

Eder Fabrício Santos Souza

**Instrumentos eletrônicos de pagamento:
seu uso e suas implicações na política
monetária**

Belo Horizonte, MG
UFMG/Cedeplar
2010

Eder Fabrício Santos Souza

Instrumentos eletrônicos de pagamento: seu uso e suas implicações na política monetária

Dissertação apresentada ao curso de Mestrado em Economia do Centro de Desenvolvimento e Planejamento Regional da Faculdade de Ciências Econômicas da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do Título de Mestre em Economia

Orientador: Prof. Dr. Gilberto de Assis Libânio

Co-orientadora: Prof^a. Dr^a. Sueli Moro

Belo Horizonte, MG
Centro de Desenvolvimento e Planejamento Regional
Faculdade de Ciências Econômicas - UFMG
2010

Folha de Aprovação

*Trabalho dedicado à alegria com que
minha saudosa avó materna (1910-2007)
sempre encarou a vida, independente dos
percalços.*

AGRADECIMENTOS

Primeiramente, ao Banco Central do Brasil e ao Cedeplar/FACE/UFMG, por viabilizarem materialmente esse mestrado e essa dissertação.

Aos professores Gilberto Libânio, Sueli Moro, Mauro Sayar e Gary Dymski, bem como aos colegas de Banco Central José Ricardo Costa e Silva, Gabriel Garber, Adriana Sales e Samuel Bracarense, pelas contribuições enriquecedoras e frutíferas discussões.

Aos que contribuíram com a formação dos alicerces de quem sou:

- minha mãe Ercila, que me ensinou o valor de compreender e respeitar as pessoas, e meu pai Edmilson, que me ensinou o valor da educação;
- meus eternos mestres-amigos Cineas Santos, Pedro Reis (*in memoriam*), Silvano “Gauss” Falcão, Edwar “Ratão” Castelo Branco e Marcílio Flávio Rangel de Faria (*in memoriam*), que me ensinaram, respectivamente, a escrever, a comunicar-me, a brincar de matemática tão divertidamente quanto acampar ou andar de caiaque, a ler História e, mais uma vez, que se existe algo matematicamente incalculável é o valor da educação;
- minha saudosa avó “Mulatinha” (*in memoriam*), meus avós paternos Ferreira e Marieta, minha madrinha Valmira, meus primos Hideraldo e Ana Lydia, meus tios Ailson, Philes e Tarcila, e os “amigos de infância” Marcio Pio e Mateus Barros, todos eles familiares e amigos que são referências para minha vida e cuja sabedoria transcende, em muito, nossa vã filosofia;
- o Sport Club Corinthians Paulista, que me ensinou, em diversas ocasiões, a tratar a adversidade como um convite à superação.

Aos colegas de turma ou “de copo e de cruz”, que fizeram minha estadia em BH mais acolhedora, divertida e instigante: Weslem “Carioca”, Thiago Caliari, Ricardo, “Don Alejo”, Clarissa “Tita”, Carlinha, Aline, Sibelle, Lissinho, Tiago Prata, Ramon, Fefê, Fafá, Luiz, colegas de Montes Claros e de outras turmas; Wescley, Mariana e Rafaela; Renata e Alice; Jaque e Lili; Carol e Cris; Tim e Leo.

LISTA DE ABREVIATURAS E SIGLAS

ADF – *Augmented* Dickey-Fuller

ARIMA – *Auto-Regressive Integrated Moving Average* (modelo econométrico)

ATM – *Automatic Teller Machine*

BC – Banco Central

BCB – Banco Central do Brasil

BIS – *Bank for International Settlements*

CEDEPLAR – Centro de Desenvolvimento e Planejamento Regional

CPSS/BIS – *Committee of Payments and Settlements System*

DOC – Documento de Crédito

DS – *Difference Series* (modelo econométrico)

IEPs – Instrumentos eletrônicos de pagamento

IF – Instituição Financeira

IPS – Im, Pesaran e Shin (teste econométrico)

LLC – Levin, Lin e Chu (teste econométrico)

M2L – Mercado de Dois Lados

MQO – Mínimos Quadrados Ordinários

NFC – *Near Field Communication*

OLS – *Ordinary Least Squares*

PL – Patrimônio Líquido

TED – Transferência Eletrônica Disponível

TS – *Trend Series* (modelo econométrico)

UFMG – Universidade Federal de Minas Gerais

SUMÁRIO

1 INTRODUÇÃO	2
2 OS MERCADOS DE CARTÕES DE CRÉDITO E DÉBITO.....	4
2.1 Sistemas de pagamento: atacado e varejo	4
2.2 A “inovação tecnológica” em instrumentos de pagamentos	5
2.3 A estrutura dos mercados de cartões de crédito e débito	9
2.3.1 A estrutura de preços em um mercado de dois lados	16
2.3.2 Desenhos do negócio.....	20
2.3.3 Regras e regulamentos específicos no mercado de cartões.....	22
2.4 Trabalhos empíricos.....	26
2.5 Considerações gerais sobre o mercado de cartões	28
3 EXERCÍCIOS ECONÔMICOS SOBRE OS MERCADOS DE CARTÕES DE CRÉDITO E DÉBITO NO BRASIL.....	30
3.1 Uma modelagem sobre o lado comerciante/credenciador	30
3.2 Metodologia: raízes unitárias.....	32
3.3 Metodologia: estimações em painéis por efeitos fixos, por efeitos aleatórios e primeira diferença.....	34
3.4 Metodologia: modelagem VAR.....	39
3.5 Base de dados, análise descritiva e evolução histórica	40
3.6 Raízes unitárias e efeitos fixos e aleatórios	48
3.7 Modelo explicativo final: estimações em primeira diferença.....	53
3.8 Modelo preditivo VAR.....	56
3.9 Considerações gerais sobre os exercícios econômicos.....	62
4 IMPACTOS DOS INSTRUMENTOS ELETRÔNICOS DE PAGAMENTO SOBRE A POLÍTICA MONETÁRIA	64
4.1 A inter-relação entre as atividades de um banco central.....	64

4.2 A política monetária e seus mecanismos de transmissão	65
4.3 A discussão sobre moeda endógena	75
4.4 Efeitos dos instrumentos eletrônicos de pagamento sobre a política monetária	80
4.5 Notas conclusivas sobre política monetária e instrumentos eletrônicos de pagamento	83
5 CONCLUSÃO.....	86
REFERÊNCIAS BIBLIOGRÁFICAS	ERRO! INDICADOR NÃO DEFINIDO.

LISTA DE ILUSTRAÇÕES

FIGURA 2.1 – Dinâmica do mercado de cartões de crédito e débito	14
FIGURA 2.2 – Fluxo de pagamentos no mercado de cartões de crédito e débito	18
TABELA 3.1 – Análise descritiva dos dados de cartões de crédito	41
TABELA 3.2 – Análise descritiva dos dados de cartões de débito	42
TABELA 3.3 – Correlações entre variáveis para cartões de crédito.....	43
TABELA 3.4 – Correlações entre variáveis para cartões de débito.....	43
GRÁFICO 3.1 – Evolução da quantidade de transações e do volume transacionado em cartões de crédito no Brasil entre o primeiro trimestre de 2003 e o segundo trimestre de 2008	45
GRÁFICO 3.2 – Evolução da quantidade de transações e do volume transacionado em cartões de débito no Brasil entre o primeiro trimestre de 2003 e o segundo trimestre de 2008	46
GRÁFICO 3.3 – Evolução do volume transacionado em cartões de crédito no Brasil e do agregado monetário M1 entre o primeiro trimestre de 2003 e o segundo trimestre de 2008	47
GRÁFICO 3.4 – Evolução do volume transacionado em cartões de débito no Brasil e do agregado monetário M1 entre o primeiro trimestre de 2003 e o segundo trimestre de 2008	48
TABELA 3.5 – Correlogramas das variáveis para o modelo de cartões de crédito	49
TABELA 3.6 – Correlogramas das variáveis para o modelo de cartões de débito	49
TABELA 3.7 – Testes de raízes unitárias no painel de cartões de crédito.....	50
TABELA 3.8 – Testes de raízes unitárias no painel de cartões de débito.....	50
GRÁFICO 3.5 – Resíduos da regressão do modelo de cartões de crédito após estimação por efeito fixo	51
GRÁFICO 3.6 – Resíduos da regressão do modelo de cartões de crédito após estimação por efeito aleatório.....	52
GRÁFICO 3.7 – Resíduos da regressão do modelo de cartões de débito após estimação por efeito fixo	52

GRÁFICO 3.8 – Resíduos da regressão do modelo de cartões de débito após estimação por efeito aleatório	53
GRÁFICO 3.9 – Resíduos da regressão do modelo de cartões de crédito após estimação em primeira diferença	54
GRÁFICO 3.10 – Resíduos da regressão do modelo de cartões de débito após estimação em primeira diferença	55
TABELA 3.9 – Parâmetros estimados para os modelos de cartões de crédito e de débito – coeficiente e p-valor	55
TABELA 3.10 – Parâmetros estimados para os modelos VAR em primeira diferença para cartões de crédito e de débito – coeficientes e erros-padrão.	58
GRÁFICO 3.11 – Simulações impulso-resposta no modelo de cartões de crédito em primeira diferença com variável de sazonalidade q4	59
GRÁFICO 3.12 – Simulações impulso-resposta no modelo de cartões de débito em primeira diferença com variável de sazonalidade q4	60
GRÁFICO 3.13 – Simulações impulso-resposta no modelo de cartões de crédito em primeira diferença com 4 defasagens	61
GRÁFICO 3.14 – Simulações impulso-resposta no modelo de cartões de débito em primeira diferença com 4 defasagens	62
FIGURA 4.1 – Canais de transmissão da política monetária	73

RESUMO

Esta dissertação propõe duas discussões sobre o uso de instrumentos eletrônicos de pagamento: análises econométricas sobre o uso de cartões de crédito e débito no Brasil e uma discussão teórica acerca dos efeitos do uso de tais instrumentos sobre a política monetária. A análise econométrica consiste em dois modelos, um de caráter explicativo e outro preditivo do tipo VAR, que objetivam estimar a relação entre instituições financeiras credenciadoras de cartões de crédito e débito e estabelecimentos comerciais. Os resultados evidenciam a relevância da taxa de intercâmbio no mercado de cartões de crédito. A discussão teórica consiste nas possíveis implicações que a adoção cada vez maior de tais instrumentos de pagamento tenha sobre a eficácia da política monetária, a partir de uma análise de seus canais de transmissão. As conclusões, nesse caso, variam em função de premissas sobre mecanismos de transmissão de política monetária e endogeneidade da moeda.

Palavras-chave: instrumentos eletrônicos de pagamento, cartões de crédito, cartões de débito, política monetária.

ABSTRACT

This dissertation proposes two discussions on the use of electronic instruments of payment: econometric analysis on the use of credit cards and debit cards in Brazil and a theoretical discussion about the effects of the use of these instruments on monetary policy. The econometric analysis consists of two models, an explanatory panel and a predictive VAR, which aim to estimate the relationship between acquirers and merchants on credit cards and debit cards markets. The results show the importance of exchange rate on the credit cards market. The theoretical discussion focuses on the implications of the increasing adoption of such instruments of payment on the effectiveness of monetary policy, based on an analysis of its transmission channels. The findings, in this case, vary depending on assumptions about the transmission mechanism of monetary policy and money endogeneity.

Keywords: electronic instruments of payment, credit cards, debit cards, monetary policy.

1 INTRODUÇÃO

Uma das atribuições específicas de bancos centrais, dentro da sua função geral de zelar pela moeda, é a observância do sistema de pagamentos em seus aspectos de segurança e eficiência. A prática usualmente adotada é olhar para o sistema de pagamentos em dois blocos: atacado, referindo-se a transações interbancárias de grandes valores, e varejo, referindo-se a pagamentos que envolvem menores valores com maior giro. É nesse segundo bloco que se situa uma das grandes tendências da economia contemporânea: o uso de instrumentos eletrônicos de pagamento, sendo que, dentre estes, destacam-se, pelo acelerado crescimento, os cartões de crédito e de débito.

Este trabalho tem por objetivos discutir duas questões que emergem dessa tendência:

1. Que fatores explicam a crescente adoção de cartões de crédito e débito e a que ritmo se espera que tal adoção cresça no Brasil nos próximos anos?
2. Que implicações esse fenômeno traz às funções da moeda e/ou à eficiência da política monetária?

Para isso, em relação à primeira questão, o trabalho apresenta a dinâmica do setor e propõe exercícios econométricos sobre ela. Quanto à segunda questão, levanta discussões sobre a natureza e as implicações da moeda, segundo diversas vertentes de pensamento econômico.

Além desta introdução, esta dissertação está organizada em outros quatro capítulos. O Capítulo 2 explica o mercado de cartões, sua dinâmica e seus agentes. O Capítulo 3 apresenta os exercícios econométricos propostos, suas análises e resultados. O Capítulo 4 discute canais de transmissão de política monetária, moeda endógena e efeitos da crescente adoção de instrumentos eletrônicos de pagamento sobre a política monetária. Finalmente, o Capítulo 5 reúne as conclusões gerais do trabalho.

2 OS MERCADOS DE CARTÕES DE CRÉDITO E DÉBITO

2.1 Sistemas de pagamento: atacado e varejo

A função geral de bancos centrais de zelar pela moeda pode ser desmembrada em três funções específicas: a gestão da política monetária; a supervisão dos bancos (ou, de modo mais geral, do sistema financeiro como um todo); e a observância do sistema de pagamentos¹.

Entende-se por sistema de pagamentos todo o conjunto de mecanismos pelos quais um montante financeiro debitado de um agente é creditado em igual valor a outro agente, com o mínimo grau de incerteza e menores tempo e custo possíveis². Dentro da arquitetura desses sistemas, bancos centrais costumam diferenciar duas categorias de pagamentos, atacado e varejo, cujas naturezas distintas geram preocupações igualmente distintas às autoridades monetárias, resultando, via de regra, em desenhos distintos para os sistemas que as regem.

Sistemas de pagamentos ditos de atacado são mecanismos voltados para efetivação de transações interbancárias, notadamente transações com títulos e demais ativos financeiros entre instituições financeiras (IFs) com objetivos de obter rentabilidade ou suprir necessidade de liquidez. Usualmente são transações de grandes valores mas não muito giro.

Os sistemas de pagamentos de varejo referem-se a pagamentos efetuados por pessoas físicas e jurídicas, em geral para consumo. Os principais instrumentos para tal modalidade de pagamento, além do papel-moeda propriamente, são cartões de débito, cartões de crédito, cheques, bloquitos de cobrança e, em

¹ Blinder (1998) comenta que atividades ligadas a sistema de pagamentos e a supervisão do sistema financeiro ocupam mais tempo e esforço na agenda cotidiana de bancos centrais do que propriamente atividades ligadas a política monetária (mais visadas nas discussões acadêmicas).

² No glossário do CPSS/BIS (2003, p. 38), o verbete “*payment system*” é definido como “*a payment system consists of a set of instruments, banking procedures and, typically, interbank funds transfer systems that ensure the circulation of money*”.

alguns casos, DOCs e TEDs. Em oposição aos pagamentos de atacado, os de varejo geralmente envolvem pequenos valores e bem maior giro.

Dadas tais características, bancos centrais costumam concentrar sua observância sobre os sistemas de atacado em aspectos de segurança (minimização da incerteza na liquidação das operações), enquanto a dos sistemas de varejo focam-se mais na sua eficiência (rapidez e baixo custo das operações).

O mercado de cartões de crédito e débito, abordado neste trabalho, situa-se, portanto, no âmbito dos sistemas de pagamento de varejo, e sua observância por parte da autoridade monetária é mais focada na sua eficiência.

2.2 A “inovação tecnológica” em instrumentos de pagamentos

Este tópico faz um breve apanhado histórico da evolução do mercado de cartões, tanto em seus aspectos de desenho de mercado quanto em seus aspectos tecnológicos.

Se, até meados do século 20, a humanidade levou milênios para transitar do escambo para a moeda metálica, o papel-moeda, os cheques (ordens de crédito preenchidas em papel) e os ativos financeiros em papel (títulos e ações ao portador), o ritmo de inovação dos instrumentos de pagamentos desde então tem sido vertiginoso, favorecido tanto pelas novas possibilidades tecnológicas nas áreas da microeletrônica e das telecomunicações quanto pela consolidação dos bancos e do setor financeiro em geral (com suas inovações financeiras) como os grandes motores de uma nova fase do capitalismo.

Até a década de 1950, o uso de cartões em pagamentos limitava-se a cartões de loja, que não cobravam tarifa de adesão e cobravam juros apenas dos usuários que atrasassem o pagamento da fatura (oferecendo, assim, um breve período livre de juros entre a compra e a data da fatura), mas só podiam ser usados nos estabelecimentos comerciais que os emitiam³. Em 1950, foi lançado o Diners

³ Acredita-se que o primeiro produto com esse formato tenha sido lançado em 1914, pela Western Union Telegraph Company, nos EUA. (Eagleton e Williams, 2007)

Club⁴, primeiro cartão de crédito aceito em diversos estabelecimentos⁵ e o primeiro a operar em plataforma de dois lados (*2-sided platforms*), ou seja, cobrando tarifa tanto do comerciante quanto do usuário. Ainda era, no entanto, uma plataforma chamada de 3-partes (*unitary system*)⁶, em que o mesmo intermediador emite cartões aos usuários e credencia os estabelecimentos a aceitá-lo. Em 1958, surgiu a American Express, com o mesmo modelo de negócio, mas que se tornaria a principal bandeira em plataformas de 3-partes. Também em 1958, o Bank of America começou a operar o BankAmericard, inicialmente com o mesmo formato, mas que traria uma inovação em 1966, quando o banco começou a trazer outros bancos para um sistema de franquia de seu cartão de crédito, formato que possibilitou expandir tanto a emissão de cartões aos usuários quanto o credenciamento de estabelecimentos que os aceitassem. Pela primeira vez, isolava-se a figura da “bandeira” (proprietário do esquema) das atividades de credenciamento e emissão. Mas um inconveniente persistia: o usuário de um cartão BankAmericard emitido por um banco X só poderia usá-lo em estabelecimentos credenciados pelo mesmo banco X. Aproveitando-se dessa falha de mercado, uma cooperativa de bancos lançou o MasterCharge, que seria o predecessor do MasterCard, como a primeira plataforma de 4-partes (*multiparty system*), ou seja, separando as figuras do emissor de cartões aos usuários (papel dos bancos individualmente) e do credenciador de estabelecimentos a aceitá-los (papel da cooperativa). Foi, assim, o primeiro modelo a aproximar-se da regra hoje chamada “*honor all cards*”, que será discutida adiante. Poucos anos depois, em 1970, em resposta à nova concorrência, o sistema de franquias BankAmericard replicou o modelo, convertendo-se na cooperativa NBI, que seria a predecessora da Visa. No ano

⁴ A criação do Diners Club acompanha uma história pitoresca. O empresário Frank McNamara jantava em um restaurante quando se deu conta de que não portava dinheiro em espécie nem cheque consigo. Desconsertado, idealizou um produto proposto a evitar tal situação. (*op. cit.*)

⁵ Já em 1967, havia mais países que aceitavam Diners do que quantos eram então países-membros da ONU. (*op. cit.*)

⁶ A terminologia usada em português (plataformas de “3-partes” e “4-partes”), que se refere ao aspecto de a “bandeira” acumular ou não para si as atividades de credenciadora e emissora de cartões, não deve ser confundida com a terminologia “mercado de 2-lados”, referente ao desenho microeconômico em que um “produtor” precisa equilibrar “curvas de demanda” de 2 grupos consumidores distintos. Tanto plataformas de 3-partes quanto de 4-partes funcionam como mercados de 2-lados.

seguinte, essa cooperativa acrescentou ao modelo uma tarifa a ser paga pelo credenciador ao emissor, chamada tarifa de intercâmbio, entendendo que ela seria fundamental para o equilíbrio de tal mercado. Pouco depois, sua principal concorrente replicou o modelo. A tarifa de intercâmbio será discutida adiante nesse trabalho⁷.

Se os cartões de crédito foram a grande inovação financeira na área de pagamentos entre as décadas de 1950 e 1980, na década de 1990, popularizaram-se os cartões de débito. Valendo-se da estrutura já montada nas décadas anteriores, de bancos emissores e credenciadores com suas bases de clientes e das possibilidades tecnológicas de redes eletrônicas interligando bancos e estabelecimentos comerciais em todo o mundo, os proprietários de esquemas de cartões de crédito criaram um novo produto: cartões que, ao invés de oferecerem ao usuário um crédito (sem juros até a data da fatura), debitam suas contas no próprio ato da compra⁸. O novo produto objetivava substituir os cheques e atrair usuários com dificuldade em administrar liquidez e crédito em suas reservas no uso de cartões de crédito.

No fim da década 2000, duas tendências mercadológicas para cartões de crédito e débito são o *co-branding* e a convergência com serviços que facilitem a ampliação da base de clientes (cartões de loja ou cartões de serviços). A prática do *co-branding* refere-se à emissão de cartões em parceria com uma empresa (companhia aérea ou clube esportivo, por exemplo), com objetivo de fidelizar o portador por intermédio da sua afinidade com a marca parceira. A convergência com cartões de loja ou de serviços (cartões de benefício-alimentação ou de uso em transporte público, por exemplo) objetiva facilitar ao emissor a captura, para sua rede, de uma base de clientes já consolidada.

Em termos tecnológicos, as inovações que se seguiram às tradicionais tarjas magnéticas dos cartões foram os “*smart cards*”. Neles, ao invés da simples tarja com dados do portador gravados (em que a leitura e processamento das

⁷ A maior parte da história contada neste parágrafo está em Evans e Schmalensee (2005c).

⁸ A rigor, o débito pode acontecer até poucos dias depois da compra, a depender do arranjo de compensação e liquidação usado em cada sistema de pagamentos.

informações é feita pelo terminal de captura), existe um *chip* com capacidade própria não só de armazenamento mas também de processamento. Esse armazenamento e processamento distribuídos (entre o servidor *mainframe* da administradora do cartão e o *chip* embutido no próprio cartão) aumentam tanto a segurança (já que permite ao cliente deixar uma senha armazenada no *chip*, e não em servidores mais vulneráveis ao ataque de piratas digitais) quanto a rapidez da transação (já que a etapa de checagem da senha digitada no terminal de captura é feita no próprio *chip*, para só então o terminal de captura comunicar-se com o servidor, informando os dados da transação). Uma modalidade específica de *smart cards* são os que trazem o *chip* acoplado a uma antena (não-visíveis na superfície do cartão) e, por isso, não requerem inserção direta em terminais de captura, mas apenas uma aproximação física – daí serem chamados de *NFC* (*near field communication*) ou *contactless*. Sua principal vantagem é a agilidade na transação. Para serviços que costumem envolver filas, tais como transporte público e eventos esportivos, em esquemas de pagamento do tipo pré-pago (ou de crédito “pré-aprovado”), o uso de cartões *NFC* agiliza o fluxo.

A fronteira subsequente que emerge é o uso da telefonia celular em esquemas de pagamento eletrônico (“*mobile payment*”). Há vários desenhos operacionais possíveis para esse mercado, mas os pontos comuns entre eles são o consumidor usar seu número de celular como identificador (ao invés do número do cartão) e valer-se do próprio aparelho para informar a senha (via mensagem de texto, por exemplo) que completa a operação. O principal argumento favorável a tal indústria é a maciça penetração que a telefonia celular tem mostrado em todo o mundo, inclusive nas faixas de renda mais baixas, em oposição à bancarização, que ainda tem índices baixos em muitas nações. A principal discussão entre reguladores acerca de tal mercado é o papel das operadoras de telefonia, já que elas tomariam parte no esquema, sobrepondo-se a funções exercidas por credenciadores ou emissores (a depender do exato desenho do esquema) no mercado de cartões, mas, em princípio, não deveriam assumir funções específicas de entidades financeiras (sem que fossem assim reguladas).

2.3 A estrutura dos mercados de cartões de crédito e débito

Pode-se intuir, de imediato, a participação de três agentes no mercado de cartões – o consumidor (portador do cartão), o estabelecimento comercial e o banco (ou instituição financeira intermediadora). Para que tal mercado prospere, é necessário que os três agentes aceitem participar dele.

Para o *consumidor*, o interesse pelos cartões estaria na comodidade que oferecem (função utilidade do consumidor). Tal comodidade manifesta-se principalmente na forma de menor necessidade de fazer saques de moeda-manual (o conhecido “custo de sola de sapato”) e de maior agilidade nos pagamentos – dispensadas as necessidades de preenchimento de cheques e de manipulação de cédulas e moedas tanto para o pagamento pela mercadoria quanto para o recebimento de troco⁹. Além disso, usuários de cartões costumam receber bônus como incentivo, tais como milhas em companhias aéreas e, no caso de cartões de crédito, têm ainda o incentivo natural de postergar um pagamento e auferir juros sobre o capital durante o período de postergação.

Por outro lado, alguns fatores podem desestimular o uso de cartões pelos consumidores. O primeiro deles seria uma relutância cultural, esperada principalmente em grupos sociais menos adaptados à adoção de novas tecnologias (indivíduos mais velhos e/ou menos escolarizados)¹⁰. Tal barreira, no entanto, vai enfraquecendo à medida que indivíduos mais jovens e mais escolarizados vão passando a representar uma fatia maior nas estatísticas demográficas e à medida que os próprios indivíduos mais velhos e menos escolarizados vão superando a relutância cultural. Tal percepção, se vista isoladamente, levaria a intuir uma tendência natural ao aumento na adoção de cartões. Ao lado da relutância cultural, outros fatores limitantes da expansão desse mercado seriam baixa renda dos indivíduos e alto custo dos cartões (tanto

⁹ A esse respeito, ver Borzekowski e Kiser (2006) – *The Choice at the Checkout* – que faz uma análise empírica para os EUA apontando como variável relevante na escolha pelo instrumento de pagamento “*elapsed time at the checkout*”.

¹⁰ A influência de fatores demográficos nesse aspecto tem suas elasticidades medidas em Borzekowski e Kiser (2006).

anuidades, cada vez mais em desuso, quanto taxas de juros cobradas sobre crédito rotativo).

Em regiões com elevados índices de criminalidade, é possível ainda que usuários considerem tal variável no momento de decidir entre portar moeda-manual (e correr risco de assalto) ou portar cartões (e correr risco de sequestros-relâmpago). Nesse caso, a percepção de violência pode funcionar tanto como mecanismo de incentivo quanto de desincentivo da adoção do uso de cartões.

