

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE CIÊNCIAS ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

Análise empírica da formação de expectativas de inflação no Brasil
uma aplicação de redes neurais artificiais a dados em painel

Andreza Aparecida Palma
Orientador: Prof. Dr. Marcelo Savino Portugal

Porto Alegre
2007

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE CIÊNCIAS ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

Análise empírica da formação de expectativas de inflação no Brasil
uma aplicação de redes neurais artificiais a dados em painel

Andreza Aparecida Palma
Orientador: Prof. Dr. Marcelo Savino Portugal

Dissertação submetida ao programa de Pós-Graduação em Economia Aplicada da Faculdade de Ciências Econômicas da UFRGS, como quesito parcial para a obtenção do grau de mestre em Economia – Ênfase em Economia Aplicada.

Porto Alegre
2007

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Responsável: Biblioteca Gládis W. do Amaral, Faculdade de Ciências Econômicas da
UFRGS

P171a Palma, Andreza Aparecida
Análise empírica da formação de expectativas de inflação no Brasil /
Andreza Aparecida Palma. – Porto Alegre, 2007.
111 f. : il.

Orientador: Marcelo Savino Portugal.

Ênfase em Economia Aplicada.

Dissertação (Mestrado em Economia) - Universidade Federal do Rio Grande do Sul, Faculdade de Ciências Econômicas, Programa de Pós-Graduação em Economia, Porto Alegre, 2007.

1. Inflação : Brasil. I. Portugal, Marcelo Savino. II. Universidade Federal do Rio Grande do Sul. Faculdade de Ciências Econômicas. Programa de Pós-Graduação em Economia. III. Título.

CDU 336.748.12 (81)

Análise empírica da formação de expectativas de inflação no Brasil
uma aplicação de redes neurais artificiais a dados em painel

ANDREZA APARECIDA PALMA

Dissertação submetida ao programa de Pós-Graduação em Economia Aplicada da Faculdade de Ciências Econômicas da UFRGS, como quesito parcial para a obtenção do grau de mestre em Economia – Ênfase em Economia Aplicada.

Aprovada em: Porto Alegre, 20 de julho de 2007.

Prof. Dr. Marcelo Savino Portugal – orientador
UFRGS

Prof. Dr. Alexandre Sartoris Neto
UNESP

Prof. Dr. Flávio Augusto Ziegelmann
UFRGS

Prof. Dr. Ronald Otto Hillbrecht
UFRGS

*Aos meus pais, Cláudio e Maria
À minha irmã, Débora*

*When the only tool you have is a hammer,
every problem you encounter tends to resemble a nail.*

Autor desconhecido

AGRADECIMENTOS

O curso de mestrado foi um período de aprendizado muito intenso, em todos os sentidos. E muitas pessoas participaram desse processo.

As palavras são insuficientes para agradecer toda a atenção dispensada pelo professor Marcelo Portugal à mim durante toda a minha trajetória no curso de mestrado na UFRGS, especialmente durante a realização dessa dissertação. Ele me “aturou” desde a primeira semana em que cheguei em Porto Alegre (e essa não foi uma tarefa das mais fáceis...rs..). Agradeço por todos os seus ensinamentos, conselhos e ajuda. E deixo registrado aqui o quanto eu o admiro como economista, professor e como pessoa.

Aos demais professores do curso de mestrado, especialmente Flávio Ziegelmann, Eduardo Pontual, Ronald e Jorge Araújo.

Agradeço também, de forma especial, ao prof. Alexandre Sartoris, sempre disposto a ajudar e a esclarecer qualquer dúvida. À ele e à sua esposa Fabiana, pelo apoio constante, sem o qual eu nem teria feito o mestrado em POA.

Não poderia deixar de agradecer ao professor Jeffrey Racine, pela atenção aos meus e-mails e pelo esclarecimento de pontos importantes em relação ao método de inferência estatística utilizado aqui. Sem esta ajuda, este trabalho não teria sido viável.

Aos colegas do curso de mestrado: Tanara, Nina, Mari Hauer, Túlio, Naka (Maurício), Gustavo, Anchieta, Aline, Felipe, Dani, Gilson, Admir, Marcelo, Carla, Jean, Erik, Jaime, Volnei, Luís, Juliana. À Nina, cabe um agradecimento muito especial por ter me fornecido o acesso a referências bibliográficas importantíssimas. Ao Túlio, por além de tudo, sempre estar disposto a traduzir o que lhe pedisse. Todos eles fizeram a minha passagem pelo solo gaúcho bem mais divertida.

Ao pessoal do xerox: Eduardo e Cia. pelo sempre eficiente e alegre atendimento e às secretárias do PPGE (Iara, Raquel, Lourdes, Claudia e Aline) pelo também eficiente atendimento a todas as solicitações.

Ao pessoal da pensão em que morei durante quase um ano e que foram essenciais na minha adaptação em POA: Jô, Mara, irmã Anísia, irmã Sueli. E posteriormente à Raquel, pela convivência no segundo ano em Porto Alegre.

À Patrícia Ueda, que dividia comigo as tardes na FEE, agradeço a sua amizade.

Ao Banco Central por ter fornecido os dados necessários para a realização deste trabalho. Em particular, ao auxílio de Katherine Hennings para solucionar dúvidas importantes em relação a esses dados.

À CAPES e, posteriormente, ao CNPQ, agradeço a bolsa concedida, sem a qual eu não teria nenhuma condição de ter feito o mestrado em Porto Alegre.

Por último, mas não menos importante, agradeço à minha família: meus pais Cláudio e Maria e minha irmã Débora, por apoiarem todas as minhas decisões, por terem suportado a ausência (apenas física) nos quase dois anos em que estive no solo gaúcho e por me darem forças para suportar todas as adversidades neste período. Sem o apoio de vocês mais esta etapa da minha formação não teria sido concluída.

Evidentemente, muitas pessoas que foram essenciais para a minha chegada até aqui não foram citadas. A todas elas, que estiveram comigo em POA ou em Araraquara, meus sinceros agradecimentos.

RESUMO

O objetivo principal deste trabalho é estudar empiricamente o processo de formação de expectativas inflacionárias no Brasil, no período recente (pós metas de inflação), através de um modelo conexionista, que aproxima a forma como os agentes fazem previsões. A coordenação das expectativas do mercado em relação à inflação futura é um aspecto crucial do regime de metas de inflação. Dessa forma, entender os fatores que afetam tais expectativas é de grande relevância para o direcionamento adequado da política monetária. Os dados para expectativas de inflação utilizados são provenientes da pesquisa FOCUS do Banco Central do Brasil, e constituem um painel de dados não balanceado. Os resultados obtidos nos permitem afirmar que a maior influência sobre as expectativas inflacionárias no período como um todo foi da volatilidade cambial, seguida pela variação no preço das *commodities*, pela defasagem de ordem um das expectativas, pela variação cambial e pela meta. Em menor magnitude, afetam as expectativas o resultado primário do governo, a defasagem de ordem dois e a taxa Selic. O comportamento desse efeito ao longo do tempo foi verificado através da análise de sensibilidade do produto da rede em resposta a cada uma das variáveis. No período de crise de confiança, há um expressivo descolamento das expectativas em relação à meta, com um aumento do efeito das demais variáveis. Resultado inverso ocorre no período pós-crise: o efeito da meta de inflação aumenta e das demais variáveis tende a se reduzir, ainda que em alguns casos tais efeitos sejam expressivos (como da defasagem de ordem um e da volatilidade cambial). Isso nos leva a concluir que o Banco Central vem consolidando sua credibilidade ao longo do tempo, mas que há ainda espaço para melhorias.

Palavras-Chave: expectativas inflacionárias, metas de inflação, redes neurais artificiais, *bootstrap*, dados em painel.

JEL: C23, C45, E52

ABSTRACT

This work aims to empirically study the formation process of inflationary expectations in Brazil, in the recent period (after the introduction of the inflation targeting policy) by a connexionist model that approaches the way agents forecast. The coordination of market expectations in relation to the future inflation is a crucial aspect of the inflation targeting. This way, understanding the factors that affect such expectations has great relevance for the adequate aiming of the monetary policies. The data for inflation expectations used in this work are from the FOCUS research of the Brazilian Central Bank, and it constitutes a unbalanced data panel. The results obtained allow us to affirm that the biggest influence on the inflationary expectations in the period as a whole was from exchange rate volatility, followed by the commodities prices variation, by the first order lag of the expectations, by the exchange rate variation and by the target. In lesser magnitude, the primary result of the government, the second order lag and the Selic tax affect the expectations. The behavior of this effect throughout the time was verified through the analysis of sensitivity of the product of the network in reply to each one of the inputs. In the period of reliable crisis, there is an expressive shift of the expectations in relation to the target, with an increase of the effect of the other variables. Inverse result occurs in the after-crisis period: the effect of the inflation target increases and of the other variables tend to be reduced, despite in some cases such effect are expressive (as the first order lag and exchange rate volatility). Thus we may conclude that the Brazilian Central Bank has been consolidating its credibility throughout the time, but there is still an open space for improvements.

Key Words: inflation expectations, inflation targeting, artificial neural networks, bootstrap, panel data.

JEL: C23, C45, E52

SUMÁRIO

1. INTRODUÇÃO	13
2. Expectativas e Macroeconomia	16
2.1 Formação de Expectativas	16
2.1.1 Expectativas estáticas ou ingênuas (<i>naive expectations</i>) e o modelo da teia-de-aranha	16
2.1.2 Expectativas extrapolativas.....	17
2.1.3 Expectativas adaptativas	18
2.1.4 Expectativas racionais.....	19
2.1.5 Expectativas e aprendizado adaptativo	20
2.2 Expectativas de inflação e o regime de metas de inflação	21
2.2.1 O arcabouço teórico do regime de metas de inflação.....	22
2.3 Expectativas de inflação e mercados financeiros: a estrutura a termo da taxa de juros	23
2.4 Estudos empíricos sobre a formação de expectativas inflacionárias no regime de metas de inflação: o caso brasileiro	24
3. Redes Neurais Artificiais	31
3.1 Introdução e breve Histórico dos modelos de RNAs	31
3.1.1 Desenvolvimento das RNAs	35
3.2 Aplicações de RNAs à Economia	42
3.3 Aspectos Teóricos das Redes Neurais Artificiais	44
3.3.1 Função de ativação.....	44
3.3.2 Arquitetura de RNAs	47
3.3.3 Aprendizagem em RNAs - paradigmas e algoritmos de treinamento	50
3.4 RNAs e Econometria	58
3.5 Inferência estatística válida em RNAs	62
3.5.1 Procedimento de Racine-White para dados i.i.d.: teste baseado em derivadas parciais.....	62
3.5.2 Inferência para dados dependentes: o uso do bootstrap em bloco.....	64
3.5.3 Inferência em um painel de dados com dependência temporal e ausência de dependência espacial	66
3.6 Aplicação de RNAs a dados em painel	66
4. Análise Empírica da Formação de Expectativas no Brasil	68
4.1 Descrição dos dados	68
4.1.1 Expectativas de inflação.....	68
4.1.2 O comportamento das expectativas de inflação entre fev-2002 e jun-2006.....	70
4.1.2 Variáveis explicativas	73
4.2 Arquitetura da rede	76
4.3 Treinamento e avaliação das redes	78
4.4 Resultados	78
5. Considerações Finais	87
REFERÊNCIAS	89
APÊNDICE A. Variáveis explicativas	99
APÊNDICE B. Alguns Modelos intermediários	105
APÊNDICE C. Sensibilidade do produto da rede	108

APÊNDICE D. Expectativas Fiscais111

1. INTRODUÇÃO

Na ciência econômica, as expectativas dos agentes a respeito dos valores futuros das variáveis são de fundamental importância. É a presença de variáveis expectativas uma das características que distingue os modelos econômicos dos modelos matemáticos usados em ciências naturais. De fato, a teoria econômica moderna reconhece que a diferença central entre a economia e as ciências naturais está nas decisões “*forward-looking*” feitas pelos agentes econômicos (EVANS ; HONKAPOHJA, 2001, p. 5)¹. Além disso, de acordo com a crítica de Lucas (1976), mudanças na política econômica alteram as expectativas formadas pelos agentes, que por sua vez, alteram os parâmetros dos modelos econômicos. Assim, entender como se dá o processo de formação de expectativas é de fundamental importância na modelagem econômica. E esse é um importante desafio para os economistas: compreender como as pessoas interpretam o mundo e formam as expectativas que irão influenciar variáveis econômicas relevantes.

De maneira mais específica, em um país que adota o regime de metas de inflação, as expectativas dos agentes em relação à inflação futura exercem um papel de grande destaque. Altas taxas de inflação alimentam expectativas futuras de inflação mais elevada. Para ancorar essas expectativas, a autoridade monetária se compromete a conter a inflação dentro dos intervalos da meta previamente anunciados. Se a autoridade monetária conta com grande credibilidade, não existe razão para esperar uma taxa de inflação mais alta no futuro, mesmo se a inflação for elevada no período atual. Assim, a autoridade monetária tem a missão de conduzir a sua política de forma a atingir o nível de preços previamente estabelecido, forçando a convergência das expectativas inflacionárias dos agentes econômicos para a meta. Como resultado, o regime de metas de inflação enfraquece a ligação entre a inflação realizada e a inflação passada, com a variável expectacional tomando um papel de grande destaque para o funcionamento adequado do regime (expectativas deixam de ser *backward-looking* e passam a ser *forward-looking*). Quando a política monetária perde credibilidade, o Banco Central aumenta a taxa de juros a fim de controlar a inflação. Pode-se dizer, assim, que o aspecto fundamental do regime de metas de inflação é a coordenação das expectativas dos agentes.

¹ Arthur (1997, p.2), por exemplo, afirma que se não fosse a necessidade de se fazer previsões para os valores futuros das variáveis econômicas na tomada de decisão dos agentes, a economia se tornaria matemática.

Ao longo do desenvolvimento da teoria econômica, foram assumidas algumas regras de formação para as expectativas: estáticas (*naive expectations*), extrapolativas, adaptativas e racionais. O conceito de expectativas racionais foi introduzido por Muth (1961) e é a suposição dominante em macroeconomia. Tal conceito estabelece que os agentes formam suas expectativas eficientemente com base em toda a informação relevante disponível, ou seja, os agentes não cometem erros sistemáticos em suas previsões². Apesar de largamente utilizado, tal processo de formação de expectativas tem sido bastante questionado, pois assume que os agentes não apenas usam toda a sua informação de maneira ótima, mas também que possuem toda a informação requerida para a previsão. Tais hipóteses são muito restritivas, pois implicam que os agentes no modelo tenham mais conhecimento que o próprio economista que os estima (SARGENT, 1993). Assim, uma maneira mais apropriada para modelar a formação de expectativas seria supor que os agentes no modelo agem, eles próprios, da forma como trabalha um economista ou um estatístico.

Seguindo essa abordagem, o uso de redes neurais artificiais (RNAs) se mostra uma ferramenta bastante apropriada para a modelagem de expectativas. A análise de regressão paramétrica, que é a ferramenta mais utilizada em econometria para verificar relações entre variáveis, bem como para fazer previsões, possui algumas limitações, como por exemplo, a necessidade de se assumir uma distribuição de probabilidade para os dados e uma forma funcional para a relação entre as variáveis. No processo de formação de expectativas, é inadequado assumir que isso seja de conhecimento dos agentes, visto que, geralmente, não é de conhecimento dos próprios economistas. Os modelos de RNAs possuem a característica atraente de reconhecer trajetórias altamente complexas (não-lineares) sem que se especifique a forma funcional dessa relação³, aprendendo através de exemplos e possuindo a capacidade de generalização, numa tentativa de imitar o funcionamento do cérebro humano. Assim, modelando a formação de expectativas com o uso de RNAs, estaremos tentando aproximar a forma como os agentes formam suas expectativas.

Dado o exposto acima, o objetivo principal do presente trabalho é estudar empiricamente a formação de expectativas inflacionárias no Brasil no período recente (pós metas de inflação). O principal diferencial do presente trabalho em relação aos

² Assim, o erro de previsão teria comportamento de ruído branco.

³ E como a maioria das relações econômicas é provavelmente não-linear (GRANGER, 1991 apud SHACHMUROVE, 2002, p. 7), a utilização das RNAs na economia tem crescido substancialmente nos últimos anos.

anteriores está na metodologia empregada: a modelagem através de RNAs. Além disso, os dados utilizados também são distintos, já que aqui estamos empregando como variável dependente as expectativas para o IPCA acumulado em 12 meses de cada instituição em cada ponto no tempo, constituindo um painel de dados não balanceado. Os estudos anteriores usavam alguma medida agregada, tal como a média ou mediana do mês ou então os dados diários (como foi o caso de Yoshiro(2005)). A principal motivação para o uso do painel de dados foi a reduzida amostra que se teria utilizando os dados agregados. Além disso, como bem salientado por Hsiao (2005), a utilização de dados em painel permite uma maior capacidade para capturar a complexidade do comportamento humano, quando comparado a séries de tempo ou *cross-section* simples.

O restante deste trabalho está organizado como segue. O capítulo 2 apresenta os principais processos de formação de expectativas utilizados em macroeconomia e faz uma discussão sobre o papel das expectativas inflacionárias no regime de metas de inflação e também nos mercados financeiros (através da estrutura a termo da taxa de juros). São revisados os estudos empíricos que trataram do problema da formação de expectativas no regime de metas de inflação no caso brasileiro. No capítulo 3, fazemos uma breve apresentação do histórico das RNAs, explicando inicialmente como é o funcionamento das redes neurais biológicas e de sua célula básica (neurônio) e, então, descrevendo o progresso das RNAs, a partir do trabalho de McCulloch e Pitts (1943). É descrito brevemente o desenvolvimento da utilização das RNAs na economia e são apresentados os seus aspectos teóricos, limitando-nos apenas a redes MLP (*multilayer perceptron*). Neste capítulo apresentamos também a metodologia para inferência estatística proposta por Racine e White (2001) para dados i.i.d. e duas modificações sugeridas para o caso de dados dependentes: o *bootstrap* em bloco e, para o caso de dados em painel com ausência de dependência espacial, a reamostragem apenas na dimensão transversal. Por fim, o capítulo 4 apresenta os resultados obtidos.

2. Expectativas e Macroeconomia

2.1 Formação de Expectativas

Durante o desenvolvimento da teoria econômica, quatro foram os principais processos assumidos para a formação de expectativas, os quais serão apresentados a seguir. Também será feita uma breve descrição sobre o processo de aprendizado na formação de expectativas.

2.1.1 Expectativas estáticas ou ingênuas (*naive expectations*) e o modelo da teia-de-aranha

Nesse caso, assume-se que os agentes formam sua expectativa sobre alguma variável y_t considerando o valor observado mais recentemente. Assim, pode-se escrever:

$$E_t(y_{t+1}) = y_t$$

Esse é o processo mais simples de formação de expectativas e foi largamente utilizado no início da literatura. Embora simples, tal processo gera previsões eficientes quando os dados observados não apresentam grandes variações no tempo ou quando essas variações repetem-se de forma freqüente e contínua.

O modelo da teia-de-aranha foi desenvolvido para explicar possíveis relações dinâmicas na economia e baseia-se na análise da reação de fazendeiros aos preços. A oferta do produto depende do seu preço passado enquanto a demanda é função do preço corrente. A decisão da produção deve ser feita de forma antecipada, ou seja, um período antes da venda real (o plantio antecede a colheita e a venda por um período de tempo significativo). Dessa forma, a decisão deve ser baseada em preços esperados. O teorema da teia de aranha considera que os preços esperados são os preços correntes na época da decisão quanto à produção. O modelo desse mercado (em sua versão linear) pode então ser descrito pelas seguintes equações:

$$\text{Demanda: } Q_t^D = a - bP_t + \mu_t$$

$$\text{Oferta: } Q_t^S = c + dP_t^e + v_t$$

$$\text{Equilíbrio: } Q_t^D = Q_t^S$$

com $b, d > 0$

Assumindo-se a condição de equilíbrio, este modelo pode ser escrito na seguinte forma reduzida:

$$P_t = \alpha + \beta P_t^e + \eta_t$$

$$\text{onde: } \alpha = \frac{a+c}{b}; \beta = -\frac{d}{b}; \eta_t = \frac{\mu_t - v_t}{b}$$

A equação acima afirma que o preço presente depende do preço futuro esperado, ilustrando, de forma simples, o papel central exercido pelas expectativas. No caso das expectativas ingênuas, temos:

$$P_t = \alpha + \beta P_{t-1} + \eta_t$$

Desconsiderando o choque aleatório, essa é uma equação em diferenças de primeira ordem. Para a análise de estabilidade desse modelo, são possíveis três resultados, os quais dependem das elasticidades da demanda e da oferta:

- a) elasticidade da demanda igual à elasticidade da oferta: flutuações cíclicas contínuas;
- b) elasticidade da demanda menor que a elasticidade da oferta: flutuações explosivas divergentes;
- c) elasticidade da demanda maior que a elasticidade da oferta: flutuações amortecidas convergentes.

2.1.2 Expectativas extrapolativas

O processo estático de formação de expectativas não é uma hipótese muito plausível a não ser que a variável que esteja sendo prevista siga um processo de passeio

aleatório (*random walk*), ou seja, que a informação contida nos demais preços passados seja nula. No modelo das expectativas extrapolativas, leva-se em consideração também a tendência mais recente das variáveis. Dessa forma, temos que o valor esperado de uma variável y será dado por:

$$y_t^e = y_{t-1} + \eta(y_{t-1} - y_{t-2})$$

onde $\eta > 0$ é o coeficiente de expectativa de Meltzer⁴.

2.1.3 Expectativas adaptativas

As expectativas adaptativas originaram-se a partir dos trabalhos de Irving Fisher na década de 1930 (SANTOS, 2003, p. 14) e foram introduzidas formalmente na década de 1950 por Cagan (1956), Friedman (1957) e Nerlove (1958) (EVANS ; HONKAPOHJA, 2001, p.10). Esse processo de formação de expectativas teve um papel fundamental na macroeconomia nas décadas de 1960 e 1970. De acordo com essa abordagem, os agentes formam suas expectativas sobre o futuro a partir dos erros das previsões passadas:

$$E_t(y_{t+1}) = E_{t-1}(y_t) + \lambda(y_t - E_{t-1}(y_t))$$

onde $0 < \lambda < 1$ é um parâmetro de ajustamento.

Fazendo uso do operador defasagem L ($Ly_t = y_{t-1}$), a equação acima pode ser escrita também como:

$$E_t(y_{t+1}) = \frac{\lambda}{1 - (1 - \lambda)L} y_t$$

O que implica:

$$E_t(y_{t+1}) = \lambda \sum_{i=0}^{\infty} (1 - \lambda)^i y_{t-1-i}$$

⁴ Note que se $\eta = 0$, o processo de formação de expectativas iguala-se ao estático.

Tal equação representa um modelo de defasagens distribuídas com pesos exponencialmente declinantes. Assim, experiências mais distantes exercem um efeito menor que experiências mais recentes, o que estaria de acordo com o bom senso. Cabe notar que, neste caso, os agentes estão sujeitos a cometer erros sistemáticos em suas previsões.

Sua principal vantagem é a facilidade de se estimar econometricamente. Note-se, porém, que se a variável de interesse estiver mudando, os indivíduos irão continuamente piorar a sua previsão. Apenas se a economia se encontrar num estado estacionário é que os indivíduos irão prever de forma satisfatória o futuro.

2.1.4 Expectativas racionais

Os modelos de expectativas adaptativas forneciam previsões bastante inadequadas em certos contextos e regras de previsão bastante superiores estavam facilmente disponíveis (EVANS ; HONKAPOHJA, 2001, p. 11). Diante dessa observação, buscou-se uma nova abordagem para a formação de expectativas que deu origem à chamada “revolução das expectativas racionais”.

O conceito de expectativas racionais foi introduzido por Muth (1961). Tal conceito implica que os agentes econômicos formam suas expectativas com base em todas as informações disponíveis, e fazem isso de maneira eficiente, ou seja, eles compreendem como essas informações afetam a variável que estão tentando prever. Assim, de acordo com a hipótese de expectativas racionais, as previsões dos agentes são essencialmente iguais às previsões da teoria econômica relevante (MUTH, 1961, p. 316). Note-se, porém, que isso não significa que suas previsões serão sempre corretas. Se houver incerteza, elas poderão desviar do seu valor verdadeiro de forma completamente aleatória, o que significa que, **na média**, as previsões serão corretas. A hipótese das expectativas racionais se torna sinônimo de previsão perfeita apenas no caso em que a informação é completa e a incerteza inexistente⁵.

Considerando expectativas racionais, podemos escrever:

$$y_{t+1} = y_{t+1}^e + \eta_t$$

⁵ É interessante notar que se a variável seguir um processo do tipo ARIMA(0,1,1), as expectativas adaptativas serão também racionais.

$$y_{t+1}^e = E_t(y_{t+1} | I_t)$$

onde I_t é o conjunto de informação disponível em t e η_t é ruído branco.

A abordagem das expectativas racionais é a hipótese padrão dos modelos econômicos recentes. De acordo com essa hipótese, os agentes não cometem erros de forma sistemática, ou seja, em média, as suas previsões estão corretas.

Uma crítica importante dessa hipótese é que ela não permite a aprendizagem ao longo do tempo. Se, por exemplo, o governo muda suas regras de política, os indivíduos poderiam levar um certo tempo para aprender a nova regra, o que não é considerado na hipótese de expectativas racionais.

2.1.5 Expectativas e aprendizado adaptativo

A abordagem das expectativas racionais pressupõe que os agentes tenham um grande conhecimento sobre a economia, a estrutura de seu funcionamento e seus parâmetros. Talvez seja mais sensato que os agentes enfrentem as mesmas limitações dos economistas sobre o conhecimento do funcionamento da economia. Uma visão mais factível de racionalidade é considerar os agentes como estatísticos ou econometristas que fazem previsões sobre o estado futuro da economia. E esta é a idéia básica da abordagem da aprendizagem adaptativa para a modelagem da formação de expectativas, na qual os agentes ajustam suas regras de previsão conforme novos dados se tornam disponíveis.

Para uma ilustração, tomemos novamente o exemplo do modelo da teia de aranha. Assumindo que os agentes acreditem que o processo estocástico para o preço de mercado assuma a forma a seguir, temos:

$$P_t^e = \mu + \varepsilon_t$$

A média amostral é a forma padrão para se estimar uma constante desconhecida. Assim, as expectativas dos agentes serão dadas por:

$$P_t^e = t^{-1} \sum_{i=0}^{t-1} P_i$$

Suponha agora que uma variável exógena conhecida, W_{t-1} , seja introduzida no modelo da teia de aranha:

$$P_t = \alpha + \beta P_t^e + \delta W_{t-1} + v_t$$

Nesse caso, a previsão para o preço seria naturalmente uma função linear de W_{t-1} . Assim, considerando a aprendizagem, os agentes devem prever de acordo com a seguinte equação:

$$P_t^e = a_{t-1} + b_{t-1} W_{t-1}$$

onde a_{t-1} e b_{t-1} são as estimativas obtidas de uma regressão por mínimos quadrados ordinários, por exemplo.

Seja y_t um vetor de variáveis que os agentes devam prever e y_t^e as expectativas formadas por tais agentes⁶. Quando estão fazendo suas previsões, os indivíduos dispõem de um vetor de variáveis X_t , que pode conter um número finito de defasagens dos componentes de y_t , de defasagens de y_t^e , bem como outras variáveis exógenas ou endógenas observáveis. Assim, a previsão seria dada por:

$$y_t^e = \Theta(X_t, \theta_{t-1})$$

θ_{t-1} : vetor de parâmetros, que pode se modificar ao longo do tempo.

Sob a abordagem estatística para o aprendizado, a regra de previsão $\Theta(X_t, \theta_{t-1})$ é baseada na especificação de um modelo econométrico.

2.2 Expectativas de inflação e o regime de metas de inflação

O regime de metas de inflação foi adotado pioneiramente na Nova Zelândia, em 1990 e no Brasil em junho de 1999. A característica principal desse regime é o anúncio de metas a serem cumpridas para determinado indicador de inflação. Essa é a informação básica a ser divulgada para o setor privado como um todo. A intuição desse modelo é dada a seguir. Altas taxas de inflação alimentam expectativas futuras de inflação mais elevada. Para ancorar essas expectativas, a autoridade monetária se

⁶ Se os agentes são heterogêneos, isto poderia ser tratado fazendo $y_t^e(k)$ as expectativas do agente k.

compromete a conter a inflação dentro dos intervalos da meta previamente anunciados. Se a autoridade monetária conta com grande credibilidade, não existe razão para esperar uma taxa de inflação mais alta no futuro, mesmo se a inflação for elevada no período atual. De maneira análoga, não existe razão para esperar uma deflação futura, mesmo se a inflação for muito baixa correntemente.

A implementação do regime de metas de inflação envolve a definição e divulgação oficial de limites para um dado índice de inflação em determinado horizonte de tempo. Assim, a autoridade monetária tem a missão de conduzir a sua política de forma a atingir o nível de preços previamente estabelecido, forçando a convergência das expectativas inflacionárias dos agentes econômicos para a meta. Como resultado, o regime de metas de inflação enfraquece a ligação entre a inflação realizada e a inflação esperada, com a variável expectacional tomando um papel de grande destaque para o funcionamento adequado do regime.

