, the 2SLS estimator is biased towards the OLS estimator.
Instrument exogeneity	Sargan test of the exogeneity of instruments. If the instruments are not exogenous, then the 2SLS estimator is inconsistent.[58]
OLS estimator	
Functional form	Ramsey RESET test of the linear specification of the regression equation. If this test is failed, the functional form of the equation may need to be reconsidered.

The outputs reported in Section A5 relate to preferred model specifications.[59] Alternative model specifications are discussed in Section A6.

A5.2 Individual payment methods

The 2SLS estimates for the regressions of individual payment instruments are outlined in Table 17 below.

Across all payment methods, the impact of an additional instant transaction per capita is negative. The null hypothesis of $H_0: \gamma_{\mathrm{m2}} = 0$ for payment instruments is rejected in the cheque and non-instant payment credit transfer equations.[60] However, the level impact from the introduction of real-time payments is positive for all payment instruments except cash.

The positive level impacts suggest that there are complementarities between real-time payments and the existing payment instruments. However, the coefficients on the variable $\mathbf{1}(IP_{it} \geq 0)$ may capture the impact of other concurrent macroeconomic events. For example, the introduction of FPS in the UK in 2008 coincided with the global financial crisis. The latter event may have contributed to an unobserved structural break in the transaction volume of payment instruments. Therefore, these coefficients may not be able to isolate the level impact of real-time payments from that of other events.

The focus of the analysis is on the marginal impact of real-time payments instead of a combination of marginal and level impacts. While the marginal impacts of real-time payments are negative, there is variation in the magnitudes. In particular, the magnitudes of the marginal impacts in the cheque, non-instant payment credit transfer, and direct debit equations suggest that real-time payments substitute less than 1 – 1 for these payment instruments. As for cash and card, the estimates suggest that real-time payments substitute more than 1 – 1 for these payment instruments.

Table 17 – 2SLS estimates in regressions of individual payment instruments

	Cash	Cheque	Card	NIP-CT	DD
Real-time payments					
Level effect	-23.32	56.24*	1247.48	23.47	47.76*
Marginal effect	-15.02	-0.57*	-95.92	-0.38*	-0.31
Economic					
Log real GDP per capita	252.20	33.34*	177.94	-11.66*	9.62*
Savings ratio	-5.11*	0.60	3.17	2.07*	-0.51
Structural					
POS terminal per capita	76.61		11972.75		
ATM per capita	25826.43				
Card issued per capita			-29.57		
Old population		42.17*			
Log real GDP per capita * Old populations		-4.21*			
Internet population				0.59*	
Statistical tests					
Number of observations	219	285	283	301	277
Exogeneity of real-time payments	Endogenous	Endogenous	Endogenous	Endogenous	Endogenous
Weak instruments (relative bias or rejection rate in Wald test)	Not weak (15% rejection rate)[61]	Not weak (20% relative bias)	Weak	Not weak (15% rejection rate)	Not weak (25% rejection rate)
Exogeneity of instruments	N/A[62]	Exogenous	N/A	Exogenous	Endogenous

* Statistically significant at 5% level

In terms of statistical tests, the test results suggest that instruments may be weak in the cheque and direct debit equations. As a result, standard errors in those equations may be overstated. The analysis estimated a number of model specifications to overcome these issues, however this has not been possible.

A5.3 Real-time payments

The OLS estimates for the regression of real-time payments are outlined in Table 18 below.

Table 18 – OLS estimates

Regression of real-time payments	
Log real GDP per capita	-3.07
Card issued per capita	0.28
Broadband per capita	54.56*
Years since IP implemented	0.59*
Statistical tests	
Number of observations	77
Ramsey RESET test	Misspecification

*Statistically significant at 5% level

Instant payment volumes per capita are estimated to increase by 0.59 every year following implementation. Intuitively, the take-up of real-time payments may be expected to grow as more consumers and businesses become aware of the service over time. The impact of broadband subscriptions per capita is estimated to be positive and significant. The impacts of real GDP per capita and cards issued were found to be insignificant.

In terms of statistical tests, linear misspecification was detected in the regression equation.[63] The preferred specification may not yield the best model fit. Various model specifications were estimated to satisfy this test, however this has not been possible.

A6 Sensitivity

A number of model specifications were estimated to assess the sensitivity of estimates. These alternative specifications involve variations to the sets of explanatory variables and dummies for unobserved fixed effects. This section reports findings from a few alternative model specifications.

A6.1 Total existing payment instruments

The following variations to the model specification outlined in Section A5.1 were estimated:

- Explanatory variables: Explanatory variables from all individual payment instrument equations were included, such as POS terminals per capita. The estimate of γ_2 is -44.37 while the impact of POS terminals per capita is statistically insignificant. When this variable is excluded, a more conservative estimate of γ_2 is obtained. This finding suggests that the marginal impact of real-time payments in this equation is sensitive to the inclusion of POS terminals per capita. Since this variable was not found to be statistically significant, it was excluded from the preferred model specification in Section A5.1.
- Dummies: Country dummies were included instead of regional dummies to control for unobserved fixed effects at a more granular level. The estimate of γ_2 in this specification is statistically insignificant. This finding may be driven by an overfitted model as the additional number of dummies reduces the degrees of freedom available for statistical tests given the modest sample size.

A6.2 Individual payment instruments

The following variations to the model specification outlined in Section A5.2 were estimated:

- Explanatory variables: Additional variables which are hypothesised to affect cheque transaction volumes were included, such as the share of self-employed and bank branches per capita. The estimate of γ_{cheque2} is -0.25, which is similar to the estimate in the preferred specification (-0.38). However, statistical tests found that instruments are not exogenous in this specification. Therefore, the specification outlined in Section A5.2 is preferred for statistical reasons.
- Dummies: The inclusion of country dummies instead of regional dummies is not possible for all individual payment instrument equations since the 2SLS estimator is reliant on country dummies as instruments for some equations, such as non-instant payment credit transfer. When this specification is tested for the cheque equation (where this is possible), the estimate of γ_{cheque2} is statistically insignificant. Once again, this finding may be driven by an overfitted model.

A6.3 Real-time payments

The following variation to the model specification outlined in Section A5.3 was estimated:

- Explanatory variables: Interaction terms between the number of years since real-time payments were implemented and outlier dummies (Korea, Japan and Turkey) were included to account for potentially different growth rates. This sensitivity check is based on a qualitative analysis of the time series plots of real-time payments transactions. The impact of the number of years since real-time payments were implemented is statistically insignificant. This finding may be driven by an overfitted model, which reduces the explanatory power of this variable for other countries. When these interactions are excluded, the estimate is positive, reflecting organic growth in the take-up of real-time payments over time.

A7 Conclusion

The econometric analysis developed a panel dataset to estimate the impact of real-time payments on different payment instruments. The model outputs provide some evidence that real-time payments substitute for existing payment instruments, in particular cheque and non-instant payment credit transfers. However, these outputs are subject to the following limitations:

- Parsimony: The impact of real-time payments is likely to be affected by various macro-level and micro-level factors. It is not practically plausible to obtain an exhaustive list of potential factors.
- Sensitivity: For some equations, the sign and significance of the parameters are sensitive to the set of explanatory and dummy variables. These include the non-instant payment credit transfer equation with respect to the inclusion of savings ratio as an explanatory variable.
- Statistical tests: Some tests were not passed, such as the exogeneity of instruments in the regression of total existing payment instruments. The analysis has attempted a number of model specifications to satisfy these tests, however in some cases this has not been possible.
- Sample: The estimates may be driven by the small number of observations with real-time payments across time and the unbalanced nature of the panel dataset. It is not practically possible to overcome the first limitation given that real-time payments were generally introduced in latter stages of the sample period. Further, there are considerations around the degrees of freedom if a balanced panel dataset is desired.
- External validity: For predictive purposes, a different selection of data granularity (such as quarterly data instead of annual data) or countries may produce different estimates of model parameters in the future.

B Impacts

B1 Payment system costs

For each payment instrument, a relationship between the sample of volumes per capita and unit payment system cost estimates from the literature is estimated using the best-fitting functional form. Table 19 provides a list of the studies used to develop the sample for estimating the relationships, Table 20 provides an outline of the estimated relationships and functional forms, and Figure 17 plots the fitted line for the relationship against the actual data.

