

6

NICHOLAS ECONOMIDES

*Competition Policy Issues in the
Consumer Payments Industry*

At the completion of a sale, money changes hands. Money changing hands could be in cash or checks, and, for the last few decades, also be in electronically transmitted funds or a guarantee of prompt electronic payment to the merchant. Such electronic payments could come from a company that provides credit to customers (such as a bank organized under the Visa or MasterCard trade names¹), from one that facilitates transactions but typically does not provide credit (such as American Express), or directly from the bank where the customer has demand deposits.² The payment system intermediary facilitates the payment to the merchant by guaranteeing that the merchant receives the money, and at the same time can also offer a variety of services to the cardholder, ranging from credit services to frequent flyer miles.

Credit and other bank cards facilitate transactions between merchants and consumers. Card networks collect significant fees from merchants to facilitate those transactions.³ The market for facilitation is dominated by the Visa and MasterCard networks. Visa had a 42 percent share of the U.S. credit card market in 2007, MasterCard 29 percent, American Express 24 percent, and Discover 5 percent.⁴

I thank Bob Litan and participants in the Future of Consumer Payments Conference at the Brookings Institution.

Both Visa and MasterCard charge fees (primarily to merchants) that are significantly above costs—some report that total card costs are only 13 to 15 percent of the fees charged and that total fees are about \$30 to \$48 billion per year.⁵ This combination of fees that are significantly above cost and high market shares suggests that current fees reflect market power.⁶

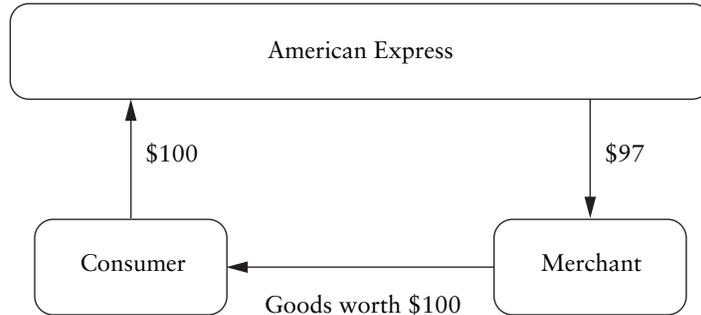
Setups of Three- and Four-Party Card Networks

The intermediation of American Express involves three parties, the cardholder, the merchant, and American Express, hence the name three-party card network. The basic structure of this setup is presented in figure 6-1. It is important that the network (American Express) can charge fees on both sides of the market, or can charge only one side and subsidize the other. This two-sidedness is a fundamental feature of network structure and can be exploited to support high transaction fees.

In a multiparty credit card association, such as Visa or MasterCard, merchants deal directly with acquiring banks that intermediate transactions to issuing banks that issue cards to consumers and ultimately send them bills as well. A transaction between a customer and a merchant conducted through Visa or MasterCard is intermediated by both the acquiring bank and the issuing bank. Figure 6-2 shows the intermediation in a Visa or MasterCard network where the functions of acquiring (a merchant) and issuing (a card to a customer) can be handled by different banks. Thus, in this setup we have four parties: the merchant, the acquiring bank, the issuing bank, and the cardholder.⁷ The two-sidedness remains important in more complex networks such as those of MasterCard and Visa.

In four-party networks, such as MasterCard and Visa, three markets are connected in sequence in each transaction, and the surplus of each end-to-end transaction is divided among the markets (figure 6-3). The three markets are between the issuer and the consumer (market 1), between the acquirer and the issuer (market 2), and between the merchant and the acquirer (market 3).

Figure 6-1. A Three-Party Card Network



The interchange fee is the amount an acquiring bank pays an issuing bank when a merchant accepts a Visa or MasterCard for a purchase, that is, the fee that changes hands in market 2. The acquiring bank pays the merchant the amount of the transaction less both the interchange fee and an additional fee that the acquiring bank keeps for itself. Visa and MasterCard set maximum interchange fees, and almost no banks deviate from them.⁸ Interchange fees in the United States are approximately 1.8 percent on average.⁹ The transaction fees the merchants pay are at

