INTERNATIONAL MONETARY FUND

Instant Payments:

Regulatory Innovation and Payment Substitution Across Countries

Tanai Khiaonarong and David Humphrey

WP/22/228

2022
NOV

WORKING PAPER

 WP/22/228

IMF Working Paper
Monetary and Capital Markets Department

Instant Payments: Regulatory Innovation and Payment Substitution Across Countries

Prepared by Tanai Khiaonarong and David Humphrey*

Authorized for distribution by Tommaso Mancini-Griffoli
November 2022

ABSTRACT: Instant, or fast, payments are credit transfers completed and settled within seconds or minutes. They have low costs, reduce payment risk, and have significantly replaced the use of cash, cards, or check and direct debit payments. We note the role played by regulators in promoting instant payments and identify instances of significant payment instrument substitution across 12 advanced and emerging market economies. This substitution reflects the realized demand for attributes offered by instant payments. As these attributes are quite similar to those for CBDC, the demand for retail CBDC (if issued) may be less compelling.

JEL Classification Numbers:	E41, E42, E58, E59
Keywords:	Instant payments, payment substitution, regulatory innovation
Author's E-Mail Address:	tkhiaonarong@imf.org; dhumphrey@cob.fsu.edu

* We thank the following central bank officials who have provided assistance in obtaining payment statistics for this study: Ryan Rizaldy and Jacky, Bank Indonesia; Sharon Seah Shing Bey, Bank Negara Malaysia; Othón Martino Moreno González, Bank of Mexico; Orchuma Prachasaisoradej, Bank of Thailand. For comments on the paper, we thank Itai Agur, Parma Bains, Arvinder Bharath, Alexis Boher, Aquiles Farias, Arif Ismail, Minsuk Kim, Harry Leinonen, Marco Pani, Rangachary Ravikumar, Oz Shy, Manmohan Singh, Jeanne Verrier, Hou Wang, and Kjetil Watne. Erica Sandoval and Tatiana Lenshina provided editorial assistance.

WORKING PAPERS

Instant Payments:

Regulatory Innovation and Payment Substitution Across Countries

Prepared by Tanai Khiaonarong and David Humphrey

Contents

1. Introduction

Instant payments have been set in motion by regulatory authorities and/or central banks, at times working with the private sector or offering the payment service themselves. The overall goal is to reduce current delays and expense in making a domestic payment transaction. This improves the cost efficiency and operating effectiveness of a country's payment system while at the same time reducing payment system risk. While some 55 countries have implemented instant payments between 2001 and 2019 and at least 10 more are projected to do so (Bech and Hancock, 2020), available data on instant payments and other instruments are limited. Hartmann, et al. (2019) have earlier investigated the use of instant payments across six countries (four of which are also covered here) and assess the factors driving their adoption and use.

The benefits of instant payment (IP) systems and their effect on other payment instruments are illustrated by examining a selection of twelve countries--both advanced and emerging market economies, where sufficient data are available. We cover Australia, China, India, Indonesia, Malaysia, Mexico, Norway, Russia, Singapore, Sweden, Thailand, and United Kingdom. Additional benefits are possible if IP systems in different countries can be connected so lower value instant payments become possible for cross-border retail payments, travel, and remittances. Such arrangements are currently in the testing stage and actually offered between some countries (BISIH, 2021; BOT-MAS, 2022).

The benefits of instant payment systems are outlined relative to the use of cash, cards, and debit instruments (checks and direct debits) in Section 2. The fact that payment transactions are completed and settled (via a central bank) in seconds or minutes, along with the associated reduction in payment risk, is an important reason why regulatory authorities have been involved in the development of instant payments (CPMI, 2016). In Section 3, the role of these authorities in catalyzing, overseeing, or operating instant payment systems is illustrated for the twelve countries noted above.

Section 4 illustrates the change in payment composition using domestic values of different payment instruments in each country over 2005-2020 (16 years). While these figures show how values of certain payment instruments have expanded while others have fallen, it is difficult to visually hold other things constant in order to more accurately determine payment substitutions and their statistical significance. This requires a more formal analysis.

An analysis using an econometric model of payment substitution is provided in Section 5. This uses data on payment shares in order to "normalize" the differences in relative payment values to facilitate cross-country comparisons. Our approach is similar to past econometric efforts to estimate the elasticity of physical capital substitution for labor in producing national output. Such substitution indicates the distribution of national income between these two factors of production.

Similarly, estimates of payment substitution, shown in Section 6, reflect the underlying, but unknown, distribution of payment costs, revenues, and non-monetary benefits that drive the use of different payment instruments at the national level. This includes purchases of the many inputs involved in the production of GDP as well as purchases associated with the distribution of GDP to their various final demands.

The degree to which instant payments have been adopted and substituted for other payment instruments reflects the extent that the productive efficiency of a country's payment system has improved by supplying (usually) lower cost payment methods. It also indicates the extent that risk in the payment system has been reduced. Both are important payment policy objectives. Our conclusions are in Section 7 and an Appendix contains additional information on payment substitution and notes the sources of our data.

2. Comparing Instant Payments with Other Payment Instruments

Individuals and firms in our selected countries have five main ways to make a payment: physical cash, debit and credit cards (which includes e-money), instant payments (IP), the use of generally higher value check and direct debit transactions, and finally, wire transfers (large value credit transfers). In the empirical analysis, check and direct debit transactions are grouped together (denoted as ChecksDD) as they are both debits to an account and are used for similar types and values of payments.

Historically, the US and Europe used checks and paper giro payments to replace cash while direct debits (via automated clearinghouses) and electronic giro payments are basically just the electronic form of checks and earlier paper giro transactions. Large value wire transfers appear to be functionally separable from the four other payment instruments and we exclude them from our analysis.[1] Functional separability of large value wire transfers from other payments allows us to focus statistically on the four remaining payment instruments where substitution is occurring.

Depending on how instant payments are implemented, they can substitute for any or all of the other payment instruments. The main difference between instant payments (IP) and cash, cards, or debit payment instruments (ChecksDD) is that they are initiated and settled (usually with a confirmation to both parties) within a few seconds or minutes on a 24/7 basis (CPMI, 2016, Annex 2). This means that receivers of IP funds have them available for immediate reuse for new out-payments through their bank. The other instruments all have a delay of from 1 to 3 days before point-of-sale income or revenues from accounts receivable can be reused through their bank.

Cash received by merchants at the point-of-sale usually has to be deposited at a bank before it is available for reuse. Thus cash is one-day to two-day money. The past practice of paying cash wages to workers was replaced by check, ACH, and giro payments in most countries decades ago which, in turn, is now is being replaced by credit transfers to deposit accounts. Although checks are still paper instruments, the information on a check in some countries can be scanned and sent electronically for collection and settlement for next-day or, sometimes, two-day funds availability after being deposited at a bank.

When checks were physically collected, transported, and presented for payment, the payment delay could be much longer depending on the size of the country and the concentration of the banking system. This was a result of commercial laws that required the physical transportation of the original check (by plane, train, or van) to the paying bank to verify the signature on the check prior to paying it. Today, signature verification only occurs for high value checks.

[1] According to a recent analysis (CPMI, 2021, p. 3), "there is little evidence hither to that IPS [Instant Payment Systems] are 'cannibalising' RTGS system volumes and values". This is likely due to limits currently placed on the transaction value of instant payments.

In the past, cards generated paper slips for each transaction. They were deposited at a bank and collected and processed in a manner similar to paper checks of the time. Today, card payments are initiated with debit or credit cards swiped, inserted into, or tapped on electronic terminals used to transmit payment information between the user's and the merchant's bank. The interbank transaction is then settled through the central bank (or commercial bank). Mobile phones can now be used to initiate card-based payments (as well as instant payments). Even so, merchant revenues received from card transactions are not credited to a merchant's bank account immediately but may incur a delay of up to three days.