2.2.1 O arcabouço teórico do regime de metas de inflação

O regime de metas de inflação pode ser visualizado como a avaliação por parte da autoridade monetária de uma função objetivo, na qual tanto a estabilização da inflação quanto do produto são importantes:

$$L_t = E_{t-1} \sum_{i=0}^{\infty} \beta^i \left[(\pi_{t+i} - \pi^{meta})^2 + \lambda x_{t+i}^2 \right]$$

onde:

π^{meta} = meta de inflação

λ = peso dado ao hiato do produto

O objetivo da autoridade monetária é escolher a taxa de juros i_t (instrumento de política monetária) que minimiza a função perda acima⁷. E a escolha de i_t está sujeita às restrições representadas pelas equações abaixo:

$$E_{t-1} x_t = E_{t-1} x_{t+1} - \frac{1}{\sigma} (i_t - E_{t-1} \pi_{t+1})$$

⁷ Deve-se notar que o controle da autoridade monetária sobre as variáveis é imperfeito. Há defasagem temporal entre a alteração no instrumento e o seu impacto na inflação, incertezas sobre a extensão dessa defasagem, além de choques inesperados.

$$E_{t-1}\pi_t = \beta E_{t-1}\pi_{t+1} + \kappa E_{t-1}x_t + \rho e_{t-1}$$

A condição de primeira ordem para a escolha de i_t implica em⁸:

$$E_{t-1}x_t = -\left(\frac{\kappa}{\lambda}\right)E_{t-1}(\pi_t - \pi^{meta})$$

A previsão da inflação, portanto, é um elemento crucial no regime de metas de inflação, pois ela é quem guiará as decisões da política monetária. Se a previsão de inflação da autoridade monetária para o período t exceder a meta, a política deve ser ajustada de forma a assegurar que o a previsão do hiato do produto seja negativa. Dessa forma, a política monetária é *forward-looking*, estando preocupada com a inflação esperada e não com a inflação corrente e, assim, respondendo às pressões inflacionárias que possam afetar o alcance da meta previamente estabelecida. Essa característica constitui a principal diferença entre o regime de metas de inflação e os demais regimes monetários.

2.3 Expectativas de inflação e mercados financeiros: a estrutura a termo da taxa de juros

As expectativas inflacionárias têm um papel fundamental na dinâmica inflacionária não apenas por serem um componente da formação de preços, salários e contratos da economia, mas também por determinarem a estrutura a termo da taxa de juros real, afetando assim as decisões privadas sobre alocação de recursos em consumo e investimento. A estrutura a termo das taxas de juros mostra a evolução destas com o alongamento do prazo das aplicações em renda fixa, ou seja, estabelece a relação entre as taxas sobre diferentes horizontes. De forma geral, quanto maior for o prazo da aplicação, maior será a taxa recebida⁹. O modelo mais simples e popular desta relação é

⁸ Veja Walsh (2003, p. 543).

⁹ Quando isso ocorre, diz-se que a estrutura a termo é normal (curva de juros positivamente inclinada). Se as taxas caírem com o alongamento do prazo da aplicação, diz-se que a estrutura a termo das taxas de juros é invertida (curva negativamente inclinada). Isso pode ocorrer quando a autoridade monetária eleva a taxa de juros visando conter a inflação futura. Contando com credibilidade por parte dos agentes, isto fará com que os agentes prevejam uma inflação futura mais baixa o que, simultaneamente a uma elevação da taxa de juros, fará com que as taxas de juros de longo prazo não se elevem (ou mesmo caiam).

conhecido como teoria das expectativas da estrutura a termo, a qual estabelece que a taxa de juros de longo prazo é formada como uma média das taxas de juros de curto prazo esperadas para o futuro, mais um prêmio de risco invariante no tempo¹⁰, o que pode ser representado como:

$$i_{t,n} = \frac{i_{t,1} + E_t i_{t+1,1} + \dots + E_t i_{t+n-1,1}}{n} + \theta$$

A inclinação da curva de juros reflete as expectativas do mercado sobre as mudanças futuras na taxa de juros. As expectativas sobre o estado futuro da economia exercem, portanto, um papel fundamental da determinação da estrutura a termo.

A estrutura a termo pode ser interpretada como um indicador das expectativas de inflação. Como a taxa de juros do mercado constituem a soma de um retorno real esperado e um prêmio de inflação esperado, a taxa de juros nominal pode ser expressa como:

$$i_n^t = \frac{\sum_{i=0}^n E_t r_{t+i} + E_t \pi_{t+n}}{n}$$

Se as taxas reais forem estacionárias em torno de uma média \bar{r} , temos:

$$i_n^t \approx \bar{r} + \frac{E_t \pi_{t+n}}{n}$$

O que ilustra que, nesse caso, variações na taxa de juros de longo prazo serão causadas principalmente por variações na inflação esperada.

2.4 Estudos empíricos sobre a formação de expectativas inflacionárias no regime de metas de inflação: o caso brasileiro

O processo de formação de expectativas no Brasil tem recebido uma atenção maior no período recente, devido à implementação do sistema de metas de inflação em junho de 1999, no qual as expectativas inflacionárias têm um papel central, como vimos acima¹¹.

¹⁰ Um prêmio de risco variante no tempo pode também ser incorporado. Veja Walsh (2003, p. 490).

¹¹ O primeiro trabalho que buscou estudar o processo de formação de expectativas no Brasil foi Garcia (1994). Ele estudou a estrutura estocástica da série de expectativas inflacionárias usando dados do

Desde então, o Banco Central do Brasil vem coletando diariamente as expectativas do mercado a respeito das principais variáveis macroeconômicas¹². Com esse banco de dados, é possível realizar estudos empíricos sobre a formação de expectativas no Brasil, ainda que de forma limitada, dado o tamanho reduzido da amostra.

Esses estudos têm como objetivo central verificar o papel exercido pela meta de inflação na formação de expectativas, controlando para outros determinantes. Se o coeficiente desta variável for estatisticamente significativo e próximo de um, teríamos evidências que a política monetária é crível. Os estudos realizados para o período utilizam como variável dependente a expectativa de mercado para o IPCA 12 meses à frente e têm como principal método de estimação o de mínimos quadrados ordinários, em um contexto de regressão linear univariada para séries temporais. Foge a essa regra de metodologia o estudo de Mello e Moccerro (2006), o qual utiliza um *framework* multivariado. As evidências apresentadas, em geral, apontam para o forte papel exercido pela meta de inflação sobre as expectativas, no período como um todo. Quando é utilizada a estimação recursiva, o comportamento observado é de influência crescente da meta e decrescente das demais variáveis. Entretanto, quando a amostra inclui o período de crise de confiança (eleições de 2002), os resultados apontam para um descolamento das expectativas em relação à meta e uma influência maior das demais variáveis durante a crise. Dessa forma, os estudos sugerem que a política monetária vem consolidando a sua credibilidade ao longo do tempo e que caminha na direção certa. Em uma abordagem distinta, contudo, Lowerkron e Garcia (2005) mostram que, no Brasil, a surpresa de curto prazo da inflação afeta o desvio das expectativas em relação à meta, o que seria evidência da falta de credibilidade, contrariamente aos demais trabalhos.

Minella et al. (2002), buscando avaliar a credibilidade da política monetária do Banco Central do Brasil, estimaram uma função de reação para as expectativas de mercado 12 meses à frente. O objetivo principal seria avaliar o papel da meta de inflação na formação de expectativas. Para o período de 2000:01 a 2002:06 eles estimaram a função de reação através do método de MQO, usando como variáveis explicativas *lags* de ordem 1 e 2 das expectativas, a meta de inflação 12 meses à frente e

mercado futuro. Mais precisamente, Garcia (1994) extraiu as expectativas inflacionárias usando a estrutura a termo de contratos futuros para o nível de preços no período de agosto de 1987 a março de 1990 e verificou a existência de raiz unitária na série de expectativas. Os testes realizados não rejeitaram a hipótese nula de raiz unitária durante o ano de 1987, que foi um período de inflação extremamente alta. Porém, no período posterior a 1987, foi possível rejeitar a hipótese de raiz unitária na série da inflação esperada.

¹² Para maiores detalhes a respeito do *survey* de expectativas de mercado do Banco Central do Brasil, consulte Marques, Fachada e Cavalcanti (2003).

a taxa de juros. Os coeficientes dessas variáveis foram todos estatisticamente significantes a 5% e com o sinal esperado. O resultado mais importante foi o coeficiente da meta de inflação, estimado em 0.96, mostrando que as expectativas reagem de forma significativa à meta¹³.

Assim como Minella et al (2002), o objetivo de Schmidt-Hebbel e Werner(2002) foi verificar a credibilidade da política monetária analisando o papel das metas de inflação sobre as expectativas dos agentes. Os autores apresentaram evidências empíricas para os casos do Brasil, Chile e México. Aqui nos deteremos a descrever apenas os resultados encontrados para o Brasil, que é foco do presente trabalho. Os regressores utilizados foram a depreciação da taxa nominal de câmbio, a diferença entre a inflação e a sua expectativa, o desvio da inflação esperada em relação à meta e a diferença da inflação em relação ao seu núcleo, para o período de 1999:12 a 2001:10. As equações estimadas por MQO indicaram que apenas o desvio da inflação em relação à meta é significativa para o caso do Brasil. Os autores concluem, com isso, que a meta de inflação exerce uma influência forte e credível nas expectativas do setor privado. Utilizando a técnica de estimação recursiva, os autores mostram que o coeficiente da taxa de depreciação do câmbio declina de forma constante ao longo do período enquanto o coeficiente da meta de inflação é positivo, grande, estatisticamente significativo e relativamente estável; o coeficiente da inflação não antecipada é próximo de zero durante todo o período, bem como a diferença entre o núcleo e a inflação cheia.

Minella et al (2003) basicamente atualizam o estudo realizado por Minella et al. (2002), utilizando o período 2000:01 a 2003:02. Através do método dos MQO, os autores encontram que a expectativa de inflação 12 meses à frente é influenciada por seus *lags* (de primeira e segunda ordem), pela meta de inflação (12 meses), pela taxa de juros e pela taxa de inflação de 12 meses. O coeficiente da meta estimado é, novamente, próximo de 1, o que seria uma evidência forte de que as metas têm um papel importante na formação de expectativas dos agentes¹⁴. Utilizando estimação recursiva, os autores também encontram que o coeficiente do termo de inflação passada começou a aumentar a partir do final de 2001 (o que seria explicado pelos choques significantes que afetaram a economia) e, mais expressivamente, nos últimos meses de 2002 e início de 2003, o

¹³ Para evitar que este resultado fosse consequência de correlação entre metas e inflação passada, os autores incluíram também a defasagem da taxa de inflação 12 meses atual na regressão, porém esta não se mostrou estatisticamente significativa a 10%, excluindo essa possibilidade.

¹⁴ Note que, novamente, esse resultado poderia ser consequência de alguma correlação entre as metas e a inflação passada, como ressaltam os autores. Porém, a regressão também inclui a taxa de inflação 12 meses atual com o intuito de “expurgar” esse efeito (MINELLA et. al 2003, p. 18).

que seria reflexo da crise de confiança neste período. De fato, com a amostra finalizando em setembro de 2002, o coeficiente do termo de inflação passada é insignificante. Foram adicionadas também as variáveis EMBI Plus e variação da taxa de câmbio em 12 meses, esta última não se mostrando estatisticamente significativa. Para a amostra finalizando em setembro de 2002, o risco-Brasil também não é significativo.

Cerisola e Gelos (2005) investigam os principais determinantes macroeconômicos das expectativas inflacionárias no Brasil, usando também a base de dados da pesquisa FOCUS. O modelo inclui como variáveis explicativas das expectativas o *lag* de ordem 1 da taxa de inflação de 12 meses, a meta de inflação de 12 meses, superávit primário consolidado (% PIB) em t-3, taxa de juros real em t-3, desvios da taxa de câmbio efetiva real e do salário real de seus valores históricos, em t-3. Os resultados obtidos sugerem que o regime de metas de inflação tem sido eficiente em ancorar as expectativas. A estimação recursiva evidencia que desde 2003 e, especialmente durante 2004, as expectativas têm sido crescentemente ancoradas pela meta de inflação e menos dependentes da inflação passada. A política fiscal se mostrou importante instrumento para ancorar as expectativas, especialmente durante os períodos de elevada incerteza (como em 2002).

Yoshihiro (2005) verifica a influência da variação cambial nominal, nível de IGP-M atual, preço do petróleo futuro, IPCA atual, meta de inflação 12 meses, do EMBI Brasil e da crise eleitoral de 2002 sobre as expectativas do IPCA 12 meses à frente¹⁵, durante o período de novembro de 2001 a março de 2005. Neste trabalho, foram utilizados os dados diários das expectativas e todas as variáveis foram consideradas em primeira diferença. O preço do petróleo futuro, o risco-Brasil, o IPCA e o IGPM não se mostraram estatisticamente significantes. A meta de inflação foi significativa estatisticamente e com coeficiente positivo, refletindo que atua como guia para a formação de expectativas dos agentes (o coeficiente estimado foi de 0,76). Para analisar a evolução dos coeficientes ao longo do tempo, os autores realizaram o mesmo exercício a 3 subperíodos: 2002, 2002:2003 e 2002:2004. Os autores concluem que a credibilidade da política monetária ainda está incompleta, mas que caminha na direção correta.

Carvalho e Bugarin (2005) estudaram empiricamente a formação de expectativas inflacionárias no Brasil, no período de junho 1999 a novembro de 2003, utilizando a

¹⁵ O autor utilizou a série suavizada das expectativas.

inflação esperada 12 meses à frente¹⁶. Os dados utilizados são provenientes do *survey* de expectativas de mercado do Banco Central do Brasil. As equações foram estimadas usando MQO. Os resultados mostram que no período anterior a 2002, o coeficiente da meta inflacionária do Bacen era altamente significativo e próximo de um, mostrando alta credibilidade da política monetária. Analisando a amostra após novembro de 2002, os autores encontram um comportamento adaptativo forte na formação de expectativas inflacionárias. Além disso, nesse período, a meta de inflação do Bacen deixa de ser estatisticamente significativa.

Carvalho e Bugarin (2005) verificam a validade da hipótese de expectativas racionais na série de inflação esperada brasileira, tanto utilizando os dados agregados quanto utilizando dados em painel. Para os testes de racionalidade, foram utilizadas as inflações esperadas para 3, 6 e 12 meses à frente. Os testes para os dados agregados mostram que as previsões medianas são não viesadas para todos os horizontes de previsão, desde que seja levada em consideração a autocorrelação dos erros. Os autores também mostram que as previsões foram eficientes no uso da informação passada, considerando como variáveis explicativas o IPCA e a taxa de juros Selic. Porém, previsões de curto prazo se mostraram ineficientes no uso do IGP-DI e da taxa de câmbio. Já as previsões de longo prazo não foram eficientes no uso do *gap* do produto. Usando dados em painel, Carvalho e Bugarin (2005) encontram que apenas o painel para as previsões 12 meses à frente, com uma amostra que se inicia em novembro de 2001, passa no teste de ausência de viés. Porém, a hipótese de eficiência no uso das informações foi rejeitada para todas as variáveis investigadas.

Carvalho e Souza Júnior (2006) estudam os determinantes das expectativas de inflação 12 e 6 meses à frente. No primeiro caso, as expectativas foram modeladas considerando a política monetária em $t-2$ (taxa nominal de juros Selic efetiva), a meta de inflação para 12 meses, o risco-país em $t-2$ (logaritmo do Embi Plus), o *lag* de primeira ordem da expectativa, a taxa de câmbio (variação entre a taxa corrente e a de 3 períodos anteriores) e a inflação mensal corrente anualizada. No caso das expectativas 6 meses à frente, encontrou-se que essa pode ser explicada pela política monetária, pelas metas de inflação, risco-país e pelo *lag* de primeira ordem. Para as expectativas 12

¹⁶ A série de inflação esperada 12 meses à frente foi construída utilizando-se a mediana das expectativas no último dia de negócios de cada mês, a fim de garantir que os agentes já conheçam a inflação do mês anterior, mas ainda não tinham informação a respeito da inflação do mês atual (CARVALHO ; BUGARIN, 2005, p. 8).

meses à frente, a variável defasada mostrou-se significativa, bem como a meta de inflação, a taxa Selic e o câmbio.

Lowerkron e Garcia (2006) verificam o efeito de surpresas de curto prazo nas expectativas de inflação de médio prazo no Brasil, comparando os resultados com outros países do mundo. Para tanto, é estimado um modelo em que a variável dependente é o desvio da meta de 12 meses ($IPCA_{12m}^e - META_{12m}^{Bacen}$) e as variáveis explicativas consistem da primeira defasagem da variável dependente, a surpresa inflacionária de curto prazo ($IPCA_t - IPCA_t^e |_{t-1}$) e a variação do câmbio. A principal conclusão dos autores é que as surpresas de inflação de curto prazo induzem uma variação significativa nas expectativas dos agentes, mesmo “expurgando” o efeito da taxa de câmbio. O mais interessante é que o Brasil seria um dos poucos países em que as surpresas de curto prazo da inflação afetam as expectativas de médio prazo. Segundo os autores, isso seria um sintoma de um de dois possíveis problemas: inércia inflacionária causada pela indexação da economia e/ou falta de credibilidade da política monetária. O modelo desenvolvido¹⁷ afirma que a falta de credibilidade pode ser verificada através do comportamento dos co-movimentos no prêmio de risco da inflação¹⁸. Essa análise mostra que esta é uma das causas no Brasil¹⁹, já que foi encontrada uma forte relação entre as surpresas de inflação de curto prazo e o prêmio de risco da inflação.

Mello e Moccero (2006) utilizaram a análise de co-integração para estimar a função de reação do Banco Central e os determinantes da inflação esperada no Brasil, Chile, Colômbia e México. Os autores também testaram a presença de *spillovers*²⁰ de volatilidade entre a taxa de juros (instrumento de política monetária) e as expectativas de inflação, através dos modelos M-GARCH. No caso brasileiro, seus resultados apontam que, no longo prazo, as expectativas têm sido influenciadas pelas metas de inflação e taxa de juros, e que uma maior volatilidade da taxa de juros leva a maior volatilidade das expectativas de inflação, concluindo que uma suavização nas mudanças

¹⁷ De maneira intuitiva, isso pode ser explicado da seguinte forma: se o efeito da surpresa de inflação de curto prazo nas expectativas de inflação 12 meses é causado apenas pela indexação da economia, não existiria razão para um aumento da incerteza caso a economia fosse afetada por um choque de inflação positivo (LOWENKRON; GARCIA, 2006, p. 9).

¹⁸ O prêmio de risco da inflação é a diferença entre a inflação de mercado (implícita nos ativos financeiros) e a expectativa de inflação dos participantes do mercado.

¹⁹ O que não significa que a indexação também seja a causa. O ponto importante é que a falta de credibilidade piora a situação.

²⁰ Ou seja, os autores verificaram se a volatilidade da taxa de juros afeta a volatilidade da inflação esperada.

do instrumento de política monetária ajudaria a reduzir a volatilidade das expectativas de inflação.

Bevilaqua, Mesquita e Minella (2007) também realizaram um estudo econométrico a respeito do papel exercido pela meta de inflação nas expectativas do IPCA 12 meses à frente. Eles encontram evidências do papel crítico exercido pela meta como atratores das expectativas inflacionárias. Contudo, tomando o período da crise de confiança em meados de 2002 e início de 2003, o coeficiente da meta é bastante baixo ou não significativo, dependendo da versão do modelo utilizada, ocorrendo o contrário para o período mais recente. Quanto ao comportamento dos demais coeficientes dos modelos apresentados, os autores ressaltam dois resultados mais importantes em relação ao período recente. O primeiro é que o efeito da taxa de câmbio é mais baixo ou até não significativo, o que indicaria uma possível redução do *pass-through* na economia. E o segundo é que o coeficiente da inflação passada é próximo de zero e não significativo. Com a inclusão do risco-Brasil, a inflação defasada deixa de ser significativa. Como ressaltado pelos autores, a presença de regressores estatisticamente significantes além da meta de inflação não é um indício de falta de credibilidade do Banco Central, mas sim de quão rico e complexo é o mecanismo de transmissão. Dessa forma, mesmo autoridades monetárias com credibilidade alta não são capazes de eliminar completamente as flutuações da inflação.

3. Redes Neurais Artificiais

3.1 Introdução e breve Histórico dos modelos de RNAs

As redes neurais artificiais (RNAs) originaram-se da tentativa de imitar o funcionamento do cérebro humano, objetivando criar um modelo que possuísse características peculiares deste, tais como a robustez e tolerância a falhas, flexibilidade, capacidade de aprendizagem, processamento paralelo de informação e a grande habilidade para reconhecer trajetórias (ABELÉM, 1994, p. 9; MISRA ; WARNER, 1996, p. 2). Apesar de terem surgido com esse propósito, os modelos de RNAs estão ainda muito longe das redes neurais biológicas.

A estrutura e funcionamento do sistema nervoso passaram a ser melhor compreendidos a partir do trabalho do histologista espanhol Santiago Ramón y Cajal (1852-1934). Baseado na doutrina celular de Theodor Schwann e Mathias Jakob Schleiden²¹ e em técnicas histológicas desenvolvidas pelo histologista italiano Camillo Golgi²², Cajal propôs que o tecido nervoso era formado por unidades individuais (o que passou a ser denominado de doutrina neuronal), e não por uma "teia única" (doutrina reticularista), como a maioria dos neurocientistas acreditava naquela época²³.

O neurônio é a unidade fundamental do sistema nervoso. É composto por três partes: corpo celular (ou soma), dendritos e axônio. O *corpo celular* aloja o núcleo e as principais organelas intracelulares, como mitocôndria, complexo de Golgi, retículo endoplasmático, ribossomos, entre outras. O *axônio* é um grande prolongamento de cor branca cuja porção final apresenta ramificações (chamadas de telodendro, palavra de origem grega que significa "ramos distantes") e é o responsável pela transmissão dos sinais de saída do neurônio, ou seja, pelas informações geradas e que devem ser conduzidas a outras células (a outros neurônios ou a outros tipos celulares). Muitos dos axônios possuem uma cobertura isolante composta por lipídios e proteínas, denominada

²¹ A doutrina celular proposta em 1838 por estes dois alemães afirmava que todos os tecidos orgânicos eram constituídos por células.

²² Cabe notar que ambos os histologistas, Golgi e Cajal, receberam o prêmio Nobel de Medicina e Fisiologia em 1906. O primeiro pelo desenvolvimento de um método histológico revolucionário, e o segundo, pela sua aplicação. O que é surpreendente é que Cajal utilizou o método desenvolvido por Golgi exatamente para negar a teoria reticularista defendida por este. Veja detalhes em Lent (2001, p. 72-73) e Sabbatini (2003).

²³ Para conhecer a história da descoberta do neurônio e das sinapses, consulte Sabbatini (2003).

bainha de mielina. Os *dendritos*²⁴, inúmeras ramificações do corpo celular, são responsáveis em receber as informações que chegam de outros neurônios ou de outros tipos celulares, funcionando assim, como verdadeiras antenas. A Figura 3.1 mostra o neurônio do córtex cerebral de um roedor visto do microscópio. De acordo com Lent (2001, p. 14), muitos neurônios do córtex²⁵ humano são parecidos com esse.

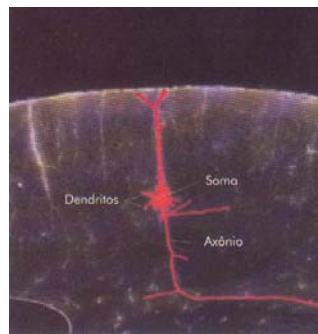


Figura 3.1 - Neurônio do córtex cerebral de um roedor
Fonte: Lent (2001, p. 14)

Os neurônios diferenciam-se das demais células animais por sua morfologia adaptada ao processamento de informações e pela grande variedade de seus tipos morfológicos. Existem neurônios de todas as formas possíveis e imagináveis, a maioria deles localizada no sistema nervoso central, dentro da caixa craniana e da coluna vertebral. A Figura 3.2 mostra alguns dos diversos tipos morfológicos de neurônios.

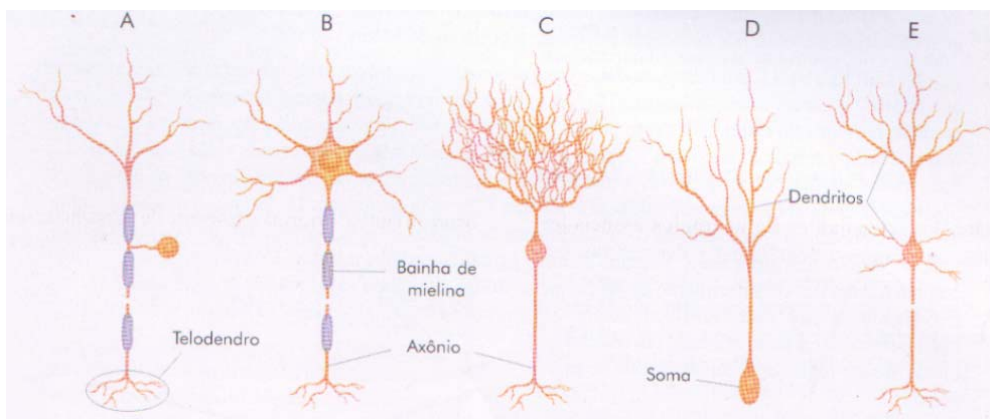


Figura 3.2 - Alguns tipos morfológicos de neurônios: (A) pseudounipolar, (B) estrelado, (C) de Purkinje, (D) unipolar e (E) piramidal
Fonte: Lent (2001, p. 15)

²⁴ Do grego, "pequenas ramificações de árvore" ou, para quem preferir, galhinhos.

²⁵ O córtex cerebral é uma superfície enrugada, cheia de giros e sulcos, responsável pelas funções neurais e psíquicas mais complexas.

No cérebro humano existem aproximadamente cem bilhões de neurônios, que realizam aproximadamente cem trilhões de conexões interneuronais simultaneamente. O conjunto de neurônios associados forma os circuitos ou *redes neurais*, como mostra a Figura 3.3.

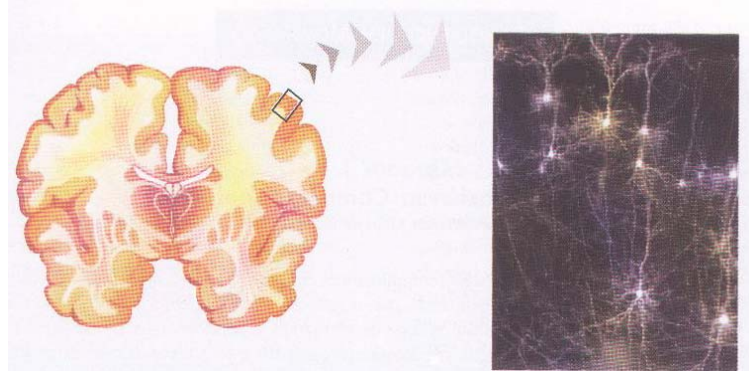


Figura 3.3 - Redes neurais vistas ao microscópio
Fonte: Lent (2001, p. 15)

Os neurônios são agrupados em grandes conjuntos no sistema nervoso, cada um deles com identidade funcional. Assim, as diferentes funções são localizadas em áreas distintas, ou seja, cada função é realizada por uma parte do sistema nervoso, mas todas operam em paralelo. É possível, dessa forma, realizarmos várias atividades ao mesmo tempo, como, por exemplo, falar, ver, pensar, sentir, mantermo-nos em pé, etc. A Figura 3.4 mostra o córtex cerebral assinalando algumas regiões funcionais.

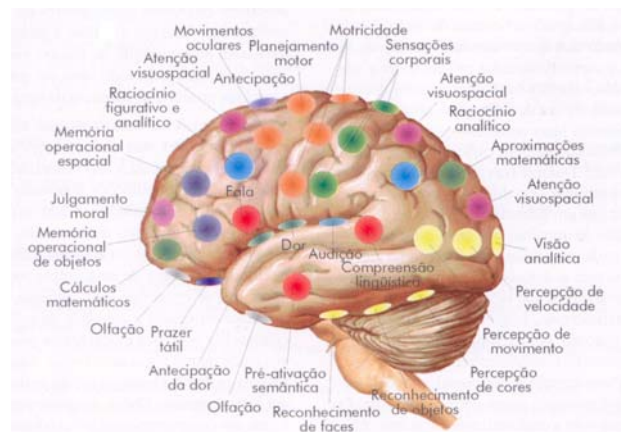


Figura 3.4 - Funções cerebrais
Fonte: Lent (2001, p. 23)

Os neurônios conectam-se entre si e aos órgãos através de um tipo especial de junção, denominada *sinapse*²⁶. A sinapse é o *chip* do sistema nervoso. É uma região de contato entre um terminal de uma fibra nervosa e um dendrito ou o corpo de uma outra célula. O tipo mais comum de sinapse é a sinapse química. A informação elétrica que chega ao elemento pré-sináptico é convertida em informação química através da liberação, na fenda sináptica (que tem de 20 a 50 nm de espessura), de uma substância armazenada no interior de vesículas, chamada de *neurotransmissor* (ou neuromediador). As moléculas dos neurotransmissores difundem-se, então, para a membrana pós-sináptica, onde a informação química é reconvertida em informação elétrica, como mostra a Figura 3.5.

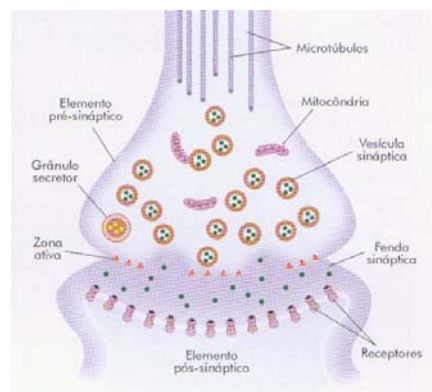


Figura 3.5 - Representação de uma sinapse
Fonte: Lent (2001, p. 103)

Essa dupla conversão de informação²⁷ permite que haja interferência sobre seu conteúdo, podendo ser comparada assim, a um verdadeiro microcomputador biológico. Na sinapse, a informação não é apenas transferida, mas transformada. Como cada neurônio pode receber milhares de sinapses ao mesmo tempo, toda essa informação deve ser combinada para orientar os sinais que a célula nervosa enviará a outras células.