Figure 6-2. A Four-Party Card Network

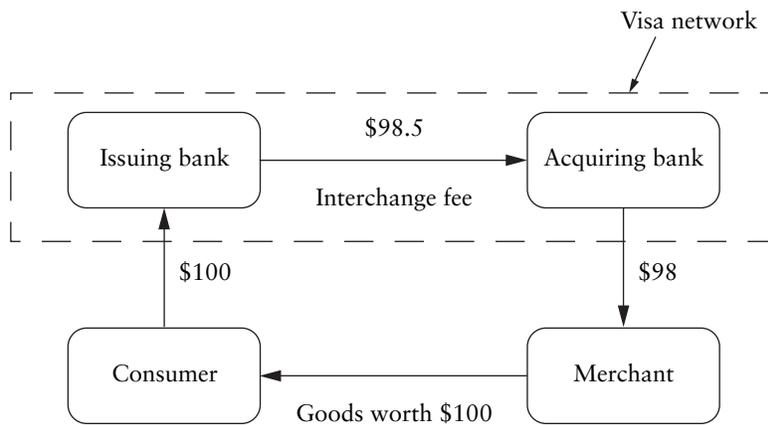
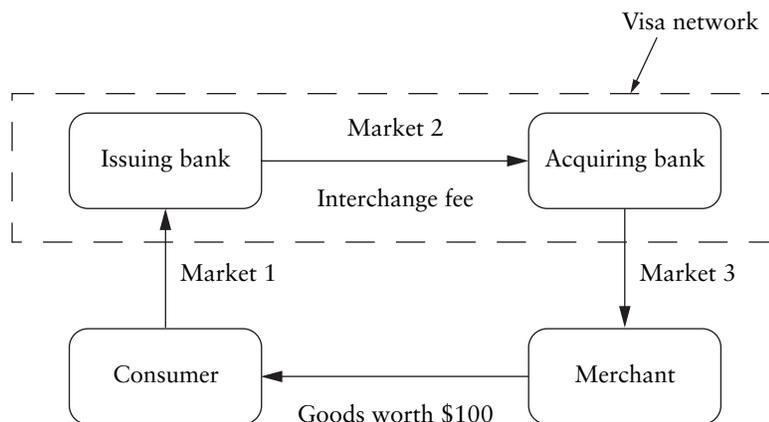


Figure 4-3. *Three Sequential Markets in Four-Party Network*

least as high as the interchange fees. Even if the market between acquirers and merchants were perfectly competitive, acquirers would have to charge merchants at least as much as the interchange fee because the interchange fee is their marginal cost (which they of course need to pay). Most commentators agree that the market between the acquirers and the merchants (market 3) is effectively competitive. Thus, if there is market power in the four-party network, it has to be in markets 1 and 2, although its final effects manifest in market 3 as well.

Both issuing and acquiring banks can charge (or be charged) on both sides of the market they intermediate.¹⁰ That is, decisions affecting pricing on one side of the market will have consequences on the other side. For example, a decision by Visa or MasterCard, the issuing banks themselves, or regulators to reduce the fees that merchants pay may either increase the fees customers pay for the card or may reduce awards or other incentives that issuing banks offer to customers. The extent to which this will occur depends on whether the issuers are passing to their customers the (interchange) fees they receive from the merchants through the acquiring banks. That is, it depends on how competitive the card issuing market is. Of course, consumers will likely benefit when a merchant reduces prices to reflect lower intermediation fees. Because of the

complexity of the market structure and the varying degrees of market power in the three markets identified, policy recommendations need to be carefully examined in terms of their impact on all sides of the markets.

How Card Networks Keep Transaction Facilitation Fees High

The card networks impose various contractual restrictions, such as those against surcharges, steering, and discrimination, as well as, until recently, one to honor all cards, which collectively prohibits or discourages merchants from favoring cards that offer better terms. This both reduces competition among card networks in getting merchants' business and supports high interchange fees. These restrictions are critical.

Credit card networks have high price-to-cost markups despite non-dominant market shares. There is evidence of very significant markups of price above cost, with total costs representing only 15 percent of revenue. It is highly unlikely that consumers receive from card networks anything approaching the fee level charged to merchants. The implied profit rates are comparable to those of Microsoft and Intel, which each have a dominant and almost monopoly market share. So the interesting question is how Visa with a 42 percent market share and MasterCard with a 29 percent share achieve such high markups and market power. An answer will also suggest ways in which distortions can be reduced in the market for transaction facilitation.

