

ANTONIO CARLOS FERNANDES MAIA FILHO

**PREVISÃO DE RECESSÕES BRASILEIRAS COM INDICADORES
ANTECEDENTES: UMA ABORDAGEM HEURÍSTICA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Economia, para obtenção do título de *Magister Scientiae*.

VIÇOSA

MINAS GERIAS - BRASIL

2013

ANTONIO CARLOS FERNANDES MAIA FILHO

**PREVISÃO DE RECESSÕES BRASILEIRAS COM INDICADORES
ANTECEDENTES: UMA ABORDAGEM HEURÍSTICA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Economia, para obtenção do título de *Magister Scientiae*.

Erly Cardoso Teixeira

Luciano Dias de Carvalho
(Coorientador)

Newton Paulo Bueno
(Orientador)

Aos meus pais , dedico.

“O peso da evidência de uma afirmação extraordinária deve ser proporcional à sua estranheza.”

Pierre Simon Laplace

SUMÁRIO

LISTA DE TABELAS	vi
LISTA DE FIGURAS	viii
LISTA DE QUADROS.....	xii
RESUMO.....	xiii
ABSTRACT.....	xiv
1 INTRODUÇÃO	1
1.1 Considerações Iniciais.....	1
1.2 O Problema e sua Importância.....	8
1.3 Objetivos	12
1.3.1 Objetivo Geral	12
1.3.2 Objetivos Específicos	12
1.4 Escopo do Trabalho	12
2 ESTUDO DOS CICLOS ECONÔMICOS E A DEFINIÇÃO DE RECESSÃO	14
3 PREVISÃO COM INDICADORES ANTECEDENTES	21
3.1 Métodos heurísticos de previsão.....	25
3.1.1 Previsão de Recessões do <i>The Conference Board</i>	25
3.1.2 Previsão de Recessão com um Evento Extremo	31
3.2 Métodos estatístico-econométricos de previsão	36
3.2.1 Previsão de Recessões da Indústria de Bens de Capital Brasileira	36
3.2.2 Previsão de Recessões da Atividade Agregada Brasileira.....	39
4 EVENTOS EXTREMOS	45
5 METODOLOGIA	52
5.1 Análise Discriminante	53
5.2 Aspectos Metodológicos da Previsão	57

6 RESULTADOS E DISCUSSÕES.....	63
6.1 Análise de janeiro de 1975 até dezembro de 2012	63
6.2 Análise de janeiro de 1991 até dezembro de 2012	68
6.3 Seleção dos indicadores antecedentes.....	79
6.3.1 Análise de Correlação e o Método Stepwise de Seleção dos Indicadores Antecedentes	79
6.3.2 Análise Individual dos Indicadores Antecedentes	82
6.4 Análise discriminante: previsões	91
6.4.1 Análise discriminante com toda informação disponível	92
6.4.1.1 Regra de classificação com o ponto de corte para grupos com tamanhos distintos	94
6.4.1.2 Regra de classificação através da distância euclidiana quadrática	97
6.4.2 Análise Discriminante para Dados Fora da Amostra.....	101
6.4.2.1 Amostra de desenvolvimento	101
6.4.2.2 Amostra de teste	105
6.5 Recessões brasileiras como eventos extremos.....	109
6.5.1 Análise do Expoente de Hurst	111
6.5.2 Evolução do Componente Auto-Regressivo de Primeira Ordem.....	114
6.6 Recessões na indústria de bens de capital	120
6.6.1 Análise de previsão no curto prazo	122
7 CONCLUSÃO	128
Referências Bibliográficas	131
APÊNDICE.....	137

LISTA DE TABELAS

Tabela 1 - <i>Turning points</i> dos <i>business cycles</i> da atividade agregada brasileira datados pelo CODACE - 1980 a 2012	17
Tabela 2- Recessões dos Estados Unidos - 1960 a 2001.....	33
Tabela 3 - Recessões da indústria de bens de capital - 1991 a 2012.....	37
Tabela 4 - Recessões brasileiras desde 1975 até 2002.....	39
Tabela 5 - Valores de Escores Discriminantes Hipotéticos	58
Tabela 6 - <i>Turning points</i> dos <i>business cycles</i> da atividade agregada brasileira datados pelo CODACE - 1975 a 2012	65
Tabela 7 - <i>Turning points</i> dos <i>business cycles</i> da produção industrial brasileira datados pelo filtro polinomial - 1975 a 2012.....	65
Tabela 8 - Diferenças mensais entre o início das recessões na indústria e o início das recessões na atividade agregada brasileira - 1975 a 2012	66
Tabela 9 - <i>Turning points</i> dos <i>business cycles</i> da atividade agregada brasileira datados pelo CODACE - 1991 a 2012	71
Tabela 10 - <i>Turning points</i> dos <i>business cycles</i> da produção industrial brasileira datados pelo filtro polinomial - 1991 a 2012.....	72
Tabela 11 - <i>Turning points</i> dos <i>business cycles</i> de um índice coincidente composto por quatro séries: emprego, produção, renda, vendas - 1991 a 2012.....	73
Tabela 12 - Diferenças mensais entre o início das recessões na indústria com os inícios das recessões na atividade agregada brasileira - 1991 a 2012.	74
Tabela 13 - <i>Turning Points</i> dos <i>business cycles</i> da atividade agregada americana datados pelo NBER - 1959 a 2010.....	76
Tabela 14 - <i>Turning Points</i> dos <i>business cycles</i> da produção industrial americana datados pelo filtro polinomial - 1959 a 2010.....	77
Tabela 15 - Diferenças mensais entre o início das recessões na indústria e o início das recessões na atividade agregada americana - 1959 a 2012.	78
Tabela 16 - Defasagens e correlações das variáveis antecedentes com a variável <i>dummy</i> das recessões	81
Tabela 17 - Valores dos coeficientes discriminantes e da constante discriminante. .	93
Tabela 18 - Acertos e Erros das Previsões	95
Tabela 19 - Duração dos alarmes antes de cada recessão	96

Tabela 20 - Acertos e Erros das Previsões	98
Tabela 21 - Duração dos alarmes antes dos inícios de cada recessão	99
Tabela 22 - Previsão para 2013	100
Tabela 23 - Valores dos coeficientes discriminantes e da constante discriminante	102
Tabela 24 - Acertos e erros das previsões	103
Tabela 25 - Duração dos alarmes antes dos inícios de cada recessão	104
Tabela 26 - Acertos e erros das previsões	105
Tabela 27 - Duração dos alarmes antes dos inícios de cada recessão	106
Tabela 28 - Previsão para 2013	108
Tabela 29 - Recessões da Indústria de bens de capital desde 1991 até 2012 datadas pelo filtro polinomial.....	121
Tabela 30 - Valores dos Coeficientes Discriminantes e da Constante Discriminante	123
Tabela 31 - Acertos e erros das previsões	124
Tabela 32 - Acertos e erros das previsões	125
Tabela 33 - Duração dos alarmes antes dos inícios de cada recessão para a Indústria de Bens de Capital	126
Tabela 34 - <i>Turning Points</i> dos <i>business cycles</i> da produção industrial americana datados pelo filtro polinomial - 1959 a 2013	138
Tabela 35 - Defasagens e correlações das variáveis antecedentes com a variável <i>dummy</i> das recessões	141
Tabela 36 - Valores dos coeficientes discriminantes e da constante discriminante	142
Tabela 37 - Acertos e erros das previsões	143
Tabela 38 - Duração dos alarmes antes do início de cada recessão	145
Tabela 39 - Duração dos alarmes antes dos inícios das recessões da atividade agregada datadas pelo NBER.....	146
Tabela 40 - Acertos e erros das previsões	146
Tabela 41 - Duração dos alarmes antes do início de cada recessão	148
Tabela 42 - Duração dos alarmes antes dos inícios das recessões da atividade agregada datadas pelo NBER.....	148

LISTA DE FIGURAS

Figura 1 - Comportamento de um indicador antecedente hipotético.....	5
Figura 2 - Identificação dos períodos recessivos a partir da série filtrada pelo <i>gaussian kernel filter</i>	18
Figura 3 - Tipos de previsão de recessões econômicas.....	22
Figura 4 - Índice coincidente e índice antecedente para a os Estados Unidos.....	29
Figura 5 - <i>Local linear detrending</i> da série de produção de manufaturados dos Estados Unidos desde 1960 até 1996.....	34
Figura 6 - Alarmes das recessões americanas (barras pretas) e recessões americanas (barras cinzas).....	35
Figura 7 - Gráficos (a) e (b): probabilidades de recessões (linhas azuis) e recessões (áreas hachuradas) para a indústria de bens de capital. Gráficos (c) e (d): probabilidades de recessão do probit simples (linhas azuis), probabilidades de recessão do probit autorregressivo (linhas vermelhas) e a recessão de 2008 da indústria de bens de capital (linha vertical preta).....	38
Figura 8 - Probabilidades de recessão (linhas azuis) e cronologia das recessões (áreas hachuradas delimitadas com linhas vermelhas) no pós-Plano Real.....	41
Figura 9 - Probabilidades de recessão (linhas azuis) e cronologia das recessões (áreas hachuradas delimitadas com linhas vermelhas) - 1975 a 2002.....	43
Figura 10 - Representação matemática da função $g(x) = e^{\frac{-x^2}{2}}$: aproximação de uma distribuição gaussiana.....	47
Figura 11 - Representação matemática da função $g(x) = e^{-2x^2}$	47
Figura 12 - Funções matemáticas análogas a leis de potência. (a) Representação matemática da função $P(x) = x^{-2}$: distribuição de um evento semelhante a uma distribuição gaussiana (b) Representação matemática da função $P(x) = x^{-1.5}$ (c) Representação matemática da função $P(x) = x^{-1}$: evento sujeito a Lei de Zipf.....	48
Figura 13 - Comparação de uma distribuição gaussiana (linha azul) com uma distribuição seguindo uma lei de potência (linha vermelha).....	50
Figura 14 - Alarme referente aos dados da Tabela 5. O retângulo cinza representa o alarme e a barra vertical representa o início da recessão.....	59

Figura 15 - Alarme referente a uma recessão de oito meses. O retângulo cinza representa o alarme e a barra vertical representa o início da recessão.....	59
Figura 16 - Alarme referente aos dados da Tabela 5. O retângulo todo representa o alarme. A parte branca representa o erro Tipo II. A barra vertical representa o início da recessão.....	60
Figura 17 - Possíveis tipos de previsão. As barras verticais contínuas representam o início de uma recessão e a barra vertical pontilhada representa o final de uma recessão. Os quatro retângulos brancos representam erros Tipo II. O retângulo branco somado com o retângulo cinza que aparece antes do início da recessão indica um alarme que obteve sucesso na previsão. O retângulo cinza somado com o retângulo branco que aparece dentro da recessão não são considerados alarmes.	61
Figura 18 - Valores filtrados da produção industrial brasileira (linha vermelha) de 1975 a 2012; datação de seus <i>business cycles</i> (áreas hachuradas) e a datação dos <i>business cycles</i> pelo CODACE para a atividade agregada (linhas verticais).	64
Figura 19 - Dados brutos da série de produção industrial brasileira (linha vermelha) e da série do PIB brasileiro (linha preta). Os valores do eixo esquerdo são para o PIB e os do eixo direito para a produção industrial.	70
Figura 20 - Valores filtrados da produção industrial brasileira (linha vermelha) de 1991 a 2012; datação de seus <i>business cycles</i> (áreas hachuradas) e a datação dos <i>business cycles</i> pelo CODACE para a atividade agregada (linhas verticais).	71
Figura 21 - Índice coincidente composto por quatro séries: emprego, produção industrial, renda, vendas - 1991 a 2012.	73
Figura 22 - Valores filtrados da produção industrial americana (linha vermelha) de 1959 a 2010; datação de seus <i>business cycles</i> (áreas hachuradas) e a datação dos <i>business cycles</i> pelo NBER para a atividade agregada (linhas verticais).	76
Figura 23 - Índice Ibovespa em taxas de crescimento.	84
Figura 24 - Exportações de caminhões em taxas de crescimento.	85
Figura 25 - Dívida Pública Mobiliária Federal interna (DPMFi) em taxas de crescimento.	86
Figura 26 - Pontos Bovespa em taxas de crescimento.	87
Figura 27 - Produção de material elétrico para veículos em taxas de crescimento... Figura 28 - Extração de carvão mineral em taxas de crescimento.	87
Figura 29 - Taxa de câmbio comercial para a venda (R\$/US\$).	89
Figura 30 - Momentos de desaceleração da produção industrial americana.	90

Figura 31 - Momentos de desaceleração da produção industrial brasileira.	90
Figura 32 - Alarmes (barras pretas) e recessões (barras azuis), utilizando a regra de classificação através da fórmula (06).....	96
Figura 33 - Alarmes (barras pretas) e recessões (barras azuis), utilizando a regra de classificação através da fórmula (07).....	98
Figura 34 - Alarmes (barras pretas) e recessões (barras azuis), utilizando a regra de classificação através da fórmula (07).....	104
Figura 35 - Alarmes (barras pretas) e recessões (barras azuis), utilizando a regra de classificação através da fórmula (07)	106
Figura 36 - Índice antecedente das recessões brasileiras.	107
Figura 37 - Produção industrial brasileira.....	110
Figura 38 - Distribuição de frequência da série de produção industrial.....	110
Figura 39 - Evolução do expoente de Hurst para os resíduos da tendência de longo prazo da produção industrial brasileira.....	113
Figura 40 - Evolução do AR(1) antes da recessão de 1980.....	115
Figura 41 - Evolução do AR(1) antes da recessão de 1987.....	116
Figura 42 - Evolução do AR(1) antes da recessão de 1989.....	116
Figura 43 - Evolução do AR(1) antes da recessão de 1995.....	116
Figura 44 - Evolução do AR(1) antes da recessão de 1997.....	116
Figura 45 - Evolução do AR(1) antes da recessão de 2001.....	117
Figura 46 - Evolução do AR(1) antes da recessão de 2003.....	117
Figura 47 - Evolução do AR(1) antes da recessão de 2008.....	117
Figura 48 - Evolução do AR(1) antes da recessão de 2012.....	117
Figura 49 - Evolução do AR(1) da série de produção industrial em relação a sua tendência de longo prazo entre os anos de 1975 e 2012.....	119
Figura 50 - Valores filtrados da produção de bens de capital desde 1991 até 2012 e datação de seus <i>business cycles</i>	121
Figura 51 - Alarmes (barras pretas) e recessões (barras azuis) para a industria de produção de bens de capital.	125
Figura 52 - Valores filtrados da produção industrial americana (linha vermelha) de 1959 a 2013 e a datação de seus <i>business cycles</i> (áreas hachuradas).	137
Figura 53 - Momentos de desaceleração da produção industrial americana - 1959 a 2012.	139

Figura 54 - Alarmes (barras pretas) e recessões (barras azuis) para a amostra de desenvolvimento.	144
Figura 55 - Alarmes (barras pretas) e recessões (barras azuis) para a amostra de desenvolvimento.	147

LISTA DE QUADROS

Quadro 1 - Indicadores antecedentes do modelo pós-Plano Real	40
Quadro 2 - Indicadores Antecedentes do Modelo de 1975 até 2002.	42

RESUMO

MAIA FILHO, Antonio Carlos Fernandes, M.Sc., Universidade Federal de Viçosa, julho de 2013. **Previsão de recessões brasileiras com indicadores antecedentes: uma abordagem heurística.** Orientador: Newton Paulo Bueno. Coorientadores: Luciano Dias de Carvalho e Evaldo Henrique da Silva.

Neste trabalho foi empregado um método alternativo para determinar os momentos de recessão da economia brasileira, utilizando um filtro polinomial na série de produção industrial. Foram realizadas previsões para estes momentos de recessão a partir de uma análise discriminante, com a variável dependente indicada dicotomicamente representando as recessões e as variáveis explicativas em forma de taxas de crescimento, defasadas em pelo menos seis meses, chamadas de indicadores antecedentes. Estes indicadores foram escolhidos a partir da literatura já existente. O arcabouço teórico que a pesquisa está inserida tratou a recessão como sendo um evento extremo que, por sua vez, defende o uso de métodos alternativos de previsão, uma vez que as técnicas econométricas tradicionais incorrem no erro de prever séries de tempo que não são independentes e identicamente distribuídas, como a maior parte das séries econômicas. Os resultados da pesquisa indicaram ser possível prever com certa confiabilidade, com uma média de seis meses de antecedência, os momentos recessivos da economia brasileira. O poder da previsão foi testado para dados não-amostrais que abrangeram as duas últimas recessões brasileiras, de 2008 e 2011, as quais foram bem antecipadas pelo modelo de previsão heurístico aqui construído.

ABSTRACT

MAIA FILHO, Antonio Carlos Fernandes, M.Sc., Universidade Federal de Viçosa, July, 2013. **Forecast brazilian recessions with leading indicators: a heuristic approach.** Adviser: Newton Paulo Bueno. Co-advisers: Luciano Dias de Carvalho and Evaldo Henrique da Silva.

In this work, it was employed an alternative method to determine the moments of recession of the Brazilian economy, using a polynomial filter on the series of industrial production. They were made previews to these moments of recession from a discriminant analysis, with a dependent variable indicated dichotomous representing the recessions and the explanatory variables in form of growth rate, lagged on, at least, six months, called leading indicators. These indicators were chosen from the previously existent literature. The theoretical outline on which this research is based treated recession as being an extreme event, which, in its turn, advocates the use of alternative methods of previewing, since traditional econometric techniques incurs the error of previewing series of time that are not independent and identically spread. The results of the research indicate it's possible to preview, with certain reliability and with an average antecedence of six months, the recessive moments of Brazilian economy. The power of previewing was tested for out-of-sample data, which comprehend the last two Brazilian recessions, of 2008 and 2011, which were well anticipated by the heuristic model of previewing here constructed.

1 INTRODUÇÃO

1.1 Considerações Iniciais

Desde o início da década de 1990 até o final do ano de 2012 o Brasil passou por seis períodos completos de recessão, que foram: a recessão de 1995, provocada pela crise financeira mexicana, a recessão de 1998, provocada pela crise das bolsas asiáticas, a recessão de 2001, provocada pela recessão americana daquele ano, a recessão de 2003, provocada pela transição de governo, a recessão de 2008, provocada pela bolha imobiliária americana e a recessão de 2011, fruto de um processo de desindustrialização nacional.

Todas essas recessões brasileiras foram oficialmente datadas pelo Comitê de Datação de Ciclos Econômicos (CODACE, 2013)¹ vinculado a Fundação Getúlio Vargas (FGV), estabelecendo uma cronologia mensal para os momentos de recessão e expansão. A metodologia utilizada para tal datação segue a utilizada pelo Comitê de Datação norte-americano, criado pelo *National Bureau of Economic Research* (NBER, 2013). Essencialmente esta metodologia segue o procedimento sugerido por Bry e Boschan (1971), que tem como característica básica datar as recessões a partir do estudo da série do Produto Interno Bruto (PIB) em nível. Ao final do processo de datação são encontrados os momentos de picos e vales da série, ou seja, seus *turning points* e, consequentemente definidos os períodos de recessão e expansão da atividade econômica. Desta forma, obtém-se o movimento cíclico da série em nível, movimento este também conhecido como *business cycle*.

A definição aqui adotada para caracterizar uma recessão difere em alguns termos da definição de recessão formulada pelo NBER (2013). O departamento americano deixa claro que a definição comumente adotada para uma recessão, que foi elaborada por Julius Shiskin em artigo publicado no *New York Times* em 1974, é incompleta. A definição de Shiskin se resume em dizer que uma recessão é

¹ A recessão de 2011 ainda está em processo de datação pelo CODACE, o qual já indicou um período recessivo para o ano de 2011 através de uma nota em seu sítio *online* em: <http://portalibre.fgv.br>. O título da apresentação desta nota é: "Datação de séries conjunturais com algoritmos tradicionais". A recessão de 2011 foi datada pelo comitê através do Índice de Atividade Econômica do Banco Central (IBC-Br), utilizando para a datação a metodologia de Bry e Boschan (1971). No decorrer deste trabalho, essa datação será considerada como oficial para a recessão de 2011.

caracterizada após dois trimestres consecutivos de queda no PIB. Para o NBER (2013), uma recessão é "um declínio significativo na atividade econômica que se espalha em toda a economia, por alguns meses, normalmente visíveis no PIB real, renda real, emprego, produção industrial e vendas no atacado/varejo". Cabe enfatizar que o NBER não deixa claro nenhuma regra fixa ou algum tipo de modelo estatístico/matemático que guie suas decisões, nem mesmo quais séries exatamente servirão de guias para a confirmação de uma recessão.

Este trabalho propõe-se primeiramente a identificar os períodos de recessão para a economia brasileira através de uma metodologia alternativa, utilizando um filtro polinomial na série de produção industrial, que é considerada na literatura² como uma boa indicadora coincidente da atividade agregada e que aqui será nossa variável-alvo de estudo, ou seja, a *target variable*. O intuito de utilizar este método alternativo de datação é que ele será de grande utilidade para se alcançar o objetivo geral deste estudo, que é tentar prever os inícios de cada recessão brasileira desde o início da década de 1990 até o final do ano de 2012. Não será objetivo deste estudo é tentar prever os finais das recessões. A metodologia de previsão seguirá uma série de passos para então ser atingido o principal objetivo específico, que é a construção de um algoritmo para antecipar recessões em nível nacional com uma média de seis meses de antecedência.

O pilar central do método de previsão é supor que economias comportam-se de forma diferente em momentos que precedem as recessões e que esse comportamento anormal apresenta um padrão sistematicamente recorrente, possível de ser identificado e eventualmente codificado. A identificação de padrões recorrentes antes das recessões como método de previsão teve como precursor o estudo elaborado por Burns e Mitchell em 1946, os quais desenvolveram a metodologia para o sistema de indicadores antecedentes que foi utilizado pelo NBER por muitos anos.

Atualmente, o acompanhamento da atividade econômica através de indicadores antecedentes é um método tradicional bem sucedido nos Estados Unidos. O *The Conference Board* (TCB), um grupo americano privado de pesquisa econômica construiu um índice de dez indicadores considerados como antecedentes da atividade agregada americana. Entre estes indicadores, estão, por exemplo: a média dos pedidos iniciais semanais de seguro-desemprego; novas encomendas de

² Ver, por exemplo, Issler e Hollauer (2006).

bens e materiais de consumo; diferencial da taxa de juros de longo e curto prazo ou o índice de expectativas do consumidor. Naquele País coexistem hoje métodos heurísticos³ e estatístico-económéticos de construção de modelos de previsão. Exemplos do primeiro são as técnicas heurísticas do TCB, e do segundo, os trabalhos económéticos de Stock e Watson (1988, 1989, 1993a, 1993b, 2002).

Nesta mesma área de pesquisa sobre indicadores antecedentes, Reinhart e Rogoff, 2009 em livro intitulado ironicamente "*This time is Different - eight centuries of financial folly*" mostram, relatando vários estudos na área, que sempre existiram sinalizações de que grandes recessões iriam ocorrer. Estas sinalizações (entre outras formas de previsão heurísticas) não fornecem a data exata que a recessão irá se iniciar nem proporcionam uma indicação da gravidade da recessão iminente, porém, oferecem informações valiosas quanto a se a economia está apresentando um ou mais sintomas clássicos que se manifestam antes de uma recessão. A metodologia de previsão aqui apresentada seguirá os preceitos de uma abordagem heurística⁴, dado que esta tem se mostrado promissora em sinalizar a ocorrência de recessões.

O primeiro passo a ser feito é datar os momentos recessivos da *target variable* utilizando o *gaussian kernel filter*, que é um filtro capaz de captar com eficiência os *turning points* de uma série de tempo. Os momentos em que os valores filtrados da série apresentarem uma trajetória de queda serão os momentos das recessões. Para facilitar a datação, os valores filtrados serão transformados em taxas de crescimento⁵ e quando essa taxa for negativa significa que a trajetória de queda está ocorrendo. Desta forma, por este método, define-se uma recessão o período, maior ou igual a seis meses, em que as taxas de crescimento da série de produção industrial filtrada apresentem-se negativas.

³ Entende-se por heurística um conjunto de regras e métodos que visam à descoberta, à invenção ou à resolução de problemas. É um procedimento simplificador (embora não simplista) que, em face de problemas difíceis envolve a substituição de resoluções complexas por outras de caráter mais intuitivo a fim de encontrar respostas viáveis (KAHNEMAN, 2012).

⁴ Mais especificamente, a metodologia segue um subgrupo dentro do grupo dos modelos heurísticos, que é o subgrupo que trata as recessões como sendo um evento extremo. Porém, como será visto no decorrer do trabalho, o método de previsão aqui adotado pode-se dizer um modelo híbrido entre a abordagem heurística e a abordagem estatístico-económética, com uma maior influência dos modelos heurísticos.

⁵ Já adiantando, a forma de cálculo que será um procedimento básico adotado em diversas séries desta pesquisa, a fórmula para o cálculo da taxa de crescimento é:

$$\left[\left(\frac{\text{valor em } t+1}{\text{Valor em } t} \right) - 1 \right] * 100.$$

O segundo passo consiste em identificar séries que apresentem precocemente o mesmo comportamento da *target variable*. Ou seja, se determinada série apresenta seus respectivos *turning points* meses antes dos *turning points* da *target variable* ela será um bom indicador antecedente. Na literatura já existem algumas pesquisas que estudaram alguns indicadores antecedentes tanto para atividade agregada como para a produção industrial⁶, as quais apontaram bons indicadores antecedentes para ambas as séries. Dentre alguns destes indicadores estão, por exemplo, a taxa de câmbio, a taxa de juros real de curto prazo, o índice Bovespa e a taxa crescimento da economia americana. Deste modo, boa parte das séries aqui usadas foi identificada a partir de trabalhos já realizados sobre este tema.

O terceiro passo consiste em testar se as variáveis apontadas na literatura apresentam um comportamento antecedente em relação aos *turning points* da *target variable* deste estudo. Para isso, todas as séries indicadas na literatura como antecedentes foram suavizadas pelo mesmo filtro utilizado na *target variable* e posteriormente foram calculadas suas taxas de crescimento. A partir daí defasou-se em doze⁷ meses as séries de taxa de crescimento de cada variável e testou-se a correlação desses valores com os valores da *target variable*. No teste de correlação a *target variable* foi transformada numa série binária, na qual os zeros representaram as recessões e os valores unitários os momentos de expansão. O motivo é que, na análise de previsão os valores a serem previstos são dicotômicos, ou seja, será prevista a ocorrência ou não de uma recessão. As variáveis que apresentaram correlação significativa com pelo menos seis meses de defasagem foram selecionadas como boas indicadoras antecedentes⁸, sendo, portanto, candidatas a comporem o algoritmo de previsão. A Figura 1 abaixo, inspirada em artigo de Zarnowitz e Moore (1982), exemplifica o comportamento de um indicador antecedente hipotético da produção industrial e através dela será introduzido o conceito do comportamento normal do indicador antecedente.

⁶ Ver Chaveut e Silva (2004); Issler e Hollauer (2006); Chaveut e Morais (2009).