Note que isso só é possível porque a membrana plasmática do neurônio possui a propriedade de excitabilidade: o interior da célula é negativo e o exterior é positivo. Assim, certos estímulos externos, ou mesmo produzidos internamente, podem provocar a abertura de canais moleculares na membrana, que permitem a passagem seletiva de certos íons, de dentro para fora e de fora para dentro. A diferença de potencial entre um

²⁶ A prova morfológica da existência das sinapses foi dada em 1954, com a ajuda do microscópio eletrônico e novas técnicas de coloração. Porém, já em 1897, Charles Sherrington havia postulado que os neurônios comunicavam-se com outros tipos celulares através de sinapses. Aliás, foi ele o responsável por essa denominação que, em grego, significa "prender".

²⁷ De acordo com Haykin (2001, p. 33), um elemento com essa propriedade é chamado de *dispositivo de dois terminais não-recíproco* na terminologia elétrica.

lado e outro da membrana pode mudar de valor e até mesmo, inverter a sua polaridade. As sinapses podem ser classificadas funcionalmente em inibitórias e excitatórias, dependendo da natureza e quantidade dos neurotransmissores liberados. Uma sinapse impõe ao neurônio receptivo excitação ou inibição, mas não ambas. Essa é a chamada propriedade de *tudo-ou-nada* do neurônio.

Em geral, os neurônios são bem mais lentos do que as portas lógicas em silício. Porém, essa velocidade mais lenta é compensada pelo imenso número de neurônios existentes no cérebro, o que explica a sua eficiência. A tabela 2.1 mostra as principais diferenças entre o computador e o cérebro humano.

Tabela 3.1 - Diferenças entre o computador e o cérebro humano

Elementos computacionais	Computador	Cérebro
Velocidade de processamento	10 ⁻⁹ segundos	10 ⁻³ segundos
Tipo de processamento	serial	paralelo
Confiabilidade dos elementos	confiável	não-confiável
Tolerância a falhas	pequena	grande
Tipo de sinal	preciso, simbólico	impreciso
Tipo de controle	centralizado	distribuído
Armazenamento de informação	substituível	adaptável

Fonte: Cottrell (1985) e Simpson (1990), extraída de Fernandes e Portugal (1996, p. 4)

3.1.1 Desenvolvimento das RNAs

O primeiro trabalho que procurou cumprir a tarefa de imitar o funcionamento do cérebro foi McCulloch e Pitts (1943)²⁸. O modelo de neurônio proposto era bem simples e seu funcionamento estava baseado na hipótese que o neurônio é um dispositivo binário, ou seja, dispara (saída igual a 1) ou não dispara (saída igual a 0), e que a rede neural é constituída por conexões sem pesos ajustáveis (portanto, não há aprendizado), que podem ser inibitórias ou excitatórias. Uma conexão inibitória envia um sinal negativo para o neurônio enquanto uma conexão excitatória envia um sinal positivo. Cada neurônio possui um limiar fixo e só dispara se a sua saída (a soma dos sinais enviados) for maior ou igual a este limiar. A Figura 3.6 mostra o neurônio artificial de McCulloch e Pitts.

²⁸ Warren McCulloch era um neuroanatomista e psiquiatra e Walter Pitts um matemático que se juntou a ele em 1942 (BRAGA, CARVALHO; LUDERMIR, 2000, p. 2).

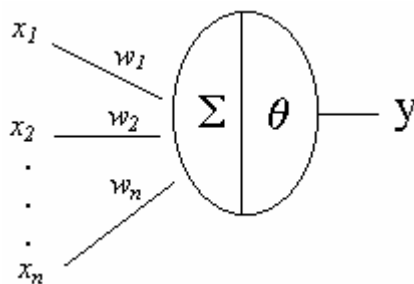


Figura 3.6 - O neurônio artificial de McCulloch e Pitts

Fonte: Elaborado pela autora

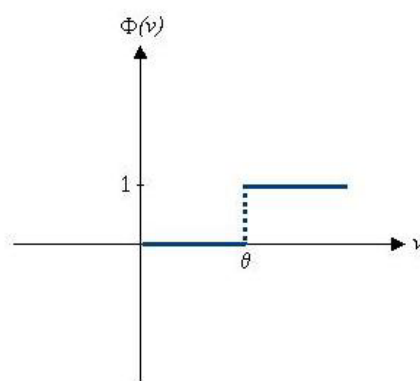
Esse neurônio pode ser escrito matematicamente como:

$$y = \Phi \left(\sum_{i=1}^n w_i x_i \right)$$

onde y é o produto, w_i são os pesos, x_i as entradas do neurônio, θ é o limiar e Φ é chamada função de ativação, e nesse caso, é a função limiar²⁹:

$$\Phi(v) = \begin{cases} 1, & \text{se } v \geq \theta \\ 0, & \text{se } v < \theta \end{cases}$$

Cujo gráfico é dado por:



Podemos fazer a seguinte analogia entre o neurônio artificial e o neurônio biológico:

Tabela 3.1 – Neurônio Biológico × Neurônio Artificial

Neurônio biológico	Neurônio artificial
estímulos	entradas
dendritos	conexões que ligam as entradas ao neurônio
sinapses	pesos
corpo celular	função de ativação
axônios	saída (produto)

Fonte: Elaborada pela autora

²⁹ Também conhecida como *função de Heaviside* (HAYKIN, 2001, p. 39).

Uma rede de neurônios de McCulloch e Pitts pode computar qualquer função lógica³⁰. A seguir, veremos alguns exemplos. Serão consideradas apenas duas entradas, x_1 e x_2 , e uma saída, y . O número 1 está associado à resposta "verdadeiro" e o número 0 à resposta "falso".

a) *Função e*:

A função *e* dá a resposta "verdadeiro" se todos os valores forem verdadeiros e a resposta "falso", caso contrário. Temos, então, a seguinte *tabela verdade*:

x_1	x_2	y
1	1	1
1	0	0
0	1	0
0	0	0

A Figura 3.7 mostra um neurônio de McCulloch e Pitts para computar esta função.

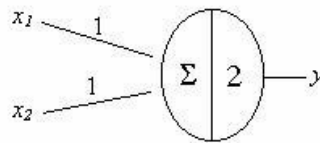


Figura 3.7 - Função *e*
Fonte: Elaborada pela autora

Nesse caso, temos que:

$$\Phi(x_1+x_2) = \begin{cases} 1, & \text{se } x_1 + x_2 = 2 \\ 0, & \text{se } x_1 + x_2 < 2 \end{cases}$$

b) *Função ou*:

A função **ou** dá a resposta verdadeiro se uma das entradas é verdadeira e a resposta falso, caso contrário. A *tabela verdade* desta função é a seguinte:

³⁰ Para detalhes, consulte Rojas (1996, cap.2).

x_1	x_2	y
1	1	1
1	0	1
0	1	1
0	0	0

A Figura 3.8 mostra um neurônio de McCulloch e Pitts para computar esta função:

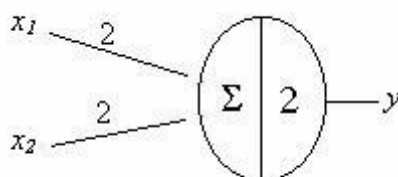


Figura 3.8 - Função ou
Fonte: Elaborada pela autora

Assim sendo, a função de ativação é:

$$\Phi(2x_1 + 2x_2) = \begin{cases} 1, & \text{se } 2x_1 + 2x_2 \geq 2 \\ 0, & \text{se } 2x_1 + 2x_2 < 2 \end{cases}$$

c) *Função e não*:

A função **e não** dá a resposta "verdadeiro" se a primeira entrada é verdadeira e a segunda é falsa, e dá a resposta "falso", caso contrário. A *tabela verdade* é, então:

x_1	x_2	y
1	1	0
1	0	1
0	1	0
0	0	0

A Figura 3.9 mostra um neurônio de McCulloch e Pitts para computar esta função.

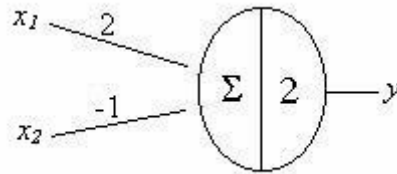


Figura 3.9 - Função e não
Fonte: Elaborada pela autora

E a função de ativação é dada por:

$$\Phi(2x_1 - x_2) = \begin{cases} 1, & \text{se } 2x_1 - x_2 = 2 \\ 0, & \text{se } 2x_1 - x_2 < 2 \end{cases}$$

d) *Função ou exclusivo*:

A função **ou exclusivo** (XOR) dá a resposta "verdadeiro" se apenas uma das entradas é verdadeira e dá a resposta "falso", caso contrário. A sua *tabela verdade* é:

x_1	x_2	y
1	1	0
1	0	1
0	1	1
0	0	0

Nesse caso, não podemos computar esta função utilizando apenas um neurônio de McCulloch e Pitts. Precisaremos de mais neurônios. Isso pode ser facilmente feito, lembrando que a função *ou exclusivo* pode ser escrita como:

$$x_1 \text{ XOR } x_2 = (x_1 \text{ e não } x_2) \text{ ou } (x_2 \text{ e não } x_1)$$

A figura 3.10 mostra uma rede de neurônios de McCulloch e Pitts para computar esta função:

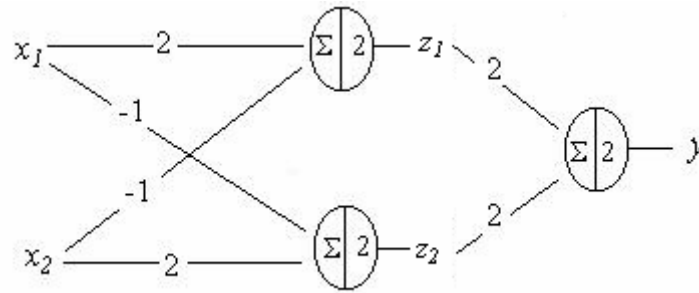


Figura 3.10 - Função ou exclusivo
Fonte: Elaborada pela autora

Os dois neurônios intermediários (que produzem as saídas z_1 e z_2) computam a função *e não*, enquanto o último neurônio computa a função *ou*. Assim, temos que:

$$\Phi(2x_1 - x_2) = \begin{cases} 1, & \text{se } 2x_1 - x_2 = 2 \\ 0, & \text{se } 2x_1 - x_2 < 2 \end{cases}, \text{ para } z_1$$

$$\Phi(-x_1 + 2x_2) = \begin{cases} 1, & \text{se } -x_1 + 2x_2 = 2 \\ 0, & \text{se } -x_1 + 2x_2 < 2 \end{cases}, \text{ para } z_2$$

$$\Phi(2z_1 + 2z_2) = \begin{cases} 1, & \text{se } 2z_1 + 2z_2 = 2 \\ 0, & \text{se } 2z_1 + 2z_2 < 2 \end{cases}, \text{ para } y$$

Em 1949, o psicólogo Donald Hebb apresentou a primeira regra de aprendizado para RNAs. Essa regra baseava-se em estudos biológicos do cérebro e consistia basicamente em aumentar a força da conexão sináptica entre dois neurônios que estivessem simultaneamente ativos.

Em 1958, o psicólogo Frank Rosenblatt introduziu um novo modelo, o *perceptron*. Ele criticou o modelo de McCulloch e Pitts por não ser realisticamente biológico. A idéia básica de seu trabalho era a de ensinar a rede até que ela aprendesse uma função desejada, assim como ocorre nos sistemas nervosos biológicos. Rosenblatt propôs uma estrutura de rede com neurônios de McCulloch e Pitts organizados em camadas e apresentou também uma regra de aprendizagem: o procedimento de convergência do perceptron (PCP). Posteriormente, em 1962, Rosenblatt apresentou o teorema de convergência do perceptron, demonstrando que seu algoritmo sempre converge, desde que o problema seja linearmente separável (BRAGA; CARVALHO; LUDERMIR, 2000, p. 35). O modelo de Rosenblatt avançou em relação ao de McCulloch e Pitts no sentido de estender a possibilidade dos neurônios assumirem

valores contínuos e não apenas binários e na introdução de uma lei de treinamento (KOVÁCS, 1996, p. 48). O perceptron causou grande entusiasmo na área de RNAs na época. De acordo com Haykin (2001, p. 65), acreditava-se que este modelo pudesse resolver qualquer problema.

Bernard Widrow e Marcian Hoff, em 1960, introduziram um modelo neural linear bastante simples denominado ADALINE (*ADaptive LINear Element*) e, posteriormente, o MADALINE (Múltipla ADALINE), cuja principal contribuição foi a criação de um princípio de treinamento denominado de *Regra Delta* ou *algoritmo dos mínimos quadrados médios*. Essa regra de aprendizagem consiste em mudar os pesos na direção que faça o erro diminuir. De acordo com Fauset (1994, p. 23), este princípio de aprendizagem foi o precursor do método de *retropropagação (backpropagation) do erro*, utilizado para o treinamento de redes de neurônios de múltiplas camadas.

Minsky e Papert (1969), uns dos líderes da escola de pensamento chamada Inteligência Artificial³¹ (SMITH, 1993, p.8) , criticaram o modelo *perceptron* de Rosenblatt mostrando matematicamente que ele era incapaz de resolver problemas relativamente simples que não fossem linearmente separáveis (como a implementação do *ou exclusivo - XOR*). Além disso, segundo eles, não havia razão para supor que qualquer uma das limitações do perceptron de única camada poderia ser superada pela versão de múltiplas camadas (HAYKIN, 2001, p. 65). Depois da publicação desse trabalho, a pesquisa em RNAs ficou praticamente estagnada. Mas, apesar do descrédito e da falta de financiamento, alguns pesquisadores continuaram trabalhando, como por exemplo, Teuvo Kohonen, Christoph von der Malsburg, Kunihiko Fukushima, Stephen Grossberg, Gail Carpenter, James Anderson, e foram realizados avanços significativos na área.

Atribui-se o passo inicial para a retomada de interesse em RNAs ao trabalho do físico John Hopfield em 1982, que chamou a atenção para as propriedades associativas desses modelos. De acordo com Yim (2001, p. 14), este trabalho mostrou de forma clara e matemática como era o funcionamento destas RNAs e o que elas eram capazes de fazer.

³¹ Para uma discussão sobre RNAs e Inteligência Artificial, veja Haykin (2001, p. 59-63).

Em 1986, David Rumelhart, Geoffrey Hinton e Ronald Williams apresentaram o algoritmo de retropropagação dos erros (*backpropagation of errors*)³², capaz de treinar redes de múltiplas camadas, o que contribuiu fortemente para o interesse renovado em RNAs. Vários outros fatos contribuíram para isso. Podemos citar, por exemplo, o trabalho de Charles Rosenberg e Terry Sejnowski, que usaram o algoritmo de retropropagação para desenvolver o modelo NET talk, que transforma textos escritos em inglês em uma representação fonética (SMITH, 1993, p. 13). Além disso, o avanço da tecnologia e o fato da escola simbolista não ter conseguido os avanços prometidos, contribuíram também para o "renascimento" das RNAs.

A partir desse renovado entusiasmo, a área de RNAs vem se desenvolvendo nos últimos anos, e o número de suas aplicações tem crescido substancialmente. Atualmente, os modelos de RNAs têm sido aplicados à medicina, geologia, oceanografia, engenharia, economia, entre várias outras. Novos modelos foram e continuam sendo desenvolvidos e melhorados, como as redes ART, o modelo neocognitron, as máquinas de comitê, etc. Nas palavras de Haykin (2001, p. 69):

As RNAs certamente trilharam um longo caminho desde os dias iniciais de McCulloch e Pitts. De fato, elas se estabeleceram como um tema interdisciplinar com raízes profundas em neurociências, psicologia, matemática, ciências físicas e engenharia. É desnecessário dizer que elas estão aqui para ficar e que continuarão a crescer em teoria, projeto e aplicações.

3.2 Aplicações de RNAs à Economia

Os primeiros trabalhos com aplicações das RNAs na economia parecem ter sido aqueles apresentados na 1ª. *IEEE International Conference on Neural Networks*³³, em 1988. Foram três os trabalhos com aplicação econômica apresentados nessa ocasião:

- White, Halbert. *Economic Prediction using neural networks: the case of IBM daily stock returns.*

³² Como salienta Smith (1993, p. 12), a retropropagação dos erros é um exemplo de invenção múltipla, já que David Parker e Yann Le Cun, independentemente, publicaram descobertas semelhantes. Além disso, Paul Werbos, já em 1974, em sua tese de doutoramento, havia exposto as técnicas envolvidas.

³³ Essa opinião é compartilhada por Simon (2002).

- Dutta, Soumitra; Shekhar, Shashi. *Bond rating: a non-conservative application of neural networks*.
- Collins, Edward; Ghosh, Sushmito; Scofield, Christopher. *An application of a multiple neural network learning system to emulation of mortgage underwriting judgements*.

Os dois primeiros trabalhos citados acima, utilizaram as RNAs para previsão de séries econômicas. Já o último aplicou as RNAs de uma forma mais conservadora para a época, utilizando-a como um modelo para tomada de decisão.

Desde então, os modelos de RNAs têm sido bastante aplicados na economia, especialmente na área de finanças. E o seu uso não se restringe apenas a modelos de previsão de séries econômicas. Wong e Selvi (1998) fazem uma revisão de parte da literatura empírica publicada entre 1990 e 1996 (97 artigos) e concluem que os temas mais visados pelos pesquisadores são: previsão de falência de empresas, precificação de opções, previsão de séries temporais financeiras, seleção de ações, previsão da performance de portfólios, *rating* de títulos, sistemas de negociações e apoio à tomada de decisão, avaliação de risco de crédito, dentre outros.

Mais recentemente, tem sido explorado o uso das RNAs para construção, por exemplo, de testes de heterocedasticidade condicional, não-linearidade, raiz unitária, co-integração e quebra estrutural, os quais aproveitam-se principalmente da capacidade desses modelos em extrair relações não-lineares complexas e a da não necessidade de especificação de uma forma funcional.

Um outro campo recente de aplicação das RNAs é na teoria econômica propriamente dita, mais especificamente, na área de economia computacional, onde são construídas economias habitadas por agentes artificiais. A maioria dos modelos econômicos existentes dá aos agentes muito mais conhecimento do que eles realmente possuem. Por exemplo, assume-se que os agentes têm pleno conhecimento da verdadeira estrutura da economia e que possuem informação completa. Ao abandonarmos tal hipótese, estaremos tratando de forma mais real o processo pelo qual os agentes tomam suas atitudes e, por exemplo, formam suas expectativas sobre o futuro. Além disso, os agentes têm diferentes graus de informação, diferentes experiências de vida e diferentes atitudes em relação ao risco, ou seja, são heterogêneos, ao contrário do tratamento dado pela hipótese do agente representativo. Para uma

introdução a esse interessante campo de pesquisa, veja Beltratti, Margarita e Terna (1996).

3.3 Aspectos Teóricos das Redes Neurais Artificiais

Uma RNA pode ser caracterizada por três elementos essenciais (FAUSET, 1994, p. 3): função de ativação, arquitetura da rede, aprendizagem (treinamento), os quais serão detalhados a seguir.

3.3.1 Função de ativação

A função de ativação (ou função de transferência) produz a saída de um neurônio, utilizando como argumento a soma de suas entradas, ponderada pelos pesos das conexões, como mostra a Figura 3.11.

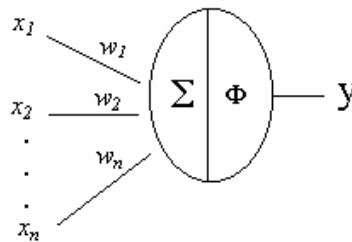


Figura 3.11 - Função de ativação em um neurônio artificial
Fonte: Elaborada pela autora

Dessa forma, podemos escrever a saída y de um neurônio como:

$$y = \Phi(v)$$

onde $v = \sum_{i=1}^n x_i w_i$ e $\Phi(\cdot)$ é a função de ativação.

As principais funções de ativação são as seguintes:

a) função identidade (ou “linear pura”):

$$\Phi(v) = v$$

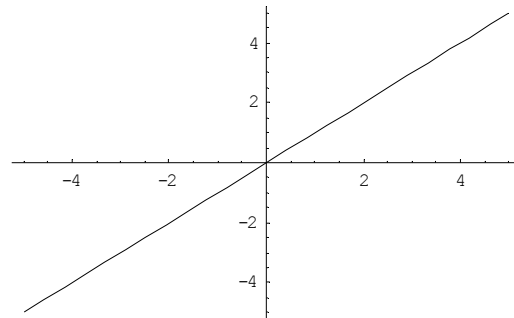


Figura 3.12- Função identidade
Fonte: Elaborada pela autora

Essa função é geralmente utilizada no neurônio de saída.

b) função de limiar ou de *Heaviside*:

$$\Phi(v) = \begin{cases} 1 & \text{se } v \geq \theta \\ 0 & \text{se } v < \theta \end{cases}$$

A Figura 3.13 mostra a função de limiar para $\theta = 2$:

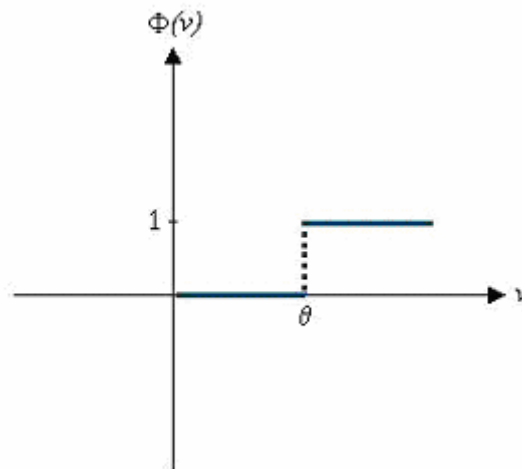


Figura 3.13 - Função de limiar
Fonte: Elaborada pela autora

Esta função é a utilizada pelo neurônio de McCulloch e Pitts. É bastante adequada em problemas de classificação de padrões.

c) função de limiar bipolar

$$\Phi(v) = \begin{cases} 1 & \text{se } v \geq \theta \\ -1 & \text{se } v < \theta \end{cases}$$

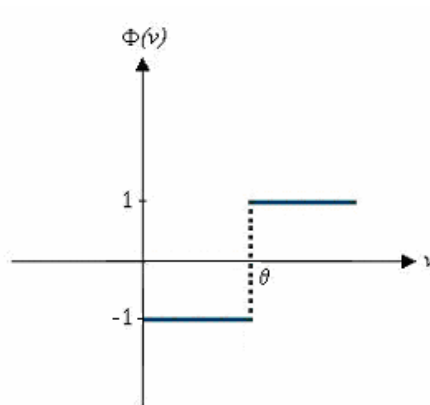


Figura 3.14 - Função de limiar bipolar

Fonte: Elaborada pela autora

d) função logística:

$$\Phi(v) = \frac{1}{1 + e^{-av}},$$

onde a é o parâmetro de inclinação dessa função. A Figura 3.15 mostra o gráfico da função logística para dois diferentes valores de a .

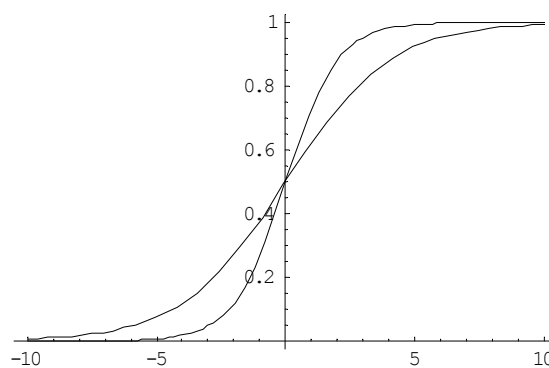


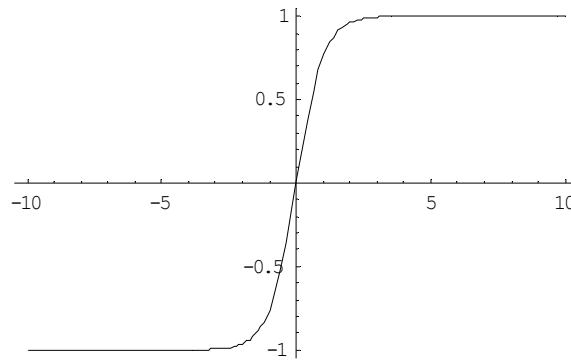
Figura 3.15 - Função logística

Fonte: Elaborada pela autora

A função logística é a mais utilizada para redes com múltiplas camadas treinadas com o algoritmo de retropropagação dos erros. Como veremos adiante, essa função possui características bastante interessantes, como por exemplo, a não-linearidade suave, continuidade, monotonicidade e diferenciabilidade.

e) função tangente hiperbólica (ou sigmóide bipolar)

$$\Phi(v) = \frac{1 - e^{-av}}{1 + e^{-av}}$$



Função 3.16 - Função tangente hiperbólica

Fonte: Elaborada pela autora

Esta função possui as mesmas características da função logística, com a diferença que seus valores de saturação são -1 e 1, sendo, assim, bastante utilizada em redes multicamadas treinadas com o algoritmo de retropropagação dos erros.

3.3.2 Arquitetura de RNAs

A arquitetura da rede refere-se à maneira pela qual os neurônios estão estruturados. Na definição da arquitetura, uma série de escolhas deve ser feita, como por exemplo, o número de variáveis de entrada, o número de neurônios em cada camada e o número de camadas ocultas.

Em relação à arquitetura, as RNAs podem ser classificadas:

a) *Quanto ao número de camadas*

- redes de camada única: não existe nenhuma camada de neurônios entre as entradas e os neurônios de saída, como mostra a Figura 3.17. Esse tipo de rede está limitado a resolver apenas problemas que sejam linearmente separáveis.

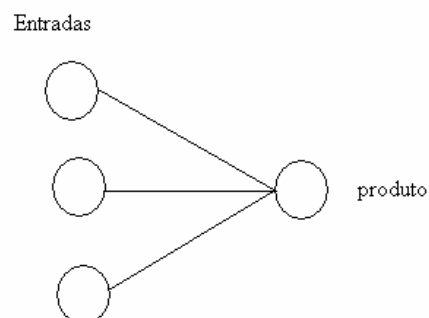


Figura 3.17 - RNA com uma única camada

Fonte: Elaborada pela autora

- redes de múltiplas camadas: existe uma ou mais camadas de neurônios entre as entradas e os neurônios de saída, como mostram as Figuras 3.18 e 3.19. Este tipo de rede tem um alto poder computacional.

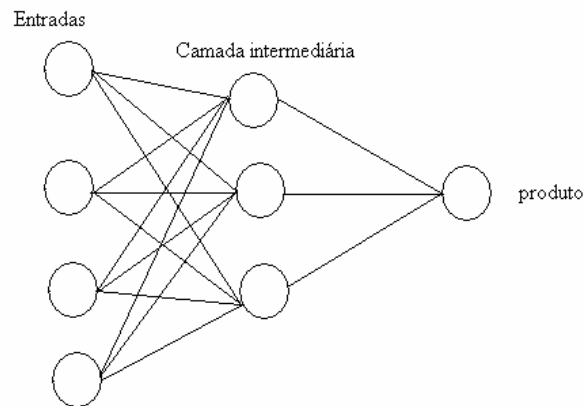


Figura 3.18 - RNA com uma camada oculta
Fonte: Elaborada pela autora

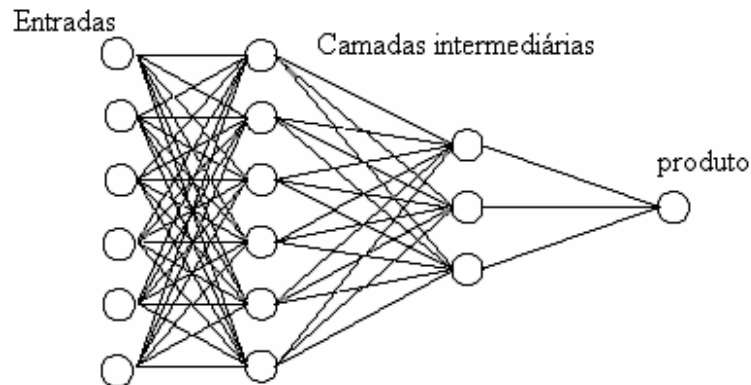


Figura 3.19 - RNA com 2 camadas ocultas

b) *Quanto ao tipo de conexão*

- redes acíclicas ou de propagação para frente (*feedforward*): a saída de um neurônio não pode servir como entrada para neurônios que estão numa mesma camada. Essas redes, dessa forma, são ditas “sem memória”. As Figuras 3.17 a 3.19 são exemplos de redes de propagação para frente.
- redes cíclicas ou recorrentes: nesse tipo de rede, a saída de um neurônio pode servir como entrada para ele novamente, ou para um neurônio da mesma camada, assim

como de uma camada anterior. Como este tipo de rede possui realimentação, diz-se que “possui memória”. Uma RNA recorrente não é necessariamente organizada em camadas, como pode ser visto a seguir.