Instant payments reduce payment risk in several ways. First, these payments are credit transfers which will not occur unless adequate funds are available in a deposit account to cover the transfer, although some countries are considering making debit transfers “instant” as well. Also, settlement of instant payments by the central bank is (effectively) immediate and guaranteed. This reduces the time gap between when a payment is made and when it is settled by a trusted intermediary with adequate liquidity, reducing the likelihood of a settlement failure.[2]

In some countries final settlement of an instant payment is at present delayed to later in the day via net settlement. Even so, the receiver can still immediately reuse the funds received. While some risk of settlement failure is retained, it is contained by placing limits on the value of customer transactions, capping the net value of payments an institution can make prior to settlement, and/or by the posting of liquid collateral for delayed same-day net settlement (CPMI, 2021).[3] Importantly, countries have adjusted their bankruptcy laws, or passed new laws, so that payments made and received cannot be unwound or reversed in the event of the bankruptcy of the sending party in a payment transaction.

To summarize, the main benefits of instant payments compared to other payment instruments involve: immediate use of funds versus delayed availability; typically, lower payment costs; no or little restriction on payment operating hours; and no, or minimum settlement risk. Instant payments can also facilitate straight-through electronic processing of business account payable and receivable operations. This can further reduce back-office business payment expenses when the payment, and the internal account payable and receivable invoice information, are fully electronic between payor and payee.[4]

[2] Even so, private sector arrangements exist that are almost as good where participating banks jointly contribute funds to handle instant or same-day settlement outside of the central bank (e.g., the Real Time Payment service from The Clearing House in New York).

[3] In a country where kidnapping and forcing a person to transfer funds to a separate account under duress has occurred, this fraud risk was reduced when an instant payment service was turned off during the night and/or transaction limits on the amount of funds that could be transferred were reduced.

[4] While some/many large firms already have straight-through processing in some countries, the availability of instant payments makes it possible to extend straight-through processing to other firms as well.

3. Regulatory Innovation in Payments

Three developments have made instant payments possible. These are: (1) a general lowering of telecommunication and computer processing costs, which increased the penetration rate of the Internet and mobile devices; (2) higher expectations from users for accessible, low-cost, and around-the-clock payment services; and (3), regulatory changes that enable, and at times directly promote, efforts to reduce delays in completing payments. A key factor has been the willingness of the central bank to extend operating hours to improve liquidity management and to clear and settle interbank transfers on a real-time, 24/7 basis (CMPI, 2022). While in some countries same-day deferred net settlement was originally used instead of RTGS, with appropriate safeguards to ensure delayed settlement, many have since implemented real-time settlement.

Central bank and relevant regulatory authorities have played different roles in the implementation of instant payment systems—as catalyst, overseer, and operator.[5] As a catalyst, it is common in many countries for authorities to be at the center in modernizing national payment systems (CPSS, 2006). This can take the form of central banks setting national strategies, such as payment roadmaps (India, Indonesia, Singapore, and Thailand), to involving the finance ministry and the competition authority in the review of payment competition, innovation, and efficiency (UK). The strategic direction for developing instant payments in Australia was set by the Reserve Bank of Australia's (RBA) Payment Systems Board which has a mandate relating to efficiency and competition in the payments system. The catalyst role can also involve coordinating linkages between privately-operated instant payment systems to facilitate cross-border payments between countries (Indonesia, Malaysia, Singapore, and Thailand). Apart from IP, central banks have stimulated payments innovation with the establishment of regulatory sandboxes to assess new financial technologies, experiment with distributed ledger technology in payments and settlements, and pilot test CBDC.

As overseer, or payment system regulator, authorities have sought to stimulate innovation in downstream payment applications by accommodating or directly providing instant payments which serves as the basis for new payment applications by banks and fintech firms, while also reducing payment risk. This may also involve having differential transaction limits on existing large value wire transfer (credit transfer) systems to accommodate instant payments and shortening the settlement cycles on existing real-time gross settlement systems (India). Another possibility is setting requirements for all standing orders to be settled within a day of submission (UK).

Finally, the role of the central bank as an operator varies from (1) being limited to just providing real-time settlement (and access to an intraday liquidity facility), to (2) being active in the governance of the firm providing the instant payment service, to (3) owning and running the instant payment service themselves. The first case applies to Norway, Thailand, and the UK. For other countries, in addition to providing real-time settlement, central banks are also involved in the governance of instant payment systems through shareholdings and as a representative on the board of directors with some say to day-to-day operational and other issues. This applies to Australia, India (for IMPS and UPI), and Malaysia. For Australia, the RBA was involved in governance arrangements initially when the instant payment system was first established and no longer held shares or appointed directors on the board following a merger. In the third case, the central bank itself owns the instant payment service provided to users and has full operational control over all aspects of this service. This applies to China, Indonesia, Mexico. In sum, the role of central banks in retail payment systems has in some countries

[5] See CPMI (2021) for detailed survey results of 31 instant payment systems across CPMI member jurisdictions, which includes examples on central bank operational roles.

expanded to do much more than just provide final settlement. Some have sought to more directly lower payment costs, reduce payment risks, and likely stimulate innovative "add-on" payment services by the private sector, in addition to providing a timelier 24/7 settlement service.

4. Changes in Payment Composition Across Twelve Countries

Figures 1 and 2 show changes in payment composition for our twelve countries. As shown in Section 4, seven countries experienced statistically significant payment substitution of instant payments for cash, for cards, or for checks. This includes China, Indonesia, Mexico, Norway, Russia, Singapore, and the UK. Countries experiencing significant substitution of a different type--cards replacing cash or checks, or checks replacing cash, are also shown. In the figures, direct debits (DD) are separated from checks.

Table 1 presents an overview of what the figures indicate regarding the use of cash and payment substitution. For this analysis, the two debit instruments--checks and direct debits--are combined (ChecksDD).[6] In the table, countries are ranked by their average annual compound growth rate of the value of cash withdrawn from ATMs--our measure of cash use (Column 1).[7]

Eight countries have seen cash use rise over 2005-2020 while only four find that cash use is on average constant or falling between 2005 and 2020.[8] Looking more closely, however, it is evident that cash use over the period often has something of an inverted "U-shape", first rising and, toward the end, falling. This is shown by the negative annual growth rate of cash over the last three years of the sample (Column 2). Only two countries (Russia and Indonesia) show a positive annual average growth rate in cash over this more recent time period.

Looking at the percentage point change in the share of cash in total payments, the (relative) reduction in cash use (Column 3) is more pronounced than its absolute reduction in value (Column 1). Here eight countries show a reduction or no change in the use of cash compared to four (India, Mexico, Malaysia, and Singapore) showing a rise in their cash share. However, with one exception (Mexico), the share of cash in the last three years of the sample are uniformly negative (Column 4), mirroring the last three years result for payment values.

