Working Paper

UNDERSTANDING MARGINAL COSTS IN A TWO-SIDED MARKET: IMPLICATIONS FOR DEBIT CARD INTERCHANGE REGULATION

Bruce Levinson
The Center for Regulatory Effectiveness
1601 Connecticut Avenue, NW
Washington, DC 20009
202.265.2383
www.TheCRE.com

Center for Regulatory Effectiveness

UNDERSTANDING MARGINAL COSTS IN A TWO-SIDED MARKET: IMPLICATIONS FOR DEBIT CARD INTERCHANGE REGULATION

I. Introduction to Marginal Costs in Two-Sided Markets

Two-sides markets occur in industries where firms need to appeal to – and bring together – two distinct groups of customers. Obvious examples of two-sided markets include shopping malls which have to attract both retailers and customers, economics journals that need both authors and readers, and payment cards that seek acceptance by consumers and merchants. Multiple academicians have noted that nightclubs that try to attract a similar number of women and men also constitute a two-sided market.[1]

The defining characteristic of firms operating in two-sided markets is that they have two distinct supply and demand curves, one set for each side of the market. This is not surprising as no one would expect newspaper subscribers and advertisers to share the same elasticity of demand since advertisers value and pay more for access to the newspaper (and its readers) than readers are willing to pay for the paper and access to its advertisers. The existence of two sets of supply and demand curves means that a firm's costs of successfully serving each side of the market are different even if they may superficially appear to be the same. For example, Faratin and Wilkening (2004) explain that "[n]ightclubs do not charge women (who in fact may even be subsidized with free drinks), whereas men can be charged both entrance and usage fee."

The economic literature's treatment of marginal costs in two-sided markets is discussed below with particular emphasis on literature discussing the payment card market. An underlying issue explored in this paper is the how financial reform legislation's use of the term "incremental costs" compares with the use of the term in the literature and with the definition of marginal cost. Also discussed is the significance of the two-sided market phenomena to efficient pricing decisions, the need to consider costs incurred on both sides of the market, and estimates of the marginal cost to banks of processing debit card payments. The paper will conclude with a summary of five key principles that the Federal Reserve should apply when developing their proposed debit card interchange fee regulation.

II. Incremental Costs Are Not Marginal Costs

The first point that should be addressed in the discussion of marginal cost is that such costs are a subset of incremental costs which are a more expansive measure of the resources a firm requires to increase production. Incremental cost, for example, may include some cost elements that are considered as "fixed costs." As Lacker and Weinberg explained in the Federal Reserve Bank of Richmond's *Economic Quarterly*,

---

[1] See, J. Wright, (2004), "One-Sided Logic in Two-Sided Markets," Review of Network Economics, vol. 3, no. 1, pp. 44-64. See also, D. S. Evans, R. Schmalensee (2007), "The Industrial Organization of Markets with Two-Sided Platforms," Competition Policy International Vol. 3, No. 1, pp. 151-179; P. Faratin, T. Wilkening (2006), Interconnection Discrimination: A Two-Sided Markets Perspective, in Proceedings of Fifth Hot Topics in Networks (HotNets-V '06), Irvine, CA, US, November 29-30, 2006.

> *It is important to distinguish between incremental and marginal costs. Marginal cost is the added cost of the last unit of a good produced. Incremental cost is all of the additional costs that arise from extending a particular set of services to a particular set of users. This may include costs that are fixed with regard to the quantity of services provided, such as the costs of connecting a group of users to an existing network.*[2] [Emphasis added]

The debit card interchange section of the Dodd-Frank Wall Street Reform and Consumer Protection Act refers to "the incremental cost incurred by an issuer for the role of the issuer in the authorization, clearance, or settlement of a particular electronic debit transaction...." The use of the term "incremental cost" in the legislation is ambiguous as it does not seem to be used in the classical meaning discussed in the literature nor does it seem to adhere to the traditional concept of marginal cost as it is not referring to an additional or "last" unit of production. As will be discussed in Section IV, the cost of processing a "particular transaction" depends on the particular transaction in question which is why this paper recommends using a probabilistic approach to determining transaction-specific costs.