Sob a ótica do *comerciante*, a decisão será fundamentada na elevação ou manutenção dos lucros – ou minimização das perdas, caso estas sejam inevitáveis. As taxas cobradas pela instituição financeira intermediadora (notadamente, aluguel de equipamento e taxa de desconto) são um fator de desestímulo. Por outro lado, o risco de perder clientes para o concorrente que aceite cartões pode impor ao agente a necessidade de adoção desse instrumento. Uma solução possível para esse dilema é a discriminação de preços entre pagamentos em espécie (moeda-manual) e pagamentos em cartão. Em alguns países¹¹, associações de cartões de crédito estabelecem regras contrárias à discriminação de preços. Em outros¹², tal imposição é proibida pelas autoridades monetárias. No entanto, observa-se empiricamente que, mesmo em regiões onde a discriminação é permitida, ela não costuma ser adotada pelos comerciantes, temerosos de perderem mercado para seus concorrentes¹³.

Do ponto de vista da *instituição financeira intermediadora*, a decisão também será fundamentada na elevação do lucro. Para ela, tal mercado significa receitas sob a forma de taxa de desconto e aluguel de equipamento, cobradas do comerciante, e anuidade (exclusiva de cartões de crédito, mas cada vez mais em desuso) e taxas de juros sobre crédito rotativo (também exclusivas do caso de cartões de crédito). Significa ainda custos com emissão e distribuição de cartões e equipamentos,

¹¹ EUA e a maioria das economias mais desenvolvidas.

¹² Holanda, Suécia, Reino Unido, Austrália.

¹³ Chakravorti e To (2006) apresentam um modelo para o qual comerciantes ficam em uma condição de “dilema do prisioneiro” quanto à aceitação de cartões: embora a aceitação seja pior para o conjunto dos agentes, ela é preferível para cada um individualmente, e acaba se configurando como a situação de equilíbrio.

manutenção de rede, análise de crédito e gestão de risco, gestão e divulgação da marca e do produto e acordos comerciais para oferta de bônus. Como o número de participantes intermediadores é bem menor que o número de participantes das pontas da intermediação (consumidores e comerciantes), o agente intermediador tem poder de mercado suficiente para ajustar suas variáveis de custo e receita de modo a maximizar seu lucro. Ou seja, para esse agente, não há razões contrárias ao desenvolvimento desse mercado.

Por fim, cumpre observar também o interesse de um *agente externo regulador*, que poderia atuar no sentido de limitar tal mercado. Sob a hipótese de que o interesse do agente regulador esteja no “ganho social” da existência e expansão do mercado de cartões. Análises empíricas aplicadas a diversos países concluem que o pagamento eletrônico custa entre 1/3 e 1/2 do custo do pagamento em cheque e que a migração total para instrumentos eletrônicos produziria economia anual de 3% do PIB de uma região¹⁴. Como, ademais, não se percebem externalidades negativas resultantes desse mercado, não há interesse do agente externo regulador em reprimi-lo.

Mesmo sob uma hipótese que vincule o interesse do regulador a *lobbies* ou corrupção, o agente que exerceria influência mais forte sobre o regulador – e, conseqüentemente, teria seus interesses privilegiados – seria o intermediador, que concentra maior poder de mercado. Como o mercado de cartões é interesse do intermediador, o regulador não teria razões para impor barreiras.

Um apanhado das intuições expostas até aqui leva a crer que o mercado de cartões é de pleno interesse de reguladores e intermediadores (instituições financeiras), é acatado por comerciantes (em condição de “dilema do prisioneiro”) e apresenta diversos incentivos aos consumidores, tendo como principais empecilhos à sua expansão baixos níveis de renda dos potenciais usuários e altos custos – notadamente taxas de juros sobre crédito rotativo, para o caso de cartões de crédito.

¹⁴ Banco Central do Brasil (2007) e Banco Central do Brasil (2005). Nestes trabalhos, o cálculo dos custos sociais engloba apenas o lado da oferta e processamento desses instrumentos (ou seja, custos bancários), metodologia mais robusta por não depender de variáveis qualitativas auto-declaráveis.

Visto que o mercado de cartões apresenta como característica essencial a necessidade de dois lados distintos (consumidor e comerciante) aderirem à plataforma de negócio, ambos na forma de “demanda” pelo produto, sua modelagem deve basear-se em uma fundamentação teórica conhecida como “mercado de dois lados” (M2L).

Mercados de dois lados (M2L) são plataformas com a presença de dois grupos distintos de participantes e existência de externalidades positivas de rede, que geralmente não são internalizadas diretamente pelos grupos¹⁵. As variáveis econômicas centrais que os caracterizam são a estrutura de preços, as decisões quanto ao desenho do negócio, e a existência de regras e regulamentos específicos.

O exemplo mais comumente citado para exemplificar tal conceito é o de uma boate, que pretende atrair para sua clientela tanto homens quanto mulheres, ou seja, “grupos distintos de participantes”. Os homens serão mais atraídos para a boate quanto mais mulheres a frequentarem, e vice-versa, o que caracteriza “existência de externalidades positivas de rede”. Por fim, supõe-se que, em geral, os casais eventualmente formados na boate não se conheceriam caso a boate não existisse, ou seja, “as externalidades não seriam diretamente internalizadas pelos grupos”.

Um resultado importante no exemplo da boate, portanto, é que seu proprietário deve preocupar-se em atrair ambos os lados em tal proporção que maximize a satisfação conjunta de ambos. E a forma de buscar esse equilíbrio é através da estrutura de preços do mercado, isto é, cobrando preços diferenciados p_H e p_M , condizentes com as elasticidades-preço de cada grupo de maneira que atraia a proporção ideal de homens e mulheres.

Assim, Rochet e Tirole (2004) caracterizam os mercados de dois lados por:

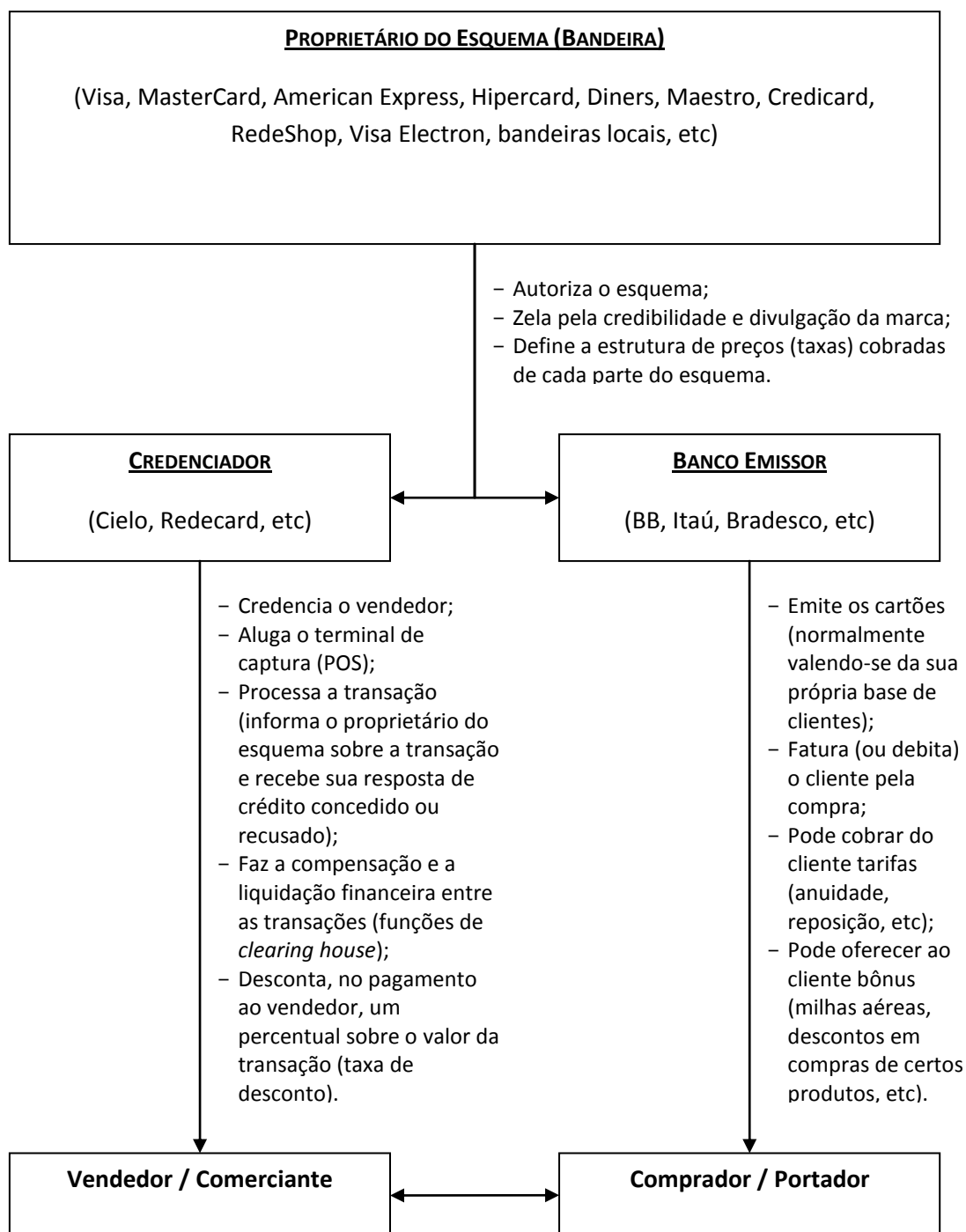
¹⁵ Essa definição é usada por Evans e Schmalensee (2005b). Embora o trabalho seminal sobre mercado de dois lados seja Rochet e Tirole (2002), a maior parte da discussão apresentada no presente trabalho baseia-se em Evans e Schmalensee (2005b), que fizeram uma compilação da teoria formulada até então.

- presença de *externalidade de rede*, ou seja, o valor de um bem para um grupo-agente depender diretamente da demanda dos demais grupos-agentes por esse bem;
- demanda pelo bem depender da *estrutura de preço* no mercado, ou seja, da relação entre preços p_A/p_B cobrados aos grupos-agentes A e B, respectivamente (enquanto que mercados de um único lado dependem apenas do *nível agregado de preço* $p = p_A + p_B$); e
- negociação direta entre os grupos-agentes não ter eficácia na internalização das externalidades.

No caso dos cartões, a “bandeira” deve atrair consumidores e comerciantes em uma proporção tal que o grau de adesão de um grupo estimule o grau de adesão do outro. E o equilíbrio é buscado na estrutura de preços que engloba o quanto consumidor e comerciante pagam (ou recebem) por aderirem ao esquema de cartões.

Especificamente na indústria de cartões, existe, no entanto, uma particularidade. Conforme já mencionado no breve histórico apresentado no tópico anterior, esse mercado buscou, em muitos casos, um arranjo alternativo ao simples “esquema de 3-partes” (*unitary system*), em que o intermediador entre consumidor e comerciante é um único agente. O formato alternativo – e majoritariamente empregado – é o “esquema de 4-partes” (*multiparty system*), em que um agente se responsabiliza pela adesão dos consumidores e outro pela adesão dos comerciantes. O desenho geral é o da Figura 2.1.

FIGURA 2.1 – Dinâmica do mercado de cartões de crédito e débito



O proprietário do esquema (também chamado “bandeira” ou “*label*”) – Visa, MasterCard, American Express e outras – administra a marca, inclusive autorizando seu uso, definindo a estrutura de preços e, em alguns casos, financiando os credenciadores.

Quando um estabelecimento comercial decide aceitar pagamentos via cartões, recorre a um credenciador (também chamado “adquirente” ou “*acquirer*”), tal como Cielo¹⁶ ou RedeCard, que irá cadastrar o estabelecimento em sua rede, prover o equipamento necessário (mediante um aluguel) e cobrar-lhe uma taxa percentual sobre cada transação efetuada, chamada taxa de desconto¹⁷. Anos atrás, no Brasil, cada banco era um credenciador, responsável por filiar seus estabelecimentos comerciais, ou seja, cada banco operava simultaneamente como credenciador (banco do comerciante) e emissor (banco do cliente), criando sua própria rede de aceitação. Tal modelo obrigava os estabelecimentos comerciais interessados em participar das redes de cartões a terem vários terminais (grosso modo, um para cada grande banco) e também obrigava cada banco participante a ter sua própria central de autorização. A multiplicidade de custos levou associações de bancos a criarem a VisaNet (atual Cielo) e a Redecard, centralizando essas operações e responsabilizando-se pelas bandeiras Visa e Mastercard, respectivamente. Por isso, hoje, credenciadores e bandeiras têm relação de praticamente um-para-um¹⁸, ou seja, cada “banco” credenciador fica responsável por representar uma bandeira. Ao mesmo tempo em que, para cada transação, o credenciador “recebe” do estabelecimento comercial uma taxa de desconto, ele “paga” à administradora uma tarifa de intercâmbio¹⁹, cuja razão da existência será explicada mais adiante.

Quando um consumidor decide adotar um cartão de crédito ou débito, recorre a um banco emissor (“*issuer*”), que fornecerá o cartão ao cliente, fará a gestão de

¹⁶ Cielo é a antiga VisaNet. A empresa optou por mudar de nome para desassociar sua imagem da bandeira Visa, depois que a Secretaria de Desenvolvimento Econômico determinou o fim do contrato de exclusividade entre credenciador e bandeira a partir de junho de 2010.

¹⁷ Conforme será mostrado adiante, as taxas de desconto praticadas no Brasil variaram, no período analisado (2003Q1-2007Q2 para cartões de crédito e 2001Q4-2007Q2 para cartões de débito), entre 1,82% e 3,09% para cartões de crédito e entre 0,59% e 1,63% para cartões de débito.

¹⁸ Não há relação estritamente 1:1. O único contrato de exclusividade, vigente até junho de 2010, é entre Cielo e Visa. A bandeira MasterCard, embora tenha como grande credenciadora a RedeCard, autoriza outros credenciadores. E a própria RedeCard também credencia estabelecimentos para outras bandeiras menores.

¹⁹ Conforme será mostrado adiante, as taxas de intercâmbio praticadas no Brasil variaram, no período analisado, entre 0% e 1,87% para cartões de crédito e entre 0,24% e 0,93% para cartões de débito.

risco – definindo limites de crédito (para o caso de cartões de crédito) e taxa de juros para o caso de inadimplência – e poderá ainda cobrar tarifas de anuidade (cada vez mais em desuso) e conceder bônus em função do montante transacionado via cartões.

Uma terminologia comum usada pelos bancos emissores para o mercado de cartões de crédito refere-se a “clientes de conveniência” e “clientes-revólver”. Os primeiros usam o cartão de crédito tão-somente para postergar o pagamento, sem efetivamente fazerem uso do crédito que lhes é autorizado – ou seja, pagam integralmente suas faturas²⁰. O segundo grupo usa os cartões para efetivamente tomarem crédito, pagando apenas parcialmente suas faturas e incorrendo em juros chamados de valor rotativo²¹.

Conforme já mencionado, as variáveis econômicas centrais que caracterizam um mercado de dois lados (M2L) são:

- a estrutura de preços;
- as decisões quanto ao desenho do negócio; e
- a existência de regras e regulamentos específicos.

Os itens seguintes discutem cada uma dessas variáveis.

2.3.1 A estrutura de preços em um mercado de dois lados

Os preços de cada lado do mercado dependem não só de seu custo marginal, mas também da elasticidade-preço da demanda de cada lado, do valor que a entrada de um membro em um dos lados gera para os participantes do outro lado (externalidade de rede) e do grau de competição inter-plataforma e intra-plataforma.

²⁰ Parte da literatura em língua inglesa refere-se aos cartões usados dessa forma como *charge cards* e ao período livre de juros entre a compra e o pagamento da fatura como *grace period*.

²¹ Conforme será mostrado adiante, no Brasil, no período analisado, a razão entre o volume financiado no rotativo e o volume total movimentado em cartões de crédito teve média ponderada de 10,8%, mas variou desde zero até 97%, dependendo do emissor do cartão.

As elasticidades-preço da demanda de cada lado cumprem papel determinante na alocação dos preços. O lado com maior elasticidade tende a ter seu preço reduzido de forma a viabilizar a entrada de participantes desse lado na rede. Em geral, o lado menos elástico tende a pagar a maior parte dos custos da indústria.

A estrutura de apreçamento deve promover incentivo a pelo menos um dos lados sem que isso resulte em punição severa (pesada o bastante para desestimular a participação) ao outro lado. O apreçamento em mercados de dois lados não necessariamente reflete sua estrutura de custos relativos, mas também o excedente gerado para um dos grupos quando o mercado incorpora um indivíduo a mais do outro grupo (Rochet e Tirole, 2003a; 2003b; 2004).

No exemplo do mercado de cartões, consumidores têm um ganho excedente (chamemo-lo de G_A) quando mais um estabelecimento comercial é incorporado ao mercado. Da mesma forma, comerciantes têm um ganho excedente (G_B) quando mais um consumidor adere ao uso de cartões. Na estrutura de apreçamento do mercado para os dois lados, tudo o mais constante, o grupo que tiver ganho excedente maior irá “subsidiar”²² o outro grupo. O lado que “subsídia” é chamado centro de lucro, enquanto o outro é chamado centro de perda (ou, pelo menos, financeiramente neutro).

O que se observa empiricamente no mercado de cartões é os estabelecimentos comerciais “subsidiarem” os consumidores, ou seja, arcarem com taxas elevadas enquanto os consumidores muitas vezes recebem bônus.

Nos esquemas de 3-partes (também chamados “cadeias fechadas”), em que um mesmo intermediador acumula as funções de credenciador, “bandeira” e emissor²³, o apreçamento dos dois lados do mercado é feito de forma direta, com vistas a maximizar a receita com as transações realizadas. Já no caso mais comum, o dos esquemas de 4-partes (ou “cadeias abertas”), em que

²² Embora se questione a adequação conceitual da terminologia “subsídio” a esse fenômeno – já que não se trata de um subsídio deliberado –, ela é amplamente usada na literatura.

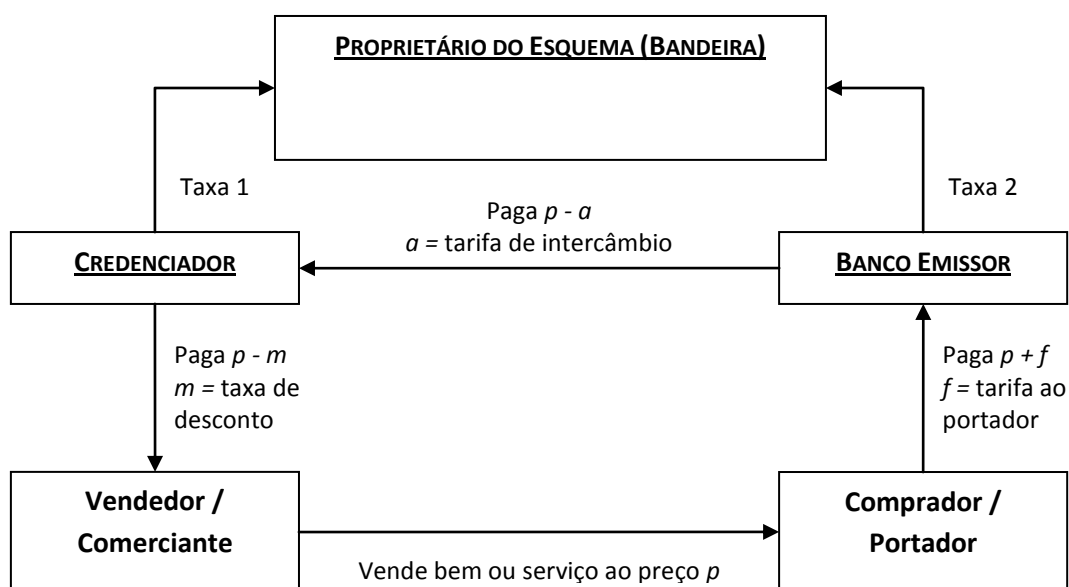
²³ O caso mais notório mundialmente é o da American Express, cuja operação no Brasil pertence à Bradesco Cartões.

credenciador, “bandeira” e emissor são agentes distintos, o instrumento usado para equilibrar o apreçamento nas duas pontas é a tarifa de intercâmbio.

A tarifa de intercâmbio é “paga” pelo credenciador para a administradora a cada transação realizada no cartão²⁴. É esperado, portanto, que o credenciador repasse esse custo para o estabelecimento comercial. Na outra ponta, como a tarifa de intercâmbio é uma forma de receita para a administradora, ela pode repassar o benefício para o consumidor, na forma de bônus.

A Figura 2.2 esquematiza a estrutura de preços padrão em um mercado de cartões²⁵.

FIGURA 2.2 – Fluxo de pagamentos no mercado de cartões de crédito e débito



A tarifa ao portador cresce com o custo dos bancos emissores e decresce com o grau de competição entre emissores, com a tarifa de intercâmbio e com o nível de resistência dos portadores de cartão – quanto maior a elasticidade-preço da demanda dos portadores por serviços de cartão, maior a sua resistência.

²⁴ A rigor, ela não é paga, mas descontada do repasse que a administradora faz ao credenciador durante o fluxo monetário consumidor-estabelecimento.

²⁵ Figura extraída de Banco Central do Brasil; Secretaria de Acompanhamento Econômico; Secretaria de Direito Econômico (2009).

O estabelecimento arca com dois tipos de custos. O custo variável, conhecido como taxa de desconto, que é calculado sobre o valor da transação. O custo para permanecer na plataforma é o custo fixo, que geralmente corresponde ao aluguel de equipamentos e aos custos de manutenção de *software*.

A taxa de desconto cobrada dos comerciantes também é função de diversas variáveis. Ela é crescente com o custo dos credenciadores e com a tarifa de intercâmbio, e geralmente é decrescente em função do grau de competição entre credenciadores e do nível de resistência dos estabelecimentos. O grau de competição entre emissores e entre credenciadores está também ligado ao grau de integração vertical que se observa de cada lado do mercado.

O nível de resistência dos comerciantes, além de ser função crescente da elasticidade-preço da demanda por serviços de cartão dos estabelecimentos, também está ligado ao grau de competição existente entre os estabelecimentos.

A tarifa de intercâmbio é, em geral, paga pelo credenciador ao emissor. A sua forma de cobrança é bastante heterogênea. Mais comumente ela corresponde a uma porcentagem do valor da transação, mas também pode assumir a forma de valor fixo por transação, ou mesmo de combinação de tarifa fixa com tarifa variável.

Percebe-se que a tarifa de intercâmbio é o instrumento usado pela administradora (“bandeira”) para otimizar a estrutura de preços cobrados aos dois lados do mercado – consumidores e estabelecimentos comerciais. Se a maximização de receita apontar para um “alto preço” para o estabelecimento e “baixo preço” para o consumidor, a administradora cobrará uma tarifa de intercâmbio elevada. Se apontar no sentido contrário, a tarifa de intercâmbio cobrada será baixa.

O entendimento da tarifa de intercâmbio como “preço de equilíbrio” do mercado (Chakravorti e To, 2006; Rochet e Tirole, 2004) será discutido no modelo proposto para o exercício econométrico deste trabalho.

Os preços podem ainda ser divididos em preço de entrada (*access price*) e preço de utilização (*variable price*). Os preços de utilização influenciam diretamente o

nível de utilização da plataforma pelos participantes. Os preços de entrada, dados os de utilização, determinam quantos usuários aderirão à plataforma.

2.3.2 Desenhos do negócio

O desenho do negócio torna-se importante no M2L em função da presença de externalidades de rede, das elasticidades-preço da demanda e devido à necessidade de se trazer os dois lados para o mercado.

O primeiro modelo proposto para o mercado de cartões (Baxter, 1983) supôs uma concorrência perfeita intra-plataforma. Nele, a tarifa de intercâmbio era posta como preço fundamental capaz de internalizar integralmente a externalidade de rede que existe no mercado.

Em um modelo alternativo de concorrência perfeita intra-plataforma (Carlton e Frankel, 1995), a internalização da externalidade de rede pode ocorrer por meio de algum instrumento alternativo à tarifa de intercâmbio. Para tal, basta que não exista regra de não-sobrepreço, que proíbe diferenciação de preços conforme o instrumento de pagamento utilizado. Nesse caso, as transações são realizadas normalmente e os custos do instrumento de pagamento são arcados por aqueles que efetivamente o utilizem. Se, a princípio, esse mecanismo pareça induzir a escolha do instrumento de pagamento pelo portador, o apreçamento direto dos instrumentos de pagamento não é trivial, pois outros instrumentos de pagamento (cheques e papel moeda, por exemplo) não são diretamente apreçados (Palley, 2002a), e, assim, o apreçamento direto de apenas um instrumento de pagamento poderia gerar um equilíbrio com bem-estar social menor.

Outra modelagem possível é manter uma concorrência intra-plataforma, mas atribuindo poder de mercado no lado dos emissores (Rochet e Tirole, 2002). Nesse caso, quanto maior o grau de competição entre emissores, maior o nível de resistência dos estabelecimentos e seu poder de barganha, o que tende a acarretar diminuição na tarifa de intercâmbio. Ainda nesse caso, outros fatores que afetariam o nível de resistência dos estabelecimentos são a divisão da tarifa ao portador em uma parte fixa e outra variável, o grau de informação dos

consumidores (inversamente proporcional ao nível de resistência), a existência de competição entre plataformas e a existência de regra de não-sobrepreço.

Outros modelos relevantes tratam a concorrência intra-plataforma com poder de mercado nos dois lados. No primeiro deles aqui analisado (Schmalensee, 2002), buscam-se os determinantes da tarifa de intercâmbio sob três abordagens “ótimas”: a que maximiza lucro; a que maximiza bem-estar social; e a que maximiza o número de transações da plataforma. Para tanto, tomam-se como hipóteses: heterogeneidade de consumidores e estabelecimentos; tomada de decisões estratégicas (via jogos) pelos estabelecimentos quanto à aceitação de cartões de pagamento (ao contrário de modelos que impõem essa aceitação como dada); e possibilidade de poder de mercado tanto de emissores quanto de credenciadores. A primeira conclusão é que a magnitude da tarifa de intercâmbio que maximiza o valor da rede depende dos seguintes fatores: (i) função-objetivo da plataforma; (ii) diferença de custos entre emissores e credenciadores; (iii) diferença de elasticidades-preço de portadores e comerciantes; (iv) diferença de grau de competição entre os dois lados do mercado; e (v) diferença do efeito da externalidade de rede entre ambos os lados. A segunda conclusão é que as três abordagens para a tarifa de intercâmbio “ótima” não levam necessariamente a resultados distintos. Elas se igualam caso as demandas por serviços de cartão sejam lineares e os emissores e credenciadores repassem custos na mesma proporção (desde que os estabelecimentos não tenham motivo estratégico para aceitar cartões).

Por fim, a literatura também traz algumas abordagens sobre competição entre plataformas. Em um dos modelos (Rochet e Tirole, 2002), a competição entre plataformas é analisada independentemente da configuração da competição intra-plataforma. A conclusão é que, caso os portadores de cartão sejam *singlehoming*, ou seja, usuários de uma única plataforma (única “bandeira”, no caso), a tarifa de intercâmbio não é afetada pela competição interplataformas. Cada plataforma isolada seria vista como um monopólio. Caso os portadores sejam *multihoming*, isto é, usuários de várias plataformas, a tarifa de intercâmbio tende a ser menor.

