Fed Now and Faster Payments in the US

Talk for FRB Boston Research Conference

Charles M. Kahn version 11/13/2024

1 Thanks for inviting me to talk. The paper is about FedNow, the Federal Reserve's new instant payment system.

2 It's a little worrying presenting a paper on payments at a general academic conference. It's very worrying doing so at the Fed. On the academic side are people saying "what are you talking about?" On the Fed side are people saying "why are you telling us all this old stuff?"

My solution is to start generally and work down to details. 3 I'll begin with a bit of an introduction to payments, and then get back to the paper itself. So Fed People, at the beginning feel free to get out your cell phones. You can jump back in later.

Really, it is a lot easier to talk to an economics audience about payments now than it used to be, when the whole subject was dismissed as plumbing, unworthy of serious study. We know payments are important for an economy for a number of reasons. 4 First they are big, even by the standards of macroeconomics: non-cash retail payments in the US are something like six times GDP—as you would expect, since it takes multiple steps to get from raw material to final good. 5 But wholesale payments between financial institutions are much, much larger: 40 times GDP for Fedwire, the Federal Reserve's backbone system for large value payments, or $4 trillion per day, and almost $2 trillion a day for CHIPS, the privately provided counterpart.

Anything that big is automatically a concern for financial stability. Incidents in the late 20th century showed vulnerabilities in large value payments systems and led to reforms internationally to protect them from breakdown, and to reduce the enormous liquidity requirements it takes to run them.

6 But for the general public, awareness of payment system has arisen because of the big changes happening in retail payments—often related to e-commerce. 6a The US has been payment card centric country for many years—but some argue that they are 6b perhaps saturating their markets. What has begun to arrive and grow quickly is mobile payment. 7 New platforms: 7a Venmo, Zelle, 7b Cash App—and if you don't know about all of them, I guarantee your children do—7c Paypal, wallets from Google and Apple. 8 In other countries changes have been even more dramatic. 8a The poster child is the mobile payment system M-Pesa which revolutionized life for many in Africa. 8b Alipay and WeChat, originating in giant webstores in China, have become primary avenues for retail payments there. 8c Extraordinary recent projects in India and 8d Brazil are also having transformative effects.

9 Innovations have raised concerns about the industrial organization of the payments industry. With all the network externalities in payments, what are the effects on competition? Will interactions between banks and new payments platforms threaten to safety of payment for customers and safety of financial institutions? Will the newcomers fragment the payment system increasing transactions costs and demands for liquidity as the systems become siloed? And for other countries, if not necessarily for the US, what happens if the Central Bank gets sidelined, especially by foreign payments platforms?

10 The threat of cryptocurrencies and other private digital currencies has led many countries to experiment with Central Bank Digital Currencies-CBDCs. But so far they've proved difficult to implement and adoption has been disappointing. So attention has recently transferred internationally to faster

payments systems based on bank accounts. These pose less immediate threats of disintermediation or financial institution instability.

11 In the area of fast payments, the US is behind. In many other countries, people as a matter of course expect that they can make person to person or person to business payments between bank accounts in a couple of seconds. 11a In the UK the Faster Payments Service has been around since 2008 and by some estimates it accounts for 10% of retail payments. In Netherlands a "pay by bank" service handles a majority of e-commerce transactions. 11b In Brazil the Pix service allows instant account-to-account payments, and in the four years since its launch has become the dominant retail service in the country.

12 In the US, since 2009 you can move money instantly from your Venmo account to your friend's Venmo account. But before instant payments services were available it took time to move money between Venmo and your bank account—one to three business days. 12a RTP, the first US faster payment service for bank accounts, was introduced by The Clearing House (the private service which also runs CHIPS) in 2017. 12b FedNow, introduced a year and a half ago, is its competitor.

That's the landscape.