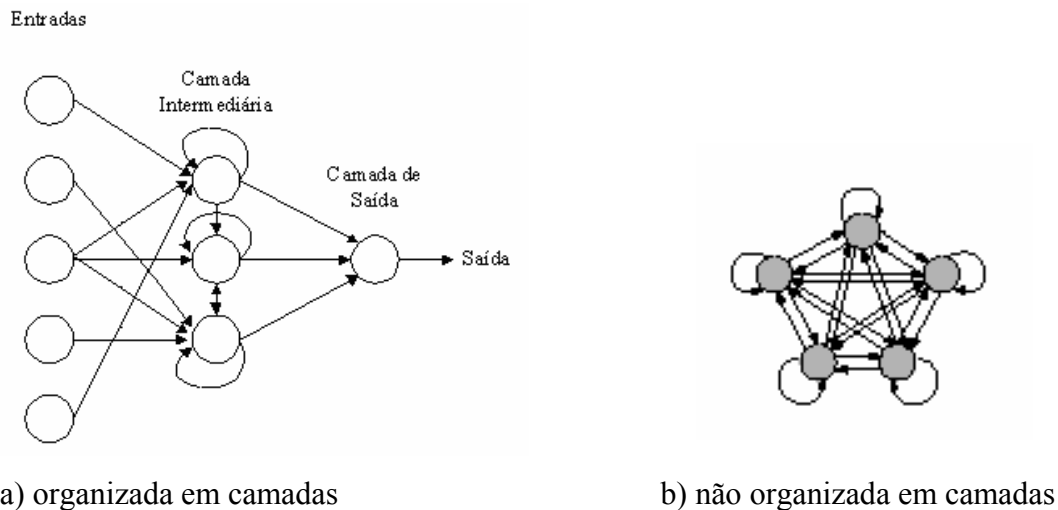


Figura 3.20 - RNAs recorrentes

Fonte: Elaborada pela autora

c) *Quanto ao tipo de conectividade*

- redes parcialmente conectadas: nem todos os neurônios de uma camada estão conectados a todos os outros neurônios da próxima camada. Um exemplo disso é mostrado na Figura 3.20 –(a).
- redes inteiramente conectadas: todos os neurônios de uma camada estão conectados aos neurônios da camada seguinte. As Figuras 3.17 a 3.19 mostram redes inteiramente conectadas.

d) *Quanto à estrutura*

- redes estáticas: a estrutura da rede não é alterada durante o treinamento.
- redes auto-organizadas: o número de neurônios e o número de camadas são alterados durante o treinamento.

A arquitetura de rede mais utilizada em aplicações econômicas é a rede multicamadas *feedforward* inteiramente conectada que, de agora em diante, será chamada apenas de MLP (*multi-layer perceptron*).

3.3.3 Aprendizagem em RNAs - paradigmas e algoritmos de treinamento

Uma das características mais atraentes das RNAs é a sua capacidade de “aprender”³⁴ através de exemplos e fazer generalizações, ou seja, produzir respostas adequadas para novas entradas. Existem dois paradigmas de aprendizagem em RNAs: *supervisionado* e *não-supervisionado*. Na aprendizagem supervisionada, a rede conta com um "professor", ou seja, as saídas desejadas são conhecidas. Na aprendizagem não supervisionada, não existe um "professor", i.e., não há conhecimento das saídas desejadas. A rede apenas recebe entradas que descrevem o ambiente e, a partir disso, tenta codificar essas características. Esse paradigma de aprendizagem só é possível quando existe redundância nos dados de entrada. Para os problemas de previsão, obviamente, utiliza-se o paradigma de aprendizagem supervisionada.

O conjunto de regras para a solução de um problema de aprendizagem é chamado de algoritmo³⁵ de aprendizagem. Existem diversos algoritmos de aprendizagem, cada um com vantagens e desvantagens específicas e que se diferenciam basicamente pela forma como é realizado o ajuste dos pesos. A seguir, veremos as primeiras regras de aprendizagem em RNAs, terminando com o algoritmo de retropropagação dos erros e uma de suas variações.

3.3.3.1 Regra de Hebb

A primeira regra de aprendizagem em redes neurais foi proposta pelo psicólogo Donald Hebb e consiste basicamente em aumentar a força da conexão sináptica entre dois neurônios que estão simultaneamente ativos. Em seu livro *The Organization of Behavior*, ele escreve:

Quando um axônio da célula A está perto o suficiente para excitar uma célula B e participa do seu disparo repetida ou persistentemente, então algum processo de crescimento ou modificação metabólica acontece em uma das células ou em ambas, de tal forma que a eficiência de A como uma das células que dispara B é aumentada (HEBB, 1949, p. 62 apud HAYKIN, 2001, p. 80).

³⁴ Na seção 3.4 será feita uma discussão sobre o significado de “aprender”. A aprendizagem em RNAs ocorre através de um processo iterativo de ajustes dos parâmetros livres de sua arquitetura.

³⁵ Palavra derivada do nome do matemático persa *Mohammed al-Kowârisimi*, que desenvolveu as regras para adição, subtração, multiplicação e divisão para números decimais ordinários (HAYKIN, 2001, p. 132).

Apesar de não ter formulado sua regra em termos matemáticos, podemos escrevê-la matematicamente como³⁶:

$$\Delta w(n) = \eta y(n)x(n)$$

Portanto, os valores atualizados dos pesos sinápticos serão dados por:

$$w(n+1) = w(n) + \Delta w(n)$$

$$w(n+1) = w(n) + \eta y(n)x(n)$$

3.3.3.2 Regra de aprendizagem do perceptron

O perceptron atualiza seus pesos de acordo com a seguinte regra:

$$\Delta w = \eta ex(t)$$

Dessa forma, os valores atualizados dos pesos serão dados por:

$$w(n+1) = w(n) + \Delta w_{ij}(n)$$

$$w(n+1) = w(n) + \eta ex(n)$$

3.3.3.3 Regra delta (ou de Widrow-Hoff)

A regra delta (ou algoritmo do mínimo quadrado médio) consiste em encontrar o mínimo global da função de erro quadrático médio através da descida do gradiente, utilizando o *algoritmo da descida mais íngreme (steepest descent algorithm)* ou do *gradiente decrescente*. Este método de otimização consiste em ajustar os pesos na direção da descida mais íngreme, ou seja, na direção contrária ao vetor gradiente dos pesos w_i . O vetor gradiente é dado por:

$$g = \nabla \varepsilon(w)$$

³⁶ Essa expressão matemática para a regra de Hebb é também conhecida como *regra do produto das atividades*.

Dessa forma, podemos escrever o algoritmo da descida mais íngreme como:

$$\boxed{w(n+1) = w(n) - \eta g(n)}$$

- *A regra Delta de Widrow-Hoff*

O erro de uma rede linear (ADALINE) é dado por:

$$e(n) = y^d(n) - wx(n)$$

onde:

y^d = resposta desejada (alvo)

$w.x(n)$ = produto gerado pela rede (pesos×entradas)

E o quadrado do erro³⁷:

$$\varepsilon(w) = \frac{1}{2} e^2(n)$$

Derivando $\varepsilon(w)$ em relação ao vetor de pesos w , temos:

$$\frac{\partial \varepsilon(w)}{\partial w} = e(n) \times [-x(n)]$$

$$\frac{\partial \varepsilon(w)}{\partial w} = -x(n)e(n)$$

Temos agora uma estimativa para o vetor gradiente. Pelo método da descida mais íngreme, a regra de atualização dos pesos é dada então por:

$$\boxed{\Delta w(n) = -\eta x(n)e(n)}$$

E os pesos atualizados:

$$w(n+1) = w(n) + \Delta w(n)$$

$$w(n+1) = w(n) - \eta x(n)e(n)$$

3.3.3.4 Regra de retropropagação dos erros (*backpropagation*)

³⁷ O valor $\frac{1}{2}$ é utilizado apenas para facilitar os cálculos posteriores.

O algoritmo de retropropagação dos erros é uma generalização da regra delta, sendo, por isso, também conhecido como *regra delta generalizada*. É utilizado para o treinamento de redes de propagação para frente (*feedforward*) com múltiplas camadas. Nesse tipo de rede, os pesos de uma unidade contribuem para o erro dessa. Porém, se existe um erro, quais desses muitos pesos podem ter o causado (já que um neurônio recebe sinais de todos os outros que estão na camada anterior)? Esse é o chamado problema de atribuição de crédito.

O erro no neurônio de saída j , no tempo n , é dado por:

$$e_j(n) = y_y^d(n) - y_j(n)$$

E a soma dos erros ao quadrado:

$$\varepsilon(n) = \frac{1}{2} \sum e_j^2(n)$$

O que desejamos é minimizar $\varepsilon(n)$ médio (ou seja, o erro quadrático médio - $\frac{1}{N} \sum_{n=1}^N \varepsilon(n)$). Pelo método do gradiente decrescente, sabemos que devemos ajustar os pesos na direção contrária ao gradiente dos pesos:

$$\Delta w \propto - \frac{\partial \varepsilon(n)}{\partial w_{ji}(n)}$$

$$\Delta w = -\eta \frac{\partial \varepsilon(n)}{\partial w_{ji}(n)}$$

A constante η dá o tamanho do passo que será dado na superfície de erro. Se for muito pequena, o aprendizado será lento. Se for muito grande, os resultados poderão oscilar muito e a rede não convergir.

Devemos calcular agora $\frac{\partial \varepsilon(n)}{\partial w_{ji}(n)}$. Pela regra da cadeia temos que³⁸:

$$\frac{\partial \varepsilon}{\partial w_{ji}} = \frac{\partial \varepsilon}{\partial v_j} \frac{\partial v_j}{\partial w_{ji}}$$

³⁸ Para não carregar a notação, iremos suprimir o índice (n) na derivação das fórmulas.

$$\text{com } v_j = \sum_{i=1}^n x_i w_{ji} .$$

A segunda derivada da expressão acima é dada por:

$$\frac{\partial v_j}{\partial w_{ji}} = \frac{\partial \sum_{i=1}^n x_i w_{ji}}{\partial w_{ji}} = x_i$$

E a primeira, pela regra da cadeia novamente:

$$\delta_j = \frac{\partial \varepsilon}{\partial v_j} = \frac{\partial \varepsilon}{\partial y_j} \frac{\partial y_j}{\partial v_j}$$

Comecemos pela derivada mais simples:

$$\frac{\partial y_j}{\partial v_j} = \frac{\partial \Phi(v_j)}{\partial v_j} = \Phi'(v_j)$$

onde $\Phi(\cdot)$ é a função de ativação do neurônio.

O cálculo de $\frac{\partial \varepsilon}{\partial y_j}$ depende da camada onde está o neurônio j . Se for um

neurônio da camada de saída, temos:

$$\frac{\partial \varepsilon}{\partial y_j} = \frac{\partial \left(1/2 \sum_{i=1}^n (y_i^d - y_i)^2 \right)}{\partial y_j} = (y_j^d - y_j)$$

Portanto δ_j será dado por:

$$\delta_j = (y_j^d - y_j) \times \Phi'(v_j)$$

Mas, se o neurônio j estiver na camada oculta, temos:

$$\frac{\partial \varepsilon}{\partial y_j} = \sum_{l=1}^M \frac{\partial \varepsilon}{\partial v_l} \frac{\partial v_l}{\partial y_j} = \sum_{l=1}^M \frac{\partial \varepsilon}{\partial v_l} \frac{\partial \sum_{i=1}^n w_{il} y_i}{\partial y_j} = \sum_{l=1}^M \frac{\partial \varepsilon}{\partial v_l} w_{jl} = \sum_{l=1}^M \delta_l w_{jl} \quad)$$

Assim:

$$\delta_j = \frac{\partial \varepsilon}{\partial v_j} = \Phi'(v_j) \times \sum_{l=1}^M \delta_l w_{jl}$$

A variação dos pesos é dada então por:

$$\Delta w_{ji} = \eta \delta_j x_i$$

E os pesos atualizados:

$$w_{ji}(n+1) = w_{ji}(n) + \eta \delta_j(n) x_i(n)$$

Lembrando que:

$$\delta_j = \begin{cases} \Phi'(v_j) \times \sum_{l=1}^M \delta_l w_{jl}, & \text{se o neurônio estiver na camada oculta} \\ \Phi'(v_j) \times (y_j^d - y_j), & \text{se o neurônio estiver na camada de saída} \end{cases}$$

Podemos, então, distinguir dois passos de computação na aplicação do algoritmo de retropropagação. O passo para frente, que consiste simplesmente em calcular os produtos para cada neurônio de saída, utilizando os pesos atuais e calcular o erro da rede (e_j). E o passo para trás, onde os erros da camada de saída são retropropagados para as camadas anteriores e os pesos são atualizados.

Assim, para calcularmos δ_j precisamos da derivada da função de ativação associada ao neurônio j . Para que isso seja possível, a função de ativação deve ser contínua.

3.3.3.4.1 Problemas com o algoritmo de retropropagação

O maior problema com o algoritmo de retropropagação é a possibilidade de obter um mínimo local da função de erro. Dessa forma, a melhor solução pode não ser encontrada. De acordo com Fernandes e Portugal (1996, p. 15), existem várias

alternativas para diminuir a possibilidade de ocorrência disso, tais como a diminuição da taxa de aprendizagem, adição de neurônios na camada oculta e introdução de um termo de momento para a atualização dos pesos.

Um outro problema com esse algoritmo é a definição dos seus parâmetros. Não existe uma regra estabelecida para escolhermos a taxa de aprendizagem, os valores iniciais dos pesos, etc. Esse problema tem recebido bastante atenção nos últimos anos e alguns avanços foram feitos. Porém, essa tarefa continua sendo basicamente um processo de tentativa e erro.

O aprendizado pela retropropagação dos erros pode ser muito lento para superfícies mais complexas, sendo necessárias milhares de épocas para a sua convergência. Vários métodos têm sido propostos para acelerar a convergência desse algoritmo. O mais popular é a adição de um termo de momento para a atualização dos pesos, que também é usada para evitar a ocorrência de mínimos locais. A atualização dos pesos com o termo de momento é dada por:

$$w_{ji}(n+1) = w_{ji}(n) + \eta \delta_j(n) x_i(n) + \alpha (w_{ji}(n) - w_{ji}(n-1))$$

em que α é o chamado termo de momento e está geralmente entre 0 e 1.

Apesar de sua grande popularidade, a retropropagação dos erros é criticada por não ter plausibilidade biológica. Mas, de acordo com Haykin (2001, p. 253-254), tais receios neurobiológicos não tiram a importância desse algoritmo como ferramenta para o processamento de informação, já que tem sido aplicado com sucesso em vários campos, inclusive na simulação de processos neurobiológicos.

3.3.3.4.2 O algoritmo do gradiente conjugado

O algoritmo de retropropagação dos erros ajusta os pesos na direção contrária ao gradiente, ou seja, na direção em que a função decresce mais rapidamente, o que não necessariamente produz uma convergência rápida. O algoritmo do gradiente conjugado, apesar de usar apenas informações do gradiente de primeira ordem³⁹ (não se constituindo num método de segunda ordem como é o método de otimização de Newton, por exemplo), é quase tão rápido quanto os métodos de segunda ordem, com a vantagem de não ser necessária a manipulação da matriz hessiana. O algoritmo do

³⁹ É importante notar que o método do gradiente conjugado faz uso da hessiana, porém, de forma implícita, ou seja, seu cálculo não é necessário.

gradiente conjugado assume a otimização de uma função quadrática e usa a propriedade das direções conjugadas, como veremos a seguir.

Considere a minimização da seguinte função quadrática:

$$f(x) = \frac{1}{2} x' A x - b' x + c$$

em que:

x é um vetor de parâmetros $W \times 1$

A é uma matriz $W \times W$, simétrica, definida positiva.

b é um vetor $W \times 1$

c é um escalar.

Um conjunto de vetores $\{p_k\}$ é um conjugado da matriz A se a seguinte condição for satisfeita⁴⁰:

$$p_i' A p_j = 0 \quad \forall i \text{ e } j \text{ tal que } i \neq j$$

Para um dado conjunto de vetores conjugados, o método da **direção conjugada** para a minimização da função quadrática $f(x)$ é dado por:

$$x(n+1) = x(n) + \eta(n)s(n) \quad n = 0, 1, \dots, W-1$$

onde:

$x(0)$ = vetor inicial

$\eta(n)$ é um escalar obtido através de: $\min_{\eta} f(x(n) + \eta s(n))$

A escolha de η , para n fixo, corresponde a uma busca em linha, que representa um problema de minimização unidimensional.

No algoritmo do **gradiente conjugado**, a busca é iniciada na direção negativa ao gradiente, ou seja:

$$d(0) = -g(0)$$

⁴⁰ Uma propriedade interessante dos vetores conjugados é que eles são linearmente independentes. Veja a prova em Haykin (2001, p. 264).

Uma busca em linha é então realizada para determinar a distância ótima. A próxima direção de busca será aquela que é conjugada à direção de busca anterior. O procedimento para determinar a nova direção de busca é combinar a nova direção do gradiente descendente com a direção de busca anterior. Dessa forma, os vetores de direção são determinados de forma seqüencial, conforme o método avança:

$$d(n+1) = -g(n+1) + \beta(n+1)d(n)$$

O algoritmo do gradiente conjugado ajusta o tamanho do passo a cada iteração. A busca é feita juntamente com a direção conjugada, para determinar o tamanho do passo que minimiza a função performance, através de uma busca em linha. Existem algumas diferentes formas para se calcular $\beta(n)$, algumas das quais são apresentadas a seguir:

Fórmula de Polak-Ribière:

$$\beta(n) = \frac{\Delta \mathbf{g}'_{n-1} \mathbf{g}_n}{\mathbf{g}'_{n-1} \mathbf{g}_{n-1}}$$

Fórmula de Fletcher e Reeves:

$$\beta(n) = \frac{\mathbf{g}'_n \mathbf{g}_n}{\mathbf{g}'_{n-1} \mathbf{g}_{n-1}}$$

Fórmula de Hestenes e Steifel:

$$\beta(n) = \frac{\Delta \mathbf{g}'_{n-1} \mathbf{g}_n}{\Delta \mathbf{g}'_{n-1} \mathbf{s}_{n-1}}$$

3.4 RNAs e Econometria

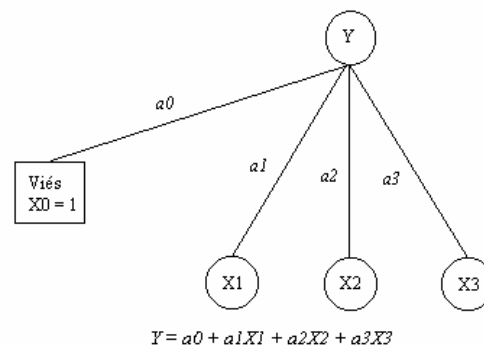
Apesar de terem sido inspiradas no funcionamento do cérebro, as ligações dos modelos de RNAs com os modelos econométricos são bem grandes. Na Tabela 3.2 apresentamos um pequeno dicionário de alguns termos utilizados em RNAs e seus correspondentes em estatística.

Tabela 3.2 - Redes Neurais × Econometria

Redes Neurais	Econometria
Pesos	Parâmetros
Conjunto de treinamento	Amostra
Entradas	Variáveis independentes/explicativas
Saídas/produtos	Valor estimado da variável dependente
Alvos	Variáveis dependentes/endógenas
Retropropagação	Aproximação estocástica
Treinamento/aprendizado	Estimação

Fonte: Adaptado de Fernandes, Portugal e Navaux (1996, 257)

A construção de uma RNA é equivalente à construção de uma fórmula matemática, só que de forma gráfica (DEBOECK, 1994, p. 4). O funcionamento de um modelo neural consiste, basicamente em, dados os valores de entrada (*inputs*), ajustar os pesos de forma que estes produzam o resultado desejado. A Figura 3.23 mostra uma RNA (ou seja, uma "fórmula gráfica") para o caso de uma regressão linear múltipla.

**Figura 3.23 - Exemplo de RNA para uma regressão linear múltipla**

Fonte: Deboeck (1994, p. 5)

As entradas X_i da RNA correspondem às variáveis explicativas do modelo de regressão linear múltipla, os pesos a_i , aos coeficientes que devem ser estimados e o produto (Y), aos valores estimados. Dessa forma, uma RNA de camada única, com função de ativação linear, é análoga a um modelo de regressão linear múltipla.

Uma RNA de camada única com mais de um produto pode ser equivalente a um modelo VAR (vetor auto-regressivo), se tivermos dados de séries temporais. Ou, se introduzirmos os próprios alvos como variáveis de entrada (tomando o cuidado, obviamente, de não fazer a conexão de uma variável com ela mesma), obteremos uma RNA equivalente a um sistema de equações simultâneas.

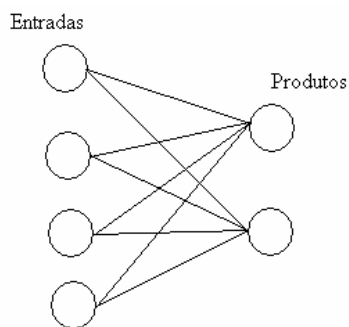


Figura 3.24 - RNA com dois produtos
Fonte: Elaborada pela autora

Uma RNA de camada única com função de ativação não-linear também tem o seu correspondente em econometria. Considere que na figura 3.23, a função de ativação seja a logística. Sabemos que o produto da rede, nesse caso, estará limitado a assumir valores entre 0 e 1. A rede resultante é, dessa forma, equivalente ao modelo logit. E se a função de ativação for a função densidade de probabilidade normal acumulada, temos o conhecido modelo Probit (ou Normit). E usando outras funções limitadas, teremos redes equivalentes a modelos econométricos que se aplicam a variáveis dependentes limitadas.

O poder computacional das RNAs em trabalhar com problemas complexos está na inclusão de camadas de neurônios intermediários entre as entradas e a camada de saída. E essas camadas ocultas não possuem um significado real, não têm interpretação. São apenas intermediárias no processo do cálculo do produto da rede. Dessa forma, essa “parte” de uma RNA não tem paralelo em econometria.

As RNAs podem ser vistas como métodos não paramétricos de estimação, já que não se preocupam em modelar o processo que gerou os dados, e apresentam algumas vantagens em relação aos métodos paramétricos tradicionais. Entre elas, podemos citar a não necessidade de se especificar uma forma funcional para a relação entre as variáveis (já que uma RNA pode aproximar bem qualquer função) e a sua grande habilidade em modelar relações não-lineares. Porém, as RNAs apresentam algumas desvantagens em relação aos métodos tradicionais. Uma delas é a não existência de uma metodologia padronizada para a construção desses modelos, sendo desenvolvidos, portanto, num processo de tentativa e erro (SHACHMUROVE, 2002, p. 27-28). Além disso, um modelo de RNA é considerado uma "caixa-preta", já que o processo gerador do resultado não pode ser explicitamente conhecido, sendo, portanto, incapaz de

explicar como e por que determinado produto foi gerado. Assim, os pesos estimados não possuem um significado, como no caso dos modelos de regressão.

Devido a sua grande popularidade nos anos recentes, as RNAs não ficaram livres de mitos. Algumas das vantagens e desvantagens freqüentemente citadas das RNAs não passam de pseudo-vantagens e pseudo-desvantagens. Por exemplo, é comum encontrar a afirmação que RNAs são “sistemas inteligentes que aprendem”. Porém, como já vimos, a capacidade de aprender das RNAs é simplesmente o resultado da aplicação de um algoritmo de otimização, como o gradiente descendente, por exemplo. Dessa forma, o modelo de regressão linear também deveria ser considerado um “sistema inteligente que aprende”, assim como vários outros métodos estatísticos⁴¹. Uma pseudo-desvantagem das RNAs freqüentemente citada é a sua quebra com a teoria econômica. Isso, no nosso entender, não é verdade⁴². A teoria econômica pode fornecer ao pesquisador informações sobre as variáveis explicativas a serem consideradas em um modelo e não necessariamente a forma funcional a ser utilizada. A quebra com a teoria na escolha de uma forma funcional, portanto, não deve ser considerada como uma fraqueza das RNAs.

Até pouco tempo atrás, uma das desvantagens freqüentemente citadas das RNAs era a não existência de métodos que permitissem a realização de inferência estatística formal com estes modelos (veja, por exemplo, Zhang et al., 1998, p. 56). Entretanto, Racine e White (2001) mostraram que é possível utilizar RNAs, mais precisamente, a rede MLP, para realizar inferência formal, utilizando técnicas de *bootstrap*⁴³. Dessa forma, podemos realizar testes de hipóteses a respeito da significância individual e conjunta das variáveis de entrada (*inputs*) da RNA. Como uma rede MLP tem a habilidade de extrair relações não-lineares complexas, a inferência baseada em tais modelos é um poderoso instrumento adicionado à caixa de ferramentas não só de economistas e econometristas, mas de pesquisadores em geral.

⁴¹ Essa questão foi ressaltada por Sarle (1994).

⁴² Essa opinião é compartilhada por Gonzalez (2000).

⁴³ A técnica de *bootstrap* consiste em utilizar a própria amostra para obter uma descrição das propriedades amostrais dos estimadores, ao invés de utilizar resultados teóricos amplos (GREENE, 2003, p. 924)

3.5 Inferência estatística válida em RNAs

Racine e White (2001) propuseram um método para verificar a significância das variáveis de entrada (*inputs*) da rede através de inferência estatística formal. Abaixo, serão apresentadas duas formas de se testar a significância das entradas da rede: uma baseada em derivadas parciais e outra em comparação de redes. Cabe salientar, porém, que esta metodologia é adequada para dados i.i.d. (com algumas exceções, como veremos). La Roca e Perna (2005) propuseram uma modificação que torna este método aplicável também para dados dependentes, como será visto abaixo. A sugestão proposta no presente trabalho é a modificação do procedimento de *bootstrap* no caso de dados em painel sem dependência espacial conforme apresentada em Kapetanios (2004).

3.5.1 Procedimento de Racine-White para dados i.i.d.: teste baseado em derivadas parciais

Dado o seguinte modelo MLP com uma camada oculta:

$$f(x, w) = w_{00} + \sum_{j=1}^h w_{0j} \psi(\tilde{x}' w_{1j})$$

onde:

ψ = função de ativação da rede (que, em nosso caso, é a função logística);

$$\tilde{x} = (1, x^T)^T$$

Testar a hipótese que certas entradas da rede não possuem efeito qualquer no produto dessa rede, equivale a testar a hipótese que as derivadas parciais em relação a essas entradas são iguais a zero, ou seja:

$$\frac{\partial f(x, w^*)}{\partial x_i} = 0 \quad i \in I_0 \quad (\text{I})$$

onde:

w^* são os pesos "ótimos" encontrados para a rede durante o treinamento;

I_0 é o conjunto de entradas cuja relevância deseja-se testar.

Quando trabalhamos com funções não-lineares, a derivada parcial não é constante, mas dependente dos valores assumidos pelas variáveis explicativas. Assim, uma medida de relevância de um conjunto de entradas deve ser baseada em alguma função das derivadas parciais, como a dada abaixo, por exemplo:

$$m^* = \sum_{i \in I_0} \int f_i(x, w^*)^2 d\mu(x) \quad (\text{II})$$

onde:

$$f_i(x, w^*) = \frac{\partial f(x, w^*)}{\partial x_i}$$

$\mu(x)$ = distribuição de probabilidade de X_t .

Temos que (I) será verdadeira se e somente se $m^* = 0$. Note, porém, que w^* e $\mu(x)$ são desconhecidos, mas o peso \hat{w}_n encontrado durante o treinamento da rede e a distribuição empírica de X são consistentemente estimados por w^* e $\hat{\mu}_n$, respectivamente. Uma estatística factível é então:

$$\hat{m}_n = \frac{1}{n} \sum_{t=1}^n \sum_{i \in I_0} f_i(X_t, \hat{w}_n)^2 = \int \sum_{i \in I_0} f_i(x, \hat{w}_n)^2 d\hat{\mu}_n(x)$$

Racine e White (2001, p. 660) derivaram a distribuição assintótica desta estatística, com a suposição principal que X_t é uma seqüência de vetores aleatórios, mostrando que:

$$n\hat{m}_n \rightarrow N_2(0, C^*; M^*)$$

Esta distribuição é a composição de χ^2 s independentes⁴⁴, onde C^* é a matriz de covariância (semidefinida positiva) e $M^* = E(\nabla^2 m(X_t, w^*)) / 2$. Tal distribuição pode ser aproximada pela técnica de *bootstrap*. A estatística *bootstrap* é dada por:

$$\bar{\beta}_n^* \equiv \sum_{t=1}^n m(X_t, \hat{w}_n^*) - \sum_{t=1}^n m(X_t, \hat{w}_n) - \sum_{t=1}^n \nabla^T m(X_t, \hat{w}_n) (\hat{w}_n^* - \hat{w}_n)$$

⁴⁴ Se $Z \sim N(\mu, V)$, então a forma quadrática $Z'KZ \sim N_2(\mu, V; K)$. A distribuição χ^2 é, portanto, um caso particular da distribuição N_2 (WHITE, 1994, p.170).

onde:

\hat{w}_n^* = pesos "ótimos" da rede treinada com a reamostra de X_t e Y_t .

$$m(x, w) = \sum_{i \in I_0} f_i(x, w)^2$$

Pelo Teorema 2.3 (RACINE; WHITE, 2001, p. 662):

$$\bar{\beta}_n^* \xrightarrow{d} N_2(0, C^*; M^*)$$

O procedimento para o teste de significância das entradas da rede é sumarizado abaixo.

- i) Será calculada uma reamostra com reposição de $\{X_t, Y_t\}$, que chamaremos de $\{X_t^*, Y_t^*\}$. Estimaremos, então, um novo modelo com $\{X_t^*, Y_t^*\}$. O algoritmo de treinamento será iniciado a partir dos valores iniciais \hat{w}_n encontrados para a rede treinada originalmente. Com esses pesos reamostrados e os pesos iniciais, podemos então computar a estatística *bootstrap*:

$$\bar{\beta}_n^* \equiv \sum_{t=1}^n m(X_t, \hat{w}_n^*) - \sum_{t=1}^n m(X_t, \hat{w}_n) - \sum_{t=1}^n \nabla^T m(X_t, \hat{w}_n)(\hat{w}_n^* - \hat{w}_n)$$

- ii) O procedimento descrito em (i) será repetido 500 vezes. Portanto, obteremos 500 valores para a estatística *bootstrap*. Dessa forma, poderemos computar a região de aceitação da hipótese nula (c_α).
- iii) A estatística $n\hat{m}_n$ será calculada. Se for maior que c_α , a hipótese nula deverá ser rejeitada.

3.5.2 Inferência para dados dependentes: o uso do *bootstrap* em bloco

A aplicação do *bootstrap* padrão a dados dependentes não fornece resultados consistentes, já que desconsidera a estrutura de dependência dos dados. Uma possível alternativa é a utilização do método de *subamostragem*, no qual blocos de observações consecutivas são obtidos da série original, tratando cada subsérie como uma série temporal válida. A estatística de teste é avaliada então em todas as subséries.