It should be noted that even when only marginal costs are being considered: 1) there is no basis in economics for thinking that price in a two-sided market should be equal to marginal cost; and 2) in estimating marginal cost, the costs relevant to both sides of the market need to be tallied, as explained in Section IV. It should also be noted that interchange costs are distinct from the "price" a merchant pays for debit card processing since interchange is a subset of the costs to the merchant of accepting a payment card.[3]

III. Price Does Not Equal Marginal Cost in Two-Sided Markets

Developing cost-oriented debit card interchange fee regulations, such as those required by the financial reform legislation, is a challenging exercise since there is a broad consensus among economists that such price controls are deeply misguided. Irrespective of the wisdom of the electronic debit payment interchange restrictions, however, the Federal Reserve is required by law to enact the specified regulations under a tight timetable. This paper provides a brief review of the primary objections to cost-based regulation in a two-sided market because it would be beneficial to policy officials to be aware the objections so that they may try to avoid as many pitfalls as possible in developing the new rules.

The biggest potential problem that policymakers need to avoid is analyzing the marginal costs of each side of the market in isolation. For debit card payments, the expenses associated with servicing both the merchant and consumer, as well as the dynamics between both sides of the market, are fundamental to understanding

---

[2] J. M. Lacker, J. A. Weinberg (1998) "Can the Fed be a Payment System Innovator?," Federal Reserve Bank of Richmond *Economic Quarterly*, Spring 1998, Footnote 4.

[3] For a detailed discussion of non-interchange payment card costs, please see , "A Practical Guide to Reducing Merchant Payment Card Processing Costs" found at http://thecre.com/pdf/CRE%20Transaction%20Processing%20Cost%20Reduction%20Paper.pdf

the marginal cost of debit card payments. The problems associated with examining only a single side of the payment card equation was noted in a 2005 paper by Eric Emch and T. Scott Thompson, the Assistant Chiefs of the Competition Policy Section and the Economic Regulatory Section of the Justice Department's Antitrust Division,

> *One robust finding of this line of research has been that welfare-maximizing and profit maximizing prices on each side of the market depend on cost and demand on both sides of the market. Thus, the conventional wisdom that pricing close to marginal cost is efficient does not hold when each side of a two-sided market is examined in isolation.*[4] [Emphasis added.]

The economic literature widely cautions against the notion concept that price should be equal to marginal cost in a two-sided market. For example, the Congressional Research Service ("CRS") examined the issue of interchange fees regulation and noted, "in the case of interchange fees, economic theory also suggests that cost-based regulation would not be expected to produce the optimal interchange fee."[5]

As CRS explains, citing an influential study by David Evans and Richard Schmalensee and legal research published by Steven Semeraro,

> "*Maximizing output requires issuers and acquirers to set prices in a way that will provide proper incentives for cardholders to use and merchants to accept the payment card. Balancing costs in some fashion would achieve this result only if the elasticity of demand on both sides were equal. Furthermore, setting the fee to zero would maximize output only if on both sides of the two-sided market costs and demand were equal. Because neither is likely to be true, one should not expect either a cost based or zero interchange fee to be optimal.*"[6] [Emphasis added.]

CRS' report on interchange fees builds on long-standing economic analyses of two-sided markets in general and interchange fees ("IFs") in specific as well as analysis of the problems resulting from cost-based regulation. For example, the CRS paper is in accord with a European study of payment card interchange fees by Jean-Charles Rochet and Jean Tirole in which the authors state,

> "*In agreement with Katz (2001), we in particular explain why there is no economic rationale for cost-based regulation of IFs.*"[7]

---

[4] E. Emch and T. S. Thompson (2005), "Market Definition and Market Power in Payment Card Networks," p. 4 found at http://www.ny.frb.org/research/conference/2005/antitrust/emch_thompson.pdf.

[5] W. W. Eubanks, CRS Report to Congress, "Payment Card Interchange Fees: An Economic Assessment," September 3, 2008, p. CRS-6. [Emphasis added.]