⁷ Cada variável foi defasada em até doze meses dado que o objetivo é uma previsão de curto prazo. Dadas estas defasagens, a melhor defasagem será escolhida a partir do teste de correlação com a *target variable*.

⁸ A defasagem calculada para cada indicador antecedente será chamada defasagem ótima.

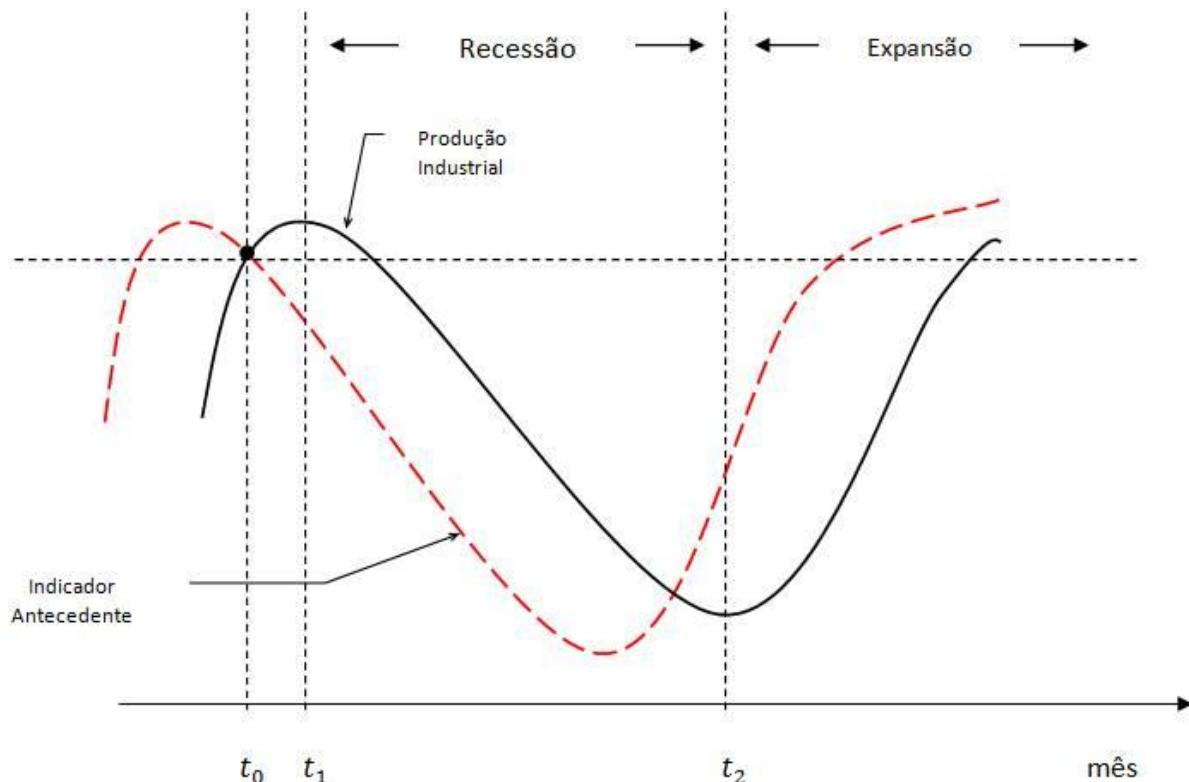


Figura 1 - Comportamento de um indicador antecedente hipotético.

As duas curvas representam os valores filtrados de cada variável. Desse modo, quando a curva da produção industrial começa a apresentar uma trajetória de queda, ou seja, no mês seguinte ao mês de pico t_1 , tem início uma recessão que se estende até t_2 , que é o próximo vale seguinte ao momento de pico. A partir dele, no mês seguinte, tem início um período de expansão. O ponto localizado sobre a linha vertical do tempo t_0 , em que o indicador antecedente começa a cair abaixo do patamar assinalado pela linha tracejada horizontal, indica uma sinalização que a produção industrial poderá entrar em recessão em pelo menos doze meses. A linha tracejada horizontal representa, portanto, o limite do comportamento normal do indicador antecedente⁹.

Kaminsky, Lizondo e Reinhart (1998) fazem um excelente estudo classificando determinados indicadores antecedentes através da abordagem ruído/sinal. Essa abordagem consiste basicamente em determinar o ponto limite do comportamento normal de cada indicador antecedente e verificar se, quando o indicador apresenta um comportamento fora do normal, existirá uma crise em no

⁹ A figura acima mostra os valores filtrados de ambas as séries para fins de melhor ilustração e entendimento. Quando for feito o estudo do limite do comportamento normal do indicador antecedente serão usadas as séries em taxa de crescimento dos valores filtrados.

máximo doze meses. Quando o indicador envia um sinal e eclode uma recessão no ano seguinte, o sinal é exato; se nenhuma recessão se segue ao final, trata-se de um alarme falso, ou ruído. Assim, os melhores indicadores são aqueles com os mais baixos índices de ruído/sinal. É de se esperar que os indicadores antecedentes que melhor se correlacionaram com a *target variable* apresentem os mais baixos índices de ruído/sinal.

Como veremos, o próximo passo, que utilizará o método da análise discriminante através do *software* estatístico SPSS®, fará a seleção final dentre todos os indicadores antecedentes escolhidos no passo três e escolherá, através do método *stepwise* o melhor conjunto de indicadores com poder de discriminar se existirá ou não uma recessão. Basicamente este método escolherá os indicadores que apresentem baixos níveis de ruído/sinal, dado que estes possuirão melhor poder discriminador. A análise ruído/sinal serve para ilustrar o procedimento do limite do comportamento normal para indicadores observados individualmente. Como esta pesquisa está interessada na construção de um algoritmo que combine diversos indicadores antecedentes, será construído um limite do comportamento normal para os valores assumidos por este algoritmo, limite que será chamado de regra de classificação e que no decorrer do trabalho será mais bem explicado¹⁰. Mesmo sabendo que a análise dos indicadores se dará de forma conjunta, será feita uma análise individual de cada indicador antecedente depois que forem escolhidos pelo método *stepwise* da análise discriminante, ou seja, uma análise *ex-post* dos indicadores.

O quarto passo consiste, dessa forma, em utilizar todos os indicadores antecedentes encontrados no passo três e testá-los através de uma análise discriminante, que, por via do método *stepwise*, serão escolhidos os indicadores antecedentes capazes de discriminar com maior eficiência se determinado mês futuro será ou não de recessão. Esta análise discriminante baseia-se essencialmente no trabalho de Altman (1968) e tem como objetivo principal atribuir determinados pesos a cada indicador escolhido como bom discriminador. Para cada mês, a soma dos valores de todos os indicadores escolhidos multiplicados por seus respectivos pesos mais o valor da constante calculado pelo *software* formam a

¹⁰ Na metodologia deste estudo será especificada formalmente como esta regra de classificação é construída para determinar o limite do comportamento normal do valor que o algoritmo assume. Serão mostradas duas regras de classificação, que são concorrentes e no decorrer do trabalho uma delas será escolhida para ser aplicada no modelo. Além do mais, nesta própria introdução, será relatado o procedimento intuitivo de uma destas regras de classificação.

função discriminante ou, como será aqui chamado, o algoritmo de previsão. Com esse algoritmo são calculados os escores discriminantes Z_t para cada mês, que serão comparados com uma regra de classificação. A intuição da regra de classificação trata de comparar o valor de cada escore discriminante de cada mês com o valor do escore médio do grupo dos zeros (recessões) e com o valor do escore médio do grupo dos uns (não-recessões) e classificá-lo de acordo com sua proximidade. Ou seja, se determinado escore se encontra mais próximo da média do grupo dos zeros do que da média do grupo dos uns ele será classificado como um mês de recessão. A diferença principal de analisar os indicadores de forma conjunta ao invés de individualmente é que aqui os indicadores se unem para formar um índice composto por vários indicadores. A essência é que este índice consiga suavizar algumas volatilidades dos indicadores individuais, sendo assim mais convincente e claro.

Por fim, mas não menos importante, a previsão será feita com dados não-amostrais. Existem basicamente dois métodos de realizar previsão. O primeiro utiliza toda a informação disponível da amostra para elaborar o modelo e então analisa o seu poder preditivo em relação às observações dentro da própria amostra. O segundo utiliza informações de uma parte da amostra, em torno de 70% das observações, para estimar o modelo e então realiza as previsões para o restante da amostra, os 30% finais. A primeira metodologia, como considera toda a informação disponível para estimar o modelo, apresenta maior potência. Porém, as previsões fora da amostra possuem maior aplicação prática, pois se aproximam melhor das restrições de informações que observamos na vida real. Será assim o segundo tipo de previsão, com dados não-amostrais, utilizado para testar o poder preditivo do algoritmo de previsão aqui construído.

1.2 O Problema e sua Importância

O problema aqui levantado pode ser posto da seguinte forma: economias comportam-se diferente em momentos que precedem uma recessão? A resposta a essa pergunta nos levará a possibilidade de codificar esse comportamento anormal para assim conseguirmos antecipar em alguns meses o início de uma recessão.

A justificativa para tal intento é bastante clara. É de grande importância a previsão de recessões econômicas. Monitorar e antecipar alterações no estado geral da economia, com foco nos *turning points* é de grande interesse por empresários, *policy-makers* e agentes econômicos em busca de emprego. Especificamente aos *policy-makers*, cuja tarefa de efetuar intervenções corretas na economia é bastante dificultosa. Essa dificuldade decorre principalmente das defasagens internas e externas das políticas fiscais e monetárias. Ou seja, leva-se tempo entre a implementação e os efeitos de uma política econômica. Assim, políticas de estabilização que são adotadas quando ocorre uma recessão podem gerar efeitos contrários dos pretendidos, já que o resultado da política pode vir a se realizar apenas quando a recessão já se findou, gerando assim, por exemplo, um processo inflacionário.

Uma política de estabilização bem sucedida exige a capacidade de prever com precisão as condições econômicas no futuro. Caso não seja possível a previsão de se a economia estará em recessão ou expansão daqui a seis meses ou um ano, não se pode avaliar se a política monetária e fiscal deve agora tentar expandir ou contrair a demanda agregada. Na literatura nacional, apesar de grande esforço de pesquisas recentes (Contador e Ferraz, 1977; Picchetti e Toledo, 2002; Carvalho e Hermanny, 2003; Chauvet e Silva, 2004; Lima, Moro e Júnior, 2006; Issler e Hollauer, 2006; Campelo Júnior, 2008; Chauvet e Morais, 2009; Issler et al., 2013), a previsão de recessões somente despertou atenção após o fim do processo inflacionário, de forma que os estudos aplicados ainda são escassos e a avaliação das diferentes técnicas de previsão ainda é preliminar. Os trabalhos destacados em itálico são os que têm como objetivo prever recessões. Os demais então inseridos sobre o tema, construindo índices coincidentes de atividade agregada ou estudando

bons indicadores antecedentes para as recessões, porém, não realizando a previsão de fato.

Contador e Ferraz (1977) se basearam na metodologia dos indicadores antecedentes do NBER e utilizaram várias outras metodologias, como a análise de componentes principais para a construção do índice antecedente, previsor da atividade agregada. Na época de sua publicação os resultados foram considerados satisfatórios. Em trabalho posterior, Contador e Ferraz (2001) afirmaram que até o ano de 1968 pouca atenção era dada ao acompanhamento dos ciclos econômicos brasileiros. Eles ressaltam que só a partir daquele ano, devido a um trabalho pioneiro da Fundação Getúlio Vargas, através do centro de Estatística e Análise Econômica, do Instituto Brasileiro de Economia (IBRE), é que começaram a serem realizadas pesquisas conjunturais junto a empresas, sobretudo sobre o nível de estoques e expectativas. A esta pesquisa deu-se o nome de "Sondagem Conjuntural", que é regularmente publicada na *Conjuntura Econômica*¹¹.

Carvalho e Hermanny (2003) construíram um sistema de indicadores antecedentes para o Brasil. Entre 1980 e 1999 os autores identificaram trinta *turning points* na série do PIB e o indicador antecedente teria falhado em 14 destes pontos. Nos outros 16 casos o indicador conseguiu antecipar os movimentos da série do PIB com uma média de três meses de antecedência. A conclusão dos autores foi que, mesmo com a alta volatilidade das séries econômicas brasileiras, existem sinais promissores que modelos de previsão com indicadores antecedentes podem ser construídos e utilizados com certo nível de confiabilidade.

Chauvet e Silva (2004), utilizando uma variável latente binária para determinar o estado da economia em momentos de expansão e recessão, conseguiram, através de indicadores antecedentes e o uso de um modelo probit, construir previsões com horizontes de curto e médio prazo (que serão mais bem especificados no decorrer do trabalho). Os autores identificaram dez períodos de recessão para o período entre 1975 e 2002. Destes dez períodos os modelos preditivos conseguiram antecipar, com uma antecedência variando de cinco a oitos meses, quatro períodos recessivos de forma clara. Foi feita uma análise com dados não-amostrais, que

¹¹ Os dados históricos destas pesquisas são apenas disponibilizados através de pagamento de mensalidade, o que impossibilitou seu uso neste estudo. Como já demonstrado por Campelo Junior (2008), algumas das séries calculadas pela Sondagem Conjuntural são bons indicadores antecedentes para o nível de atividade agregada, servindo assim como possibilidade futura para serem utilizadas através da metodologia aqui proposta.

mostrou-se menos robusta que os dados amostrais, porém ainda conseguindo prever quatro das dez recessões identificadas pelos autores.

Lima, Moro e Júnior (2006), utilizando a metodologia de componentes principais, constrói um índice antecedente para a economia brasileira para captar os pontos de reversão da série do PIB com até três meses de antecedência. A variável do PIB foi prevista em valores quantitativos. Calculados os fatores que melhor explicavam o PIB, os autores estimaram uma regressão linear, baseando-se no trabalho de Stock e Watson (2002) e encontraram um R^2 de 0,48 para os dados amostrais. Em relação aos *turning points*, 17% deles foram previstos de forma incorreta, indicando um falso sinal ou um erro de omissão. Apesar de os autores não deixarem claro as recessões identificadas, nem exatamente com quantos meses foram previstas, concluíram que existem sinais promissores de que modelos de indicadores cíclicos possam ser construídos com certa confiabilidade para o Brasil.

Chauvet e Morais (2009) dataram as recessões na série de bens de capital brasileira utilizando o mesmo algoritmo utilizado pelo NBER e pelo CODACE, o algoritmo de Bry e Boschan (1971). Através dele foram identificados cinco períodos recessivos entre 1990 e 2009. Assim, os autores transformaram a variável a ser prevista numa variável dicotômica identificando os períodos de recessão e expansão. Para prever estes estados da economia os autores utilizaram seis indicadores antecedentes inseridos em um modelo probit dinâmico autorregressivo. O modelo foi testado para dados amostrais e não-amostrais. No caso dos dados amostrais o modelo obteve uma previsão média com seis meses de antecedência para quatro das cinco recessões identificadas. No caso da previsão para os dados não-amostrais, que foi a recessão de 2008, o modelo conseguiu antecipá-la em cinco meses. Dentre todos os modelos de previsão aplicados ao Brasil, este foi o que se mostrou com maior capacidade preditiva, mas que, porém, se restringiu a prever recessões na indústria de bens de capital, a qual os autores não fizeram demonstração de ser uma boa previsora da atividade agregada.

Relatado tais estudos, é notório que, apesar do esforço destes autores, ainda não foi construído um modelo de previsão cíclica da atividade agregada brasileira que conseguisse prever com certa antecedência as recessões brasileiras. Soma-se a isto o fato que ainda não foi publicado nenhum trabalho de previsão cíclica que se baseasse na datação das recessões estabelecidas pelo CODACE (2013) para a

atividade agregada, que este trabalho também se propõe a fazer. Outro ponto é que, no Brasil, os modelos de previsão elaborados até o presente momento seguem essencialmente uma modelagem estatístico-econométrica, não existindo assim uma abordagem heurística de previsão de recessões, que, como será visto no decorrer deste estudo, tem se mostrado eficiente nas previsões de recessões americanas¹²

Não existindo, no Brasil, métodos heurísticos de previsão de recessões, não existe a abordagem do problema de previsão focando numa crítica pertinente aos modelos econométricos tradicionais, que é a crítica de que as distribuições dos valores do PIB e de suas variáveis explicativas geralmente não seguem uma distribuição normal, não sendo portanto independentes e identicamente distribuídos (i.i.d.), o que é uma pressuposição básica para se poder construir e aplicar os modelos econométricos tradicionais. Assim, como as recessões mais severas tendem a se localizar fora da abrangência de uma distribuição normal, alguns grupos de autores, entre eles os *econofísicos*, consideram o estudo das recessões inserido na teoria dos eventos extremos, teoria esta que caracteriza um evento extremo como sendo raro e de grande impacto¹³.

Assim, elaborar um modelo preditivo com base em modelos heurísticos de previsão e assimilando técnicas estatístico-econométrica mais simples, como as que serão aqui apresentadas (análise de correlação de Spearman e análise discriminante, que não pressupõe distribuição normal das variáveis) é pertinente para a economia brasileira, dado a carência de trabalhos na área e principalmente ao fato de não existir na literatura um estudo que tenha conseguido prever com clareza, com pelo menos seis meses de antecedência as recessões brasileiras.

Neste sentido, a hipótese central a ser testada é que, como já citado, economias comportam-se diferente em momentos que precedem as recessões e que esse comportamento anormal apresenta um padrão sistematicamente recorrente, possível de ser identificado e eventualmente codificado. Definida a hipótese, seguem os objetivos desta pesquisa.

¹² Ver TCB (2000) e Keilis-Borok et al. (2000).

¹³ Ver Meyers (2009).

1.3 Objetivos

1.3.1 Objetivo Geral

Prever o início de cada recessão brasileira, desde o início da década de 1990 até o final do ano de 2012.

1.3.2 Objetivos Específicos

- a) Datar os momentos recessivos da economia brasileira entre 1991 e 2012 a partir de um método alternativo, utilizando como *proxy* da atividade agregada a série de produção industrial;
- b) Comparar as recessões datadas pelo método alternativo proposto neste estudo com as recessões brasileiras datadas pelo CODACE;
- c) Construir um algoritmo com indicadores antecedentes para antecipar recessões em nível nacional com uma média de seis meses de antecedência, entre os anos de 1991 e 2012;
- d) Analisar de forma particular os indicadores antecedentes que fizeram parte do algoritmo preditivo elaborado no objetivo anterior;
- e) Analisar se as recessões brasileiras enquadram-se no escopo teórico dos eventos extremos;
- f) Comparar em termos de resultados o método de previsão construído nesse estudo com o método elaborado por Chauvet e Morais (2009) para prever recessões na indústria de bens de capital.

1.4 Escopo do Trabalho

O presente estudo será dividido em seis capítulos, desconsiderando-se esta introdução. O Capítulo 2 irá apresentar uma visão pragmática do que é uma recessão, inserida nos estudos dos *business cycles*. O Capítulo 3 irá mostrar como é possível prever recessões com indicadores antecedentes e comparar as duas principais metodologias de previsão inseridas nesta temática. O Capítulo 4 irá procurar explicar estatisticamente como se comporta

um evento extremo, para posteriormente analisar se as recessões brasileiras podem ser inseridas neste tipo de evento. O Capítulo 5, da metodologia, irá procurar mostrar como, através da análise discriminante, é possível antecipar as recessões brasileiras. Por fim, o Capítulo 6 mostrará os resultados da pesquisa e o Capítulo 7 apresentará as conclusões.

Entrando nos detalhes do Capítulo 6, ele será dividido em seis seções. As duas primeiras, 6.1 e 6.2, buscarão justificar o método alternativo de datação de recessões aplicado neste estudo, bem como determinar os alvos de previsão. A seção 6.3 irá mostrar o processo de escolha dos indicadores antecedentes para fazer parte do modelo preditivo. A seção 6.4 mostrará os principais resultados do trabalho, ou seja, todas as previsões analisadas. A seção 6.5 vai procurar analisar estatisticamente se as recessões brasileiras podem ser estudadas no arcabouço teórico dos eventos extremos. Por fim, a seção 6.6 irá fazer uma comparação do modelo preditivo construído neste estudo com um modelo preditivo elaborado por Chauvet e Moraes (2009).

2 ESTUDO DOS CICLOS ECONÔMICOS E A DEFINIÇÃO DE RECESSÃO

Antes de definir o que seja uma recessão econômica, é necessário diferenciar os dois tipos de estudo de ciclos econômicos. O primeiro estudo, chamado de *business cycle*, estuda o agregado econômico em nível e assim pode ser definido como: flutuações recorrentes e alternantes de expansão e recessão que caracterizam as séries temporais medidas em nível da maioria dos agregados econômicos. Nesse caso, os *turning points* da série definem os períodos de crescimento e decrescimento absoluto da variável. Esta abordagem foi criada pelo NBER em 1920, data de sua fundação, e continua sendo uma referência.

O segundo estudo, chamado de *growth cycle*, estuda o agregado econômico em desvios em relação à tendência de longo prazo ou em taxas de crescimento da série filtrada de sua tendência de longo prazo (ZARNOWITZ e OZYILDIRIM, 2006) e tem a Organização para Cooperação Econômica e Desenvolvimento (OECD) a frente das pesquisas. Nesse tipo de ciclo, um decrescimento na série sinaliza um declínio na taxa de crescimento, mas não necessariamente uma recessão. Como todas as recessões são precedidas por momentos de desaceleração, os *growth cycles* ocorrem com maior frequência e tendem a ser mais simétricos do que os *business cycles*. Aqui a dicotomia do movimento é expressa em duas palavras: aceleração e desaceleração.

Nos *business cycles* a dicotomia do movimento é expressa em: expansão e recessão. O NBER, que atualmente é o datador oficial dos *business cycles* americano, fez a seguinte declaração sobre a conceituação de uma recessão:

The NBER does not define a recession in terms of two consecutive quarters of decline in real GDP. Rather, a recession is a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales (NBER, 2013).

Ou seja, a definição mais disseminada, formulada por Julius Shiskin em artigo no *The New York Times* no ano de 1974, de que uma recessão é

caracterizada após dois trimestres consecutivos de queda no PIB, é incompleta. A decisão do NBER em definir um período como recessivo baseia-se na análise de quatro importantes agregados econômicos: emprego, renda pessoal, produção industrial e vendas dos setores industrial e atacadista. O conjunto desses quatro indicadores forma o que se chama na literatura de índice coincidente da atividade agregada. É através deste índice que são confirmadas as datações dos momentos de expansão e recessão da economia observados no PIB real.

Essencialmente a metodologia de datação do NBER segue o procedimento sugerido por Bry e Boschan (1971), que tem como característica básica datar as recessões a partir do estudo da série em nível. Ao final do processo de datação, são encontrados os momentos de picos e vales da série, ou seja, seus *turning points* e, consequentemente, são definidos os períodos de recessão e expansão da atividade econômica. Cabe enfatizar que o NBER não deixa claro nenhuma regra fixa ou algum tipo de modelo estatístico-matemático que guie suas decisões, nem mesmo quais séries exatamente servirão de guias para a confirmação de uma recessão.

No Brasil, em documento lançado em maio de 2009, o Comitê de Datação de Ciclos Econômicos (CODACE, 2013) identificou oito períodos de recessões na economia brasileira entre 1980 e 2009, seguindo o modelo adotado pelo NBER. A datação do CODACE (2013) considerou como recessão os períodos em que houve expressivo declínio no nível de atividade econômica espalhado por diversos segmentos da economia brasileira, perceptível durante ao menos dois trimestres consecutivos. A principal variável usada na datação pelo CODACE (2013) foi PIB trimestral, dessazonalizado, calculado pelo IBGE. Segundo o comitê, a periodicidade trimestral escolhida para a primeira datação decorreu da escassez de estatísticas de qualidade contínuas e com periodicidade mensal no País.

Em documento lançado em fevereiro de 2010, o CODACE (2013) estabeleceu uma cronologia mensal de recessões para a economia brasileira entre 1980 e 2009. Os integrantes do comitê basearam a datação mensal na análise da evolução das séries econômicas mensais que melhor retratavam, a cada momento, o estado da produção industrial, vendas no comércio, emprego e rendimento do trabalho. Os *turning points* da datação mensal são

coincidentes ou situam-se muito próximos aos *turning points* da cronologia trimestral calculada na reunião anterior. O descasamento eventual decorre da própria mudança de frequência das séries usadas em cada datação. Ambas as datações, trimestral e mensal, seguiram os mesmos procedimentos adotados pelo NBER, ou seja, utilizando o algoritmo sugerido por Bry e Boschan (1971).

Em documento lançado em março de 2013, o CODACE (2013) estabeleceu diversas datações de séries conjunturais, utilizando-se do algoritmo de Bry e Boschan. Entre estas séries, está o Índice de Atividade Econômica do Banco Central (IBC-Br), que, através da datação, foram encontrados meses de recessão para o ano de 2011. Essa recessão de 2011, como já afirmado em nota de rodapé na introdução, será somada às oito recessões encontradas pelo CODACE (2013), totalizando assim nove recessões desde 1980 até o ano de 2012¹⁴. As datações podem ser verificadas através da Tabela 1 abaixo, que mostra os picos e vales das recessões, bem como a duração de cada recessão.

¹⁴ Com a intenção de justificar o uso da datação do IBC-Br na recessão de número nove para suprir o atraso que o CODACE está tendo em datar esta recessão oficialmente, comparou-se os *turning points* da recessão de 2008 do IBC-Br com os *turning points* da recessão de 2008 (recessão nº 8) da datação oficial. Como pode ser visto na Tabela 1, a recessão nº 8 da datação oficial tem seu pico no mês de julho e o vale no mês de janeiro de 2009. Já o momento de pico da datação com o IBC-Br é em junho e o vale no mês de dezembro de 2008. Ou seja, as datações são semelhantes, diferindo apenas em um mês no início e um mês no final da recessão.

Tabela 1 - *Turning points* dos *business cycles* da atividade agregada brasileira datados pelo CODACE¹⁵ - 1980 a 2012

Recessão	Picos	Vales	Meses em recessão
1	Outubro/1980 (IV)	Fevereiro/1983 (I)	28
2	Fevereiro/1987 (II)	Outubro/1988 (IV)	20
3	Junho/1989 (II)	Dezembro/1991 (I/92)	30
4	Dezembro/1994 (I/95)	Setembro/1995 (III)	9
5	Outubro/1997 (IV)	Fevereiro/1999 (I)	16
6	Dezembro/2000 (I/01)	Setembro/2001 (IV)	9
7	Outubro/2002 (IV)	Junho/2003 (II)	8
8	Julho/2008 (III)	Janeiro/2009 (I)	6
9	Março/2011 (I)	Outubro/2011 (IV)	7
Duração total de meses			133

Fonte: CODACE (2013).

Dada tais informações, este presente estudo se propôs a identificar os *business cycles* para a economia brasileira de um modo diferente do proposto pelo NBER ou pelo CODACE. As diferenças estão centradas em duas especificidades.

A primeira delas é que a identificação dos *turning points* realizou-se através do estudo da série filtrada e não em nível. O filtro utilizado para a identificação foi o *gaussian kernel filter*¹⁶. A justificativa de utilização reside no fato de que este filtro é construído especificamente para captação de *turning points* (MARSH e CORMIER, 2002). Com esse filtro, os momentos em que os valores filtrados da série apresentarem uma trajetória de queda serão os momentos das recessões. Dois trabalhos serviram de base para a adoção desta metodologia de datação.