Em particular, pelo *Teorema 1* e o *Corolário 2* de La Rocca e Perna (2005, p. 420-421), o procedimento de Racine e White pode ser estendido a dados dependentes com o requerimento adicional que o processo seja estacionário e *strong mixing*⁴⁵. O procedimento de *bootstrap*, no entanto, deve ser modificado para levar em consideração a dependência dos dados. A técnica sugerida é a do *bootstrap* em bloco citada acima. Dessa forma, o procedimento de reamostragem deve ser alterado conforme descrito a seguir:

- i) Fixa-se uma subsérie (bloco) em b , fazendo $Y_{b,t} = (Y_t, \dots, Y_{t+b-1})$ e $X_{b,t} = (X_t, \dots, X_{t+b-1})$ subséries de b observações consecutivas da amostra original.
- ii) Seja $\hat{W}_{T,b,t}$ a estimativa do vetor de parâmetros calculados para essas subséries e $\hat{m}_{T,b,t}$ a estatística de teste calculada para cada uma dessas subamostras. Repetindo esse processo para cada subsérie, nós podemos aproximar a distribuição amostral da estatística de teste através de:

$$\hat{G}_{T,b}(x) = \frac{1}{T-b+1} \sum_{i=1}^{T-b+1} I(b.m_{T,b,t} \leq x)$$

em que $I(\cdot)$ é a função indicador.

- iii) o valor crítico para o teste é obtido como o quantil de ordem $(1-\alpha)$ de $\hat{G}_{T,b}(\cdot)$.

O principal problema em aplicar este tipo de procedimento está na escolha da duração do bloco, ou seja, na variável b . Contudo, pelo Teorema 2.7.1, apresentado em Politis et al. (1999), resultados assintóticos são válidos para um amplo intervalo de escolha de b .

⁴⁵ É um tipo de independência assintótica. Assim, embora exista uma dependência entre eventos recentes, quando separados por um grande intervalo de tempo eles são quase independentes. Veja Spanos (1986) e Davidson e MacKinnon (1993).

3.5.3 Inferência em um painel de dados com dependência temporal e ausência de dependência espacial

A literatura em dados em painel tem focado principalmente na dimensão tempo. Kapetanios (2004) considera a questão do procedimento de *bootstrap* no contexto de dados em painel⁴⁶ e sugere uma possibilidade que não tem sido explorada pela literatura no assunto. A sua sugestão é reamostrar com substituição apenas as unidades *cross-section*. Nos casos onde os dados apresentam dependência temporal, mas não espacial, é possível com esse procedimento aplicar abordagens do *bootstrap* para dados *i.i.d.* em vez da aplicação do *bootstrap* em bloco, por exemplo. O autor mostra que o *bootstrap* baseado apenas nas unidades *cross-section* tem performance superior quando comparado ao *bootstrap* em bloco.

Neste trabalho, sugerimos adotar o procedimento de reamostragem exposto acima. É razoável supor, nesse caso, que as expectativas inflacionárias apresentem dependência temporal porém não apresentem correlação no corte transversal, evitando assim o uso do *bootstrap* em bloco que, como salientam Lahiri (2003) e Andrews (2002), tem fornecido aproximações menos acuradas que o *bootstrap* “padrão”.

3.6 Aplicação de RNAs a dados em painel

A aplicação de RNAs é mais freqüente em problemas que envolvam séries temporais, sendo a análise com dados em painel ainda limitada, já que estes são geralmente analisados por meio de técnicas de estimação apropriadas e específicas (LONGHI et. al, 2005).

Alguns poucos trabalhos aplicaram RNAs no contexto de dados em painel. Longhi et. al (2005), por exemplo, a utilizam em um painel não-balanceado para a previsão de padrões de emprego regional na Alemanha e concluem que esta técnica parece oferecer previsões mais acuradas que as tradicionais. Outras aplicações incluem a estimação de fluxos de investimento direto externo (PLIKINAS; AKBAR, 2005), política de endividamento de firmas (PAO; CHIH, 2006), previsão de crises financeiras (PELTONEN, 2006). A maioria desses estudos usa o método da retropropagação dos erros para estimar um modelo do tipo *pooled cross-section*.

⁴⁶ No contexto de dados em painel, as seguintes abordagens de reamostragem são possíveis: reamostragem temporal, *cross-sectional*, temporal e *cross-sectional* simultaneamente.

No caso de dados em painel, torna-se interessante modelar efeitos não-observados relacionados às unidades *cross-section* ou então efeitos de tempo específicos. Isso poderia, por exemplo, ser realizado por meio da inclusão de *dummies* o que, porém, aumentaria de maneira expressiva o número de *inputs*. Existem, na literatura, outras maneiras de tratar o problema da heterogeneidade não-observada (efeitos fixos e efeitos aleatórios, por exemplo). No presente trabalho, seguiremos a estratégia de modelagem adotada em Longhi et. al (2005). Os efeitos específicos das unidades de corte serão capturados por meio de uma variável discreta, definida entre os valores de 0 e 1, assumindo um valor diferente para cada indivíduo, calculada da seguinte forma:

$$\text{Efeitos individuais} = \frac{i}{N}$$

onde:

i varia de 1 a N

N = tamanho da amostra no corte transversal

De maneira similar, os efeitos específicos de tempo serão capturados a partir da seguinte variável:

$$\text{Efeitos específicos de tempo} = \frac{t}{T}$$

onde:

t varia de 1 a T

T = tamanho da amostra no corte temporal

4. Análise Empírica da Formação de Expectativas no Brasil

O objetivo do presente capítulo é estudar empiricamente a formação de expectativas inflacionárias no Brasil no período recente (pós metas de inflação), usando para isso um modelo conexionista que aproxima a forma como os agentes fazem suas previsões⁴⁷. A formação de expectativas de inflação vem recebendo uma maior atenção por parte dos pesquisadores devido à implementação do sistema de metas de inflação no país, em junho de 1999. Os estudos empíricos que tratam dessa questão foram revisados no capítulo 2 e ainda sofrem de limitações devido ao número reduzido de informações disponíveis.

O principal diferencial deste trabalho em relação aos anteriores está na metodologia empregada: a modelagem através de RNAs. Além disso, os dados utilizados também são distintos, já que aqui estamos empregando como variável dependente as expectativas para o IPCA acumulado em 12 meses de cada instituição em cada ponto no tempo, constituindo um painel de dados não balanceado. Os estudos anteriores usavam alguma medida agregada, tal como a média ou mediana do mês ou ainda os dados diários (como foi o caso de Yoshihiro(2005)). A principal motivação para o uso do painel de dados foi a reduzida amostra que se teria utilizando os dados agregados. Além disso, como bem salientado por Hsiao (2005), a utilização de dados em painel permite uma maior capacidade para capturar a complexidade do comportamento humano, quando comparado a séries de tempo ou *cross-section* simples.

4.1 Descrição dos dados

4.1.1. Expectativas de inflação

Os dados para as previsões de inflação são referentes às expectativas mensais para o IPCA acumulado 12 meses à frente, no período de fevereiro/2002 a junho/2006.

⁴⁷ Cabe salientar, porém, que o mais adequado seria modelar a formação de expectativas de forma individual, visto que também não é apropriado assumir que os agentes interpretem o mundo e formem suas expectativas de forma homogênea. Os agentes dentro dos modelos econômicos são heterogêneos e a suposição do agente representativo vai contra isso (BELTRATTI; TERNA; MARGARITA, 1996). Além de haver heterogeneidade, há também interação entre os agentes, que poderia ser modelada usando-se algoritmos genéticos. Porém, o interesse desse trabalho é de natureza empírica e não teórica. Desejamos apenas entender como as expectativas são formadas de forma “agregada”.

Para cada mês, temos a informação da expectativa de cada instituição pesquisada pelo BACEN, referente ao período posterior à data de referência⁴⁸ (cinco dias). Os dados foram gentilmente cedidos pelo GERIN⁴⁹. Foram excluídas da amostra as instituições que contassem com menos de 10 observações durante o período inteiro. Com esse procedimento, a amostra passou a 80 instituições.

O uso de dados de *survey* como medida de expectativas de inflação é algumas vezes questionado. De acordo com Garcia (1994, p. 48, 67), a análise realizada com dados de inflação esperada obtidos através do preço de ativos financeiros são superiores às que se utilizam de *surveys* de expectativas de inflação. Isso porque, nesse último caso, os agentes podem não ter incentivos para revelar suas verdadeiras expectativas. Além disso, Lima e Céspedes (2003) realizam um estudo de habilidade preditiva das expectativas e concluem que estas “não espelham as verdadeiras expectativas do mercado e que, portanto, não deveriam ter qualquer influência na formulação da política monetária do país” (LIMA; CÉSPEDES, 2003, p. 83). Carvalho (2004) não compartilha de tal opinião, mostrando que o Bacen possui mecanismos de incentivos para que as instituições revelem suas expectativas verdadeiras. Além disso, como salientado em Carvalho e Bugarin (2005, p. 2, nota de rodapé n. 3), o procedimento para extrair as expectativas de instrumentos financeiros no caso brasileiro é seriamente comprometido por questões relacionadas à liquidez e ao prêmio de risco (que não é constante no tempo e muito menos negligenciável).

⁴⁸ A data crítica para o IPCA consiste na data de divulgação do IPCA-15. A data de referência refere-se ao dia anterior à data crítica, e é a data-base para a verificação das instituições TOP-5.

⁴⁹ Gerência Executiva de Relacionamento com Investidores do Banco Central do Brasil.

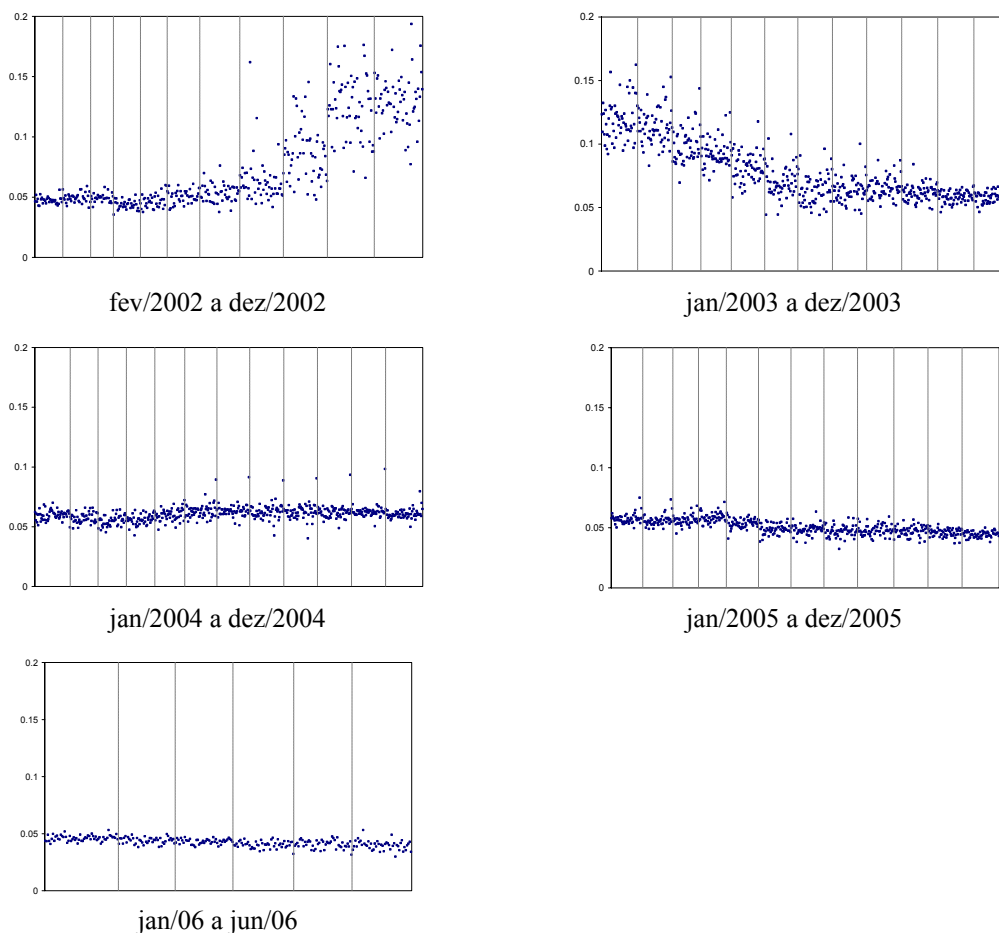


Gráfico 4.1. Expectativas de inflação (IPCA) para os próximos 12 meses
 Fonte: Dados do Banco Central do Brasil

4.1.2 O comportamento das expectativas de inflação entre fev-2002 e jun-2006

O regime de metas de inflação foi adotado no país em junho de 1999, após a implementação da flutuação cambial, em janeiro do mesmo ano, em meio a uma crise relacionada à remoção da âncora cambial, o que exigiu que o Banco Central perseguisse a credibilidade de suas ações com maior vigor. O instrumento operacional utilizado para alcançar as metas de inflação é a taxa de juros de curto prazo (Selic). Assim, quando a política monetária perde credibilidade, o Banco Central deve aumentar a taxa de juros a fim de controlar a inflação.

O Gráfico 4.1 mostra a evolução das expectativas inflacionárias entre fev-2002 e jun-2006. Como pode ser observado, a partir de ago-2002, a magnitude das expectativas de inflação, bem como a sua dispersão, iniciam um processo de aumento considerável. Essa situação começa a se reverter em meados de 2003.

Esse comportamento pode ser explicado basicamente por dois fatores: a elevada aversão ao risco no cenário internacional⁵⁰ e, principalmente, a crise de confiança que ocorreu no país em 2002⁵¹. Nesse período, tivemos um ambiente de credibilidade limitada, já que o excesso de conservadorismo da política monetária foi foco constante de críticas durante a campanha eleitoral (BEVILAQUA; MESQUITA; MINELLA, 2007). O risco-Brasil aumentou de 750 pontos em abril de 2002 para 2.400 pontos no final de setembro. A demanda por títulos domésticos diminuiu consideravelmente. A inflação anual, medida pelo IPCA, alcançou 12,53% em 2002 (em 2001, havia sido de 7,67%), ficando acima do intervalo de tolerância para a meta (3,5% com intervalo de $\pm 2\%$). Houve uma súbita interrupção nos influxos de capital que levou a uma depreciação forte do câmbio. De fato, a depreciação acentuada da taxa de câmbio fez com que as expectativas de inflação aumentassem.

O Gráfico 4.2 apresenta a evolução do erro de previsão da inflação para os próximos 12 meses. Entre fevereiro e agosto de 2002, os agentes subestimaram a inflação. A partir de agosto de 2002, os agentes passam a ter expectativas de inflação mais altas, tendendo a superestimar a inflação. Nesse contexto, o Banco Central, que vinha até então reduzindo a taxa de juros, decidiu por um aumento nesse instrumento, passando de 18%a.a. (julho-2002) para 21%a.a. (outubro-2002). A taxa Selic sofreu então aumentos sucessivos, atingindo o pico de 26,5% em fevereiro-2003.

Com a instalação do novo governo em 2003 e a efetiva continuidade da política macroeconômica do governo anterior, a incerteza diminuiu e as expectativas começaram a se reverter, ou seja, tanto sua magnitude quanto dispersão diminuíram⁵².

⁵⁰ Provocada pelo baixo crescimento mundial, surgimento de problemas contábeis em grandes empresas americanas, crises em mercados emergentes e a perspectiva de mais uma guerra do Golfo.

⁵¹ A crise de confiança foi ocasionada pela incerteza em relação às eleições presidenciais, que apontavam o candidato Luis Inácio Lula da Silva como o favorito, o qual, em seu passado, havia defendido o não pagamento da dívida.

⁵² Para uma avaliação do período de metas de inflação no Brasil consulte Garcia (2006) e Bevilaqua, Mesquita e Minella (2007).

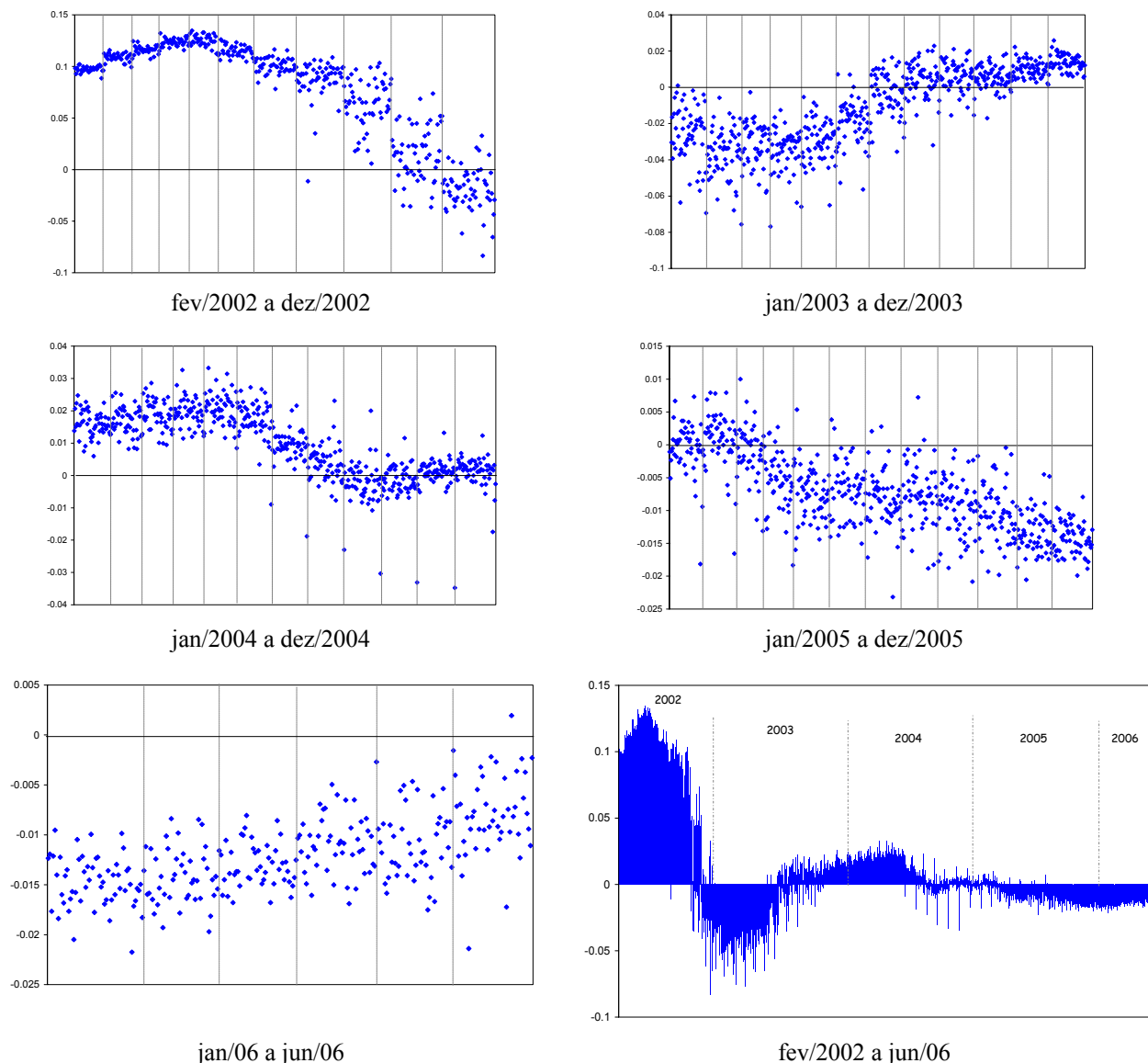


Gráfico 4.2. Erro de Previsão - IPCA para os próximos 12 meses
 Fonte: Dados do Banco Central do Brasil e IBGE

A tendência dos agentes em superestimar a inflação acumulada se mantém até meados de 2003, quando o erro de previsão passa a ser predominantemente positivo, ou seja, os agentes novamente passam a subestimar a inflação acumulada em 12 meses, tendência que se mantém até meados de 2004. A partir de então, os agentes passam novamente a superestimar a inflação. Em meados de 2004 houve um aumento da inflação: o IPCA acumulado em 12 meses passou de 5,15% em maio-2004 para 7,18% em agosto do mesmo ano, o que pode explicar essa revisão para cima das expectativas. De acordo com Garcia (2006), os preços administrados foram os principais responsáveis por esse crescimento do índice de preços, já que os choques de oferta que afetaram a economia em 2002 e 2003 foram transmitidos a estes pela regra de indexação. Os

preços livres também se elevaram, o que indica que o aumento dos preços de produtos administrados foi repassado aos preços livres. Dessa forma, em setembro de 2004, a taxa Selic foi novamente elevada para conter o aumento da inflação, atingindo o ápice em maio de 2005 (19,75%) e voltando a cair apenas em setembro de 2005 (19,5%).

É claro a partir dos Gráficos 4.1 e 4.2, o período de consolidação da desinflação a partir de 2005. O erro de previsão vem diminuindo consideravelmente, bem como a incerteza sobre a inflação futura (medida pelo desvio-padrão das expectativas). Esse comportamento parece refletir um ambiente mais estável e previsível, com um maior entendimento por parte do mercado da dinâmica da inflação.

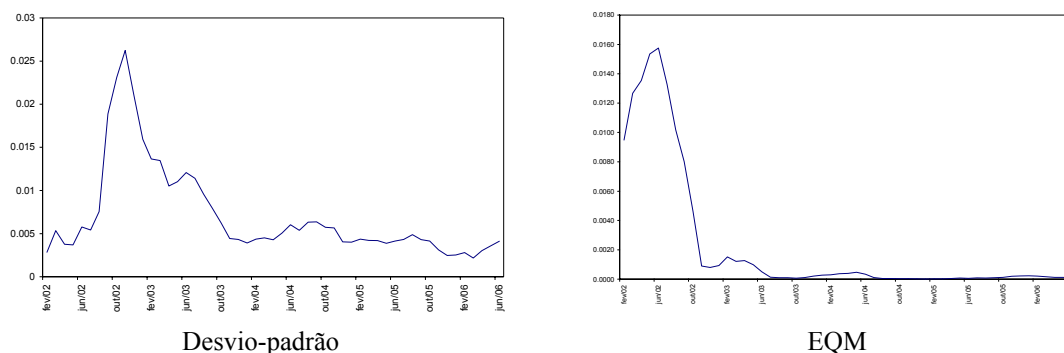


Gráfico 4.3. Desvio-Padrão dos erros de previsão e EQM – fev/02 a jun/06

4.1.2 Variáveis explicativas

Abaixo apresentamos o conjunto de variáveis que podem, a princípio, ser explicativas do processo de formação de expectativas no Brasil. As variáveis foram escolhidas com base teórica, em estudos anteriores e de acordo com as observações realizadas na seção anterior.

- Defasagens de primeira e segunda ordem das expectativas de inflação.

Fonte: Banco Central do Brasil

- Surpresa de curto prazo do IPCA-15⁵³: $IPCA - 15_t - E_{t-1}(IPCA - 15)$.

Fonte: Banco Central do Brasil e IBGE

⁵³ Com a inclusão dessa variável, a amostra diminui expressivamente.

- Surpresa de curto prazo da inflação: $IPCA_t - E_{t-1}(IPCA_t)$.

Fonte: Banco Central do Brasil e IBGE

- IPCA acumulado em 12 meses com um período de defasagem ($IPCA_{t-1}^{12M}$).

Fonte: Banco Central do Brasil e IBGE

- Variação do Risco – Brasil (EMBI Brasil) com um período de defasagem.

Fonte: Macrodados

- Hiato do produto com três períodos de defasagem⁵⁴: desvio entre o produto real e o potencial, calculado através do filtro de *Hodrick-Prescott* e com a utilização da série da produção industrial dessazonalizada.

Fonte: Dados da produção industrial – IBGE

- Taxa de juros selic efetiva (% a.a.), com um período de defasagem.

Fonte: Banco Central do Brasil

- Meta de inflação: a meta para a inflação acumulada em 12 meses foi obtida por interpolação linear⁵⁵ das metas anuais:

$$\pi_{12m,t+12}^T = \frac{12-i}{12} \pi_t^T + \frac{i}{12} \pi_{t+1}^T$$

Fonte: Banco Central do Brasil

- “Efeito fixo”: variável discreta no intervalo (0;1), que assume um valor diferente para cada instituição. Dado que temos 80 instituições na amostra, essa variável será dada por:

$$\frac{1}{80} \times j; \text{ onde } j \text{ varia de } 1 \text{ a } 80 \text{ e refere-se ao número da instituição.}$$

⁵⁴ Como bem salientado por Bevilaqua, Mesquita e Minella (2007), dois problemas com o uso dessa variável no presente caso é que estamos utilizando dados revisados e o filtro HP foi aplicado à série da produção industrial completa.

⁵⁵ As metas de inflação foram ajustadas nos períodos em que o Bacen alterou formalmente a meta ser perseguida, devido à ocorrência de choques adversos.

- “Efeito de tempo específico”: construída da mesma forma que a variável de efeito fixo:

$$\frac{1}{53} \times t; \text{ onde } t \text{ varia de } 0 \text{ a } 53 \text{ e refere-se ao ponto no tempo.}$$

- Variação Cambial com um período de defasagem.

Fonte: Banco Central do Brasil

- Volatilidade cambial: Desvio-padrão do câmbio no mês anterior; desvio padrão do câmbio acumulado em 3 meses; desvio-padrão do câmbio acumulado em 6 meses.

- Resultado primário do Governo Central (% PIB).

Fonte: STN – Secretaria do Tesouro Nacional

- Dívida total (% PIB).

Fonte: Banco Central do Brasil

- Variação dos preços das commodities com um período de defasagem.

Fonte: IPEA

- Desvio-padrão médio das expectativas com um período de defasagem.

Fonte: Banco Central do Brasil

- Média das previsões médias das instituições Top 5 de longo prazo com um período de defasagem.

Fonte: Banco Central do Brasil

- *Dummy* : foram utilizadas algumas variáveis *dummies* para o período de crise de confiança, com diversas datas de início e fim entre março de 2002 e fevereiro de 2003 e março de 2003 a agosto de 2003 (pós-crise).

4.2 Arquitetura da rede

A definição da arquitetura da rede é uma tarefa bastante complicada, não existindo ainda uma metodologia estabelecida para isso. Como salientam Zhang, Hu e Patuwo (1998, p. 42), a definição desses parâmetros é fortemente dependente do problema para o qual se está utilizando as RNAs. O processo de determinação da arquitetura de uma RNA é chamado por muitos de “magia negra”, dada a sua pouca compreensão. Alguns autores como Anders e Korn (1999), Ripley (1995), Pizarro, Guerrero e Galindo (2000), Medeiros e Teräsvirta (2001), entre outros, propõem procedimentos estatísticos para esta tarefa. La Rocca e Perna propuseram um procedimento bastante simples, no qual a escolha das variáveis explicativas é feita através de uma medida de relevância e posterior teste de significância destas variáveis. Abaixo, detalhamos como escolhemos a arquitetura da rede utilizada no presente trabalho.

- seleção de variáveis de entrada (“inputs”)

A seleção será feita basicamente de forma empírica. Usaremos como guia o procedimento sugerido por La Rocca e Perna (2005), ou seja, um teste estatístico formal para remover aquelas entradas que forem irrelevantes, com algumas adaptações ao problema específico que se tem aqui. A rede geral, com todas as variáveis será treinada e com isso, poderemos computar a medida de relevância de cada entrada⁵⁶. As entradas que possuírem tal medida próxima de zero serão tidas como candidatas à exclusão do modelo. Para tanto, a significância estatística de cada uma será testada. Se não for significativa, será treinada uma nova rede sem essa variável. As redes serão comparadas através do CIS.

- número de camadas ocultas (*hidden layers*)

Existem trabalhos teóricos que mostram que uma camada oculta é suficiente para aproximar qualquer função não-linear complexa (veja, por exemplo, Hornik, Stinchcombe e White, 1989). A maioria dos trabalhos que utilizam RNAs para previsão

⁵⁶ A medida de relevância sugerida por La Rocca e Perna (2005) é a estatística m dada por: $m = \sum_{i=1}^n f_i^2(x, w)$

de séries temporais faz uso de apenas uma camada oculta⁵⁷. De acordo com Zhang, Hu e Patuwo (1998, p. 44), uma camada oculta parece ser adequada para a maioria dos problemas de previsão. Sendo assim, utilizaremos redes com apenas uma camada intermediária.

- número de neurônios na camada oculta (hidden nodes)

Geralmente, o número de neurônios na camada oculta é definido de forma empírica⁵⁸. Se utilizarmos neurônios demais, a rede poderá memorizar excessivamente os dados de treinamento, levando a *overfitting* (excesso de ajustamento) e, conseqüentemente, a uma baixa capacidade de generalização. Se utilizarmos poucos neurônios, o treinamento poderá ser extremamente lento.

Segundo Racine e White (2001), a complexidade da rede será determinada com o auxílio do Critério de Informação de Schwarz (CIS). Para tanto, utilizaremos de 1 a $2i+1$ unidades ocultas (i = número de *inputs*) e dez reinícios aleatórios de pesos iniciais para cada conjunto de unidades de entrada⁵⁹. A configuração que produzir o menor valor para o CIS, para cada rede, será então escolhida como a melhor, dado o número de variáveis de entrada.

- número de neurônios de saída (output)

Como estamos interessados em prever expectativas inflacionárias um passo à frente, devemos utilizar apenas um neurônio de saída.

- função de ativação

Será utilizada a função sigmóide (logística)⁶⁰ para os neurônios da camada intermediária e a função linear “pura” (ou identidade) para o neurônio de saída.

⁵⁷ Existem, porém, trabalhos mostrando os benefícios de se utilizar duas camadas ocultas. Veja, por exemplo, Barron (1994).

⁵⁸ Existem também algumas “regras de bolso” para a determinação desse parâmetro. Porém, como salientado por Zhang, Hu e Patuwo (1998, p. 44), nenhuma delas trabalha bem para todos os problemas.