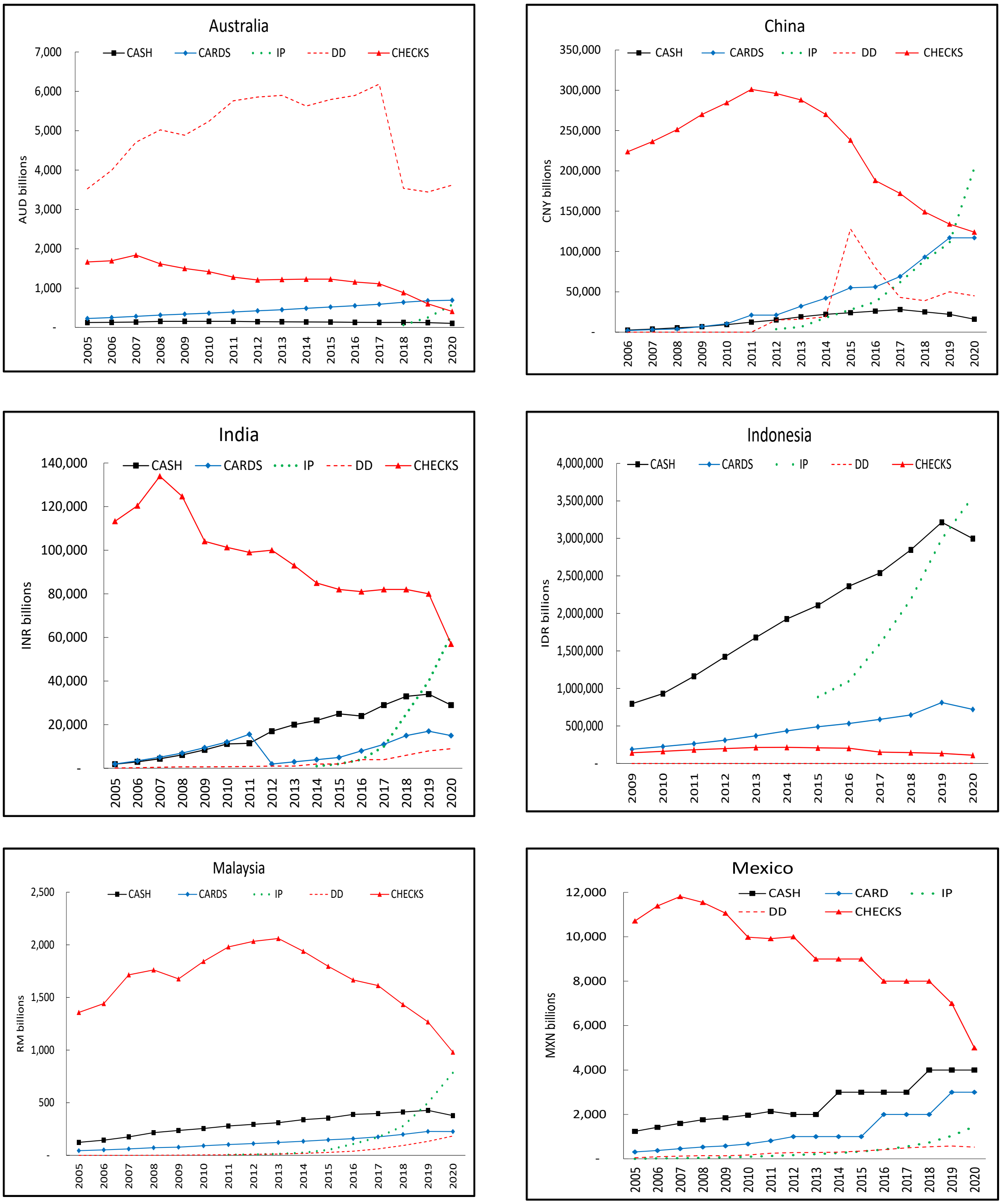
The difference between changes in the absolute value of cash use and the use of cash relative to other payment instruments is because there are two ways the share of cash can fall. The first is that cash falls absolutely (Column 1). The second is that other instruments are displacing cash by growing faster (Column 3). In both cases, the result is that the use of cash is falling.

[6] The only real difference between them is that direct debits are initiated, processed, and collected electronically while paper checks are manually initiated but can be processed and collected electronically. Both instruments are not instant and are risky to accept for payment. If there are insufficient funds to cover either type of debit, they can be rejected.

[7] Other ways of obtaining cash do exist but are almost never publicly reported. This includes cash-back at the point-of-sale or obtaining cash over-the-counter (OTC) at a banking office. OTC withdrawals, when reported, include consumer and business withdrawals together but only consumer withdrawals are relevant. Business withdrawals are used for making change, not purchases. Including business withdrawals would be inconsistent with other payment instruments--especially cards, a strong competitor to cash, which has no need to make change.

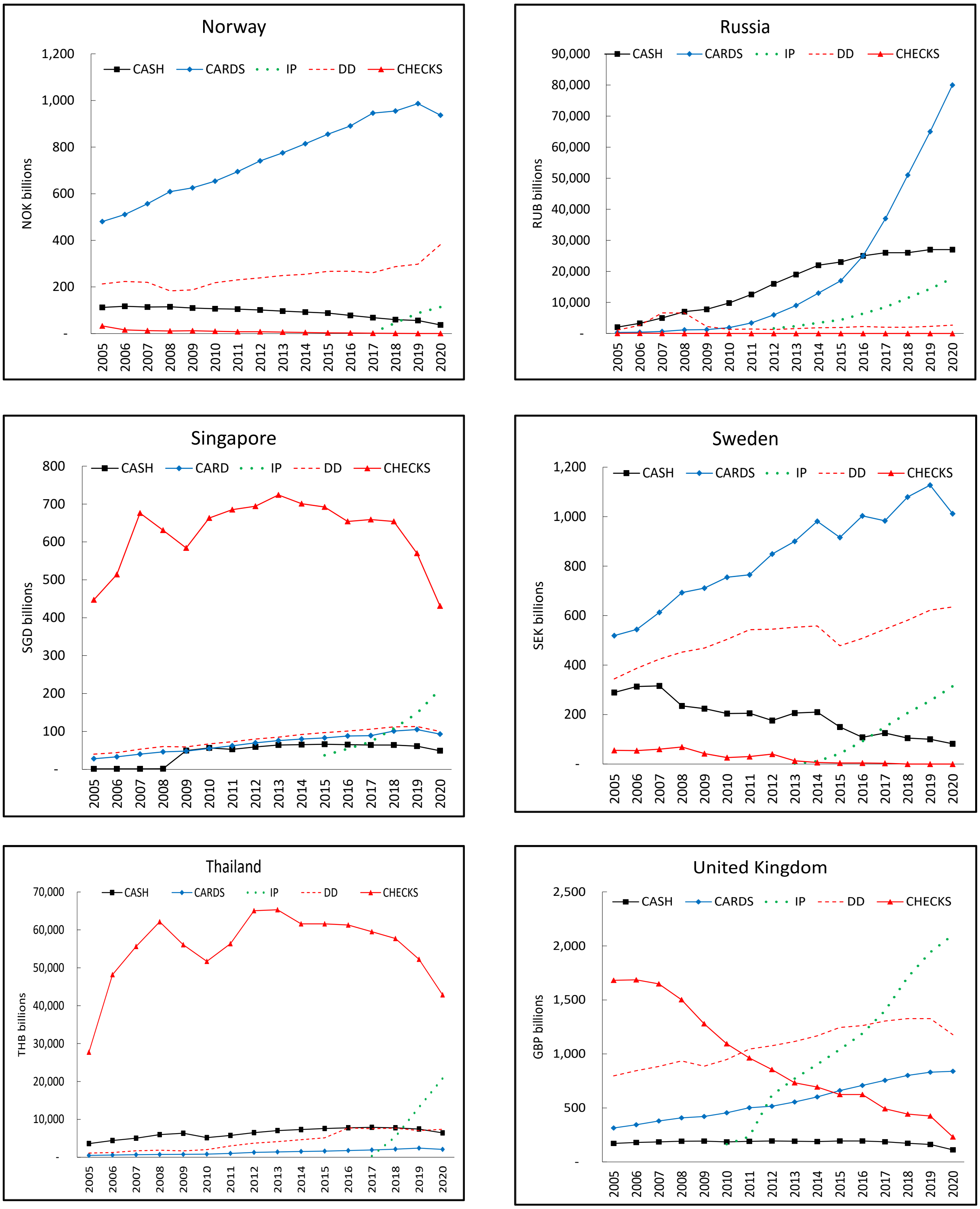
[8] Japan and Korea were excluded from our analysis. Their payment data on lower value retail instant payments (that are often subject to limits on the value of a transaction in other countries) are comingled with large value instant payments (like wire transfers) and thus are not comparable to data for countries we cover. Argentina was also excluded as the data available was insufficient.

Figure 1. Payment Values in Selected Countries



Source: BIS, National Central Banks, UK FPS

Figure 2. Payment Values in Selected Countries (continued)



Source: BIS, National Central Banks, UK FPS

Table 1. Changes in Payment Values and Payment Shares Over Time

	Cash Payment Values % Change 2005-20		All Payment Shares % Point Change 2005-20				
	Annual	Last 3-Yrs	Cash	Last 3-Yrs	Cards	IP	ChecksDD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
India	20%	- 6%	15	- 4	7	36	- 58
Russia	19	2	- 38	- 8	53	14	- 30
China	14	- 20	2	- 3	22	40	- 65
Indonesia	13	3	- 30	- 8	- 7	48	- 11
Mexico	8	0	19	2	19	10	- 48
Malaysia	8	- 4	7	- 2	6	31	- 44
Thailand	4	- 9	- 3	- 2	1	26	- 24
Australia	1	- 10	0	- 1	9	11	- 19
Singapore	0	- 13	5	- 1	5	24	- 34
UK	- 3	- 20	- 3	- 1	8	47	- 52
Norway	- 7	- 22	- 11	- 2	6	8	- 3
Sweden	- 8	- 12	- 20	- 1	7	15	- 2

Table Notes: Data are rounded. ChecksDD represents checks and direct debits together. The time period covered was 2005-2020 for all countries except China (2006-20) and Indonesia (2009-20).