[6] Ibid.

[7] Jean-Charles Rochet and Jean Tirole, "An Economic Analysis of the Determination of Interchange Fees in Payment Card Systems," Review of Network Economics, vol. 2, No. 2., June 2003, pp. 69-79.

Rochet and Tirole bluntly explain why the economics of the payment cards markets needs to be understood before attempts are made to impose regulatory solutions on perceived problems:

> "*Misunderstanding the economics of the problem and imposing cost-based regulation could impose substantial distortions in the industry.*"[8]

CRS' opposition to cost-based regulation in two-sided markets was consistent with Rochet-Tirole. The Rochet-Tirole paper explains why cost-based regulation of IFs is inappropriate and goes on to explain some of the implications of applying cost-based regulation in other two-sided markets, thus illustrating the law of unintended consequences.

> "*A cost-based regulation of the IF would be an unfortunate precedent for two-sided markets. The same logic would then imply that advertisers' fees paid to TV networks, newspapers and portals should be regulated on a cost basis so as to stop the subsidization of eyeballs by advertisers...and the social gatherings should be regulated so as to prevent payments to or free entry for attractive participants (e.g., celebrities) while others pay for entry.*"[9]

Similarly, former Federal Trade Commission Chairman Murris, in an article published in Columbia Business Law Review concluded that,

> *Because the participants on each side of the market simultaneously generate costs and benefits for one another, traditional notions of setting prices according to marginal cost and other measures of market efficiency are irrelevant*.[10] [Emphasis added.]

The question then becomes, how should costs (marginal or incremental) be addressed in interchange regulation when such consideration is mandated by law?

IV. How to Estimate Marginal Costs in a Two-Sided Market

The simplest although incomplete approach for addressing two-sided market marginal costs is found in an official US government position paper for an OECD Roundtable on Two-Sided Markets. In the document's discussion of pricing in a two-sided market, the USG explains,

> *Let the marginal cost of a transaction be $c = c_I + c_A$ where c_I is the marginal cost of providing network services to the issuing bank and c_A is the marginal cost of providing network services to the acquiring bank. A basic feature of payment networks is that it*

---

[8] Ibid. [Emphasis added.]

[9] Ibid.

[10] T. J. Murris (2005), "Payment Card Regulation and the (Mis)Application of the Economics of Two-Sided Markets," Columbia Business Law Review, Vol. 2005, No. 3, p. 131.

> *may be efficient for price to be below marginal cost on one side of the market (e.g., $p_I < c_I$) and above marginal cost on the other side of the market ($p_A > c_A$).*[11]

A key concept in this statement is that the marginal cost of processing a transaction includes the marginal costs associated with both sides of the market. In a more complete form, c_I and c_A would be written out to specify their various subcomponents – including a proportionate share of the "lumpy" marginal costs, such as repairing a network connection used to authorize transactions or resolving a specific customer dispute which prevents an additional transaction from being processed. These are examples of lumpy costs that must be incurred in order to process an additional debit transaction since any given transaction, *i.e.*, a "particular" transaction will probabilistically be responsible for a certain share of these marginal costs.

Therefore, $c_I = \sum_{i=1}^{n} c_I 1p, c_I 2p \ldots$ where each n is a cost element associated with providing network services to the issuing bank, including the issuing bank's incremental costs, and p is the probability of each cost element occurring during a given incremental transaction. Thus, for cost elements that occur in each transaction, p=1 while for cost elements that are incurred to complete one in every thousand incremental debit transactions, p=.001.