¹⁵ Os trimestres estabelecidos na datação trimestral são apresentados entre parênteses.

¹⁶ Este filtro é facilmente disponibilizado pelos *softwares* estatísticos. Trata-se de uma regressão polinomial local, que é uma abordagem não paramétrica para análise de regressão, na qual a estimativa é feita em cada ponto aplicando a função de regressão linear a certa quantidade de observações vizinhas ao ponto. Nesta pesquisa utilizou-se o *software* Stata®. Especificamente, o comando segue os seguintes passos: *Graphics/Smoothing and Densities/Local Polynomial Smoothing*. A função *kernel* é: *gaussian* e o grau do polinômio é 1. Estas especificidades seguiram o mesmo procedimento de suavização adotado por Dakos et al. (2008) em trabalho para detectar momentos de mudança em séries temporais climáticas.

O primeiro foi um artigo elaborado por Keilis-Borok (2005) para identificar momentos de aceleração e desaceleração da taxa de desemprego para a economia norte-americana. Os momentos de aceleração e desaceleração foram distinguidos após a filtragem da série de desemprego, filtragem este que segundo os autores teve o papel de retirar as sazonalidades da série, com o intuito de facilitar o processo de delimitação dos momentos de aceleração e desaceleração. Os momentos em que a série dos valores filtrados apresentavam-se em queda foram identificados como de desaceleração e quando os valores filtrados apresentavam-se em ascensão foram identificados como momentos de aceleração. O segundo trabalho, como explicado na nota de rodapé de número 11, é um artigo elaborado por Dakos et al. (2008).

Para facilitar a datação das recessões, os valores filtrados serão transformados em taxas de crescimento e quando essa taxa for negativa significa que a trajetória de queda está ocorrendo. Desta forma, por este método, define-se uma recessão o período, maior ou igual a seis meses, em que as taxas de crescimento da série de produção industrial filtrada apresentem-se negativas. A Figura 2 abaixo faz uma exemplificação do que é considerado um período de recessão ou de expansão utilizando os valores filtrados da variável a que se deseja datar seus *business cycles*.

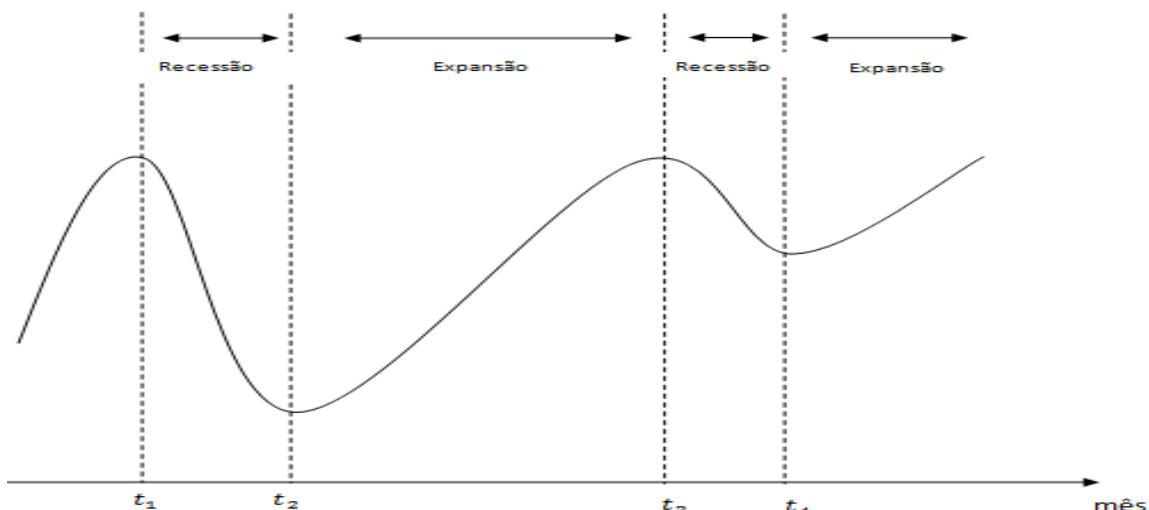


Figura 2 - Identificação dos períodos recessivos a partir da série filtrada pelo *gaussian kernel filter*.

Como pode ser observado, os períodos que são delimitados como recessivos correspondem aos períodos em que os valores da série filtrada apresentaram uma trajetória de queda, ou seja, no mês seguinte ao mês de pico t_1 até o vale t_2 e no mês seguinte ao mês de pico t_3 até o vale t_4 .

A justificativa para usar um filtro como datador dos momentos de recessão em detrimento de métodos mais tradicionais já consolidados na literatura será exposta nos resultados desta pesquisa, especificamente no final da seção 6.2.

A segunda especificidade que diferencia o método de datação aqui proposto do método adotado pelo CODACE (2013) é que a série de referência para datação não leva em conta as vendas no comércio o nível de emprego, o rendimento do trabalho, nem a variável principal de datação, que é o PIB. Apenas faz-se uso da série de produção industrial, que como será visto nos resultados deste estudo, seções 6.1 e 6.2, é uma boa *proxy* tanto para o PIB como para o índice das quatro séries que formam o índice coincidente da atividade agregada.

Por fim, cabe ainda nesta seção apresentar uma contribuição que este método alternativo de datação oferece, especificamente no sentido de utilizar o *gaussian kernel filter* como datador de recessões em tempo real.

Hamilton (2011) apontou um problema nas datações do NBER para a economia americana, que reside no fato de que instituto não faz a datação em tempo real, ou seja, a datação só é lançada quando meses já se passaram do final da recessão. Este fato já vem ocorrendo também no Brasil através do CODACE, dado que para a recessão de 2011 ele ainda não se pronunciou oficialmente.

Ainda, segundo Hamilton (2011), uma possível mecanização do processo de datação traria três importantes benefícios. O primeiro é que diminuiria a morosidade do lançamento do momento em que a recessão se iniciou. O segundo é que, caso se consiga obter um meio de datação claro e objetivo que possa ser implementado em tempo real, esse processo de datação se livraria de possíveis influências políticas, dado que, por vezes, a liberação da informação que a economia se encontra em recessão não seja de grande interesse pela classe política.

O terceiro motivo da mecanização do processo de datação é que ele pode ajudar a esclarecer exatamente como e por que uma recessão é datada. Ou seja, poder-se-ia fechar com mais elucidação o conceito sobre o que realmente queremos dizer quando uma economia está em recessão. Ou seja, o caminho seria tornar o processo de datação o mais claro e acessível possível.

Desta forma, o método de datação utilizando o *gaussian kernel filter* permite uma razoável mecanização do processo de datação. Outras formas de suavização, como o filtro HP e a regressão *spline*, não obtiveram resultados tão bons quanto. Porém, sua implementação em tempo real é possível com uma ressalva, que será explicada a seguir, relatando o processo de datação.

Dado que a série alvo a qual se deseja datar as recessões seja a da produção industrial disponibilizada mensalmente. Quando surgir uma nova observação, o procedimento é aplicar o *gaussian kernel filter* em toda a série. Se o valor filtrado desta ultima observação for menor do que o valor filtrado da penúltima observação significa que a série de produção industrial entrou em momento recessivo, e é neste fato que reside a ressalva anteriormente colocada.

O fato é que, dado que existiu essa queda, para se poder afirmar que está ocorrendo uma recessão em tempo real seria contrassenso esperar que essa queda se desse por um período mínimo de seis meses, o que de encontro com a definição de recessão aqui apresentada. O que se pode afirmar com base nos resultados deste trabalho (seções 6.1 e 6.2), é que quando as séries dos valores filtrados passaram a entrar em declive, essa queda sempre durou mais que seis meses, isso tanto no processo de datação das recessões brasileiras como americanas.

3 PREVISÃO COM INDICADORES ANTECEDENTES

O objetivo inicial desta seção é diferenciar os tipos de previsão de recessões antes e depois dos estudos de Stock e Watson (1988; 1989; 1993), que passaram a focalizar as recessões a partir de modelos de mudança de regime, ou seja, passaram a prever, através de indicadores antecedentes, o estado da economia através de uma variável latente dicotômica que indicava os momentos de recessão e expansão. Objetiva também diferenciar os dois tipos de métodos de previsão que coexistem hoje nos Estados Unidos utilizando uma variável latente dicotômica da atividade agregada combinada com indicadores antecedentes, que são: os métodos heurísticos e os métodos estatístico-econométricos.

Dadas as diferenças entre os dois métodos, será feita uma revisão de quatro trabalhos que utilizaram destes métodos para prever recessões e que serviram de base para elaboração deste presente estudo. Os dois primeiros se enquadram no método heurístico de previsão, que são: o sistema de indicadores antecedentes elaborado pelo *The Conference Board* (TCB) e um modelo que trata as recessões como sendo um evento extremo, elaborado por Keilis-Borok et al. (2000). Ambos são previsões de recessões para a economia Norte-americana. Os outros dois estudos se enquadram no método estatístico-econométrico, que são: Chauvet e Morais (2009) e Chauvet e Silva (2004). Ambos são previsões para a economia brasileira, sendo que o primeiro trata de prever recessões na indústria de bens de capital e o segundo de prever recessões na atividade agregada. A Figura 3 abaixo ajuda a sintetizar os tipos de previsão de recessão que serão aqui discutidos:

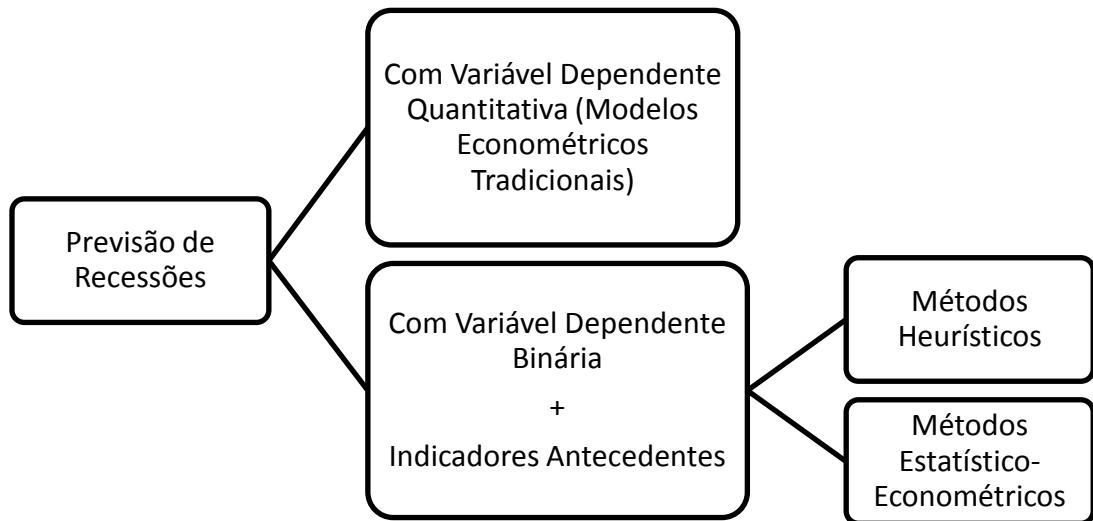


Figura 3 - Tipos de previsão de recessões econômicas.

Antes de diferenciar o marco dos estudos de previsão, estabelecido através dos estudos de Stock e Watson (1988; 1989; 1993) e que se concretiza no esquema acima retratado como "com variável dependente binária + indicadores antecedentes", cabe uma breve descrição da influência que estes autores sofreram para chegar a tal fim. Essa influência foi advinda do NBER que inaugurou, por volta de 1930 através de Burns e Mitchell, uma forma de previsão não usual, que tratava a variável agregada de forma cíclica, através das datações dos momentos de recessão e expansão. Essa forma de enxergar a variável do PIB como uma variável latente fez e faz parte do método de previsão com mudança de regime. Neste tipo de método, a atenção é voltada apenas para o comportamento cíclico da série, através de seus *turning points*. Em outras palavras, este método procurou monitorar e antecipar alterações no estado geral da economia, com foco nos pontos de reversão. Lahiri e Moore (1991) afirmaram que a ênfase nos momentos de reversão seria o motivo do grande sucesso destes modelos, uma vez que sinais antecipados dos momentos de expansão e recessão são de grande interesse por empresários, *policy-makers* e agentes econômicos em busca de emprego.

Concomitantemente e anteriormente a essa forma de previsão do NBER, existiam os modelos econométricos tradicionais usados no monitoramento de curto prazo da economia, os quais voltavam a atenção para a previsão quantitativa do PIB. Entre estes modelos, estão: o Processo Autoregressivo Integrado de Média Móvel (ARIMA); os Vetores Autoregressivos (VAR); o

Heterocedasticidade Condicional Autoregressiva (ARCH); o Heterocedasticidade Condicional Autoregressiva Generalizada (GARCH). Estes modelos buscavam estabelecer pressupostos sobre o curso das variáveis exógenas, como política monetária, política fiscal ou os preços do petróleo para então realizarem suas previsões. Diebold (1997) revela que tais modelos não obtiveram sucesso em prever o estado da economia.

Foram Stock e Watson (1988; 1989; 1993) que inauguraram uma segunda geração de modelos econométricos para prever recessões. Estes autores buscaram explorar a sincronia cíclica do PIB por meio do uso intensivo de técnicas econométricas. Cabe ressaltar, como já afirmado, que o NBER já fazia este trabalho de previsão de uma forma heurística. Assim, quando é citado, no começo desta seção que os trabalhos de Stock e Watson foram um marco entre os modelos preditivos, este marco se refere especificamente aos modelos econométricos, que antes previam a variável dependente de forma quantitativa e depois passaram a prevê-la de forma qualitativa. A ideia central dos modelos de Stock e Watson foi tratar a variável do PIB como uma variável binária que indicava os momentos de recessão e expansão. Ao mesmo tempo que se passou a tratar a variável dependente de forma binária, os estudiosos passaram a fazer as previsões utilizando indicadores antecedentes. Dessa forma, os trabalhos de previsão que antes focavam o estudo especificamente na variável do PIB, fazendo previsões particularmente através de estudos econométricos sobre esta variável, passaram a focar a previsão numa variável latente da atividade econômica agregada, que poderia ser prevista através de um conjunto de séries antecedentes. Os estudos de Stock e Watson foram patrocinados pelo NBER, pelos quais os autores desenvolveram índices coincidentes e antecedentes de atividade econômica e um indicador de recessões.

Os trabalhos de Stock e Watson (1988; 1989; 1993) não obtiveram grande sucesso na previsão das recessões americanas, já que falharam em prever a recessão de 1990/1991. Os autores afirmaram que o problema principal se encontrou no processo de escolha dos indicadores antecedentes e não qualquer erro de especificação do modelo. A ideia dos autores é que a importância destes indicadores muda ao longo do histórico de recessões dos Estados Unidos (STOCK e WATSON, 1993). Chauvet (1998) utilizando a

mesma lógica de previsão de Stock e Watson (1988; 1988; 1989; 1993) acrescentou uma especificação de mudança de regime na previsão, ou seja, tratou de captar as assimetrias existentes entre os momentos de expansão e recessão, dado que esta última tende a ser mais brusca e curta, enquanto as expansões tendem a ser mais suaves e longas. Segundo Chauvet (1998), essa falha de não captar as assimetrias levou Stock e Watson a não conseguirem prever a recessão de 1990/1991. Estrella e Mishkin (1999) usando um probit e testando para observações dentro e fora da amostra chegaram à conclusão que o *spread* das taxas de juros de longo e curto prazo e os indicadores da bolsa de valores são bons previsores da atividade econômica americana. Este modelo com apenas dois indicadores foi capaz de prever a recessão de 1990/1991. Não é sabido na literatura modelos de previsão estatístico-econométricos que conseguiram prever as recessões americanas de 2001 e 2008.

Essa nova forma de previsão que relaciona uma variável latente dependente e um conjunto de indicadores antecedentes formaram um novo grupo de previsão que está em voga até o presente momento. Observando o esquema anteriormente apresentado, nota-se que esse novo grupo de previsão (com variável dependente binária + indicadores antecedentes) se bifurca em duas metodologias diferentes. Uma delas, em termos gerais, trata de dar continuidade os estudos de Stock e Watson, evidentemente com algumas variações no decorrer do tempo, mas sempre focando a construção de indicadores antecedentes e aplicando modelos de previsão com um forte apelo aos modelos econométricos específicos para estes fins, como, por exemplo, os modelos de probabilidade não-linear. A outra metodologia, que segue preceitos heurísticos, também faz uso de indicadores antecedentes e trata a variável a ser prevista como uma variável latente dicotômica, porém, utiliza técnicas menos sofisticadas para tais fins, como por exemplo, tratar não só a variável dependente de forma binária, mas também as variáveis antecedentes.

Um certo grupo de autores inseridos neste grupo de previsão heurística, chamados de econofísicos, tratam as recessões como sendo um evento extremo. Para eles, a justificativa principal do uso, na economia, destes modelos heurísticos é que os modelos econométricos comumente utilizados não levam em consideração o fenômeno da complexidade quando realizam

suas previsões. Segundo o matemático Mandelbrot (1963; 1963), em eventos complexos a curva que descreve a probabilidade de flutuações de determinada variável tem uma “cauda mais pesada” que a das distribuições gaussianas. Ou seja, as séries estudadas não são i.i.d. e, portanto, não são passíveis de uma boa análise de previsão econométrica. Se, por acaso, o comportamento da série temporal do nível de produto seguisse uma distribuição gaussiana, seria extremamente improvável a ocorrência de flutuações violentas como a da crise 1929 e da crise de 2008. O fato é que elas ocorreram e esse comportamento não-normal dos dados coloca em questão os modelos de previsão econométricos tradicionais.

Na próxima seção, intitulada "Métodos Heurísticos de Previsão", será descrito com detalhes como funciona a previsão com indicadores antecedentes, especificamente através do sistema de indicadores antecedentes elaborado pelo *The Conference Board* (TCB) e de um modelo que trata as recessões como sendo um evento extremo, elaborado por Keilis-Borok et al. (2000). Posteriormente, em outra seção, intitulada "Métodos Estatístico-Econôméticos de Previsão" serão descritos dois trabalhos de previsão de recessões brasileiras utilizando modelos estatístico-econôméticos, que são os modelos de Chauvet e Silva (2004) e Chauvet e Morais (2009).

3.1 Métodos heurísticos de previsão

3.1.1 Previsão de Recessões do *The Conference Board*

O pilar central do método de previsão utilizando indicadores antecedentes é supor que economias comportam-se diferente em momentos que precedem recessões (DIEBOLD e RUDEBUSH, 1989). Esse comportamento anormal apresenta um padrão sistematicamente recorrente, possível de ser identificado e eventualmente codificado, tornando-se possível a antecipação das recessões (ZALIAPIN e KEILIS-BOROK, 2002). A identificação de padrões recorrentes antes das recessões como método de previsão teve como precursor o livro clássico elaborado por Burns e Mitchell

(1946) intitulado "Measuring Business Cycles", pelo qual é sintetizada e esquematizada a metodologia para o sistema de indicadores antecedentes (SIA), que até hoje serve como base referencial de estudos na área. O livro foi um marco do estudo empírico e pragmático sobre o fenômeno dos ciclos econômicos, condensando praticamente meio século de pesquisas e estabelecendo um dos métodos de previsão cíclica mais consagrados desenvolvidos no século vinte e que serviu de base metodológica para os estudos do NBER, quando este ainda elaborava estudos de previsão centrados nos indicadores antecedentes. A partir de 1995, o Departamento de Comércio dos Estados Unidos escolheu o *The Conference Board* (TCB), uma entidade privada e sem fins lucrativos, para dar continuidade à tarefa de medição de séries relacionadas com a atividade econômica, ou seja, atualmente é o TCB que produz o sistema de indicadores antecedentes para a economia Norte-americana.

A lógica do SIA do TCB se baseia na perseguição de uma *target variable* e na construção de três tipos de índices: um coincidente, um antecedente e um defasado. A *target variable* utilizada para se estudar recessões é o PIB, ou seja, é a variável final a qual se deseja prever. Os índices construídos são relacionados com esta variável, ou seja, se um índice é considerado como coincidente, seu movimento é concomitante ao movimento da *target variable*; se o índice é considerado como antecedente ele precede o movimento da *target variable* e se o índice é defasado seu movimento ocorre depois do movimento da *target variable*. Sobre o porquê da construção dos índices para realizar as previsões e não utilizar o método de avaliar cada indicador antecedente de forma individual o TCB (2013) faz a seguinte citação:

The leading, coincident, and lagging economic indexes are essentially composite averages of several individual leading, coincident, or lagging indicators. They are constructed to summarize and reveal common turning point patterns in economic data in a clearer and more convincing manner than any individual component, primarily because they smooth out some of the volatility of individual components.

O índice coincidente é composto por um conjunto de séries coincidentes com a série do PIB. Ele é útil por dois motivos. O primeiro é que sua

construção pode ser feita antes da divulgação do PIB, visto que as séries que o integram são mais rapidamente divulgadas do que a série do PIB e assim pode facilitar o acompanhamento da economia em tempo real. O segundo motivo é que o índice coincidente, como já citado na seção sobre recessões, serve como um confirmador de que uma recessão está ocorrendo. Alguns trabalhos de previsão utilizam o índice coincidente como a *target variable*, já que ele retrata fielmente o estado atual da economia. O TCB utiliza na construção de seu índice coincidente as mesmas quatro variáveis coincidentes que o NBER usa para confirmar e datar os *turning points* dos *business cycles* americano. São variáveis representantes do emprego, da renda pessoal, da produção industrial e das vendas dos setores industrial e atacadista, como já citado na seção sobre recessões.

O índice antecedente é composto por um conjunto de séries antecedentes em relação à série do PIB e é útil pois apresenta precocemente o mesmo comportamento da *target variable*. Desse modo, se o índice antecedente apresentar um *turning point*, provavelmente, num futuro próximo, a *target variable* também apresentará esse *turning point*. Séries que representam estágios iniciais do processo produtivo, como, por exemplo, novas encomendas de bens de capital e de bens duráveis de consumo, em princípio precedem séries que representam a produção de bens de capital e de bens finais (que andam juntas com a série do PIB) e, assim, podem ser consideradas como séries antecedentes da evolução da atividade agregada. Do mesmo modo, séries que refletem importantes relações fluxo-estoque envolvendo a demanda e oferta domésticas de bens e serviços são influenciadas por variações na taxa de crescimento das economias líderes, nos estoques de produtos acabados das empresas e na oferta de moeda e crédito.

Atualmente, o TCB constrói seu índice antecedente através de dez séries de tempo consideradas como bons indicadores antecedentes para a atividade agregada norte-americana¹⁷. Cada observação de cada série antecedente entra no cálculo mensal do índice através de dois procedimentos

¹⁷ São elas: 1) a semana de trabalho média dos operários da indústria; 2) a média dos pedidos iniciais semanais de seguro-desemprego; 3) novas encomendas de bens e materiais de consumo; 4) ISM índice de novas encomendas; 5) novas encomendas de bens de capital que não seja para defesa; 6) concessões de novas autorizações de construção; 7) índices de preços de títulos imobiliários; 8) oferta monetária (M2) ajustada à inflação; 9) diferencial da taxa de juros de longo e curto prazo; 10) índice de expectativas do consumidor.

básicos. O primeiro é transformar cada observação em diferença percentual em relação ao mês anterior, ou, dizendo de outra forma, calculando sua taxa de crescimento. O segundo procedimento trata de ajustar essa taxa de crescimento através de um fator de normalização (*standardization factor*), que representa apenas uma ponderação que cada série possui no índice.

O procedimento para o cálculo do *standardization factor* é simples. Primeiramente calcula-se o desvio padrão v_x dos valores em taxa de crescimento de cada série antecedente. Em seguida inverte-se o valor de v_x , encontrando-se assim o valor do desvio-padrão invertido w_x para cada série. A soma de todos os w_x de cada série antecedente é dada por k . Assim, para cada série, o valor do *standardization factor* é dado através da seguinte fórmula:

$$\text{standardization factor} = \frac{1}{k \cdot v_x} \quad (01)$$

A soma dos *standardization factors* de todas as séries antecedentes iguala-se a um. O TCB deixa claro que estes fatores apenas tem a função de amenizar a volatilidade relativa de cada série tomada individualmente. Séries que apresentem um *standardization factor* menor indicam que são mais voláteis e por isso receberam um peso menor no índice. Ou seja, um *standardization factor* menor não significa necessariamente que a série tenha menos importância para o desempenho do índice. Anualmente os desvios-padrão de todas as séries são atualizados, mudando assim os *standardization factors* para cada indicador.

As duas últimas recessões norte-americanas, de 2001 e de 2008, oficialmente datadas pelo NBER, foram bem antecipadas pelo índice antecedente do TCB. A Figura 4 abaixo mostra o índice coincidente e o índice antecedente calculado pelo TCB para a economia norte-americana, na qual a linha vermelha indica a evolução do índice coincidente e a linha azul a evolução do índice antecedente e as áreas hachuradas indicam os períodos de recessão datados pelo NBER.

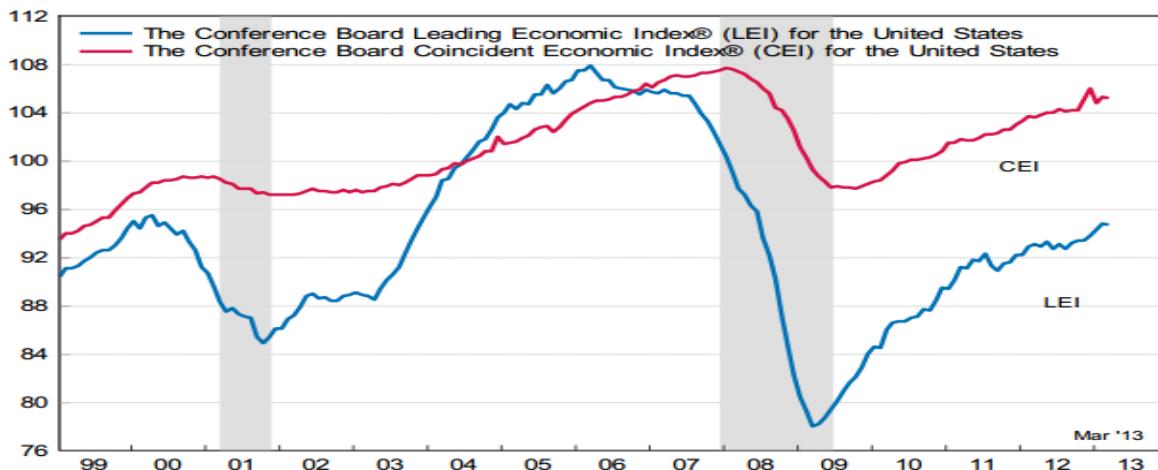


Figura 4 - Índice coincidente e índice antecedente para a os Estados Unidos.