To date, the standard way of assessing cash use has been to compute the ratio of cash in circulation (CIC) to GDP. If this ratio falls, as when CIC is reduced absolutely or when GDP grows faster than CIC, the conclusion is that cash use is falling. Our focus on percentage point changes in cash and other payment shares in the empirical section below uses a similar relative measure (payment shares) to assess the use and substitution of different payment instruments.[9]

The percentage point changes in cash, cards, instant payments, and checks and direct debits (ChecksDD), indicate that the overall change in payment composition has involved a large rise in the share of instant payments (Column 6) and a smaller expansion in the current share of cards (Column 5), versus a large reduction in the share of debit instruments (Column 7) and a smaller reduction in the share of cash (Column 3). This suggests that cards and instant payments have likely substituted for both debit instruments and cash over the period. To pin down which payment instruments may be substituting for others, a more formal analysis of the apparent negative covariation among any two instruments is needed. This is necessary to hold constant other covariations among the instruments. Visual assessments of the degree of payment substitution among payment instruments are suggestive but need confirmation.

[9] Reductions in cash use shown using payment shares (Column 3) are more common than when the ratio of CIC to GDP, or a less-used ratio of the value of small denomination currency and coin to GDP or household consumption, are used (c.f., Khiaonarong and Humphrey, 2022).

5. A Measure of Payment Substitution Across Countries

Earlier efforts to assess economic substitution have used a production function.[10] Given existing technology, a production function relates the mix of observed quantities of inputs of a firm or an economy to the quantity of output being produced under the assumption that output is being maximized. Once estimated, a production function can determine which inputs are substitutes or complements in production and identify statistically significant input substitution. We use a production function model to determine the substitution or complementary relationship among payment instruments in producing and consuming the total output of an economy.[11]

The underlying determinants of payment use and substitution are not specified in our econometric model. Instead, we rely on the revealed preference resulting from these various determinants to generate observed payment use. Although of some interest, we do not need to determine the details of why one payment instrument (Pi) may have been chosen versus another (Pj) to make a payment. Revealed preference of Pi versus Pj over time gives us the same result without the need for payment demand information in a decision-based theoretical model. A change in revealed preference would not be observed unless the change was beneficial to the user and/or receiver in terms of improved convenience, lower cost, greater safety, faster speed, less risk, or some other positive payment attribute compared to other payment instruments.[12]

A production function for an economy can be expressed as the quantity of GDP being produced which is a function of the quantities of labor, physical capital, and intermediate inputs used in that production: GDP = f(L, K, intermediate inputs). Except for labor inputs, expressed as number of workers or labor hours, all these "quantities" are typically in value terms or (in difficult to construct) price weighted quantity indices. Final demand (FD) in a country is merely the sum of the value of the various outputs that are purchased for consumption, investment, etc., after being produced.

Consequently, the gross value of all inputs that comprise value added GDP plus the value of FD is effectively equivalent to the total value of payments (TP) in an economy. In turn, this is composed of the reported value of cash, card, instant payment, and checks and direct debits (ChecksDD), plus wire transfers (i.e., $\sum_i P_i$), and this equals TP. Instead of looking at the actual value of the various inputs and produced outputs in the national accounts and the corresponding output components of final demand, we instead use their direct correspondence with the value of different payment instruments used to purchase them. However, as wire transfers are deemed to be functionally separable from the other four types of payment instruments, they have

[10] See, for example, Berndt and Christensen (1973) or Humphrey and Moroney (1975).

[11] A cost function is dual to a production function and could in principle be used instead. However, data on a nation's cost of different payment instruments are not available. Such information would cover the preparation, transmission, and receipt expense of the various parties to different types of payment transactions. This is needed for total cost (the dependent variable of a cost function), as well as the separate costs of all parties to different types of payment transactions.

[12] Alternative approaches to payment substitution do exist. Snellman, et al. (2001) model the substitution between cash and non-cash instruments for Europe while Koulayev, et al. (2016) estimate a two-stage adoption and use decision model for the U.S. Both require information on the determinants of payment demand (which is not widely available) as well as payment use. We need only the latter.

been deleted from both sides of the relationship TP = k($\sum_i P_i$). Also, the value of these payments differ markedly across countries and including them would hinder comparisons across-countries.

The translog expression for a production function has been commonly used in production analysis for identifying input substitution and is used here to identify payment substitution. It is expressed in logarithmic terms of all the variables and has been used to obtain an elasticity of input substitution among factor and other inputs, among other results.[13] Thus we specify the following (translog) quadratic function:

$$\ln TP = \ln\alpha_o + \sum_i \alpha_i \ln P_i + 1/2\sum_i \sum_j \beta_{ij} \ln P_i \ln P_j + \varepsilon \quad (1)$$

where:

TP = the total value of all Pi payments (noted below) observed annually in a country over time; and

P_i = the value of four payment instruments demanded by users because of the services provided: 1. cash, 2. debit and credit cards (including e-money), 3. instant payments (IP), and 4. Checks and direct debits (ChecksDD) that comprise TP. This excludes wire transfers which are assumed to be functionally separable from the other four instruments.

Using (1), we focus on the degree that P_i may substitute for any Pj over time, assuming that businesses and individuals are maximizing the benefits of their revealed payment choices and thus the mix of payment instruments used over time. To determine the substitute or complement relationships among different payment instruments, we assume coefficient symmetry and take partial derivatives of (1) with respect to each $\ln P_i$:

$$\begin{aligned}
\partial\ln TP/\partial\ln P_1 &= \alpha_1 + \beta_{11} \ln P_1 + \beta_{12} \ln P_2 + \beta_{13} \ln P_3 + \beta_{14} \ln P_4 \\
\partial\ln TP/\partial\ln P_2 &= \alpha_2 + \beta_{21} \ln P_1 + \beta_{22} \ln P_2 + \beta_{23} \ln P_3 + \beta_{24} \ln P_4 \\
\partial\ln TP/\partial\ln P_3 &= \alpha_3 + \beta_{31} \ln P_1 + \beta_{32} \ln P_2 + \beta_{33} \ln P_3 + \beta_{34} \ln P_4 \\
\partial\ln TP/\partial\ln P_4 &= \alpha_4 + \beta_{41} \ln P_1 + \beta_{42} \ln P_2 + \beta_{43} \ln P_3 + \beta_{44} \ln P_4\ .
\end{aligned} \quad (2)$$

Here $\partial\ln TP/\partial\ln P_i$ is the elasticity of total payment output to the change in each ith payment input expressed as a set of four equations, one for each payment instrument. The elasticities are expected to be overall positive, just as marginal products in production are required to be strictly positive. And they are not constants, but depend on the levels of payment services demanded (lnPi).

In the first equation in (2), $\partial\ln TP/\partial\ln P_1$ refers to the use of cash in total payments. The contribution of cash (lnP1) to the level of total payments (lnTP) is expected to be positive ($\beta_{11} > 0$). The total effect on lnTP in the cash equation can be reduced, however, if card use ($\ln P_2$) or instant payments ($\ln P_3$) replaces cash in some transactions (so β_{12} or β_{13} would be < 0). Card use can substitute for cash use at the point-of-sale. More recently, instant payments can substitute for cash in low value P2P and other transactions (as occurs in Indonesia, Russia, UK) or for higher value cards (in Norway, UK), or for checks plus direct debits (in China, Indonesia, Mexico, Singapore, UK).

[13] In situations where the value of a payment may be zero in a country, such as the value of instant payments before they were issued or the disappearance of checks in Sweden after a certain date, a small positive value is inserted in their place to take natural logs of the variables. A possible "work-around" for zero or negative values in the data has been devised by Bos and Kotter (2011) but is not applied here.