Table 19 – Literature used to gather estimates of unit payment system costs and volumes by payment type

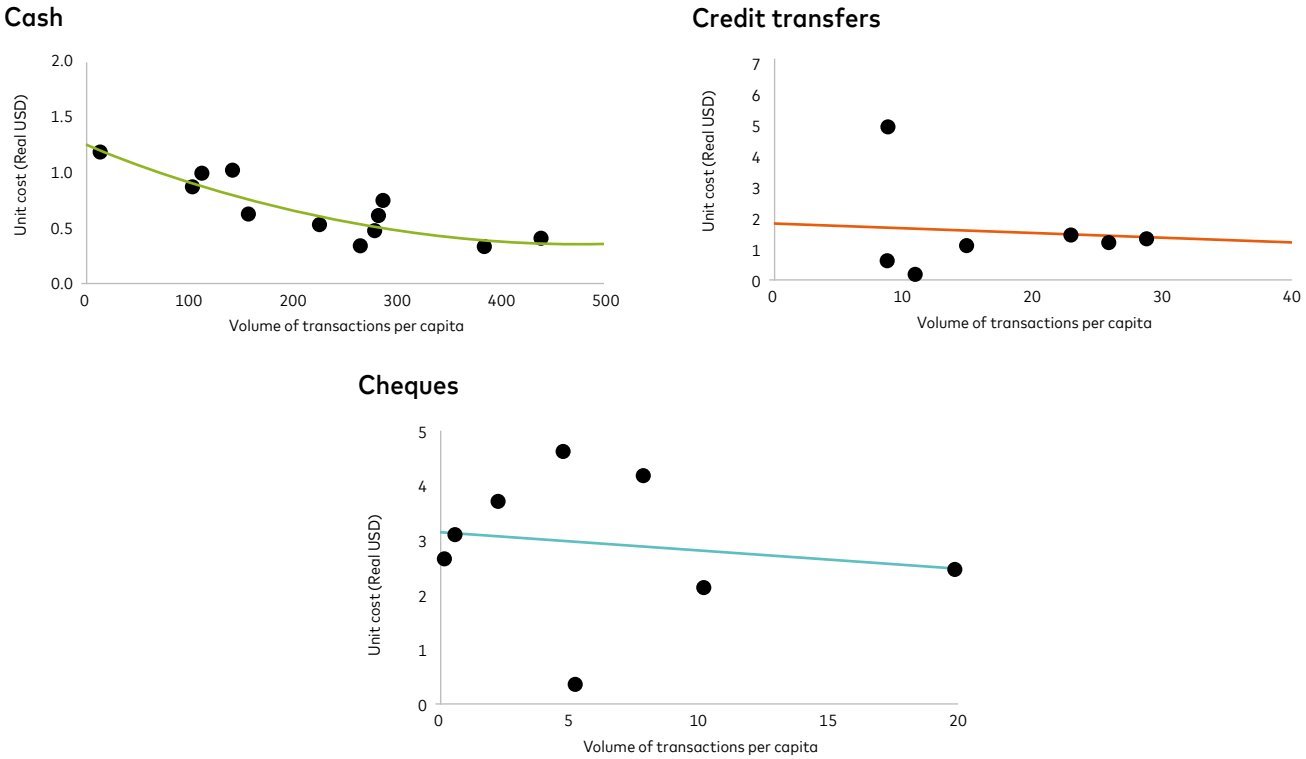
Year of study	Author / Organisation	Payment types considered
1993	Wells, 1996	Cheques
1994	Humphrey, Kim & Vale, 2001	Cheques
2001	Gresvick & Owre, 2002	Cash, Cheques, Credit transfers, Direct debits
2002	Brits & Winder, 2005	Cash
2002	Bergman, Guidborg & Segendorf, 2007	Cash, credit transfers, Direct debits
2002	Guibourg & Segendorf, 2004	Cheques
2003	National Bank of Belgium, 2006	Cash
2005	Banco de Portugal, 2007	Cheques, Credit transfers, Direct debits
2006	Schwartz et al., 2008	Cash, Cheques, Direct debits
2007	Gresvik & Haare, 2009	Cash, Cheques, Credit transfers, Direct debits
2009	Segendorf & Jansson, 2012	Cash, Credit transfers, Direct debits
2009	Jonker, 2013	Cash
2009	Turjan et al., 2011	Cash, Credit transfers, Direct debits
2009	Danmarks Nationalbank, 2012	Cash, Credit transfers, Direct debits
2012	Jonker, 2013	Cash
2013	Stewart et al., 2014	Cash, Cheques, Direct debits
2013	Banco de Portugal, 2016	Cheques, Credit transfers, Direct debits

Table 20 – Coefficient estimates and functional forms for the estimated relationships between volume per capita and unit payment system costs, by payment instrument[64]

Quadratic functional form: $\alpha+\beta(\text{volume per capita})+\gamma(\text{volume per capita})^2$			
	Cash	**Cheque**	**Credit transfer**
α	1.25	3.129037	NA
β	-0.003664	-0.03597	NA
γ	0.000003801	0.000176	NA
Reciprocal functional form: $\alpha+\beta(\text{volume per capita})$			
α	NA	NA	0.4565
β	NA	NA	14.9622

Source: Deloitte analysis based on data from the World Bank, FRED, and the studies in Table 19.

Figure 17 – Estimated curve for the relationships between volume per capita and unit payment system costs and actual data, by payment instrument



Source: Deloitte analysis based on data from the World Bank, FRED, and the studies in Table 19.

The impact on the net payment system costs can then be calculated by estimating differences in costs for the appropriate volumes of transactions for each payment method, given the substitution due to the take-up of real-time payments. The expected impact on net payment system costs for a particular payment instrument could be calculated as:

Impact on net payment system costs for the payment instrument =

$$[\boldsymbol{V_{0,i}} * \boldsymbol{C}\left(\boldsymbol{V_{0,i}}\right) - \boldsymbol{V_{1,i}} * \boldsymbol{C}\left(\boldsymbol{V_{1,i}}\right)] * \boldsymbol{P_i} \quad \text{(Equation 4)}$$

where:
$V_{i,t}$ = Volume of the payment instrument per capita at time *t* for country *i*
P_i = The population for country i (assumed to be constant for the purposes of an indicative calculation)
$C\left(V_{i,t}\right)$ = Unit payment system cost of the payment instrument for the specific volume, as determined by the functional form and estimates in Table 20.

As it is expected that the introduction of real-time payments would lead to a substitution effect, V_1 would be less than V_0 leading to a net cost saving. Impacts on each payment method's system costs can then be aggregated to estimate the total impact on net payment system costs across payment types. The cost of real-time payments, estimated using the cost curve for credit transfers more widely, would then be added to the total impact on system costs to conclude on net benefits to the payment system. Therefore:

Total impact on net payment system costs of existing payment types= $\boldsymbol{\Sigma_{p=1}^{n} Impact_p + C_{IP}}$ (Equation 5)

where:
p = Each different payment method
$Impact_p$ = The impacts on net payment system costs, as calculated above
C_{IP} = The estimated costs of real-time payments, as estimated using cost curve for credit transfers.

B2 Financial system efficiency

Valuation of the reduction in financial system inefficiency requires estimating the total daily float for cheques and for credit transfers separately, which are the values of the money locked in the system due to delays in these payment instruments:

Total daily float value = $\boldsymbol{F_{i,t}} = \frac{\boldsymbol{Total\ number\ of\ transactions_{i,t} * Avg\ value\ of\ transactions_{i,t}}}{\boldsymbol{260}}$ (Equation 6)

where:
$T_{i,t}$ = Total number of transactions in country *i* at time *t*
$AV_{i,t}$ = The average value of transactions in country *i* at time *t*

And the reduction in the value of the float from transaction displacement:

Reduction in float value = $\boldsymbol{F_{i,t} * D_{i,t}}$ (Equation 7)

where:
$D_{i,t}$ = The displaced volume in country *i* at time *t*, a quantification of benefits in terms the overall "time value of money" can be estimated by applying the weighted average delay and using discount rates generally assumed to account for time value, i.e.:

$$\boldsymbol{D_{i,t} * WD_{i,t} * TV_{i,t}} \quad \text{(Equation 8)}$$

where:
$WD_{i,t}$ = The weighted average delay in country *i* at time *t*
$TV_{i,t}$ = The time value of money in country *i* at time *t*

B3 Indicative impacts for hypothetical countries

Table 21 presents the characteristics of the hypothetical countries considered in the indicative impacts exercise. These are based on a combination of the economic characteristics of multiple countries within each of the high- middle- and low-income groupings and, as such, do not reflect the characteristics of any one country. GDP is normalised for comparison purposes, with economic characteristics adjusted to have an aggregate GDP of USD 1.5tn for each country.