13 So the paper starts with some theory, providing a classification of payments systems, looking at the fundamental motives for making a payment, and relating these to the demand for faster payments. I use this structure to distinguish between Pix, UK faster payments, and FedNow and provide a little political economy as to why FedNow (and RTP) haven't been the widely touted runaway success that Pix has been. Finally the paper examines the actual use cases that FedNow does and doesn't address, how FedNow might interact with new payments services, and what is needed to spur development of faster payments in the US.

14 A couple of definitions first. A faster payment is one where

1) Transaction is instantly observable to payer and payee
2) Funds immediately available to payee as soon as the payer initiates the transaction
3) Transaction is irrevocable – neither payer nor payer's institution can reverse it

15 Throughout we are talking about account-based payments systems--ones in which payments are effected by crediting and debiting accounts on some other institution's books. This is in contrast to what is sometimes called token based, or store of value based payment epitomized by physical cash or payment by a bearer instrument.

16 The simplest case of an account-based payment is one where the accounts of both payer (Alice) and payee (Bob) are on the books of a single institution. Think of an "on us transaction" like a check drawn on the bank where it is deposited, or a Venmo payment.

17 By the way, a payment from one financial institution to another works like on-us transactions on the central bank. The most important of those reforms in large value payment systems I mentioned before was the institution of Real Time Gross Settlement, which can be understood as requiring that money was actually available in the payer's central bank account before funds could be credited to the payee.

18 Back to Alice and Bob. Things get more complicated if the accounts are with different institutions. In that case, one way that payment could occur is through the maintenance of reciprocal balances. The payment from Alice to Bob is effected by messages between Bank A and Bank B: Bank A's balance at bank B goes down and B's balance at A goes up.

19 But that arrangement gets messy if there are lots of banks: you have to tie up liquidity and credit all over the place, and that could lead to problems of trust and financial instability if the interbank holdings get too high.

So more frequently what happens is a hierarchical or "tiering" arrangement, with a "settlement" on the central bank's books.

19a When the settlement occurs, what happens on the books of bank A and bank B is an adjustment of their reserve balances, in line with an adjustment of customer account balances. 20 We won't need to keep track of that internal accounting; we can simply focus on the banks' interactions with each other and with their customers.

The rules and processes by which interactions among the institutions occur are often referred to as the payments rails.

21 There are two aspects to those rails – on the one side there are the messages sent and the protocols for validating the messages. In faster payments systems the steps are generally as follows:

1. Alice asks Bank A to Pay $5 to Bob at bank B
2. Bank A verifies Alice's identity and adequacy of Alice's funds
3. Bank A notifies Bank B of the payment
4. Bank B verifies Bob has an account that can legally receive payments
5. Bank B sends confirmation to Bank A
6. Both banks send confirmation to their respective customers.

22 But on top of these messages, the rules of the rail must address the timing of funds transfers to and from customers *and* between banks. There are three key steps:

1. 22a Encumbering Alice's account—also referred to as "funding" the payment. At what point does bank A place restrictions on Alice's use of the funds in her account? Relatedly, at what point does the payment become *irrevocable*? That is, when can Alice no longer halt the payment because she changed her mind, or accidently provided the wrong information? And at what point can the bank no longer recall the payment because it was not in fact made by her, but by a fraudster?

2. 22b Releasing the funds to Bob--How much time lag is permitted between Bank B's receipt of the notification and its release of funds to Bob's account, and under what conditions? Is release *final?* Or are there conditions under which the money can be clawed back?

   Now as far as customers are concerned, those are the crucial issues: when does the payer lose control over the funds, and when does the payee receive control? 22c But settlement on the central bank's books is distinct from encumbering and release of funds. What is the relation between the time of settlement and the time of appearance of funds in Bob's account?

23 Distinct payment rails exist mainly because of the different costs attached to each of these protocols or transfers in differing use cases.
We can also distinguish between rails based on which of three levels they operate: 23a do they deal with end users, 23b do they deal with settlement, or 23c just with clearing (which is usually defined to include all interbank activities except settlement itself).