A standard translog production function can be written as $\ln q = \ln \alpha_o + \sum_i \alpha_i \ln X_i + 1/2 \sum_i \sum_j \beta_{ij} \ln X_i \ln X_j$ where q is output and (Xi, Xj) are inputs. The marginal product of the ith input is $f_i = \partial q/\partial X_i = (q/X_i)(\partial \ln q/\partial \ln X_i) = (q/X_i)(\alpha_i + \sum_j \beta_{ij} \ln X_j)$. The output elasticity $\partial \ln q/\partial \ln X_i$ thus equals $(f_i X_i)/q$. Profit maximization in a competitive market requires that prices of inputs (pi) reflect their marginal productivity so $p_i = f_i$. Substituting these prices into $(f_i X_i)/q$ yields $(p_i X_i)/q$ which is a ratio of the value share of input i to the value of output q or V_i/V_q which equals the output elasticity.[14]

Assuming that maximizing the benefits of using the mix of observed payments is equivalent to maximizing profits in a standard production function, we replace our payment elasticities $\partial \ln TP/\partial \ln P_i$ in (2) with the observed share of the use of each Pi in TP, denoted as Si. This is shown in (3) which is estimated as a single regression consisting of four equations using a seemingly unrelated regression estimation (SURE) framework and where cross equation coefficient symmetry is imposed using a stacked regression:[15]

$$(3) \qquad \begin{array}{lcl} S_1 & = & 1\ 0\ 0\ 0\ \ \ln P_1\ \ 0\ \ 0\ \ 0\ \ \ln P_2\ \ \ln P_3\ \ \ln P_4\ \ 0\ \ 0\ \ 0 \\ S_2 & = & 0\ 1\ 0\ 0\ \ 0\ \ \ln P_2\ \ 0\ \ 0\ \ \ln P_1\ \ 0\ \ 0\ \ \ln P_3\ \ \ln P_4\ \ 0 \\ S_3 & = & 0\ 0\ 1\ 0\ \ 0\ \ 0\ \ \ln P_3\ \ 0\ \ 0\ \ \ln P_1\ \ 0\ \ \ln P_2\ \ 0\ \ \ln P_4 \\ S_4 & = & 0\ 0\ 0\ 1\ \ 0\ \ 0\ \ 0\ \ \ln P_4\ \ 0\ \ 0\ \ \ln P_1\ \ 0\ \ \ln P_2\ \ \ln P_3 \end{array}$$

The stacked 4 x 14 matrix in (3) is post-multiplied by a column vector of coefficients $(\alpha_1\ \alpha_2\ \alpha_3\ \alpha_4\ \beta_{11}\ \beta_{22}\ \beta_{33}\ \beta_{44}\ \beta_{12}\ \beta_{13}\ \beta_{14}\ \beta_{23}\ \beta_{24}\ \beta_{34})^T$ and also contains four error vectors (one for each equation) in a single stacked column (not shown) giving our full model.[16] Stacked regression is explained in McFadden (1999). There it is noted that if there is strong collinearity among the RHS variables over time in a separate equation-by-equation OLS estimation of each of the four equations in (2), where payment shares (S_i) replace $\partial \ln TP/\partial \ln P_i$, the stacked regression framework (3) can reduce this multicollinearity.[17]

Imposing coefficient symmetry ($\beta_{ij} = \beta_{ji}$) to obtain (3) means that if substitution of cards for cash occurs in the first equation ($\beta_{12} < 0$), the reverse substitution of cash for cards (β_{21}) in the second equation in (2), is expected to be the same as the original substitution of cards for cash. That is, $\beta_{21} = \beta_{12} < 0$ in (3).[18] We compute the

[14] In practice, at the national or industry level, both Xi and q are measured in value terms and thus represent a value share. Neither exist as a pure quantity measure.

[15] Exogeneous shocks to the share of cash (S_1) in the first equation in (3), due to changes in economic growth or differential output price inflation over time, for example, are likely to be correlated with similar shocks to the card or instant payment shares (S_2 or S_3), so the error terms in these equations may be correlated. This is additional explanatory information that may improve the efficiency of the parameter estimates when estimated within a SURE framework.

[16] When the translog production function was first developed, it was common to estimate input substitution using a system of equations similar to those shown in (3). While linear homogeneity could also be imposed here, reducing the number of estimated parameters, this leads to estimation difficulties when payment shares (S_i) are exceptionally small or asymptotically zero. This occurs in some countries for checks, in others for cards or direct debits (DD). It occurs in all countries for instant payments as they were zero in 2005--the start of our sample--before they were implemented. According to available data, most instant payment arrangements were implemented after 2011. Degrees of freedom are df = (4 x n) - 14, where n = the sum of observations across all four equations together in (3) for 2012-2020 or 2005-2020.

[17] All estimations use the LSQ command in the TSP econometric program (c.f., McFadden, 1999, p. 11).

[18] Substitution symmetry between two inputs satisfies Young's Theorem on integrable functions: $\partial^2 f(P_i, P_j)/\partial P_i \partial P_j = \partial^2 f(P_i, P_j)/\partial P_j \partial P_i$. It is commonly applied in translog production or cost functions with more than two inputs and virtually all following analyses using the translog form. It is a testable restriction, but this rarely occurs.

marginal rate of payment substitution between two payment types at a time holding all other types constant. This is shown in the sign and value of the estimated coefficients (β_{ij}).[19]

To deal with possible contemporaneous correlation, (3) is re-estimated with all RHS variables lagged by one year. Technically, this addresses the potential problem.[20] Although we are not interested in forecasting the possible long-run equilibrium composition of payment instruments across countries, we conducted unit root and cointegration tests which indicated non-stationarity of our variables: the variable means and variances are non-constant and vary with time.[21] For illustration, we re-estimated (3) with first differences of all the variables.

If this gives stationarity, it likely also lowers our chance of observing payment substitution. An upward trend in one payment instrument and a downward trend in another instrument would be "removed" and the difference in the means of both instruments would be rather constant over time. Identifying payment substitution would then rely on the covariation of the time pattern of observations above and below a roughly constant mean. As seen in the earlier figures, the composition of the payment systems in our countries are not anywhere near being in equilibrium. Equilibrium would be when payment levels and payment shares are very stable, giving means that do not vary much with time.[22] In any case, concerns about estimating a spurious regression if the data are not differenced should not arise. Our measurement of total payment value (TP) is directly related to the individual values of the different payment instruments (P_i). This is not a hypothesis: TP was defined as $\sum_i P_i$, just as the total value of produced output in a country is the sum of the value of all inputs used to produce it (including a return on invested capital).

Finally, we estimate the ad hoc linear equation system in (4). This removes the four column vectors containing P_1 to P_4 in (3) and replaces the four share (S_1) dependent variables with their corresponding payment values (Pi). No variables are logged. This is a much simpler equation system compared to the quadratic translog specification in (3):

$$\begin{array}{lcccccccccc}
P_1 = & 1 & 0 & 0 & 0 & P_2 & P_3 & P_4 & 0 & 0 & 0 \\
P_2 = & 0 & 1 & 0 & 0 & P_1 & 0 & 0 & P_3 & P_4 & 0 \\
P_3 = & 0 & 0 & 1 & 0 & 0 & P_1 & 0 & P_2 & 0 & P_4 \\
P_4 = & 0 & 0 & 0 & 1 & 0 & 0 & P_1 & 0 & P_2 & P_3.
\end{array} \tag{4}$$

The system of equations in (4) is post-multiplied by a column vector of parameters to be estimated $(\alpha_1\ \alpha_2\ \alpha_3\ \alpha_4\ \beta_{12}\ \beta_{13}\ \beta_{14}\ \beta_{23}\ \beta_{24}\ \beta_{34})^T$. This tells us how the level of the value of each P_i payment instrument varies with

[19] As noted in Miller, et al. (2019), coefficient symmetry applies when computing Allen but not Morishima elasticities of input substitution (c.f., Blackorby and Russell, 1989). Neither of these elasticities are computed here and, more importantly, we do not empirically see cards, instant payments, or checks replacing cash and then see the reverse occurring at a later time. Empirically, it is always one way.