If confronted with the cost of their transactions, consumers would most likely use the card with the lowest fee in businesses where multiple cards are accepted. Of course, price is only one of a number of consumer considerations. Everything else equal, however, consumers are more likely to use cards that impose lower direct costs to them. So, if consumers faced directly the costs of intermediation for their transactions, they would choose to use lower-cost cards. Competition among the card networks would therefore drive fees down.

The networks use a multipronged strategy to achieve an equilibrium with less competition. The first part of their strategy is to ensure that a cardholder does not directly face the cost of using a particular card for

payment. This requires two conditions, that the consumer not pay more to the issuer for a card that has higher costs, and that the consumer not pay more to the merchant when using a card that has higher costs. The two-sidedness of the card network can easily ensure the first condition because costs can be recovered from the merchant side. The second condition is more complicated to implement.

Focusing on the first condition, we note that cardholders do not need to face the merchant's cost of their transactions because of the two-sidedness of the network. As long as the network can collect from one side (the merchants), it does not need to collect from the other (the cardholders), and can in fact even subsidize the cardholders. Therefore, unless the merchants impose additional costs on the cardholders when the network imposes such costs on the merchants, the cardholders will not face transaction costs directly and therefore will not in general use the lowest-cost card.

Second, by imposing contractual obligations on merchants, networks make certain that merchants cannot charge different prices (to reflect the different card fees) for the same item to consumers who use different cards.

Card networks have used a number of instruments to make it difficult for merchants to respond to card fee differences. This of course facilitates high fees. The first such instrument is the no-surcharge rule, a contractual restriction imposed on merchants. The second was the honor-all-cards rule, which was abolished in 2003 after an antitrust suit by the merchants.¹¹

The No-Surcharge, No-Discrimination, and Most-Favored-Customer Rules

Essentially the no-surcharge rule says that a merchant can charge the same amount for a Visa transaction as for cash, but if a merchant offers a discount for cash payments, he cannot offer the same discount to a comparable card (MasterCard). Additionally, if a merchant offers a discount to a comparable card, he must offer it to Visa as well.¹² This, in economics, is called a most-favored-customer rule. The effect of the no-

surcharge rule is that the merchant cannot offer better terms to customers who buy with MasterCard than with Visa, although it would make sense to do so if MasterCard's fees to the merchant were lower. This rule allows no price flexibility in the merchant's pricing. It is as if Coca-Cola were to impose the requirement that a can of Pepsi be sold at the same price as a can of Coke. The only option for the merchant who does not like the fees of a particular network is to not accept that network's card. An additional restriction is the no-discrimination rule, which MasterCard phrases this way: "Merchants may not engage in acceptance practices or procedures that discriminate against, or discourage the use of, MasterCard cards in favor of any other card brand."¹³

Industrial organization theory has established that most-favored-customer rules can be used to increase prices to collusive levels.¹⁴ The intuition for this result is simple. Most-favored-customer rules impose on a merchant the requirement to cut prices to all customers with whom it has agreed on this rule if it cuts the price to any one customer. Thus the loss of revenue implied by a price cut to one customer is multiplied in the presence of the most-favored-customer rule. It follows that a firm is less likely to decrease a price under the most-favored-customer rule. This effect is strengthened when a number of firms put these rules into effect.

The Honor-All-Cards Rule

High merchant fees were threatened by technological change. Debit networks, typically with PIN verification, offered lower merchant fees than traditional card networks. Debit cards in the MasterCard and Visa networks also offered much lower fees than signature-based cards. To avoid loss of profits in credit cards, the networks imposed an honor-all-cards rule. This required that if a merchant accepted one Visa card, he had to accept all Visa cards, both credit and debit, issued by any bank in the Visa network.

There were two aspects of this rule. First, if a merchant accepted a certain type of card (say, Visa debit) issued by one bank (say, Citibank), he was required to accept the same type of card (in this case, Visa debit) issued by another bank. The rule also imposed the requirement that a

merchant accept any other Visa products (such as Visa credit cards) if he accepted one (such as a Visa debit card). Visa's rules stated that "the Merchant shall promptly honor all valid Visa cards when properly presented as payment."

The second requirement, that is, to accept different types of cards of the same brand, was essentially tying and has anticompetitive consequences. To put this in context, it would be anticompetitive were Microsoft to say, "If your corporation buys Windows, it must also buy MS Office," or were Dell to say, "If you buy Dell servers you must also buy Dell laptops."