⁵⁹ O reinício aleatório dos pesos é recomendável, já que o treinamento de uma RNA é bastante sensível aos pesos iniciais. Têm sido propostos, também, alguns métodos para a inicialização dos pesos como, por exemplo, inicialização baseada nos resultados de uma regressão múltipla (CHAN, LAM e WONG, 2000), entre outros.

⁶⁰ Como salientam Zhang, Hu e Patuwo (1998, p. 47) e Swingler (1996, p. 62), a utilização da função logística nos neurônios ocultos é a mais popular escolha para aplicações de RNAs.

4.3 Treinamento e avaliação das redes

Para o treinamento das redes, utilizamos o algoritmo do gradiente conjugado. Para evitar a ocorrência de um mínimo local, o algoritmo é imediatamente reiniciado a partir do ponto em que o treinamento foi encerrado. Além disso, reinícios aleatórios dos pesos iniciais são também utilizados. As redes foram treinadas com o auxílio do programa *Matlab 6.5*.

Uma questão bastante relevante quando se inicia o treinamento é decidir se as variáveis serão ou não “normalizadas”⁶¹. Se sim, resta ainda decidir qual tipo de “normalização” será utilizada. Não há ainda um consenso sobre a necessidade ou não de se realizar a normalização das variáveis. Shanker, Hu e Hung (1996), por exemplo, avaliam os benefícios no treinamento da rede de se realizar a normalização linear e estatística nas variáveis⁶². Eles concluem que, em geral, há sim benefícios em se realizar a normalização, mas estes diminuem quando o tamanho da rede e da amostra aumentam. Palma (2004) por exemplo, encontrou melhores resultados com a normalização estatística, a qual será empregada no presente trabalho.

Outra questão que aqui deve ser ressaltada é em relação à estacionariedade das séries. Como a duração temporal que temos aqui é pequena, a robustez dos testes de estacionariedade ficaria comprometida. Por isso, eles foram omitidos da análise.

4.4 Resultados

Como a amostra diminui consideravelmente com a inclusão da surpresa de curto prazo do IPCA-15, fizemos algumas tentativas de treinamento para testar a significância desta variável. Em nenhuma das tentativas esta se mostrou estatisticamente significativa a 10% (o valor-p girava em torno de 40%). A inclusão dessa variável era a razão de usarmos esse modelo com redução significativa da amostra. Dessa forma, a análise foi realizada com a amostra completa, sem a utilização da surpresa de curto prazo do IPCA-15. Além disso, fizemos um experimento para verificar a correlação entre a variação

⁶¹ Aqui cabe uma ressalva: a palavra normalização, na terminologia de RNAs, é utilizada como sinônimo de transformação dos dados originais, de uma maneira que facilite a detecção de características dos dados de entrada.

⁶² Normalização linear: $\frac{y_t - y_{\text{mínimo}}}{y_{\text{máximo}} - y_{\text{mínimo}}}$; normalização estatística: $\frac{y_t - \bar{y}}{dp(y_t)}$.

das expectativas entre cinco dias antes e cinco dias depois da data de referência com a variação do IPCA-15. O coeficiente de correlação encontrado foi de 0.18, ou seja, bastante baixo.

Várias tentativas de treinamento foram feitas com o modelo mais geral (o que inclui todas as variáveis), seguindo o método exposto na seção anterior⁶³. Cabe ressaltar que, como o efeito do hiato do produto se mostrou insignificante em todas as tentativas de estimação, tentamos substituir essa variável pela taxa de crescimento do produto (também defasada), mas esta se mostrou também insignificante. O modelo final é apresentado a seguir.

Tabela 4.1 Função de Reação das Expectativas de Inflação

Variáveis	
Lag1, lag2, taxa selic, meta, efeito fixo, resultado primário, volatilidade câmbio, câmbio, <i>commodities</i> , erro curto prazo IPCA	
Conjunto de Treinamento	
Number of neurons:	5
Number of weights:	61
Goodness of fit:	93.0972 %
RMSE:	0.0063473
MAE:	0.00372672
MAPE:	0.0547819
SIGN:	1
CORR:	0.964869
CIS:	-6.2495596426681095
Conjunto de Validação	
Goodness of fit:	24.5576 %
RMSE:	0.0034069
MAE:	0.00272535
MAPE:	0.0675678
SIGN:	1
CORR:	0.523611

⁶³ Os resultados dessas tentativas podem ser obtidos através de solicitação aos autores. Algumas das tentativas se encontram no apêndice.

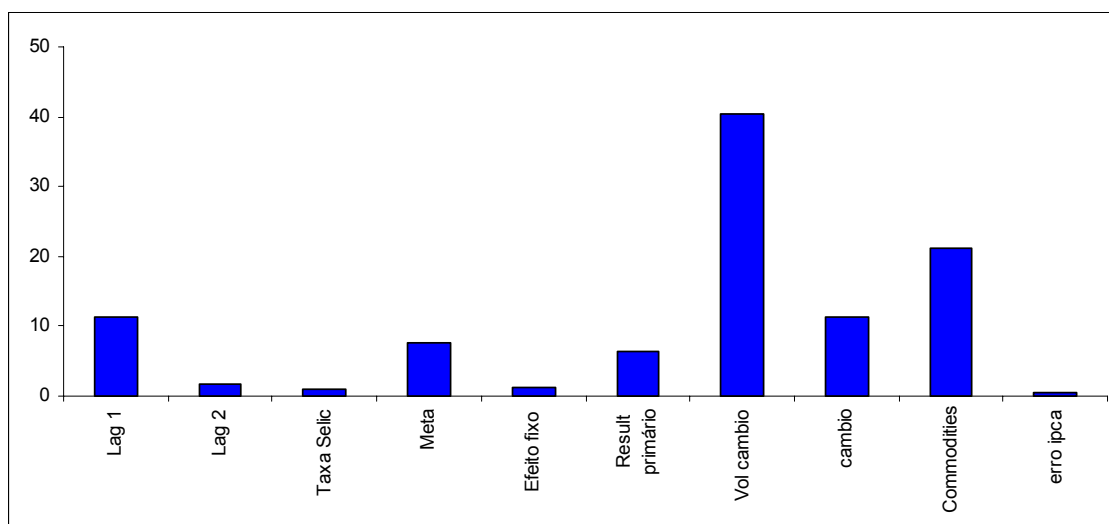


Gráfico 4.4. Medida de Relevância

Como pode ser observado pelo Gráfico 4.4, a maior influência sobre as expectativas inflacionárias no período como um todo foi da volatilidade cambial, seguida pela variação no preço das *commodities*, pela variação cambial e pela meta. A tabela abaixo mostra que todas as variáveis, com exceção da surpresa de curto prazo do IPCA, são estatisticamente significantes a 5%. Como a surpresa de curto prazo se mostrou insignificante, treinamos uma nova rede sem tal variável, mas o CIS diminuiu e, dessa forma, continuamos com o modelo apresentado acima.

Tabela 4.2. Inferência sobre a relevância das variáveis⁶⁴

Variáveis	Estatística de teste	Valor crítico (5%)	Valor-p
Lag1	11.39417306	4.97181385	0.00750000
Lag2	1.60194240	0.72832546	0.01500000
Taxa Selic	0.97217262	0.55913907	0.03000000
Meta12m	7.53851883	3.17884365	0.02750000
Efeito fixo	1.26685331	0.53758574	0.01000000
Resultado primário	6.48979855	2.72684811	0.00500000
Volatilidade câmbio	40.40148104	15.92434283	0.01000000
Câmbio	11.31297167	4.09519099	0.00500000
Commodities	21.28378370	10.17932675	0.01500000
Surpresa curto prazo IPCA	0.59742535	14.31063722	0.60750000
Teste conjunto (todas as variáveis)	102.85912054	37.49401074	0.00750000

Para avaliar o efeito de cada uma das variáveis acima sobre as expectativas, procedemos à análise de sensibilidade do produto da rede. Para facilitar esse estudo,

⁶⁴ O procedimento de *bootstrap* foi realizado com 500 repetições.

fizemos a distinção entre três períodos específicos: anterior à crise de 2002 (abr-02 a ago-02), período de crise (set-02 a ago-03) e posterior à crise (set-03 a abr-06). A sensibilidade do produto da rede a variações de 1% em cada uma das entradas é apresentada na tabela a seguir⁶⁵.

Tabela 4.3. Sensibilidade do produto da rede à variações de 1% nos inputs

Variáveis		Abr/02 – ago/02	Set/02- ago/03	Set/03-abr/06
Lag1	Mínima	0.540017	-26,9524	0.566717
	Média	0.704789	0,824882	0.757721
	Máxima	0.888142	39,14938	0.926067
	Mediana	0.683552	0,710823	0.758913
	Desvio-padrão	0.081364	3,485624	0.050709
Lag2	Mínima	0.016649	-7.619765	-0.087761
	Média	0.078753	0.122593	0.032956
	Máxima	0.149144	11.439488	0.148087
	Mediana	0.081183	0.073168	0.045896
	Desvio-padrão	0.032610	1.261339	0.051368
Taxa Selic	Mínima	-0.610709	-18.988264	-0.969732
	Média	0.109173	0.616301	-0.296184
	Máxima	1.067402	35.155207	0.985158
	Mediana	-0.023903	0.280336	-0.338012
	Desvio-padrão	0.465596	3.295717	0.326615
Meta12m	Mínima	-0.470666	-1.85975	-0.726204
	Média	0.089931	-0.33134	0.161364
	Máxima	0.555521	0.496298	0.772871
	Mediana	0.249244	-0.26701	0.361683
	Desvio-padrão	0.322930	0.448563	0.436041
Resultado primário	Mínima	-0.013639	-10.845458	-0.260583
	Média	0.052125	0.133719	0.019614
	Máxima	0.296288	12.076498	0.174237
	Mediana	0.025418	0.012132	0.028840
	Desvio-padrão	0.065885	1.098675	0.079422
Volatilidade câmbio	Mínima	0.001811	-113.196815	-0.016540
	Média	0.122444	1.425033	0.087537
	Máxima	0.291662	585.620676	0.220753
	Mediana	0.113399	0.018602	0.095008
	Desvio-padrão	0.091683	28.310890	0.044285
Câmbio	Mínima	-0.013358	-21.693820	-0.256298
	Média	0.140254	0.119045	0.009140
	Máxima	0.426242	40.219463	0.193218
	Mediana	0.049999	-0.002685	0.006240
	Desvio-padrão	0.151930	3.766278	0.047228

⁶⁵ No apêndice são apresentados os respectivos gráficos.

<i>Commodities</i>	Mínima	-0.004089	-22.519365	-0.206895
	Média	0.017308	0.007203	-0.006856
	Máxima	0.074578	25.927903	0.082748
	Mediana	0.005188	-0.008099	-0.002028
	Desvio-padrão	0.020183	2.025257	0.033323
Surpresa curto prazo IPCA	Mínima	-0.040390	-0.269077	-0.049492
	Média	0.009917	-0.003510	0.000023
	Máxima	0.100874	0.213415	0.086320
	Mediana	0.001830	-0.003683	0.000000
	Desvio-padrão	0.021827	0.035609	0.016029

É interessante notar o comportamento da influência da meta de inflação sobre as expectativas⁶⁶. Durante o período de crise de confiança, o efeito da meta sobre a inflação esperada é expressivamente menor que nos demais e mesmo declinante. O contrário ocorre nos outros períodos, atentando para o fato que as expectativas vêm crescentemente sendo influenciadas pela meta no período recente, o que é um indício de que a credibilidade da autoridade monetária vem aumentando⁶⁷. Esse resultado é consistente com o obtido, por exemplo, em Bevilaqua, Mesquita e Minella (2007). Carvalho e Bugarin (2005) apresentam também evidências que durante o final de 2002 e início de 2003, as previsões de inflação foram feitas ignorando totalmente a meta de inflação. O coeficiente de variação (médio) aqui obtido é ainda menor que aqueles encontrados pelos trabalhos anteriores, o que sugere que a credibilidade é imperfeita, havendo espaço para melhoria. Para o período de abril-2006, por exemplo, o efeito médio de uma variação de 1% na meta é de 0.65%. Se a credibilidade do Banco Central fosse perfeita, esse coeficiente deveria ser igual a 1 (ou bem próximo de 1).

A taxa Selic tem também um comportamento interessante. No período da crise de confiança, as expectativas de inflação dos agentes reagem predominantemente de maneira positiva a variações da taxa de juros. No período de início da implementação do regime de metas de inflação (pré-crise de confiança), há episódios de reação negativa (abril-maio de 2002 e ago-2002) e episódios de reação positiva (junho-julho/2002), com a reação positiva tendo um efeito maior sobre as expectativas. No período pós-crise, esse efeito é predominantemente negativo e maior em magnitude. Os resultados obtidos por Carvalho e Bugarin (2005), apontavam para um efeito positivo da taxa Selic sobre as expectativas, que foi justificado pelos autores como sendo proveniente de uma

⁶⁶ Esse comportamento é melhor visualizado no Gráfico C.4 (Apêndice C).

⁶⁷ Seria interessante obter os intervalos de confiança para o produto da rede para verificar se em alguns períodos as variáveis deixam de ser significativas. Porém, isso foge ao escopo deste trabalho.

possível relação não-linear entre estas variáveis. Minella et. al (2002, 2003) também encontram um efeito positivo da taxa de juros defasada sobre as expectativas de inflação e explicam que isso pode ser resultado da reação das taxas de juros a pressões inflacionárias: se é prevista uma inflação mais alta pela autoridade monetária e pelo mercado, a taxa de juros é aumentada, mas, apesar disso, os efeitos inflacionários não são completamente eliminados, já que existem defasagens nos mecanismos de transmissão da política monetária e o Banco Central leva em consideração a volatilidade do produto em suas decisões; além disso, os movimentos na taxa de juros refletem também as expectativas de inflação do Banco Central, já que este reage às suas próprias expectativas inflacionárias também.

Nosso argumento para o efeito ambíguo da taxa de juros sobre as expectativas, contudo, é de que este efeito pode estar associado à influência também ambígua da taxa de juros sobre o câmbio⁶⁸, a qual é derivada do efeito da taxa de juros sobre a dívida. A paridade descoberta da taxa de juros implica que um aumento nos juros torna os títulos domésticos mais atrativos, o que aumenta os influxos de capital, aumentando a oferta de dólares no mercado de câmbio e, assim, apreciando a taxa de câmbio (dólar fica mais barato), o que leva à diminuição das expectativas de inflação. Porém, o aumento da taxa de juros pode ser interpretado como uma elevação do risco de *default*, o que levaria a uma depreciação do câmbio e conseqüente aumento das expectativas. Com isso, um aumento da taxa de juros pode levar a expectativas de inflação mais altas e não mais baixas. Pelos resultados obtidos aqui, e de acordo com a argumentação anterior, teriam existido períodos (junho-02 a julho-02 e maior parte do período de crise) em que o efeito de dominância fiscal cobriu o efeito usual da taxa de juros sobre o câmbio (dado pela condição de paridade-juros descoberta). No entanto, o contrário ocorre no período mais recente, ou seja, a transmissão da taxa de juros sobre as expectativas inflacionárias se dá pelos canais usuais.

O efeito da variação cambial sobre as expectativas é predominantemente positivo (ou seja, a depreciação do câmbio leva a uma elevação da inflação esperada) e também mais baixo no período mais recente (pós-crise de 2002), indicando uma

⁶⁸ Veja Gonçalves e Guimarães (2006). Os autores apresentam evidências empíricas de que o efeito usual que liga aumentos da taxa de juros a apreciação cambial foi superado pelo efeito da dominância fiscal (o qual liga aumentos da taxa de juros a depreciação cambial) no período de 2000 a 2005.

redução no *pass-through*⁶⁹ na economia, como também evidenciado em Bevilaqua, Mesquita e Minella (2007) e Schmidt-Hebbel e Werner (2002). Esse comportamento é consistente com o papel crescente exercido pela meta de inflação. É interessante notar que os resultados obtidos por Minella et. al (2002, 2003) indicavam a não significância da variação do câmbio sobre as expectativas de inflação dos agentes⁷⁰. A influência da variação do câmbio sobre as expectativas de inflação apresentada aqui é, em média, também menor que a encontrada em trabalhos anteriores.

A volatilidade cambial exerce um papel significativo e positivo na formação de expectativas. Assim como a variação cambial, a volatilidade do câmbio vem exercendo um papel declinante sobre as expectativas, porém, ainda bastante expressivo. De nosso conhecimento, nenhum outro estudo tinha levado em consideração o efeito da volatilidade cambial sobre as expectativas inflacionárias. Um dos principais desafios do regime de metas de inflação em economias emergentes, como o Brasil, é trabalhar com a elevada volatilidade da taxa de câmbio, já que essas economias são mais sensíveis aos efeitos de crises financeiras (MINELLA et al. 2002 e PASTORE et al., 2004). Uma volatilidade maior significa maior incerteza, o que deve impactar as expectativas dos agentes. A volatilidade do câmbio gera freqüentes revisões das expectativas inflacionárias podendo resultar em um não cumprimento da meta. Dado o resultado aqui obtido de que as expectativas reagem fortemente à volatilidade cambial, é interessante manter a estabilidade da moeda nacional. Albuquerque (2005) encontrou uma relação positiva entre a volatilidade cambial e a volatilidade inflacionária, ou seja, a variabilidade do câmbio aumenta a incerteza; essa maior incerteza afeta as expectativas de inflação dos agentes acerca da inflação futura. Porém, seus resultados indicavam para uma resposta menor da inflação frente à volatilidade cambial em períodos de crise, onde a volatilidade fosse muito alta. Os resultados aqui obtidos sugerem uma influência maior da volatilidade cambial sobre as expectativas de inflação no período de crise cambial.

A variação do preço das *commodities* é significativa e seu efeito apresenta comportamento predominantemente positivo no período de crise e pré-crise e predominantemente negativo no período pós-crise. Um aumento no preço das *commodities* pode ser interpretado como um choque adverso de custos, levando a um

⁶⁹ Repasse das variações cambiais à inflação. As desvalorizações cambiais afetam diretamente a inflação através dos produtos domésticos que utilizam insumos importados e através dos insumos importados que entram no cálculo dos índices de preços e indiretamente por seus efeitos sobre a demanda agregada.

⁷⁰ Minella et. al (2002,2003) usou a variação cambial acumulada em 12 meses.

aumento nas expectativas de inflação. Contudo, uma variação positiva nessa variável afeta os termos de troca (relação entre os preços das exportações e os preços das importações) de maneira positiva em países exportadores de *commodities*, levando a uma apreciação da taxa futura de câmbio. Assim, uma variação positiva no preço das *commodities* levaria a uma valorização da taxa cambial, diminuindo as expectativas inflacionárias o que, de fato, vem ocorrendo no período recente.

O efeito do resultado primário é predominantemente positivo, em quaisquer dos períodos considerados. Poderíamos esperar que esse efeito fosse negativo, ou seja, um aumento no resultado primário reduziria as expectativas de inflação, já que poderia ser entendido como a existência de uma maior coordenação entre as políticas monetária e fiscal (a política fiscal ajuda a política monetária no combate à inflação). Além disso, o aumento do superávit primário pode ser visto também como uma redução no risco de *default* da dívida, levando à diminuição do prêmio de risco, o que afetaria de maneira inversa as expectativas indiretamente via o efeito da taxa de câmbio⁷¹. Porém, há uma certa preocupação ainda em relação à evolução das contas públicas, com evidências que o resultado primário positivo no Brasil vem sendo conseguido à custa de impostos elevados e corte nos investimentos. No período como um todo, a política monetária esteve sobrecarregada com um quadro fiscal problemático persistente. Como bem salientado por Werneck (2006), houve uma melhora importante na situação fiscal, mas esta foi bem mais favorável pelo lado das contas externas. Ainda de acordo com Werneck (2006), o ajuste fiscal à brasileira é caracterizado pelo aumento do superávit primário consolidado (incluindo estados e municípios) e expansão simultânea de gastos primários. Uma das causas para esse efeito positivo do resultado primário sobre as expectativas inflacionárias pode ser a percepção de que o governo não conseguirá continuar gerando superávits primários pelo modelo corrente de ajuste fiscal. Dessa forma, o corte dos gastos públicos primários seria relevante para extinguir eventuais dívidas quanto à sustentabilidade da dívida pública.

Porém, a análise das expectativas fiscais (resultado primário e razão da dívida sobre o PIB) realizada por Pires (2006), mostra que quando o horizonte de tempo é mais longo, a expectativa de superávit é menor e a expectativa da dívida é também menor, o que poderia contradizer o argumento acima⁷². Esta análise, contudo, carece de maior

⁷¹ Ou seja, a redução no risco de *default* apreciaria o câmbio levando a uma diminuição das expectativas.

⁷² Já que significaria que os agentes esperam esforços menores do ponto de vista fiscal, pois esses se tornam menos necessários.

cuidado. Um ponto interessante é que quando o horizonte se alonga, o desvio-padrão das expectativas fiscais (especialmente da razão da dívida em relação ao PIB) aumenta consideravelmente também, sinalizando uma maior incerteza⁷³.

Os resultados da análise de sensibilidade, contudo, mostram alguns períodos onde esse efeito é negativo no período pós-crise, atentando para o fato que o efeito negativo é consideravelmente maior em magnitude do que o efeito positivo. Uma dúvida que poderia ser resolvida através do cálculo do intervalo de confiança do produto da rede (que foge do escopo deste trabalho) é se apenas os efeitos negativos são significativos.

A surpresa de curto prazo do IPCA não se mostrou estatisticamente significativa a 10%, ao contrário do que ocorre em Lowenkron e Garcia (2006), onde esta variável impacta o desvio das expectativas em relação à meta. Contudo, o resultado aqui obtido é semelhante ao encontrado por Schmidt-Hebbel e Werner (2002).

Quanto à primeira defasagem das expectativas, o seu efeito positivo aumenta no período de crise e, apesar de diminuir em magnitude no período pós-crise de confiança, permanece em um nível maior que no período anterior à crise. De todas as variáveis, a defasagem de primeira ordem das expectativas de inflação é a que apresenta o maior efeito médio sobre as expectativas, evidenciando um forte comportamento adaptativo. O efeito da defasagem de ordem dois é menor e parece que vem diminuindo no período pós-crise. Dessa forma, o ajustamento das expectativas em função de choques em suas variáveis explicativas se dará de forma gradual.

É interessante notar também que, no período de crise, há uma maior dispersão dos efeitos de cada variável sobre as expectativas de inflação, medida pelos **desvios-padrão** dos efeitos das variações de cada variável em cada período. Isto é um indício de maior incerteza a respeito da inflação futura por parte dos agentes.

Dados os resultados expostos acima, podemos concluir que o Banco Central vem consolidando sua credibilidade ao longo do tempo, já que o efeito da meta sobre as expectativas inflacionárias vem crescendo e os efeitos das demais variáveis diminuindo. Porém, há ainda espaço para melhoria, dado o efeito expressivo de variáveis como a volatilidade cambial e o preço das *commodities* sobre as expectativas dos agentes, bem como a magnitude do efeito da própria meta.

⁷³ Confira os gráficos no Apêndice B.

5. Considerações Finais

O presente trabalho buscou analisar empiricamente os determinantes das expectativas de inflação no Brasil, no período recente (pós metas de inflação), utilizando para isso os modelos de RNAs, que possuem a característica atraente de reconhecer trajetórias altamente complexas (não-lineares) sem que se especifique a forma funcional dessa relação.

O aspecto fundamental do regime de metas de inflação é a coordenação das expectativas dos agentes. Se a autoridade monetária possui credibilidade, as metas de inflação servem como âncora na formação de expectativas. Dessa forma, entender os fatores que afetam tais expectativas é de extrema importância para o adequado direcionamento da política monetária.

De maneira distinta dos trabalhos anteriores que trataram do mesmo assunto, utilizamos os dados das expectativas desagregados por instituição. A principal motivação para o uso do painel de dados foi a reduzida amostra que se teria com os dados agregados. E como bem salientado por Hsiao (2005), a utilização de dados em painel permite uma maior capacidade para capturar a complexidade do comportamento humano, quando comparado a séries de tempo ou *cross-section* simples.

A utilização de RNAs no caso de dados em painel é ainda bastante recente. Alguns poucos trabalhos fizeram uso dessa combinação, a maioria sem nenhum tratamento específico. Desenvolvimentos nesta metodologia podem, no futuro, enriquecer a análise que aqui foi realizada. Uma alternativa bastante interessante para trabalhos futuros é a metodologia de *Panel Smooth Transition Regression Models*, proposta por González, Teräsvirta e van Dijk (2005). Além disso, um estudo mais cuidadoso do método de inferência aqui adotado, no que se refere à estrutura de dependência dos dados, será bem-vindo. Outra limitação, que pode ser resolvida por trabalhos futuros, é o cálculo do intervalo de confiança para o produto da rede na análise de sensibilidade, o que permitiria verificar se em determinado período a variável deixou de ser significativa ou não.

A análise dos dados referentes às expectativas inflacionárias mostra que o erro de previsão vem diminuindo consideravelmente, bem como a incerteza sobre a inflação futura. Esse comportamento parece refletir o forte comprometimento do Banco Central percebido pelos agentes em alcançar a meta de inflação, além de um maior

entendimento por parte do mercado da dinâmica da inflação (já que o mercado tem acertado mais no período recente).

Os resultados obtidos aqui nos permitem afirmar que a maior influência sobre as expectativas inflacionárias no período como um todo foi da volatilidade cambial, seguida pela variação no preço das *commodities*, pela defasagem de ordem um das expectativas, pela variação cambial e pela meta. Em menor magnitude, afetaram as expectativas o resultado primário do governo, a defasagem de ordem dois e a taxa Selic. O comportamento desse efeito ao longo do tempo foi verificado através da análise de sensibilidade do produto da rede em resposta a cada uma das variáveis. No período de crise de confiança, há um expressivo descolamento das expectativas em relação à meta, com um aumento do efeito das demais variáveis. Resultado contrário ocorre no período pós-crise: o efeito da meta de inflação é maior e das demais variáveis menor, ainda que em alguns casos expressivos (como da defasagem de ordem 1 e da volatilidade cambial). Isso nos leva a concluir que o Banco Central possui credibilidade e esta vem se consolidando ao longo do tempo, mas que há ainda espaço para melhorias.

REFERÊNCIAS

- ABELÉM, A.J.G. *Redes neurais artificiais na previsão de séries temporais*. 1994. Dissertação (Mestrado em Engenharia Elétrica) – Faculdade de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 1994.
- ALBUQUERQUE, C. R. *Taxas de câmbio e inflação no Brasil: um estudo econométrico*. 2005. Tese (Doutorado em Economia) – Faculdade de Ciências Econômicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005.
- ALBUQUERQUE, C. R.; PORTUGAL, M. S. Testing nonlinearities between brazilian exchange-rate and inflation volatilities. *Revista Brasileira de Economia*, Rio de Janeiro, v. 60, n. 4, 2006.
- ANDERS, U.; KORN, O. Model selection in neural networks. *Neural Networks*, Oxford, v.12, n.2, p. 309–323, 1999.
- ANDREWS, D.W.K. Higher order improvements of a computationally attractive k-step bootstrap for extremum estimators. *Econometrica*, v. 70, n. 1, p. 119-162, 2002.
- ANG, A.; BEKAERT, G.; WEI, M. Do macro variables, asset markets or surveys forecast inflation better? *NBER Working Paper*, Cambridge, n. 11538, p. 1-62, 2005.
- ARTHUR, W.B. The end of certainty in economics. *Mathematical Social Sciences*, v. 33, n.1, p. 96-96, Feb. 1997.
- AZOFF, E.M. *Neural Network time series forecasting of financial markets*. Chichester: John Wiley & Sons, 1994.
- BALL, L.; CROUSHORE, D. Expectations and the effects of monetary policy. *Journal of Money, Credit and Banking*, Columbus, v. 34, n. 4, p. 473-484, 2003.
- BARRON, A.R. A comment on “Neural networks: a review from a statistical perspective”. *Statistical Science*, v.9, n.1, p. 33-35, 1994.
- BAXT, W.G.;WHITE, H. Bootstrapping confidence intervals for clinical input variable effects in a network trained to identify the presence of acute myocardial infarction. *Neural Computation*, Cambridge, v. 7, p. 624-638, 1995.
- BEGG, D.K.H. *The rational expectations revolution in macroeconomics*. Oxford: Philip Allan, 1982.
- BELTRATTI, A.; MARGARITA, S.; TERNA, P. *Neural networks for economic and financial modelling*. London: International Thomson Computer, 1996.
- BEVILAQUA, A.S.; MESQUITA, M.; MINELLA, A. Brazil: taming inflation expectations. *Working Paper Series: Banco Central do Brasil, Brasília*, n. 129, p. 1-39, 2007.

BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. *Redes neurais artificiais: teoria e aplicações*. Rio de Janeiro: LTC, 2000.

BURGESS, A.N. Non-linear model identification and statistical significance tests and their application to financial modelling. In: *IEEE Proceedings of the 4th International Conference on Artificial Neural Networks*, Cambridge, p. 312-317, 1995.

CAGAN, P. The monetary dynamics of hyper-inflation. In: FRIEDMAN, M. (Ed.). *Studies in the quantity theory of money*, Chicago: University of Chicago Press, 1956.

CAMPBELL, J.Y.; LO, A.W.; MACKINLAY, A.C. *The econometrics of financial markets*. Princeton, NJ: Princeton University Press, 1997.

CARVALHO, F.A. Testing the rationality of market expectations on Brazilian inflation: evidence from an unbalanced panel. *Working Paper UNB*, Brasília, n. 307, p. 1-24, Apr. 2004. Disponível em: <<http://www.unb.br/face/eco/cpe/TD/307abril2004Fabia.pdf>>.

CARVALHO, F.A.; BUGARIN, M.S. Inflation expectations in Latin América. *Economía*, v. 6, n. 2, p. 101-145, 2006.