[20] Specifying instrumental variables that are correlated with the dependent variables S_i in (3) but are uncorrelated with the error terms, is not really feasible. Such non-spurious instruments would most reasonably come from detailed survey data showing a change in user preference for one type of payment instrument over another due to their different attributes. Not just once, but for multiple time periods. The lack of such information is the reason why we specify payment values in terms of revealed preference and do not attempt to determine the various incentives that led to the revealed preference itself.

[21] Dickey-Fuller and Weighted Symmetric unit root tests were used. If we had wanted to forecast the likely short-term future composition of payments, we would do so by specifying a logistic curve for payment instruments that are currently rising and an inverse logistic curve for those that are falling, rather than specifying (for example) an error correction model.

[22] It is true that “constant” means of a time-series are not required for stationarity, but a test for cointegration among the variables was also not stationary.

all the other P_j payment instruments ($i \neq j$), rather than the share (S_i). It is estimated in a stacked SURE regression with symmetry imposed. As always, coefficient values $\beta_{ij} < 0$ indicate payment substitution while positive values indicate complementarity among payment instruments.

6. Substitution Results

Six types of payment substitutions are possible for our four classes of payments. With twelve countries, there could be seventy-two (symmetric) substitutions. Table 2 shows eighteen statistically significant payment substitutions for the full model. If an estimated parameter is negative and significant for at least one of the three models presented in the table, we report the sign of the other estimated parameters even though they are not significant.

For example, instant payments in Indonesia, Russia, and the UK all significantly replace cash using our preferred translog quadratic model (Column 1). This occurs only for Sweden when all RHS variables are lagged by 1-year to possibly address contemporaneous correlation issues (Column 2). First differencing of all variables, since unit root and cointegration tests were rejected, yields significant substitution of instant payments for cash in China, Indonesia, Russia, and UK (Column 3). The other cells for this type of substation have a + or – sign indicting the direction of the estimated relationship, but the parameter was not significant.

Using our quadratic payment model, instant payments (IP) have significantly replaced cards in Norway and the UK and have replaced debit instruments in China Indonesia, Mexico, Singapore, and UK. Similarly, cards have replaced cash in Russia and debit instruments in India, Mexico, Norway, Russia, Sweden, and the UK.[23] In this process, cards and instant payments overlap in their replacement of cash in only one country (Russia) but overlap in two (Mexico, and UK) when replacing checks or direct debits.[24]

Cards replacing checks, especially at the point-of-sale, is well known and has been on-going for some time. This type of substitution will likely continue for India, and Mexico as use of debit instruments, while falling, are still very much in use. For Norway, Russia, Sweden, and the UK, however, the situation is reversed. Check use in these countries is now close to or at zero so this type of substitution is effectively in equilibrium by being (almost) no longer used.

One surprise is that instant payments in Sweden (which includes Swish) do not appear to have significantly replaced cash over 2012-2020. This is despite anecdotal news stories about how instant payments are used in place of cash for settling up a restaurant bill among individuals, making church donations, or replacing cash in other P2P transactions. Looking at Figure 1 for Sweden over a longer period of 2005-2020, cash withdrawn from ATMs has been slowly falling since 2006, some seven years before instant payments were implemented

[23] Our sample for each country is small, so it is possible that an outlier in the data could change our results. For one country in our sample, data for 2020 was viewed by authorities as being an outlier. However, the outlier was small and, as our R2 was > .95, there was no meaningful change the reported results after deleting 2020.

[24] For completeness, Table 3 in the Appendix shows the statistically significant substitution results if equation (3) had been applied to the entire 16-year period shown in the figures. Here instant payments would have significantly substituted for cash, cards, or ChecksDD eight times using our quadratic model (rather than ten) using the longer time period. Parameter estimates are mean values and in the early years of 2005-2020, instant payments were only just being introduced (UK) or were zero. Hence our focus on 2012-2020.

(according to our data) in 2013. It appears that the value of cash replaced by instant payments has not been large enough to become statistically significant at this time.[25]

Table 2. Statistically Significant Payment Substitution Over 2012-2020

Type of Substitution	Country	Full Model Eq. 3 (1)	RHS Variables Lagged 1-yr (2)	1st Differencing All Variables (3)
IP replaces Cash	China	-	-	-.071º
	Indonesia	-.010**	+	-.012**
	Russia	-.040**	-	-.063*
	Sweden	-	-.006*	+
	UK	-.041**	-	-.041**
IP replaces Cards	China	-	+	-.069*
	Indonesia	-	+	-.003*
	Norway	-.011**	-.032**	-.547º
	UK	-.062**	+	-.500*
IP replaces ChecksDD	Australia	-	-.066**	-.034º
	China	-.058**	-	-
	Indonesia	-.006º	-	-
	Mexico	-.038**	-.225**	-.074**
	Singapore	-.013º	-	-
	Thailand	-	-.011*	+
	UK	-.120**	-	-.109**
Cards replace Cash	Mexico	-	-	-.023**
	Norway	-	-	-.113*
	Russia	-.170**	-	-.191**
	Sweden	+	+	-.042**
Cards replace ChecksDD	India	-.168º	-	+
	Mexico	-.063**	+	-.071**
	Norway	- .205**	+	-
	Russia	-.034º	-	-.033º
	Singapore	-	-.365*	-
	Sweden	-.301**	-	-.190**
	UK	-.127**	-	-.101**
ChecksDD replace Cash	Mexico	-.101**	+	-.127**
Significance levels are for p-values .01 (**), .05 (*), and .10 (º).				

[25] However, in Column 2 of Table 2, it seems that when all variables in equation (3) are lagged by one year, cash is significantly replaced by instant payments in Sweden, although the estimated effect on payment shares (the estimated parameter) is quite small.

If one believes that contemporaneous correlation is a problem with our preferred translog model in Column 1, then Column 2 in the table indicates that only six possible substitutions are statistically significant. Alternatively, since unit root and cointegration tests were all rejected, perhaps the translog model should be estimated using differenced data. With first differenced data, twenty substitutions (Column 3) are significant, 13 of which duplicate the sign and statistical significance found with using undifferenced data in the full model.

Finally, we estimated the simple linear model specified in equation (4). Only five payment substitutions were significant (not shown) and none involved instant payments. It appears that the quadratic nature of the translog specification captures something the linear specification does not. This would be that payment substitution, as visually implied in the figures (which reflect scatter diagrams of the underlying data), is more quadratic than it is linear. That is, the expectation regarding payment substitution is that a new or an increasingly used payment instrument is rising like a standard logistic curve, while the instrument being replaced is falling like an inverse logistic curve (namely falling at a decreasing rate). The quadratic nature of the data appears to be an essential characteristic of payment substitution needed to locally identify it.

While Table 2 indicates that instant payments have effectively substituted for existing cash, card, and debit transfer payment instruments, what about the potential for instant payments to substitute for a payment instrument that has yet to be issued: namely, retail CBDC? The similarities between instant payments and retail CBDC, especially when operated by and settled through a central bank, would be quite strong. Both would provide instantly available good and final funds backed by a central bank. Both would have a zero-user cost, like Swish (for individuals), other country arrangements, or the recently initiated central bank instant payment service in Brazil (Duarte et al., 2022; Lobo and Brandt, 2022), or have only a de minimis fee (as is envisioned for the US FedNow instant payment service). Both could accommodate P2P payments, as cash does, and could even be adjusted so that electronic wallets (cards or mobile phones) could transfer funds without the need for an external terminal (as CBDC enabled mobile phones can do in China). As well, both could be modified to transfer funds anonymously if within a certain value limit.

