

Why Do Debit Card Networks Charge Percentage Fees?

By David A. Price and Zhu Wang

Why do debit card networks base their fees on a percentage of transaction amounts when the marginal cost of executing a transaction does not vary by amount? Research suggests that this type of fee structure, a linear *ad valorem* fee, maximizes profits for card networks by allowing price discrimination. Also, because percentage fees make card usage more economical for lower-value transactions, such a fee structure tends to increase social welfare.

Americans have been using debit cards more in recent years. According to the *Nilson Report*, a newsletter of the payments industry, the share of U.S. consumer transactions based on debit cards grew from 19 percent in 2006 to 31 percent in 2011. Debit card payments also have been growing in terms of total dollar volume: They handled 14 percent of the value of consumer transactions in 2006 and 22 percent in 2011. Most of these increases came at the expense of paper checks and, to a lesser extent, at the expense of cash transactions. (During the same period, the share of payments based on credit cards stayed essentially flat, in terms of both the number of transactions and dollar volume.)

Merchants that accept debit cards pay fees known as merchant discounts, which are composed mainly of interchange fees paid to card issuers (that is, cardholders' banks); interchange fees are set by card networks on behalf of their issuers.¹ Interchange fees generally increased over time to the point where many merchants contended that the fees reached excessive

levels and amounted to an abuse of the market power wielded by card networks and issuers. Merchant groups fought successfully in Congress for the Durbin Amendment, a provision of the Dodd-Frank Act of 2010 requiring the Federal Reserve Board of Governors to ensure that interchange fees on debit transactions are "reasonable and proportional to the cost incurred by the issuer."

In addition to the size of the fees, many merchants and some policymakers have criticized the fee structure, which is based on a linear *ad valorem* model—a fixed base fee plus a percentage of the transaction amount, with the percentage component often accounting for most of the interchange revenue. Signature-based debit card networks, which rely on the Visa and MasterCard credit card infrastructures, charged *ad valorem* interchange fees from the outset. In contrast, PIN-based debit card networks began charging merchants a fixed fee per transaction in the mid-1990s, then moved to *ad valorem* fees, mostly in the 2000s.² The prevalence of this fee structure presents a question for economists: Can *ad*

valorem fees be rationalized in the context of debit card transactions when the marginal cost of executing each transaction does not vary with the amount and when debit card transactions carry little fraud risk and—unlike credit-card transactions—give rise to no credit risk?³

One of the authors of this *Economic Brief* (Wang of the Richmond Fed) and Julian Wright of the National University of Singapore attempt to answer this question in a recent working paper. They consider why debit card networks favor linear *ad valorem* fee arrangements, and whether such fees increase or decrease social welfare.⁴ (They also apply the same analysis to auction and shopping websites that accommodate third-party merchants, such as eBay and Amazon.) They construct a model of a payment platform—similar to a debit card network—that handles transactions between buyers and sellers. In their model, the platform is a monopoly that charges transaction fees to merchants. The platform facilitates trade in various goods with differing costs and valuations that it does not observe (apart from the transaction amount), and the merchants sell in a competitive market.⁵ Wang and Wright then consider three scenarios: one based on an unfettered, profit-maximizing platform operator; a second scenario with a Ramsey regulator (that is, a regulator who seeks to maximize social welfare, subject to the constraint of allowing the network to recover its costs, including fixed costs); and a third scenario with a regulator who can choose only between allowing or forbidding an *ad valorem* fee structure, while the operator retains the discretion to set the fee level.

In the unconstrained profit-maximizing scenario, the researchers find that for a broad class of demand functions, a linear *ad valorem* fee determined by a fixed base rate plus a constant percentage of the sales price is, in fact, profit-maximizing. It represents a form of third-degree price discrimination, one in which the fee for every transaction on the platform meets the classic Ramsey principle that a monopolist's markup varies inversely with the price elasticity of demand. While the platform operator does not observe the costs and valuations of each good traded, a linear *ad valorem* fee enables the opera-

tor to obtain the same profit that it would if it had perfect information and could set an optimal fee for each transaction. In addition to shedding light on payment platforms, this finding also may help explain the prevalence of fixed-percentage fees used in many other areas, such as auction houses, stock exchanges, and real estate brokerages.