Como pode ser observado, na recessão que começa em abril de 2001 o índice antecedente conseguiu antecipá-la em mais de um ano. Na recessão que começa em janeiro de 2008, o índice coincidente conseguiu antecipá-la em cerca de um ano de meio. A antecipação exata depende da análise que o TCB faz sobre o limite do comportamento normal do índice antecedente, ou seja, nem toda queda do índice antecedente representa uma sinalização que ocorrerá uma recessão, ele deve cair abaixo de certo patamar. Na Introdução deste estudo, através da ilustração lá representada e da análise do ruído/sinal, foi elaborado o conceito do limite do comportamento normal de um indicador antecedente.

Como o TCB não divulga qual seja o limite do comportamento normal para o seu índice antecedente, não se pode informar com exatidão em quantos meses o índice antecipou as recessões. Analisando com mais atenção a evolução do índice antecedente, pode ser notado que no ano de 2006 o índice apresentou forte queda. Se for feita uma análise de antecipação de um semestre¹⁸, essa queda em todo o ano de 2006 irá se refletir em reduções da taxa de crescimento do PIB no segundo semestre de 2006 e no primeiro semestre de 2007. Em março de 2006, após uma queda de 0,2% do índice antecedente, o economista Ken Goldstein, do TCB, já adiantou:

¹⁸ Sobre essa antecipação, o economista Ataman Ozyildirim, do TCB, afirmou que o procedimento para analisar como a economia estará indo nos próximos três a seis meses trata-se em analisar o comportamento do índice antecedente dos últimos seis meses.

No fundo, o que tivemos foi um crescimento moderado no primeiro trimestre. Podemos ter um ligeiro aumento no segundo e um ligeiro recuo no terceiro; o crescimento vai ser um pouco mais lento no segundo semestre do ano¹⁹.

De fato, o crescimento no segundo semestre foi mais lento que no primeiro, cresceu 2% contra 3% do primeiro semestre. No primeiro semestre de 2007, assim como apontou o índice antecedente um semestre antes, a atividade agregada continuou o desaquecimento do segundo semestre de 2006, apresentando um crescimento de 1,4%.

No ano de 2007 existem dois tipos de comportamento do índice antecedente. No primeiro semestre ele apresenta-se constante com uma leve tendência de alta e no segundo semestre é apresentada uma vertiginosa queda do índice. Isto significa que o segundo semestre de 2007 apresentaria taxas de crescimento ao menos iguais ao do primeiro semestre de 2007 e que o primeiro semestre de 2008 apresentaria uma desaceleração ou o início de uma recessão. De fato, o crescimento do segundo semestre de 2007 superou o crescimento do primeiro, apresentando uma taxa de crescimento de 2,3% contra os 1,4% de crescimento do semestre anterior. O primeiro semestre de 2008 também atendeu as expectativas do índice antecedente, com uma forte desaceleração da taxa de crescimento, apresentando 1,3% de crescimento e levando o NBER, que data as recessões a partir do estudo da série do PIB em nível, a apontar uma recessão com início no primeiro semestre de 2008.

Segundo o TCB, nos últimos seis meses de 2007 o índice antecedente acumulou uma queda de 0,8%. Pode-se notar que, em meados de outubro de 2007, o índice antecedente começa a decair a taxas crescentes, que foi exatamente quando começou a recessão da atividade industrial americana²⁰, sinalizando uma queda vertiginosa da atividade agregada, que viria a ocorrer em janeiro de 2008. Portanto, conclui-se que o método de previsão heurístico utilizado pelo TCB através de indicadores antecedentes mostrou-se eficaz para antecipar as duas últimas recessões da atividade norte-americana. O TCB

¹⁹ Essa afirmação do economista Ken Goldstein pode ser encontrada em Folha Online (2006).

²⁰ Esta recessão da atividade industrial americana será melhor apresentada nos resultados deste estudo, que procurou mostrar que tal série de atividade industrial é uma boa candidata a proxy da atividade agregada americana, visto que seus *turning points* ocorrem praticamente na mesma época, com pequenas diferenças de meses, dado que a recessão na indústria tende a surgir pouco antes do que a recessão na atividade agregada.

ainda afirma que os dados históricos também indicam o sucesso do indicador antecedente, que conseguiu captar sempre com alguma antecedência os *turning points* da atividade agregada.

3.1.2 Previsão de Recessão com um Evento Extremo

Ainda sobre trabalhos de previsão utilizando indicadores antecedentes, alguns autores (pertencentes ao grupo dos *econofísicos*) têm tratado alguns eventos econômicos, tais como a recessão e o alto índice de desemprego, como eventos extremos²¹. Deste modo, o método de previsão para antecipar recessões deve seguir a mesma metodologia com a qual eventos extremos são previstos.

Como já afirmado, uma das justificativas que os *econofísicos* usam de utilizar um método alternativo de previsão para prever recessões é que as séries econômicas não são i.i.d. e por isso os modelos econométricos tradicionais não são adequados. A outra justificativa é que quando uma recessão é reconhecida *ex-post*, geralmente o órgão oficial datador analisa quatro séries econômicas para confirmar se houve mesmo um período recessivo. Assim, uma análise para estabelecer previsões de momentos recessivos que leve em consideração apenas a variável do PIB é incompleta²². Segundo Keilis-Borok et al. (2000), deve ser feita uma análise com várias séries de tempo em um ambiente econômico possivelmente não-estacionário, no qual estas séries de tempo não necessitam ter uma relação de causalidade com a variável a ser prevista.

A previsão de eventos extremos nos vários campos da ciência tem sido influenciada pelas técnicas desenvolvidas pelos sismólogos para prever terremotos de grandes proporções. Está bem estabelecido na literatura que

²¹ Será dedicado um capítulo deste trabalho (Capítulo 4) para caracterizar o que é um evento extremo. Porém, em uma definição mais simples, um evento extremo possui duas características básicas: ser raro e causar um grande impacto quando ocorre.

²² Essa crítica aos modelos econométricos refere-se aos modelos antes dos trabalhos de Stock e Watson que tratam a previsão do PIB de forma quantitativa, com uma função contínua e com as variáveis explicativas possuindo uma relação de causa e efeito com a variável dependente. Como será relatado mais a frente, neste método alternativo de previsão as variáveis explicativas não necessitam ter uma relação de causa e efeito com a variável dependente. As variáveis explicativas necessitam apresentar o mesmo movimento da série dependente alguns meses, além de alguma justificativa teórica para esta antecedência.

terremotos de grandes proporções são precedidos por aumentos simultâneos de atividades sísmicas em locais adjacentes próximos a falhas geológicas, formando um padrão. Com base nesses resultados, os sismólogos têm construído algoritmos altamente eficazes que preveem a eclosão de grandes terremotos quando a porcentagem de locais próximos com movimentação sísmica acima de certo patamar ultrapassa um limiar crítico (ZALIAPIN e KEILIS-BOROK, 2002). Os econofísicos, ao encararem alguns eventos econômicos como eventos extremos têm tentado desenvolver algoritmos semelhantes ao dos sismólogos capazes de gerar, por exemplo, alarmes de ocorrência de recessões.

Keilis-Borok et al. (2000) elaboraram um excelente estudo de previsão dos inícios das recessões americanas aplicando a mesma lógica de previsão utilizada na antecipação de terremotos, utilizando apenas seis indicadores antecedentes²³. Os autores afirmaram que o objetivo do trabalho é explorar o modelo de previsão e não comparar as propriedades de previsão de diferentes indicadores antecedentes. Afirmaram também que a seleção dos indicadores antecedentes se deu antes de serem realizado quaisquer tipos de transformações nas séries, transformações estas que serão relatadas mais a frente.

Este artigo foi atualizado e publicado no livro *Complex Systems in Finance and Econometrics*, organizado por Meyers (2009) e está inserido no capítulo *Extreme Events in Socio-economic Political Complex Systems, Predictability of*. A previsão se deu utilizando dados mensais de 1961 até 2001 (no trabalho original, os dados vão até o ano de 1996) que serviram para construção do algoritmo, o qual foi testado para prever as recessões americanas datadas pelo NBER, mostradas na Tabela 2 abaixo:

²³ Os indicadores foram: diferencial de juros de longo e curto prazo; taxa de juros de três meses, produção de manufaturados; número de pedidos de seguro desemprego; número de anúncios de empregos nos jornais e um índice coincidente da atividade agregada. Como será explicado mais a frente, este índice coincidente da atividade agregada é considerado um indicador antecedente devido a transformação que ele sofre. Basicamente, a transformação realizada na série consegue captar os momentos de aceleração e desaceleração do índice, o que é bastante útil na previsão de recessões, visto que antes de toda recessão sempre existirá uma desaceleração da economia.

Tabela 2 - Recessões dos Estados Unidos - 1960 a 2001

Recessão	Pico	Vale
1	1960:04	1961:02
2	1969:12	1970:11
3	1973:11	1975:03
4	1980:01	1980:07
5	1981:07	1982:11
6	1990:07	1991:03
7	2001:03	2001:11

Fonte: NBER (2013).

A lógica da previsão se baseia em determinar o comportamento normal para cada indicador antecedente, determinação esta semelhante ao procedimento citado na introdução deste estudo. Para determinação desse comportamento normal é necessário efetuar algumas transformações nas séries, transformações estas, segundo os autores, que requerem julgamento crítico do pesquisador para realizá-las. Estas transformações podem ser de dois tipos. No primeiro, os valores dos indicadores antecedentes são analisados em desvios da tendência, ou seja, a série de cada indicador é filtrada de sua tendência de longo prazo e depois subtraída destes valores filtrados, obtendo assim o *local linear detrending*²⁴, uma nova série, portanto. As séries que sofreram esse tipo de transformação foram: taxa de juros de três meses, produção de manufaturados e a série do índice coincidente. No segundo tipo, as séries são suavizadas através de uma média móvel (no caso deste estudo, uma média móvel de cinco meses para a série de solicitações de seguro desemprego e uma média móvel de 10 meses para a série de número de anúncios de empregos nos jornais) e depois calculadas suas taxas de crescimento. A série de diferencial de juros de longo e curto prazo não sofreu nenhuma alteração.

²⁴ Formalmente, a nova série *local linear detrending* é dada por $y = y_t - y_f$ em que $y_t - y_f$ são, respectivamente os valores originais da série e seus valores filtrados. Esse tipo de procedimento é o mesmo que é usado para se determinar os *growth cycles* de uma série de tempo, ou seja, através dele obtém-se os momentos de aceleração e desaceleração da série em estudo, assim como apontado por Zarnowitz e Ozyildirim (2006).

Dadas as transformações, são analisados os comportamentos dos valores das séries transformadas antes dos momentos recessivos e a partir desta análise é determinado o limite do comportamento normal do indicador. Os autores não entram em detalhes para o procedimento de determinação do limite. A Figura 5 abaixo ilustra os valores do *local linear detrending* para a produção de manufaturados, bem como o limite do comportamento normal deste indicador que é representado através da linha tracejada horizontal. As linhas tracejadas verticais delimitam as recessões datadas pelo NBER. Para fins de melhor visualização do comportamento do indicador, no gráfico, os autores definiram arbitrariamente para os meses de recessão o valor zero, o que facilita a visualização do comportamento do indicador antes e depois das recessões.

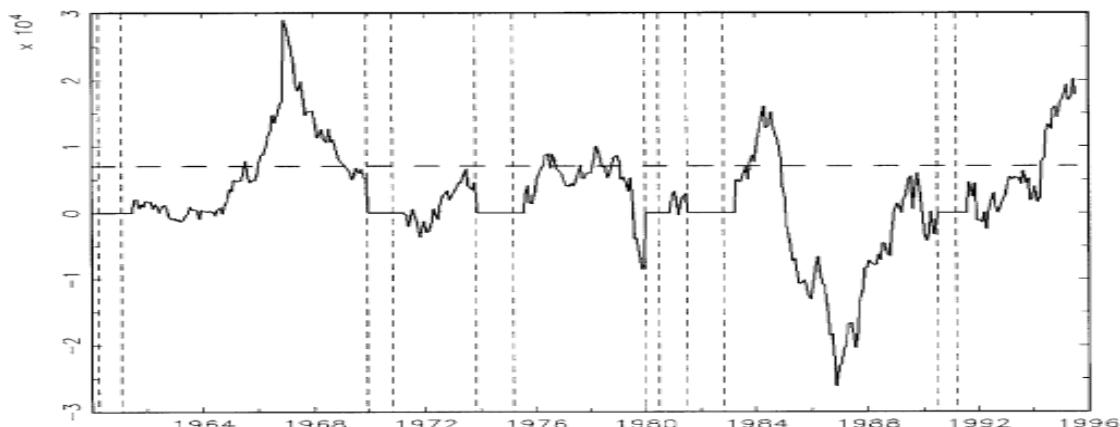


Figura 5 - *Local linear detrending* da série de produção de manufaturados dos Estados Unidos desde 1960 até 1996.

O indicador de produção de manufaturados indica uma recessão quando os valores do *local linear detrending* estiverem abaixo do nível determinado pela linha tracejada horizontal. Observando a Figura 5, é possível notar que o indicador sinalizou bem todas as recessões. Porém, houve vários falsos indicativos de recessão, que, segundo os autores, são compensados e suavizados quando é feita a análise conjunta dos indicadores. O procedimento para os outros tipo de transformação é o mesmo, são plotados os valores da variável em um gráfico e então se analisa seus comportamentos antes dos períodos recessivos. As séries que representam a taxa de juros de três meses e o número de pedidos de seguro desemprego apresentam um comportamento inverso em relação às outras quatro séries, ou seja, espera-se que seus

valores estejam acima e não abaixo do limite que é considerado normal antes das recessões. O cálculo para achar o valor exato de do limite não é especificado pelos autores.

Para um dado mês, quando a metade ou mais da metade dos indicadores estiverem fora do comportamento normal o alarme de recessão é acionado. Formalmente, o alarme soa quando a *distância de hamming* de determinado vetor passa de um limite considerado normal. Essa *distância de hamming* também é chamada de *distância do sinal*. Basicamente, ela compara um vetor limite, chamado de vetor *kernel*, com os demais vetores a serem analisados. No referido trabalho, como são seis indicadores, um vetor qualquer pode ser apresentado como (0,0,0,0,1,1), no qual os zeros representam os indicadores que fugiram de seu comportamento normal e os valores unitários representam os indicadores que estão de acordo com o comportamento normal. O vetor limite é determinado a partir do que se considera normal para períodos não-recessivos, que no caso do referido trabalho é apresentado na forma (0,0,1,1,1,1), não importando a ordem que os vetores se apresentem. Caso um determinado vetor apresente três ou mais zeros ele estará fugindo do comportamento normal e então o alarme é acionando. A Figura 6 abaixo mostra através das barras pretas os alarmes das recessões e através das barras cinzas os momentos de recessão oficialmente datados pelo NBER.

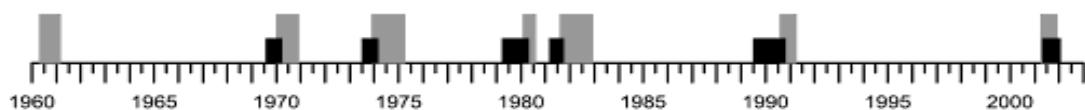


Figura 6 - Alarmes das recessões americanas (barras pretas) e recessões americanas (barras cinzas).

Como pode ser observado, das seis recessões possíveis de serem previstas, cinco o alarme soou alguns meses antes, alarmando as recessões com cinco a quinze meses de antecedência. Não houve falsos alarmes. A recessão de 2001 foi a única na qual o alarme só foi acionado após seu início. Os autores testaram a capacidade do algoritmo para dados fora da amostra. O método para construção do algoritmo e testá-lo para dados não-amostrais baseia-se num experimento recursivo, pelo qual os valores dos limites dos

comportamentos normais dos indicadores são calculados com base nas observações de uma janela deslizante das recessões, ou seja, constrói-se os limites até a recessão de 1969, por exemplo, e testa-se o algoritmo para prever as próximas recessões, em seguida constrói-se os limites até a recessão de 1973 e testa o algoritmo para as recessões seguintes, e assim por diante. Por esse método, a previsão apresentou capacidade preditiva tão boa quanto para os dados amostrais, a diferença se dá na duração dos alarmes, que para os dados amostrais têm uma duração um pouco maior. Por fim, foi feita uma comparação com modelos clássicos de previsão²⁵ e, segundo, os autores, foi surpreendente que um modelo tão simples em que as informações dos dados são reduzidas a indicadores binários tivesse conteúdo preditivo comparável, ou em muitos casos, melhores que modelos clássicos de previsão.

3.2 Métodos estatístico-econométricos de previsão

3.2.1 Previsão de Recessões da Indústria de Bens de Capital Brasileira

O estudo de previsão com indicadores antecedentes mais recente no Brasil foi elaborado por Chauvet e Moraes (2009) e é intitulado "Indicadores Antecedentes para a Indústria de Bens de Capital", o qual foi concebido em 2009 com medalha de honra pela Confederação Nacional das Indústrias (CNI). Marcelle Chauvet é membro do CODACE e, desde 1998, vem trabalhando com previsão de recessões econômicas, com diversos artigos internacionais publicados na área. Neste trabalho citado, é utilizado um modelo probit simples e um modelo probit com dinâmica autorregressiva para prever os momentos recessivos da série de produção de bens de capital. A variável a ser prevista foi uma variável binária que representava os ciclos da produção de bens de capital entre janeiro de 1991 e abril de 2009, na qual os zeros representavam os períodos de recessão e os valores unitários os períodos de não-recessão. Essa datação dos momentos de recessão e expansão foi feita utilizando a

²⁵ Os autores utilizaram um modelo de probabilidade linear e um modelo de probabilidade não-linear, no caso, um probit.

metodologia de datação de *turning points* sugerida por Bry e Boschan (1971), a qual captou cinco momentos de recessão para a série estudada, que são sintetizados na Tabela 3 a seguir:

Tabela 3 - Recessões da indústria de bens de capital - 1991 a 2012

Recessão	Pico	Vale
1	1991:07	1992:03
2	1995:03	1996:07
3	1997:10	1999:07
4	2001:08	2003:03
5	2008:10	2009:05

Fonte: Chauvet e Moraes (2009).

Foram utilizados seis indicadores antecedentes²⁶, todos eles séries da própria indústria de bens de capital. Estas séries foram selecionadas dentre várias outras séries de bens de capital através da análise espectral cruzada, que permite identificar se existe característica cíclica comum através da proporção da variância explicada entre o indicador antecedente e a variável alvo de previsão. Escolhidas estas variáveis antecedentes, realizou-se as previsões com os dois tipos de modelo probit. Os resultados indicaram que o modelo probit com dinâmica autorregressiva apresentou um desempenho de previsão bem superior ao modelo probit simples, tanto para previsões dentro ou fora da amostra. Para as previsões dentro da amostra foram utilizadas todas as observações. Para as previsões fora da amostra, o modelo foi construído de janeiro de 1991 até dezembro de 2006, enquanto que as demais observações, de janeiro de 2007 até abril de 2009, foram utilizadas para a previsão fora da amostra.

A Figura 7 a seguir mostra os resultados da pesquisa utilizando o modelo probit dinâmico autorregressivo, em num cenário de previsão de curto prazo e outro de longo prazo. Os gráficos (a) e (b) representam as previsões de curto e longo prazo para dentro da amostra, as linhas azuis mostram as probabilidades de recessão e as áreas hachuradas delimitam as recessões. Os

²⁶ Os seis indicadores foram: extração de carvão; produção têxtil; produção de produtos químicos inorgânicos; produção de tratores, máquinas e equipamentos agrícolas; produção de material elétrico para veículos e a produção de carrocerias e reboques.

gráficos (c) e (d) representam as previsões de curto e longo prazo para fora da amostra, as linhas vermelhas mostram as probabilidades de recessão para o probit dinâmico autorregressivo, as linhas azuis representam as probabilidades de recessão para um probit simples e as barras verticais o início da recessão de 2008. Os quatro gráficos possuem em seu eixo vertical os valores das probabilidades. Os períodos recessivos são caracterizados como aqueles nos quais a probabilidade de recessão ultrapassa o limite de 50%, sem que se exija a decorrência de um período mínimo de meses para que seja declarada uma recessão.

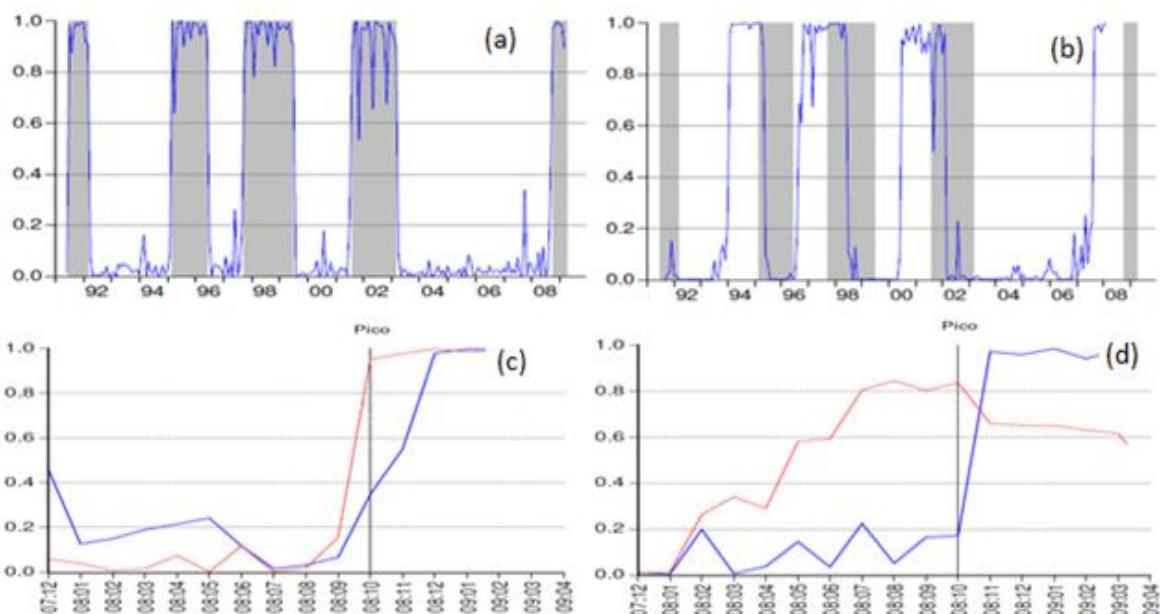


Figura 7 - Gráficos (a) e (b): probabilidades de recessões (linhas azuis) e recessões (áreas hachuradas) para a indústria de bens de capital. Gráficos (c) e (d): probabilidades de recessão do probit simples (linhas azuis), probabilidades de recessão do probit autorregressivo (linhas vermelhas) e a recessão de 2008 da indústria de bens de capital (linha vertical preta).

Como pode ser observado através dos gráficos, para as previsões dentro da amostra, gráficos (a) e (b), o modelo de curto prazo, gráfico (a), não conseguiu nenhuma antecipação das recessões. O modelo de longo prazo para previsões dentro da amostra, gráfico (b), conseguiu antecipar as recessões com um horizonte médio de seis meses, utilizando para isto indicadores antecedentes defasados em cerca de vinte meses. Para as previsões fora da amostra, gráficos (c) e (d), o modelo de curto prazo do probit

dinâmico, linha vermelha do gráfico (c), conseguiu antecipar a recessão de 2008 em apenas dois meses. O modelo de longo prazo do probit dinâmico para previsões fora da amostra, linha vermelha do gráfico (d), conseguiu antecipar a recessão de 2008 em cinco meses. O modelo probit simples, linhas azuis dos gráficos (c) e (d), não se mostrou eficaz nas previsões fora da amostra. Cabe salientar que as previsões do modelo probit simples para dentro da amostram também não se mostraram eficazes, por isso não foram aqui apresentadas.

3.2.2 Previsão de Recessões da Atividade Agregada Brasileira

Em trabalho anterior, Chauvet e Silva (2004) construíram um modelo de previsão para as recessões brasileiras utilizando um modelo probit. A *target variable* utilizada foi a taxa de crescimento do PIB real entre janeiro de 1975 e dezembro de 2002. Ela foi transformada em valores binários, os quais os zeros representaram as recessões e os valores unitários as expansões. Basicamente, o modelo identificou as recessões como os períodos nos quais o PIB decresce a uma taxa igual a sua média menos um desvio padrão. Conforme esse critério, foram identificadas dez recessões entre janeiro de 1980 e dezembro de 2002, que segundo os autores, a causa da maioria delas está relacionada a períodos de contrações econômicas mundiais. A Tabela 4 abaixo mostra as datações das recessões.

Tabela 4 - Recessões brasileiras desde 1975 até 2002

Recessão	Pico	Vale
1	1975:02	1976:02
2	1979:01	1979:05
3	1980:02	1981:09
4	1982:10	1983:03
5	1987:04	1987:09
6	1988:07	1989:02
7	1990:01	1991:06
8	1991:12	1992:05

9	1995:05	1995:10
10	1997:11	1998:12

Fonte: Chauvet e Silva (2004).

A previsão se dá em dois cenários, um pós-Plano Real que vai de agosto de 1994 até maio de 2002 e outro com todos os dados, que vai de janeiro de 1975 até maio de 2002. Foram realizadas previsões de curto prazo e de médio prazo. No caso do modelo pós-Plano Real, que contém apenas as recessões de 1995 e 1997, são analisados quatro horizontes de previsão, utilizando para cada uma deles indicadores antecedentes distintos, que podem ser vistos no Quadro 1, abaixo, no qual os meses mostrados na primeira coluna indicam as defasagens dos indicadores.

	Balança comercial, papel ondulado, taxa de juros americana com maturidade de 3 meses, taxa de juros americana com maturidade de 10 meses e produção de máquinas e equipamentos
a) Curto Prazo - 2 meses	Retorno de ações do Ibovespa, risco Brasil JP Morgan e balança comercial
b) Curto Prazo - 4 meses	Taxa de juros selic, índice de preços de máquinas e equipamentos, risco Brasil JP Morgan, indicador antecedente americano
c) Médio Prazo - 8 meses	Risco Brasil JP Morgan, taxa de juros americana com maturidade de três meses, indicador antecedente americano
d) Médio Prazo - 10 meses	

Quadro 1 - Indicadores antecedentes do modelo pós-Plano Real

Os indicadores antecedentes foram escolhidos de acordo com a capacidade em antecipar os *turning points* da série do PIB, capacidade esta medida através de um probit simples, no qual a variável dependente era a variável binária do PIB e a variável explicativa o indicador antecedente. Como pode ser observado, variáveis que remetem ao setor externo estão presentes nos quatro tipos de previsão do modelo pós-Plano Real.

A Figura 8 abaixo mostra as probabilidades de recessão, obtidas através de um modelo probit, para os quatro tipos de horizonte de previsão. Os gráficos (a) e (b) remetem ao curto prazo e os gráficos (c) e (d) remetem ao médio

prazo, seguindo a mesma indicação apresentada no Quadro 1. As áreas hachuradas delimitadas pelas linhas vermelhas representam as recessões e as linhas azuis representam as probabilidades das recessões. Os quatro gráficos possuem em seu eixo vertical os valores das probabilidades. Os períodos recessivos são caracterizados como aqueles nos quais a probabilidade de recessão ultrapassa o limite de 50%, sem que se exija a decorrência de um período mínimo de meses para que seja declarada uma recessão.