Table 21 – Hypothetical country characteristics

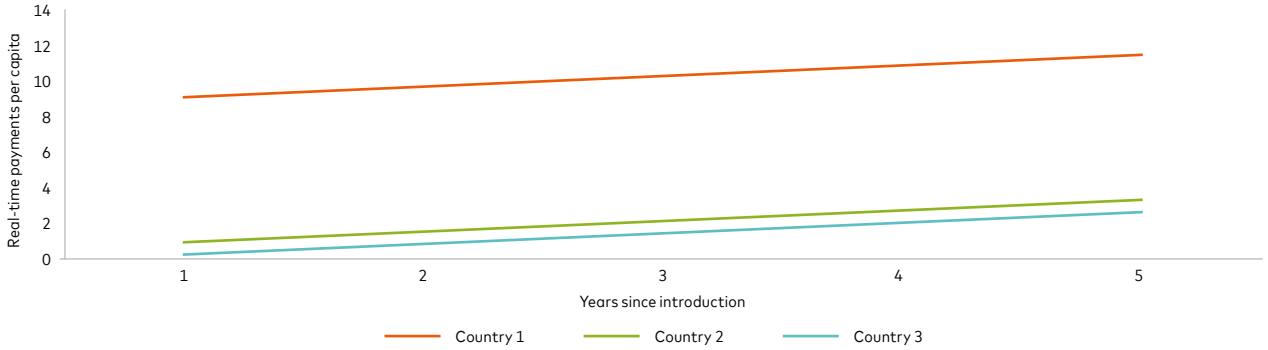
	Country 1	Country 2	Country 3
GDP per capita	USD 50,000	USD 10,000	USD 6,000
Population	30,000,000	150,000,000	250,000,000
Tax as a % of GDP[65]	20%	15%	10%
Total number of broadband subscriptions	11,000,000	18,000,000	20,000,000
Total number of cards issued	71,000,000	180,000,000	200,000,000
Total number of transactions	25,000,000,000	54,000,000,000	95,000,000,000
Average value of credit transfers	USD 9,000	USD 10,000	USD 5,000
Average value of cheque transactions	USD 9,000	USD 1,500	USD 1,000
Initial payment instrument mix	Cash: 70% Cheques: 1% Credit transfers: 9% Direct debits: 10% Cards: 10%	Cash: 90% Cheques: 1% Credit transfers: 2% Direct debits: 1% Cards: 6%	Cash: 95% Cheques: 0% Credit transfers: 1% Direct debits: 0% Cards: 4%
Assumed impact on cash usage[66]	A decrease of 0.50 in cash transactions per capita for every additional instant payments transaction per capita		
Interest rate	1%	2%	5%

Source: Deloitte illustration based on data for high-, middle-, and low-income countries from the World Bank and BIS.

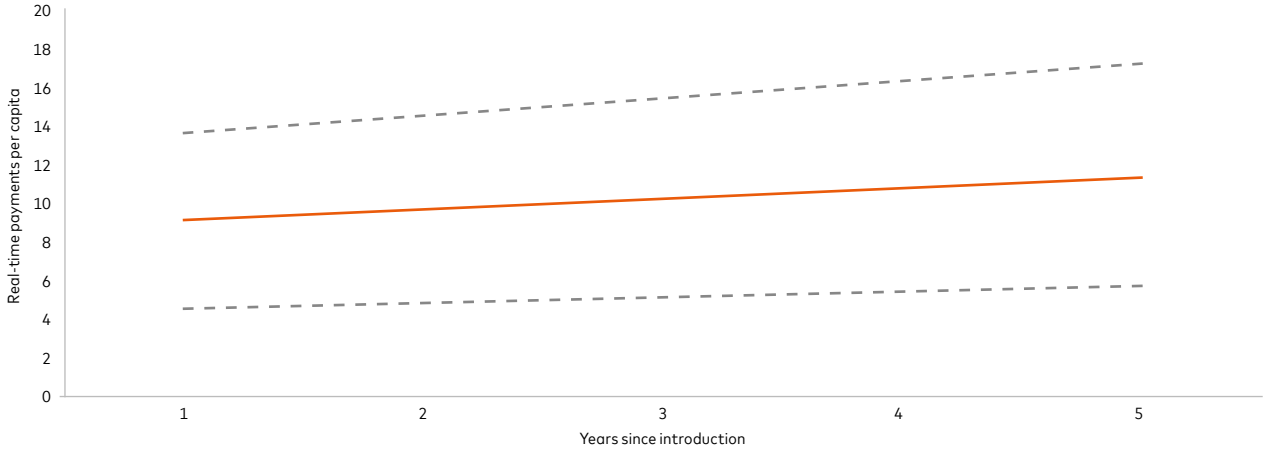
Note that the average transaction values for cheques and credit transfers are very high. This reflects that usage of these instruments is lower for smaller transactions in these economies and that this is a mean figure, which will include large business to business payments.

Figure 18 presents the predicted take-up of real-time payments for Country 1, Country 2, and Country 3 as estimated using the econometric model. It also presents "high" take-up and "low" take-up scenarios modelled as 50% above and below the predicted take-up to simulate the potential impacts of having or not having take-up-supporting implementation features.

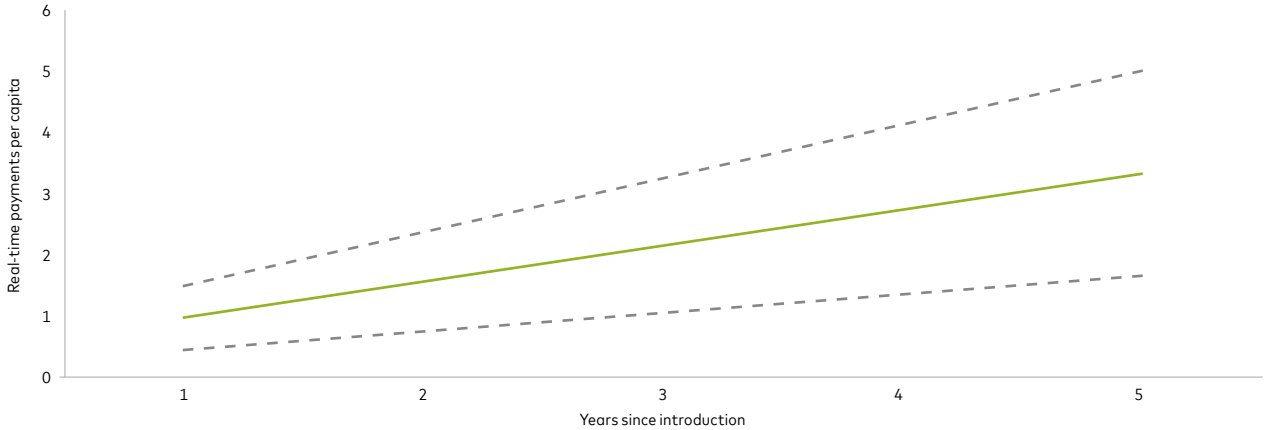
Figure 18 – Predicted take-up of real-time payments over five years following scheme introduction, with high and low take-up scenarios



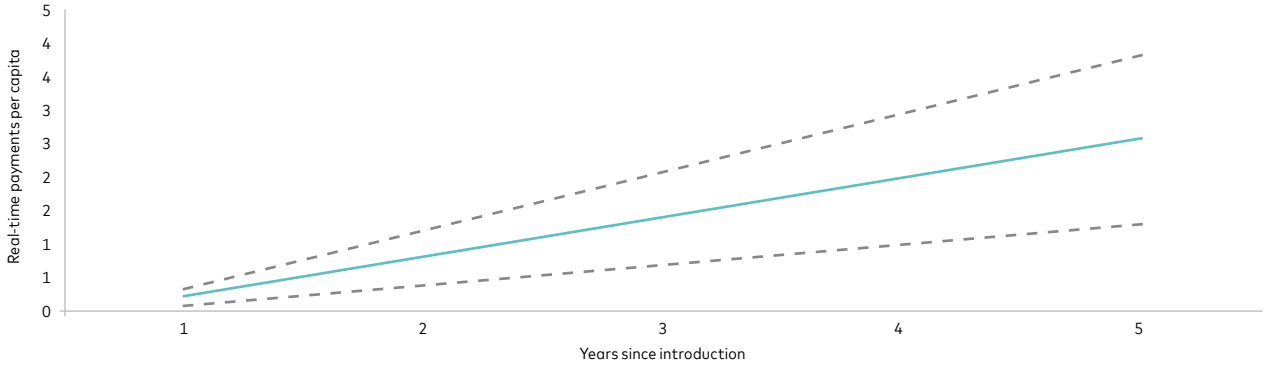
Country 1



Country 2



Country 3



Source: Deloitte analysis

Glossary

Authorisation – A step in the payments process. This refers to the first step whereby a payer initiates a payment to a payee, generally by providing the payee's information and a type of payer identification.