If instant payments are structured in a country such that they are a feasible alternative to cash at the retail level, this may well limit the incentive for users to adopt retail CBDC, if issued. To date, the case for issuing retail CBDC has been found to be less compelling in some advanced economies (Australia, Singapore, UK) compared to emerging markets and developing economies (Hayashi and Lei Toh, 2022; Lowe, 2021, MAS, 2021; Menon, 2022; Economic Affairs Committee, 2022). For the latter, the motivations include promoting financial inclusion, enhancing payment system efficiency, competition, security, resiliency, and cross-border payments.[26]

Lastly, the percentage point changes in payment shares shown in Columns 3 and 5 to 7 in Table 1 can be used to indicate the <u>change</u> in each country's level of payment system efficiency in meeting user and receiver payment needs. Payment shares for each payment type in a country would show no change over 2005-2020 if all the presumed benefits from using one payment type over another was already fully realized. Changes in payment shares, as shown in Table 1, indicate that this is not the case. Some are closer to having fully realized their country-specific benefits and corresponding payment efficiencies than others.

For example, in Russia (Row 2) the sum of the positive (negative) changes in the value shares of cash, cards, instant payments and debit instruments (checks and direct debits) in Table 1 are +67 (-67) percentage

[26] See Auer, et al. (2022) for a summary of views from nine central banks on CBDC being a new tool in the financial inclusion toolkit. For other first-mover countries that are not in the scope of this study, earlier CBDC programs were discontinued with opposition from private banks in Ecuador and continued preference for debit cards in spite of e-money alternatives that offered anonymity in Finland (Arauz, et al., 2021; Grym, 2020).

points while for Norway they are +14 (-14) percentage points. Although revealed payment preferences in any two countries need not be the same, if they were then Russia has undergone the greatest relative change in payment composition while Norway has experienced the least. A country ranking of most to least change over the period would be Russia, China, India, UK, Indonesia, Mexico, Malaysia, Singapore, Thailand, Sweden, Australia, and Norway.[27]

The same procedure can be used across countries to identify which payment instrument has experienced the greatest percentage point share reduction in use in the countries sampled. The greatest reduction in payment shares occurs for checks and direct debits (-390) and then cash (-105). These reductions in percentage point shares are almost balanced out by the increase in the rise in payment shares of instant payments (+310) and cards (+143).[28] Thus instant payments have seen the greatest increase in use and debit instruments have seen the greatest reduction.

7. Conclusions

Instant (or fast) payments are a relatively new way of making domestic payments. Many of these arrangements have only been implemented since 2012 but have spread across a range of countries that differ in their level of economic development. Through this process, regulatory authorities and/or central banks have sought to reduce the current delays and expense in making domestic payments. They have at times worked with the private sector to do this or have offered the instant payment service themselves. Either way, this improves the cost efficiency and operating effectiveness of a country's payment system while at the same time importantly reducing payment system risk.

We illustrate the progress of instant payment (IP) systems using the experience of twelve advanced and emerging market economies where instant payments have been adopted and data are available. The change in payment composition across countries over 2005-2020 (16 years) is illustrated visually in figures using domestic currency values of cash, debit and credit cards (including e-money), instant payments, and two debit instruments--checks and direct debits. Cash withdrawn from ATMs experienced positive annual growth in eight of our twelve countries. However, in the last 3-years of the sample, nine countries show a decline in cash use, which seems likely to continue.

Percentage point changes in payment shares indicate that Russia, China, and India had the greatest change in payment composition while Sweden, Australia, and Norway, had the smallest. The largest reduction in payment shares occurs for checks plus direct debits, followed by cash. The largest expansion in shares occurred for instant payments, closely followed by cards.

In the figures presented, the apparent negative covariation observed between one instrument that is rising with another that is falling suggests possible payment substitution. Confirmation of such substitution, however, involves an economic model that holds constant contemporaneous changes in other payment instruments and permits a test of significance of payment substitution. With six possible types of payment substitution across

[27] Changes in predicted payment shares of statistically significant payment substitutions, shown in Table 2, could alternatively be used for the same task.

[28] They do not completely balance since cash, although mostly falling, also had a rise of +48 percentage points in some countries while cards had a reduction of -7. Overall, the total percentage point change in shares was +501/-502 (with some rounding error).

twelve countries, there are seventy-two opportunities for substitution. Our model identifies eighteen that are statistically significant.

The most numerous significant substitution of one instrument for another involves cards replacing checks. However, the most numerous substitution of one instrument or all other instruments occurred with instant payments replacing checks, cash, as well as cards. Although checks are little used today in Australia, Indonesia, Russia, Norway, Sweden, and the UK, they are still extensively used in India, Malaysia, Mexico, Singapore, and Thailand. However, this use is falling, sometimes rapidly.

Overall, check and cash payment shares fell. This was due to direct substitution by instant payments and cards, as well as by more rapid growth by these two instruments compared to checks and cash. Should current trends continue, all check and cash use seem destined to fall to zero or de minimis levels over time. Many countries are already at that level for checks, less so for cash. A more current issue involves retail CBDC. If instant payments provide P2P transactions at zero cost and P2B payments only pose very low costs to business payment receivers, the adoption of CBDC may be importantly compromised, if issued. The similarities between instant payments and CBDC, noted in the text, are too strong to be ignored. As well, instant payments have first mover advantage.

Importantly, the level of adoption of instant payments brought about through its substitution for more expensive payment methods, and their correspondingly slower speed of settlement, indicates the extent that the cost efficiency of a country's payment system has improved and the degree that payment system risk has been reduced. Both are important payment policy objectives and measuring payment substitution provides a way to monitor both objectives.

Appendix

Table 3. Statistically Significant Payment Substitution for Translog Model (2005-2020).

Type of Substitution	Country	Full Model Eq. 3
IP replaces Cash	Russia	-.010**
	Indonesia	-.008**
IP replaces Cards	Norway	-.015**
	Sweden	-.004°
IP replaces ChecksDD	Australia	-.024**
	India	-.013**
	Singapore	-.015**
	Thailand	-.014**
Cards replace Cash	Mexico	-.033*
	Norway	-.063**
	Russia	-.192**
Cards replace ChecksDD	China	-.127**
	India	-.065**
	Mexico	-.073**
	Norway	-.165**
	Russia	-.011*
	Singapore	-.137**
	Sweden	-.182**
	UK	-.335**
ChecksDD replace Cash	Mexico	-.119**
	Russia	-.171**
	Singapore	-.015**

Table Notes: Negative parameter estimates indicate payment substitution. Significance levels are p-values .01 (**), .05 (*), and .10 (°). The estimated parameter reflects the percentage point change in the payment share due to significant substitution. No parameters were significant for Malaysia.

Data Sources

Data used in this study were collected from the Bank for International Settlements and the national central banks in Indonesia, Malaysia, Mexico, Norway, and Thailand. Country notes are as reported in the statistics on payments and financial market infrastructures in the CPMI countries (Red Book statistics) and supplementary information from national central banks and private operators. For further details, see CPMI (2021).

Country	System Name (Year)	Notes
Australia	New Payments Platform (NPP) (2018)	Fast payments refer to transactions processed through the NPP.
China	Internet Banking Payment System (IBPS) (2010)	Fast payments refer to transactions processed through the IBPS fast payment system.
India	Immediate Payment Service (IMPS) (2010) United Payments Interface (UPI) (2016)	Fast payments refer to transactions processed through the IMPS and UPI payment systems. As of December 2019, National Electronic Funds Transfer payments could be considered instant payments as they operate on a 24x7 basis and take place in near real time (within 30 minutes).
Indonesia	ATM and online banking	Fast payments consist of credit transfers through ATMs and online banking (internet and mobile) reported by four domestic switches and global principals (VISA, Mastercard, CUP, JCB). Individual credit transfers processed through BI-RTGS are excluded. Data before 2015 is not available. Fast payment data includes the value of interbank transfers below Rp100 million per transaction sent through ATMs, mobile banking, and internet banking. Funds transfer amounts exceeding Rp100 million is legally recognized as large value payments. Direct debits were introduced in 2017 using the ACH system. Check data before 2009 is not available. E-money includes the total value of chip-based e-money and server-based e-money issued by commercial banks and non-bank fintech firms.
Malaysia	Instant Interbank Fund Transfer (2006) Real-Time Retail Payments Platform (RPP) (2018)	Fast payments refer to instant credit transfers which include DuitNow transactions from the beginning of December 2018. The value of check payments is based on clearance through eSPICK.
Mexico	SPEI (2015)	Fast payments refer to low value payments which are less than MXP 8,000 per transaction from the SPEI payment system.
Norway	Straksbetalinger (2019)	Fast payments refer to instant payments. The value of all card transactions includes the use of Norwegian cards for debit, billing, credit, and e-money functions. The value of ATM cash withdrawals includes the use of Norwegian cards at Norwegian terminals. The value of e-money is included in the use of Norwegian cards. See Norges Bank Retail payment services reports for further details. Available at: https://www.norges-bank.no/en/news-events/news-publications/Reports/Norges-Bank-Papers/2022/memo-22022-payment-services/
Russia	Faster Payments System (2019)	Fast payments refer to 'on-us' payments.