CARVALHO, L.M.; SOUZA JÚNIOR, J.R. DE C. Modelo para previsão de inflação. *Boletim de Conjuntura IPEA*, Rio de Janeiro, n. 72, p. 97-102, 2006.

CERISOLA, M.; GELOS, R.G. What drives inflation expectations in Brazil? An empirical analysis. *IMF Working Paper*, Washington, n. 05/109, p. 1-22, 2005.

CHAN, M.-C.; LAM, C.-C.; WONG, C.-C. Financial time series forecasting by neural network using conjugate gradient learning algorithm and multiple linear regression weight initialization. *Computing in Economics and Finance 2000*, Barcelona, n. 61, 2000.

CHATFIELD. C. Neural networks: forecasting breakthrough or passing fad? *International Journal of Forecasting*, 9, p.1-3, 1993.

CHENG, B.; TITTERINGTON, D.M. Neural networks : a review from a statistical perspective. *Statistical Sciences*, v. 9, n. 1, p.2-54, 1994.

COLLINS, E.; GHOSH, S.; SCOFIELD, C. An application of a multiple neural network learning system to emulation of mortgage underwriting judgements. *Proceedings of the IEEE International Conference on Neural Networks*, San Diego, v. 2, p. 459-466, 1988.

CÔRREA, W.R.; PORTUGAL, M.S. Previsão de séries de tempo na presença de mudança estrutural: redes neurais artificiais e modelos estruturais. *Economia Aplicada*, Ribeirão Preto, v. 2, n. 3, p. 487-514, 1998.

DAVIDSON, R.; MACKINNON, J. G. *Estimation and inference in econometrics*. New York: Oxford University Press, 1993.

_____; _____. Bootstrap tests: how many bootstraps? *Econometric Reviews*, 19, p. 55-68, 2000.

_____; _____. The power of bootstrap and asymptotic tests. *Journal of Econometrics*, v. 127, n. 2, p. 421-441, 2006.

DEBOECK, G. J. *Trading on the edge*. New York: John Wiley & Sons, 1994.

DEMUTH, H.; BEALE, M. *Neural Network Toolbox for use with Matlab*. Natick: The MathWorks, 2003.

DENTON, J.W. How good are neural networks for causal forecasting? *The Journal of Business Forecasting*, v. 14, n.2, p. 17-20, Summer 1995.

DIAZ, M. D. M.; ARAÚJO, L. J. S. Aplicação de redes neurais à economia. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 2, n. 2, 1998.

DINIZ, H. *Integração de Redes Neurais Artificiais & métodos estocásticos para previsão de séries temporais*. 1999. Dissertação (Mestrado) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 1999.

DUTTA, S.; SHEKHAR, S. Bond rating: a non-conservative application of neural networks. *Proceedings of the IEEE International Conference on Neural Networks*, v. 2, San Diego, p. 443-450, 1988.

EFRON, B.; TIBSHIRANI, R.J. *An introduction to the bootstrap*. New York: Chapman and Hall, 1993.

EVANS, G.; HONKAPOHJA, S. *Learning and expectations in macroeconomics*. Princeton: Princeton University Press, 2001.

FAUSET, L. *Fundamentals of neural networks: architectures, algorithms and applications*. Upper Saddle River: Prentice Hall, 1994.

FERNANDES, L.G.L.; PORTUGAL, M.S. Redes neurais artificiais e previsão de séries econômicas: uma introdução. *Nova Economia*, Belo Horizonte, v. 6, n. 1, p. 51-74, 1996.

_____; _____. NAVAUX, P.O.A. Previsão de séries de tempo: redes neurais artificiais e modelos estruturais. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 26, n.2, p. 253-276, 1996.

FREEMAN, J.A.; SKAPURA, D.M. *Neural networks: algorithms, applications, and programming techniques*. New York: Addison-Wesley, 1992.

FRIEDMAN, M. *Theory of the consumption function*. Princeton: Princeton University Press, 1957.

FRIEDMAN, M. The role of the monetary policy. *American Economic Review*, Nashville, vol. 58, p. 1-17, 1968.

FUJITA, O. Statistical estimation of the number of hidden units for feedforward neural networks. *Neural Networks*, v. 11, n.5, p. 851-859, 1998.

GALLANT, A.R.; WHITE, H. There exists a neural network that does not make avoidable mistakes? *Proceedings of the Second Annual IEEE Conference on Neural Networks*, San Diego, v. 1, p. 657-664, 1988.

GARCIA, M. G. P. The formation of inflation expectations in Brazil: a study of the futures market for the price level. *Revista Brasileira de Economia*, Rio de Janeiro, v. 48, n. 1, p. 47-70, Jan./Mar. 1994.

_____. Inflation Targeting in Brazil: Evaluation and Policy Lessons for Latin American countries. Rio de Janeiro: PUC, jan. 2006.

GONCALVES, C. E.; GUIMARÃES, B. Monetary policy and the exchange rate in Brazil. In: XXVIII ENCONTRO BRASILEIRO DE ECONOMETRIA, 2006, Salvador. *Anais do XXVIII Encontro Brasileiro de Econometria*, 2006.

GONZÁLEZ, A.; TERÄSVIRTA, T.; VAN DIJK, D. Panel Smooth Transition Regression Models. *Research Paper: Quantitative Finance Research Centre of the University of Technology Sydney*, n. 165, p. 1-14, 2005.

GONZALEZ, S. Neural Networks for macroeconomic forecasting: a complementary approach to linear regression models. *Working Papers: Department of Finance: Canada, Ottawa*, n. 07, p.1-40, 2000.

GRANGER, C.W.J.; TERÄSVIRTA, T. *Modelling nonlinear economic relationships*. Oxford: Oxford University Press, 1993.

GREENE, W.H. *Econometric Analysis*. Upper Saddle River: Prentice Hall, 2003.

HAGAN, M.T.; DEMUTH, H.B.; BEALE, M. *Neural network design*. New York: PWS Publishing Company, 1995.

HÄRDLE, W.; HOROWITZ, J. L.; KREISS, J.P. Bootstrap methods for time series. *International Statistical Review*, v. 71, n. 2, p. 435-459, 2003.

HAYKIN, S. *Redes neurais: princípios e prática*. Porto Alegre: Bookman, 2001.

HONORÉ, B.E. Non-linear models with panel data. *CEMMAP Working Paper*, London, n. CWP13/02, p. 1-21, 2002.

HORNIK, K.; STINCHCOMBE, S.M.; WHITE, H. Multilayer feedforward networks are universal approximators. *Neural Networks*, v. 2, n. 5, p.359-366, 1989.

HOROWITZ, J. L. Bootstrap methods in econometrics: theory and numerical performance. In: KREPS D.M.; WALLIS, K.F. (Eds.). *Advances in Economics and Econometrics: Theory and Applications*. New York: Cambridge University Press, 1997, p. 188-222.

HSIAO, C. *Analysis of panel data*. New York: Cambridge University Press, 2003.

HSIAO, C. Why panel data? *IEPR Working Paper*, Los Angeles, n. 05-33, p. 1-17, 2005.

KAPETANIOS, G. A bootstrap procedure for panel datasets with many cross-sectional units. *Working Papers*: Department of Economics of the University of London, Cidade, n. 523, p.1-15, 2004.

KEANE, M.P.; RUNKLE, D.E. Testing the rationality of price forecasts: new evidence from panel data. *American Economic Review*, Nashville, v. 80, n. 4, 714-735, 1990.

KOVÁCS, Z.L. *Redes neurais artificiais: fundamentos e aplicações*. São Paulo: Collegium Cognitio, 1996.

KUAN, C.M.; WHITE, H. Artificial neural networks: an econometric perspective. *Econometric Reviews*, v. 13, n. 1, p.1-91, 1994.

LA ROCCA, M.; PERNA, C. Subsampling in artificial neural networks for hydrological data. *Metodi statistici e matematici per l'analisi delle serie idrologiche*, p. 103-114, 2003. Disponível em: www.irpi.pg.cnr.it/grimaldi/papers/LaRoccaPerna2003.pdf.

_____; _____. Variable selection in neural network regression models with dependent data: a subsampling approach. *Computational Statistics & Data Analysis*, v. 48, n.2, p. 415-429, 2005.

LAHIRI, S.N. *Resampling methods for dependent data*. New York: Springer, 2003.

LEE, T.H.; WHITE, H.; GRANGER, C.W.J. Testing for neglected nonlinearity in time series models - a comparison of neural network methods and alternative tests. *Journal of Econometrics*, v. 56, n.3, p. 269-290, 1993.

LENT, R. *Cem bilhões de neurônios: conceitos fundamentais de neurociência*. São Paulo: Atheneu, 2001.

LIMA, E.C.R.; CÉSPEDES, B.J.V. O desempenho do Mercado (FOCUS) e do Bacen na previsão da inflação: comparações com modelos lineares univariados. *Boletim de Conjuntura IPEA*, Rio de Janeiro, n. 60, p. 75-83, mar. 2003.

LONGHI, S.; NIJKAMP, P.; REGGIANI, A.; MAIERHOFER, E. Neural network modeling as a tool for forecasting regional employment patterns. *International Regional Science Review*, v. 28, n. 3, p.330-346, 2005.

LOWENKRON, A.; GARCIA, M. Monetary Policy Credibility and Inflation Risk Premium: a model with application to Brazilian data. *Textos para Discussão*: Departamento de Economia PUC-RJ, Rio de Janeiro, n. 543, p. 1-21, 2006.

LUCAS, R.E. Econometric policy evaluation: a critique. In: CARNEGIE-ROCHESTER CONFERENCE SERIES ON PUBLIC POLICY. Amsterdã: North-Holland, 1976, p. 19-46.

- LUCAS, R.E. Expectations on the Neutrality of Money. *Journal of Economic Theory*, v. 4, p. 103-124, 1972.
- MACKINNON, J.G. Bootstrap inference in econometrics. *Canadian Journal of Economics*, v. 35, n.4, p. 615-645, 2002.
- MADALLA, G.S. *Introduction to Econometrics*. New York: Macmillan, 1988.
- MARQUES, A.; FACHADA, J.; CAVALCANTI, D. Sistema Banco Central de Expectativas de Mercado. *Notas Técnicas do Banco Central do Brasil*, Brasília, n. 36, p. 1-27, 2003.
- MARQUEZ, L.; HILL, T.; WORTHLEY, R.; REMUS, W. Neural network models as an alternative to regression. *Proceedings of the 24th Hawaii International Conference on System Sciences*, p.129-135, 1991.
- McCULLOCH, W.S; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, v. 5, p. 115-133, 1943.
- MEDEIROS, M.C.; TERÄSVIRTA,T. Statistical methods for modelling neural networks. *Texto para discussão*: PUC-RJ, Rio de Janeiro, n. 445, p.1-24, 2001.
- _____; VEIGA, A. Diagnostic checking in a flexible nonlinear time series model. *Journal of Time Series Analysis*, v. 24, n. 4, p. 461-482, 2003.
- _____;_____; RECH, G. Building neural network models for time series: a statistical approach. *Journal of Forecasting*, v. 25, n.1, p. 49-75, 2006.
- MELLO, L.M.; MOCCERO, D. Monetary policy and inflation expectations in Latin America: long-run effects and volatility spillovers. *OECD Working Paper*, n. 518, p.1-26, 2006.
- MINELLA, A.; FREITAS, P.; GOLDFAJN, I.; MUINHOS, M. Inflation Targeting in Brazil: Lessons and Challenges. *Working Paper of the Central Bank of Brazil*, Brasília, n. 53, 2002.
- MINELLA, A.; FREITAS, P.; GOLDFAJN, I.; MUINHOS, M Inflation targeting in Brazil: constructing credibility under exchange rate volatility. *Journal of International Money and Finance*, v. 22, n.7, p. 1015-1040, 2003.
- MINSKY, M.L.; PAPERT, S.A. *Perceptrons*. Cambridge, MA: MIT Press, 1969.
- MISRA, M.; WARNER, B. Understanding neural networks as statistical tools. *The American Statistician*, v. 50, n.4, p. 284-293, 1996.
- MUTH, J. Rational Expectations and the theory of price movements. *Econometrica*, v. 29, n. 3, p. 315-335, 1961.
- NELSON, M.M.; ILLINGWORTH, W.T. *A Practical Guide to neural nets*. New York: Addison-Wesley, 1991.

NERLOVE, M. *The dynamics of supply: estimation of the farmers' response to price*. Baltimore: Johns Hopkins University Press, 1958.

ORPHANIDES, A.; WILLIAMS, J.C. Imperfect knowledge, inflation expectations, and monetary policy. *Federal Reserve Board Working Papers*, Washington, n. 27, p. 1-35, 2002.

ORR, G.B.; MÜLLER, K.-R. (Ed.). Neural Networks: tricks of the trade. *Lectures Notes in Computer Science*, v.1524, p. 93-112, 1998.

PALMA, A.A. *Inferência estatística em redes neurais artificiais: uma aplicação à taxa de câmbio brasileira*. 2004. Monografia (Graduação em Economia) - Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2004.

PALMA, A.A.; SARTORIS, A. Inference in artificial neural networks: testing market efficiency for the brazilian exchange rate. *Anais do Encontro Latino-Americano de Econometria – LAMES*. Cidade do México, México. 2006.

PAO, H-T; CHIH, Y-Y. Comparison of TSCS regression and neural network models for panel data forecasting: debt policy. *Neural Computing & Applications*, London, v. 15, n.2, p. 117-123, 2006.

PASTORE, A. C.; PINOTTI, M.C.; ALMEIDA, L.P. Public debt, exchange rate shocks and inflation targets. *Seminário 5 anos de Metas de Inflação*, Banco Central do Brasil, 2004. Disponível em: www.bcb.gov.br/Pec/Seminario5Anos/port/pastorePaper.pdf. Acesso em mar. 2007.

PATUELLI, R.; REGIANNI, A.; NIJKAMP, P.; BLIEN, U. New neural network methods for forecasting regional employment: an analysis of German labour markets. *Spatial Economic Analysis*, v.1, n.1, p.7-30, 2006.

PELTONEN, T.A. Are emerging market currency crises predictable? A test. *Working Paper Series: European Central Bank*, n.571, p. 1-49, 2006.

PEREIRA, R.R. Comunicação em Política Monetária. In: ENCONTRO NACIONAL DE ECONOMIA, 32., 2004, João Pessoa. *Anais...* João Pessoa: ANPEC, 2004.

PIRES, M.C.C. Credibilidade na política fiscal: uma análise preliminary para o Brasil. *Economia Aplicada*, São Paulo, v. 10, n.3, p. 367-375, 2006.

PIZARRO, J.; GUERRERO, E.; GALINDO, P.L. A statistical model selection strategy applied to neural networks. In: EUROPEAN SYMPOSIUM ON ARTIFICIAL NEURAL NETWORKS, 8th., 2000, Bruges. *Proceedings...*, Bruges: D-Facto, 2000. p. 55-60.

PLIKYNAS, D.; SAKALAIUSKAS, L.; POLIAKOVA, A. Analysis of foreign investment impact on the dynamics of national capitalization structure: a computational intelligence approach. *Research in International Business and Finance*, v. 19, n. 2, p. 304-332, 2005.

PLIKYNAS, D.; AKBAR, Y. Neural network approaches to estimating FDI flows: evidence from East and Central Europe. *Eastern European Economics*, v. 44, n. 3, p. 29-59, 2006.

PLIKYNAS, D.; AKBAR, Y. Application of modified MLP input weights' matrices: an analysis of sectorial investment distribution in the emerging markets. *Neural Computing & Applications*, London, v. 15, n. 3, p. 183-196, 2006.

POLITIS, D.N.; ROMANO, J.P.; WOLF, M. *Subsampling*. New York: Springer, 1999.

PORTUGAL, M.S. Rational Expectations and Economic Policy. University of Warwick, 1988.

PORTUGAL, M.S. Neural networks versus time series methods: a forecasting exercise. *Revista Brasileira de Economia*, Rio de Janeiro, v. 49, n. 4, p. 611-629, 1995.

PRATAP, R. *Getting started with Matlab 7: A quick introduction for scientists and engineers*. New York: Oxford University Press, 2006.

RACINE, J. WHITE, H. Statistical inference, the bootstrap, and neural network modeling with application to foreign exchange rates. *IEEE Transactions on Neural Networks*, v. 12, n. 4, p.1-19, 2001.

RESLER, D.H. The formation of inflation expectations. *Federal Reserve Bank of St. Louis Review*, St. Louis, p. 2-12, Apr. 1980.

RIPLEY, B.D. Statistical aspects of neural networks. In: BARNDORFF-NIELSEN, O.E.; JENSEN, J.L.; KENDALL, W.S. (Ed.). *Networks and chaos-statistical and probabilistic aspects*. London: Chapman and Hall, 1993. p. 40-123.

ROJAS, R. *Neural networks: a systematic introduction*. New York: Springer-Verlag, 1996.

ROMERO, R.A.F.; SANTOS, V.P. *Utilização do Matlab para o desenvolvimento de redes neurais*. São Carlos: ICMCSC – USP, 1994. (Relatório Técnico n. 18).

SABBATINI, R.M.E. Neurônios e sinapses: a história de sua descoberta. *Cérebro & Mente*, n. 17, mai-jun 2003.

SANTOS, N.C. *A hipótese das expectativas racionais: teoria e testes*. 2003. Dissertação (Mestrado Profissionalizante em Economia) - Faculdade de Ciências Econômicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2003.

SARGENT, T. *Bounded Rationality in Macroeconomics*. Oxford: Oxford University Press, 1993.

_____; WALLACE, N. Some unpleasant monetarist arithmetic. *Federal Reserve Bank of Minneapolis Review*, p. 1-19, Fall 1981.

SARLE, W.S. Neural networks and statistical models. *Proceedings of the SAS Users Group International Conference*, p. 1528-1550, 1994.

SCHMIDT-HEBBEL, K.; WERNER, A. M. Inflation Targeting in Brazil, Chile and Mexico: performance, credibility, and the exchange rate. *Central Bank of Chile Working Paper*, Santiago, n. 171, p. 1-51, 2002.

SHACHMUROVE, Y. Applying artificial neural networks to business, economics and finance. *Penn CARESS Working Paper*: University of Pennsylvania, Philadelphia, n. 8, p. 1-47, 2002.

SHANKER, M.;HU, M.Y.; HUNG, M.S. Effect of data standardization on neural network training. *Omega*, v. 24, n. 4, p. 385-397, 1996.

SIMON, E. Forecasting foreign exchange rates with neural networks. *Project Report*: Artificial Intelligence Laboratory, Swiss Federal Institute of Technology. 2002.

SICSÚ, J. Expectativas Inflacionárias no Regime de Metas de Inflação: uma análise preliminar do caso brasileiro. *Economia Aplicada*, Ribeirão Preto, v. 5, n. 1, p. 129-158, 2002.

SMITH, M. *Neural networks for statistical modeling*. New York: van Nostrand Reinhold, 1993.

SPANOS, A. *Statistical foundations of econometric modelling*. New York: Cambridge University Press, 1986.

STERN, H.S. Neural networks in applied statistics. *Technometrics*, v. 38, n.3, p. 205-214, 1996.

SWANSON, N.R.; WHITE, H. A model selection approach to real time macroeconomic forecasting using linear models and artificial neural networks. *Review of Economics and Statistics*, v. 79, p. 540-550, 1997.

SWINGLER, K. *Applying neural networks: a practical guide*. London: Academic Press, 1996.

TRAPLETTI, A. *On neural networks as statistical time series models*. 2000. Thesis (PhD) - Institut für Statistik, Technische Universität Wien. 2000.

VON NEUMANN, J. *The Computer and the brain*. New Haven: Yale University Press, 1958.

WALDE, J.F. Valid Statistical inference based on feedforward artificial neural networks models. In: YOUNG STATISTICIANS MEETING, 7th., 2003, Cidade. *Proceedings of the Seventh Young Statisticians Meeting*, 2003. Disponível em: <http://mrvar.fdv.uni-lj.si/pub/mz/mz21/walde.pdf>. Acesso em: 29 jan. 2004.

WALSH, C.E. *Monetary theory and policy*. Cambridge: The MIT Press, 2003.

WERNECK, R.L.F. Importância da agenda fiscal. In: SEMINÁRIO ANUAL DO BANCO CENTRAL DO BRASIL, 2006, Rio de Janeiro. Disponível em: <http://www.bcb.gov.br/Pec/Seminarios/SemMetInf2006/port/RogérioWerneck.ppt>. Acesso em mar. 2007.

WHITE, H. Some asymptotic results for learning in single hidden layer feedforward network models. *Journal of the American Statistical Association*, 84, p. 1003-1013, 1989a.

WHITE, H. Connectionist nonparametric regression: multilayer feedforward networks can learn arbitrary mappings. *Neural Networks*, Oxford, v. 3, n. 5, p. 535-549, 1990.

WHITE, H. Learning in artificial neural networks: a statistical perspective. *Neural Computation*, Cambridge, v.1, n. 4, p. 425-464, 1989b.

WHITE, H. *Estimation, inference, and specification analysis*. New York: Cambridge University Press, 1994.

WOOLDRIDGE, J.M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press, 2002.

WONG, B. K.; SELVI, Y. Neural network applications in finance: a review and analysis of literature. *Information & Management*, v. 34, n. 3, p. 129-139, 1998.

YIM, J. *Previsão de séries de tempo: modelos ARIMA, modelos estruturais e redes neurais artificiais*. 2001. 87 f. Dissertação (Mestrado) - Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2001.

YOSHIHIRO, S. *A formação da expectativa inflacionária no regime monetário de metas de inflação e a credibilidade do Banco Central do Brasil*. 2005. Dissertação (Mestrado em Economia) – Escola de Economia de São Paulo, Fundação Getúlio Vargas, São Paulo, 2005.

ZHANG, G.; HU, M.Y.; PATUWO, B.E. Forecasting with artificial neural networks: the state of the art. *International Journal of Forecasting*, v. 14, n.1, p. 35-62, 1998.

APÊNDICE

APÊNDICE A. Variáveis explicativas

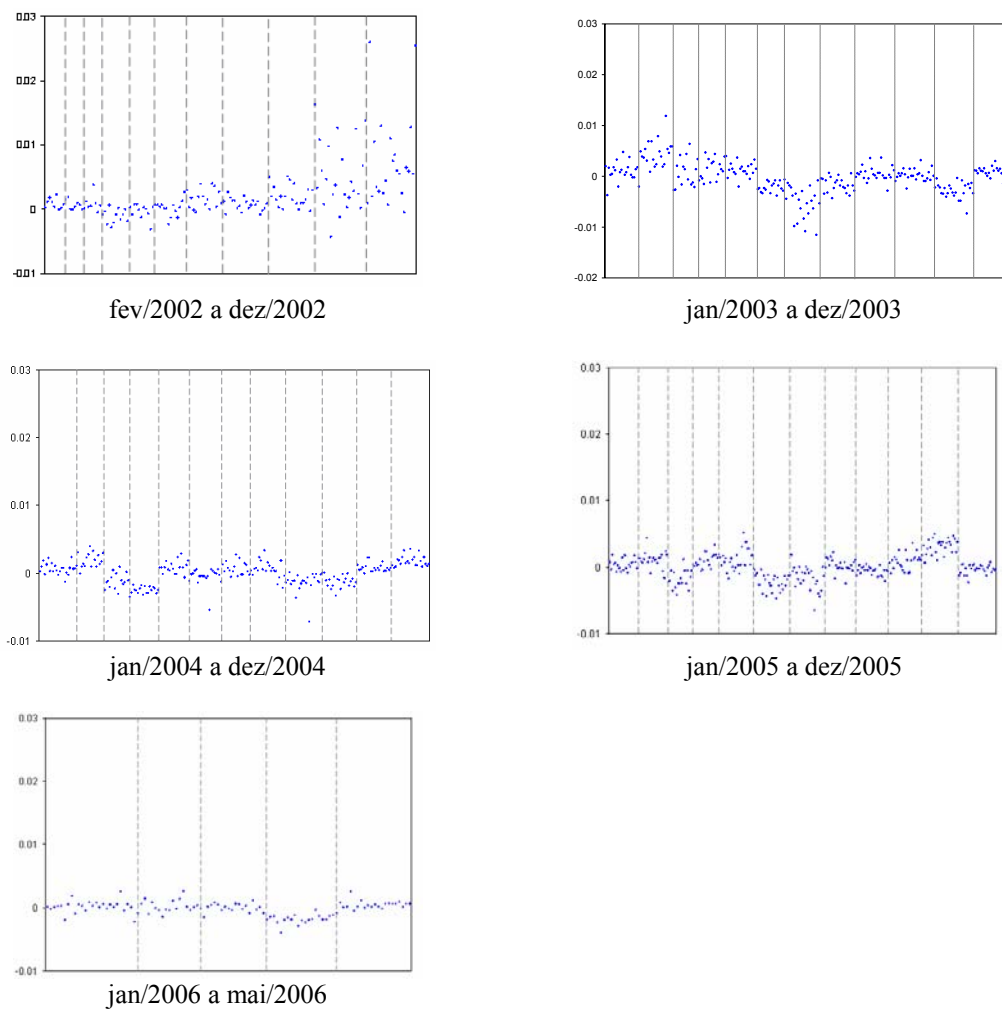


Gráfico A.1. Surpresa de curto prazo do IPCA-15
Fonte: Banco Central do Brasil e IBGE

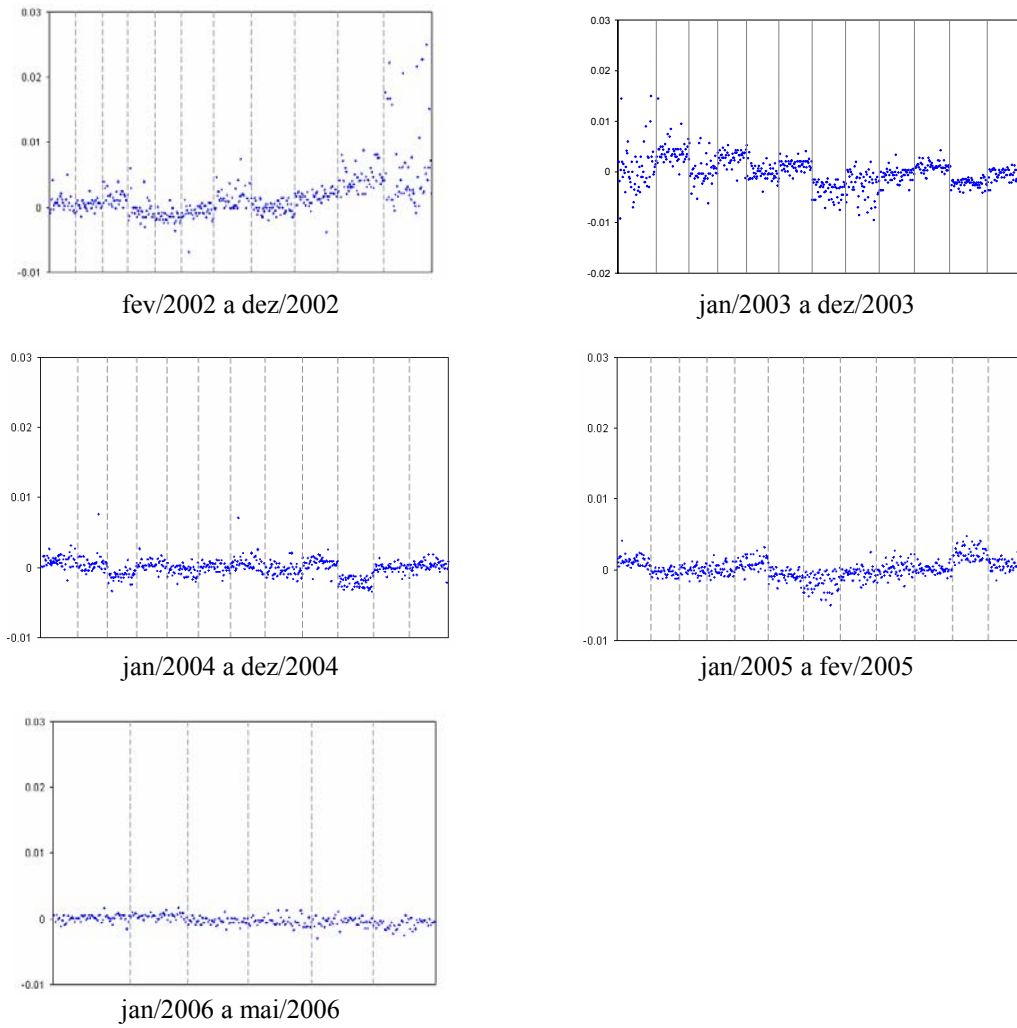


Gráfico A.2. Surpresa de curto prazo IPCA
 Fonte: Banco Central do Brasil e IBGE

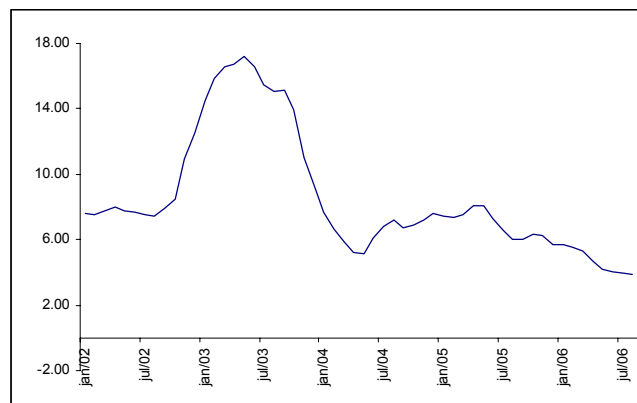


Gráfico A.3. Inflação (IPCA) acumulada em 12 meses
 Fonte: Dados do IBGE

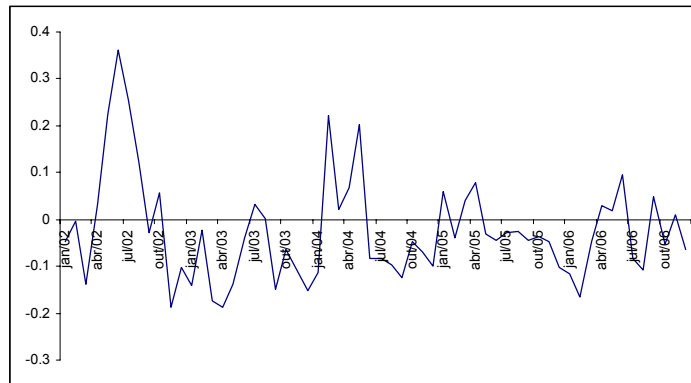


Gráfico A.4 Variação do Risco-Brasil (EMBI Brasil)
Fonte: Macrodados

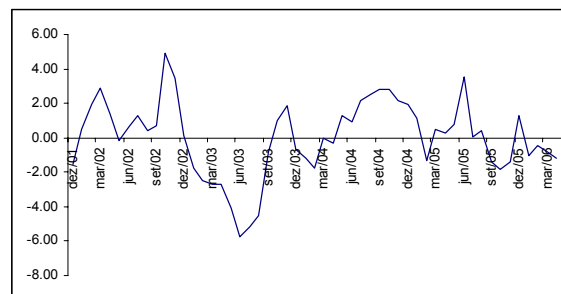


Gráfico A.5. Hiato do produto
Fonte: Dados da produção industrial - IBGE

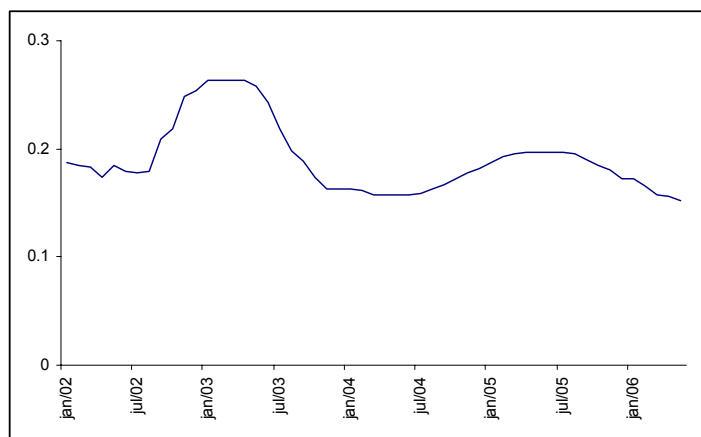


Gráfico A.6. Taxa de juros Selic
Fonte: Banco Central do Brasil

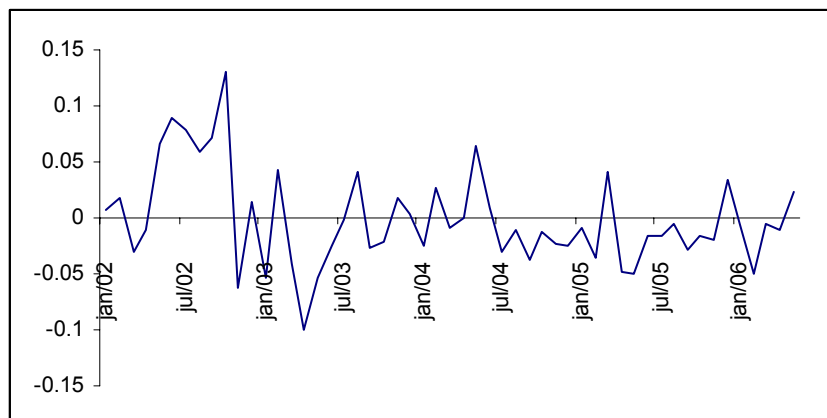
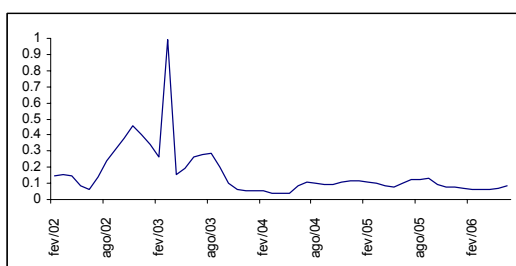
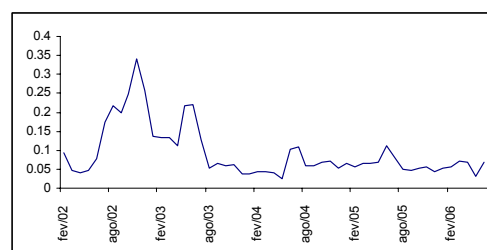