B2B – Business-to-Business payments, such as a payment to a supplier.

B2P – Business-to-Person payments, such as a paycheque.

BIS – The Bank for International Settlements.

Cheque – A written order requiring the payer's financial institution to pay a specified sum on demand from the payer's financial account to a specified payee, with this instruction typically deposited by the payee's financial institution and the transaction settled through a clearing and settlement system.

Clearing – A step in the payments process. This refers to a number of activities undertaken prior to payment settlement, including "transmitting, reconciling,...confirming transfer orders,...netting of orders and the establishment of final positions for settlement" (BIS, 2003).

Credit transfers – A payment order or a sequence of payment orders to transfer funds from a payer's bank account directly to a payee's bank account.

Direct debit - A preauthorised payment instruction by a payer, allowing the payee to set a payment amount and initiate payments. Once this arrangement is set up, funds are deducted automatically from the payer's bank account on a set date, with the payee allowed to set the value of the funds to be deducted.

FPTF – The Faster Payments Task Force, which was set up by the Federal Reserve bank in May 2015 to "identify and evaluate alternative approaches for implementing safe, ubiquitous, faster payments capabilities in the United States." (Faster Payments Task Force, 2016)

Real-time payments – Payments whereby "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-hour and 7-day (24/7) basis as possible." (BIS, 2016)

Net deferred settlement – A process for settlement whereby a number of payment obligations between two or more financial institutions are offset to calculate a net amount for settlement. The settlement system then allows for settling net obligations in batches (e.g. at the end of business hours, multiple times throughout the business day, etc.).

Netting – When all payment obligations between the financial institutions are offset to provide final positions for settlement.

NFC – Near-field Communication, a technology used for communication between two devices within four centimetres of each other.

Non-instant payment credit transfers – Credit transfers utilising legacy wholesale payment systems, generally with processing cycles of one working day or more.

Glossary

Notification – Confirmation by the financial institution that a payment has been made to the payer and/or payee.

P2B – Person-to-Business payments, such as a retail purchase.

P2P – Person-to-Person payments, such as a payment to an acquaintance, or in some cases, to a sole trader.

QR Code – A matrix barcode that is machine-readable, such as by a smartphone, to transfer information.

Release of funds - When funds are made available to the payee.

RTGS – Real-time Gross Settlement, whereby obligations for payments are settled, and funds are transferred between the payer and payee's financial institutions, in real-time on a gross (i.e. non-netted) basis.

Settlement – Discharging of payment obligations between the payer and payee's financial institutions, with the transfer of funds occurring on a gross basis or on a net basis.

Bibliography

Amromin, G. & Chakravorti, S. (2007). Debit card and cash usage: a cross-country analysis.

Ang, J. B. (2010). Finance and inequality: the case of India. Southern economic journal, 76(3), 738-761.

ASEAN Today (2018, 2 February). Thailand could go cashless within three years. Available at https://www.aseantoday.com/2018/02/thailand-could-go-cashless-within-three-years/

Ayyagari, M., Juarros, P., Peria, M. & Singh, S. (2016). Access to Finance and Job Growth: Firm-Level Evidence across Developing Countries. Available at https://home.gwu.edu/~ayyagari/Finance-growth.pdf

BACS (2017). Direct debit consultation outcomes. Available at https://www.bacs.co.uk/documentlibrary/dd_consultation_outcomes.pdf

Banco de Portugal (2007). Retail payment instruments in Portugal: Costs and benefits. Available at https://www.bportugal.pt/sites/default/files/anexos/pdf-boletim/study%20-%20july%202007.pdf

Banco de Portugal (2016). Social costs for retail payment instruments in Portugal. Banco de Portugal. Available at https://www.bportugal.pt/sites/default/files/anexos/pdf-boletim/estudo_custossociaisinstrumentospagamentos-2016-en.pdf

Bangkok Post (2018). BoT Tells E-Money Services to Register. Available at https://www.bangkokpost.com/business/finance/1448215/bot-tells-e-money-services-to-register

Bank for International Settlements (2003). A glossary of terms used in payments and settlement systems. Available at https://www.bis.org/cpmi/glossary_030301.pdf

Bank for International Settlements (2011). Payment, clearing and settlement systems in the CPSS countries. Available at https://www.bis.org/cpmi/publ/d97.pdf

Bank for International Settlements (2016). Fast payments – Enhancing the speed and availability of retail payments. Available at: https://www.bis.org/cpmi/publ/d154.pdf

Bank of England (2017). Stress Testing the UK Banking System: 2017 Results. Available at https://www.bankofengland.co.uk/-/media/boe/files/stress-testing/2017/stress-testing-the-uk-banking-system-2017-results

Bank of Thailand (2016). Payment Systems Annual Report 2016. Available at https://www.bot.or.th/English/PaymentSystems/Publication/PS_Annually_Report/Documents/Payment_2016_E.pdf

Bank of Thailand (2018). Payment Systems: Standardized QR Code in Thailand. Available at https://www.bot.or.th/Thai/AboutBOT/Activities/event/Documents/ADBI_bancha.pdf

Beck, T. (2007). Financing Constraints of SMEs in Developing Countries: Evidence, Determinants and Solutions. Available at http://siteresources.worldbank.org/DEC/Resources/Financing_Constraints_of_SMEs.pdf

Beck, T., Demirgüç-Kunt, A., & Levine, R. (2007). Finance, inequality and the poor. Journal of economic growth, 12(1), 27-49.

Bergman, M., Guibourg, G., & Segendorf, B. (2007). The costs of paying-private and social costs of cash and card payments. Swedish Riksbank. Available at http://archive.riksbank.se/Upload/Dokument_riksbank/Kat_publicerat/WorkingPapers/WP212.pdf

Bloomberg. (2018). Airtel Kenya bites into Safaricom market share in second quarter. Available at https://www.bloomberg.com/news/articles/2018-07-18/airtel-kenya-bites-into-safaricom-market-share-in-second-quarter

Bolt, W., Humphrey, D., & Uittenbogaard, R. (2008). Transaction Pricing and the Adoption of Electronic Payments: A Cross-Country Comparison, International Journal of Central Banking. Available at https://www.ijcb.org/journal/ijcb08q1a3.pdf

Borzekowski, R. & Kiser, E. K. (2006). The Choice at the Check-out: Quantifying Demand across Payment Instruments. International Journal of Industrial Organisation. Available at https://www.federalreserve.gov/pubs/feds/2006/200617/200617pap.pdf

Bounie, D., & François, A. (2006). Cash, check or bank card? The effects of transaction characteristics on the use of payment instruments.