Outro trabalho (Guthrie e Wright, 2003) conclui que a competição entre plataformas não resulta em tarifas de intercâmbio mais baixas quando o portador

possui apenas um cartão. Mas, no caso de portadores *multihoming*, o equilíbrio da tarifa de intercâmbio ocorre em nível inferior ao do caso de plataforma monopolista.

Hayashi (2005) analisa qual a porcentagem necessária de portadores *multihoming* para que a competição entre plataformas faça com que a tarifa de intercâmbio diminua (e de quanto seria a redução). Neste caso, conclui-se que a competição entre plataformas não necessariamente diminui a tarifa de intercâmbio caso os estabelecimentos aceitem cartão por motivo estratégico. A porcentagem de portadores *multihoming* precisa ser grande o suficiente para que as plataformas fixem tarifas de intercâmbio menores do que a tarifa fixada por uma plataforma monopolista.

Em trabalho posterior, Wang (2006) argumenta que, em plataformas maduras, as externalidades de rede resultantes da adoção e uso de cartões de pagamento já não seriam tão relevantes. Os resultados são os mesmos para plataformas monopolistas e oligopolistas.

2.3.3 Regras e regulamentos específicos no mercado de cartões

A existência de regras e regulamentos específicos definidos pelo proprietário do esquema ou pelo regulador tem por objetivo a construção de padrões que promovam as externalidades positivas e limitem as negativas, e possam ainda trazer benefícios para o conjunto dos usuários finais.

No Brasil, o mercado de cartões até então tem-se apoiado em quatro regras: não-sobrepreço; *honor-all-cards* em sua “versão fraca”; vínculo a instituição financeira (supervisionada pelo Banco Central); e uso de tarifa de intercâmbio do credenciador para o emissor definida unilateralmente pela bandeira.

A regra do não-sobrepreço é uma interpretação sobre o Código de Defesa do Consumidor que proíbe (e pune com multa) a diferenciação entre os preços de bens e serviços em função dos custos dos instrumentos de pagamento utilizados. Embora proibida, a prática do sobrepreço (no caso, para uso de cartões de crédito) é adotada em diversos setores da economia, sendo justificada pelos

comerciantes pela necessidade de compensar as perdas de margens oriundas das elevadas taxas de desconto e da demora no recebimento do pagamento por parte do credenciador (média de 30 dias no Brasil, contra média de 2 dias no exterior).

Reforçando a discussão apresentada no tópico anterior, sobre o grau de competição entre estabelecimentos e as implicações de regras como a do não-sobrepço, Wright (2003) mostra que, caso os estabelecimentos sejam monopolistas, a possibilidade de sobrepço diminui o bem-estar social, pois os estabelecimentos podem cobrar um sobrepço maior sem sofrer redução na demanda, e uma consequência possível seria a sub-utilização de cartões, por conta do aumento excessivo no preço. Caso haja competição de oligopólio na forma de Bertrand, o resultado esperado é a especialização. Estabelecimentos que vendem bens e serviços a preços baixos aceitariam somente dinheiro e estabelecimentos que vendem bens e serviços a preços altos, somente cartão. Nessa configuração, o lucro dos participantes e o bem-estar social seriam independentes tanto da tarifa de intercâmbio quanto da regra de não-sobrepço.

A regra *honor all cards*, em sua “versão fraca”, é a exigência de aceitação de cartões pertencentes a um mesmo esquema de pagamentos, independentemente do banco emissor, da modalidade e do produto. Em sua “versão forte”, a regra obriga lojistas e prestadores de serviços que desejem ser credenciados para a aceitação de um determinado produto de uma bandeira a aceitarem todos os produtos desta, particularmente seus cartões de crédito e de débito, o que caracteriza venda casada (*tying*) entre mercados distintos. No Brasil, a regra é usada apenas na sua “versão fraca”.

Sobre a “versão forte” da regra *honor all cards* (ou seja, incluindo atrelamento entre as funções crédito e débito), Rochet e Tirole (2006) concluem que a presença ou ausência de tal regra pode repercutir de formas distintas sobre a tarifa de intercâmbio e o bem-estar social. Tomam-se como hipóteses iniciais a concorrência nas transações de débito e a substitubilidade entre as funções de crédito e débito, e depois flexibiliza-se essa segunda hipótese. O primeiro resultado é que a ausência da regra *honor all cards* faz com que a tarifa de intercâmbio dos cartões de débito tenda a ser inferior ao valor socialmente ótimo,

enquanto a tarifa de intercâmbio dos cartões de crédito tenda a ser igual ou maior que o valor socialmente ótimo (dependendo do poder de mercado de emissores e de credenciadores). Nesse sentido, a venda casada das duas funções seria uma maneira de se estruturar as tarifas de intercâmbio de modo a beneficiar o bem-estar social. No entanto, apenas flexibilizando-se a hipótese de substitubilidade entre as funções de crédito e débito, chega-se ao resultado de que a presença da regra *honor all cards* poderia levar a uma redução do bem-estar social.

A exigência de que o emissor e, em geral, o credenciador sejam instituições financeiras supervisionadas pelo Banco Central é comum entre proprietários de esquemas de cartões, como forma de gestão de risco. No Brasil, ainda está sendo formalizada a responsabilidade do BC sobre a fiscalização dos participantes no mercado de cartões.

Chama-se, ainda, *net issuer acquirer* a regra que limita a participação no credenciamento aos emissores da plataforma. Uma justificativa para essa regra é evitar que credenciadores selecionem apenas poucos estabelecimentos de alta rentabilidade, sem contribuir de fato para a expansão do sistema (prática conhecida como *cherry picking*).

Por fim, tem-se o uso de tarifa de intercâmbio, que pode ser acordada de forma multilateral, bilateral ou definida pelo proprietário do esquema. No Brasil, ela é paga pelo credenciador ao emissor e definida unilateralmente pelo proprietário do esquema (o que coloca as tarifas praticadas no Brasil no mesmo patamar das praticadas internacionalmente). As variáveis que refletem a diferenciação entre tarifas são segmento de mercado, número de parcelas, forma de captura e tipo de produto.

Quanto ao portador de cartão, ele pode ter cobrada anuidade, taxa de reposição do plástico e de uso de programa de recompensas, mas nenhuma dessas cobranças é comumente observada na prática. Ao contrário, é mais comum ao portador receber do emissor bônus em programas de recompensa.

Especificamente sobre o mercado de cartões de débito no Brasil, uma regulamentação importante definiu que, a partir de dezembro de 2007, o Banco

Central tornou obrigatória a emissão de cartões de débito gratuitos associados às contas correntes pelos bancos comerciais²⁶.

Para finalizar a discussão teórica sobre mercado de dois lados em geral e mercado de cartões em particular, ressaltamos dois outros resultados relevantes da literatura.

O primeiro (Evans e Schmalensee, 2005b) diz que a estrutura e o tamanho de um mercado de dois lados dependem de cinco grupos de fatores: externalidades de rede; ocorrência de congestionamento; presença de economias de escala; possibilidade de diferenciação de produtos (tanto vertical quanto horizontal, favorecendo o *multihoming*); e possibilidade de participar em mais de uma rede ao mesmo tempo (*multihoming*).

O segundo (Katz, 2001) conclui que existem dois mecanismos capazes de internalizar a externalidade de rede existente no mercado de cartões de pagamento: a tarifa de intercâmbio e a possibilidade de sobrepreço. O problema da tarifa de intercâmbio é poder trazer distorções ao mercado de cartões, já que pode afetar as decisões de entrada e utilização por consumidores e estabelecimentos. A depender da estrutura do mercado, a tarifa de intercâmbio pode ser posta em nível que acarrete efeitos adversos sobre a eficiência econômica e o bem-estar social. Assim, a tarifa de intercâmbio seria um mecanismo promotor da internalização das externalidades de rede ao custo de possíveis ineficiências e perdas de bem-estar. Já a possibilidade do sobrepreço neutraliza possíveis efeitos negativos da tarifa de intercâmbio, permitindo a internalização das externalidades de rede.

A supressão da regra do não-sobrepreço permitiria que os estabelecimentos sinalizassem, através de seus preços, os custos de cada instrumento de pagamento, promovendo maior eficiência econômica. Além disso, aumentaria o consumo de consumidores que não possuem cartão, já que eles pagariam um preço menor, corrigiria a quantidade de transações realizadas com cartões em direção a um nível socialmente ótimo e aumentaria o número de estabelecimentos que aceitam cartões de pagamento.²⁷

²⁶ Resolução 3.518 do Conselho Monetário Nacional (art. 2º, inciso I, alínea a), dezembro de 2007.

²⁷ Banco Central do Brasil; Secretaria de Acompanhamento Econômico; Secretaria de Direito Econômico (2009).

2.4 Trabalhos empíricos

As pesquisas empíricas que se propunham a dimensionar as causas do uso dos diversos tipos de instrumento de pagamento – cartões de crédito e débito inclusive – são basicamente trabalhos do Federal Reserve que usam como base de dados²⁸ levantamentos aplicados a indivíduos sobre suas razões para a escolha do instrumento de pagamento²⁹.

Os principais resultados desses estudos reportam que a decisão das famílias em portar e usar cada instrumento de pagamento (aí incluídos cartões de crédito e débito) depende de renda, idade e características demográficas, e as correlações com essas variáveis não parece alterar-se ao longo do tempo. Famílias com maior renda, menor faixa etária e maior escolaridade são mais propensas a adotar novos instrumentos de pagamento. Além disso, ao depararem-se com novas tecnologias de instrumentos de pagamentos, as famílias tendem a adotá-las como uma alternativa a mais ao invés de simplesmente substituir os instrumentos anteriormente usados (*multihoming*). A única relação significativa observada de substitubilidade entre instrumentos de pagamento foi entre cartões de débito e cartões de crédito com uso por conveniência – ou seja, os usuários tendem a optar ou exclusivamente por débito “instantâneo” em suas contas ou exclusivamente por uma postergações de pagamento que não incorram em juros (Klee, 2006)³⁰.

Como no Brasil não há esse tipo de levantamento, foi necessário adotar outra metodologia. A fundamentação teórica apoiou-se em resultados de trabalhos sobre o mercado de dois lados, já discutidos na seção anterior, e em um modelo dinâmico de 2-períodos (intertemporal) e 3-agentes (um consumidor, um vendedor monopolista e um proprietário do esquema) sobre o uso de cartões de crédito (Chakravorti e To, 2006).

²⁸ A base de dados mais comum nos trabalhos é a *Survey of Consumer Finances (SCF)*.

²⁹ Nesses trabalhos, as análises são eminentemente regressões *logit* e *probit* sobre variáveis discretas e completas (*dummies*) e variáveis contínuas.

³⁰ Resultados similares podem ser encontrados em Borzekowski e Kiser (2006), Borzekowski, Kiser e Ahmed (2006), Gerdes e Walton (2002), Hayashi e Klee (2002).

O modelo tem três aspectos restritivos, mas que não comprometem sua generalidade. O primeiro é ater-se somente aos cartões de crédito. A diferença básica para os cartões de débito é que não haverá o período de postergação do pagamento sem cobrança de juros (chamado *grace period*). A segunda restrição é limitar-se ao desenho de um esquema de 3 partes. E a terceira é a premissa do vendedor monopolista. Tal premissa não enfraquece o modelo, pois pode-se considerar a existência de uma infinidade de vendedores (portanto, sem poder de barganha sobre a taxa de desconto), cada qual atuando como monopolista de um produto bastante diferenciado. As conclusões do modelo são igualmente pertinentes para o mercado de cartões em geral.

O modelo elabora um jogo bayesiano dinâmico em que:

- No período $t = 0$, o proprietário do esquema tem que decidir qual renda mínima exigir do consumidor para que haja oferta de crédito e qual taxa de desconto cobrar do comerciante. No mesmo período, o comerciante tem que decidir entre aceitar ou não cartões de crédito.
- No período $t = 1$, a natureza determina a renda e a demanda de cada consumidor e ele tem que tomar duas decisões: comprar ou não comprar e usar ou não cartão de crédito.

O modelo chega a três conclusões:

- Se os custos do proprietário do esquema não forem muito elevados e as margens do vendedor forem suficientemente altas, o proprietário do esquema ampliará o crédito para usuários qualificados, o comerciante aceitará cartões e o consumidor correrá risco de inadimplência.
- A capacidade do proprietário do esquema em cobrar taxas de desconto elevadas depende da quantidade de consumidores incorporada quando se aceitam cartões.
- Cada vendedor depara-se com um dilema do prisioneiro, em que cada um escolhe individualmente aceitar cartões, embora os lucros em dois períodos de todos eles sejam reduzidos por isso.

2.5 Considerações gerais sobre o mercado de cartões

Esse capítulo apresentou uma evolução histórica dos instrumentos eletrônicos de pagamento, em seus aspectos de inovação econômica e tecnológica, com a finalidade de esclarecer tanto o cenário atual quanto as expectativas para o futuro próximo nessa área.

Além disso, discutiu ainda a importância do estudo e observância pelos bancos centrais do uso de tais instrumentos, dentro do contexto do sistema de pagamentos. Sendo uma forma de “uso da moeda”, pode acarretar efeitos econômicos relevantes, que serão discutidos em capítulo posterior desse trabalho.

Por fim, para dar suporte a modelagens, abordou a dinâmica dos mercados dos mais populares dentre estes instrumentos (cartões de crédito e débito), com seus conceitos e variáveis fundamentais.

3 EXERCÍCIOS ECONOMETRÍCOS SOBRE OS MERCADOS DE CARTÕES DE CRÉDITO E DÉBITO NO BRASIL

3.1 Uma modelagem sobre o lado comerciante/credenciador

Idealmente, a parametrização de um mercado de dois lados deve-se dar via parâmetros dos agentes que estão nas duas pontas da estrutura desse mercado. No entanto, este trabalho concentra-se, por dois motivos, particularmente no lado comerciante/credenciador, em detrimento do lado emissor/portador.

O primeiro motivo é a premissa aqui adotada de que os consumidores tendam a uma adoção natural desses instrumentos de pagamentos, seja por conta de serem o lado “subsidiado” na estrutura de mercado de dois lados (via, por exemplo, recebimento de bônus como milhas aéreas pelo uso dos cartões ou via parcelamentos “sem juros”), seja por conta das comodidades oferecidas (redução do “custo sola de sapato” e maior agilidade e exatidão nos pagamentos), seja pela tendência natural de redução ao longo do tempo na relutância cultural de adoção a tais instrumentos. Assim, adotada tal premissa, o “tamanho” desse mercado seria determinado basicamente pelo lado comerciante/credenciador.

A segunda razão a orientar essa abordagem de foco nesse lado do mercado é a dificuldade de conciliação de bases de dados. A grande maioria dos dados utilizados neste trabalho são provenientes de levantamento feito trimestralmente pelo Banco Central do Brasil entre instituições financeiras credenciadoras e emissoras de cartões de crédito e débito. Ou seja, pode-se tratar essa base de dados na forma de um painel em que a dimensão *cross-section* seja cada instituição financeira e que contenha todas as variáveis relevantes para estimação do lado comerciante/credenciador. Por sua vez, o lado emissor/portador fica associado a aspectos demográficos (nível de renda, idade, escolaridade). Como não se tem aqui um painel trimestral segmentado por instituição financeira com características demográficas de seus clientes, fica impraticável conciliar bases de

dados voltadas para os dois lados e tentar, por exemplo, estimar um modelo de equações simultâneas.

Com base no modelo de Chakravorti e To (2006), as decisões do comerciante e da instituição financeira quanto à aceitação na participação do mercado são basicamente reativas a preço. Assim, o primeiro esboço da curva consiste em observar o comportamento do volume financeiro movimentado em cartões de crédito e de débito (em equações distintas) em relação aos “preços” praticados (taxas de desconto e de intercâmbio). A variável empregada para tentar captar a adesão ao mercado pelo lado emissor/portador é o número de estabelecimentos credenciados. E, no caso particular do mercado de cartões de débito, considera-se ainda a particularidade de que, por uma convenção de mercado, a tarifa de intercâmbio é fixada em 50% da taxa de desconto. Assim, as equações do modelo são:

$$\log(valor) = \alpha_0 + \alpha_1 \log(estab) + \alpha_2(txitc) + \alpha_3(txdesc) + \varepsilon \quad (\text{eq. 3.1})$$

para cartões de crédito e

$$\log(valor) = \beta_0 + \beta_1 \log(estab) + \beta_2(txitc) + \varepsilon \quad (\text{eq. 3.2})$$

para cartões de débito, em que *valor* é o valor movimentado em transações com cartões; *estab* é o número de estabelecimentos credenciados; *txitc* a tarifa de intercâmbio e *txdesc* a taxa de desconto.

A escolha da variável dependente como sendo o valor financeiro movimentado e não a quantidade de transações deve-se ao interesse do trabalho em focar-se no *fluxo de moeda* que esses instrumentos envolvem e não no nível de utilização dos instrumentos propriamente.

Conforme será mostrado adiante, os dados utilizados foram estruturados na forma de um painel balanceado, com um número maior de períodos³¹ do que de

³¹ Totais de 22 períodos no caso de cartões de crédito (2003:1 a 2008:2), e de 27 períodos no caso de cartões de débito (2001:4 a 2008:2).

indivíduos³². Assim, antes da regressão propriamente dita, são feitas avaliações de tendência e sazonalidade nas séries, bem como testes de autocorrelação e raízes unitárias. A questão da sazonalidade pode ser tratada com simples variáveis binárias para os trimestres em que o fenômeno se manifestar. Já para o problema da tendência, é necessário avaliar se ela é melhor modelada como determinística ou estocástica. Para tendências determinísticas, o melhor recurso é a remoção de tendência por meio do uso de uma variável de tendência na regressão. Caso existam raízes unitárias (tendências estocásticas), o procedimento adotado é observar os resíduos resultantes de regressões em primeira diferença, em painel por efeitos fixos e em painel por efeitos aleatórios.

As seções seguintes explicam as metodologias de análise de raízes unitárias e estimações em painel por efeitos fixos, efeitos aleatórios e primeira diferença.

3.2 Metodologia: raízes unitárias

Para séries que apresentem um nítido comportamento crescente (ou decrescente) ao longo do tempo, duas possibilidades muito comuns de modelagem são a forma de *trend series* (TS, como na eq. 3.3 a seguir) ou a forma de *difference series* (DS, como na eq. 3.4 a seguir)³³.

$$y_t = a + bt + e_t \quad (\text{eq. 3.3})$$

$$y_t = a + y_{t-1} + e_t \quad (\text{eq. 3.4})$$

Na eq. 3.4, o coeficiente unitário na variável defasada é o que se chama “raiz unitária”. Modelos com mais de uma defasagem também podem apresentar raízes unitárias se os coeficientes das defasagens obedecerem a certas

³² Total de 6 combinações credenciador/bandeira no caso de cartões de crédito e de 3 combinações credenciador/bandeira no caso de cartões de débito.

³³ *Trend series* podem ser “*detrended*” (regressão com variável de tendência entre os controles) mas não devem ser diferenciadas, sob o risco de o desenvolvimento do modelo levar a conclusões completamente errôneas. *Difference series* podem ser diferenciadas mas não devem ser “*detrended*”, também sob o risco de o desenvolvimento do modelo levar a conclusões completamente errôneas. Ver Hamilton (1994).

condições³⁴. A presença de raízes unitárias em uma série temporal tem implicações teóricas críticas, especialmente para modelos preditivos. A raiz unitária significa que um choque aleatório sobre a variável em questão, ocorrido em qualquer momento da série histórica, tem seu efeito carregado por todo o restante da série, ou seja, não é atenuado com o passar do tempo³⁵. Isso caracteriza *não-estacionariedade* da série, ou seja, uma ausência de tendência de que ela retorne para um caminho determinístico projetado *a priori* para o longo prazo. Outra consequência – a mais crucial para modelos com propósitos preditivos – é que, na presença de raízes unitárias, a variância das previsões para os períodos futuros fica proporcional ao “avanço” do período previsto em relação ao atual (Hamilton, 1994). Ou seja, previsões de “longo prazo” passam a ter variância infinita, tornando-se completamente ineficazes. Em modelos com múltiplas variáveis, a presença de raiz unitária pode ainda resultar no problema chamado “regressão espúria”, em que o resultado do modelo sugere relação de causalidade entre variáveis que não tenham de fato essa relação, levando a conclusões equivocadas. Nesses casos, a aparente causalidade é simples consequência de as variáveis “crescerem juntas”.

Para testar a presença de raiz unitária em uma série, o teste mais popular é o *augmented Dickey-Fuller* (ADF)³⁶, que escreve o modelo na forma diferencial generalizada da eq. 3.3 e testa a hipótese nula de existência de raiz unitária ($u=1$).

$$\Delta y_t = a + (u - 1)y_{t-1} + bt + e_t \quad (\text{eq. 3.5})$$

Para o uso das estatísticas dos testes ADF, dois aspectos fundamentais são a especificação do modelo quanto à presença ou não de intercepto e tendência e quanto ao número de defasagens.

³⁴ Como regra geral, fala-se na condição matemática de “raízes fora do círculo unitário”. No entanto, a caracterização dessa “regra geral” difere segundo o número de defasagens no modelo. Os detalhes não são escopo deste trabalho e podem ser encontrados em Hamilton (1994).

³⁵ A única possibilidade de o efeito desse choque “desaparecer” nos períodos seguintes é ele ser contrabalançado por outros choques que tenham, somados, a mesma dimensão mas o sinal oposto.

³⁶ Dickey e Fuller (1979; 1981) elaboraram 4 tabelas de estatísticas do teste (com e sem intercepto e com e sem tendência) a partir de simulações de um processo de Monte Carlo.

Outros testes já foram formulados, mas o ADF segue sendo o mais amplamente empregado nos trabalhos. No entanto, cumpre ressaltar que testes de raiz unitária conhecidos até então ainda são incapazes de caracterizar perfeitamente se uma série é uma DS (ou seja, possui raiz unitária, é não-estacionária) ou se é uma TS com coeficiente próximo de 1.

Como uma tendência cada vez mais dominante nos trabalhos empíricos tem sido o emprego de dados em painel, foram formulados recentemente testes de raízes unitárias para esse tipo de dados. Os testes aplicáveis³⁷ a esse trabalho são o LLC (Levin, Lin e Chu, 2002) e o IPS (Im, Pesaran e Shin, 2003).

O teste LLC trata o painel como uma *cross-section* empilhada, para testar a hipótese nula de que cada série temporal individualmente contenha uma raiz unitária contra a hipótese alternativa de que cada série temporal não contenha raiz unitária. Ou seja, o procedimento pressupõe comportamento idêntico de cada série temporal quanto à presença de raiz unitária (hipótese bastante restritiva) e independência entre os indivíduos da amostra. Apesar de tal limitação, esse teste tem maior poder do que testes de raiz unitária separados para cada série temporal. Isso porque, além de sua estatística de teste convergir para uma distribuição normal à medida que as dimensões de *cross-section* ou de séries temporais aumentam, simulações de Monte Carlo mostram que o mesmo resultado vale para amostras relativamente pequenas.

O teste IPS faz uma média das estatísticas Dickey-Fuller entre os indivíduos da amostra. Nesse teste, a hipótese nula é que todas as séries (a de cada indivíduo) tenham raiz unitária. A hipótese alternativa, naturalmente, é que ao menos uma das séries (de ao menos um dos indivíduos) não tenha raiz unitária.

3.3 Metodologia: estimações em painéis por efeitos fixos, por efeitos aleatórios e primeira diferença

Considere-se a seguinte especificação geral para um modelo de dados em painel:

³⁷ O teste de Breitung não se aplica porque se baseia na convergência assintótica em painéis com grande número de indivíduos e poucos períodos, formato oposto ao painel tratado neste trabalho. O teste PP-Fisher também não convém porque não há quebra estrutural nas séries.

$$y_{it} = x_{it}\beta_{it} + u_{it}; \text{ sendo } i = 1, \dots, N \text{ e } t = 1, \dots, T \quad (\text{eq. 3.6})$$

em que β_{it} é o vetor ($k \times 1$) de parâmetros desconhecidos relativos ao indivíduo i no momento t e x_{it} é a matriz ($k \times 1$) de variáveis explicativas, cuja primeira coluna, no caso do modelo ter termo independente, será integralmente constituída por 1's. O erro u_{it} é chamado de erro idiossincrático ou erro de variação temporal, porque ele representa fatores não observados que mudam ao longo do tempo e afetam y_{it} . Num modelo estático, assume-se que as variáveis explicativas são independentes dos termos de perturbação. No que se refere à heterogeneidade, pode-se assumir que ela reside nos coeficientes de regressão (que podem variar no tempo ou de indivíduo para indivíduo) ou na estrutura dos termos de perturbação. A escolha de uma especificação de validade universal é impossível, restando escolher uma que, ante os dados analisados, melhor se enquadre. Concebem-se então quatro tipos de especificações básicas de modelos: MQO empilhado, modelos em primeira diferença, modelos de efeito fixo e modelos de efeito aleatório.

O modelo de MQO empilhado (*pooled OLS*) parte da suposição de que os parâmetros comportamentais ao longo do período e do espaço são rígidos. Ou seja, assume-se que todas as observações das *cross-sections* têm o mesmo comportamento ao longo de todos os anos. Sendo assim, o MQO empilhado praticamente desconsidera a heterogeneidade existente entre as variáveis, o que é uma hipótese bastante restritiva. Ressalta-se o fato de que os parâmetros estimados por este procedimento podem ou não ser diferentes dos parâmetros estimados fazendo-se regressões individuais para cada *cross-section*, uma vez que tal especificação considera que a variância do termo de erro não varia entre os indivíduos.

Quanto aos modelos em primeira diferença, pode-se dizer inicialmente que sua principal vantagem é a eliminação de efeitos não observados da análise. A diferenciação faz com que esses efeitos que não variam ao longo da série sejam suprimidos, mantendo apenas os efeitos das variáveis explicativas. A especificação básica de modelos em primeira diferença é a seguinte:

$$\Delta y_{it} = \Delta x_{it} \beta_{it} + \Delta u_{it}; \text{ sendo } i = 1, \dots, N \text{ e } t = 1, \dots, T \quad (\text{eq. 3.7})$$

no qual Δ representa a diferenciação entre as variáveis no tempo. Na verdade, tem-se apenas uma equação em que cada variável é diferenciada ao longo do tempo. Para que o mesmo seja válido, é preciso que todas as hipóteses fundamentais sejam satisfeitas. A mais importante delas é que Δu_{it} seja não correlacionado com Δx_{it} . Essa hipótese será naturalmente satisfeita se u_{it} for não correlacionado com x_{it} em nenhum período, o que se chama exogeneidade estrita das variáveis explicativas.