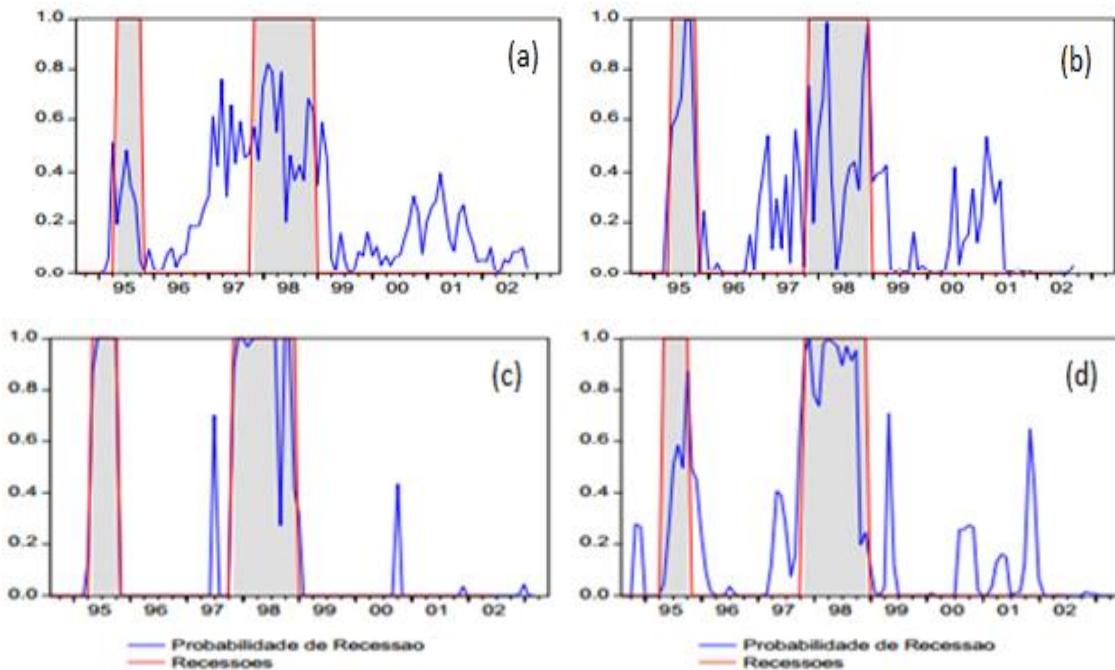


Figura 8 - Probabilidades de recessão (linhas azuis) e cronologia das recessões (áreas hachuradas delimitadas com linhas vermelhas) no pós-Plano Real²⁷.

Como pode ser observado, os dois modelos de curto prazo, gráficos (a) e (b), só conseguem prever a recessão de 1997, com cerca quatro meses de antecedência, dado que para a recessão de 1995 as probabilidades sobem acima dos 50% apenas quando a recessão se inicia. Nos dois modelos de médio prazo, gráficos (c) e (d), apenas o modelo mostrado no gráfico (c) consegue alguma antecipação de uma recessão, que foi a recessão de 1997, com cerca de três meses de antecedência.

No caso do modelo que vai de janeiro de 1975 até maio de 2002, também são analisados quatro horizontes de previsão, utilizando para cada uma deles indicadores antecedentes distintos, que podem ser vistos no Quadro

²⁷ Os quatro gráficos remetem ao Quadro 1.

2 abaixo, no qual os meses mostrados na primeira coluna indicam as defasagens dos indicadores.

	Pessoal ocupado, balança comercial, diferencial de juros americanos de longo e curto prazo, índice de preços de máquinas e equipamentos.
e) Curto Prazo (3 a 6 meses)	
f) Curto Prazo (3 a 6 meses)	Pessoal ocupado, balança comercial, índice de preço de máquinas e equipamentos, retorno de ações.
g) Médio Prazo (9 a 11 meses)	Taxa de juros selic, índice de preços de máquinas e equipamentos, taxa de câmbio efetiva, indicador antecedente americano, diferencial de juros americanos de longo e curto prazo.
h) Médio Prazo (9 a 11 meses)	Taxa de juros selic, índice de preços de máquinas e equipamentos, taxa de câmbio efetiva, indicador antecedente americano, pessoal ocupado, taxa de juros americanas com maturidade de 3 meses

Quadro 2 - Indicadores Antecedentes do Modelo de 1975 até 2002.

Os indicadores antecedentes foram escolhidos de acordo com o mesmo procedimento relatado para o modelo do pós-Plano Real. Como pode ser observado, variáveis que remetem ao setor externo também estão presentes nos quatro horizontes de previsão deste cenário.

A Figura 9 abaixo mostra as probabilidades de recessão obtidas através de um modelo probit, para os quatro tipos de horizonte de previsão. Os gráficos (e) e (f) remetem ao curto prazo e os gráficos (g) e (h) remetem ao médio prazo, seguindo a mesma indicação apresentada no Quadro 2. As áreas hachuradas delimitadas pelas linhas vermelhas representam as dez recessões deste período e as linhas azuis representam as probabilidades das recessões. Os quatro gráficos possuem em seu eixo vertical os valores das probabilidades.

Os períodos recessivos são caracterizados como aqueles nos quais a probabilidade de recessão ultrapassa 50%, sem que se exija a decorrência de um período mínimo de meses para que seja declarada uma recessão.

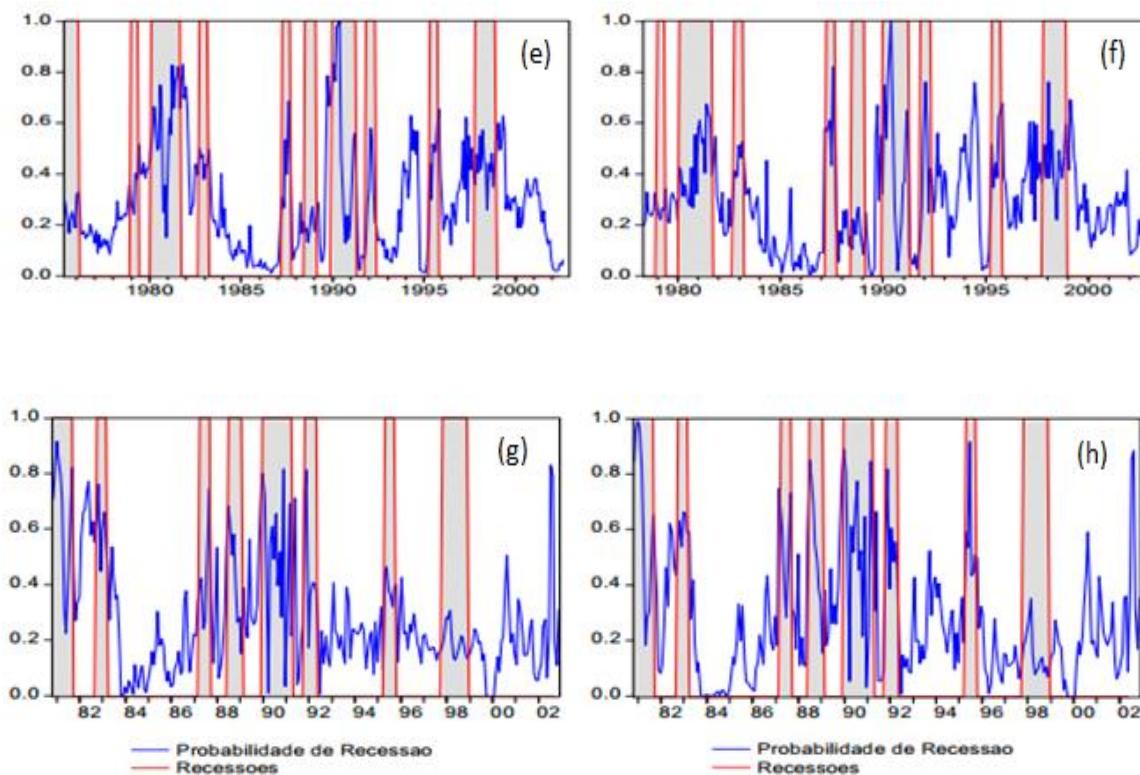


Figura 9 - Probabilidades de recessão (linhas azuis) e cronologia das recessões (áreas hachuradas delimitadas com linhas vermelhas)²⁸ - 1975 a 2002.

Como pode ser observado, os dois modelos de curto prazo, gráficos (e) e (f), só conseguiram prever com alguma antecedência as recessões de 1995 e 1997, dado que as probabilidades de recessão sobem acima dos 50% cerca de três a seis meses antes destas recessões. Nas demais recessões, as probabilidades só passam a aumentar quando estas já se iniciaram. Nos dois modelos de médio prazo, gráficos (g) e (h), existem antecipações para as recessões de 1982, 1988, 1990, 1991, nas quais as probabilidades sobem acima dos 50% meses antes das recessões. A análise para a década de oitenta nota-se confusa, devido a, segundo os autores, o baixo desempenho da economia brasileira nesta época, ocasionando portanto regulares momentos de ascensão das probabilidades de recessão. O gráfico da previsão (h) também conseguiu antecipar a recessão de 1995. Já a recessão de 1997 não conseguiu ser prevista em nenhum dos modelos de médio prazo.

Em relação às previsões com dados não-amostrais, foi utilizado o método da previsão recursiva. Dessa forma, os modelos são recursivamente

²⁸ Os quatro gráficos remetem ao Quadro 2.

re-estimados, permitindo que se reproduza a informação que estaria disponível para se fazer previsões a cada mês em tempo real. Os resultados não são apresentados com detalhes pelos autores, porém, estes afirmam que o cenário pós-Plano Real respondeu de melhor forma, se aproximando do modelo de previsão gerado com os dados amostrais, especificamente para os modelos (a) e (c), ou seja, um de curto prazo e outro de médio prazo.

4 EVENTOS EXTREMOS

Eventos extremos são eventos raros que causam um grande impacto. Tais eventos também são conhecidos como fenômenos críticos, desastres, catástrofes e crises. Eles surgem persistentemente em *sistemas complexos*. São criados pela natureza ou pela sociedade, ou às vezes, por ambos (MEYERS, 2009). Ele engloba uma série de atributos como raro, excepcional, catastrófico, surpreendente, dentre outros. É considerado raro devido à sua baixa probabilidade de ocorrência e catastrófico, pois gera graves consequências no meio que ocorre. Desse modo, um furacão de grandes proporções só é considerado um evento extremo se causa perda de vidas e gera danos materiais. Caso ele atinja áreas desabitadas ele não é considerado um evento extremo²⁹. Outro exemplo é a queda de um asteroide que só se caracteriza um evento extremo quando atinge a Terra e muda o curso da evolução, como ocorreu há 65 milhões de anos (ALBEVERIO et al., 2005).

Eventos extremos podem ser individuais, como o nascimento de uma criança, a morte de um conjugue ou a atribuição de um prêmio Nobel; podem ser desastres sociais como a Gripe Suína ou a AIDS; desastres naturais, como inundações, secas e ciclones; crises de mercado, como grandes perdas ou ganhos no mercado de ações ou grandes quedas na atividade econômica em geral (ALBEVERIO et al., 2005). Nota-se, portanto, que a definição de evento extremo ainda está sujeita a várias interpretações. É necessária uma definição mais rígida do conceito. O que se pode fazer é caracterizar um evento extremo por suas propriedades estatísticas.

Deste ponto de vista, eventos extremos ocorrem na cauda de distribuições de probabilidade. Em uma distribuição gaussiana (distribuição normal) estas caudas são exponenciais (como pode ser visto na Figura 10), ou seja, à medida que a curva vai se distanciando da média ela cai rapidamente. Isto ocorre pois a maioria dos valores se concentram na parte central da curva,

²⁹ Essa caracterização de eventos extremos foi definida para fins de política, ou seja, no sentido de poder melhorar as condições de vida dos seres humanos. Se for levar em conta o ponto de vista extremamente científico, por vezes basta um evento ser muito raro para ser considerado como extremo, como, por exemplo, uma grande chuva magnética de poeira solar na nossa atmosfera.

em torno da média, enquanto que valores muito longe da média são estatisticamente pouco prováveis e, depois de certo limite, impossíveis.

Até certo ponto, é correto afirmar que distribuições gaussianas são a forma da natureza, cuja larga aplicabilidade resulta do *teorema do limite central*: em qualquer caso onde um grande número de eventos aleatórios independentes contribuem para um determinado resultado, este seguirá a distribuição normal. Através da regra dos 3-sigmas (três desvios-padrão em relação a média) praticamente todos os valores estão inseridos dentro da distribuição, representando 99,7% do conjunto de dados. Grandes variações em relação a média (mais que três desvios-padrão) possuem baixíssima probabilidade de ocorrência, com aproximadamente 0,3% de probabilidade. Uma distribuição gaussiana representa bem, por exemplo, a variação nas alturas das pessoas, com um valor médio bem definido e uma variação relativamente pequena. Matematicamente, a função densidade para a distribuição gaussiana é:

$$f(x) = \left(\frac{1}{\sqrt{2\pi}}\right) e^{-x^2/2} \quad (2)$$

na qual o núcleo dessa distribuição é dado por:

$$g(x) = e^{-x^2/2} \quad (3)$$

Ou seja, o gráfico da função $g(x)$ é um gráfico da distribuição de probabilidade usual em forma de sino. Como a função f é tão somente g vezes $\frac{1}{\sqrt{2\pi}} \approx 0,39$, fazendo com que o gráfico de f seja semelhante ao gráfico de g . O comportamento do gráfico da função g é não-escalável, já que possui uma escala característica determinada pela constante $^{-1/2}$. Se tal constante for aumentada em módulo, ou seja, seu valor for para -2 por exemplo, a curva apresentará um comportamento assintótico mais cedo, como pode ser visto na Figura 11. O objetivo desta comparação é apenas para familiarizar o leitor com a ideia de escalabilidade que será melhor apresentada posteriormente.

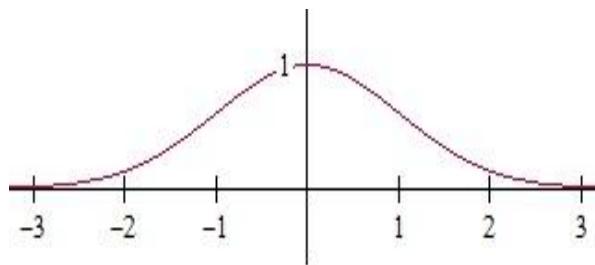


Figura 10 - Representação matemática da função $g(x) = g(x) = e^{-\frac{x^2}{2}}$: aproximação de uma distribuição gaussiana.

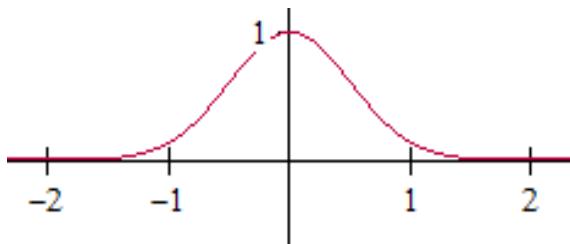


Figura 11 - Representação matemática da função $g(x) = e^{-2x^2}$

Vários trabalhos na literatura vêm mostrando que diversos eventos não seguem uma distribuição gaussiana. O precursor destes trabalhos foi o matemático francês Benoit Mandelbrot. Ele estudou uma série histórica dos preços do algodão e notou que a ocorrência de grandes variações era muito maior que a prevista pela distribuição gaussiana. A curva que descreveria a probabilidade de flutuações dos preços tinha uma "cauda mais pesada" que aquela distribuição. Foi o ponto de partida para mostrar que mesmo eventos não-gaussianos podem apresentar algum tipo de regularidade (Gleria et al. 2004).

Eugene Stanley e Rosário Mantegna, autores do livro *Econophysics*, de 1999, analisaram uma enorme quantidade de dados sobre ações na bolsa de valores. A análise mostrou que a distribuição gaussiana, assim como relatou Mandelbrot, subestimava a ocorrência de grandes flutuações. A curva que melhor descreveria as flutuações nos preços parecia ser formada por uma *lei de potência* (*power law* ou *lei de escala*), que é a marca registrada de um

sistema complexo, pois representa sistemas que não possuem precisamente uma escala característica.

Uma *lei de potência*, diferentemente da distribuição gaussiana, é incompatível com a noção de que a média representa a escala característica. As *leis de potência* conseguem representar a distribuição de riqueza (conhecida como Lei de Pareto), as distribuições de intensidade dos terremotos (Richter-Gutenberg); a ocorrência de incêndios e avalanches, dentre outros exemplos. As chamadas *leis de potência* descrevem melhor a ocorrência de eventos extremos pois possuem suas caudas de distribuição mais pesadas, caindo mais lentamente do que distribuições gaussianas. Matematicamente, segue um exemplo de uma *lei de potência*:

$$P(x) = Kx^{-\alpha}, \alpha > 0 \quad (4)$$

Quando $\alpha = 2$ tal lei se aproxima de uma distribuição normal (ver, por exemplo, Koedijk et al., 1990), como mostrado na Figura 12 através do gráfico (a). Dessa forma, valores iguais ou maiores que dois não atribuem grandes probabilidades a eventos extremos, dado que, matematicamente, a curva começa a se comportar assintoticamente mais cedo. Para efeito de comparação, a Figura 10 comparara três valores de α : 2; 1,5 e 1. Quanto menor for o valor do alfa significa a presença de caudas mais pesadas na distribuição, ou seja, contêm uma porcentagem maior de eventos extremos.

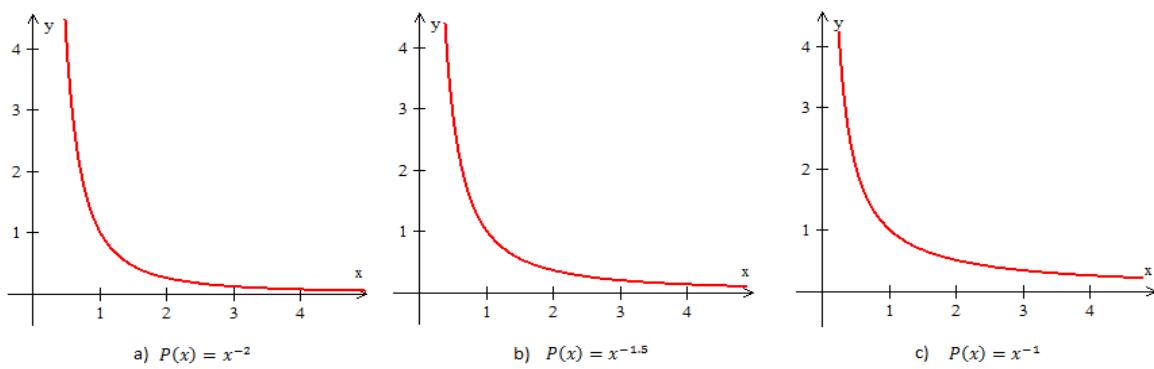


Figura 12 - Funções matemáticas análogas a leis de potência. (a) Representação matemática da função $P(x) = x^{-2}$: distribuição de um evento semelhante a uma distribuição gaussiana (b) Representação matemática da função $P(x) = x^{-1.5}$ (c) Representação matemática da função $P(x) = x^{-1}$: evento sujeito a Lei de Zipf.

No gráfico (b), no qual o valor de alfa foi de 1,5, representa, por exemplo, o estudo do economista Pareto para explicar a concentração de renda na Inglaterra, em diversas cidades da Itália e estados da Alemanha, em Paris e no Peru, utilizando dados do final do século XIX. Seu estudo indicou que 10% dos domicílios mais ricos recebiam 46% da renda total e os 20% mais ricos recebiam 58% da renda total, o que implicaria um alfa de 1,5 e um índice de Gini de 0,5.

No gráfico (c), quando o valor do alfa assume o valor unitário, a *lei de potência* gerada é conhecida como Lei de Zipf. Esta lei, formulada por Zipf, linguista da *Harvard University*, diz respeito ao comportamento humano e o princípio do menor esforço. O pesquisador identificou que a probabilidade de ocorrência de pequenas palavras no decorrer dos textos é maior quanto menor for a palavra. Em um exemplo hipotético, no caso de uma palavra com uma letra, a probabilidade de sua ocorrência, dada a equação do gráfico (c), é de $\frac{1}{1}$ e no caso de uma palavra com 100 letras a probabilidade de ocorrência é de $\frac{1}{100}$, o que significa que a probabilidade do evento com tamanho 1 é cem vezes maior. Ou seja, a distribuição de probabilidade não segue a forma normal. Como regra geral, quanto menor for o expoente α , maior a frequência de casos na cauda superior da distribuição, i.e. maior a probabilidade de eventos extremos (TALEB, 2007).

Dado o que é uma *lei de potência* e sua aplicabilidade, faz-se importante entender as diferenças entre escalabilidade de uma distribuição gaussiana e a escalabilidade de uma distribuição seguindo uma *lei de potência*. Esta comparação pode ser feita transformando-se as equações de ambas as distribuições em logaritmos. Primeiramente, a transformação da *lei de potência* (04):

$$\log(P(x)) = \log(K) - \alpha \log(x) \quad (4.1)$$

Faremos um valor de $\alpha = 1.5$ e para simplificação, um valor de $K = 1$:

$$\log(P(x)) = -1.5 \log(x) \quad (4.2)$$

Transformando a equação (03) do núcleo da distribuição gaussiana temos:

$$\log(g(x)) = \log(e^{-x^2/2}) \quad (3.1)$$

As equações (4.2) e (3.1) foram plotadas no gráfico da Figura 13 que é um gráfico em escala log-log. O eixo y, das probabilidades, foi restringido a escala de 10^{-5} até um. Já o eixo da variável independente x foi restringido de um até cem. A *lei de potência* é expressa através de uma reta vermelha com inclinação de 1,5. A linha azul representa a distribuição gaussiana, seu comportamento declinante exponencialmente informa claramente a não-escalabilidade dessa distribuição, ao contrário da *lei de potência* que é escalável até o infinito.

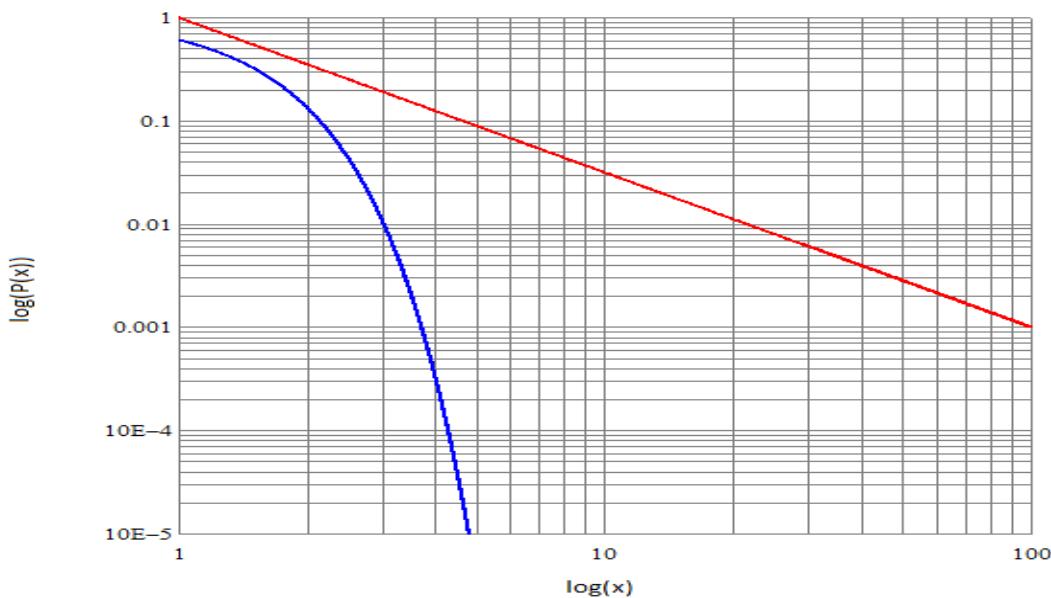


Figura 13 - Comparaçāo de uma distribuição gaussiana (linha azul) com uma distribuição seguindo uma lei de potência (linha vermelha).

Como pode ser observado, grandes valores são probabilisticamente impossíveis de ocorrer em uma distribuição gaussiana. Desta forma, para captar valores extremos é necessário outra distribuição de probabilidade, seguindo os preceitos de uma *lei de potência*, que consegue captar a ocorrência de eventos extremos.

Segundo Taleb (2007), a ocorrência de eventos extremos vem aumentando, principalmente depois da Revolução Industrial, quando o mundo começou a se tornar mais complexo. Eric Hobsbawm (1994), em seu livro *Era dos Extremos: o Breve Século XX*, descreve bem a dinâmica complexa geradora de eventos extremos, no período que se estende da Primeira Guerra Mundial em 1914 até a queda da União Soviética, em 1991. Taleb (2007) afirma também que eventos extremos podem ser subdivididos em duas categorias: passíveis de previsão ou totalmente imprevisíveis. Os totalmente imprevisíveis são chamados de cisnes negros e os passíveis de previsão de cisnes cinzentos. Assim, para o autor, alguns eventos extremos, que são raros e consequenciais, são relativamente previsíveis. Estes eventos são também chamados de quase-cisnes negros, ou de cisnes cinzentos de aleatoriedade mandelbrotiana, ou cisne negro mandelbrotiano. O que estes cisnes cinzentos possuem de diferente dos cisnes negros são uma espécie de natureza fractal, na qual é possível observar algum padrão de recorrência e assim ser realizada a previsão. Deste modo, as ocorrências de recessões econômicas devem se enquadrar na categoria dos cisnes cinzentos, dado que seja possível identificar padrões de suas ocorrências.

Nos resultados desta pesquisa, especificamente na seção 6.5, serão realizados testes estatísticos para saber até que ponto as recessões brasileiras podem ser inseridas no arcabouço teórico dos eventos extremos, ou seja, possam ser caracterizadas como cisnes cinzentos.

5 METODOLOGIA

Na introdução desta pesquisa já foi adiantado o processo metodológico para a construção do modelo preditivo aqui adotado. Lá foram citados quatro passos que serviram de balizamento para se alcançar o objetivo geral da pesquisa. O primeiro passo tratou de explicar de forma intuitiva como seria o método de definição de uma recessão. Este método de definição foi explicado no Capítulo 2, que tratou sobre recessões. Os demais esclarecimentos sobre como as recessões foram identificadas serão tratados nas duas primeiras seções do Capítulo 6, que mostrará os resultados da pesquisa, e especificamente nestas duas seções a aplicação deste método de definição de recessão para alguns exemplos de séries de produção industrial.

O segundo e o terceiro passo relatados na introdução trataram de explicar de forma intuitiva como as variáveis antecedentes foram escolhidas pelo autor, os procedimentos realizados nestas variáveis para verificar o quanto eram eficientes em anteceder os movimentos da *target variable*. Especificamente, o procedimento seguiu quatro etapas: identificação das variáveis na literatura; transformações nas variáveis; teste de correlação com a *target variable* e a de seleção final dos indicadores antecedentes através do método *stepwise*. Somando-se uma quinta etapa encontra-se uma análise *ex-post*, que é a análise gráfica das séries escolhidas para a determinação do que seria o comportamento normal de cada indicador antecedente. Estas cinco etapas serão compiladas e explicadas dentro do Capítulo 6, especificamente na seção 6.3.