24 In the US, payments between bank accounts at different banks can travel a variety of rails, including rails provided by payments card companies, and several rails provided by the Federal Reserve—notably Fedwire, FedNow, and the batch service FedACH (generally used for repeat payments like utility billing or company payrolls).

So that's the structure for payment rails

25 Next issue: Why do we even want faster payment? 26 Schuh and Stavins estimated that for typical payments consumers care more about cost than speed. In general speed doesn't matter in situations with enduring relationships; the agents can simply adjust the timing of the initiation. Now there are circumstances where speed is important, but to examine them more carefully we need to consider two basic and distinct motivations for making payments: 27

1. Is the payment the extinction of an obligation (bill payment, debt repayment)
2. Or is it the monetary leg of a spot transaction. (DvP)

28 In the first case speed is valuable if the two parties have different effective discount rates. That's kind of a bloodless way of putting it: It includes liquidity constrained situations, occasional emergencies, situations in which there is no alternative funding. It is particularly likely to happen to unbanked individuals with irregular employment. Note that many of these examples while socially important are unlikely on their own to stimulate private investment in faster payments.

29 In the second case, DvP, the source of need for faster payment is the lack of trust—casual purchases where the vendor doesn't know the buyer, major one-time purchases such as real estate, or internet purchases where it is difficult to verify the buyer's identity. In these cases, you can't hand the good over without assurance that the money will come.

30 In these two cases the significance of "faster" payments is different. For obligations, faster means faster release of the money to the payee, while for DvP, "faster" means receiving faster word that the money has been immobilized--if the vendor has assurance it is coming, the speed of its arrival is a secondary concern.

31 Notice that in both cases, these concerns are separate from the question of when settlement actually occurs. Settlement is about financial stability, specifically the risk bank B faces between the receipt of funds and the release to his customer. Real time gross settlement systems require receipt before release. 32 If the money is going to another bank for example, we want to be sure that the order of receipt and release is carefully maintained, because these amounts are large relative to the value of the banks participating. In a retail system these considerations are much less urgent. Knowing the money is coming is generally good enough, so the system can afford to let the bank take the credit risk: the financial stability considerations are much much less.

33 Now we have the components needed to distinguish between FedNow and the systems in Brazil and UK. 34 In the UK, release of funds happens in real time, but settlement occurs only once in a while—three times a day, in fact, in a batch process. 35 In the US, FedNow handles both clearing between the banks and settlement on the Federal Reserves books instantaneously. 36 Finally Brazil does it all, including providing a simple and standardized feel for end users.

And the results of this last difference are dramatic: 37 This graph shows the value of payments on Pix starting in the first full year after its launch, compared with the values on the US RTP system. FedNow

launched in 2023 and hasn't released any figures, but whatever they are wouldn't change the picture noticeably. 38 Looking at number of transactions rather than value tells a similar story.

39 Why is the structure of FedNow different? Primarily it's a question of political economy. The Banco Central do Brasil was able to require all large banks to join pix and abide by its rules including pricing—individuals are charged nothing—and timing of release of payments, and to mandate standards that all end user apps must meet. The UK system, with a limited number of players, was able to set standards requiring rapid release of payments even if settlement lagged.

The Federal Reserve must rely on banks to volunteer to join the system. 1000 have done so, accounting for a sizeable majority of the deposits in the US, but a majority of those 1000 banks only signed up for the power to receive payments, not to send them; and there are still some 8000 banks and credit unions not yet members of the system. Thus the service is still a good way from universal. Three features have probably limited initial uptake: 1) banks' concern for their own stability—what would be the consequences of allowing customers to withdraw funds so quickly (although FedNow allows banks to put caps on withdrawal size), 2) liquidity concerns—small banks lack infrastructure needed making funds available round the clock, and 3) fraud risk, specifically the inability to respond adequately in real time. Meanwhile, the Fed is required to cover costs so it has little flexibility in the prices it is allowed to charge the banks for its services. And it cannot mandate prices banks will charge customers who use the service. Any standardization of end user experience is strictly voluntary. The result—slower adoption and much less user awareness.