Embora a diferenciação de dados em painel seja eficaz em controlar efeitos não observados, ela também tem suas desvantagens. A primeira é que tal método acarreta perda de i observações (referentes ao período inicial da amostra), o que pode ser muito problemático se a série temporal disponível for curta. Essas observações perdidas reduzem os graus de liberdade e, conseqüentemente, a significância dos parâmetros. Além disso, a diferenciação utilizada para eliminar o efeito não observado pode reduzir bastante a variação nas variáveis explicativas, e pequenas variações em Δx_{it} podem levar a grandes erros-padrão MQO. Pode-se contornar esse problema empregando-se grandes cortes transversais, mas isso depende da base de dados disponível.

A primeira diferenciação não é a única maneira de eliminar o efeito fixo não observado de modelos de dados em painel. Outro método que pode funcionar melhor, sob certas hipóteses, é o método de transformação por efeitos fixos. Basicamente, o modelo consiste na estimação da seguinte equação:

$$y_{it} - \bar{y}_i = \beta_{it} (x_{it} - \bar{x}_{it}) + u_{it} - \bar{u}_i; \text{ sendo } i = 1, \dots, N \text{ e } t = 1, \dots, T \quad (\text{eq. 3.8})$$

em que $\bar{y}_i, \bar{x}_i, \bar{u}_i$ são as médias das variáveis y , x e u , respectivamente. O importante na equação acima é que o efeito fixo não observado desaparece, justamente pela introdução da subtração da média de cada variável. Isso sugere que devemos estimar agora o modelo por MQO empilhado. E um modelo MQO empilhado baseado em variáveis temporais reduzidas é o chamado estimador de efeitos fixos.

A forma com que o modelo de efeitos fixos é especificado permite que diferentes interceptos (nesse caso, captados pela diferença $u_{it} - \bar{u}_i$) possam captar todas as diferenças entre as unidades de *cross-sections*, assumindo-se as inclinações como constantes para todas elas. Todas as diferenças nas variáveis entre os indivíduos e ao longo do tempo são captadas pelo termo de intercepto. Não obstante, existe ainda a possibilidade de serem estimadas, para cada *cross-section*, diferentes inclinações. Para isso, bastaria inserir uma *dummy* para cada ano da observação, técnica entretanto não recomendada para estimações com muitos períodos, devido ao uso de muitas *dummies* e consequente perda de muitos graus de liberdade.

Sob a hipótese de exogeneidade estrita das variáveis explicativas, o estimador de efeitos fixos é não-viesado. É importante notar também que o estimador de efeitos fixos leva em consideração uma correlação arbitrária entre os efeitos não observados e as variáveis explicativas em qualquer período de tempo, como na primeira diferenciação. Assim, qualquer variável explicativa que seja constante ao longo do tempo para todo i será removida, e por esse motivo não vale a pena inserir tais variáveis na estimação.

A especificação por efeitos aleatórios pressupõe que o comportamento específico dos indivíduos e particularidades de períodos de tempo são aspectos desconhecidos, não podendo ser observados nem medidos. Assim, em amostras longitudinais de grande dimensão, pode-se sempre representar estes efeitos individuais ou temporais específicos sob a forma de uma variável aleatória normal. Dentre as vantagens dos modelos de efeitos aleatórios e para o caso particular dos modelos de componentes de erro, destacam-se: a capacidade para trabalhar com bases de dados de qualquer dimensão; o fato de a inferência estatística aplicável ser uma mera derivação dos testes de hipóteses usuais; e os fatos de serem a estrutura em painel melhor detalhada e a interpretação dos resultados de estimação mais intuitiva.

Ao contrário do modelo de efeitos fixos, a heterogeneidade não é induzida através do termo independente (ou seja, pelo valor esperado $E[y_{it}]$), mas sim pela variância da variável endógena. Uma especificação geral pode ser:

$$y_{it} = x_{it}\beta_{it} + a_i + u_{it}; \text{ sendo } i = 1, \dots, N \text{ e } t = 1, \dots, T \quad (\text{eq. 3.9})$$

que segue as mesmas especificações do modelo geral da eq. 3.6, exceto pela inclusão do efeito não observado, o termo a_i . Assume-se, sem perda de generalidade, que o efeito tem média zero. Quando se emprega o efeito fixo, supomos que o efeito não observado está correlacionado com uma ou mais variáveis da matriz x_{it} . Mas, supondo que a_i seja não-correlacionado a cada variável explicativa em todos os períodos de tempo, o resultado será estimadores ineficientes. Assim, para que a equação acima seja um modelo de efeitos aleatórios, é preciso garantir que:

$$\text{Cov}(x_{ij}, a_i) = 0; \text{ com } t = 1, \dots, T \text{ e } j = 1, \dots, K \quad (\text{eq. 3.10})$$

As hipóteses de efeitos aleatórios incluem todas as hipóteses básicas de efeitos fixos mais o requisito adicional de independência entre a_i e todas as variáveis explicativas em todos os períodos de tempo. Sendo assim, os coeficientes poderiam ser estimados por meio de um único corte transversal, ou até mesmo pelo método de MQO empilhado. Porém, isso ignora uma característica fundamental do modelo. Se for definido o termo de erro de composição como $v_{it} = a_i + u_{it}$, a especificação geral pode ser reescrita como:

$$y_{it} = x_{it}\beta_{it} + v_{it}; \text{ sendo } i = 1, \dots, N \text{ e } t = 1, \dots, T \quad (\text{eq. 3.11})$$

Como se sabe que a_i é o erro em cada período de tempo, os v_{it} são serialmente correlacionados ao longo do tempo. E essa correlação levaria a estimação MQO a coeficientes incorretos. Então, para se resolver o problema de correlação serial, faz-se a estimação por MQG. A derivação da transformação MQG que elimina a correlação serial nos erros é feita pelo fator de ponderação (Wooldridge, 2002; 2006):

$$\lambda = 1 - \left[\frac{\sigma_u^2}{\left(\sigma_u^2 + T\sigma_a^2 \right)^{1/2}} \right], \quad (\text{eq. 3.12})$$

que é um número entre zero e um. A equação transformada resultará em

$$y_{it} - \lambda \bar{y}_i = \beta_{it} (x_{it} - \lambda \bar{x}_{it}) + (v_{it} - \lambda \bar{v}_i) \quad (\text{eq. 3.13})$$

Como λ é função de variâncias desconhecidas, a estimação do modelo é feita, na verdade, a partir de um estimador de efeitos aleatórios $\hat{\lambda}$. Sob as hipóteses clássicas de regressão linear, pode-se considerar os estimadores de efeitos aleatórios consistentes e distribuídos normalmente e assintoticamente.

Em suma, se o estimador de efeitos fixos subtrai as médias temporais de cada variável, o estimador de efeitos aleatórios subtrai uma fração desta média temporal, na qual a fração dependerá das variâncias de a_i e u_i e do número de períodos de tempo T . Pode-se, inclusive, entender os modelos de MQO empilhado e de efeito fixo como particularizações do modelo de efeitos aleatórios, em que se tem $\lambda = 0$ no primeiro caso e $\lambda = 1$ no segundo. Portanto, se as estimativas por efeitos aleatórios estiverem próximas de zero, o resultado será próximo do de uma estimação por MQO empilhado. Esse é o caso quando o efeito não observado a_i é quase irrelevante. É mais comum σ_a^2 ser grande em relação a σ_u^2 , o que torna $\hat{\lambda}$ próximo da unidade. Conforme T fica maior, $\hat{\lambda}$ tende para um, o que torna próximas as estimativas por efeito fixo e efeito aleatório.

3.4 Metodologia: modelagem VAR

Outro exercício econométrico feito aqui é a aplicação de uma modelagem VAR aos mercados de cartões de crédito e débito.

A tentativa de prever eventos futuros tendo por base acontecimentos passados é algo recorrente no debate econômico. A evolução desse debate conduziu a quatro tipos de modelos bastante comuns que têm esse propósito: modelos de equação única; modelos de equações simultâneas; modelos Box-Jenkins ou ARIMA (auto-regressivos com média móvel integrados); e modelos VAR (vetores auto-regressivos).

O primeiro tipo estrutura-se como uma variável y_t sendo explicada, a partir de fundamentos teóricos, por um conjunto de variáveis exógenas X_t , simultâneas à variável explicada. Foi o caminho adotado no exercício econométrico anterior.

A abordagem de equações simultâneas, por sua vez, embute uma endogeneidade no modelo, colocando y_t como um vetor de variáveis endógenas que aparecem nas equações explicativas umas das outras.

Nos anos de 1970, foram propostos os modelos ARIMA, que propunham padrões quantitativos que determinassem os valores futuros das variáveis como combinações lineares de seus próprios valores passados e de termos aleatórios, abrindo mão da preocupação com fundamentos teóricos que justificassem correlações entre variáveis, para adotarem a noção de as variáveis “explicarem-se a si próprias”, ou seja, de seus valores atuais e futuros serem, em boa medida, reflexos de seus valores passados. Embora criticados por sua fragilidade teórica, tais modelos se consagraram por bons resultados empíricos e permanecem como motivo de alguma celeuma no debate econômico.

Por fim, os modelos VAR combinam, de certa forma, o pressuposto de endogeneidade dos modelos de equações com a noção “auto-explicativa” dos modelos ARIMA. Neles, o vetor y_t é estimado como

$$y_t = Y_{t-j} \beta_j \quad (\text{eq. 3.14})$$

em que Y_{t-j} refere-se à matriz cujas colunas são as defasagens de 1 a j do vetor y_t .

3.5 Base de dados, análise descritiva e evolução histórica

A base de dados utilizada no trabalho provém do Banco Central do Brasil e é obtida por meio de formulários enviados às instituições financeiras credenciadoras e emissoras de cartões de crédito e de débito, que são obrigadas a respondê-los trimestralmente desde o primeiro trimestre de 2003. Algumas IFs já respondiam ao levantamento desde 1999, mas ele passou a ser respondido por todas somente a partir do primeiro trimestre de 2003 para o caso de cartões de crédito e

a partir do quarto trimestre de 2001 para o caso de cartões de débito. Por isso, as análises apresentadas aqui referem-se somente a tais períodos. Os dados referentes a outros períodos foram excluídos das análises.

A base referente aos emissores foi desconsiderada porque o modelo parte da premissa de que o preço que equilibra o mercado de dois lados é a tarifa de intercâmbio (Schmalensee, 2002), definida na relação entre credenciador e proprietário do esquema. A base de credenciadores conta com 6 participantes para cartões de crédito e 3 para cartões de débito. Dessa forma, as análises apresentadas referem-se a painéis balanceados 6 x 22 para cartões de crédito e 3 x 27 para cartões de débito.

Os primeiros tratamentos dados à base de dados foram uma análise descritiva das variáveis (Tabelas 3.1 e 3.2) e a observação das evoluções históricas tanto da quantidade de transações em cartões quanto do volume financeiro transacionado com eles (Gráficos 3.1 e 3.2), a fim de se obter uma primeira impressão sobre tendências e sazonalidades presentes nas séries (consolidadas por período). Comparativamente, confrontam-se ainda as evoluções históricas da quantidade de transações e volume financeiro movimentado em cartões com o agregado monetário M1 (Gráficos 3.3 e 3.4)³⁸.

TABELA 3.1 – Análise descritiva dos dados de cartões de crédito

Variável	Obs	Média	Desvio Padrão	Mínimo	Máximo
Valor movimentado (R\$ milhões)	132	5,520.00	7,010.00	0.16	27,200.00
Quantidade de transações (mil)	132	63,000	81,000	2	290,000
Estabelecimentos credenciados (mil)	132	316	231	1	914
Taxa de desconto (%)	132	2.73	0.32	1.82	3.09
Taxa de intercâmbio (%)	132	0.99	0.72	0.00	1.87

³⁸ Dados de M1 obtidos no site do Banco Central do Brasil.

TABELA 3.2 – Análise descritiva dos dados de cartões de débito

Variável	Obs	Média	Desvio Padrão	Mínimo	Máximo
Valor movimentado (R\$ milhões)	81	4,250.00	4,000.00	49.10	14,600.00
Quantidade de transações (mil)	81	90,900	81,500	901	287,000
Estabelecimentos credenciados (mil)	81	316	259	15	884
Taxa de desconto (%)	81	1.31	0.37	0.59	1.63
Taxa de intercâmbio (%)	81	0.70	0.18	0.24	0.93

Observa-se, pela análise descritiva (Tabelas 3.1 e 3.2), uma grande heterogeneidade nas variáveis³⁹, basicamente decorrente da diferença de porte entre credenciadores das grandes bandeiras (Visa e MasterCard) e das bandeiras menores. Os únicos casos em que a heterogeneidade não é grande são a variável taxa de desconto para os dois tipos de cartões. Essa maior “homogeneidade” nos valores de taxa de desconto levanta a suspeita, já fundamentada pela teoria de que essas taxas não funcionem como “preços” que equilibrem tais mercados. Ou seja, o volume de uso de cartões (tanto medido em volume financeiro transacionado quanto em quantidade de transações) parece, à primeira vista, mais “descolado” das taxas de desconto. O mesmo não acontece com a tarifa de intercâmbio para cartões de crédito, que, como discutido, funciona como o “preço” que equilibra esse mercado.

Seguindo-se à análise descritiva, foram observadas as correlações entre as variáveis, mostradas nas Tabelas 3.3 e 3.4.

³⁹ Observada a partir das razões desvio-padrão/média e (máximo – mínimo)/média.

TABELA 3.3 – Correlações entre variáveis para cartões de crédito

	Valor movimentado (R\$ milhões)	Quantidade de transações (mil)	Estabelec. credenciados (mil)	Taxa de desconto (%)	Taxa de intercâmbio (%)
Valor movimentado (R\$ milhões)	1	0.99325	0.880044	0.509538	0.387274
Quantidade de transações (mil)	0.99325	1	0.888435	0.546205	0.373853
Estabelec. credenciados (mil)	0.880044	0.888435	1	0.763066	0.583159
Taxa de desconto (%)	0.509538	0.546205	0.763066	1	0.744334
Taxa de intercâmbio (%)	0.387274	0.373853	0.583159	0.744334	1

TABELA 3.4 – Correlações entre variáveis para cartões de débito

	Valor movimentado (R\$ milhões)	Quantidade de transações (mil)	Estabelec. credenciados (mil)	Taxa de desconto (%)	Taxa de intercâmbio (%)
Valor movimentado (R\$ milhões)	1	0.997641	0.965875	0.589767	0.682789
Quantidade de transações (mil)	0.997641	1	0.975663	0.620494	0.718807
Estabelec. credenciados (mil)	0.965875	0.975663	1	0.658346	0.784315
Taxa de desconto (%)	0.589767	0.620494	0.658346	1	0.921404
Taxa de intercâmbio (%)	0.682789	0.718807	0.784315	0.921404	1

Os resultados confirmam correlações naturalmente esperadas como altas entre valor movimentado, quantidade de transações e número de estabelecimentos credenciados. Um resultado interessante é que a correlação entre estabelecimentos credenciados e quantidade de transações (ou volume movimentado) é menor no mercado de cartões de crédito do que no de débito (mesmo sendo elevada em ambos os casos). Isso sugere que estabelecimentos com maior escala e conseqüente maior poder de barganha possam negociar um menor repasse da tarifa de intercâmbio no mercado de crédito – em virtude do custo já contraído por conta da maior demora no recebimento de pagamentos em cartões de crédito – ou até mesmo rejeitem participar desse mercado.

Já no que concerne à evolução histórica (Gráficos 3.1 e 3.2), é perceptível a tendência de elevação na quantidade de transações e no volume financeiro transacionado tanto com cartões de crédito quanto de débito. Questões que emergem daí são se essa tendência de crescimento converge para algum nível assintótico e que nível seria esse.

Observando-se o estreitamento do espaço entre as curvas de quantidade (linha cheia) e volume financeiro (linha tracejada) em ambos os Gráficos 3.1 e 3.2, conclui-se também que o valor médio por transação tem aumentado ao longo dos trimestres analisados.

Outra conclusão imediata da observação dos gráficos (e amplamente conhecida da atividade econômica em geral) é a sazonalidade presente no 4º trimestre de cada ano.

GRÁFICO 3.1 – Evolução da quantidade de transações e do volume transacionado em cartões de crédito no Brasil entre o primeiro trimestre de 2003 e o segundo trimestre de 2008

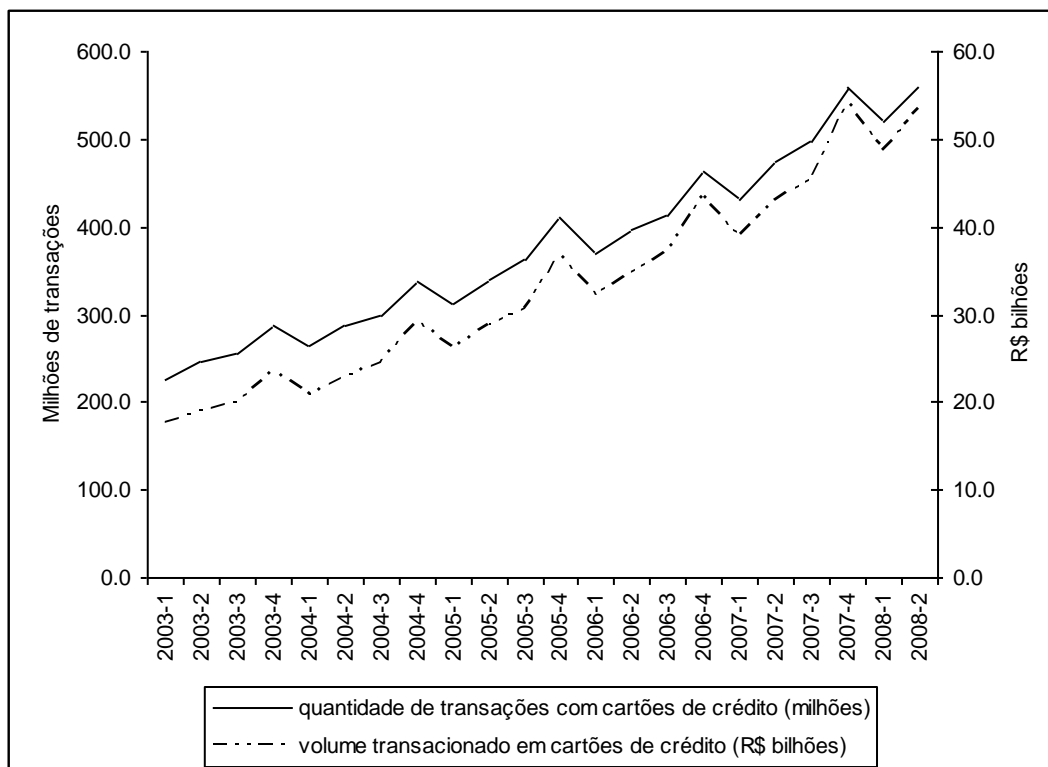
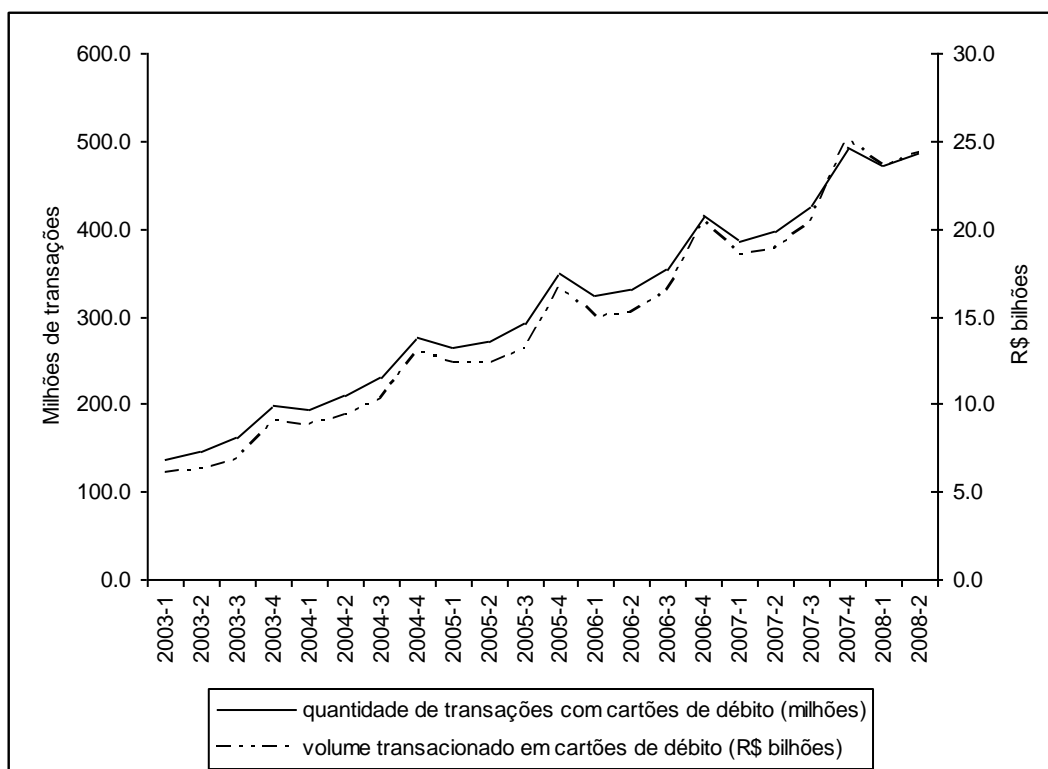


GRÁFICO 3.2 – Evolução da quantidade de transações e do volume transacionado em cartões de débito no Brasil entre o primeiro trimestre de 2003 e o segundo trimestre de 2008



Em seguida, foram traçados gráficos comparativos com o agregado monetário M1. Neles, as inclinações das curvas de cartões, tanto de crédito quanto de débito, com taxas médias de crescimento anual respectivamente de 23% e 30%, são perceptivelmente maiores que a do M1, com taxa média de crescimento anual⁴⁰ de 17%, o que sugere que a expansão do volume transacionado por tais instrumentos seja mais do que uma simples “taxa natural” de expansão monetária medida pelo crescimento do M1.

Isso sugere que o ritmo de crescimento de demanda por moeda esteja mais intenso sobre os instrumentos eletrônicos do que sobre o papel-moeda. Como os instrumentos eletrônicos permitem às instituições financeiras que os operam manter suas reservas (caso dos emissores de cartões) disponíveis para operações de crédito (ao menos operações intradia ou em D+1, para o caso da parcela de reservas que será destinada a liquidação de operações com cartões

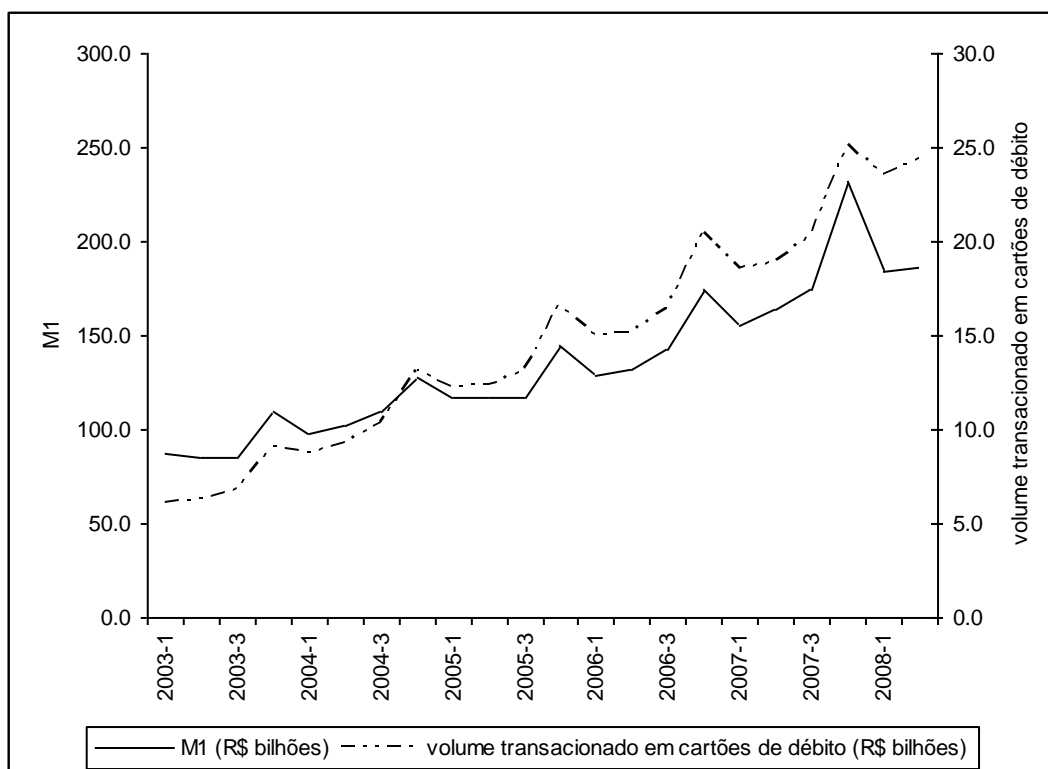
⁴⁰ Médias geométricas das taxas de crescimento do valor de um trimestre em relação ao trimestre equivalente do ano anterior.

de débito), esse fenômeno pode ter reflexos no mercado de moeda e implicações na política monetária. Tal discussão será aprofundada no capítulo 3.

GRÁFICO 3.3 – Evolução do volume transacionado em cartões de crédito no Brasil e do agregado monetário M1 entre o primeiro trimestre de 2003 e o segundo trimestre de 2008



GRÁFICO 3.4 – Evolução do volume transacionado em cartões de débito no Brasil e do agregado monetário M1 entre o primeiro trimestre de 2003 e o segundo trimestre de 2008



3.6 Raízes unitárias e efeitos fixos e aleatórios

Caracterizadas tendência e sazonalidade das variáveis, a etapa seguinte são testes de raízes unitárias.

Depois de já observadas visualmente tendências crescentes nos gráficos de diversas variáveis, pode-se observar que seus correlogramas (Tabelas 3.5 e 3.6) também revelam significativa correlação entre os períodos.