Enfim, o quarto passo tratou de explicar de forma intuitiva a construção do algoritmo de previsão através da análise discriminante. Será este quarto passo que será aqui explicado no Capítulo 5, que será dividido em duas seções. A primeira, intitulada "Análise Discriminante", mostrará o que é esta análise e como se dará sua utilização nesta pesquisa. A segunda, intitulada "Aspectos Metodológicos da Previsão", mostrará como, a partir dos resultados obtidos na análise discriminante, é possível antecipar as recessões.

5.1 Análise Discriminante

Segundo Hair (1984), a análise discriminante é uma técnica estatística que pode ser utilizada para classificação de elementos de uma amostra em grupos distintos. Para sua aplicação é necessário que os grupos sejam predefinidos, ou seja, os grupos são conhecidos a priori. No caso deste estudo, os grupos são dois: recessões (zeros) e não-recessões (valores unitários). O conhecimento *a priori* dos grupos permite que se elabore uma função matemática chamada de *função discriminante*, que é utilizada para classificar novos elementos amostrais nos grupos já existentes.

Segundo Corrar et al. (2011), a análise discriminante é semelhante a regressão múltipla, porém, difere no sentido que esta última tem a característica de sua variável dependente ser métrica, enquanto na análise discriminante a variável dependente é de natureza qualitativa. Ela também difere da regressão logística, pois, na regressão logística a variável dependente pode ser classificada em função de valores contidos num intervalo fechado, compreendido entre zero e um, enquanto na análise discriminante a variável dependente só pode assumir valores inteiros.

As variáveis independentes geralmente são métricas, mas também podem assumir valores qualitativos. Estas variáveis independentes servem para diferenciar os valores qualitativos assumidos pela variável dependente. Isso é feito, como já afirmado, através de uma função matemática chamada *função discriminante*. Tal regra pode ser definida pela seguinte função linear:

$$Z_t = a + \beta_1 X_1 + \beta_2 X_2 + \dots \beta_n X_n \quad (05)$$

em que,

Z_t é o escore discriminante para uma data t no formato mês/ano;

a é o intercepto da função;

β_n é o coeficiente discriminante de cada indicador antecedente;

X_n são os valores mensais de cada indicador antecedente.

A Equação (05), acima, é uma combinação linear de variáveis independentes que resumem suas informações em um único índice, o escore discriminante Z_t . Os coeficientes discriminantes β_n são estimados com o intuito de separar cada elemento da amostra nos grupos pré-determinados. Se a função discriminante servir para distinguir, por exemplo, se determinado mês é de recessão ou de não-recessão, os valores de Z_t devem diferir entre um mês de recessão e um mês de não-recessão. Assim, os valores dos coeficientes discriminantes são escolhidos para que consigam diferenciar cada mês da forma mais eficiente possível (HAIR, 1984).

O software estatístico SPSS® ao realizar a análise discriminante gera duas tabelas de coeficientes discriminantes, uma com os coeficientes padronizados e outra com os coeficientes não-padronizados e usa os coeficientes não-padronizados para calcular a função discriminante. Os coeficientes não-padronizados são os multiplicadores das variáveis quando estas estão expressas em suas unidades originais. Os coeficientes padronizados são utilizados quando as variáveis explicativas estão padronizadas para uma média zero e desvio padrão um. Dessa forma, os coeficientes não-padronizados são calculados para amenizar as volatilidades de cada série, enquanto que os coeficientes padronizados são calculados já levando em conta a padronização das variáveis (NORUSIS, 1990).

A interpretação dos coeficientes é similar a interpretação dos coeficientes da regressão múltipla. Por vezes é tentador interpretar as magnitudes dos coeficientes como indicadores da importância relativa das variáveis. Variáveis com coeficientes grandes são pensadas como maiores contribuintes para o valor da função discriminante. No entanto, a magnitude dos coeficientes não é um bom índice de importância relativa quando as variáveis diferem nas unidades em que estão sendo medidas, que é o caso das variáveis utilizadas neste estudo, mesmo no caso que forem anteriormente padronizadas (NORUSIS, 1990). Ressaltando o que já foi afirmado, os coeficientes discriminantes são estimados com o intuito de separar cada elemento da amostra nos grupos pré-determinados da forma mais eficientes possível.

Dados os coeficientes discriminantes de todos indicadores antecedentes e o valor da constante calculada pelo software é possível calcular, para cada

mês, o valor do escore discriminante Z_t . Com todos os escores apurados, calcula-se a média dos escores para cada grupo, ou seja, a média dos escores que representam o grupo das recessões e a média dos escores que representam o grupo das não-recessões. Estas médias servirão de base para elaborar a regra de classificação. A literatura sobre análise discriminante costuma classificar essa média com o nome de centroide. Dessa forma, a média dos escores do grupo das recessões Z_0 chama-se centróide dos grupos das recessões e a média dos escores do grupo das não-recessões Z_1 chama-se centróide do grupo das não-recessões. Estes valores serão utilizados para classificar as observações nos seus determinados grupos, de recessão ou de não-recessão.

Segundo Corrar et al. (2011), se o número de observações de cada grupo fossem iguais, a regra de classificação seria apenas somar os centroides e dividi-los por dois. Achar-se-ia assim o escore crítico ou ponto de corte Z_c que serviria para classificar novos elementos. Porém, como no caso deste estudo, o número de observações de cada grupo difere entre si. Sendo assim, é necessária uma ponderação, ou seja, uma atribuição de peso as centroides em função do número de observações de cada grupo, através da seguinte fórmula:

$$Z_c = \frac{N_0 Z_1 + N_1 Z_0}{N_0 + N_1} \quad (06)$$

em que,

Z_c = valor do escore crítico ou ponto de corte;

N_0 = número de observações do grupo das recessões;

N_1 = número de observações do grupo das não-recessões;

Z_0 = centróide do grupo das recessões;

Z_1 = centróide do grupo das não-recessões.

A lógica é que valores maiores que o escore crítico sejam classificados no grupo das não-recessões e que valores menores que o escore crítico sejam classificados no grupo das recessões.

Outra regra de classificação, baseada em Corrar et al. (2011, p.263), consiste em comparar os escores de cada observação com suas distâncias euclidianas quadráticas³⁰ em relação aos centroides de cada grupo. A decisão de classificação para cada observação será feita escolhendo-se a menor distância euclidiana quadrática em relação às centroides de cada grupo, ou seja, a distância euclidiana quadrática capta apenas a proximidade de cada escore discriminante com os referidos centroides. A regra de decisão então será:

$$Classificação = \begin{cases} 0 & \text{se } (Z_t - Z_0)^2 < (Z_t - Z_1)^2 \\ 1 & \text{se } (Z_t - Z_1)^2 < (Z_t - Z_0)^2 \end{cases} \quad (7)$$

em que:

Z_t = escore discriminante para a data t no formato mês/ano;

Z_0 = centróide do grupo das recessões;

Z_1 = centróide do grupo das não-recessões;

$(Z_t - Z_0)^2$ = distância euclidiana quadrática para o grupo das recessões;

$(Z_t - Z_1)^2$ = distância euclidiana quadrática para o grupo das não-recessões.

Estas regras de classificação aqui mostradas podem cometer erros, que são os erros Tipo I e Tipo II. No erro Tipo I rejeita-se H_0 quando esta hipótese³¹ é verdadeira, ou seja, quando deve ser aceita. O erro Tipo II, por sua vez,

³⁰ Em matemática, distância euclidiana (ou distância métrica) é a distância entre dois pontos que pode ser calculada como $\sqrt{(x - y)^2}$ para o caso de pontos unidimensionais. A distância euclidiana quadrática retira da formula anteriormente apresentada a raiz quadrada, ficando apenas como $(x - y)^2$. Quando a distância se trata entre pontos unidimensionais essa mesma distância euclidiana pode ser chamada de distância de Mahalanobis, como é o caso deste estudo.

³¹ A hipótese nula a ser testada é de um mês qualquer ser de recessão.

consiste em aceitar H_0 quando ela não é verdadeira. A próxima seção utilizará destes dois tipos de erro para mostrar como a análise discriminante será útil para construir o modelo final de previsão. Lá será mostrado, por exemplo, que dependendo do local que apareça o erro Tipo II ele pode não ser considerado um falso alarme como é considerado dentro da análise discriminante.

Um teste simples para medir o poder de previsão do algoritmo é o proposto por Diebold e Rudebusch (1989), chamado *Quadratic Probability Score* (QPS), que é análogo ao erro médio quadrado da estimativa. Segue abaixo sua equação:

$$QPS = \frac{1}{T} \sum_{t=1}^T 2(P_t - R_t)^2 \quad (08)$$

Em que P_t é a probabilidade de ocorrência de uma recessão para a data t e R_t a realização desse evento; P_t e R_t assumem o valor 1 se a recessão ocorre e 0 caso contrário. A variável T representa o total de eventos possíveis a serem previstos.

Dessa forma, toda vez que a previsão da análise discriminante cometer uma erro Tipo I ou Tipo II o valor de $(P_t - R_t)^2$ será unitário, e caso exista um acerto na previsão o valor de $(P_t - R_t)^2$ será zero. O valor do QPS se encontra no intervalo de zero a dois, de modo que quanto mais próximo de zero, maior é o poder preditivo do algoritmo.

5.2 Aspectos Metodológicos da Previsão

O objetivo desta seção é apresentar a explicação acerca do procedimento adotado para prever a *target variable* com uma antecedência de seis meses. A Tabela 5 abaixo mostra valores hipotéticos para a construção de escores discriminantes. Para facilitação da análise, adotou-se o ponto de corte Z_c como sendo zero. Valores de Z_t maiores que zero serão classificados como

períodos de não-recessão e valores menores como períodos de recessão. Como já afirmado, o valor unitário indica uma não-recessão e o zero indica uma recessão. A coluna *Dummy* indica os meses de recessão observados e a coluna previsão indica as previsões com base na regra de classificação do ponto de corte. O valor da constante α é o mesmo para todos os meses. Os valores dos coeficientes discriminantes β 's não foram adicionados na Tabela 5, já que para entendimento do procedimento preditivo eles não assumem grande importância. Porém, os valores das variáveis X_n multiplicadas pelos seus respectivos coeficientes discriminantes são importantes para o entendimento da previsão e são listados na tabela abaixo.

Tabela 5 - Valores de Escores Discriminantes Hipotéticos

Mês	α	$\beta_1 X_1$	$\beta_2 X_2$	$\beta_3 X_3$	$\beta_4 X_4$	$\beta_5 X_5$	$\beta_6 X_6$	$\beta_7 X_7$	$\beta_8 X_8$	$\beta_9 X_9$	Z_t	Previsão	Dummy
Jan	2	1	-3	-2	1	-2	3	2	-1	1	2	1	1
Fev	2	1	3	-4	-2	5	1	-3	-3	1	1	1	1
Mar	2	1	3	4	-3	-2	-4	-3	1	-3	-4	0	0
Abr	2	3	4	-3	-3	4	-4	-4	1	-4	-4	0	0
Mai	2	-2	5	2	-3	-4	3	-1	-1	-3	-2	0	0
Jun	2	1	2	2	-3	-3	2	-1	-2	-3	-3	0	0
Jul	2	1	2	-4	-5	3	-2	1	-2	2	-2	0	0
Ago	2	2	-2	3	-2	-2	-1	-2	2	-3	-3	0	0
Set	2	2	-1	-2	1	3	-2	3	-3	1	4	1	1
Out	2	1	-2	3	3	2	3	-5	-2	1	2	1	1

Fonte: Elaboração própria.

Como pode ser observado, é possível construir para cada mês uma função discriminante e seu respectivo escore Z_t . Apesar de ser um exemplo hipotético, as variáveis independentes estão dispostas de acordo com suas defasagens ótimas, mostradas na Tabela 16 da subseção 6.3.1, obtida através dos resultados desta pesquisa. Como a *Dummy* indica os períodos recessivos observados, a previsão hipotética acertou todos os casos, inclusive o mês inicial e o mês final do período recessivo, que começa em março e vai até agosto.

Como essas variáveis antecedentes possuem defasagens que variam de seis a doze meses, a antecipação máxima de uma recessão só pode ser de

seis meses. Em outras palavras, seis meses antes do início da recessão, que começou em março, todos os dados das nove variáveis antecedentes já estariam disponíveis. A Figura 14 abaixo exemplifica a antecipação de uma recessão para os dados da Tabela 5:

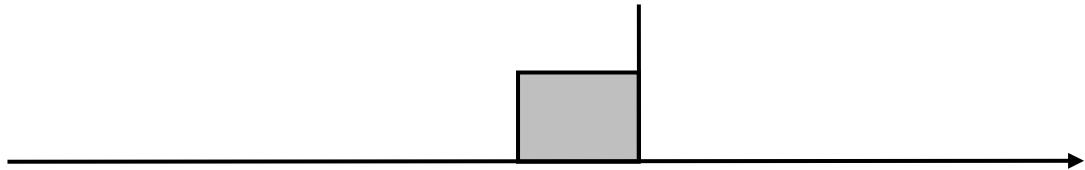


Figura 14 - Alarme referente aos dados da Tabela 5. O retângulo cinza representa o alarme e a barra vertical representa o início da recessão.

O retângulo cinza possui um comprimento de seis meses. Ele é deslocado seis meses antes, assim, o início do alarme se dá em setembro do ano anterior. Caso os dois últimos meses da Tabela 5 também fossem meses de recessões e que o modelo preditivo tenha acertado a classificação, a duração total do período recessivo seria de oito meses. Nesse caso, a Figura 15 abaixo exemplifica como ficaria o alarme:

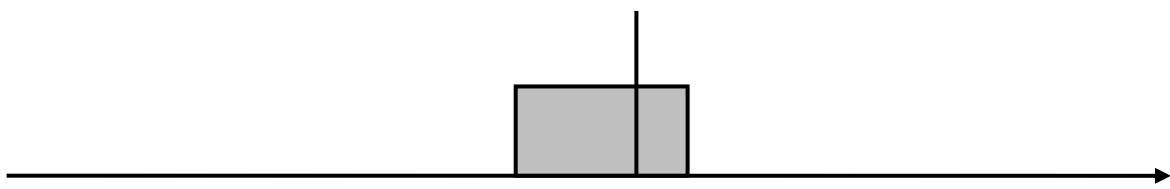


Figura 15 - Alarme referente a uma recessão de oito meses. O retângulo cinza representa o alarme e a barra vertical representa o início da recessão.

O retângulo cinza possui um comprimento de oito meses. Ele é deslocado seis meses antes, assim, o início do alarme também se dá em Setembro do ano anterior. O que difere da Figura 14 é que, como é uma recessão de oito meses, o alarme continua a soar por dois meses após o início da recessão, soando até o mês de maio.

Ainda analisando a Tabela 5, agora supondo que a previsão erre para os meses de janeiro e fevereiro, classificando-os como recessivos. Nesse caso, existirá um *falso alarme da análise discriminante* de dois meses. Como a

hipótese nula a ser testada é a de um mês qualquer ser de recessão, esse tipo de erro é caracterizado como um erro do Tipo II, que consiste em aceitar H_0 quando ela não é verdadeira. O erro Tipo II pode aparecer de três formas diferentes. Apenas uma destas três formas é prejudicial para a previsão, fato este que será explicado mais a frente. A Figura 16 abaixo mostra o alarme da recessão para a Tabela 5 supondo que para os meses de Janeiro e Fevereiro a previsão cometeu erros do Tipo II:

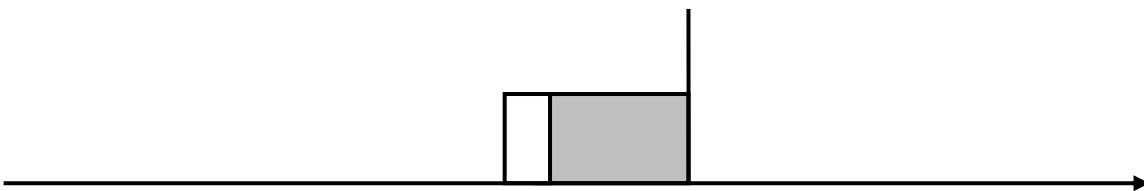


Figura 16 - Alarme referente aos dados da Tabela 5. O retângulo todo representa o alarme. A parte branca representa o erro Tipo II. A barra vertical representa o início da recessão.

O retângulo branco representa um comprimento de dois meses e o retângulo cinza um comprimento de seis meses. Neste caso, o erro do Tipo II não foi prejudicial à previsão, já que ele está junto de alarmes verdadeiros e fez o horizonte temporal da previsão aumentar em dois meses, totalizando oito meses. A condição para aceitá-los como parte do alarme é que seu comprimento seja de no máximo seis meses, já que somando os seis meses do erro Tipo II mais os seis meses dos acertos da previsão (retângulo cinza) seriam doze meses de alarme, que é o máximo aceitável como horizonte temporal de previsão. Acima deste horizonte será considerado que o algoritmo estará gerando falsas sinalizações de recessão, ou seja, *falsos alarmes preditivos*.

Dada tais informações, cabe aqui fazer a diferenciação entre o *falso alarme da análise discriminante* (erro Tipo II) e um *falso alarme preditivo*. Este último é o que realmente interessa neste estudo. São duas condições para sua existência. A primeira é ele ser um *falso alarme* da análise discriminante, que seria a condição necessária, mas não suficiente. A segunda é que esse falso alarme discriminante esteja ao menos doze meses antes do início de uma recessão. Em outras palavras, caso soe um alarme e a recessão não apareça

em ao menos doze meses, ele será considerado *um falso alarme preditivo*. A Figura 17 exemplifica a ocorrência dos três tipos do erro Tipo II:

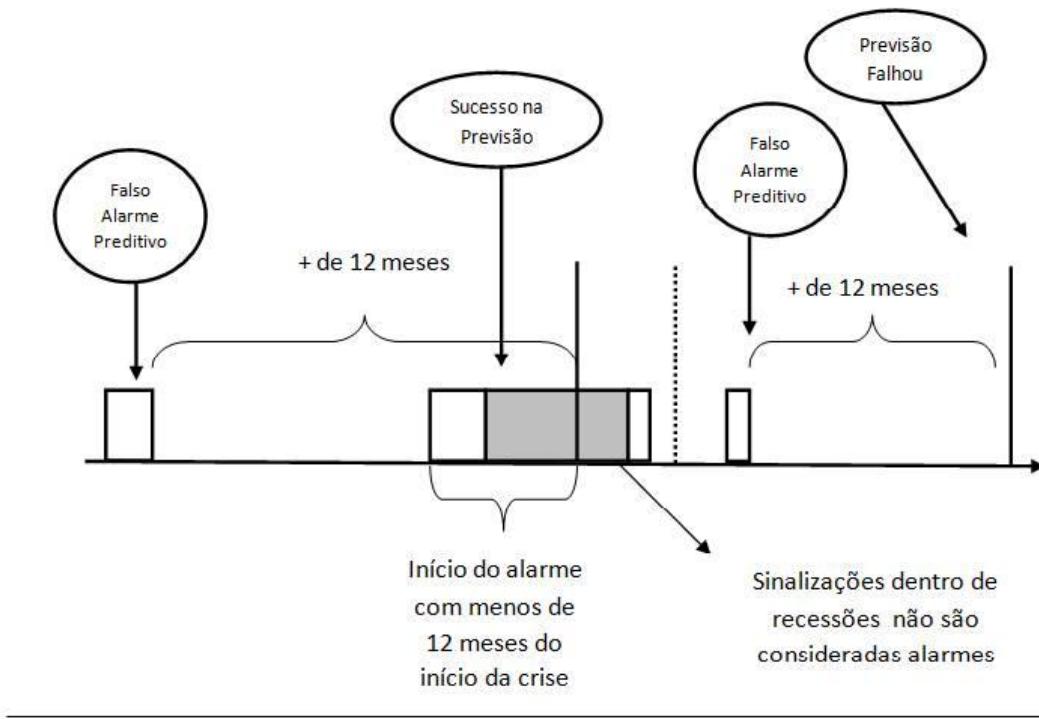


Figura 17 - Possíveis tipos de previsão. As barras verticais contínuas representam o início de uma recessão e a barra vertical pontilhada representa o final de uma recessão. Os quatro retângulos brancos representam erros Tipo II. O retângulo branco somado com o retângulo cinza que aparece antes do início da recessão indica um alarme que obteve sucesso na previsão. O retângulo cinza somado com o retângulo branco que aparece dentro da recessão não são considerados alarmes.

Como pode ser observado, o erro Tipo II, ou seja, o *falso alarme da análise discriminante* aparece em quatro casos diferentes. Apenas nos casos das extremidades é que são considerados *falsos alarmes preditivos*, já que indicaram recessões sem que estas viessem a ocorrer num período de doze meses.

O segundo retângulo branco antecede o início do alarme, assim ele não é considerado um *falso alarme preditivo*, mas apenas um *falso alarme da análise discriminante* ou um erro Tipo II. O terceiro retângulo branco não é considerado um *falso alarme preditivo*, pois sinalizações que ocorram dentro das recessões não são consideradas alarmes.

O último retângulo branco caso esteja deslocado seis meses do final da recessão (tempo mínimo que a literatura³² identifica entre uma recessão e outra) e caso, depois desse retângulo branco não apareça uma recessão em no mínimo doze meses, ele será considerado um *falso alarme preditivo*. Se esse retângulo viesse aparecer apenas quatro meses depois do final da recessão ele não seria considerado um alarme.

Ainda deve ser considerado o erro Tipo I, que ocorre rejeitando-se H_0 quando essa hipótese é verdadeira. A Figura 17 exemplifica sua ocorrência através da segunda barra vertical contínua, que ocorre uma recessão e nenhum alarme é disparado ao menos doze meses antes. Em termos práticos, resumindo ao que realmente interessa no processo preditivo e abstraindo os erros Tipo I e Tipo II da análise discriminante, só existem três possibilidades de previsão, são elas:

- a) Alarme falso preditivo: o alarme soar e a recessão não acontecer em ao menos 12 meses;
- b) Sucesso na previsão: o alarme soar e a recessão ocorrer em até doze meses;
- c) Falha na previsão: ocorrer a recessão e nenhum alarme ter soado pelo menos 12 meses antes.

³² Ver por exemplo os relatórios do CODACE em <http://portalibre.gfv.br>

6 RESULTADOS E DISCUSSÕES

6.1 Análise de janeiro de 1975 até dezembro de 2012

O objetivo principal desta seção é justificar o uso do filtro polinomial como um bom datador de *business cycles* e ao mesmo tempo justificar o uso da variável de produção industrial como uma boa série coincidente para a atividade agregada. Para se alcançar esses objetivos serão feitas comparações entre os *business cycles* da produção industrial datados pelo filtro polinomial e os *business cycles* da atividade agregada datados pelo CODACE. Não é objetivo desta seção determinar os alvos de previsão, estes serão construídos apenas na próxima seção 6.2.

A principal justificativa para o uso da produção industrial como a variável alvo a ser prevista (*target variable*) em detrimento do PIB é a disponibilidade dos dados. Mensalmente o PIB só está disponível a partir de 1991. A série de produção industrial brasileira possui uma abrangência maior, começando mensalmente desde o ano de 1975. Apesar da análise de previsão que aqui será construída só se iniciar a partir de 1991, ainda assim a de produção industrial será utilizada como a *target variable*, dado que, como será visto mais a frente, esta série tem a característica de apresentar uma maior sensibilidade a responder antes às variações de conjuntura. Quanto a justificativa de utilizar um filtro para datar as recessões em detrimento do algoritmo sugerido por Bry e Boschan (1971), ela será apresentada ao final da seção 6.2, quando forem apresentadas as conclusões do método de datação.

Dessa forma, datou-se as recessões brasileiras a partir da análise da série de produção industrial³³, que é considerada na literatura como uma série coincidente da atividade econômica agregada (ver, por exemplo, Issler e Hollauer, 2006 ou Issler et al., 2013). Os dados desta datação começam em 1975, com intuito de dar maior robustez as justificativas, ou seja, de mostrar

³³ Disposta mensalmente, dessazonalizada, transformada em índice com a média de 2002 = 100, acompanhada pelo IBGE através da PIM-PF. A série vai de janeiro de 1975 até dezembro de 2012.

que os momentos recessivos da atividade agregada e da produção industrial são semelhantes não só a partir de 1991, mas desde 1975.

A Figura 18 mostra os valores filtrados pela suavização polinomial para a série mensal da produção industrial. As áreas delimitadas por linhas verticais correspondem às datação dos *business cycles* para a atividade agregada brasileira de acordo com o CODACE, já as áreas hachuradas correspondem às datações dos *business cycles* da produção industrial datadas pelo autor pelo filtro polinomial.

As Tabelas 6 e 7 mostram os *turning points* que delimitam os *business cycles* das respectivas datações (atividade agregada e produção industrial), bem como as durações das recessões em números de meses. Na Tabela 6 encontra-se a numeração de referência de cada recessão datada pelo CODACE, numeração esta que servirá de base para citações no decorrer deste estudo.

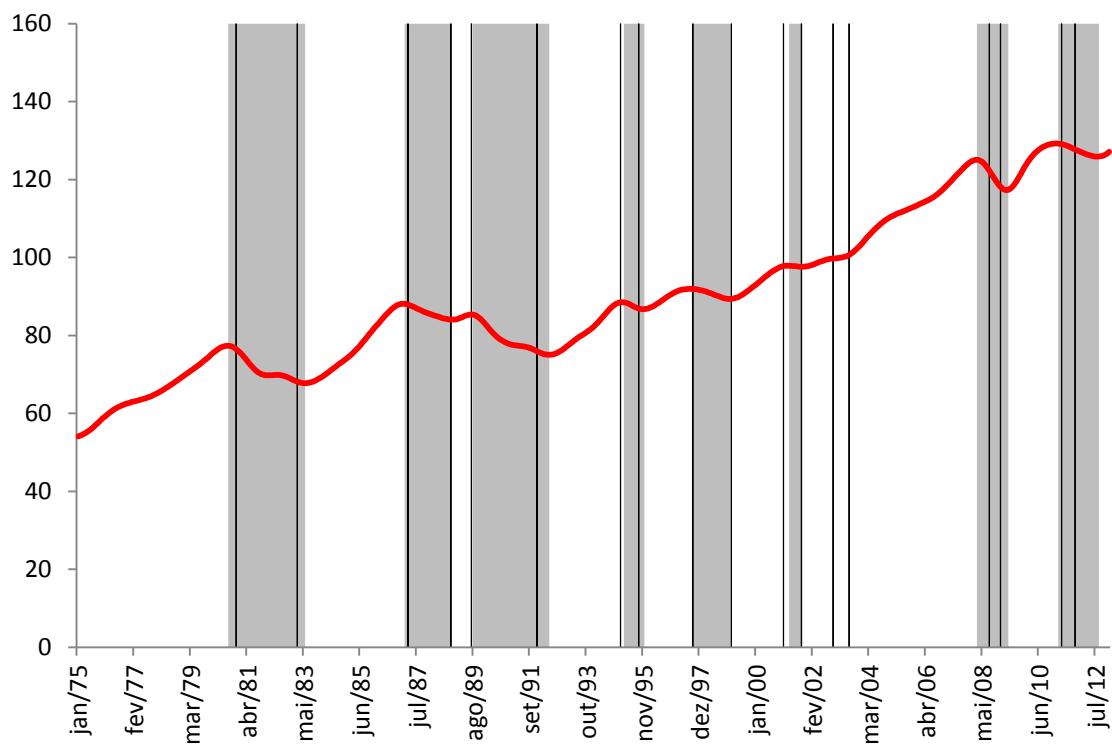


Figura 18 - Valores filtrados da produção industrial brasileira (linha vermelha) de 1975 a 2012; datação de seus *business cycles* (áreas hachuradas) e a datação dos *business cycles* pelo CODACE para a atividade agregada (linhas verticais).