The probabilistic approach to analyzing incremental costs in the consumer payment instrument industry is used in two papers by Daniel Garcia-Swartz, Robert Hahn and Anne Layne-Farrar which estimate the marginal costs to the various parties (banks, merchants) for transactions using the most popular forms of payment including cash, checks and various types of payment cards. In calculating marginal costs, the authors explain,

> *Just as with paper currency, the magnetic stripe on plastic cards wears out over time and must be replaced. About a tenth of the magnetic stripes on credit and debit cards fail within two years and most card issuing banks replace their plastic cards within this time period. The per-transaction card production cost is calculated by taking the cost per plastic card ($0.75) divided by the number of transactions per card over its two-year life span (around 66).*[12] [Notes omitted]

The total marginal cost c of a network platform processing an additional debit card transaction, considering both sides of the market, would therefore be $c = \sum_{i=1}^{n} c_I 1p, c_I 2p \ldots + \sum_{i=1}^{n} c_A 1p, c_A 2p \ldots$

---

[11] Delegation of the United States to the Competition Committee, Directorate for Financial and Enterprise Affairs, Organisation for Economic Co-operation and Development, "Roundtable on Two-Sided Markets," 04-Jun-2009, p.4.

[12] D. Garcia-Swartz, R. Hahn, A. Layne-Farrar (2006), "The Move Toward a Cashless Society: Calculating the Costs and Benefits," Review of Network Economics Vol.5, Issue 2 – June 2006, p. 205.

Two points to keep in mind when considering total marginal cost are: 1) marginal cost is a subset of incremental cost; and 2) even when both sides of the marginal cost equation are totaled, as CRS explained, there is no reason to expect that a cost based interchange fee would be optimal.

V. Estimates of Debit Card Marginal Costs

The two papers by Garcia-Swartz, Hahn, et. al. estimate the marginal costs (as well as the benefits) to each participant in sample transactions based on case studies for the most popular payment methods.[13] Thus, the papers estimate the marginal costs to merchants, consumers, central banks (for cash transactions) and commercial banks for processing transactions made by cash, verified and unverified checks, credit cards, signature debit cards and pin debit cards. The transactions reflect the average payment sizes for different payment methods and retail encounters, *e.g.*, an average size cash transaction at a grocery store, an average check transaction at an electronics store. An overview of the study and results are found in 2006a while the detailed calculations behind analysis are found in the companion paper, 2006b.

This paper will not detail the studies' analyses and conclusions since the reader can consult the papers for this information. Instead, the marginal costs to commercial banks of different payment methods for a single transaction, a grocery story purchase of $54.12, will be presented along with a conversion of those costs to a percentage of the transaction amount.

The purpose of this illustrative example is provide some quantitative guidance to the Federal Reserve in the development of their proposed rule. It should be noted that the marginal costs, as a percentage of the transaction, for non-case payment methods were higher for the case study of a grocery story purchase of $11.52 since changes in marginal cost for non-cash transactions do not vary in direct proportion to transaction size. As always, it is important to keep in mind that marginal costs are a subset of incremental costs.

Grocery Store Transaction: $54.24

	Cash	**Non-Verified Check**	**Verified Check**	**Credit/ Charge**	**Signature Debit**	**PIN Debit**
Commercial Bank Marginal Cost	$0.07	$0.12	$0.12	$0.34	$0.27	$0.27
% of Transaction	0.6%	0.2%	0.2%	1.2%	0.7%	0.7%

SOURCE: Garcia-Swartz, et. al., "The Move Toward a Cashless Society: Calculating the Costs and Benefits, Table 2-4

[13] D. Garcia-Swartz, R. Hahn, A. Layne-Farrar (2006a) "The Move Toward a Cashless Society: A Closer Look at Payment Instrument Economics," Review of Network Economics, Vol.5, Issue 2 – June 2006, pp. 175-198; D. Garcia-Swartz, R. Hahn, A. Layne-Farrar (2006b) "The Move Toward a Cashless Society: Calculating the Costs and Benefits, Review of Network Economics, Vol.5, Issue 2 – June 2006, pp. 199-228.

VI. Principles Applicable to Cost-Oriented Regulation in Two-Sided Markets

- Price does not equal marginal cost in two-sided markets.
- To estimate marginal cost in a two sided market, the costs associated with serving both sides of the market have to be tallied.
- Marginal costs for debit card processing include a probabilistic share of lumpy costs that may be incurred when processing a particular transaction.
- Marginal costs are a subset of incremental costs.
- Incremental costs include some but not all fixed costs.