TABELA 3.5 – Correlogramas das variáveis para o modelo de cartões de crédito

lag	Valor (log)		Quantidade (log)		Estabelecs (log)		Taxa de intercâmbio		Taxa de desconto	
	ACF	PACF	ACF	PACF	ACF	PACF	ACF	PACF	ACF	PACF
1	0.807	0.807	0.799	0.799	0.776	0.776	0.918	0.918	0.845	0.845
2	0.647	-0.01	0.633	-0.02	0.62	0.047	0.836	-0.04	0.709	-0.02
3	0.515	-0.01	0.495	-0.02	0.521	0.065	0.755	-0.04	0.588	-0.02
4	0.408	-0.01	0.382	-0.01	0.507	0.186	0.674	-0.05	0.489	0.003
5	0.336	0.039	0.306	0.034	0.493	0.068	0.595	-0.04	0.393	-0.05
6	0.284	0.021	0.251	0.013	0.468	0.041	0.524	0.008	0.293	-0.08
7	0.234	-0.02	0.198	-0.02	0.429	0.014	0.454	-0.05	0.223	0.034
8	0.192	0	0.154	-0.01	0.391	0.007	0.386	-0.03	0.168	0
9	0.157	0	0.118	0	0.355	0	0.32	-0.04	0.102	-0.08
10	0.124	-0.01	0.084	-0.02	0.321	-0.02	0.255	-0.04	0.046	-0.02
11	0.091	-0.02	0.051	-0.02	0.288	-0.02	0.193	-0.03	-0.01	-0.05
12	0.078	0.034	0.046	0.051	0.203	-0.16	0.173	0.214	0	0.145

TABELA 3.6 – Correlogramas das variáveis para o modelo de cartões de débito

lag	Valor (log)		Quantidade (log)		Estabelecs (log)		Taxa de intercâmbio		Taxa de desconto	
	ACF	PACF	ACF	PACF	ACF	PACF	ACF	PACF	ACF	PACF
1	0.76	0.76	0.768	0.768	0.852	0.852	0.866	0.866	0.913	0.913
2	0.578	0.002	0.591	0	0.718	-0.03	0.754	0.015	0.829	-0.03
3	0.433	-0.02	0.458	0.011	0.6	-0.02	0.678	0.085	0.756	0.026
4	0.325	0	0.352	-0.01	0.489	-0.04	0.614	0.026	0.671	-0.12
5	0.233	-0.02	0.262	-0.02	0.391	-0.02	0.551	-0.01	0.589	-0.03
6	0.148	-0.05	0.177	-0.04	0.304	-0.03	0.461	-0.14	0.506	-0.07
7	0.081	-0.02	0.112	-0.01	0.224	-0.04	0.372	-0.06	0.421	-0.06
8	0.03	-0.02	0.062	-0.01	0.147	-0.05	0.276	-0.12	0.33	-0.1
9	-0.01	-0.01	0.026	-0.01	0.08	-0.03	0.203	0.006	0.248	-0.02
10	-0.04	-0.04	-0.01	-0.03	0.02	-0.03	0.143	-0.01	0.165	-0.07
11	-0.07	-0.03	-0.04	-0.03	-0.03	-0.03	0.091	0.014	0.091	-0.01
12	-0.09	-0.02	-0.06	-0.02	-0.08	-0.04	0.054	0.039	0.008	-0.13

Os resultados dos testes de raízes unitárias⁴¹ estão nas Tabelas 3.7 e 3.8 a seguir.

⁴¹ Testes LLC e IPS, com intercepto e tendência individuais, número de defasagens calculado pelo critério de Schwarz, kernel de Bartlett e largura de banda pelo critério de Newey-West. Resultados gerados no aplicativo Eviews 5.0.

TABELA 3.7 – Testes de raízes unitárias no painel de cartões de crédito

Variável	Método	Estatística	p-valor	Cross-sections	Obs
Log do valor movimentado	LLC	3.60009	0.9998	6	106
	IPS	-0.74537	0.228	6	106
Taxa de intercâmbio	LLC	0.3789	0.6476	4	83
	IPS	0.42883	0.666	4	83
Taxa de desconto	LLC	-2.24449	0.0124	6	119
	IPS	-2.76632	0.0028	6	119

TABELA 3.8 – Testes de raízes unitárias no painel de cartões de débito

Variável	Método	Estatística	p-valor	Cross-sections	Obs
Log do valor movimentado	LLC	-1.71256	0.0434	3	71
	IPS	-0.13227	0.4474	3	71
Taxa de intercâmbio	LLC	-2.3398	0.0096	3	78
	IPS	-2.60051	0.0047	3	78
Taxa de desconto	LLC	-2.73538	0.0031	3	77
	IPS	-2.18919	0.0143	3	77

Os testes apontam para a presença de raiz unitária nas séries de valor movimentado (em log) tanto para cartões de crédito quanto para cartões de débito e na série de tarifa de intercâmbio para cartões de crédito. A ideia de raiz unitária significar “incorporação de choque”, nesse caso, reafirma a função da tarifa de intercâmbio como o “preço de equilíbrio” do mercado de cartões de crédito.

A identificação de raízes unitárias nos dados é ainda reforçada pelos gráficos de resíduos obtidos a partir de regressões por efeito fixo e por efeito aleatório do modelo:

$$\log(valor_{it}) = \alpha_0 + \alpha_1 \log(estab_{it}) + \alpha_2(txitc_{it}) + \alpha_3(txdesc_{it}) + \alpha_4 t + \alpha_5 q4 + \varepsilon \quad (\text{eq. 3.15})$$

para cartões de crédito e

$$\log(valor_{it}) = \beta_0 + \beta_1 \log(estab_{it}) + \beta_2(txitc_{it}) + \beta_3 t + \beta_4 q4 + \varepsilon \quad (\text{eq. 3.16})$$

para cartões de débito. Os gráficos dos resíduos são mostrados nos Gráficos 3.5, 3.6, 3.7 e 3.8.

GRÁFICO 3.5 – Resíduos da regressão do modelo de cartões de crédito após estimação por efeito fixo

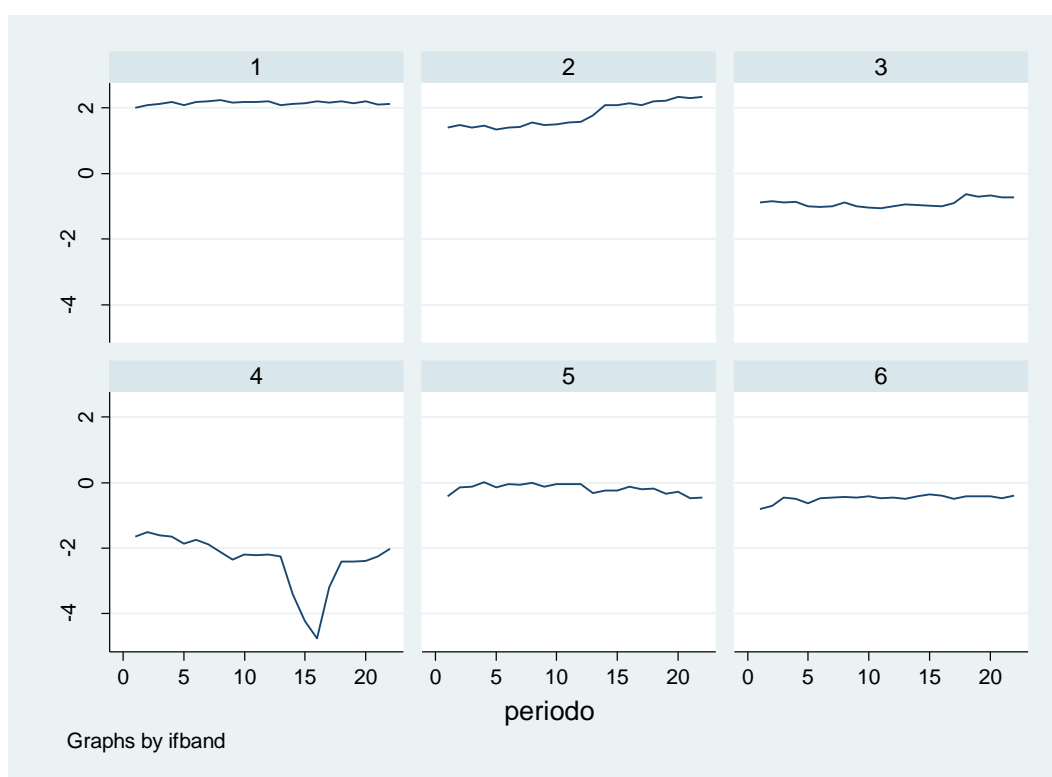


GRÁFICO 3.6 – Resíduos da regressão do modelo de cartões de crédito após estimação por efeito aleatório

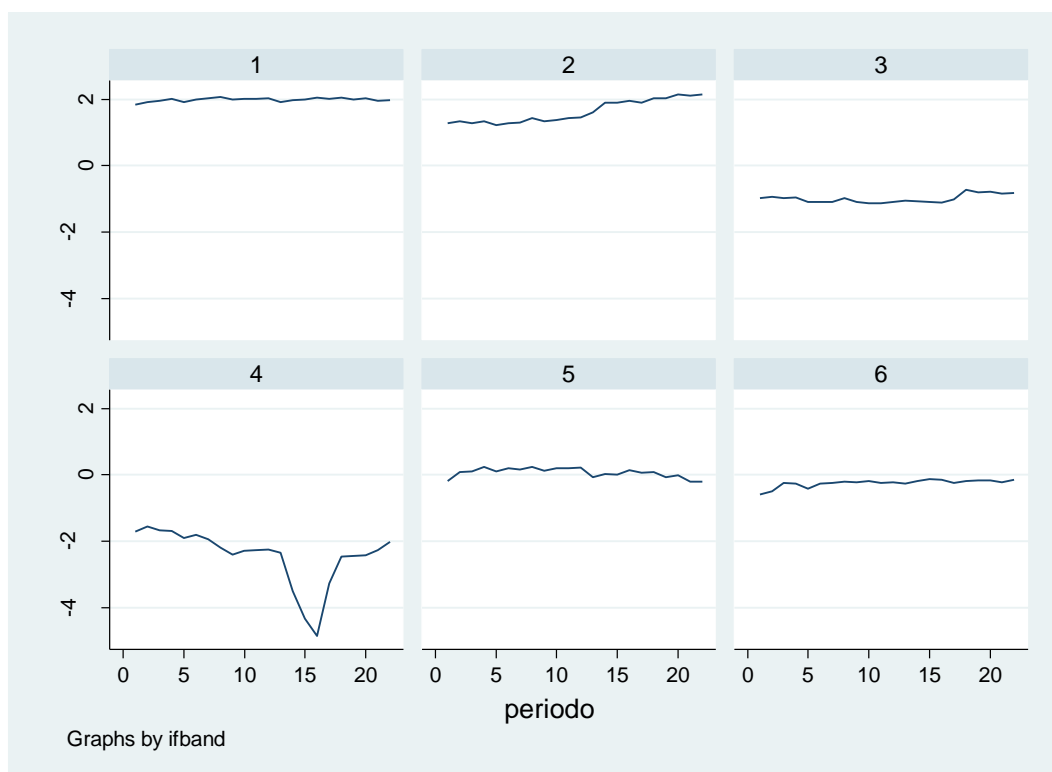


GRÁFICO 3.7 – Resíduos da regressão do modelo de cartões de débito após estimação por efeito fixo

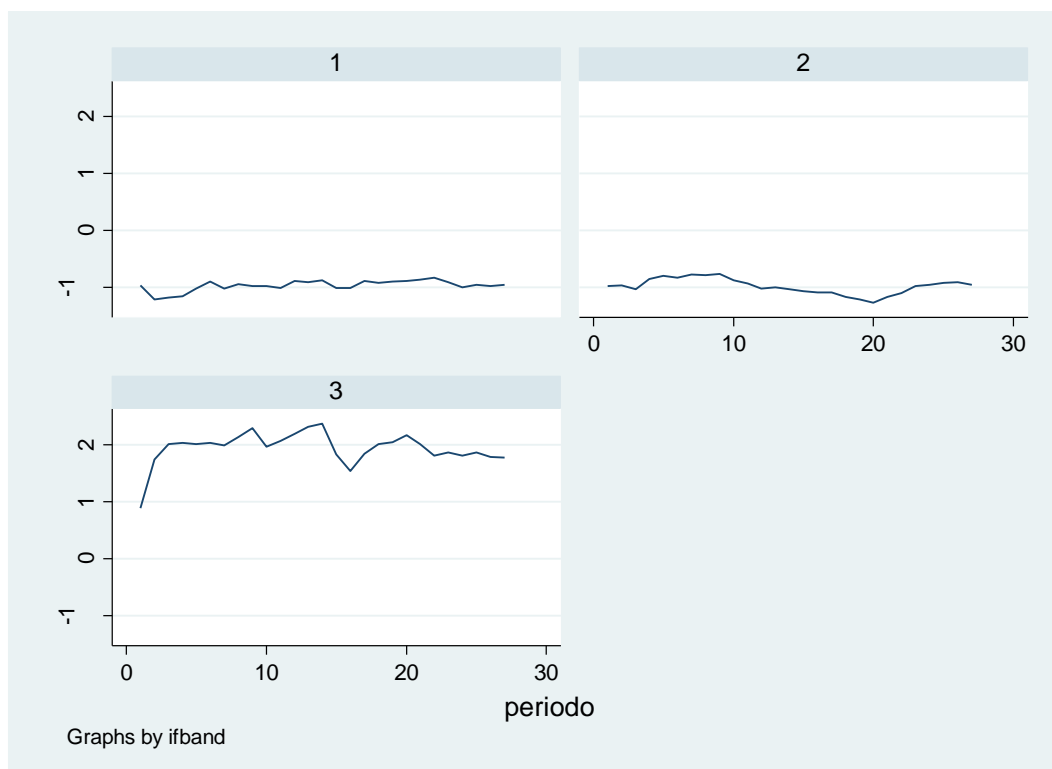
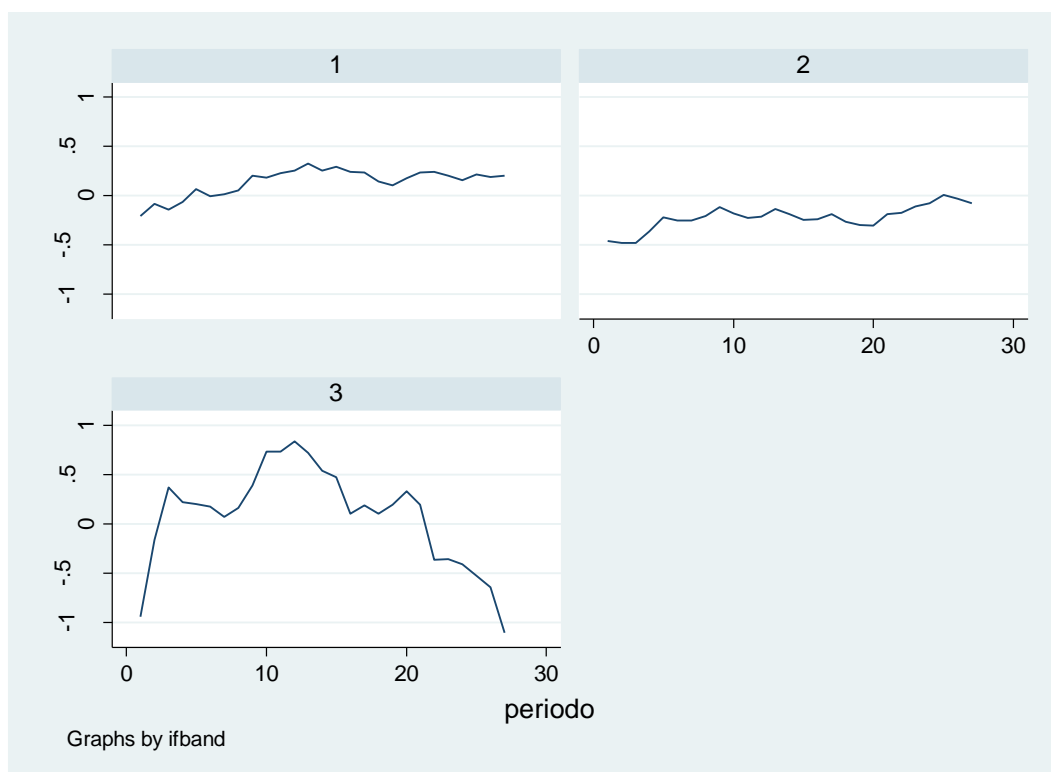


GRÁFICO 3.8 – Resíduos da regressão do modelo de cartões de débito após estimação por efeito aleatório



Teste de Hausman revela ainda não haver diferença estatística entre as estimações por efeito fixo e aleatório para o modelo proposto no caso de cartões de crédito ($\chi^2 = 0,26$ e p-valor de 99,84%). Já para o caso de cartões de débito, o teste de Hausman revelou diferença entre as estimações ($\chi^2 = 129$ e p-valor 0%), mas, como os resíduos de ambas as estimações mostraram comportamento distinto de ruído branco, a opção continua sendo pela estimação em primeira diferença.

Esse conjunto de resultados – presença de raízes unitárias e inadequação equivalente de estimações por efeito fixo e por efeito aleatório – aponta para a estimação em primeira diferença como a mais apropriada para o modelo.

3.7 Modelo explicativo final: estimações em primeira diferença

A estimação em primeira diferença é feita na forma:

$$\Delta \log(\text{valor}_{it}) = \alpha_1 \Delta \log(\text{estab}_{it}) + \alpha_2 \Delta(\text{txitc}_{it}) + \alpha_3 \Delta(\text{txdesc}_{it}) + \alpha_4 q4 + \varepsilon \quad (\text{eq. 3.17})$$

para cartões de crédito e

$$\Delta \log(\text{valor}_{it}) = \alpha_1 \Delta \log(\text{estab}_{it}) + \alpha_2 \Delta(\text{txitc}_{it}) + \alpha_3 q4 + \varepsilon \quad (\text{eq. 3.18})$$

para cartões de débito.

Nesse modelo, não se emprega tendência determinística, já que foi identificada tendência estocástica (raiz unitária) nas variáveis. Mas é utilizada uma variável de controle para sazonalidade ($q4$).

Uma inspeção visual sugere que, com esse modelo, os resíduos pós-estimação se comportam de forma mais próxima à de ruído branco, como mostram os Gráficos 3.9 e 3.10.

GRÁFICO 3.9 – Resíduos da regressão do modelo de cartões de crédito após estimação em primeira diferença

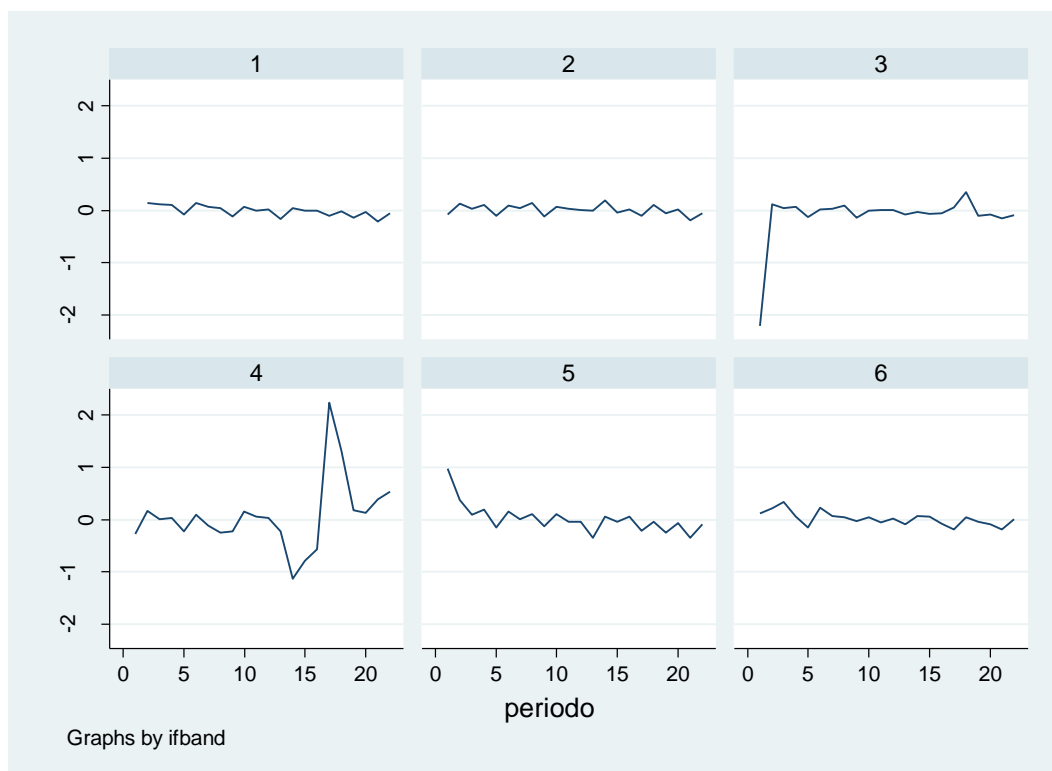
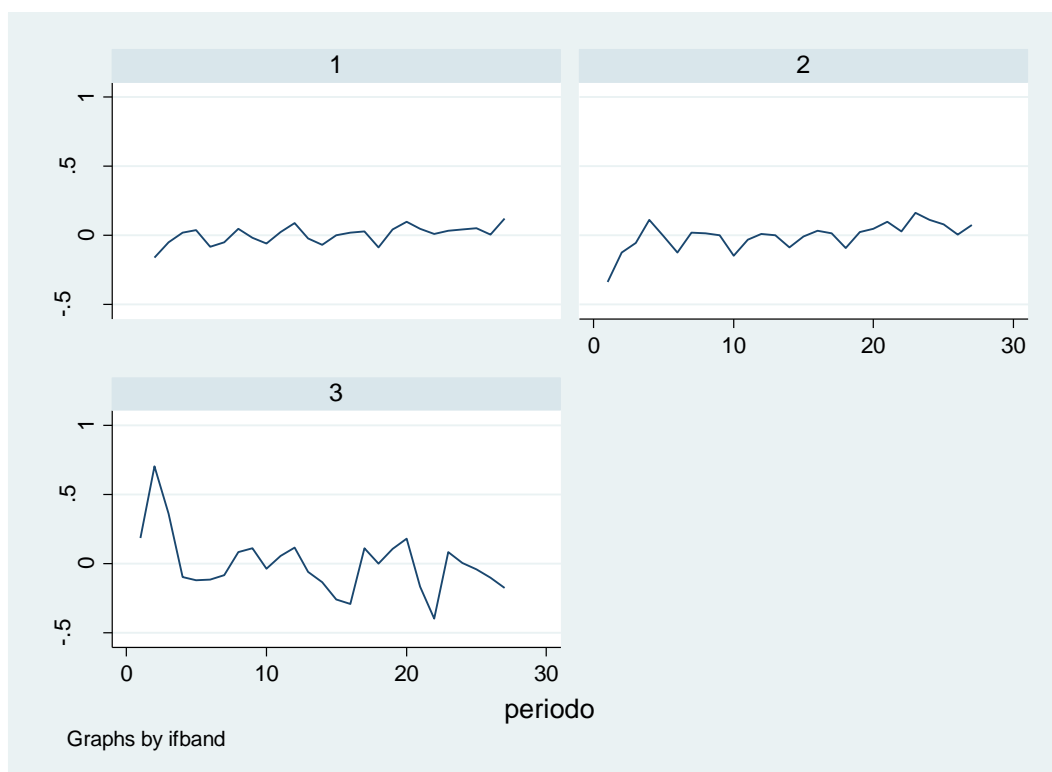


GRÁFICO 3.10 – Resíduos da regressão do modelo de cartões de débito após estimação em primeira diferença



Por fim, os parâmetros estimados para tal modelo são os mostrados na Tabela 3.9.

TABELA 3.9 – Parâmetros estimados para os modelos de cartões de crédito e de débito – coeficiente e p-valor

Variável	Modelo Crédito	Modelo Débito
Estabelecimentos (1a dif do log)	1.75 (0)	1.32 (0)
Taxa de desconto (1a dif do log)	0.137 (0.67)	
Tarifa de intercâmbio (1a dif do log)	-0.838 (0.0086)	0.216 (0.42)
<i>Dummy</i> 4o trim (1a dif do log)	0.0671 (0.4)	0.161 (0.00014)
Intercepto (1a dif do log)	-0.019 (0.62)	-0.0431 (0.03)
R2	0.806	0.918

A estimação dos parâmetros mostra, para ambos os mercados, uma forte reação do volume de utilização em relação ao número de estabelecimentos credenciados. Com relação à tarifa de intercâmbio, a reação é negativa e significativa para o caso de cartões de crédito e não-significativa para o caso de cartões débito, confirmando a ideia de que ela é o “preço que equilibra as partes” no mercado de cartões de crédito – cuja relutância à adesão pelos estabelecimentos é maior, devido aos custos implicados – mas não tem essa mesma função no mercado de cartões de débito.

Um questionamento que pode ser levantado sobre o modelo diz respeito à endogeneidade entre as variáveis volume financeiro e número de estabelecimentos credenciados – e o viés de estimação decorrido daí⁴². É presumível existir uma retroalimentação entre o volume financeiro movimentado nos mercados de cartões e o número de estabelecimentos credenciados ativos. Ou seja, assim como um maior volume movimentado incentiva mais estabelecimentos a se credenciarem, o aumento no número de credenciados faz com que o volume aumenta. E isso significa endogeneidade no modelo e viés na estimação.

O modelo alternativo proposto na seção seguinte pode solucionar esse problema.

3.8 Modelo preditivo VAR

O modelo apresentado até aqui mostrou um bom poder explicativo para a curva de aceitação de cartões por parte dos comerciantes, revelando especialmente a sensibilidade desse mercado à tarifa de intercâmbio. No entanto, o modelo recorre ao artifício de empregar o próprio número de estabelecimentos credenciados ativos como variável para captar o efeito da aceitação por parte dos consumidores, o que prejudica seu caráter preditivo. Além disso, tem-se o problema da endogeneidade comentado anteriormente.

⁴² O uso de estimação GMM (*generalized method of moments*) e Arellano-Bond não fez parte dos exercícios tratados aqui e podem ser feitos em trabalho futuro.

Para contornar tais deficiências, utilizou-se também um modelo VAR em primeira diferença que toma como variáveis endógenas o volume financeiro movimentado nos mercados de cartões e o número de estabelecimentos credenciados ativos (em log) e como variáveis exógenas as mesmas covariadas da estimação do exercício anterior. Assim, as equações do modelo foram as seguintes⁴³:

$$\Delta \log(valor_t) = \alpha_0 + \sum_{i=1}^2 \alpha_i \Delta \log(valor_{t-i}) + \sum_{j=1}^2 \alpha_{j+2} \Delta \log(estab_{t-j}) + \sum_k \alpha_k \Delta taxa_k + q4 + \varepsilon$$

(eq. 3.19)

$$\Delta \log(estab_t) = \beta_0 + \sum_{i=1}^2 \beta_i \Delta \log(valor_{t-i}) + \sum_{j=1}^2 \beta_{j+2} \Delta \log(estab_{t-j}) + \sum_k \beta_k \Delta taxa_k + q4 + \varepsilon$$

(eq. 3.20)

em que *taxa* refere-se às taxas de desconto e de intercâmbio (a primeira, apenas no caso dos cartões de crédito).

Tais modelos também foram estimados empregando-se 4 defasagens ao invés da variável de sazonalidade *q4*.

Os resultados estão apresentados na Tabela 3.10.

⁴³ O uso de variáveis exógenas ou instrumentos, ao invés de apenas as duas variáveis endógenas discutidas, também pode ser alvo de exercícios em trabalhos futuros. O uso de defasagens até 2ª ordem para as duas variáveis foi resultado de testes LR (*sequential modified likelihood ratio*), padrão do aplicativo E-Views 5.