Tabela 6 - *Turning points* dos *business cycles* da atividade agregada brasileira datados pelo CODACE - 1975 a 2012

Recessão	Picos	Vales	Meses em recessão
1	Outubro/1980	Fevereiro/1983	28
2	Fevereiro/1987	Outubro/1988	20
3	Junho/1989	Dezembro/1991	30
4	Dezembro/1994	Setembro/1995	9
5	Outubro/1997	Fevereiro/1999	16
6	Dezembro/2000	Setembro/2001	9
7	Outubro/2002	Junho/2003	8
8	Julho/2008	Janeiro/2009	6
9	Março/2011	Outubro/2011	7
Duração total de meses			133

Fonte: CODACE (2013).

Tabela 7 - *Turning points* dos *business cycles* da produção industrial brasileira datados pelo filtro polinomial³⁴ - 1975 a 2012

Recessão	Picos	Vales	Meses em recessão
1	Julho/1980	Maio/1983	34
2	Janeiro/1987	Outubro/1988	21
3	Julho/1989	Maio/1992	34
4	Fevereiro/1995	Novembro/1995	9
5	Agosto/1997	Janeiro/1999	17
6	Março/2001	Setembro/2001	6
7	n/d	n/d	n/d
8	Fevereiro/2008	Abril/2009	14
9	Fevereiro/2011	Agosto/2012	18
Duração total em meses			153

Fonte: Elaboração própria.

³⁴ A recessão de número sete não foi captada na série de produção industrial.

Abstraindo os aspectos metodológicos de cada datação, pode-se inferir algumas conclusões sobre a comparação das duas datações. Como observado, os períodos de recessão na indústria mostraram-se mais prolongados do que os da atividade agregada. As recessões na indústria duraram vinte meses a mais no total, diferença não muito expressiva se for considerado o horizonte de tempo estudado, com 456 meses. Os meses iniciais de cada recessão foram relativamente semelhantes, apresentando uma diferença máxima de cinco meses na recessão de número oito. A Tabela 8 abaixo sintetiza tais diferenças. Valores negativos indicam que a recessão na produção industrial precedeu a recessão da atividade agregada, enquanto que os valores positivos indicam uma postergação do seu início.

Tabela 8 - Diferenças mensais entre o início das recessões na indústria e o início das recessões na atividade agregada brasileira - 1975 a 2012

Recessão	Diferenças
1	-3
2	-1
3	0
4	+2
5	-2
6	+3
7	n/d
8	-5
9	-1

Fonte: Elaboração própria.

Na média, as recessões na indústria tendem a aparecer um pouco antes do que na atividade agregada³⁵. Apenas as recessões quatro e seis fogem a essa regra, aparecendo dois e três meses depois, respectivamente. A recessão número sete não foi captada na série da produção industrial. A linha suavizada da produção apresentada na Figura 18 permanece horizontal neste período. De

³⁵ Este fato faz com que o método de previsão aqui construído se torne mais eficiente em prever recessões na atividade agregada. Quando forem realizadas a previsões, o horizonte temporal destas em relação a atividade agregada podem ser aumentados observado-se os dados da Tabela 8.

fato, esta recessão foi bastante curta, com apenas oito meses e também pouco intensa, justificando o fato dela não ter sido captada.

Duas observações importantes devem ser enfatizadas. A primeira, que os momentos iniciais das recessões são bastante semelhantes com os da atividade agregada. A segunda, que de todas as recessões datadas na série da produção industrial, nenhuma representou um falso sinal de recessão da atividade agregada medida pelo CODACE. Tal informação é de bastante utilidade para o modelo de previsão, já que caso soe um alarme para uma recessão industrial, pode-se concluir que este alarme também servirá para uma recessão da atividade agregada. Portanto, concluiu-se que a série de produção industrial é um bom indicador coincidente da atividade agregada e seus respectivos *turning points* são bem captados através da método de suavização polinomial.

Assim, vistos tais resultados, este trabalho decidiu utilizar o filtro polinomial para datação dos *business cycles* da atividade industrial brasileira. Sua vantagem decorre, primeiramente, porque o mesmo procedimento de datação pode ser aplicado em larga escala para centenas de séries candidatas a antecedentes de forma rápida e eficaz. Porém, o motivo mais importante decorre de que datar os *business cycles* com as séries em nível (sem utilizar qualquer tipo de filtro) não produz uma nova série com seus picos e vales bem definidos. O método do algoritmo de Bry e Boschan (1971) apesar de ser bastante intuitivo não gera ao seu final uma série suavizada indicando claramente seus *turning points*, como faz a datação pelo filtro polinomial. O método sugerido por Bry e Boschan apenas indica uma rotina de procedimentos e manipulações para o pesquisador conseguir captar os devidos *turning points*.

A série suavizada, gerada pelo filtro polinomial, é de grande utilidade não só para a construção da *target variable*, mas também para os indicadores antecedentes que serão utilizados no modelo de previsão. O motivo é que os valores filtrados dos mais diversos indicadores antecedentes são transformados em taxas de crescimento e quando a taxa de crescimento da série suavizada estiver negativa a série realmente estará em um período de recessão. Notar que para os indicadores antecedentes não se faz necessário uma definição precisa de número de meses que as taxas de crescimento

devem estar negativas. Só fato dos valores filtrados entrarem em declínio já transmite uma importante informação de um comportamento anormal da série.

Caso fossem utilizadas as séries em nível e calculadas suas taxas de crescimento, ter-se-ia vários valores negativos de taxa de crescimento sem que necessariamente esses valores representassem uma recessão do indicador antecedente. Nota-se, neste sentido, que o papel do filtro é retirar as sazonalidade da série e mostrar os períodos de queda só quando estes realmente forem significativos.

Uma saída, caso não se usasse o filtro, seria criar variáveis *dummies* para os momentos de expansão e recessão dos indicadores antecedentes, porém se perderia o grau de intensidade entre as diversas recessões, o que não acontece quando se utilizada dos valores das taxas de crescimento. Esse fato é bastante importante, talvez central, na construção do modelo de previsão, visto que são estes valores, dos indicadores antecedentes, que servirão para discriminar se determinado mês será ou não de recessão.

6.2 Análise de janeiro de 1991 até dezembro de 2012

O objetivo principal desta seção é datar os *business cycles* para a série de produção industrial entre 1991 e 2012. As recessões encontradas nesta datação serão os alvos de previsão do modelo que mais a frente será construído. A justificativa de fazer uma outra datação, além da já realizada na seção anterior, é que houve uma mudança na série de produção industrial que, desde janeiro de 1991, reformulou-se para adequar-se a nova metodologia da PIM-PF.

O objetivo secundário desta seção é justificar, para o período aqui proposto, assim como realizado na seção anterior, o uso da série de produção industrial como sendo coincidente para com o nível de atividade agregada, bem como justificar o uso do filtro polinomial para as datações. Para se alcançar esses objetivos serão feitas comparações entre os *business cycles* da

produção industrial datados pelo filtro polinomial e os *business cycles* da atividade agregada datados pelo CODACE. Também será feita uma comparação dos *business cycles* da produção industrial com os *business cycles* de um índice coincidente para a atividade agregada, construído por Issler et al. (2013).

Por fim, com o intuito de adicionar mais uma justificativa ao uso do filtro polinomial como um bom datador de *business cycles*, será feita uma comparação entre as datações dos *business cycles* da atividade agregada americana realizadas pelo NBER com as datações dos *business cycles* da produção industrial americana realizadas pelo filtro polinomial.

Como o intuito principal deste estudo é fazer uma análise de previsão das recessões através de variáveis consideradas como antecedentes, utilizou-se a série de produção industrial de janeiro de 1991 até dezembro de 2012 em detrimento da série completa que começa desde 1975, pois muitas das séries utilizadas como indicadores antecedentes no modelo preditivo só se encontravam disponíveis a partir de janeiro de 1991.

Dado isso, com o intuito de analisar se a série de produção industrial é uma boa coincidente para o nível de atividade agregada, a Figura 19 mostra a evolução da produção industrial e do PIB brasileiro entre janeiro de 1991 e dezembro de 2012.³⁶

³⁶ As séries foram transformadas em índice média 2002 = 100 e dessazonalizadas. A série do PIB foi transformada em valores reais através do IGP-DI.

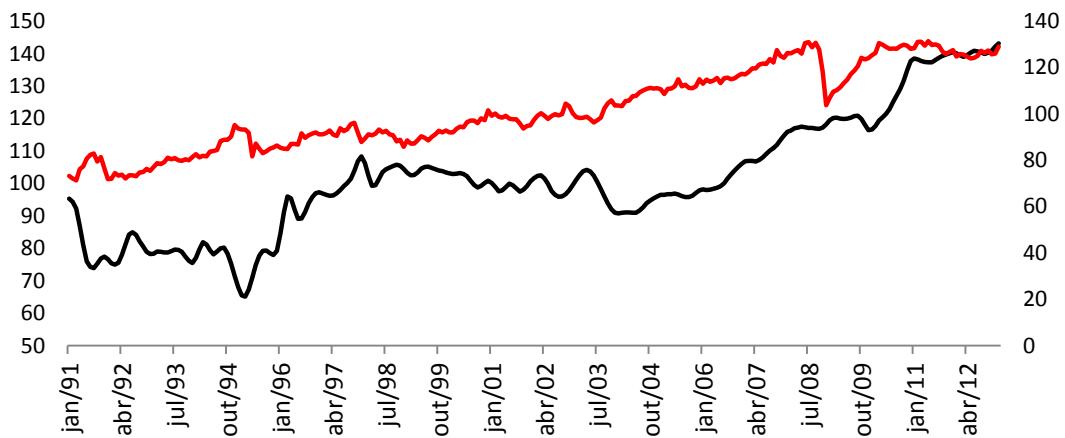


Figura 19 - Dados brutos da série de produção industrial brasileira (linha vermelha) e da série do PIB brasileiro (linha preta). Os valores do eixo esquerdo são para o PIB e os do eixo direito para a produção industrial.

Verificou-se uma correlação linear da ordem de 0,881 através do Coeficiente Pearson. Ou seja, é um primeiro indicativo que, para este período, os *turning points* das séries possam ser semelhantes. Dado isso, calculou-se os *business cycles* da série de produção industrial através do filtro polinomial.

A Figura 20 mostra os dados suavizados pelo filtro polinomial para a série de produção industrial. As recessões da produção industrial são delimitadas pelas áreas hachuradas e as recessões da atividade agregada datadas pelo CODACE são delimitadas pelas linhas verticais. Os *turning points* destas duas datações podem ser visualizados através das Tabelas 9 e 10, referentes a atividade agregada e produção industrial, respectivamente³⁷.

³⁷ A Tabela 9, referente a atividade agregada, é apenas uma adaptação da Tabela 6 (apresentada na seção anterior) para dos dados desta presente seção.

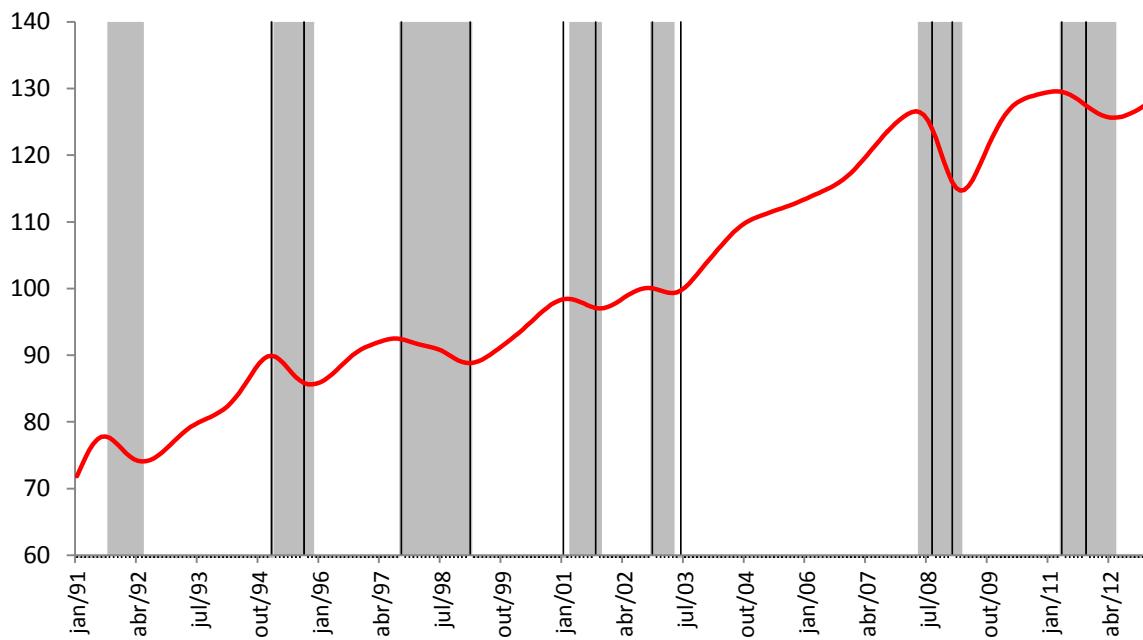


Figura 20 - Valores filtrados da produção industrial brasileira (linha vermelha) de 1991 a 2012; datação de seus *business cycles* (áreas hachuradas) e a datação dos *business cycles* pelo CODACE para a atividade agregada (linhas verticais).

Tabela 9 - *Turning points* dos *business cycles* da atividade agregada brasileira datados pelo CODACE³⁸ - 1991 a 2012

Recessão	Picos	Vales	Meses em recessão
3	Janeiro/1991	Dezembro/1991	11
4	Dezembro/1994	Setembro/1995	9
5	Outubro/1997	Fevereiro/1999	16
6	Dezembro/2000	Setembro/2001	9
7	Outubro/2002	Junho/2003	8
8	Julho/2008	Janeiro/2009	6
9	Março/2011	Outubro/2011	7
Duração total de meses			66

Fonte: CODACE (2013)

³⁸ Como pode ser observado na Tabela 6, a Recessão 3 teve início em junho de 1989, porém, como a análise desta seção é a partir de 1991, datou-se seu início em Janeiro de 1991.

Tabela 10 - *Turning points* dos *business cycles* da produção industrial brasileira datados pelo filtro polinomial - 1991 a 2012

Recessão	Picos	Vales	Meses em recessão
3	Agosto/1991	Maio/1992	9
4	Janeiro/1995	Novembro/1995	10
5	Agosto/1997	Fevereiro/1999	18
6	Fevereiro/2001	Outubro/2001	8
7	Outubro/2002	Abril/2003	6
8	Abril/2008	Março/2009	11
9	Março/2011	Maio/2012	14
Duração total de meses			76

Fonte: Elaboração própria a partir dos resultados da pesquisa.

Levando-se em conta que as comparações entre as duas datações serão realizadas observando-se os meses iniciais das recessões, a recessão três não deve sofrer esta comparação, visto que, na datação do CODACE esta se inicia realmente em junho de 1989, como pode ser observado na Tabela 6. Devido a tal fato ela não foi marcada pelas linhas verticais na Figura 20. Na Tabela 7, referente a datação da produção industrial desde 1975, a recessão três se inicia em julho de 1989 e termina em maio de 1992. Na datação realizada na série de produção industrial começando em 1991, Tabela 10, o filtro polinomial ainda conseguiu captar nove meses desta recessão.

Antes de iniciar a comparação dos meses iniciais das recessões, cabe a apresentação da Figura 21, que mostra os *business cycles* de um indicador coincidente da atividade agregada em forma de índice, que foi construído a partir da composição de quatro séries representativas dos seguintes agregados: emprego, renda, produção e de vendas. As áreas hachuradas representam os momentos recessivos do índice. Ele é fruto de um estudo de Issler et al., 2013, e os valores de seus *turning points* podem ser visualizados através da Tabela 11.

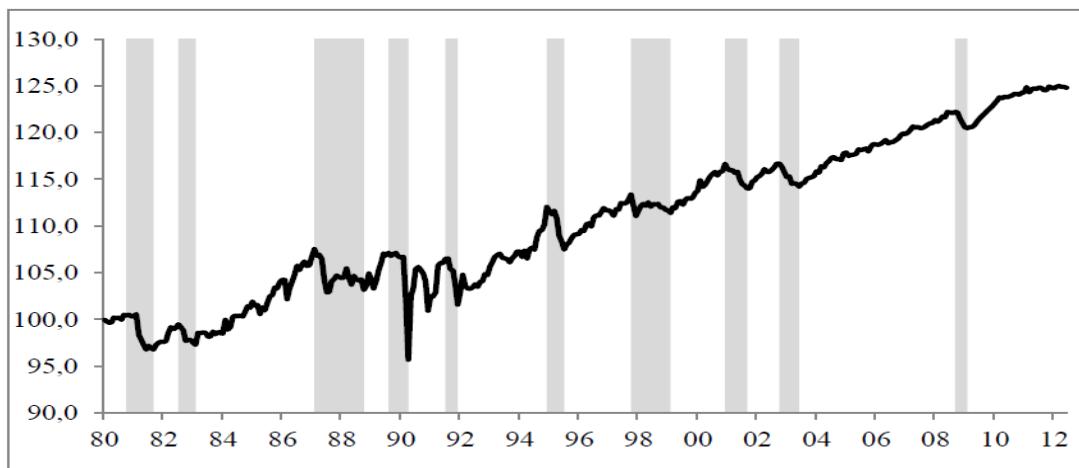


Figura 21 - Índice coincidente composto por quatro séries: emprego, produção industrial, renda, vendas - 1991 a 2012.

Tabela 11 - *Turning points* dos *business cycles* de um índice coincidente composto por quatro séries: emprego, produção, renda, vendas - 1991 a 2012

Recessão	Picos	Vales	Meses em recessão
3	Julho/1991	Dezembro/1991	6
4	Dezembro/1994	Julho/1995	8
5	Outubro/1997	Fevereiro/1999	17
6	Dezembro/2000	Setembro/2001	10
7	Outubro/2002	Junho/2003	9
8	Setembro/2008	Fevereiro/2009	6
9	n/d	n/d	n/d
Duração total de meses			56

Fonte: Isller et al. (2013).

Como pode ser observado, os autores que construíram o índice coincidente não identificaram a recessão de número nove. O motivo talvez esteja na composição deste índice, que dado o elevado nível de emprego que a economia brasileira se encontrou nos anos de 2011 e 2012, o índice não tenha apresentado uma queda significativa, mesmo com a produção industrial em decadência.

Dadas estas informações sobre as três datações aqui relatadas, a Tabela 12 resume o que se pode extrair de mais importante nesta seção. Ela compara os *business cycles* da produção industrial com os *business cycles* da atividade agregada e com os *business cycles* do índice coincidente em termos dos momentos iniciais das recessões. Valores negativos indicam que a recessão na produção industrial precedeu a recessão da atividade agregada, enquanto que os valores positivos indicam uma postergação do seu início.

Tabela 12 - Diferenças mensais entre o início das recessões na indústria com os inícios das recessões na atividade agregada brasileira - 1991 a 2012.

Recessão	Diferenças com as recessões do CODACE	Diferenças com as recessões do índice coincidente
3	-	-1
4	-1	-1
5	-2	-2
6	-2	-2
7	0	0
8	-4	-6
9	0	n/d

Fonte: Elaboração própria.

Analizando a Tabela 12, primeiramente em relação ao índice coincidente, nota-se que o início da recessão número três difere em apenas um mês da recessão na indústria. Cabe aqui ressaltar que tal índice coincidente "dividiu" a recessão três em duas partes. Observando a Figura 21 é possível notar essa descontinuidade (entre a área hachurada de 1990 e a área hachurada de 1992). Este fato não é um bom indicativo para uma série coincidente, visto que sinaliza uma possível nova recessão para a atividade agregada que na verdade já estaria chegando ao seu fim. Ademais, as recessões da atividade industrial precederam todas as recessões datadas pelo índice coincidente, exceto pela recessão sete que começou no mesmo mês. A recessão nove não pode ser comparada, visto que não foi captada pelo índice coincidente.

Em relação as datações do CODACE, as recessões quatro, cinco, seis, sete e oito surgiram antecipadamente na atividade industrial. As exceções foram as recessões sete e nove, que começaram no mesmo mês/ano das recessões do CODACE. Essa antecipação das recessões pode se dar tanto pelo fato das recessões na indústria começarem realmente um pouco antes do que as recessões da atividade agregada, por ser uma atividade mais sensível as mudanças conjunturais, ou pelo fato da metodologia de datação diferir da metodologia utilizada pelo CODACE.

Concluindo e resumindo o que já foi relatado, a série que será usada como *target variable* será a da produção industrial brasileira de janeiro de 1991 até dezembro de 2012, fornecida pelo IBGE através da PIM-PF. A série foi dessazonalizada e transformada em índice fazendo a média de 2002 = 100. Os valores desta série foram filtrados pelo filtro polinomial, produzindo 76 meses de recessão e 188 meses de expansão. Desta forma, as recessões identificadas na Figura 20 e na Tabela 10 serão os alvos de previsão deste estudo, totalizando sete recessões³⁹. Definidos os alvos de previsão, a próxima seção deixará claro o método para escolha dos indicadores antecedentes.

Com o intuito de adicionar mais uma justificativa para o uso do filtro polinomial como datador de recessões, foram datadas as recessões americanas através do filtro polinomial. O NBER disponibiliza em seu endereço online as séries temporais que utilizou para realizar a datação da última recessão (2008). Dentre estas, se encontra a série do índice da produção industrial americana, fornecida pelo *Board of Governors of the Federal Reserve System* (BGRFS). Esta série é apresentada pelo NBER desde janeiro de 1959 até julho de 2010. A Figura abaixo mostra os dados suavizados pelo filtro polinomial para a referida série⁴⁰. As áreas delimitadas por linhas verticais correspondem às datações dos *business cycles* para a atividade agregada

³⁹ Na introdução deste estudo foram relatadas seis recessões desde o início da década de noventa. Dessa forma, não foi contabilizada a recessão de número 3, visto que seu início se deu em 1989. Esta recessão de número 3 entrará, porém, no alvo de previsão do modelo, totalizando assim sete recessões desde o início da década de noventa. Mais especificamente, a *target variable* vai abranger o período entre fevereiro de 1992 até dezembro de 2012, período este que será explicado e justificado na próxima seção 6.3. Assim, a recessão três irá se resumir a apenas quatro meses dentro do modelo preditivo, visto que esta se finda em maio de 1992 e começará em fevereiro de 1992.

⁴⁰ No Apêndice deste estudo foram realizadas previsões para as recessões americanas através desta série de produção industrial.

americana de acordo com o NBER, já as áreas hachuradas correspondem às datações dos *business cycles* da produção industrial datadas pelo autor pelo filtro polinomial.

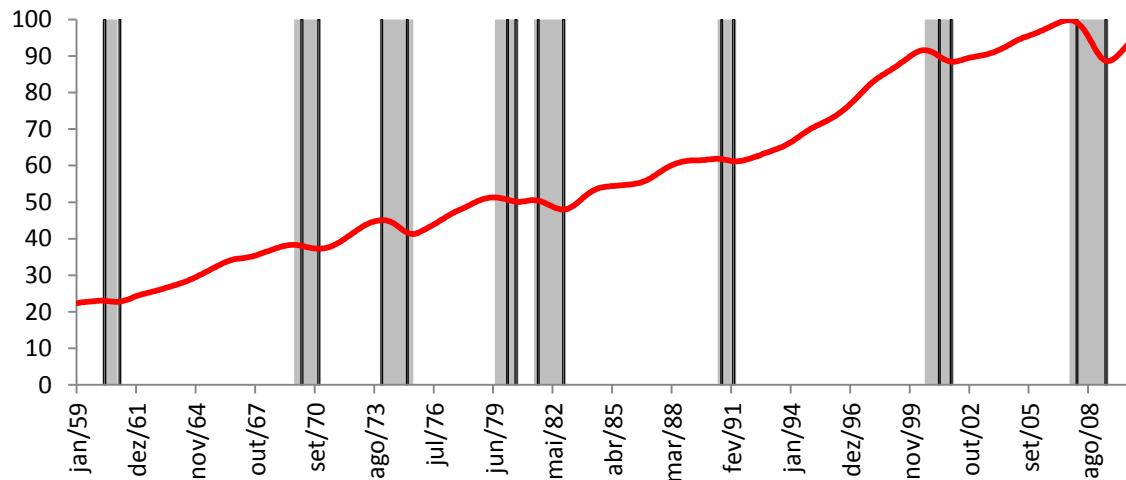


Figura 22 - Valores filtrados da produção industrial americana (linha vermelha) de 1959 a 2010; datação de seus *business cycles* (áreas hachuradas) e a datação dos *business cycles* pelo NBER para a atividade agregada (linhas verticais).

As Tabelas 13 e 14 abaixo mostram os *turning points* que delimitam os *business cycles* das respectivas datações (atividade agregada e produção industrial), bem como as durações das recessões em números de meses. Na Tabela encontra-se a numeração de referência de cada recessão americana datada pelo NBER.

Tabela 13 - *Turning Points* dos *business cycles* da atividade agregada americana datados pelo NBER - 1959 a 2010

Recessão	Picos	Vales	Meses em recessão
1	Abril/1960	Fevereiro/1961	10
2	Dezembro/1969	Novembro/1970	11
3	Novembro/1973	Março/1975	16
4	Janeiro/1980	Julho/1980	6
5	Julho/1981	Novembro/1982	16
6	Julho/1990	Março/1991	8

7	Março/2001	Novembro/2001	8
8	Dezembro/2007	Junho/2009	18
Duração total de meses			93

Fonte: NBER (2013).

Tabela 14 - *Turning Points* dos *business cycles* da produção industrial americana datados pelo filtro polinomial - 1959 a 2010.

Recessão	Picos	Vales	Meses em recessão
1	Março/1960	Dezembro/1960	9
2	Agosto/1969	Novembro/1970	15
3	Dezembro/1973	Junho/1975	18
4	Junho/1979	Agosto/1980	14
5	Maio/1981	Novembro/1982	18
6	Maio/1990	Abril/1991	11
7	Julho/2000	Dezembro/2001	17
8	Agosto/2007	Junho/2009	23
Duração total de meses			125

Fonte: Elaboração própria com dados do NBER via BGRFS.