The honor-all-cards rule is now illegal in the United States, merchants having won an antitrust suit against the card networks in 2003. The court essentially forbade the second requirement but confirmed the first—that networks can require merchants to honor all cards across all member banks for a specific type of card (such as a debit card).

Effects of the Present Equilibrium

Transaction facilitation fees charged to merchants, driven primarily by the interchange fee, are significantly above the total cost of facilitating transactions. Because most merchants do not offer discounts for paying in cash, those who primarily do use cash end up paying, through higher product prices, for the costs of card use, which for the most part is by more wealthy consumers. Card transactions are subsidized by cash transactions. Cardholders do not see the fees imposed on merchants, only retail product prices, which increase for all consumers. Additionally, as the networks try to expand by signing up more issuing bank members, they have incentives to increase their interchange fees to make entry into their network more attractive and to avoid exit.¹⁵ As acquirers "typically 'blend' their pricing and charge each merchant one overall merchant service fee based on the projected proportionate volume of cards from each scheme"¹⁶ (network), "in effect, the lower cost scheme therefore subsidizes the higher cost scheme with the merchant receiving

only perhaps some marginal benefit of the lower cost scheme's interchange rates."¹⁷ Thus, at the present equilibrium, high-cost card transactions are subsidized by low-cost transactions. And, under the present rules, interbrand (internetwork) competition does not produce lower fees—quite the opposite.¹⁸

Improving Efficiency

How can efficiency be improved in this sector? The optimal approach is to help the markets work. Recognizing that the credit card setup comprises three sequential two-sided markets, as described, we need to consider how to improve and enhance competition on the merchant side (between merchants and acquiring banks), between issuing and acquiring banks, and on the consumer side (between consumers and issuing banks).

Changes between Merchants and Acquiring Banks (Market 3)

On the merchant side, I propose that card contracts allow for merchant flexibility in acceptance and pricing depending on the card's brand and type as well as on the fees charged to the merchant. That is, a merchant should be allowed to offer different discounts (or surcharges) to consumers for using a particular card if that card offers the merchant lower (or higher) fees. This requires, of course, that the no-surcharge and no-discrimination rules be eliminated from the contracts.¹⁹

The direct consequence of changing these rules will allow the merchants to make customers face the costs of transaction facilitation, which will increase internetwork and intranetwork competition. First, the change in the rules will increase competition between the products of the same network, resulting in lower fees for all of these products. A customer faced with, say, a lower fee when using a Visa debit card rather than a Visa signature card will use Visa debit. Second, the change in the rules will increase competition between the card networks and thus lower fees across the board. A customer faced, say, with a lower fee from MasterCard than Visa will use MasterCard.

Changes between Issuing and Acquiring Banks (Market 2)

The network, not the market, now sets the maximum interchange fee between issuing and acquiring banks and practically no bank in the network deviates from it. The interchange fee is set high, leading acquirers to charge merchants an even higher fee. Card networks have built-in incentives to increase the interchange fee to attract more issuers.

To reduce the interchange fee, I propose that the network no longer set the maximum interchange fee. Let it instead be determined in bilateral negotiations between an issuer and an acquirer, starting from a zero fee basis (par). This would allow for bilateral negotiations between the banks that could result in a variety of interchange fees depending on the specific pair of issuer and acquirer and their competitive conditions. The system could start from a default zero interchange fee, with the market determining any positive or negative adjustment of the fee in a bank pair.

There are two objections to this scenario. The first is that it might lead to too many bilateral contracts. But there is significant concentration among acquirers, with 86 percent of all Visa and MasterCard volume generated by the top ten acquiring banks. Similarly, 84 percent of this volume is generated by the top ten issuers. Therefore ninety contracts generate 72 percent of all MasterCard and Visa volume.²⁰ The second objection is that an issuer can hold out for a high (monopoly) fee to an acquirer. To the extent that this is a unilateral exercise of monopoly power that was acquired legitimately, it should not be an antitrust concern. High fees by a particular issuer who brings high value transactions will hopefully attract competition by other issuers for the same customers and will, in the long run, have these customers signed by a different issuer, resulting in lower fees because of competition among issuers. Additionally, it is not clear that the imposition of high fees is not happening right now with the network setting the monopoly fee for all issuers. With bilateral negotiations, the high fees will be limited to a few issuers instead. Moreover, since the fee will not be set collectively by the network, the incentive to set a high fee across the board to attract more issuers to the network will be eliminated.