TABELA 3.10 – Parâmetros estimados para os modelos VAR em primeira diferença para cartões de crédito e de débito – coeficientes e erros-padrão

	Crédito c/ q4		Débito c/ q4		Crédito s/ q4		Débito s/ q4	
	valor*	estab*	valor*	estab*	valor*	estab*	valor*	estab*
(valor-1)*	0,41 (-0,09)	0,22 (-0,07)	0 (-0,11)	-0,03 (-0,03)	0,36 (-0,1)	0,04 (-0,06)	0,09 (-0,16)	0,03 (-0,04)
(valor-2)*	0,04 (-0,05)	0,11 (-0,04)	0,08 (-0,1)	0,08 (-0,03)	0,34 (-0,1)	0,16 (-0,06)	0,22 (-0,15)	0,12 (-0,04)
(valor-3)*	- -	- -	- -	- -	-0,17 (-0,11)	0,56 (-0,06)	-0,35 (-0,14)	-0,06 (-0,04)
(valor-4)*	- -	- -	- -	- -	-0,03 (-0,05)	-0,01 (-0,03)	0,41 (-0,12)	0,1 (-0,03)
(estab-1)*	0,34 (-0,11)	0,68 (-0,09)	1,21 (-0,34)	0,8 (-0,1)	0,38 (-0,12)	0,5 (-0,07)	0,35 (-0,63)	0,52 (-0,16)
(estab-2)*	-0,08 (-0,11)	-0,26 (-0,09)	-0,21 (-0,14)	-0,11 (-0,04)	-0,22 (-0,14)	-0,41 (-0,08)	-1,01 (-0,62)	-0,31 (-0,16)
(estab-3)*	- -	- -	- -	- -	0,26 (-0,14)	0,18 (-0,08)	2,43 (-0,52)	0,57 (-0,14)
(estab-4)*	- -	- -	- -	- -	0,1 (-0,12)	0,01 (-0,07)	-0,65 (-0,17)	-0,15 (-0,04)
tx_interc**	-0,68 (-0,56)	0,38 (-0,46)	0,39 (-0,27)	-0,05 (-0,08)	-0,66 (-0,57)	0,43 (-0,33)	0,41 (-0,31)	-0,02 (-0,08)
tx_desc**	0,64 (-0,35)	0,5 (-0,29)	- -	- -	1,37 (-0,44)	0,64 (-0,25)	- -	- -
q4	0,14 (-0,04)	0,01 (-0,03)	0,26 (-0,04)	0,05 (-0,01)	- -	- -	- -	- -
Intercepto	-0,05 (-0,02)	-0,01 (-0,02)	-0,08 (-0,02)	-0,01 (-0,01)	-0,02 (-0,02)	-0,01 (-0,01)	-0,04 (-0,02)	0 (-0,01)
R2	0,44	0,54	0,53	0,65	0,5	0,8	0,43	0,69

* Primeira diferença do log

** Primeira diferença

Os resultados de uma simulação impulso-resposta reforçam adequação do modelo⁴⁴.

⁴⁴ Para uma indicação mais acurada de tendência das variáveis, pode-se analisar sua estacionaridade, o que pode ser feito em trabalho futuro.

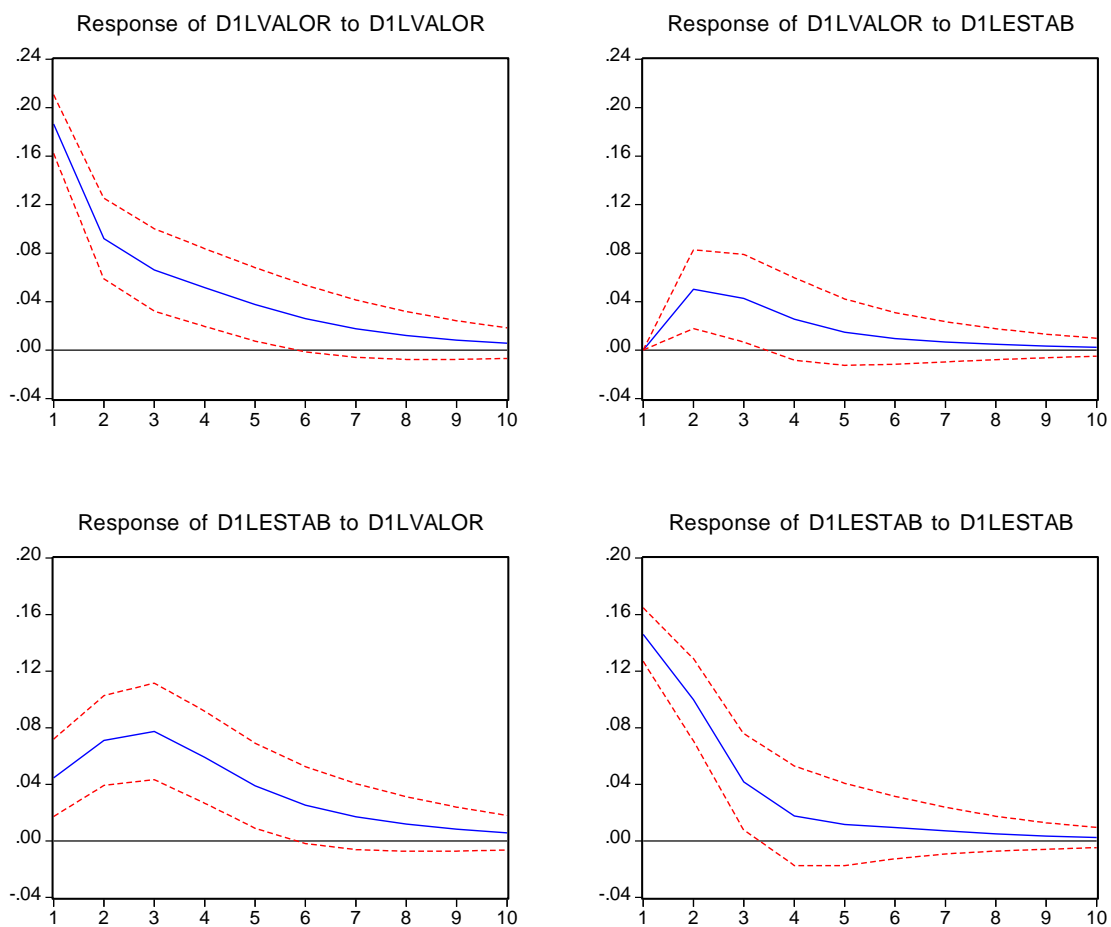
GRÁFICO 3.11 – Simulações impulso-resposta no modelo de cartões de crédito em primeira diferença com variável de sazonalidade q4Response to Cholesky One S.D. Innovations ± 2 S.E.

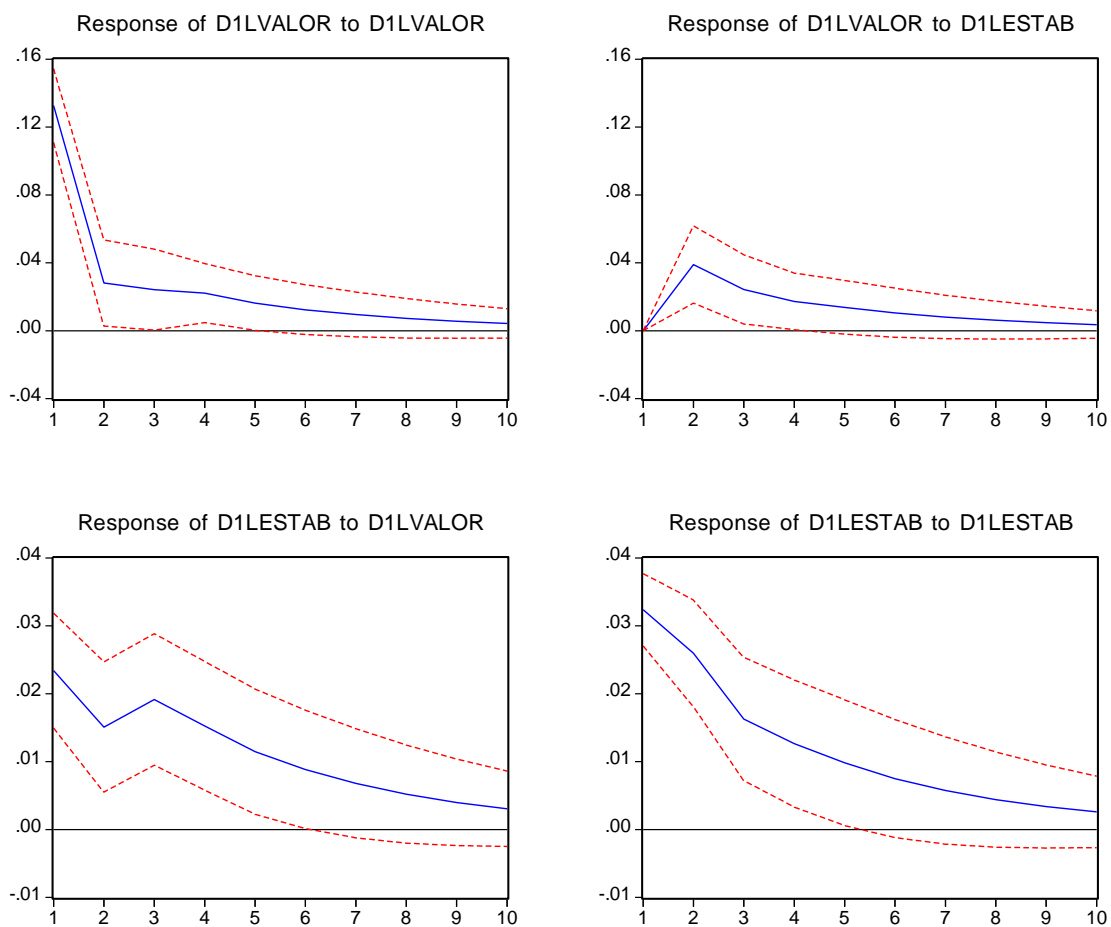
GRÁFICO 3.12 – Simulações impulso-resposta no modelo de cartões de débito em primeira diferença com variável de sazonalidade q4Response to Cholesky One S.D. Innovations ± 2 S.E.

GRÁFICO 3.13 – Simulações impulso-resposta no modelo de cartões de crédito em primeira diferença com 4 defasagens

Response to Cholesky One S.D. Innovations ± 2 S.E.

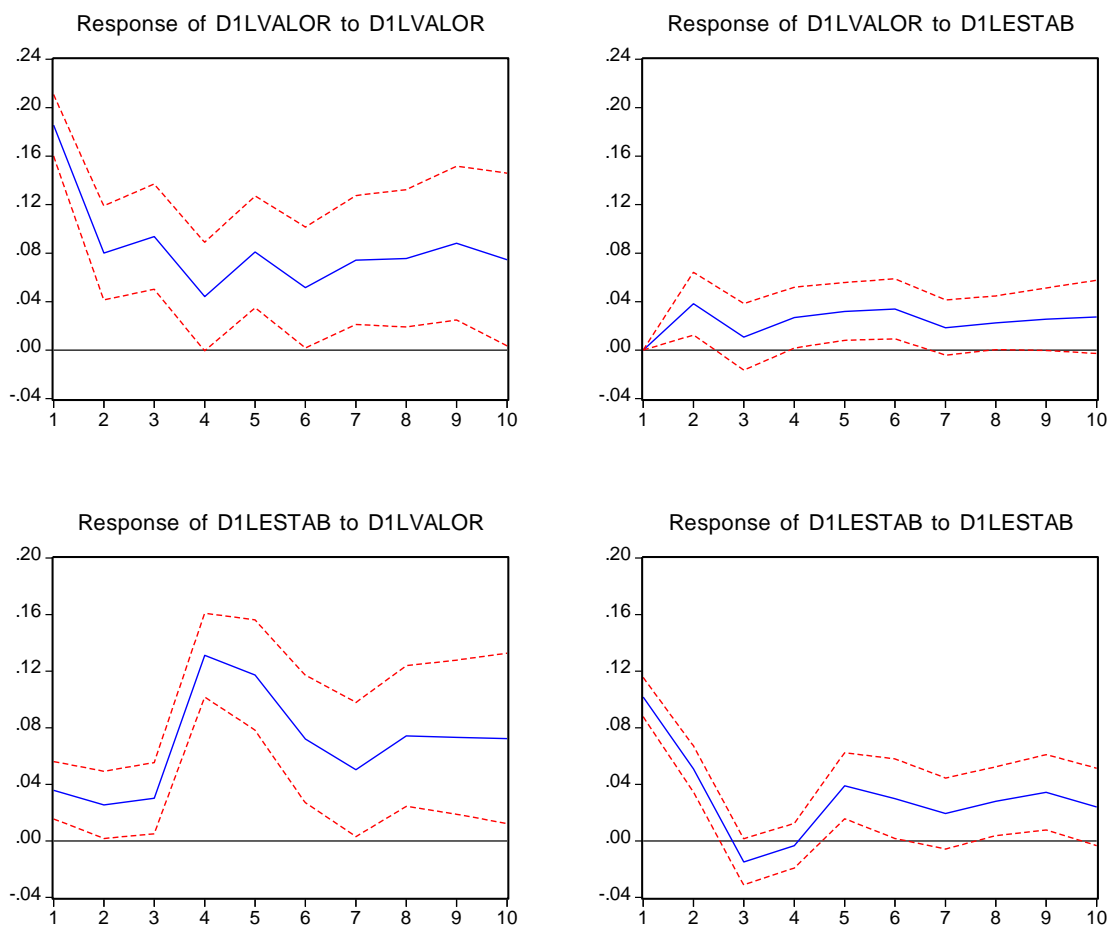
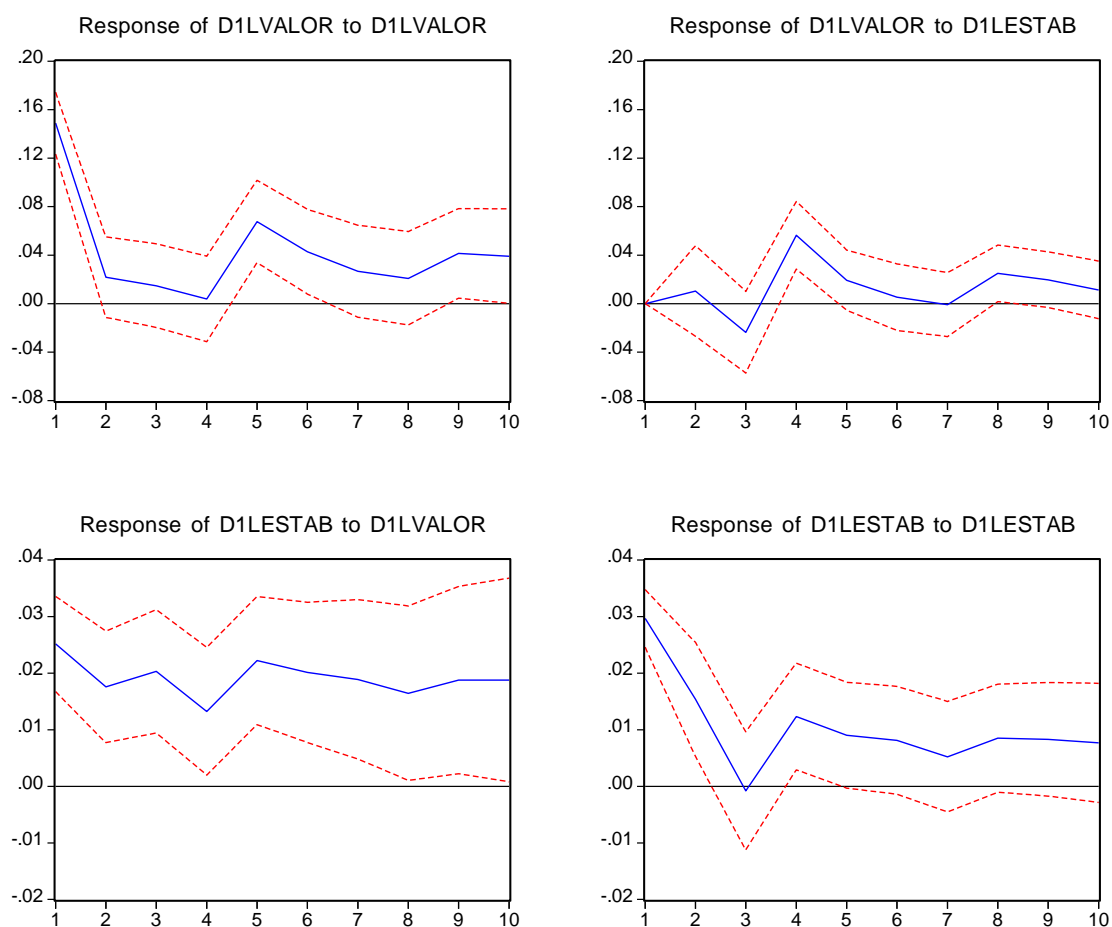


GRÁFICO 3.14 – Simulações impulso-resposta no modelo de cartões de débito em primeira diferença com 4 defasagens

Response to Cholesky One S.D. Innovations ± 2 S.E.



3.9 Considerações gerais sobre os exercícios econométricos

Esse capítulo propôs exercícios econométricos referentes aos mercados de cartões de crédito e de débito no Brasil, a partir de uma base de dados em painel, tomada de emissores e credenciadores de cartões no Brasil.

Os resultados para uma estimação em primeira diferença reforçaram empiricamente o entendimento, já abordado em diversos trabalhos teóricos, de que a taxa de intercâmbio no mercado de cartões de crédito funciona como um “preço de equilíbrio” entre os dois lados desse mercado (credenciador e emissor, ou, em última instância, comerciante e consumidor). O mesmo não se observa no caso dos cartões de débito.

Os resultados para a estimação VAR corroboraram a intuição de que não apenas o valor financeiro movimentado em cartões é função da quantidade de estabelecimentos credenciados a aceitá-los como vice-versa, isto é, maior volume financeiro movimentado induz mais estabelecimentos a aceitarem cartões. Trata-se, assim, de uma dinâmica endógena, em que as variáveis se retroalimentam.

O capítulo seguinte voltará suas atenções para a relação entre a adoção de tais instrumentos de pagamento e os eventuais reflexos sobre os mecanismos de política monetária.

4 IMPACTOS DOS INSTRUMENTOS ELETRÔNICOS DE PAGAMENTO SOBRE A POLÍTICA MONETÁRIA

4.1 A inter-relação entre as atividades de um banco central

Tipicamente, bancos centrais dedicam-se a três atividades essenciais: a execução da política monetária do país; a supervisão de seu sistema financeiro; e a arquitetura e observância de seu sistema de pagamentos (Blinder, 1998). Política monetária, aqui, é entendida como o conjunto de estratégias e mecanismos diretos ou indiretos de oferta de moeda, voltados para assegurar estabilidade ao seu poder de compra (tanto interno, na forma de combate à inflação, quanto externo, no que se costuma chamar política cambial⁴⁵) sem prejuízo ao produto real da economia. A supervisão do sistema financeiro diz respeito à regulação sobre seus agentes e ao acompanhamento de seus riscos individuais (sejam eles riscos de liquidez, de crédito, operacionais, legais ou de outra natureza). E o sistema de pagamentos consiste no desenho por onde corre o fluxo de moeda e ativos financeiros entre os agentes. Nesse caso, a preocupação do banco central é que esse desenho tenha o máximo de segurança (de forma a assegurar certeza de liquidação e que eventuais inadimplências não promovam contágio e risco de quebra ao sistema como um todo) e eficiência (mínimo custo e máxima comodidade).

A literatura econômica discute bem mais vastamente o tema política monetária do que os temas supervisão do sistema financeiro e sistemas de pagamentos ou a própria inter-relação entre eles.

Sobre a relação entre política monetária e supervisão do sistema financeiro, considerando-se que desta última faz parte a observância contábil e de portfólio dos bancos, a corrente de pensamento econômico chamada “pós-

⁴⁵ A política cambial às vezes é tratada separadamente da política monetária, mas, para os objetivos desse trabalho, será tratada conjuntamente, já que a função de fiel depositário das reservas internacionais (e, portanto, de executor da política cambial) é reconhecidamente um papel clássico de banco central.

keynesiana estruturalista” entende que a oferta de moeda (variável de política monetária) é função não apenas da taxa de juros administrada pelo governo e da demanda por crédito, mas também da gestão de ativos e passivos dos bancos, alegando que “mesmo que a autoridade monetária se recusasse a acomodar aumentos na demanda por reservas, os bancos ainda poderiam acomodar parcialmente uma elevação na demanda por empréstimos a partir de suas próprias iniciativas” (Palley, 2002).

Por sua vez, a relação entre as observâncias do sistema financeiro, tratando-se seus agentes individualmente, e do sistema de pagamentos centra-se na discussão sobre efeito-contágio e risco sistêmico⁴⁶, ou seja, sobre o risco de a eventual inadimplência de um agente ante os demais poder acarretar um “efeito dominó” de falta de liquidez que atinja dimensões graves ao sistema como um todo.

Por fim, resta a discussão sobre a relação entre política monetária e sistema de pagamentos, que é, em certa medida, o objeto deste capítulo. A discussão está dividida a seguir em três seções. A primeira delas traz uma revisão das principais abordagens sobre política monetária e seus mecanismos de transmissão. A segunda apresenta a discussão sobre moeda endógena, princípio adotado por diversos economistas mas visto por outros como controverso. A seção final tenta, então, responder se a adoção continuada de instrumentos eletrônicos de pagamento traz consequências sobre a efetividade dos mecanismos de política monetária.

4.2 A política monetária e seus mecanismos de transmissão

Desde os anos de 1960 (até o agravamento da crise financeira de 2008), o pensamento econômico predominante passou a preconizar que a estabilização de produto e inflação ficasse a cargo da política monetária. Nesse período, a

⁴⁶ Segundo Litan (1997), o efeito-contágio ocorre quando a insolvência ou a falta de liquidez de um banco dá início ao colapso de outros bancos credores do banco inadimplente. Kaufman (1999) define risco sistêmico como a probabilidade de ocorrer perdas acumuladas devido a um evento que dá início a uma série de prejuízos sucessivos ao longo de uma cadeia de instituições ou mercados, que compõem um sistema. Essas perdas ocorrem porque os bancos estão fortemente interconectados via mercado interbancário e sistema de pagamentos.

ideia de uso da política fiscal para tais fins perdeu força, tanto pela preocupação com a possibilidade de persistência de grandes déficits orçamentários quanto pela desconfiança se sistemas políticos seriam suficientemente ágeis para prover mudanças em impostos e gastos governamentais nos prazos necessários para que se sentissem os efeitos desejados⁴⁷.

Até os anos 1970, a política monetária era vista basicamente como “controle da oferta de moeda” e os trabalhos sobre o tema concentravam-se no controle do banco central sobre os agregados monetários⁴⁸. Somente a partir daí foi proposto um modelo de escolha do instrumento de meta entre as opções taxa de juros (r) e agregado monetário (M), com a conclusão de que choques sobre a curva LM favoreciam a escolha da taxa de juros, enquanto choques na curva IS favoreciam a opção pelo agregado monetário (Poole, 1970). No entanto, evidências empíricas de não-cointegração no longo-prazo entre diversos agregados monetários e o PIB nominal, bem como a própria não-conclusão teórica sobre qual dos agregados empregar ($M0$, $M1$, $M2$, etc) fizeram a alternativa pelo instrumento M ser abandonada⁴⁹.

Posto que o instrumento adequado é a taxa de juros, vem o novo dilema: a demanda agregada é sensível a taxas de juros *reais de longo prazo*, mas o banco central só pode controlar sua taxa de juros *nominal de curto prazo*. Para resolver o problema do prazo, simplesmente assume-se a hipótese de que a estrutura a termo do mercado faça eficientemente a ligação entre curto e longo prazo. Já quanto à divergência entre taxa nominal e real, é preciso um cuidado extra para evitar espirais inflacionárias e deflacionárias⁵⁰. Assim, os modelos

⁴⁷ O tema da dominância fiscal, mais recorrente a partir do final dos anos 1990 e que alerta que a política monetária só é frutífera se acompanhada de uma gestão adequada da dívida, não será tratado aqui por desviar-se do tema deste trabalho. Mas uma boa referência para o assunto é Blanchard (2004).

⁴⁸ O exemplo mais clássico é Friedman (1968).

⁴⁹ Nas palavras de Gerald K. Bouey, presidente do Banco Central do Canadá entre 1973 e 1987, “nós não abandonamos os agregados monetários, eles nos abandonaram” (Blinder, 1998).

⁵⁰ Suponha o banco central elevando a taxa nominal (aperto monetário). Isso acarreta elevação da taxa real, contração da demanda agregada e redução da inflação, que é o efeito inicialmente desejado. Acontece que, no período seguinte, se a taxa nominal for mantida, a

ortodoxos passaram a tratar como instrumento-chave de política monetária a chamada “taxa de juros real neutra” (r^*), definindo-a como a taxa que iguala o PIB na “curva IS estática” (curva IS após a acomodação das variáveis defasadas e na ausência de choques) ao PIB potencial. Três aspectos devem ser ressaltados nessa definição. Primeiro, a definição de “neutralidade” refere-se a controle da variável inflação (constante no médio prazo). Segundo, a taxa de juros real neutra não é um valor fixo. Ela responde, entre outras coisas, à política fiscal, à taxa de câmbio e a choques permanentes. Terceiro, a taxa é difícil de ser estimada, e funciona mais como um conceito para entendimento teórico do que como um parâmetro numérico.

Além da discussão sobre qual seria o principal “instrumento” da política monetária, outro debate recorrente é sobre quais seriam seus “objetivos”. E aqui a discussão recai basicamente sobre duas variáveis: inflação e produto.

Embora diversos bancos centrais no mundo tenham adotado modelos de metas de inflação, os modelos acadêmicos costumam considerar como função-objetivo a ser minimizada um somatório entre taxa de inflação (ou desvio da taxa de inflação em relação à meta buscada pelo banco central) e hiato do produto (diferença entre o produto observado e o produto potencial, a pleno emprego), com um parâmetro de ponderação que indica se a aversão do banco central à inflação é maior ou menor em relação à aversão ao hiato de produto⁵¹.

Compreendidos os instrumentos e objetivos de política monetária mais recorrentes na literatura acadêmica, a discussão seguinte recai sobre os

taxa *real* estará maior (porque a inflação estará menor) e pressionará a inflação mais pra baixo ainda. Assim, se a taxa nominal não for ajustada, chega-se a uma espiral deflacionária. Um afrouxamento monetário descuidado pode, pela mesma lógica, gerar uma espiral inflacionária. Nas palavras de Blinder (1998), “*Before too long, the central bank must adjust its nominal rate so as to guide the real rate back towards its neutral setting*”.