Brits, H., & Winder, C. (2005). Payments are no free lunch. De Nederlandsche Bank. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.477.2604&rep=rep1&type=pdf

Burapongbandhit, P. (2017). The Impact of Perceived Risk and Perceived Benefits on Adoption and Interbank Transactions of Prompt Pay Mobile Banking in Bangkok, Chulalongkorn University. Available at http://www.eba.econ.chula.ac.th/files/upload/files/research/5645936929%20Pimpatchara%20Burapongbandhit.pdf

Business Insider (2016). AXA partners with Silicon Valley startup Trov to launch insurance 'as simple as Tinder' for British millennials. Available at http://uk.businessinsider.com/axa-insurance-startup-trov-uk-millennials-insurtech-tinder-airbnb-2016-11

Bibliography

Chan, R. (2010). Financial Constraints, Working Capital and the Dynamic Behaviour of the Firm. Available at: https://blogs.worldbank.org/files/allaboutfinance/WCpaper_FPDAcademy.pdf

Cook, W. & McKay, C. (2015). Top 10 things to know about M-Shwari. The Consultative Group to Assist the Poor. Available at http://www.cgap.org/blog/top-10-things-know-about-m-shwari

Cook, W. & McKay, C. (2017). Banking in the M-Pesa age: Lessons from Kenya. The Consultative Group to Assist the Poor. Available at http://www.cgap.org/sites/default/files/Working-Paper-Banking-in-the-M-PESA-Age-Sep-2017.pdf

Cook, W. (2017). In Kenya, bank accounts again more popular than M-Pesa – Why?. The Consultative Group to Assist the Poor. Available at http://www.cgap.org/blog/kenya-bank-accounts-again-more-popular-m-pesa-%E2%80%93-why

Cornejo, C. & Madia, M. (2015). Financial Inclusion in Action: South Africa Program Achieves Payment Usage.

Danmarks Nationalbank (2012). Costs of payments in Denmark. Denmarks Nationalbank. Available at https://www.nationalbanken.dk/en/publications/Documents/2012/04/betaling_engelsk_samlet_web.pdf

Dao, M. C. & Liu, L. Q. (2017). Finance and Employment in Developing Countries: The Working Capital Channel, IMF Working Paper No. 17/189. Available at: https://www.imf.org/en/Publications/WP/Issues/2017/08/15/Finance-and-Employment-in-Developing-Countries-The-Working-Capital-Channel-45068

Deloitte (2015). Real-time payments are changing the reality of payments. Available at https://www2.deloitte.com/content/dam/Deloitte/us/Documents/strategy/us-cons-real-time-payments.pdf

Demirgüç-Kunt, A., Klapper, L., & Singer, S. (2017). What do we know about the link between financial inclusion and inclusive growth?. [World Bank blog post]. Available at http://blogs.worldbank.org/allaboutfinance/what-do-we-know-about-link-between-financial-inclusion-and-inclusive-growth

Dobes, L., Argyrous, G., & Leung, J. (2016). Social cost-benefit analysis in Australia and New Zealand, Appendix 4. Available at http://press-files.anu.edu.au/downloads/press/p345233/pdf/app04.pdf

Drehmann, M., Goodhart, C., & Krueger, M. (2002). The challenges facing currency usage: will the traditional transaction medium be able to resist competition from the new technologies?. Economic Policy, 17(34), 193-228.

Edgar, Dunn & Company (2017). 11th Advanced Payments Report. Available at http://edgardunn.com/wp-content/uploads/2017/06/EDC_AdvancedPaymentA4_2017.pdf

European Central Bank. (2018). Central banking in the era of real-time payments. Remarks by Marc Bayle de Jessé, Director General Market Infrastructure and Payments. Available at: https://www.ecb.europa.eu/paym/intro/news/shared/2018-05-14_payments_systems_forum-speech_marc_bayle.pdf

European Payments Council (2017, 11 September). The names of the first payment service providers ready for SCT Inst are revealed. Available at https://www.europeanpaymentscouncil.eu/news-insights/news/names-first-payment-service-providers-ready-sct-inst-are-revealed

EY (2017). Reducing the shadow economy through electronic payments. Available at https://www.ey.com/Publication/vwLUAssets/Report_Shadow_Economy/$FILE/REPORT_ShadowEconomy_FINAL_17.pdf

EY (2018). Real-time payments strategy: Key focus areas for financial institutions to address. Available at https://www.ey.com/Publication/vwLUAssets/ey-real-time-payments-strategy/$FILE/ey-real-time-payments-strategy.pdf

Faster Payments (2016). Request for Payment. Available at: http://www.fasterpayments.org.uk/sites/default/files/Request%20For%20Payment.pdf

Faster Payments (Accessed 24 September 2018). HSBC. Available at http://www.fasterpayments.org.uk/hsbc

Faster Payments (Accessed 26 September 2018). Request to Pay. Available at http://www.fasterpayments.org.uk/industry-news/request-pay

Faster Payments Task Force (2016). Faster Payments Effectiveness Criteria. Available at https://fedpaymentsimprovement.org/wp-content/uploads/fptf-payment-criteria.pdf

Faster Payments Task Force (2017). The U.S. path to faster payments. Available at https://www.federalreserve.gov/newsevents/press/other/US-path-to-faster-payments-pt1-201701.pdf

Federal Reserve (2014, February 19). Fedwire Funds Services. Available at https://www.federalreserve.gov/paymentsystems/fedfunds_about.htm

Federal Reserve Bank of New York (2007). Float. Available at https://www.newyorkfed.org/aboutthefed/fedpoint/fed08.html

Bibliography

Federal Reserve Financial Services (2017). FedACH Processing Schedule. Available at https://www.frbservices.org/assets/financial-services/ach/091517-same-day-schedule.pdf

Financial Times (2016). China's Zhongan sees scope for offbeat insurance. Available at https://www.ft.com/content/bee51a52-accd-11e6-9cb3-bb8207902122

Financial Times (2017). ZhongAn launches insurtech concept to world. Available at https://www.ft.com/content/c9d10ada-9eb1-11e7-8cd4-932067fbf946

Finextra (2017). New Payments Ecosystem of Real-Time Payments. Available at https://www.finextra.com/blogposting/14618/new-payments-ecosystem-of-real-time-payments

FinExtra (2018). Push Payments and ZayZoon combine to help gig economy workers get paid instantly. Available at https://www.finextra.com/pressarticle/72444/push-payments-and-zayzoon-combine-to-help-gig-economy-workers-get-paid-instantly

Firpo, J. (2009). E-Money – Mobile money – mobile banking – What's the difference?. [World Bank blog post]. Available at http://blogs.worldbank.org/psd/e-money-mobile-money-mobile-banking-what-s-the-difference.

FIS (2017). Flavors of Fast: A Trip around the World of Immediate Payments. Available at http://empower1.fisglobal.com/rs/650-KGE-239/images/FIS_Flavors_of_Fast_2017.pdf

FIS (2018). Flavors of Fast: 2018. Available at http://empower1.fisglobal.com/rs/650-KGE-239/images/FLAVOR-OF-FAST-Report-2018.pdf

Foundation for Economic & Industrial Research (2015). Digital payments and tax revenues in Greece. Available at http://iobe.gr/docs/research/en/RES_05_F_21102015_REP_EN.pdf

Gerdes, G. R., & Walton, J. K. (2002). The use of checks and other noncash payment instruments in the United States. Fed. Res. Bull., 88, 360.

Greene, C., Rysman, M., Schuh, S., & Shy, O. (2014). Costs and benefits of building faster payment systems: the UK experience and implications for the United States. Federal Reserve Bank of Boston No. 14-5. Available at https://www.bostonfed.org/-/media/Documents/cprc-presentations/schuh10232015.pdf?la=en

Gresvik, O., & Haare, H. (2009). Costs in the Norwegian payment system. Staff Memo, 4, 2009. Norges Bank. Available at https://brage.bibsys.no/xmlui/bitstream/handle/11250/2504096/costs-in-the-payment-system.pdf?sequence=1&isAllowed=y

Gresvik, O., & Owre, G. (2002). Banks' costs and income in the payment system in 2001. Norges Bank. Available at https://www.norges-bank.no/globalassets/upload/publikasjoner/economic_bulletin/2002-04/gresvik.pdf

Gretta, S. A. A. B. (2017). Financial inclusion and growth. The Business & Management Review, 8(4), 434. Available at http://www.abrmr.com/myfile/conference_proceedings/Con_Pro_67307/conference_78515.pdf

Guibourg, G., & Segendorf, B. (2004). Do Prices Reflect Costs? A study of the price-and cost structure of retail payment services in the Swedish banking sector 2002 (No. 172). Sveriges Riksbank Working Paper Series.

Harasim, J. (2016). Europe: the shift from cash to non-cash transactions. In Transforming Payment Systems in Europe (pp. 28-69). Palgrave Macmillan, London.