Consequences of the Changes in the Rules

Allowing interbrand competition is expected to increase competition between the card networks. It is difficult to estimate the extent of additional competition and the extent of the reduction in fees. The “natural experiment” of Australia might give us some insight.²¹ In 2003, the Reserve Bank of Australia (RBA) reduced interchange fees for credit cards in Australia from an average of 0.95 to 0.55 percent, and in November 2006 to 0.50 percent, and at the same time allowed merchants to impose surcharges.²² Even though surcharging was not widespread, merchant fees fell even more than the interchange fees. The Reserve Bank of Australia made the following observation in its annual report:

The fall in the average merchant service fee since the reforms is significantly larger than the decline in the average interchange fee. . . . These lower merchant costs are feeding through into lower prices for goods and services (or smaller price increases than otherwise would have occurred). While merchants would undoubtedly have hoped that these lower costs translated into increased profits, competition means that just as the banks passed on their lower costs to merchants, so too must merchants pass on their lower costs to consumers.²³

Additionally, the overall cost to the economy of facilitating transactions fell.²⁴ The reforms outlined are likely to cause significantly lower fees for facilitating transactions. The subsidy from cash transactions to credit transactions is likely to be reduced. This will help less affluent customers, who tend to pay in cash. Within credit card transactions, the subsidy from high fee cards to low fee cards will be reduced.

Conclusion

Card network fees are considerably higher than card network costs. This is facilitated by rules imposed on the merchants that do not allow merchants to steer consumers to cards that carry lower fees. The no-

surcharge and no-discrimination rules force merchants to not charge different prices to customers using different cards even though merchants may pay different fees to the card networks. Abolition of these rules would help merchants impose the cost of the payment option they use on consumers. Abolition of these rules will increase competition in payment systems, both across card networks and within each card network.

Notes

1. For many years Visa and MasterCard functioned as not-for-profit associations of member banks. They recently made initial public stock offerings.

2. Although credit and noncredit cards started as single-store cards or single-product cards (such as travel services), they quickly evolved into payment systems used for a large variety of transactions.

3. Consumers pay extra for credit.

4. This is reported in the May 2008 issue of the *Nilson Report*, a trade publication, on credit card networks, and based on credit card purchase volume, excluding both cash volume, such as advances, and debit cards. Similarly, in 2006 the market shares were almost identical: Visa 42 percent, MasterCard 29 percent, American Express 23 percent, and Discover 5 percent. Discover Financial Services was spun off by Morgan Stanley in 2007. In the U.S. debit card market that year, market shares were Visa 48 percent and MasterCard 14 percent. See also Associated Press, "How Visa Operates," *Forbes.com*, February 25, 2008 (www.forbes.com/markets/feeds/afx/2008/02/25/afx4694434.html).

5. See "Interchange fee," Wikipedia, February 21, 2009 (http://en.wikipedia.org/wiki/Interchange_fee), and "NRF Welcomes Senate Bill Requiring Visa and MasterCard to Negotiate over Hidden Annual Fee," National Retail Federation, June 5, 2008 (www.nrf.com/modules.php?name=News&op=viewlive&sp_id=530). These are fees and costs solely for transaction facilitation, not those in the credit market in which many card networks also participate.

6. That American Express has charged higher merchant fees than Visa or MasterCard likely reflects the higher income of its customers, additional services American Express offers, and the fact that until recently it did not offer credit and therefore did not make money on credit. Higher American Express fees therefore do not mean that Visa's and MasterCard's fees are or might be competitive. In any event, lower Visa and MasterCard fees would likely create competitive pressure on American Express to lower its fees.

7. In a three-party setup, such as American Express and Diners Club, a single bank handles both the acquiring and the issuing functions. American Express

now also has a four-party network, in which it is the single acquiring bank, after the restriction by the MasterCard and Visa networks prohibiting member banks from issuing American Express cards was ruled anticompetitive in *United States v. Visa U.S.A.* 344 F.3d 229.

8. Visa's current interchange fees are available at its U.S. website (<http://usa.visa.com/download/merchants/visa-usa-interchange-rates.pdf>).