⁵¹ Os modelos usualmente tomados como referência nos trabalhos são Tinbergen (1952) e Theil (1961), que propõem a abordagem do problema como a maximização de uma função-objetivo quadrática, sujeita a equações lineares gerais para a economia; Taylor (1993), que formula uma equação linear simples (a propalada “regra de Taylor”) para definir a taxa de juros a partir da taxa real de equilíbrio de longo prazo, do desvio entre a inflação observada e a meta proposta pelo banco central e do hiato de produto; e Clarida, Galí e Gertler (1999), “versão *forward-looking* da regra de Taylor”, em que a inflação contemporânea é substituída por sua expectativa presente sobre o valor futuro.

mecanismos de transmissão dessa política, ou seja, sobre como alterações nas variáveis-instrumentos repercutem nas variáveis-objetivos. Para haver efetividade na condução da política monetária, são necessárias avaliações precisas tanto do *timing* quanto dos efeitos desta sobre a economia, o que requer um entendimento dos canais pelos quais a política monetária afeta inflação e produto.

Os canais mais largamente discutidos são o canal monetarista e o canal taxa de juros. No entanto, outros quatro canais: taxa de câmbio, preços de outros ativos e dois canais de crédito (chamados “canal de crédito estreito” e “canal de crédito amplo”) são compreendidos como fundamentais nessa dinâmica.

O *canal monetarista* foca no efeito direto de mudanças na quantidade relativa de ativos (financeiros e não-financeiros), ao invés de mudança na taxa de juros. O princípio é que, como vários ativos são substitutos imperfeitos nos portfólios dos investidores, mudanças provocadas pela política monetária na composição de carteiras provocam variações nos preços relativos, que, por sua vez, acarretam efeitos reais. Por esta visão, o papel das taxas de juros é simplesmente o de mais um dos vários preços relativos de ativos.

When asked to describe how the central bank might generate inflation, a non-economist might say something about “too many dollars chasing too few goods”. Few economists now believe this simplistic monetarist explanation; rather, they envisage a complex mechanism whereby interest rates influence demand, which in turn influences inflation⁵².

O *canal taxa de juros* é o mecanismo-chave de transmissão da política monetária no modelo keynesiano básico, em que uma contração monetária (elevação das taxas de juros de curto prazo) propaga-se para as taxas de juros de longo prazo (por meio de uma combinação de rigidez de preços e expectativas racionais), elevando o custo de capital e levando firmas e famílias a postergarem seus gastos com investimento em bens de capital, investimento imobiliário, consumo de bens duráveis e investimento em estoques⁵³.

⁵² Arestis (2002).

⁵³ A rigor, o trabalho original de Keynes não discutia a propagação entre taxas de curto e longo prazo e só considerava os efeitos sobre os investimentos em bens de capital. Seu modelo foi,

$$i \uparrow \Rightarrow I \downarrow \Rightarrow Y \downarrow$$

A despeito do fundamento teórico, estudos empíricos revelam grande dificuldade em identificar quantitativamente efeitos importantes das taxas de juros sobre o custo de capital. E foi justamente essa constatação que estimulou a busca teórica por outros mecanismos de transmissão de política monetária (Bernanke e Gertler, 1989).

Outro importante mecanismo é o *canal taxa de câmbio*. O princípio em que se baseia é que uma elevação nas taxas de juros aumenta o fluxo positivo de dólares (ou os depósitos em dólar, no caso dos EUA), tornando-os “baratos” em relação às outras moedas e, assim, desfavorecendo as exportações em relação às importações, reduzindo o saldo em transações correntes e o produto.

$$i \uparrow \Rightarrow E \uparrow \Rightarrow NX \downarrow \Rightarrow Y \downarrow^{54}$$

Duas outras abordagens, que se referem aos efeitos de preços de outros ativos e que levam à mesma conclusão entre si, extrapolam os canais de transmissão da política monetária para além das taxas de juros e de câmbio. São o efeito-riqueza sobre o consumo e a teoria do q de Tobin. Ambas as abordagens passam pela compreensão de que uma contração monetária (redução na oferta de moeda) acarreta queda no valor de mercado das firmas (ações)⁵⁵ – e também, pela mesma lógica, de propriedades imobiliárias e bens de alto valor.

A primeira abordagem, também conhecida como modelo do ciclo de vida (Modigliani, 1971), começa atrelando o consumo das famílias a três variáveis acumuladas ao longo da vida: capital humano, capital real e riqueza financeira.

na verdade, sofisticado por trabalhos posteriores, tendo sido levado, no entanto, às mesmas conclusões.

⁵⁴ A notação para a taxa de câmbio está na forma $E = X\$ / US\$$.

⁵⁵ Monetaristas e keynesianos têm suas versões próprias para explicação desse efeito. Na versão monetarista, uma retração na oferta de moeda cria nas pessoas a percepção de menor poder de compra, fazendo-as cortar despesas não-essenciais, dentre as quais ações na bolsa, o que reduz o valor de mercado das empresas. Na versão keynesiana, a contração via elevação nas taxas de juros torna o investimento em títulos mais atrativo do que em ações, fazendo migrar o investimento financeiro e, igualmente, reduzindo o valor de mercado das firmas.

Se uma parte considerável da riqueza financeira está atrelada ao valor de ações, quedas no valor de mercado das firmas acarretam queda na riqueza e, portanto, no consumo. É o que se chama efeito-riqueza.

Em outra abordagem que conduz ao mesmo resultado (Tobin, 1969), define-se o parâmetro: $q = \frac{\text{valor_de_mercado_das_firmas}}{\text{custo_de_reposição_de_capital}}$. Se o valor do parâmetro q for elevado, o custo de reposição (captação) de capital é baixo comparativamente ao valor da firma e, conseqüentemente, ela pode investir em novas plantas e equipamentos a partir de emissões de ações. Já se o valor de q for baixo, as firmas preferirão expandir-se pela aquisição de outras firmas (que estarão baratas), o que significa compra de capital velho e não investimento em capital novo. Portanto, o q de Tobin e o investimento movimentam-se no mesmo sentido, de modo que uma queda no valor de mercado das firmas (e, conseqüentemente, no q de Tobin) acarreta queda no investimento.

As duas abordagens podem, então, ser resumidas no esquema a seguir:

$$i \uparrow \Rightarrow P_e \downarrow \Rightarrow \left\{ \begin{array}{l} \text{riqueza} \downarrow \Rightarrow \text{consumo} \downarrow \\ q \downarrow \Rightarrow I \downarrow \end{array} \right\} \Rightarrow Y \downarrow$$

Tal mecanismo é chamado na literatura acadêmica *canal riqueza*.

Por fim, dois canais de transmissão relacionam-se ao mercado de crédito: o canal empréstimo bancário (*canal de crédito estreito*⁵⁶) e o canal balanço patrimonial das firmas (*canal de crédito amplo*⁵⁷).

O canal empréstimo bancário baseia-se no princípio de que um aperto monetário leva a menos depósitos bancários e, conseqüentemente, a menos empréstimos. Pode ser representado no esquema a seguir:

⁵⁶ "Narrow credit channel" ou "lending channel".

⁵⁷ "Broad credit channel".

$$i \uparrow \Rightarrow \text{depósitos}_{\text{bancários}} \downarrow \Rightarrow \text{empréstimos}_{\text{bancários}} \downarrow \Rightarrow I \downarrow \Rightarrow Y \downarrow$$

Considera-se o canal empréstimo bancário especialmente relevante para pequenas empresas, uma vez que grandes firmas contam com possibilidade de acesso direto a mercados de crédito via emissões de ações ou debêntures, sem precisar recorrer a empréstimos bancários e, portanto, os bancos acabam cumprindo função especialmente adequada para negociar com tomadores sobre os quais a assimetria de informação é maior (ou seja, as pequenas firmas, sobre as quais se tem pouca informação).

O canal balanço patrimonial das firmas envolve uma lógica mais intrincada de seleção adversa e risco moral, relacionada ao que se chamam “*agency costs*” nos mercados financeiros, que se refere ao custo incorrido por uma organização decorrente do risco de divergência entre os interesses do “agente” e do “principal” (e consequentes ações “indevidas” do agente), embutindo aí o custo de controles para mitigação desse risco.

O princípio nesse caso é que menor patrimônio líquido das firmas (PL) significa que bancos emprestadores efetivamente recebem menor colateral por seus empréstimos, o que eleva o risco de maiores perdas por seleção adversa. Portanto, uma queda no PL das firmas, que eleva o problema de seleção adversa, leva a uma queda nos empréstimos para que elas obtenham financiamento para seus investimentos. Da mesma forma, menor PL das firmas significa menor patrimônio dos proprietários (acionistas controladores), incentivando-os a engajarem-se em projetos de investimento mais arriscados e terem, assim, maior probabilidade de não conseguirem honrar seus compromissos com os bancos emprestadores. Ou seja, eleva-se o risco moral e isso acarreta quedas nos empréstimos – e, portanto, no investimento. Do mesmo modo que o efeito se transmite pelo PL das firmas, também afeta o fluxo de caixa do balanço patrimonial. O mecanismo pode ser sintetizado no esquema a seguir:

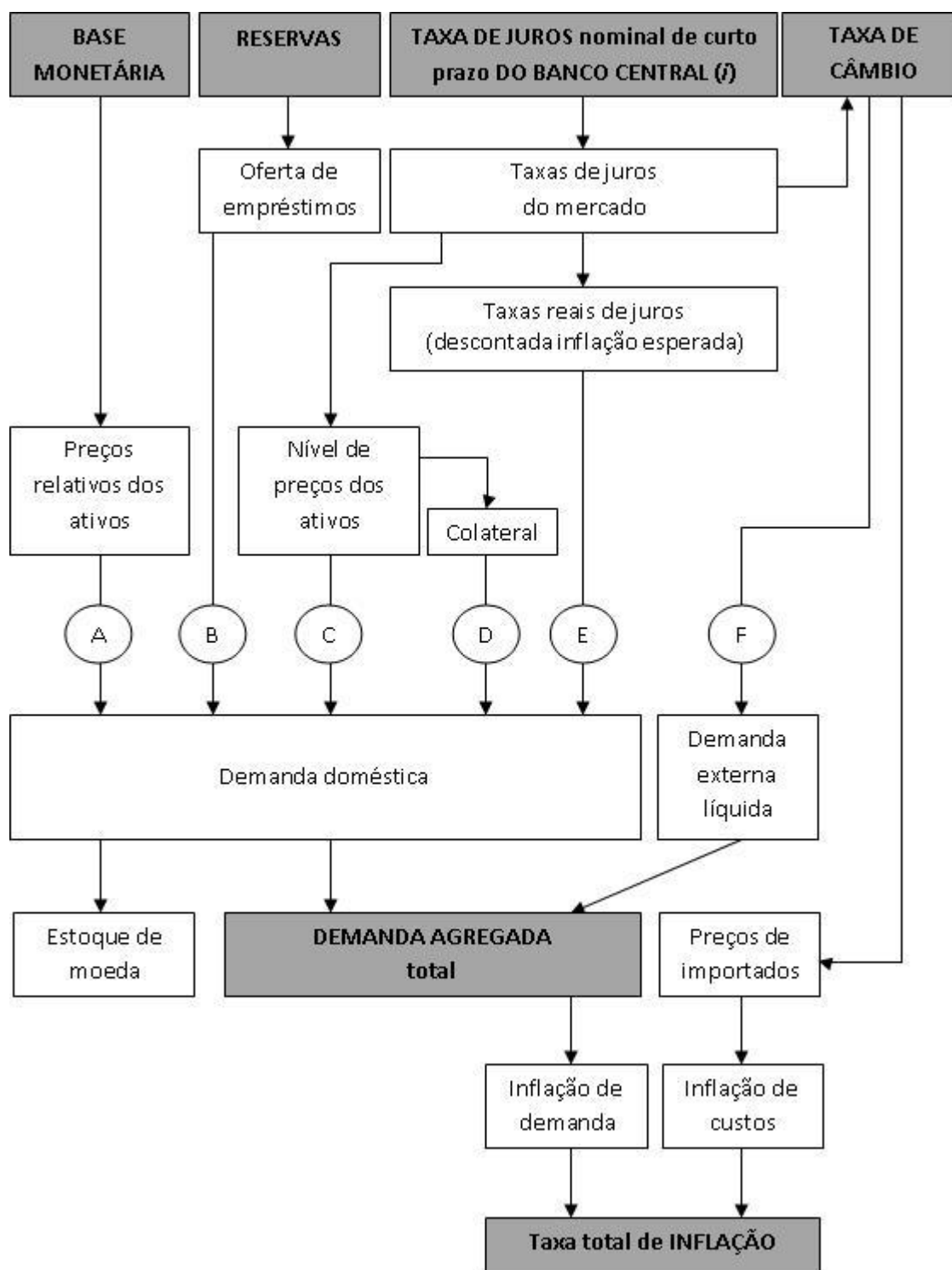
$$i \uparrow \Rightarrow \left\{ \begin{array}{l} P_e \downarrow \Rightarrow PL \downarrow \\ \text{fluxo}_{\text{de}}_{\text{caixa}} \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \text{seleção}_{\text{adversa}} \uparrow \\ \text{risco}_{\text{moral}} \uparrow \end{array} \right\} \Rightarrow \text{empréstimos} \downarrow \Rightarrow I \downarrow \Rightarrow Y \downarrow$$

Uma lógica análoga é aplicável ao “balanço patrimonial” das famílias (“patrimônio líquido” e “fluxo de caixa”), que, sob maior percepção de risco econômico oriundo de uma contração monetária, reduzem seus consumos de imóveis e bens duráveis.

$$i \uparrow \Rightarrow \left\{ \begin{array}{l} P_e \downarrow \Rightarrow PL \downarrow \\ \text{fluxo_de_caixa} \end{array} \right\} \Rightarrow \text{consumo_de_imóveis_e_bens_duráveis} \downarrow \Rightarrow Y \downarrow$$

Com base no que foi apresentado até aqui, a proposta deste trabalho é pautar o entendimento da transmissão de política monetária nos mecanismos ilustrados na Figura 4.1 a seguir.

FIGURA 4.1 – Canais de transmissão da política monetária



O esquema da Figura 4.1 exhibe os canais de transmissão explicados anteriormente (A - canal monetarista; B - canal de crédito estreito; C - canal riqueza; D - canal de crédito amplo; E - canal taxa de juros; F - canal taxa de câmbio), mas traz algumas diferenças em relação a outros trabalhos⁵⁸, a saber:

1. Coloca sob a alçada do banco central não apenas a taxa de juros e a base monetária, mas também a taxa de câmbio (mesmo considerando a influência que esta sofre das taxas de juros do mercado), já que o banco central se constitui como fiel depositário das reservas internacionais, e também as reservas bancárias, que podem ser alvo de administração de liquidez por parte do banco central, via recolhimentos compulsórios e/ou operações de redesconto;
2. Estende o mecanismo até os seus efeitos sobre a inflação, e não apenas sobre a demanda agregada, entendendo a inflação como efeito não apenas da demanda mas também de custos de importação; e
3. Entende o estoque de moeda como efeito, não como causa, da demanda agregada.

Além disso, embora isso não apareça explicitamente no diagrama, considera-se que os governos dispõem ainda de outras formas de agir sobre o canal oferta de empréstimos, que são o fortalecimento e/ou redirecionamento de empréstimos, via bancos públicos oficiais.

Uma dificuldade empírica que se impõe sobre essa discussão é o “dimensionamento” dos efeitos de cada canal, na tentativa de identificar quão relevante é cada um deles. A dificuldade dá-se basicamente por três razões:

1. Atraso nos efeitos, que podem levar bastante tempo até serem sentidos⁵⁹ e as medições precisam considerar isso.

⁵⁸ Arestis (2002) já considera os dois últimos aspectos. Kuttner e Mosser (2002), não.

⁵⁹ Estudos empíricos como Angeloni et alli (2002) já apontaram demoras de até 3 anos no efeito de uma mudança na taxa de juros do BCE sobre a inflação na zona do euro.

2. O fato de os canais funcionarem de forma simultânea. Uma mudança na taxa de juros propaga-se não só diretamente por este canal, mas também se refletindo nos canais de crédito, câmbio e riqueza, por exemplo.

3. Existência de uma dinâmica endógena pela qual as decisões de política monetária não apenas *influenciam* a economia real mas também *reagem* a ela. Bancos centrais normalmente relaxam a política monetária diante de uma realidade econômica frágil e apertam-na ante um cenário econômico forte.

No entanto, há conclusões empíricas que apontam para grande relevância dos canais taxa de câmbio, riqueza e crédito, em contraste com efeitos mais tênues dos canais taxa de juros (efeito direto) e monetarista (Angeloni et alli, 2002).

A busca de uma resposta sobre os efeitos da adoção de instrumentos eletrônicos de pagamento sobre os mecanismos de política monetária será fundamentada nesse entendimento, aliado a uma discussão sobre moeda endógena.

4.3 A discussão sobre moeda endógena

Desde os anos seguintes ao lançamento da *Teoria Geral* de Keynes, correntes distintas de pensamento econômico surgiram como desdobramentos de suas ideias. Correntes dominantes, que confluíram para o que hoje se chamam “novos keynesianos”, buscaram convergências entre a abordagem de Keynes e teorias anteriores (tais como a teoria dos fundos emprestáveis e a teoria quantitativa da moeda) e chegaram a resultados como a neutralidade da moeda (não-influência desta nos resultados de longo prazo de variáveis reais da economia) e o modelo IS-LM (capaz de determinar uma taxa de juros de equilíbrio para a economia). Outra corrente, conhecida como pós-keynesiana, propôs um entendimento que coloca o setor bancário (e sua capacidade de endogenamente ofertar moeda) no centro da dinâmica econômica e concluiu pela existência de efeitos, inclusive de longo prazo, de choques monetários sobre a economia real (a partir de seus efeitos sobre a demanda efetiva).

Esquemas pós-keynesianos preveem causalidade endógena entre demanda por crédito bancário e oferta de moeda. E basicamente oferecem duas abordagens sobre o problema, conhecidas por “acomodacionista” e “estruturalista”. Acomodacionistas modelam a oferta de moeda como função apenas da taxa de juros administrada pelo governo e da demanda por crédito – ou seja, $M^S = M^S(i, M^D)$. Estruturalistas acrescentam ainda a gestão de ativos e passivos dos bancos, alegando que “mesmo que a autoridade monetária se recusasse a acomodar aumentos na demanda por reservas, os bancos ainda poderiam acomodar parcialmente uma elevação na demanda por empréstimos a partir de suas próprias iniciativas” (Palley, 2002). As duas abordagens, no entanto, resultam em uma mesma dinâmica de endogeneidade (e não-neutralidade) da moeda.

O entendimento de endogeneidade da moeda, pelo lado dos agentes não-financeiros, relaciona-se a uma análise da teoria da preferência pela liquidez, originalmente abordada na *Teoria Geral* de Keynes. Parte-se da ideia de que a decisão de investir é basicamente uma escolha de portfólios (ou “carteiras”), onde ativos de capital sofrem a concorrência de outros ativos (inclusive monetários) como objeto de acumulação. Como a moeda oferece uma segurança contra a incerteza do futuro, mudanças em expectativas podem levar os agentes a concentrar sua demanda em ativos líquidos, como a moeda, em detrimento de ativos de capital. Assim, produção e acumulação de capital não se destinam necessariamente à maximização de uma função objetivo claramente definida⁶⁰. Mercados são incertos, o que eleva a possibilidade de decepção para o caso de investimentos em ativos de capital (que envolvem bens de alto valor e grande durabilidade). Isto justifica comportamentos defensivos e cautelosos, como a demanda por ativos que, mesmo com baixo retorno, tenham melhor liquidez (o que seria irracional em um contexto diverso).

⁶⁰ Minsky (1986) chama a teoria neoclássica de “paradigma da feira de aldeia” (em contraposição ao “paradigma de Wall Street”). Na “feira de aldeia”, os agentes isolados interagem pela troca de bens. Para ele, ao desconsiderar a presença da moeda e do sistema financeiro, tal abordagem torna-se irrealista para o estudo do capitalismo moderno.

A incerteza⁶¹, portanto, afeta os balanços dos agentes, tanto na escolha do portfólio de ativos como na escolha das formas de financiamento de suas posições. A decisão de portfólio é uma escolha entre ativos que ofereçam alta liquidez mas, possivelmente, baixos retornos esperados e os que, ao contrário, compensem sua baixa liquidez com a perspectiva de retornos elevados. Por sua vez, as expectativas com relação aos rendimentos futuros dos investimentos são importantes também na determinação do passivo dos agentes econômicos, que devem decidir se financiam suas compras de ativos com rendas próprias ou recorrendo a fundos de terceiros. E, no caso de utilização de recursos de terceiros, sob que formas, prazos e termos.

Compreendidos esse possível *trade-off* entre liquidez e retornos nas carteiras de ativos e a “irracional” preferência pela liquidez por parte dos agentes, pode-se analisar o efeito dessa dinâmica sobre os resultados de renda (demanda agregada) e taxa de juros.

Quando a preferência pela liquidez aumenta, os preços dos ativos se ajustam de modo que aqueles cujo benefício é mais relacionado à liquidez aumentam em comparação aos daqueles cujos principais benefícios são os lucros esperados. Isso reduz os retornos esperados totais do detentor de ativos pouco líquidos (ativos de capital, via de regra mais produtivos), diminuindo seus preços demandados e, portanto, a eficiência marginal do capital. Além disso, um aumento na preferência pela liquidez seria associado a uma queda na demanda por ativos de capital, queda nos investimentos e demanda agregada, retração nos empréstimos e depósitos bancários, e aumento nas taxas de juros. Ou seja, um aumento na preferência pela liquidez é “contado duplamente”: primeiro, pela queda na eficiência marginal do capital; depois, pela elevação nas taxas de juros dos empréstimos (Wray, 2004).

⁶¹ Incerteza, aqui, não se refere ao conceito tradicional de risco, de probabilidades (objetivas ou subjetivas) ou possibilidade de compra de seguros (conforme o conhecido modelo Arrow-Debreu). Incerteza, aqui, significa não saber o que ocorrerá no futuro e nem mesmo o que *poderá* ocorrer (donde funções de distribuição de probabilidades não podem sequer ser determinadas, já que o universo de eventos não é cognoscível).

Mecanismo análogo vale para uma diminuição na preferência pela liquidez, que resulta em aumento nos investimentos e demanda agregada e redução nas taxas de juros.

Pela explicação da “demanda por moeda”, corrente nos livros-texto, aumento na preferência pela liquidez significa elevação na “demanda por moeda” e, já que, por essa abordagem, a oferta de moeda é exogenamente definida pelo banco central, o resultado de maior demanda para uma mesma oferta de moeda seria um aumento no “preço da moeda”, ou seja, na taxa de juros.

O que diferencia as duas abordagens é a divergência sobre se as mudanças na preferência pela liquidez na sociedade são naturalmente “acomodadas” pelo setor bancário – e, em última instância, pela autoridade monetária –, hipótese que “endogeniza” a oferta de moeda, ou se o banco central detém para si o controle da oferta de moeda, hipótese que a torna uma variável exógena. A discrepância conclusiva entre as duas abordagens mantém-se presente no debate econômico por causa da dificuldade em se “medir” preferência pela liquidez e fazer testes empíricos a partir daí.

A lógica da “acomodação” natural pelo banco central está no cerne do pensamento pós-keynesiano e sua estrutura é a seguinte. A moeda ofertada pelo setor bancário é sempre uma relação entre dois agentes: o banco compra um compromisso de pagamento do tomador e gera um depósito à vista; o tomador emite um compromisso de pagamento e recebe um depósito bancário. Daí, há quatro destinos possíveis: o depósito pode ser voluntariamente mantido pelo recebedor; pode ser sacado (convertido em papel-moeda); pode ser transferido a outro recebedor; ou pode ser usado para cobrir um empréstimo bancário (nova operação entre dois agentes). A conversão do depósito em papel-moeda leva a uma “zeragem das reservas”⁶² mas esta “zeragem” é sempre acomodada pelo banco central, por meio de reposição (temporariamente) a descoberto⁶³.

⁶² “*Net reserve clearing*”.

⁶³ “*Overdraft*”.

Por esse entendimento, movimentos de taxas de juros nada têm a ver com forças equilibrando oferta de poupança e demanda por investimento. Tal como Keynes enfatizou, a poupança sempre se iguala ao investimento porque é o fluxo de despesas que determina a renda e, em particular, é a despesa com produto indisponível que determina a parcela de renda não consumida. Qualquer que seja a despesa, ela deve ser financiada, independentemente dos fluxos de renda que sejam gerados por ela. E mesmo que uma crescente tomada de empréstimo para financiar gastos crescentes pressione as taxas de juros dos empréstimos, tudo isso ocorre antes dos fluxos de despesas, renda e poupança sobre os quais a teoria dos fundos emprestáveis se concentra.

Em todas as economias modernas, a moeda do banco central é usada para fazer a liquidação entre as contas dos bancos e com o governo, e também para cobrir os saques de papel-moeda. Não convém ao banco central recusar-se a prover reservas (via “*overdraft*”) ante a necessidade de um sistema de liquidação entre contas⁶⁴. Além disso, a demanda por reservas é altamente inelástica a juros, donde reservas em excesso rapidamente empurram as taxas “*overnight*”⁶⁵ para zero, enquanto reservas insuficientes provocam rápidas elevações nas taxas. Como o banco central não pode permitir que as “forças de mercado” imponham taxas “*overnight*”, sua ação é definir uma meta para tal taxa e acomodar toda a demanda sob esta taxa⁶⁶.

Nessa dinâmica da preferência pela liquidez por parte dos agentes não-financeiros e da necessidade do banco central em acomodar a demanda do setor financeiro, fundamenta-se a ideia de causalidade endógena entre a demanda por crédito bancário e a oferta de moeda.

A moeda endógena tem três grandes implicações: a não-causalidade entre agregados monetários e problemas macroeconômicos; a ineficiência da tentativa de controle da economia via oferta de moeda no sentido

⁶⁴ Desde que não haja choques no sistema (ou seja, que as variações em suas variáveis sejam suaves) e que não seja objetivo do banco central operar com taxa “*overnight*” igual a zero (como no caso do Banco do Japão).

⁶⁵ Taxas de juros de curtíssimo prazo, de um dia para o outro.

⁶⁶ No caso brasileiro, trata-se da taxa Selic.

monetarista⁶⁷; e o fato de as flutuações endógenas de oferta de moeda poderem gerar instabilidade e serem importantes nos ciclos reais.

Com tais conceitos em mente, segue-se para a discussão se a adoção continuada de cartões de crédito e débito reflete-se na transmissão da política monetária.