A key part of the history of FedNow is the influence of community banks and credit unions and the tension between them and large banks. Initially small banks were the strongest supporters of the Fed's project, and large banks the key opponents. Large banks in effect introduced RTP midway through the preliminary development of FedNow. A cynic, like me, not a Federal Reserve employee, might say the introduction of RTP was a strategic response by large banks to the threat posed by FedNow. In turn, in launching FedNow the Fed argued that RTP was likely to be inadequate for the needs of smaller banks and that no private alternative was likely to arrive. 40 And so FedNow largely duplicates the services of RTP, so that the Clearing House and the Federal Reserve constitute a public/private duopoly on each of the three rails: Large value, batch, and instant payments.

41 Well, what are the uses FedNow addresses? 42 Before its implementation, many believed the service would be used initially, just like Pix was, for person to person instant payments. Instead, because of lack of standardization and uptake, it seems that that person to person demand is being met by Venmo and the like. Indeed the biggest use of FedNow according to published statements, is loading or retrieving money in those person-to-person apps.

A second apparently important use case thus far has been fast payments by businesses to individuals—refunds on purchases and payments to gig workers. The business signs up to an app which gives the individual the option to receive, for a fee, these amounts through one of the instant payment systems, if their bank allows it. A third probable success is micropayments, the payments of a few cents by a business to a customer to verify that a new customer has a valid account. Faster confirmation makes a difference in getting the new customers to take the necessary steps to sign up for purchases.

In the long term government should provide important demand for instant payments. Speed can matter when onboarding new recipients of government payments, when correcting errors, or when emergencies arise. In fact, the government rollout is going to be gradual and piecemeal. Decisions on government payment procedures are extremely decentralized, and for most government payments one

or two days faster is less important than avoiding the transition problems that will arise from changing platforms.

One potential for faster payments in the US is as a challenge to the dominance of credit cards through so-called "pay-by-bank" arrangements for consumer purchases. Cards are expensive; businesses would like to avoid them. Banks might be willing to accept lower fees in return for obtaining greater control over information about their depositors' spending; this information is an increasingly important asset in the digital age. But the US has a long way to go before "pay-by-bank" become mainstream. At present, we lack easy and standardized ways for shoppers to inform their banks to make the payment. No app has become dominant, and no standard way of providing a business's identity has arisen.

43 Procedures for QR codes, or quick look-up directories of business identities, or improved methods for a business to initiate the instant payment rather relying on the customer to initiate it (so called "pull" vs "push") are still several years away. Success will probably come through new platforms, but innovation on that front is still hampered by inadequate access of payments service providers to their own Federal Reserve accounts—again a conflict between concerns for financial stability and speed of payments.

44 So in summary: Starting with a framework for describing payments rails and the drivers for faster retail payment, the paper examined FedNow, comparing it with services in the UK and Brazil. In the talk I've focused on differences in regulatory powers between countries, but the paper also notes cross country differences in already existing payments arrangements, and in structure of financial systems.

45 I've only referred to the faster payments aspect of FedNow. FedNow is also intended to provide resiliency of payments, putting together a future-proofed system adaptable to extensions as need arises and as demand for older rails decreases in the long run. FedNow is and was intended to be essentially a back-office service dealing only with the bank-to-bank portion of the payment process. Faster payments systems pose fewer threats than CBDC to macro-financial stability and fewer alterations to liquidity needs. But in the case of FedNow, the concerns are even less in the short run because of the limits on demand for the program. Eventually FedNow and its close substitute RTP are likely to become dominant rails for payment, but the speed with which it happens, and any resultant threat to financial stability, will depend on adoption of complementary services, whether by the Fed or by other payments service providers.