In the case of the Ramsey regulator, the researchers determine that for the same broad class of demand functions that rationalizes a platform's use of linear *ad valorem* fees, the Ramsey regulator also will charge a linear *ad valorem* fee, though at a lower level than the profit-maximizing operator. Such a fee structure satisfies the Ramsey principle that to maximize social welfare under the condition of just recovering costs, including fixed costs, the regulator sets the markup of each transaction lower than a monopolist's, but still inversely proportional to the price elasticity of demand.

In the scenario of the regulator who considers whether to ban *ad valorem* fees but does not restrict the level of the fees, Wang and Wright find that the regulator generally will allow *ad valorem* fees. Indeed, they find that within the same broad class of demand functions, a linear *ad valorem* fee structure increases social welfare by making the platform service more economical for lower-value transactions. The advantage of such a fee structure can be illustrated by a simple example in which only two goods with sufficiently different values—such as a pencil and a computer—are sold through the platform. As stated above in the first scenario, a linear *ad valorem* fee essentially allows the profit-maximizing platform operator to set a separate monopoly fee for transactions involving each of the two goods. Presumably, the fee for computers is much higher than the fee for pencils. In contrast, the platform operator who is required to use a flat fee would want to forego transactions involving pencils by setting its flat profit-maximizing fee at the monopoly level for transactions involving computers. (At this fee level, pencils would not be traded through the platform because merchants would set the price of pencils at a level unacceptable to consumers.) Conversely, accommodating transactions involving pencils would require

a much lower fee and thus sacrifice too much of the platform's profit from transactions involving computers. Therefore, allowing the platform to set different fees by way of a linear *ad valorem* schedule would not only increase the platform's profit but also consumer surplus (and thus social welfare) since transactions involving pencils could be accommodated.

These results suggest that to the extent interchange fees are viewed as excessive in card payment systems, the harm to social welfare may rest in the fee level rather than the *ad valorem* fee structure. If this is the case, then public policy considerations generally should focus on the fee level rather than the *ad valorem* fee structure.

David A. Price is senior editor of *Region Focus*, the quarterly economics magazine of the Federal Reserve Bank of Richmond. **Zhu Wang** is a senior economist in the Bank's Research Department.

Endnotes

- ¹ For more on the structure of interchange fees, see Mead, Tim, Renee Haltom, and Margaretta Blackwell, "The Role of Interchange Fees on Debit and Credit Card Transactions in the Payments System," Federal Reserve Bank of Richmond *Economic Brief*, No. 11-05, May 2011.
- ² "Debit Card Interchange Fees and Routing: Final Rule," *Federal Register*, July 20, 2011, vol. 76, no. 139, pp. 43,394–43,405. With regard to issuers subject to the Durbin Amendment—banks with assets of \$10 billion or more—networks responded by setting the interchange fee at the regulatory maximum of 21 cents per transaction plus 0.05 percent of the transaction amount.
- ³ Fraud risk is much less for debit cards than for credit cards. According to industry studies, the average net fraud loss to card issuers is 0.08 percent for credit card transactions, 0.05 percent for signature debit card transactions, and 0.01 percent for PIN debit card transactions. See page 1,577, footnote 8 of Shy, Oz, and Zhu Wang, "Why Do Payment Card Networks Charge Proportional Fees?" *American Economic Review*, June 2011, vol. 101, no. 4, pp. 1,575–1,590. In addition, debit card transactions entail no credit risk because no credit is extended.
- ⁴ Wang, Zhu, and Julian Wright, "Ad-Valorem Platform Fees and Efficient Price Discrimination," Federal Reserve Bank of Richmond Working Paper No. 12-08, November 2012.
- ⁵ This research therefore differs from Shy and Wang (2011), who provide another rationale for the use of proportional fees by platforms. For a specific demand specification and considering a single good, Shy and Wang (2011) show that when a platform and sellers both have market power, the platform earns higher profit by charging a proportional fee rather than a fixed per-transaction fee. In that scenario, proportional fees are shown to increase social welfare, although sellers are worse off. Those findings, however, hinge on the presence of double marginalization (that is, the application of markups by two levels of actors with market power). If sellers are competitive, there is no difference between charging a proportional fee and a per-transaction fee in terms of profit or welfare.

This article may be photocopied or reprinted in its entirety. Please credit the authors, source, and the Federal Reserve Bank of Richmond and include the italicized statement below.

Views expressed in this article are those of the authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.