9. See "Visa Hikes Overall Interchange 0.6%, Effective April 14," *Digital-TransactionsNews*, April 12, 2007 (www.digitaltransactions.net/newsstory.cfm?newsid=1311).

10. Similarly, in the three-party setup of American Express and Diners Club, the network can charge on both sides, that is, the merchants as well as the customers.

11. See "Settled Merchant Lawsuit," MasterCardWorldwide (www.mastercard.com/us/company/en/newsroom/merch_law.html [March 3, 2009]).

12. See *Card Acceptance and Chargeback Management Guidelines for Visa Merchants* (San Francisco: Visa U.S.A., 2008), p. 10 (http://usa.visa.com/download/merchants/card_acceptance_guide.pdf): "No Surcharging. Always treat Visa transactions like any other transaction; that is, you may not impose any surcharge on a Visa transaction. You may, however, offer a discount for cash or another form of payment (such as a proprietary card or gift certificate) provided that the offer is clearly disclosed to customers and the cash price is presented as a discount from the standard price charged for all other forms of payment. The discount may not be applied to [a] 'comparable card.' A 'comparable card' is any other branded, general purpose payment card that uses the cardholder's signature as the primary means of cardholder authorization (e.g., MasterCard, Discover, American Express). Any discount made available to cardholders who pay with 'comparable cards' must also be made available to cardholders who wish to pay with Visa cards."

13. See Lloyd Constantine, Gorden Schnell, Reiko Cyrs, and Michelle Peters, "The VISA Check/MasterMoney Antitrust Litigation," Constantine Cannon LLP, 2006, p. 24 (www.constantinecannon.com/pdf_etc/THEVISACHECKMASTERMONEYANTITRUSTLITIGATION.pdf).

14. See Steven C. Salop, "Practices That (Credibly) Facilitate Oligopoly Coordination," in *New Developments in the Analysis of Market Structure*, edited by Joseph E. Stiglitz and G. Frank Mathewson (New York: Macmillan 1986).

15. See, for example, Pete Hisey, "How High Can You Go?" *Credit Card Management* (April 1999): 105: "Visa, which says it has been at a disadvantage to MasterCard in the amount of cash it can allow an issuer to earn, says that its increases in interchange rates simply level the playing field. . . . Clearly, neither Visa nor MasterCard is content to allow the other the high ground, particularly as large issuers are deciding if they even want to stay with either association."

16. See letter from Bruce Mansfield, General Manager, Visa International, Australia & New Zealand, to John Veale, Head of Payments Policy, Reserve Bank of

Australia, April 7, 2005 (www.rba.gov.au/PaymentsSystem/Reforms/CCSchemes/SubmissionsCCIStd/visa_07042005_1.pdf).

17. Ibid.

18. See also Alan S. Frankel and Allan L. Shampine, "The Economic Effects of Interchange Fees," *Antitrust Law Journal* 73, no. 3: 627–73.

19. This proposal might be implementable by enforcement of existing antitrust laws against unreasonable restraint of trade. However, to avoid the delays and uncertainties of adjudication it may be simpler to enact legislation to ensure that contracts between card networks, merchants, acquiring banks, and issuing banks do not restrict the ability of merchants to preferentially steer (through pricing or otherwise) customers to a particular card network (Visa versus MasterCard versus AMEX); to a particular product of the card network (for example, debit rather than credit card); or to a particular issuer bank of the same card network (say, Citibank Visa versus Chase Visa).

20. See Frankel and Shampine, p. 641.

21. See Howard Chang, David Evans, and Daniel D. Garcia-Swartz, "The Effect of Regulatory Intervention in Two-Sided Markets: An Assessment of Interchange-Fee Capping in Australia," *Review of Network Economics* 4, no. 4 (2005): 329.

22. See Alan S. Frankel, "Toward a Competitive Card Payments Marketplace," mimeo, 2006, p. 32.

23. See Reserve Bank of Australia, Payments System Board, 2005 Annual Report, pp. 10–11.

24. Chang, Evans, and Garcia-Swartz reported a 60 to 70 percent reduction in the overall cost of transactions in the economy since card issuers have recovered 30 to 40 percent of the lost interchange fee revenue by charging higher fees to cardholders. See also Frankel, p. 37.