About CRE

CRE is a non-partisan regulatory watchdog established in 1996 by former senior career officials from the Office of Management and Budget ("OMB").[14] CRE has identified interchange regulation as requiring watchdog oversight since:

- Payment card regulation affects virtually every facet of the economy; and
- CRE has an extensive record in acting as a watchdog on legislative and regulatory proposals affecting payment systems.[15, 16]

To promote public participation and transparency in the regulation of payment systems, CRE has established the Interchange Fees Forum ("IFF"), http://www.thecre.com/premium/ an operating component of CRE's Federal Financial Forum, http://www.thecre.com/insurance/.

The IFF is an Interactive Public Docket ("IPD") which is an eRulemaking tool that provides "the public with the capability to...publicly post data and other materials pertaining to federal proceedings on a continuous basis, including after the close of the Administrative Procedure Act comment period...." Additional information about the significance of IPDs may be found in a thoughtful commentary written by The Center for Progressive Reform.[17]

---

[14] http://www.thecre.com/emerging/Jim_Tozzi_Bio.html.

[15] http://www.thecre.com/pdf/Banking%20Report.pdf.

[16] http://www.federalreserve.gov/SECRS/2008/September/20080909/R-1298/R-1298_256_1.pdf.

[17] http://www.progressivereform.org/CPRBlog.cfm?idBlog=8CB82125-FDE7-1BB5-04DAB9B697EFEACB

Bibliography

Center for Regulatory Effectiveness (2010) "A Practical Guide to Reducing Merchant Payment Card Processing Costs."

Delegation of the United States to the Competition Committee, Directorate for Financial and Enterprise Affairs, Organisation for Economic Co-operation and Development, "Roundtable on Two-Sided Markets," 04-Jun-2009

Emch, Eric and Thompson, T. Scott (2005), "Market Definition and Market Power in Payment Card Networks," presented September 14, 2005 at the Antitrust Activity in Card-Based Payment Systems: Causes and Consequences conference sponsored by the Federal Reserve Bank of New York and the Review of Network Economics.

Eubanks, Walter W. (2008) CRS Report to Congress, "Payment Card Interchange Fees: An Economic Assessment," September 3, 2008.

Evans, Dale S. and Schmalensee, Richard (2007), "The Industrial Organization of Markets with Two-Sided Platforms," Competition Policy International Vol. 3, No. 1, pp. 151-179.

Faratin, P. and Wilkening, T. (2006), Interconnection Discrimination: A Two-Sided Markets Perspective, in Proceedings of Fifth Hot Topics in Networks (HotNets-V '06), Irvine, CA, US, November 29-30, 2006.

Garcia-Swartz, D., Hahn, R., and Layne-Farrar, A (2006a) "The Move Toward a Cashless Society: A Closer Look at Payment Instrument Economics," Review of Network Economics, Vol.5, Issue 2 – June 2006, pp. 175-198;

Garcia-Swartz, et. al. (2006b) "The Move Toward a Cashless Society: Calculating the Costs and Benefits, Review of Network Economics, Vol.5, Issue 2 – June 2006, pp. 199-228.

Lacker, Jeffrey M. and Weinberg, John A. (1998) "Can the Fed be a Payment System Innovator?," Federal Reserve Bank of Richmond *Economic Quarterly*, Spring 1998, pp. 1-25.

Murris, Timothy J. (2005), "Payment Card Regulation and the (Mis)Application of the Economics of Two-Sided Markets, Columbia Business Law Review, Vol. 2005, No. 3, pp. 515-550.

Public Law 111-203, Dodd-Frank Wall Street Reform and Consumer Protection Act, Section 1075.

Rochet, Jean-Charles and Tirole, Jean (2003) "An Economic Analysis of the Determination of Interchange Fees in Payment Card Systems," Review of Network Economics, vol. 2, No. 2., June 2003, pp. 69-79.

Wright, Julian (2004), "One-Sided Logic in Two-Sided Markets," Review of Network Economics, vol. 3, no. 1, pp. 44-64.