4.4 Efeitos dos instrumentos eletrônicos de pagamento sobre a política monetária

O primeiro passo aqui são as definições dos termos empregados. Neste trabalho, chamam-se “instrumentos eletrônicos de pagamento” os pagamentos de varejo, ou seja, feitos geralmente por pessoas físicas e com valores financeiros unitários baixos, cujo instrumento tenha seu valor fiduciariamente atrelado a um meio eletrônico (uma base de dados conectada a uma rede, um *chip* ou uma tarja magnética, por exemplo), e não a um meio físico (papel-moeda ou cheque). Cartões de crédito e débito fazem parte desse conceito, bem como cartões de loja e cartões de serviços, transferências eletrônicas, DOCs e TEDs de baixo valor. Em alguns trabalhos, esse grupo é chamado de *e-money*, *e-cash*, *e-purse* ou “moeda eletrônica”. No entanto, como o BIS define formalmente *e-money* como “valor armazenado”⁶⁸ em um dispositivo com fim específico, o que exclui do grupo cartões de crédito e débito, prefere-se aqui usar o termo mais geral “instrumentos eletrônicos de pagamento”. Outra distinção importante a ser feita é entre o instrumento de pagamento em si e seu canal de acesso. Pagamentos via internet, via ATM, ou via telefone podem certamente valer-se de algum instrumento eletrônico (cartão de crédito, de débito ou transferência eletrônica). Mas, nesse caso, a internet, o ATM ou o telefone em si não se caracterizam como o instrumento, e sim como o canal de acesso. Daí alguns trabalhos alertarem para que não se confundam *e-money* e

⁶⁷ Nas palavras de Palley (2002), “a lei de Goodhart (atribuída a Charles Goodhart), segundo a qual qualquer relação estável entre um agregado monetário e a atividade econômica cessará se a autoridade monetária tentar controlar o agregado, é a versão cínica da ideia pós-keynesiana de que é impossível controlar a oferta de moeda”.

⁶⁸ CPSS/BIS (2003): “*value stored electronically in a device such as a chip card or a hard drive in a personal computer*”.

e-banking (que seria o uso de canais de acesso eletrônicos para serviços bancários e de pagamentos).

Desde o início da década 2000, tem-se discutido muito acerca dos efeitos sobre a política monetária de um tipo de instrumento eletrônico de pagamento caracterizado por armazenar valor ("*e-money*", segundo a definição do BIS, ou "*private e-money*" como preferem chamar os autores, para distingui-lo do "*government e-money*") emitido, de modo pré-pago ou na forma de crédito, e liquidado pela própria instituição emitente do instrumento, que, dependendo da forma de adoção, poderia tirar do banco central o papel de monopolista natural na emissão de moeda (Fullenkamp e Nsouli, 2004; Woodford, 2000; Sardoni e Verde, 2002; Sardoni, 2006).

Os modelos simulam uma drástica redução da base monetária, considerando que, com o uso disseminado desses instrumentos, a população abriria mão do papel-moeda e os bancos, dos depósitos à vista. Concluem, daí, que os efeitos significativos de política monetária de bancos centrais se dão pelos mecanismos que tornam a taxa de juros praticada pela autoridade monetária (em operações de *open market* e redesconto) a taxa base da economia e que, portanto, a drástica redução na base monetária não implicaria perda, mas apenas redução, no "poder de mercado" por parte dos bancos centrais. A justificativa dessa "redução de poder" é que, com o *private e-money* (da mesma forma que com o *e-banking*), reduz-se a necessidade de as pessoas usarem moeda em espécie (*cash*) e, com isso, aumenta a elasticidade da demanda por moeda em relação à taxa de juros, tornando a curva LM mais horizontal e, assim, exigindo dos bancos centrais variações menores nas taxas de juros para obter o mesmo efeito sobre o produto.

Uma abordagem "monetarista" pode ser feita a partir da teoria quantitativa da moeda, que se baseia na identidade $Mv = PY$, em que M é um agregado monetário, v é a velocidade de circulação da moeda, P é o nível de preços e Y é o produto. O uso mais intensivo de instrumentos eletrônicos de pagamento provocaria elevação em v , obrigando a autoridade monetária a reduzir M para recompor o equilíbrio PY .

Por uma abordagem de moeda endógena, pode-se olhar especificamente para o uso de cartões de crédito como uma forma de formação de dívida (caso de usuários-revólver) e questionar os efeitos de um aumento do uso desse instrumento sobre as taxas de juros e, via canais de transmissão anteriormente discutidos, sobre demanda agregada e inflação. Por essa visão, se o uso de cartões de crédito aumentar ou reduzir o desejo por reservas no sistema bancário, o banco central acomodará a mudança. É razoável, então, supor que aumentos cíclicos no uso de cartões de crédito estejam associados a um aumento das taxas de juros de cartões de crédito, caso o uso por tomadores *sub-prime* (ou mesmo a simples percepção de risco) aumente. Entretanto, a conexão entre o uso de cartões de crédito e as taxas de empréstimo ou depósito bancário é ambígua. Se a “sabedoria popular” diz que taxas de juro devem cair por causa do “declínio na demanda por moeda” decorrente do uso de cartões de crédito, ela se baseia em uma ideia equivocada sobre a natureza da moeda bancária. Se o uso de cartões de crédito substituir empréstimos, então poucos empréstimos serão feitos mas, da mesma forma, poucos depósitos serão oferecidos. Já que a criação de moeda bancária é uma relação entre dois agentes, ela nunca leva a excesso nem insuficiência de depósitos. Contrair dívida via proprietário de esquema de cartão de crédito seria, do ponto de vista qualitativo, equivalente a fazê-lo via empréstimo bancário⁶⁹, o que leva à conclusão, por essa linha de pensamento, de que não é o maior uso de cartões que causa impacto sobre taxas de juros (e sobre seus efeitos no produto e na inflação), mas sim os mecanismos de crédito de maneira geral.

Em suma, os efeitos esperados do uso de instrumentos eletrônicos de pagamento sobre a política monetária variam segundo três abordagens aqui discutidas: preferência pela liquidez, teoria quantitativa da moeda e moeda endógena.

⁶⁹ No mundo real, dois aspectos fazem as situações “empréstimo bancário” e “uso do crédito via cartão” distinguirem-se bastante. O primeiro, muito relevante para usuários-revólver, é que diferentes estruturas de concorrência no setor bancário e no setor cartões de crédito levam a diferentes taxas de juros praticadas por cada setor. O segundo, muito importante para usuários de conveniência, é o período livre de juros entre a data da compra e o pagamento da fatura (*grace period*).

Na abordagem da preferência pela liquidez, o maior uso de tais instrumentos retrai a preferência pela liquidez e, conseqüentemente, a elasticidade da demanda por moeda em relação à taxa de juros, reduzindo a eficiência da política monetária. Ou seja, se as variações de taxas de juros praticadas pelos bancos centrais mantiverem-se nos patamares de épocas com menor uso de instrumentos eletrônicos de pagamento, o efeito dessas variações será menor.

Na abordagem da teoria quantitativa da moeda, estes instrumentos elevam a velocidade de circulação da moeda e obrigam a autoridade monetária a reduzir a oferta de moeda (ou induzir sua redução), a fim de evitar elevação de preços na economia (considerando-se o produto como fixo, dado pelo pleno emprego).

Na abordagem da moeda endógena, o maior uso de instrumentos eletrônicos de pagamento terá efeito sobre a política monetária se tal uso envolver tomada de crédito, como no caso de uso do crédito rotativo em cartões de crédito (usuários-revólver). Mas será o mesmo efeito de um aumento no crédito de maneira geral.

Medições e estimações desses efeitos são um interessante e importante tema para pesquisa futura.

4.5 Notas conclusivas sobre política monetária e instrumentos eletrônicos de pagamento

Ainda que não haja consenso sobre o reflexo do uso mais intenso de instrumentos eletrônicos de pagamento sobre as taxas de juros, algumas lições podem ser tiradas da discussão.

Uma delas é a busca de entendimento o melhor detalhado possível dos mecanismos de transmissão de política monetária, bem como da mensuração e/ou estimação do efeito de cada canal de transmissão. Uma compreensão detalhada de tais mecanismos permite concluir quais deles podem ser afetados pelo uso de instrumentos eletrônicos de pagamento, e a mensuração da “força” dos canais afetados permite conclusões acerca do grau de relevância desse efeito.

Outra conclusão é a importância de periodicamente bancos centrais mensurarem o impacto da moeda eletrônica sobre a elasticidade da demanda por moeda em relação à taxa de juros.

5 CONCLUSÃO

Este trabalho ocupou-se de discutir a crescente tendência de uso de instrumentos eletrônicos de pagamento. Para tanto:

- inseriu o tema dentro das preocupações de bancos centrais com seus sistemas de pagamento (e, em particular, pagamentos de varejo);
- apresentou uma breve evolução histórica do desenvolvimento de tais instrumentos (incluindo suas tendências esperadas para o futuro próximo);
- mostrou a dinâmica do mercado de instrumentos eletrônicos em particular (cartões de crédito e de débito) e a fundamentação microeconômica atualmente mais empregada para abordar o tema, a teoria do mercado de dois lados;
- propôs exercícios econométricos sobre um painel trimestral de dados por instituição financeira referente ao uso de cartões de crédito e de débito no Brasil;
- propôs uma discussão acerca da inter-relação entre funções de bancos centrais sobre sistema de pagamentos e política monetária (e também supervisão bancária), inserindo os instrumentos eletrônicos de pagamento nesse contexto;
- apresentou o atual estágio das pesquisas sobre transmissão de política monetária; e
- discutiu os efeitos da adoção de instrumentos eletrônicos de pagamento sobre a política monetária, à luz da teoria da preferência pela liquidez, da teoria quantitativa da moeda e da noção de moeda endógena.

Como principais contribuições, o trabalho apresentou:

- um reforço empírico sobre resultado teórico segundo o qual a taxa de intercâmbio funciona como um “preço de equilíbrio” em mercados de

cartões de crédito (mercado de 2 lados quando estruturado em plataforma de 4 partes);

- um reforço empírico sobre a ideia de retroalimentação entre valor financeiro movimentado em cartões e número de estabelecimentos credenciados a aceitá-los;
- um novo desenho esquemático geral para os canais de transmissão de política monetária, que mostra seus efeitos reais e nominais e que traz para o âmbito da autoridade monetária, além do controle sobre a base monetária e sobre a taxa de juros nominal de curto prazo, um controle sobre as reservas (via recolhimentos compulsórios e redescontos) e algum controle sobre a taxa de câmbio (ao considerar a comum prática de regimes cambiais *dirty floating*); e
- uma conclusão sobre a relevância de bancos centrais monitorarem regularmente a elasticidade da demanda por moeda à taxa de juros, a despeito das divergências teóricas sobre o impacto que a adoção de instrumentos eletrônicos de pagamento possam acarretar sobre a eficiência da política monetária.

Finalmente, como estímulos a trabalhos posteriores, tem-se:

- outras especificações de modelo, por exemplo, com outras variáveis, com instrumentos ou com variáveis exógenas no VAR;
- estimações GMM e Arellano-Bond, que possam corrigir os problemas de viés dos estimadores por conta de endogeneidade;
- análises de estacionaridade, que possam indicar de maneira mais acurada as tendências das variáveis;
- trabalho quantitativo que relacione ou teste relações de causalidade entre adoção do uso de cartões de crédito e de débito e efeitos da política monetária.

REFERÊNCIAS

ANGELONI, I. *et al.* **Monetary transmission in the Euro Area: where do we stand.** Frankfurt am Main: European Central Bank, 2002. (Working Paper, 114).

ARESTIS, Philip; SAWYER, Malcolm. **Can monetary policy affect real economy?**. The Levy Economic Institute of Bard College, 2002. (Working Paper, 355). Disponível em: <<http://www.levy.org/pubs/wp355.pdf>>. Acesso em: 8 jan. 2010.

ARESTIS, Philip; SAWYER, Malcolm. The nature and role of monetary policy when money is endogenous. **Cambridge Journal of Economics**, London, v. 30, n. 6, p. 847-860, Aug. 2006.

BANCO CENTRAL DO BRASIL. **Competição e cooperação na indústria de cartões de pagamentos: análise do mercado de credenciadores.** 2005. Trabalho não publicado.

BANCO CENTRAL DO BRASIL. **Custo e eficiência na utilização de instrumentos de pagamento de varejo.** 2007. Disponível em: <<http://www.bcb.gov.br/htms/novaPaginaSPB/Nota%20T%E9cnica%20-%20Custo%20Eficiencia.pdf>>. Acesso em: 8 jan. 2010.

BANCO CENTRAL DO BRASIL. **Modernização dos instrumentos de pagamento de varejo no Brasil.** 2005. Disponível em: <<http://www4.bcb.gov.br/pec/appron/Apres/Diagn%F3sticoImprensa1.pdf>>. Acesso em: 8 jan. 2010.

BANCO CENTRAL DO BRASIL. **Relatórios sobre a competição no mercado de cartões de pagamento – Reino Unido, Austrália, Holanda, México e Chile.** 2006. Trabalho não publicado.

BANCO CENTRAL DO BRASIL. **Uso dos instrumentos de pagamento.** 2006. Disponível em: <<http://www.bcb.gov.br/?SPBEST2005>>. Acesso em: 8.jan. 2010.

BANCO CENTRAL DO BRASIL. Secretaria de Acompanhamento Econômico. **Relatório sobre a indústria de cartões de pagamento.** 2009. Disponível em: <<http://www.bcb.gov.br/?SPBINSTPAG>>. Acesso em: 8 jan. 2010.

BANK FOR INTERNATIONAL SETTLEMENTS. Committee on Payment and Settlement Systems. **A glossary of terms used in payments and settlement systems.** 2003. Disponível em: <www.bis.org/publ/cpss00b.pdf>. Acesso em: 8 jan. 2010.

BANK FOR INTERNATIONAL SETTLEMENTS. **Implications for central banks of the development of electronic money.** 1996. Disponível em: <www.bis.org/publ/bisp01.pdf>. Acesso em: 8.jan. 2010.

BAXTER, W. F. Bank interchange of transactional paper: legal perspectives. **Journal of Law and Economics**, Chicago, v. 26, n. 3, p.541-588, Oct. 1983.

BERNANKE, Ben; GERTLER, Mark. Agency costs, net worth, and business fluctuations. **American Economic Review**, Nashville, v. 79, n. 1, p. 14-31, Mar. 1989.

BLACK, Fischer. Bank fund management in an efficient market. **Journal of Financial Economics**, Lausanne, v. 2, n. 4, p. 323-339, 1975.

BLANCHARD Olivier. **Fiscal dominance and inflation targetting**: lessons from Brazil. Cambridge: National Bureau of Economic Research, 2004. (NBER Working Paper, 10389). Disponível em: <http://papers.nber.org/papers/w10389.pdf?new_window=1>. Acesso em: 8 jan. 2010.

BLINDER, Alan. **Central bank in theory and practice**. Cambridge: MIT Press, Cambridge, 1998.

BOGDANSKI, Joel; TOMBINI, Alexandre A.; WERLANG, Sérgio R. C. **Implementing inflation targeting in Brazil**. Brasília: Banco Central do Brasil, 2000. (Working Paper Series, 1).

BORZEKOWSKI, Ron; KISER, Elizabeth K. **The choice at the checkout**: quantifying demand across payment instruments. Washington: Board of Governors of the Federal Reserve System, 2006. (Finance and Economics Discussion Series).

BORZEKOWSKI, Ron; KISER, Elizabeth K.; AHMED, Shaista. **Consumers' use of debit cards**: patterns, preferences, and price response. Washington: Board of Governors of the Federal Reserve System, 2006. (Finance and Economics Discussion Series).

BORZEKOWSKI, Ron *et al.* Trends in the use of payment instruments in the United States. **Federal Reserve Bulletin**, Washington, v. 91, n. 2, p. 180-201, Springer 2005.

BRASIL.Ministério da Fazenda. Secretaria de Acompanhamento Econômico. **Ato de concentração**. 2006

BRASIL.Ministério da Fazenda. Secretaria de Acompanhamento Econômico. **Ato de concentração 08012.002208/2004-78**. 2004.

CARLTON, Dennis W.; FRANKEL, Alan. S. The antitrust economics of credit card networks. **Antitrust Law Journal**, v. 63, n. 2, p. 643-668, May 1995.

CHAKRAVORTI, Sujit; MHCHUGH, Timothy. Why do we use so many checks? Chicago, **Economic Perspectives**, v.26, n. 3, p. 44-59, 3rd quarter, 2002.

CHAKRAVORTI, Sujit; TO, Ted. **A theory of credit cards**. **International Journal of Industrial Organization**, Amsterdam, v. 25, n. 3, p. 583-585, June, 2007. Disponível em: http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6V8P-4KM46MG-1&_user=1561396&_coverDate=06%2F30%2F2007&_rdoc=1&_fmt=high&_orig=search&_sort=d&_docanchor=&_view=c&_searchStrId=1246245162&_rerunOrigin=scholar.google&_acct=C000053641&_version=1&_urlVersion=0&_userid=1561396&_md5=a39de0db566573aa8198051be9d77f32>. Acesso em: 8 jan. 2010.

CLARIDA, Richard; GALÍ, Jordi; GERTLER, Mark. The science of monetary policy: a new keynesian perspective. **Journal of Economic Literature**, Nashville, v. 37, n. 4, p. 1661-1707, Dec. 1999.

COWEN, Tyler; KROSZNER, Randall. The development of the new monetary economics. **The Journal of Political Economy**, Chicago, v. 95, n. 3, p. 567-590. July 1987.

EAGLETON, Catherine; WILLIAMS, Johnathan. **Money: a history**. Buffalo, USA: Firefly Books, 2007.

ENDERS, Walter. **Applied econometrics time series**. 2nd ed. Hoboken, USA: Wiley, 2004.

EVANS, David S.; SCHMALENSEE, Richard. The economics of interchange fees and their regulation: an overview. Cambridge, USA: MIT Sloan, 2005. (Working Paper, 4548-05).

EVANS, David; SCHMALENSEE, Richard. **The industrial organization of markets with two-sided platforms**. Cambridge: National Bureau of Economic Research, 2005. (NBER Working Paper, 11603).

FONTANA, Giuseppe. Rethinking endogenous money: a constructive interpretation of the debate between horizontalists and structuralists. **Metroeconomica**, Oxford, v. 55, n. 4, p. 367-385, Nov. 2004

FONTANA, Giuseppe; PALACIO-VERA, Alfonso. Are long-run price stability and short-run output stabilization all that monetary policy can aim for? **Metroeconomica**, Oxford, v. 58, n. 2, p.269-298, May 2007.

FRIEDMAN, Milton. The role of monetary policy. **American Economic Review**, Nashville, v. 58, n. 1, p. 1-17. Mar. 1968.

FULLENKAMP, Connel; NSOULI, Saleh M. **Six puzzles in electronic money and banking**. Washington: International Monetary Fund, 2004. (IMF Working Paper, 04/19).

GERDES, Geoffrey R.; WALTON, Jack K. The use of checks and other noncash payment instruments in the United States. **Federal Reserve Bulletin**, Washington, v. 88, n. 8, p. 360-374, Aug, 2002.

GREENE, William H. **Econometric analysis**. 5th ed. New Jersey: Prentice Hall, 2003.

GUTHRIE, Graeme; WRIGHT, Julian. **Competing payment schemes**. Auckland: University of Auckland, Department of Economics, 2003. (Working Paper, 245).

HAMILTON, James D. **Time series analysis**. Princeton: Princeton University Press, 1994.

HAYASHI, Fumiko. **Network competition and merchant discount fees**. Kansas City: Federal Reserve Bank, 2005.

HAYASHI, Fumiko; KLEE, Elizabeth C. **Technology adoption and consumer payments: evidence from survey data**. Kansas City: Federal Reserve Bank, 2002. (Payments System Research Working Paper. PSR WP 02-01).

HESTER, Donald. Financial desintermediation and policy. **Journal of Money, Credit and Banking**, Columbus, v. 1, n. 3, p. 600-617, Aug.1969.

HESTER, Donald. Monetary policy in the "checkless" economy. **The Journal of Finance**, New York, v. 27, n. 2, p. 279-293, May 1972.

HESTER, Donald. On the adequacy of policy instruments and information when the meaning of money is changing. **American Economic Review**, Nashville, v. 72, n. 2, p. 40-44, May 1982.

HOLTHAUSEN, Cornelia; MONNET, Cyril. **Money and payments: a modern perspective**. Frankfurt am Main: European Central Bank, 2003. (Working Paper, 245).

IM, Kyung So; PESARAN, M. Hashem; SHIN, Yongcheol. Testing for unit roots in heterogeneous panels. **Journal of Econometrics**, Amsterdam, v. 115, n. 1, p. 53-74, July 2003.

KATZ, Michael. **Reform of credit card schemes in Australia II: network effects, interchange fees, and no-surcharge rules in the Australian credit and charge card industry**. Sydney: Reserve Bank of Australia, 2001. II Comissioned Report,

KAUFMAN, C. Bank contagion: a review of the theory and evidence. **Journal of Financial Services Research**, Boston, v. 8, n. 2 , p. 123-150, Apr. 1994.

KEYNES, John Maynard. **The general theory of employment, interest, and money**. Orlando: First Harvest/Harcourt, 1964.

KLEE, Elizabeth C. **Families' use of payment instruments during a decade of change in the U.S. payment system**. 2006. (Finance and Economics Discussion Series, 2006-01). Disponível em: <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=899115>. Acesso em: 8 jan. 2010.

KUTTNER, Kenneth N.; MOSSER, Patricia C. The monetary transmission mechanism: some answers and further questions. **Economic Policy Review**, New York, v. 8, n. 1 , p. 15-24, May 2002.

LEVIN; Andrew; LIN, Chien-Fu; CHU, Chia-Shang James. Unit root tests in panel data: asymptotic and finite-sample properties **Journal of Econometrics**, Amsterdam, v. 108, n. 1, p. 1-24, May 2002.

LIBÂNIO, Gilberto. Unit roots in macroeconomic time series: theory, implications, and evidence. **Nova Economia**, Belo Horizonte, v. 15, n. 3, p. 145-176, dez. 2005.

LITAN, Richard. **Institutions and policies for maintaining financial stability, research in financial services**. 1997. Disponível em: <<https://www.kansascityfed.org/publicat/sympos/1997/pdf/s97litan.pdf>>. Acesso em: 8 jan. 2010.

MINSKY, Hyman P. **Stabilizing an unstable economy**. New Haven: Yale University Press, 1986. cap. 1, 10, 12. (A Twentieth Century Fund Report).

MODIGLIANI, F. Monetary policy and consumption. In: _____. **Consumer spending and monetary policy: the linkages**. Boston: Federal Reserve Bank of Boston, 1971. p. 9-84. (Conference Series, 5).

MUINHOS, Marcelo Kafoury.; ALVES, Sergio, Afonso Lago. **Medium-size macroeconomic model for the Brazilian economy**. Brasília: Banco Central do Brasil, 2003. (Working Paper Series, 64).

PALLEY, Thomas I. Endogenous money: what it is and why it matters. **Metroeconomica**, Oxford, v. 53, n. 2, p. 152-180, Nov. 2002.

PALLEY, Thomas I. The e-money revolution: challenges and implications for monetary policy. **Journal of Post-Keynesian Economics**, Armonk, v. 24, n. 2, p. 217-239. July 2002.

POOLE, William. Optimal choice of monetary policy instruments in a simple stochastic macro model. **Quarterly Journal of Economics**, Cambridge, v. 84, n. 2, p. 197-216, May 1970.

ROCHET, Jean-Charles; TIROLE, Jean. Tying in two-sided markets and the honor all cards rule. **International Journal of Industrial Organization**, Amsterdam, v. 26, n.6, p. 1333-1347, Nov. 2008. Disponível em: , http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6V8P-4RP0MDB-1&_user=1561396&_coverDate=11%2F30%2F2008&_rdoc=1&_fmt=high&_orig=search&_sort=d&_docanchor=&view=c&_searchStrId=1246242087&_rerunOrigin=scholar.google&_acct=C000053641&_version=1&_urlVersion=0&_userid=1561396&md5=44655bd55bb3062f7cab4dc279cb2fea>. Acesso em: 08 jan. 2010.

ROCHET, Jean-Charles; TIROLE, Jean. An economic analysis of the determination of interchange fees in payment card systems. **Review of Network Economics**, Toulouse, v. 2, n. 2, p. 69-79, June, 2003.

ROCHET, Jean-Charles; TIROLE, Jean. Cooperation among competitors: the economics of payment card associations. **Rand Journal of Economics**, Toulouse, v. 33, n. 4, p. 549-570, Winter 2002.

ROCHET, Jean-Charles; TIROLE, Jean. Platform competition in two-sided markets. **Journal of European Economic Association**, Toulouse, v. 1, n. 4, p. 990-1029, June 2003.

ROCHET, Jean-Charles; TIROLE, Jean. **Two-sided markets**: an overview. 2004. Disponível em: <http://faculty.haas.berkeley.edu/HERMALIN/rochet_tirole.pdf>. Acesso em: 8 jan. 2010.

SARDONI, Claudio. **Why central banks (and money) “rule the roost”**. Roma: The Levy Economics Institute of Bard College, 2006. (Working Paper, 457).

SARDONI, Claudio; VERDE, Alessandro. The “IT revolution” and the monetary system: electronic money and its effects. Roma: University of Rome “La Sapienza”, 2002.

SCHMALENSEE, Richard. Payment systems and interchange fees. **Journal of Industrial Economics**, Oxford, v. 50, n. 2, p. 103-122, June 2002.

TAYLOR, John B. Discretion versus policy rules in practice. **Carnegie-Rochester Conference Series on Public Policy**, v. 39, n. 1, p. 195-214, Dec. 1993.

THEIL, Henri. **Economic Forecasts and Policy**. Amsterdam: North-Holland, 1961.

TINBERGEN, Jan. **On the theory of economic policy**. Amsterdam: North-Holland, 1952.

TOBIN, James. A general equilibrium approach to monetary theory. **Journal of Money Credit and Banking**, Ohio, v. 1, n. 1, p. 15-29, Feb. 1969.

WANG, Zhu. **Market structure and credit card pricing**: what drives the interchange? Kansas City: Federal Reserve Bank of Kansas City, 2006. (Working Paper, 06-04).

WOODFORD, Michael. **Monetary policy in a world without money**. 2000 (NBER Working Paper, 7853). Disponível em: <<http://www.nber.org/papers/w7853>>. Acesso em: 8 jan. 2010

WOOLDRIDGE, Jeffrey. M. **Econometric analysis of cross section and panel data**. Cambridge: Mit Press, 2002.

WOOLDRIDGE, Jeffrey. M. **Introductory econometrics**. England: Mit Press, 2006.

WRAY, L. Randall. Loanable funds, liquidity preference, and endogenous money: do credit cards make a difference? **Journal of Post-Keynesian Economics**, Armonk, v. 26, n. 2, p. 309-323, Winter 2004.

WRIGHT, Julian. Optimal card payment systems. **European Economic Review**, Auckland, v. 47, n. 4, p. 587-612. Aug. 2003.

WRIGHT, Julian. Pricing in debit and credit card schemes. **Economics Letters**, Amsterdam, v. 80, n 3, p. 305-309, Sept. 2003.