Hataiseree, R., & Banchuen, W. (2010). The effects of e-payment instruments on cash usage: Thailand's Recent Evidence and Policy Implications. Available at: https://pdfs.semanticscholar.org/047c/08f6b632b11d37904ffad034556d598f4c97.pdf.

Hayashi, F. & Keeton, W. (2012). Measuring the costs of retail payment methods. Federal Reserve Bank of Kansas City. Available at https://www.kansascityfed.org/publicat/econrev/pdf/12q2Hayashi-Keeton.pdf

He, L. (2018). Big Data and AI – The Future of Car Insurance According to China's First Internet-Only Insurer, South China Morning Post. Available at https://www.scmp.com/business/companies/article/2131948/big-data-and-ai-future-car-insurance-according-chinas-first

Hong Kong Monetary Authority (2018, 17 September). Press releases: The Launch of Faster Payment System (FPS). Available at https://www.hkma.gov.hk/eng/key-information/press-releases/2018/20180917-3.shtml

HSBC (2018). Instant payment schemes: The future of payments?. Available at https://www.gbm.hsbc.com/-/media/gbm/reports/insights/instant-payment-schemes.pdf

Humphrey, D. B., Kim, M., & Vale, B. (2001), Realizing the Gains from Electronic Payments: Costs, Pricing, and Payment Choice. Journal of Money, Credit and Banking, 33(2), 216-34.

Humphrey, D. B., Pulley, L. B., & Vesala, J. M. (1996). Cash, paper, and electronic payments: a cross-country analysis. Journal of Money, Credit and Banking, 28(4), 914-939.

Humphrey, D., Willesson, M., Bergendahl, G., & Lindblom, T. (2006). Benefits from a Changing Payment Technology in European Banking., Journal of Banking and Finance 30(6). 1631-52.

Bibliography

Iaozzolino, G., & Wasike, N. (2015). The unbearable lightness of digital money. Journal of Payments Strategy & Systems, 9(3), 229-241. Available at http://s3-eu-central-1.amazonaws.com/fsd-circle/wp-content/uploads/2015/11/30093729/The-unbearable-lightness-of-digital-money.pdf

Immordino, G., & Russo, F. F. (2016). Cashless payments and tax evasion. CSEF Working Paper No. 445. Centre for Studies in Economics and Finance. Available at http://www.csef.it/WP/wp445.pdf

Integrated Payment Services (2017, 16 February). Banking Industry Unveils PesaLink Brand. Available at https://ipsl.co.ke/2017/02/16/banking-industry-unveils-pesalink-brand/

Jonker, N. (2013), Social costs of POS payments in the Netherlands 2002 – 2012: Efficiency gains from increased debit card usage, DNB Occasional Studies 11(2).

Kalckreuth, U. V., Schmidt, T., & Stix, H. (2014). Using cash to monitor liquidity: implications for payments, currency demand, and withdrawal behavior. Journal of Money, Credit and Banking, 46(8), 1753-1786.

KPMG (2018). Consumer and retail updates: Industry in focus. Available at https://assets.kpmg.com/content/dam/kpmg/th/pdf/2018/02/th-consumer-and-retail-updates-january-2018.pdf

Kruger, M., & Seitz, F. (2014). Costs and benefits of cash and cashless payment instruments: Overview and initial estimates. Deutsche Bundesbank. Available at https://www.bundesbank.de/Redaktion/EN/Downloads/Publications/Studies/costs_and_benefits_of_cash_2014.pdf?__blob=publicationFile

Lacker, Jeffrey M. (1997). The Check Float Puzzle, Federal Reserve Bank of Richmond, Economic Quarterly 83: 1 – 25. Available at: https://www.richmondfed.org/~/media/richmondfedorg/publications/research/economic_quarterly/1997/summer/pdf/lacker.pdf

Ledger Insights (2018). ZhongAn automating health insurance claims. Available at https://www.ledgerinsights.com/zhongan-health-insurance/

Leibbrandt, G. (2010). You can't get there from here: The role of the installed base in payment system adoption. Journal of Payments Strategy & Systems, 4(1), 6-16.

Lipis Advisors (2018). Serving the underserved. Available at https://iqbusiness.net/wp-content/uploads/2018/07/LIPIS_Report1.pdf

Lipis Advisors, IQbusiness, BankservAfrica, & PASA (2017). Considerations for Modernising South Africa's Low-Value Payment Infrastructure.

Majesco. (2018). Future Trends 2018. Available at https://www.the-digital-insurer.com/wp-content/uploads/2018/05/1188-Future-Trends2018.pdf

Masinde, J. (2017). Kenyan banks have joined forces to launch a mobile money rival to M-Pesa. Quartz Africa. Available at https://qz.com/africa/913761/safaricoms-m-pesa-has-a-new-rival-in-pesalink-from-kenyas-banks/

MasterCard & SASSA. (2016). Where Government Efficiency Meets Financial Inclusion.

Mbiti, I., & Weil, D. N. (2011). Mobile banking: The impact of M-Pesa in Kenya (No. w17129). National Bureau of Economic Research. Available at http://www.nber.org/papers/w17129.pdf

McAndrews, J. & Roberds, W. (2000). The Economics of Check Float, Federal Reserve Bank of Atlanta, Economic Review, Fourth Quarter. Available at https://www.frbatlanta.org/-/media/documents/research/publications/economic-review/2000/vol85no4_mcandrews-roberds.pdf

McKinsey & Company (2015, March). Digital Banking in ASEAN: Increasing Consumer Sophistication and Openness. Available at https://www.mckinsey.com/~/media/mckinsey/business%20functions/mckinsey%20digital/pdf/digital%20banking%20in%20asean%20(4).ashx

Medina, L., & Schneider, F. (2018). Shadow economies around the world: What did we learn over the last 20 years?. IMF Working Paper No. 18/17. International Monetary Fund. Available at https://www.imf.org/~/media/Files/Publications/WP/2018/wp1817.ashx

Milligan, B. (2017, March 22). Bank cheques to be cleared within a day. Available at: https://www.bbc.co.uk/news/business-39351793

Milne, A., & LeiLei, T. (2005). An Economic Analysis of the Potential Benefits and Dis-benefits of Faster Payments Clearing, Office of Fair Trading No. 795. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.573.5017&rep=rep1&type=pdf

MobileTopUp (2018). Does PromptPay have a future? Available at https://medium.com/mobiletopup/does-promptpay-have-a-future-a3c35efa6bb5

MoneySENSE (2018, August 15). Fast and Secure Transfers (FAST). Available at http://www.moneysense.gov.sg/understanding-financial-products/banking-and-cash/types-of-products-and-services/fast.aspx

National Bank of Belgium. (2006). Costs, advantages and drawbacks of the various means of payment. Economic Review, June, 41-49. Available at https://www.nbb.be/doc/ts/publications/economicreview/2006/ecorevi2006_h3.pdf

Bibliography

National E-Payment. (2017). National E-Payment Opening Doors to Thailand 4.0 and Digital Economy. Available at: http://www.nationmultimedia.com/advertisement/business/30

New Payments Platform (2018, 13 February). New Payments Platform Media Release: The new payments platform launches. Available at https://www.nppa.com.au/wp-content/uploads/2018/02/New-Payments-Platform-Financial-Services-Media-Release.pdf

Nienhuis, J., Cortet, M., & Lycklama, D. (2013). Real-Time Financing: Extending E-Invoicing to Real-Time SME Financing. Journal of Payments Strategy & Systems, 7(3).