Singapore	Fast and Secure Transfers (FAST) (2014)	ATM cash withdrawals do not include in-house ATM data pre-2009. A large spike in ATM cash withdrawals is therefore evident from 2009.
Sweden	BiR (2012)	Fast payments refer to transactions processed by the BiR payment system.
Thailand	PromptPay (2017)	Fast payments refer to PromptPay services, which use identification numbers such as mobile phone numbers, national identity numbers, corporate tax numbers, e-Wallet numbers, and traditional banking account numbers. The value of card and e-money payments includes the following: debit cards from 2010 and domestic and overseas spending using Thai debit cards via EFTPOS and the Internet. Since 2015, data include domestic spending of foreign debit card via EFTPOS and the Internet. For credit cards from 2010, domestic and oversea spending of Thai credit cards and domestic spending of foreign credit cards via EFTPOS and the Internet. For e-money from 2010, data includes banks and non-banks which operate electronic payment service businesses and excludes top-up cards. For 2005 to 2009, figures were collected from the Bank of Thailand's Annual Payment Systems Report 2010 (page 80, table 7). The value of ATM cash withdrawals is from ATM cards and debit cards. For ATM cash withdrawals through debit cards, transactions included cash withdrawals, purchasing of goods and services via other non-EFTPOS channels, deposit and funds transfer. Publicly available statistics provide a breakdown of debit card transactions into purchasing purposes via EFTPOS and for other purposes (including cash withdrawals). The value of check payments includes in-house and interbank checks. Some financial institutions report in-house clearing in Bangkok and the metropolitan areas only. Check data from Special Financial Institutions were included from 2019. The value of e-money payments include data from banks and non-banks which operate electronic payment service business and excludes top-up cards.
United Kingdom	Faster Payments Service (FPS) (2008)	Fast payments refer to "single immediate payments" conducted through the FPS. In 2012, a new regulation was introduced in the UK for payments to reach the payee's account no later than the day after the payer's account is debited. This resulted in the migration of remaining bill payments and standing orders to payments processed by Faster Payments Service. For the purpose of this dataset, UK fast payments figures include single immediate payments, standing order payments, forward dated payments, and return payments. The value of total cash withdrawals is used as a substitute of value of ATM cash withdrawals, which is not available.

References

Arauz, A., R. Garratt., and D. Ramos (2021), “Dinero Electronico: The Rise and Fall of Ecuador’s Central Bank Digital Currency”, Latin American Journal of Central Banking, 2: 1-10.

Auer, R., H. Banka, N. Yaa Boakye-Adjei, A. Faragallah, J. Frost, H. Natarajan, and J. Prenio (2022), Central Bank Digital Currencies: A New Tool in the Financial Inclusion Toolkit?, World Bank Group and Bank for International Settlement Financial Stability Institute Insights on Policy Implementation, No 41, April.

Bank of Thailand (BOT) and Monetary Authority of Singapore (MAS) (2022), “Thailand and Singapore Launch World’s First Linkage of Real-Time Payment System”, Joint Press Release, April 29.

Bank for International Settlements Innovation Hub (BISIH) (2021), *Nexus: A Blueprint for Instant Cross-Border Payments*, July.

Bech, M., and J. Hancock (2020), "Innovations in Payments", *BIS Quarterly Review*, March: 21-36.

Berndt, E., and L. Christensen (1973), "The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing, 1929-1968", *Journal of Econometrics*, 1 (March): 81-114.

Blackorby, C. and R. Russell (1989), "Will the Real Elasticity of Substitution Please Stand Up? (A Comparison of the Allen/Uzawa and Morishima Elasticities)", *American Economic Review*, 79(4):882–888.

Bos, J. and M. Koetter (2011), "Handling Losses in Translog Profit Models", *Applied Economics*, 43 (3): 307-312.

Committee on Payments and Market Infrastructures (CPMI) (2022), *Extending and Aligning Payment System Operating Hours for Cross-Border Payments*, Bank for International Settlements, May.

Committee on Payments and Market Infrastructures (CPMI) (2021), *Developments in Retail Fast Payments and Implications for RTGS Systems*, Bank for International Settlements, December.

Committee on Payments and Market Infrastructures (CPMI) (2016), *Fast Payments--Enhancing the Speed and Availability of Retail Payments*, Bank for International Settlements, November.

Committee on Payment and Settlement Systems (CPSS) (2006), *General Guidance for National Payment System Development*, Bank for International Settlements, January.

Duarte, A., J. Frost., L. Gambacorta., P. W. Wilkens., and H. S. Shin (2022), “Central Banks, the Monetary System and Public Payment Infrastructures: Lessons from Brazil’s Pix”, BIS Bulletin, No. 52, 23 March.

Economic Affairs Committee (2022), “Central Bank Digital Currencies: A Solution in Search of a Problem?” House of Lords, January 13.

Grym, A. (2020). “Lessons Learned from the World’s First CBDC”, *Bank of Finland Economics Review*, 8: 1-20.

Hartmann, M., L. Hernandez-van Gijsel, M. Plooij, and Q. Vandeweyer (2019), Are Instant Payments Becoming the New Normal? A Comparative Study. European Central Bank, Occasional Paper Series, No. 229, August: 1-53.

Hayashi, F., and Y. Lei Toh (2022), “Assessing the Case for Retail CBDCs: Central Banks’ Considerations”, *Payment Systems Research Briefing*, Federal Reserve Bank of Kansas City, May 26.

Humphrey, D., and J. Moroney (1975), "Substitution Among Capital, Labor, and Natural Resource Products in American Manufacturing", *Journal of Political Economy*, 83 (February): 57-82.

Khiaonarong, T., and D. Humphrey (2022), "Falling Use of Cash and Demand for Retail CBDC", Working Paper, International Monetary Fund WP/22/27, February, 2022.

Koulayev, S., M. Rysman, S. Schuh, and J. Stavins (2016), “Explaining Adoption and Use of Payment Instruments by U.S. Consumers”, Rand Journal of Economics, 47 (2): 293-325.

Lobo, B., and C. Brandt (2022), “Pix: The Brazilian Fast Payments Scheme”, Journal of Payments Strategy & Systems, 15 (4): 367-375.

Lowe, P. (2021), “Payments: The Future?”, Address to the Australian Payments Network Summit 2021, Reserve Bank of Australia, December 9.

McFadden, D. (1999). "Chapter 5. Systems of Regression Equations", https://eml.berkeley.edu/~mcfadden/e240b_f01/ch5.pdf.

Menon, R (2022), "Yes to Digital Asset Innovation, No to Cryptocurrency Speculation" Opening Address by Mr. Ravi Menon, Managing Director, Monetary Authority of Singapore, at Green Shoots Seminar on August 29, 2022.

Miller, N., J. Bergtold, and A. Featherstone (2019), "Economic Elasticities of Input Substitution Using Data Envelopment Analysis", *PLoS ONE* 14(8): e0220478. https://doi.org/10.1371/journal. pone.0220478

Monetary Authority of Singapore (MAS) (2021), “A Retail Central Bank Digital Currency: Economic Considerations in the Singapore Context”, Economic Policy Group, November.

Snellman, J., J. Vesala, and D. Humphrey (2001), “Substitution of Non-Cash Payment Instruments for Cash in Europe”, Journal of Financial Service Research, 19 (April): 131-145.

PUBLICATIONS

Instant Payments: Regulatory Innovation and Payment Substitution Across Countries
Working Paper No. WP/22/228