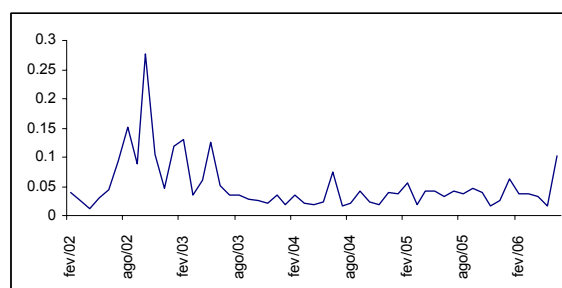
Gráfico A.7 Variação do câmbio
 Fonte: Banco Central do Brasil



6 meses



3 meses



1 mês

Gráfico A.8. Volatilidade do câmbio
 Fonte: Calculado a partir de dados do Banco Central do Brasil

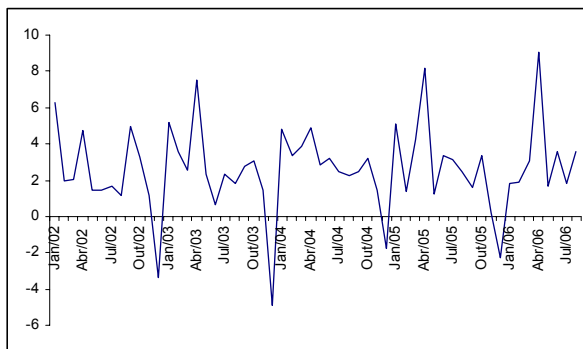


Gráfico A.9 Resultado primário do governo central
Fonte: STN – Secretaria do Tesouro Nacional

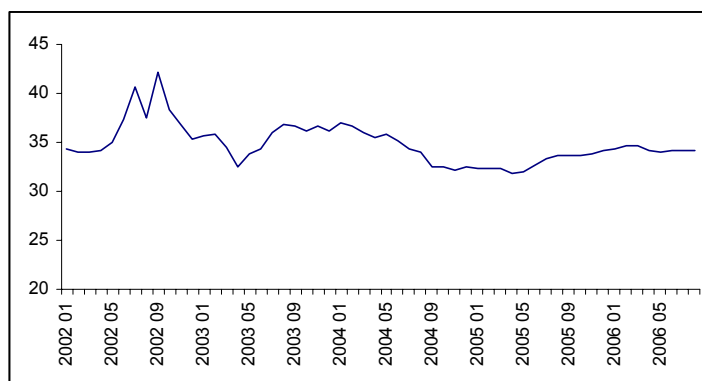


Gráfico A.10. Dívida total (% PIB)
Fonte: Banco Central do Brasil

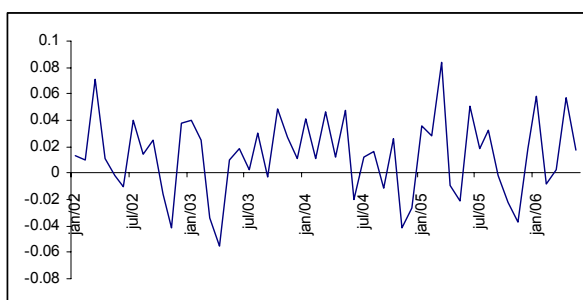


Gráfico A.11 Variação dos preços das commodities em t-1
Fonte: IPEA

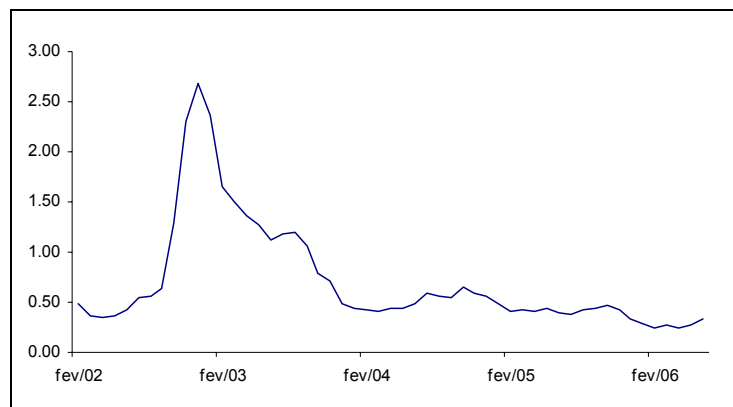


Gráfico A.12 Desvio-padrão médio das expectativas em t-1
Fonte: Banco Central do Brasil

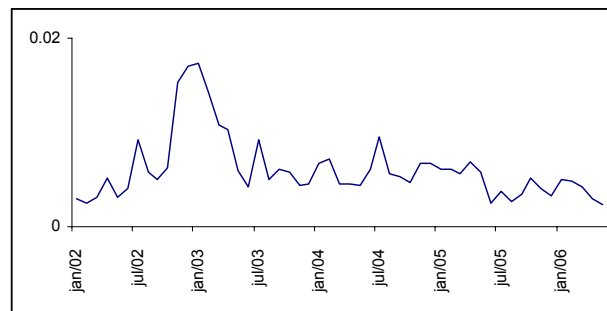


Gráfico A.13 Média das previsões médias das instituições Top 5 de longo prazo
Fonte: Banco Central do Brasil

APÊNDICE B. Alguns Modelos intermediários

Tabela B.1. Função de Reação das Expectativas de Inflação: Modelo Geral I

Variáveis	
Lag1, lag2, $IPCA_{t-1}^{12m}$, $hiato_{t-2}$, taxa selic, meta, efeito fixo, efeito tempo, resultado primário, dívida total, volatilidade câmbio, variação cambial, EMBI, <i>commodities</i> , <i>dummy</i> , σ_{t-1} , erro curto prazo IPCA	
Conjunto de treinamento	
Number of neurons:	6
Goodness of fit:	91.5826 %
RMSE:	0.00700914
MAE:	0.00387922
MAPE:	0.0556442
SIGN:	1
CORR:	0.956988
CIS:	-6.1819461455675988
Conjunto de validação:	
Goodness of fit:	39.2327 %
RMSE:	0.00286637
MAE:	0.00230181
MAPE:	0.0593879
SIGN:	1
CORR:	0.66591

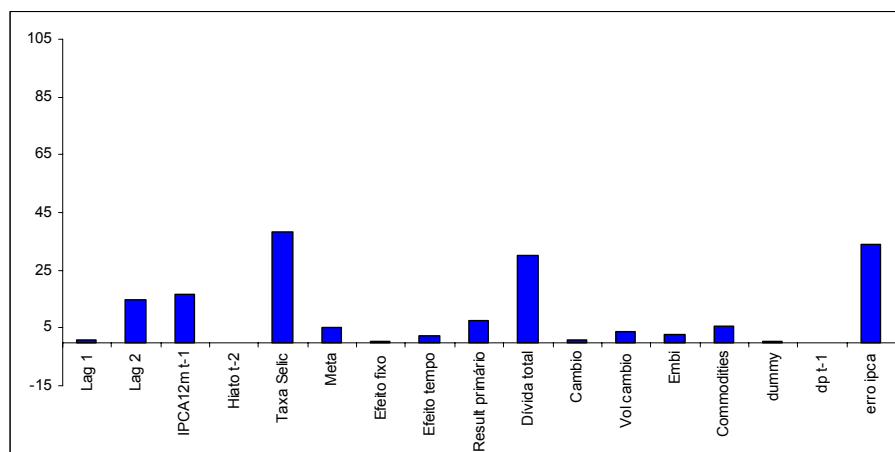


Gráfico B.1. Medida de relevância modelo I

As variáveis “menos relevantes” de acordo com os resultados acima expostos seriam: lag1, hiato do produto em t-2, efeito fixo, efeito tempo, câmbio, *dummy* e o desvio-padrão. Procedemos então ao teste de significância dessas variáveis.

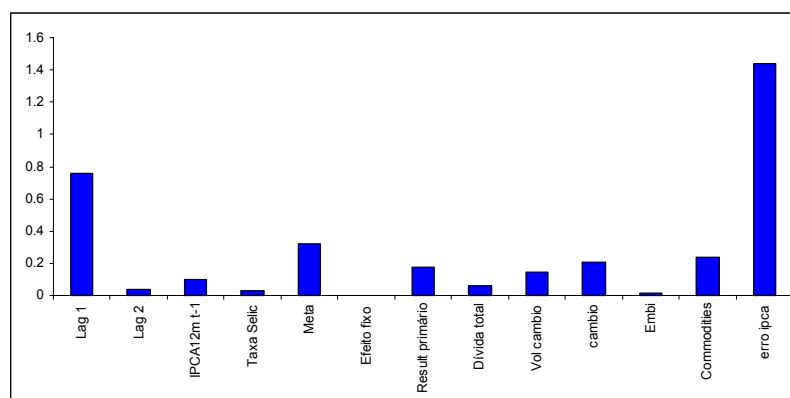
Tabela B.2. Inferência sobre as variáveis “menos” relevantes⁷⁴

Variável	Estatística de teste	Valor Crítico (5%)	Valor-p
Lag1	0.8779097669135172	0.8466624928426955	0.0460000000
Hiato do produto	0.002006474329580945	0.1297431138349556	0.3660000000
Efeito fixo	0.2480062886066492	0.1361161928648171	0.0240000000
Efeito tempo	2.093564560774443	90.4167829221479	0.4520000000
Câmbio	0.6323921661737989	1.1811299076377	0.1020000000
Dummy	0.2711741940	18.7104471704	0.4760000000
σ (expectativas)	0.1128418687931905	7.106095462147868	0.4640000000

Assim, temos que o hiato do produto, o efeito tempo, a *dummy* e o desvio-padrão das expectativas não são estatisticamente diferentes de zero. O modelo com essa restrição é apresentado a seguir.

Tabela B.3. Função de Reação das Expectativas

Variáveis	
Lag1, lag2, IPCA12mt-1, taxa selic, meta, efeito fixo, resultado primário, dívida total, câmbio, volatilidade câmbio, EMBI, <i>commodities</i> , erro curto prazo IPCA	
Conjunto de Treinamento	
Number of training observations:	2096
Number of neurons:	4
Number of weights:	61
Number of restarts:	10
Goodness of fit:	92.1314 %
RMSE:	0.00677681
MAE:	0.00404543
MAPE:	0.0590948
SIGN:	1
CORR:	0.959851
CIS:	-6.1186032351219302

**Gráfico B.2.** Medida de relevância

⁷⁴ O procedimento de *bootstrap* foi realizado com 500 repetições.

De acordo com o exposto acima, devemos agora testar a significância do lag 1, lag 2 e efeito fixo.

Tabela B.3. Inferência sobre as variáveis menos relevantes

Variável	Estatística de teste	Valor Crítico (5%)	Valor-p
Lag1	2.861111939516629	10.7692261525526	0.24400000000
Lag2	1.281777510348316	9.83364243878512	0.4125000000
Efeito fixo	0.07899553579046657	0.4514915298290749	0.4550000000

Dados os resultados acima, realizamos o seguinte exercício: treinamos um modelo sem as variáveis lag2 e efeito fixo e sem as três variáveis acima, que se mostraram insignificantes e retirando-se uma variável dentro deste conjunto por vez. As estimações não se mostraram melhores que as anteriores, de acordo com o CIS. Dessa forma, prosseguimos com a análise do modelo que inclui todas essas variáveis. Foi realizado o procedimento de inferência e calculada a sensibilidade do produto em relação a cada variável. Resultados estranhos foram obtidos: a totalidade das variáveis não se mostrou estatisticamente significativa, mesmo assim, as variáveis conjuntamente foram significantes. Suspeitamos que isso possa ter ocorrido devido à presença de multicolinearidade no modelo. Tentando minimizar esse efeito, algumas modificações foram realizadas. O IPCA acumulado em 12 meses defasado está altamente correlacionado com as variáveis incluídas no modelo. Assim, ele foi retirado. A correlação entre variação cambial e o risco-Brasil também é grande (0.65), assim como entre a variação cambial e a dívida pública. Realizamos um experimento retirando a dívida, o risco-Brasil e o IPCA, e os resultados foram animadores, dado que o efeito das entradas ficou menos sensível à retirada de variáveis. O modelo é apresentado na seção 4.4⁷⁵.

⁷⁵ Para confirmar a exclusão das variáveis hiato do produto, efeito tempo, *dummy* e desvio-padrão, fizemos alguns exercícios com a inclusão dessas variáveis no modelo final (sem o IPCA, risco-Brasil e dívida pública) e estas não se mostraram significativas a 10% em nenhum desses experimentos.

APÊNDICE C. Sensibilidade do produto da rede (variação %) a uma variação de 1% em cada *input*

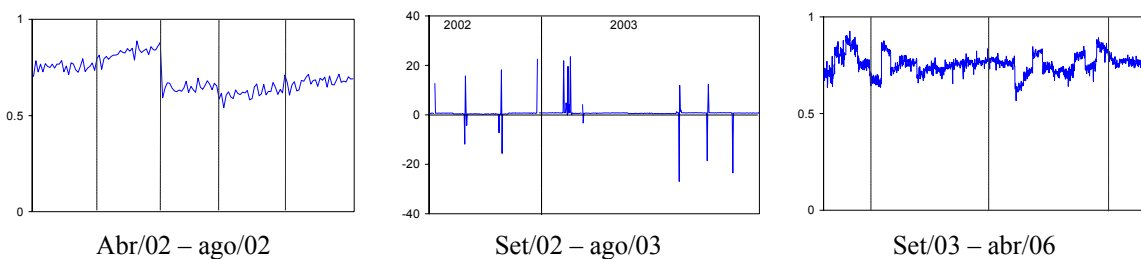


Gráfico C.1. Defasagem de ordem 1

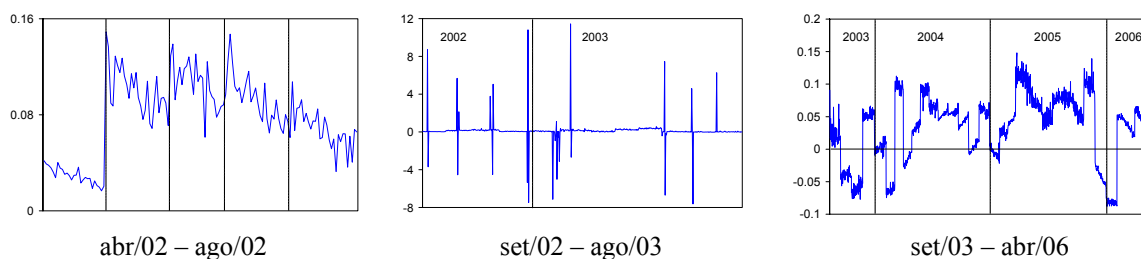


Gráfico C.2. Defasagem de ordem 2

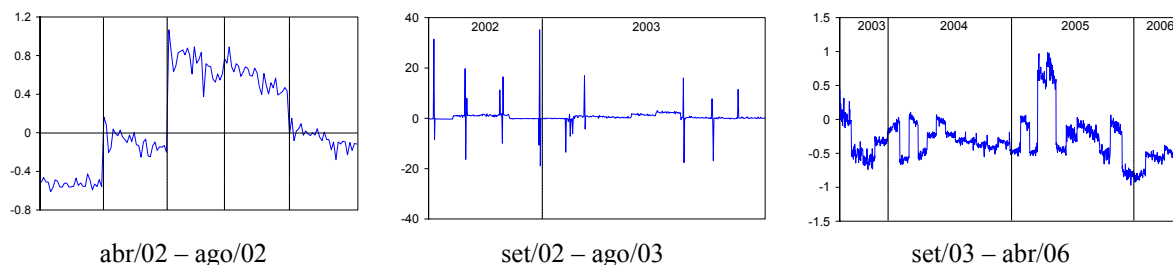


Gráfico C.3. Taxa Selic

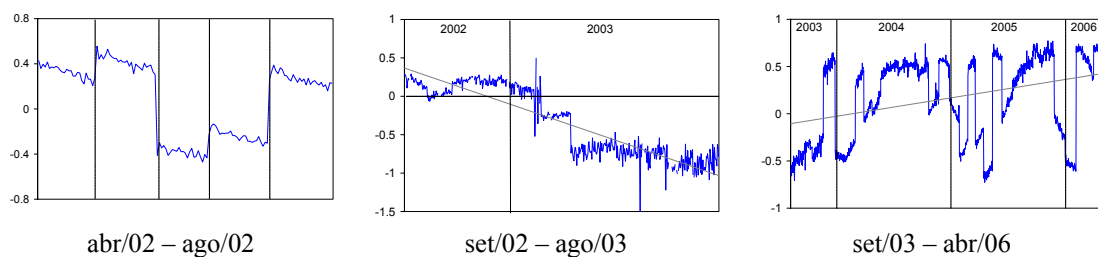


Gráfico C.4. Meta 12 meses

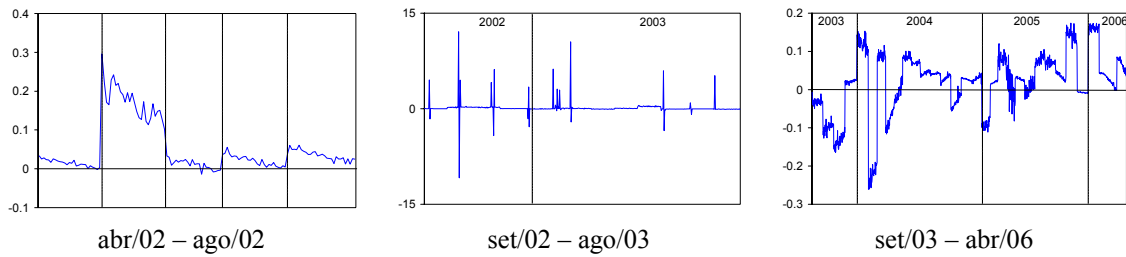


Gráfico C.5. Resultado Primário

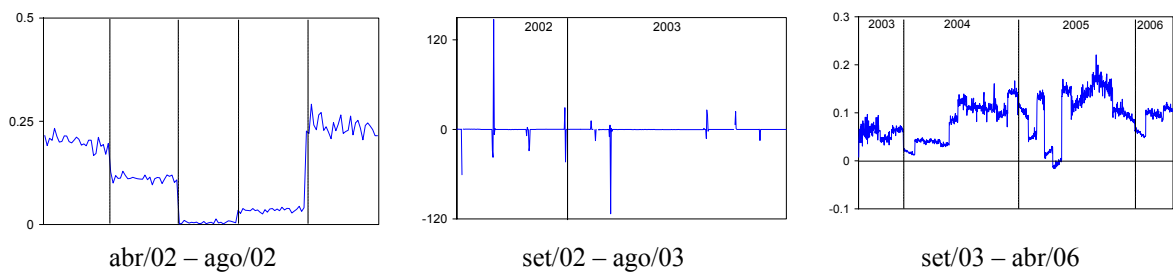


Gráfico C.6. Volatilidade cambial

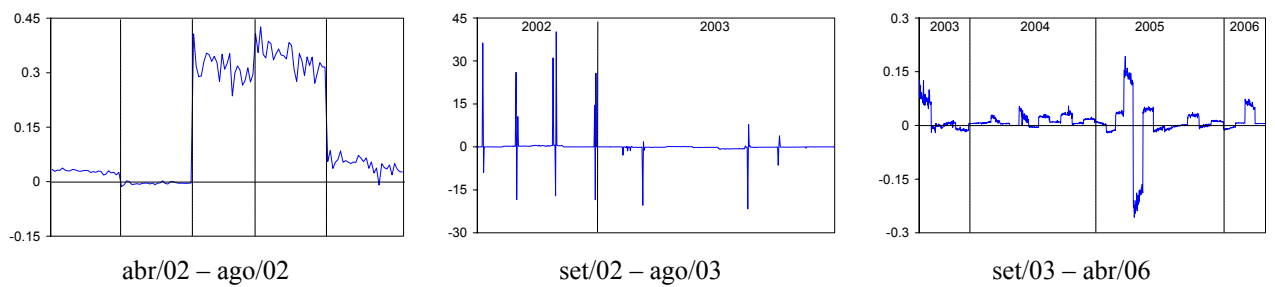


Gráfico C.7. Variação cambial

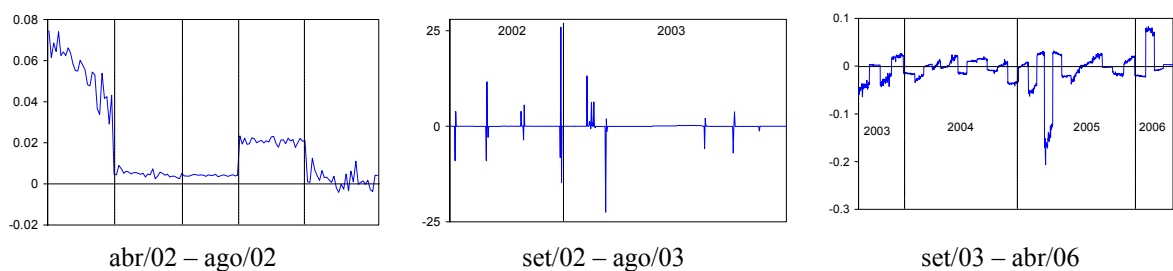


Gráfico C.8. Commodities

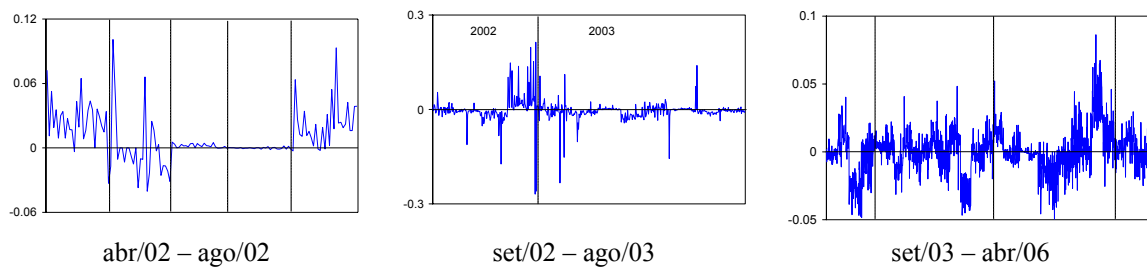
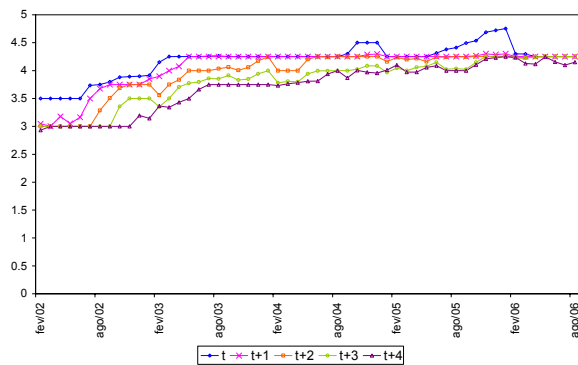
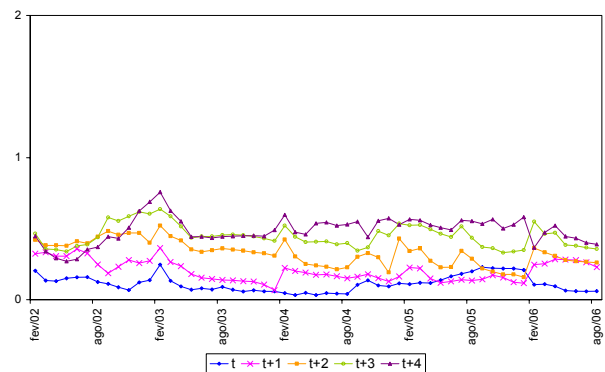


Gráfico C.9. Surpresa Curto Prazo IPCA

APÊNDICE D. Expectativas Fiscais

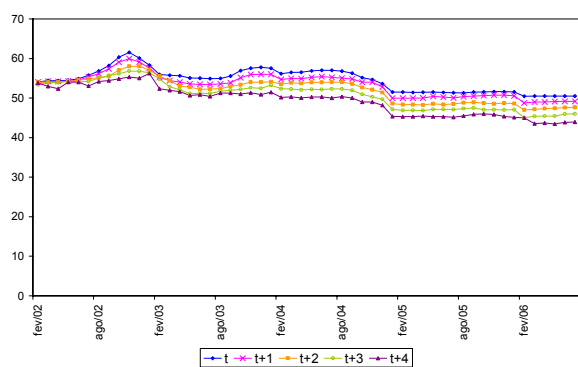


Média das medianas das expectativas

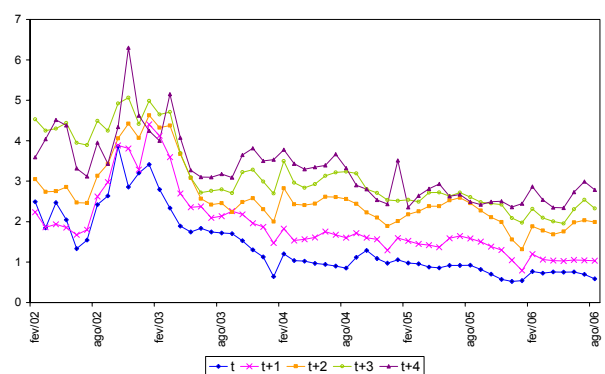


Desvio-padrão

Gráfico D.1. Expectativas para o resultado primário do governo: fev-2002 a ago-2006
Fonte: Banco Central do Brasil



Média das medianas das expectativas



Desvio-padrão

Gráfico D.2. Expectativas para a dívida líquida do setor público: fev-2002 a ago-2006
Fonte: Banco Central do Brasil