Office of Fair Trading (2005). Press release: Payment Systems Task Force announces faster payments service. Available at http://webarchive.nationalarchives.gov.uk/20131101202847/http://www.oft.gov.uk/news-and-updates/press/2005/94-05

Open Forum (2016). Open Forum on Pan-European Real-time payments. Available at https://www.abe-eba.eu/media/azure/production/1508/03rd-may-2016-presentaion-slides.pdf

Paym (2016, September). Paym statistical update. Available at https://www.paym.co.uk/wp-content/uploads/2016/09/Paym-dashboard-Aug-2016.pdf#

Payments Canada (2018, 29 June). Corporate Plan: 2018 to 2022. Available at https://www.payments.ca/sites/default/files/29-Jun-18/pc20182022_corporate_plan_en_final.pdf

PR Newswire (2017). Lemonade Sets New World Record. Available at https://www.prnewswire.com/news-releases/lemonade-sets-new-world-record-300386198.html

PYMNTS (2018). Gig Workers And Instant Pay. Available at https://www.pymnts.com/digital-payments/2018/uber-lyft-gig-economy-instant-pay/

Sadlakowski, D. (2017). The role of instant payment systems in the Polish economy. Copernican Journal of Finance & Accounting, 6(2), 59-69. Available at http://apcz.umk.pl/czasopisma/index.php/CJFA/article/viewFile/CJFA.2017.011/13548

Safaricom (Accessed 2 October 2018). M-PESA Rates. Available at https://www.safaricom.co.ke/personal/m-pesa/getting-started/m-pesa-rates

Sahay, R. A. T. N. A., Čihák, M., N'Diaye, P., Barajas, A., Mitra, S., Kyobe, A., ... & Yousefi, S. R. (2015). Financial inclusion: Can it meet multiple macroeconomic goals? IMF Staff Discussion Note, 15/17. International Monetary Fund, Washington, DC. Available at https://www.imf.org/external/pubs/ft/sdn/2015/sdn1517.pdf

Schmiedel, H., Kostova, G., & Ruttenberg, W. (2012). The Social and Private Costs of Retail Payment Instruments: A European Perspective. European Central Bank Occasional Paper Series No. 137. Available at https://www.ecb.europa.eu/pub/pdf/scpops/ecbocp137.pdf

Schneider, F. (2017). Restricting or Abolishing Cash: An Effective Instrument for Fighting the Shadow Economy, Crime and Terrorism?. Deutsche Bundesbank. Available at https://www.econstor.eu/bitstream/10419/162914/1/Schneider.pdf

Schuh, S., & Stavins, J. (2010). Why are (some) consumers (finally) writing fewer checks? The role of payment characteristics. Journal of Banking & Finance, 34(8), 1745-1758.

Schwartz, C., Fabo, J., Bailey, O., & Carter, L. (2008). Payment Costs in Australia. In Payments System Review Conference (pp. 88-138). Available at https://www.rba.gov.au/payments-and-infrastructure/resources/publications/payments-au/paymts-sys-rev-conf/2007/7-payment-costs.pdf

Segendorf, B. L., & Jansson, T. (2012). The cost of consumer payments in Sweden. Swedish Riksbank. Available at http://archive.riksbank.se/Documents/Rapporter/Working_papers/2012/rap_wp262_120619.pdf

SIX Group (2018). Annual Press Conference 2018. Available at https://www.six-group.com/dam/download/publications/0321-e-six-annual-results-2017.pdf

Stewart, C., Chan, I., Ossolinski, C., Halperin, D., & Ryan, P. (2014). The Evolution of Payment Costs in Australia. Reserve Bank of Australia. Available at https://www.rba.gov.au/publications/rdp/2014/pdf/rdp2014-14.pdf

The Clearing House (2018). Spring 2018 Updates: Faster, smarter and safer payments for everyone. Available at https://www.theclearinghouse.org/payment-systems/-/media/2667e63d23fb40d7bbf7dd6c3b5df1a0.ashx

The Economic Times (2017). Google Payment App Tez Sees a Slow Start; People Call It Buggy. Available at https://economictimes.indiatimes.com/tech/software/google-payment-app-tez-will-make-payments-as-simple-as-cash-caesar-sengupta/articleshow/60730054.cms

The Economic Times (2018). Ola, PhonePe ally for easy digital pay ride. Available at https://economictimes.indiatimes.com/small-biz/startups/newsbuzz/ola-phonepe-ally-for-easy-digital-pay-ride/articleshow/64595364.cms

The Economist (2015). Why does Kenya lead the world in mobile money?. Available at https://www.economist.com/the-economist-explains/2015/03/02/why-does-kenya-lead-the-world-in-mobile-money

Bibliography

The Telegraph (2017). Would you buy 'micro-insurance' that tracks your every move?. Available at https://www.telegraph.co.uk/insurance/car/would-buy-micro-insurance-tracks-every-move/

The UK Cards Association (2017). UK Card Payments Summary. Available at http://www.theukcardsassociation.org.uk/wm_documents/UK%20Card%20Payments%202017%20-%20Summary%20FINAL.pdf

Trütsch, T. (2014). The impact of contactless payment on spending. International Journal of Economic Sciences, 3(4), 70-98.

Turján, A., Divéki, É., Keszy-Harmath, É., Kóczán, G., & Takács, K. (2011). Nothing is free: A survey of the social cost of the main payment instruments in Hungary (No. 93). Magyar Nemzeti (Hungarian Central) Bank Occasional Papers. Available at https://www.mnb.hu/letoltes/op93-fmtk.pdf

Vertafore (2017). New research from Vertafore reveals millennial love-hate relationship with insurance, valuing coverage but prioritizing Netflix. Available at https://www.vertafore.com/resources/press-releases/new-research-vertafore-reveals-millennial-love-hate-relationship-insurance

Vocalink (2013). Mobile usage: attitudes and payments research report 2013. Available at https://www.Vocalink.com/downloads-and-media/research/mobile-usage-attitudes-payments/

Vocalink (2016, 12 December). The Future of RTGS. Available at https://connect.Vocalink.com/2016/december/the-future-of-rtgs/

Vocalink (2018a). Financial inclusion accelerated: How real-time payments are bringing millions of people into the economic fold. Available at https://www.Vocalink.com/real-time-payments/financial-inclusion/

Vocalink (2018b). Real-Time Payments Powering Thai Businesses. Available at: https://connect.Vocalink.com/2018/may/real-time-payments-powering-thai-businesses/

Vocalink (Accessed 25 September 2018). Thailand: PromptPay. Available at https://www.Vocalink.com/real-time-payments/promptpay/

Wells, K. E. (1996). Are checks overused? Federal Reserve Bank of Minneapolis Quarterly Review, 20(4), 2–12. Available at https://core.ac.uk/download/pdf/6608960.pdf

World Bank (2016). Retail Payments: A practical guide for measuring retail payment costs. Available at http://documents.worldbank.org/curated/en/255851482286959215/pdf/111216-WP-P155382-PUBLIC-ABSTRACT-SENT.pdf

Footnotes

1. BIS notes that this step may be condensed with the "settlement" stage in some sources, although this is incorrect. The distinction of these two steps is important for real-time payments services as currently set up in most existing schemes.

2. Note that BIS defines real-time payments as "real-time or near-real-time". BIS further clarifies that "the characteristics of [instant] payments may also vary by jurisdiction" and therefore the "definition is not intended to be precise in relation to the specific speed and service availability that qualifies as a [instant] payment". BIS also notes that in some cases borderline cases may exist, and that some systems may have initially not fit the definition of "instant" payments but may have been further developed to meet the criteria (BIS, 2016).

3. Year commenced refers to when the scheme offered real-time or near-real-time payment functionality at or near 24/7 availability. This may differ for some countries from the year of introduction of the scheme more generally – see Notes. For the purposes of the econometric analysis in this study, real-time payments schemes are defined according to the characteristics identified by BIS (2016), with some allowances for boundary candidates regarding 24/7 availability of the scheme in order to obtain a sufficient sample size.

4. Gaps in the figures are attributed to the unavailability of data in initial years since introduction.

5. Mobile penetration values over 100% reflect that some consumers have multiple mobile phones. The mobile penetration rate is measured as the number of mobile phone subscriptions divided by the population. Data on the percentage of the population that owns at least one mobile phone is available for a sufficient range of countries and dates.

6. 585m to 1,228m; Deloitte analysis based on data from the Bank of Thailand.

7. Central banks include the Reserve Bank of Australia, Banco Central do Brasil, Reserve Bank of India and Banco de Mexico. Further details on data sources are provided in Section A2 Background.

8. Cards are outside the scope of this study and were not found to have a statistically significant link with real-time payments.

9. Statistical significance is assessed at the 5% level.

10. For example, the UK's CHAPS system charges £20 per transaction while banks charge end-users USD 20-USD 40 per transaction using the US's FedWire system.

11.The FedWire system is available Monday through Friday from 9:00 PM on the preceding calendar day to 6:30 PM, with a deadline of 6:00 PM for making payments (Federal Reserve, 2014).

12. Note that in some cases, substitution from legacy non-instant payment credit transfers to real-time payments may be mandated. In the UK, standing orders were required from 1 January 2012 to reach payee accounts within the same day as debiting from the payer's account as part of the Payment Services Regulations 2009. This resulted in the migration of all standing orders to FPS, the UK real-time payments scheme.

13. Six day processing figure was from prior to the introduction of cheque image-based clearing.

14. More detail on the econometric results in this study can be found in the Technical appendix.

15. An alternative source of data on cash was the number of cash withdrawals from ATMs or bank counters. However, this data would not allow for direct translation to cash transactions and may exclude cash already in circulation.

16. It is worth pointing out that "social costs" are defined differently in the payments literature as compared to other fields in economics. In wider economics literature, social costs equate to the sum of private and external costs. In the payments literature, social costs are defined as the net sum of all private (resource) costs borne by the payment system. They do not include what would be considered social costs in wider economics literature, for example environmental costs from usage of paper-based vs electronic payment instruments. This study therefore avoids the use of the term where possible and uses "system costs".

17. Deloitte calculations based on estimates from the study by Segendorf and Jansson, 2012.

18. The importance of difference cost elements, and the balance between fixed and variable costs, varies across payment instruments. A higher degree of fixed costs suggests stronger economies of scale, as net payment system costs per transaction decrease more steeply with higher volumes of transactions for the payment instrument.

Footnotes

19. In its consultation paper, "A Practical Guide for Measuring Retail Payment Costs", the World Bank recommends identifying major cost elements for each payment instrument and then collecting primary data on the relevance of each payment method, along with the monetary value of costs associated with it via a survey. This allows for a direct re-calculation of fixed and variable costs before and after the change in the 'payment mix', and may be considered the most comprehensive method of estimating savings in net payment costs. Such an exercise is beyond the scope of a general paper like this, however. Direct estimates of fixed and variable costs by countries that would allow for calculation as per the World Bank's recommended method are limited. The analysis is therefore based on a sample from the existing literature of 61 estimates of net payment system costs per transaction for different payment instruments combined with the volumes corresponding to these costs (a list of studies is available in Appendix B1).

20. As discussed in Schmiedel et al., (2012), it cannot be directly concluded from the curves that a steeper slope indicates stronger economies of scale. The ECB study notes that "to draw further conclusions regarding the statistical significance of economies of scale among retail payment instruments, one would need to study these effects by conducting a more detailed econometric analysis [which would] require detailed data to be available over a longer time period and a larger number of sample countries". Similar to the ECB study, the analysis finds there is more limited evidence of a clear trend for cheques.

21. This includes all types of credit transfers, as the source studies do not differentiate between instant credit transfers and non-instant credit transfers. Note that in the majority of the studies covered, real-time payments schemes had not been introduced.

22. Axes for the volume of transactions per capita have been limited to observed data ranges.

23. The literature sometimes defines the delay in payments in the financial system as the "payment float", although there does not seem to be a generally accepted definition of float as contrasted with all other types of timing differences (Federal Reserve Bank of New York, 2007). In general, payment floats may be defined as the double-counting of money in the banking system due to the time difference in when payments are debited from the payer and credited to the payee.

24. For the purpose of this calculation, Germany's compounded social discount rate of 3% per annum was uncompounded to a daily rate (Dobes, Argyrous & Leung, 2016)

25. Tax revenue includes transfers to the central government only and excludes fines, penalties, and social security contributions.

26. The VAT gap is defined as the difference between theoretical VAT liability and actual VAT revenue.

27. More recent estimates have suggested this is between 90% and 98% of all retail transactions in Kenya (Newsweek, 2017; Business Daily, 2017).

28. In August 2017, the chairman of the Thailand E-Payment Trade Association stated that Thailand could become a cashless society within three years.

29. See Section 3.2.3 for further details.

30. In addition to direct competition on mobile payments through industry coordination on PesaLink, other forms of competition have included launching in-house mobile network and money schemes (e.g. Equitel) and collaboration with some mobile money providers to offer new banking services (e.g. M-Shwari; Cook & McKay, 2015, 2017).

31. Deloitte calculations from World Bank data.

32. Deloitte calculations from World Bank data.

33. These scenarios are defined as 50% above (for the high take-up scenario) and 50% below (for the low take-up scenario) the econometrically predicted take-up profile.

34. The unweighted average standard VAT rate in OECD countries is 19.2% (Source: OECD data). This is compared to a global average of 15.4% (Source: KPMG).

35. https://www.bis.org/statistics/payment_stats.htm?m=3%7C16%7C385

36. These countries are Australia, Belgium, Brazil, Canada, China, France, Germany, India, Italy, Japan, Korea, Mexico, Netherlands, Russia, Saudi Arabia, Singapore, South Africa, Sweden, Switzerland, Turkey, UK and USA.

37. Whilst data is available from 1980, only data from 1991 are used due to data quality issues.

38. Euromonitor cash data is not available for Belgium and Switzerland.

39. Bank for International Settlements (2016). Fast payments – enhancing the speed and availability of retail payments. Available at: https://www.bis.org/cpmi/publ/d154.pdf.

40. Further details on real-time payments schemes are discussed in Section 4.2.

41. See Footnote 5.

42. The real-time payments schemes in Thailand and Australia are PromptPay and New Payments Platform respectively.

Footnotes

43. The SPEI scheme in Mexico operated on a 21/7 basis from 2004 and on a 24/7 basis from 2015.

44. For example, cheques ceased to exist in the Netherlands from 2003.

45. An alternative variable would indicate the number of years since real-time payments were introduced. This variable is linear, deterministic, and hence implies a constant growth rate.

46. This dummy variable would distinguish actual zeros from potential zeros. Actual zeros arise when real-time payments are available but no real-time payments transactions are made. Potential zeros arise when real-time payments are not available.

47. Model parsimony provides degrees of freedom for statistical tests.

48. Mathematically, the instant payment variables are endogenous if they are correlated with the idiosyncratic error terms $\{\varepsilon_{m,it}\}_{m=1}^{5}$.

49. Regional dummies are based on continents: Africa, Asia, Australia, East Europe, Latin America, North America and Western Europe.

50. The Fixed Effect estimator utilises variation across time (within variation) rather than across country (between variation) to identify the model parameters.

51. Outlier dummies are not included in equations with country dummies as perfect multicollinearity would arise.

52. Country dummies are used in the cash equation as the cross-country variation in cash volumes appears to be more substantial than other payment methods' based on time series plots of different payment instrument volumes.

53. The statistical insignificance may be driven by the lack of variation in the bank branches over time.

54. This variable is an interaction between log real GDP per capita and proportion of old population.

55. The use of the natural logarithm of real GDP per capita obtains the impact of a 1% change in real GDP per capita on transaction volume.

56. In the presence of weak instruments, the asymptotic distribution of the 2SLS estimator is unknown, thus affecting the ability to conduct statistical tests.

57. Further details of these tests are reported in Section A6.

58. The Sargan test is only available when the number of instruments exceed the number of endogenous variables.

59. Heteroscedasticity-robust standard errors are used in all model specifications.

60. The p-value for reject the null hypothesis of $H_0: \gamma_{\text{cash2}} = 0$ is 0.052, so the estimate of γ_{cash2} is marginally insignificant.

61. If the rejection rate in a Wald test at 5% level is tolerated at 15%, the test concludes that the instruments are not weak.

62. No overidentifying restrictions available, so this test is unable to be conducted.

63. The test result suggests that non-linear combinations of fitted values can help improve the fit of the model.

64. Estimates are provided for the payment instruments with a significant econometric impact from real-time payments take-up and cash (see Section 2).

65. Excluding fines, penalties, and social security contributions.

66. As discussed in sections 3.2.1 and 4.4.2, it is not possible to econometrically estimate the impact of real-time payments on cash directly due to data limitations. In the hypothetical examples in this section, an indicative impact is assumed to estimate the impact real-time payments may have. Note that realisation of any impacts on cash are likely to depend on scheme implementation characteristics that ensure access and ease-of-use so that real-time payments can serve as a suitable substitute to cash transactions (see sections 3.2.1 and 4.4.2).

Contact us
info@vocalink.com
vocalink.com

Head Office
1 Angel Lane
London
EC4R 3AB
United Kingdom

DESIGNED BY 20BELOW

140264/03.091019