Abstraindo os aspectos metodológicos de cada datação, pode-se inferir algumas conclusões sobre a comparação das duas datações. Como observado, os períodos de recessão na indústria mostraram-se mais prolongados do que os da atividade agregada. As recessões na indústria duraram 32 meses a mais no total, diferença não muito expressiva se for considerado o horizonte de tempo estudado, com 619 meses. Os meses iniciais de cada recessão foram relativamente semelhantes, apresentando uma diferença máxima de oito meses na recessão de número sete. A Tabela 15 abaixo sintetiza tais diferenças. Valores negativos indicam que a recessão na produção industrial precedeu a recessão da atividade agregada, enquanto que os valores positivos indicam uma postergação do seu início.

Tabela 15 - Diferenças mensais entre o início das recessões na indústria e o início das recessões na atividade agregada americana - 1959 a 2012.

Recessão	Diferenças
1	-1
2	-4
3	+1
4	-7
5	-2
6	-2
7	-8
8	-4

Fonte: Elaboração própria.

As recessões na indústria tendem a aparecer um pouco antes do que na atividade agregada. Apenas a recessão três foge a esta regra, aparecendo um mês depois. Dado isso, duas observações importantes devem ser enfatizadas. A primeira, que os momentos iniciais das recessões são bastante semelhantes com os da atividade agregada. A segunda, que de todas as recessões datadas na série da produção industrial pelo filtro polinomial, nenhuma representou um falso sinal de recessão da atividade agregada medida pelo NBER. Tal informação é de bastante utilidade para o modelo de previsão que possa ser construído futuramente em outra pesquisa, uma vez que, caso soe um alarme para uma recessão industrial americana, pode-se concluir que este também servirá para uma recessão da atividade agregada. Portanto, concluiu-se que a série de produção industrial americana é um bom indicador coincidente da atividade agregada americana e seus respectivos *turning points* são bem captados através do método de suavização polinomial.

6.3 Seleção dos indicadores antecedentes

O objetivo desta seção é relatar os procedimentos realizados nas variáveis candidatas a antecedentes para verificar o quanto são eficientes em anteceder os movimentos da target variable. Especificamente, o procedimento seguiu quatro etapas: identificação das variáveis na literatura; transformações nas variáveis; teste de correlação com a target variable e a seleção final dos indicadores antecedentes através do método stepwise. Somando uma quinta etapa encontra-se uma análise ex-post, que é a análise gráfica das séries escolhidas para a determinação do que seria o limite comportamento normal de cada indicador antecedente, além de uma breve explicação teórica do motivo para aquele indicador ser considerado como antecedente.

6.3.1 Análise de Correlação e o Método Stepwise de Seleção dos Indicadores Antecedentes

O objetivo desta subseção é descrever os passos dois e três citados na introdução deste estudo, ou seja, as quatro etapas de seleção dos indicadores antecedentes. A seleção das variáveis antecedentes baseou-se em trabalhos na literatura que se propuseram a selecionar os melhores indicadores antecedentes para a produção industrial e/ou para o PIB. Foram testadas mais de cento e cinquenta variáveis candidatas a antecedentes.

Como o trabalho tratou de estudar o comportamento cíclico da produção industrial, decidiu-se dessazonalizar todas as séries. Cada variável candidata a indicador antecedente foi verificada em termos de nível de correlação com a target variable. O intuito do teste é verificar se existe significância estatística na correlação entre a variável antecedente e a target variable. O teste de correlação escolhido foi o da correlação de Spearman, que é uma medida de correlação não-paramétrica, a qual tem por característica não levar em consideração a distribuição de frequência da variável em estudo (LIRA, 2004).

A target variable foi transformada em variável dicotômica, com valores zero para os períodos de recessão e com valores unitários para períodos de não-recessão, pois a previsão se dará de forma dicotômica com a análise

discriminante. Dado que a target variable apresenta-se em formato não-linear, justifica-se o uso da correlação de Spearman em detrimento da correlação de Pearson. O teste de correlação de Pearson, por exigir linearidade nas variáveis não seria cabível neste caso. A variável dicotômica que representa as recessões sendo uma variável não-linear precisa de um teste de correlação que leve apenas em consideração o ordenamento das variáveis, que é o caso da correlação de Spearman.

A análise de correlação inicia-se a partir de Fevereiro de 1992 e vai até Dezembro de 2012. Tal período justifica-se devido as doze defasagens aplicadas em todas as séries candidatas a antecedentes. O valor máximo da defasagem de doze meses justifica-se, pois, não espera-se uma correlação significativa com uma defasagem maior que esta que não seja espúria. As variáveis antecedentes foram testadas em seu nível de correlação com a dummy através de quatro transformações diferentes, com intuito de verificar a forma mais correlacionada. As transformações foram: taxa de crescimento da variável a partir de seus valores em nível; taxa de crescimento da variável após sua filtragem ; valores filtrados; valores da série sem nenhuma transformação, ou seja, em nível.

O comportamento que melhor se correlaciona com a dummy é a taxa de crescimento da variável após sua filtragem. Fato este já esperado, uma vez os valores das recessões e não-recessões da target variable foram estabelecidos a partir da taxa de crescimento da variável filtrada. As séries que foram coletadas em valores percentuais apresentaram melhor correlação com a dummy através de seus valores filtrados ou em nível, sem necessidade de calcular suas taxas de crescimento.

Vinte variáveis apresentaram correlação significativa com pelo menos seis meses de defasagem. Tais variáveis foram utilizadas para a construção do modelo discriminante. O software estatístico SPSS® realizou, através do método stepwise , uma série de combinações entre as variáveis a fim de escolher o grupo que melhor discrimina a target variable. Variáveis como taxa de juros dos Estados Unidos, variáveis de tributação e de produção de bens de consumo não foram selecionadas pelo método. Apenas nove entraram para o modelo. Por vezes, existiram variáveis que melhor explicariam de forma individual a recessões, porém, não foram selecionadas para o modelo final por

estarem correlacionadas com outras variáveis preditoras mais importantes. A Tabela 16 mostra as melhores defasagens das nove variáveis escolhidas para fazer parte do modelo discriminante. Mostra também o valor da correlação de Spearman para cada defasagem e o tipo de transformação estabelecida na variável.

Tabela 16 - Defasagens e correlações das variáveis antecedentes com a variável *dummy* das recessões

Série	Defasagem ótima	Correlação	Transformação
Ibovespa	-11	0,458**	Taxa de crescimento Filtro
Exportações de Caminhões	-12	0,124*	Taxa de crescimento Filtro
Dívida Mobiliária	-10	0,167**	Taxa de crescimento Filtro
Pontos Bovespa	-6	0,565**	Taxa de crescimento Filtro
Produção de Materiais Elétricos para Veículos	-6	0,487**	Taxa de crescimento Filtro
Produção de Carvão	-6	0,321**	Taxa de crescimento Filtro
Taxa de Câmbio	-7	0,223**	Em nível
Desaceleração Industrial - EUA	-6	0,378**	Taxa de crescimento Filtro***
Desaceleração Industrial - Brasil	-6	0,424**	Taxa de crescimento Filtro***

Fonte: Elaboração própria.

*** Valores transformados para variáveis dicotômicas, zero para desaceleração e um para aceleração. ** Correlação significativa ao nível de 1%. * Correlação significativa ao nível de 5%.

6.3.2 Análise Individual dos Indicadores Antecedentes

Assim como afirmado por Keilis-Borok (2000) em seu estudo de previsão das recessões americanas, o objetivo deste trabalho é explorar o modelo de previsão e não comparar as propriedades de previsão de diferentes indicadores antecedentes. Apesar disso, faz-se necessário um breve estudo sobre cada variável selecionada para o modelo preditivo. Esta subseção tem como objetivo analisar cada indicador antecedente de forma separada. Foram realizados testes de estacionariedade nas nove variáveis escolhidas pelo método stepwise e apenas a variável remetente a taxa de câmbio não apresentou-se estacionária. Tal estacionariedade já era esperada nas variáveis visto que estas se apresentavam em taxas de crescimento. As variáveis que medem os momentos de desaceleração das indústrias americana e brasileira por estarem representadas dicotomicamente não foram submetidas aos testes de estacionariedade.

A análise gráfica de cada indicador antecedente não tem o mesmo sucesso preditivo que todos os indicadores juntos, já que o método de seleção stepwise realizado pelo SPSS® escolheu os indicadores para atuarem de forma conjunta, deixando de fora, como já afirmado, bons indicadores antecedentes. . A essência é que o índice formado por todos os indicadores juntos consiga suavizar algumas volatilidades dos indicadores individuais, sendo assim mais convincente e claro.

A análise individual tem a função de determinar o que seria o limite do comportamento normal do indicador, limite este que não será determinado de forma precisa (como o método do ruído/sinal apresentado na introdução deste estudo), mas apenas pela observação gráfica. A determinação precisa deste limite será necessária apenas na próxima seção, na qual o algoritmo de previsão construirá um índice formado por todos os indicadores antecedentes (expresso pelo escore discriminante Z_t) e este índice deverá ter um limite de comportamento normal que ao ser ultrapassado indique que haverá uma recessão.

Os indicadores antecedentes apresentados aqui nesta subseção da letra (a) até a letra (f) foram todos utilizados na análise de previsão a partir do

cálculo de suas taxas de crescimento dos valores filtrados através da seguinte fórmula:

$$\left[\left(\frac{\text{valor em } t+1}{\text{Valor em } t} \right) - 1 \right] * 100 \quad (09)$$

A análise gráfica das letras (a) até a letra (f) será efetuada apenas observando-se os momentos em que as taxas de crescimento encontram-se em valores negativos ou com uma tendência de decrescimento, dado que todos estes indicadores apresentam relação direta com a target variable, ou seja, quando estes indicadores apresentam-se em desaceleração e/ou em taxas de crescimento negativas a target variable também tende a apresentar-se em desaceleração ou em recessão.

A letra (g) diz respeito a taxa de câmbio (R\$/US\$) que não sofreu qualquer tipo de transformação e o estudo do seu comportamento gráfico será relatado mais a frente. As letras (h) e (i) são variáveis dicotômicas que representam os momentos de aceleração e desaceleração da produção industrial. Para determinação dos momentos de desaceleração o procedimento foi simples. Dados os valores filtrados da série de produção industrial (em valores físicos) foi calculada suas taxas de crescimento. A partir da nova série de taxa de crescimento calculou-se as diferenças da série fazendo apenas o valor em $t+1$ menos o valor em t , o que gerou uma nova série de tempo. Quando os valores desta nova série de tempo apresentam-se negativos significa que a produção industrial está desacelerando.

A observação gráfica dos valores em taxas de crescimento já daria a informação dos momentos de desaceleração, porém, para facilitação da demarcação da variável dummy, o procedimento das diferenças da taxa de crescimento foi adotado. Assim, nos gráficos das letras (h) e (i) as barras vermelhas representam os momentos de desaceleração da produção industrial física dos Estados Unidos e do Brasil. Em todas as figuras apresentadas aqui nesta subseção, da letra (a) até a letra (i), as barras azuis representam os momentos recessivos da produção industrial física brasileira, momentos estes demarcados pelo filtro polinomial e que já foram apresentados neste estudo através da Figura 20 e da Tabela 10 da seção 6.2.

Por fim, como já relatou Keilis-Borok et al. (2000), o fato de haver um co-movimento defasado entre a target variable a as séries antecedentes não implica causalidade, mas que estas são mais sensíveis a responder antes às variações de conjuntura.

a) Índice Ibovespa

Esta variável está disponibilizada pelo Banco Central do Brasil (BCB) e apresenta-se através de níveis percentuais mensais. Sua utilização baseou-se em estudo de Chauvet e Silva (2004) e Reinhart e Rogoff (2009). A lógica de sua utilização é que as recessões iniciam-se nos mercados financeiros e também que o índice seja uma proxy das expectativas dos agentes econômicos. A série foi transformada em taxas de crescimento. Como pode ser notado, antes dos momentos recessivos a taxa de variação do Índice Bovespa costuma ser negativa, indicando assim uma relação positiva com a atividade industrial.

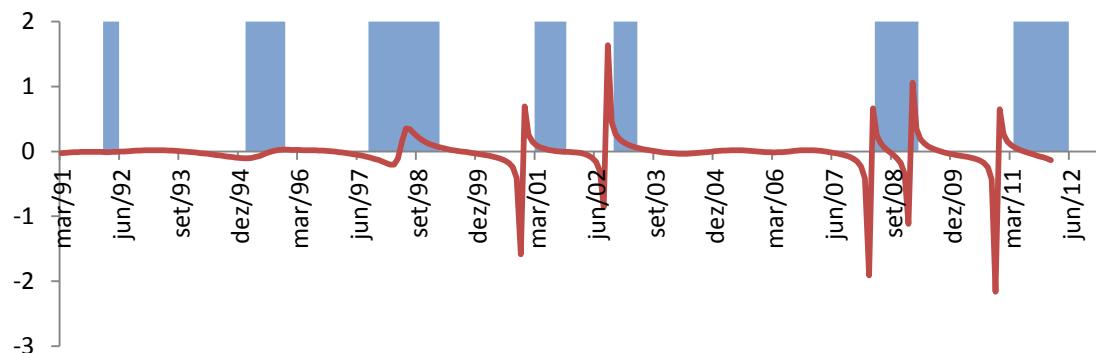


Figura 23 - Índice Ibovespa em taxas de crescimento.

b) Exportações de caminhões

Esta variável está disponibilizada pela Associação dos Fabricantes de Veículos Automotores (Anfavea) e apresenta-se mensalmente através de unidades exportadas. Sua utilização justifica-se pois é uma variável que serve de proxy para o setor de exportações da indústria automobilística, que é uma indústria de suma importância no Brasil. A série foi transformada em taxas de crescimento. Como pode ser notado, existe uma tendência de queda das exportações de caminhões alguns meses antes das recessões, indicando assim uma relação positiva com a atividade industrial.

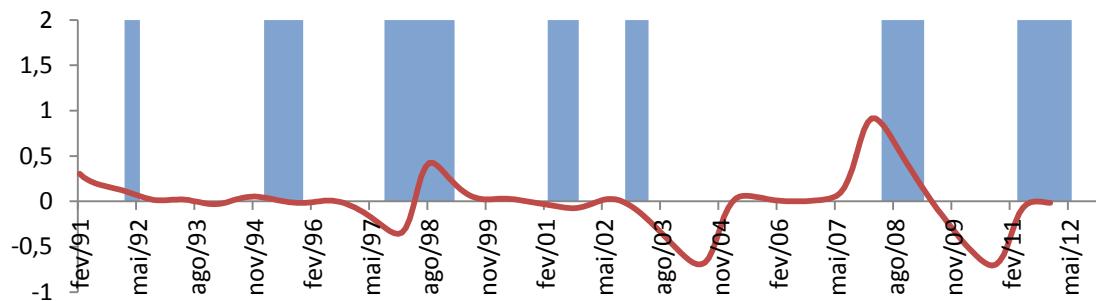


Figura 24 - Exportações de caminhões em taxas de crescimento.

c) Dívida Pública Mobiliária Federal interna (DPMFi)

Esta variável está disponibilizada pelo Banco Central do Brasil (BCB) e apresenta-se em milhões de reais e foi deflacionada através do IGP-DI. A DPMFi tende a cair quando o governo corta seus gastos (ocasionando um superávit primário) e dispensa seus recursos para pagá-la, sendo assim uma proxy de uma política fiscal contracionista. A série foi transformada em taxas de crescimento. Como pode ser notado, existe uma tendência antes das recessões da DPMFi encontrar-se com valores negativos, indicando assim uma relação positiva com a atividade industrial.

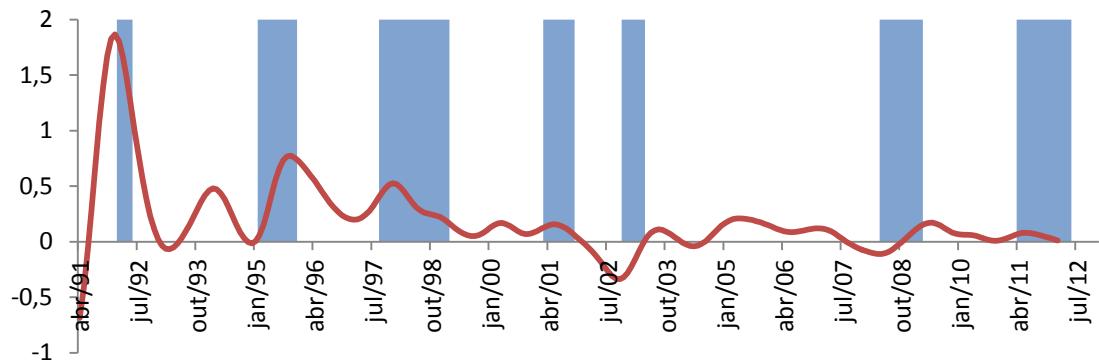


Figura 25 - Dívida Pública Mobiliária Federal interna (DPMFi) em taxas de crescimento.

d) Pontos Bovespa

Esta variável é disponibilizada pelo Banco Central do Brasil (BCB) e apresenta-se mensalmente em quantidade total de pontos. Sua utilização baseou-se em estudo de Chauvet e Silva (2004) e Reinhart e Rogoff (2009). A lógica é que as recessões do lado real da economia iniciam-se nos mercados financeiros e também que o índice seja uma proxy das expectativas dos agentes econômicos. A série foi transformada em taxas de crescimento. Como pode ser notado, antes dos momentos recessivos a taxa de variação dos Pontos Bovespa costuma ser negativa, indicando assim uma relação positiva com a atividade industrial. Apesar de já existir a variável do índice Bovespa, como o setor financeiro é central para o entendimento das recessões do lado real, justifica-se a utilização de dois tipos de séries deste setor. Deve-se notar que as séries possuem defasagens ótimas distintas, como pode ser visto na Tabela 16.

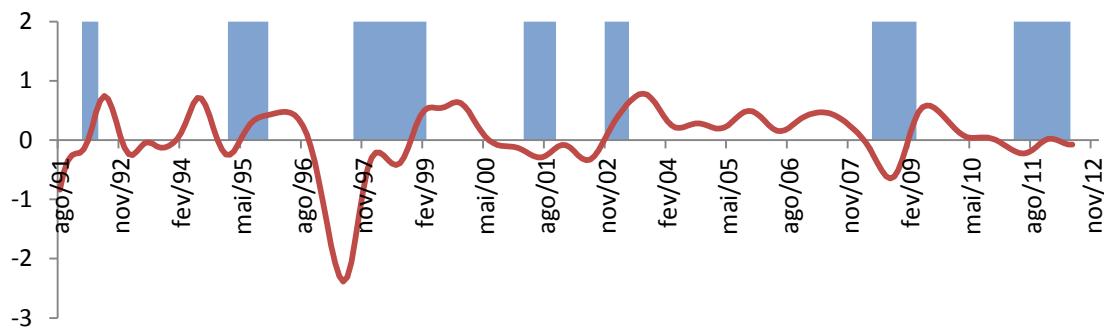


Figura 26 - Pontos Bovespa em taxas de crescimento.

e) Produção de material elétrico para veículos

Esta variável é disponibilizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e apresenta-se mensalmente em sua produção física (unidades) que é transformada em número índice com a média do ano de 2002 igual a cem. Sua utilização baseou-se em estudo de Chauvet e Morais (2009) e justifica-se por ser um setor de fornecimento de matéria prima para o setor automobilístico. A série foi transformada em taxas de crescimento. Como pode ser observado, antes de todas as recessões a taxa de variação da Produção de Material Elétrico para Veículos apresenta-se negativa. Não significa uma ideia de causação, mas apenas de que esta variável é mais sensível aos sinais de uma recessão, apresentando assim um comportamento fora do normal antes da produção industrial.

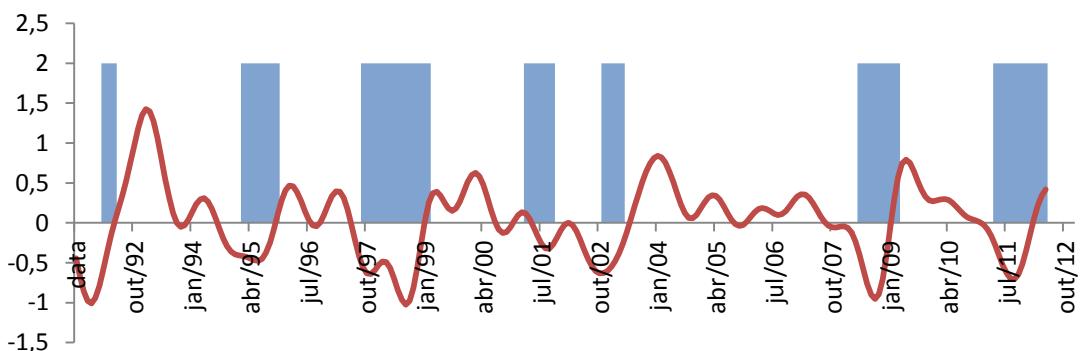


Figura 27 - Produção de material elétrico para veículos em taxas de crescimento.

f) Extração de carvão mineral

Esta variável é disponibilizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e apresenta-se mensalmente em sua produção física (toneladas) que é transformada em número índice com a média do ano de 2002 igual a cem. Sua utilização baseou-se em estudo de Chauvet e Morais (2009) e justifica-se por ser um setor fornecedor de matéria prima para diversos outros setores industriais. A série foi transformada em taxas de crescimento. Como pode ser observado, antes de todas as recessões a taxa de variação da Extração de Carvão Mineral apresenta-se negativa. Não significa uma ideia de causação, mas apenas de que esta variável é mais sensível aos sinais de uma recessão, apresentando assim um comportamento fora do normal antes que a produção industrial.

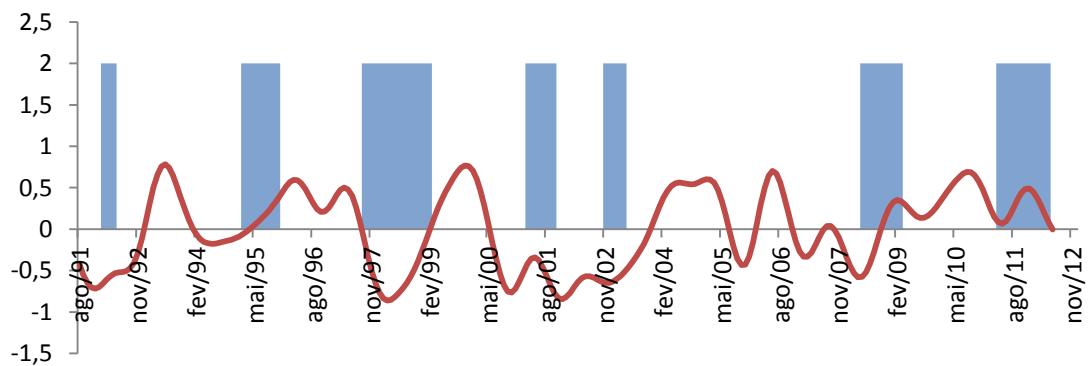


Figura 28 - Extração de carvão mineral em taxas de crescimento.

g) Taxa de câmbio comercial para a venda (R\$/US\$)

Esta variável é disponibilizada pelo Banco Central do Brasil (BCB) e apresenta-se mensalmente em termos de R\$/US\$. Sua utilização baseou-se em estudo de Issler Hollauer (2006). A lógica é que uma taxa de câmbio valorizada, ou seja, quando seu valor estiver baixo, desestimule o setor industrial, dado que este depende do desempenho de suas exportações. Como esta variável não sofreu transformações, seu limite normal pode ser analisado apenas pela observação dos momentos de crescimento e decrescimento da

variável. Assim, o valor da Taxa de Câmbio Comercial para a Venda apresenta uma relação positiva com a atividade industrial.

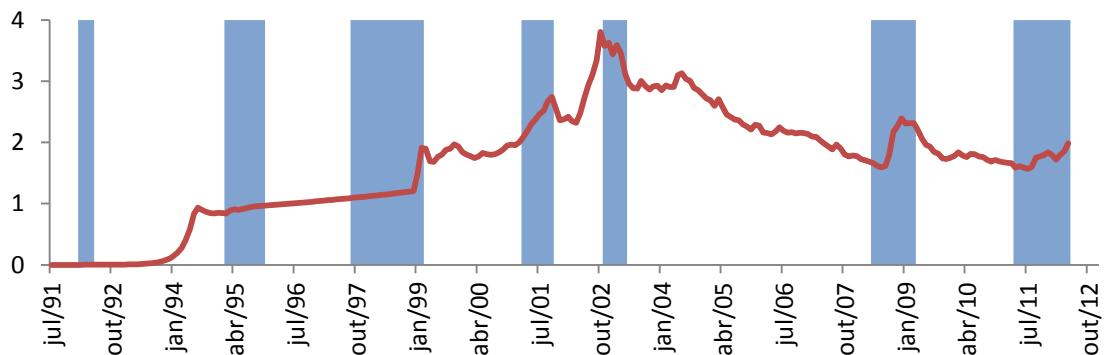


Figura 29 - Taxa de câmbio comercial para a venda (R\$/US\$).

h) Desaceleração da produção industrial americana

Esta variável é disponibilizada pelo Banco Central do Brasil (BCB) e apresenta-se mensalmente em forma de índice da quantidade total de unidades produzidas mensalmente. Os momentos de desaceleração da produção industrial americana, representados pelas barras vermelhas, ocorrerem sempre antes dos momentos de recessão da atividade industrial brasileira. Houveram alguns momentos de desaceleração da indústria americana que não provocaram recessões no Brasil, como por exemplo as desacelerações de 2005 e 2006. A maiores desacelerações, de 2000 e de 2007, provocaram as duas últimas recessões americanas, que foram as recessões de 2001 e 2008, como podem ser visualizadas na Figura 22 da seção 6.2. Como pode ser visto na Figura 30 abaixo, o Brasil também entrou em recessão depois destas duas desacelerações americanas. Assim, os momentos de desaceleração da produção industrial americana apresentam uma relação positiva com a atividade industrial brasileira.

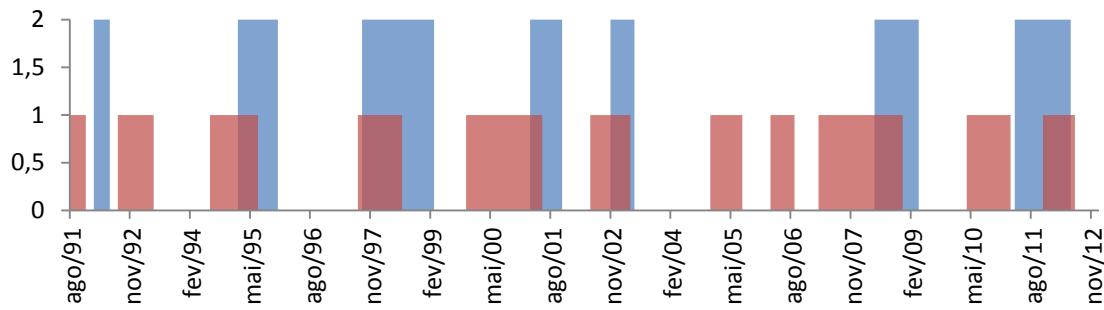


Figura 30 - Momentos de desaceleração da produção industrial americana.

i) Desaceleração da produção industrial brasileira

Esta variável é disponibilizada pelo Banco Central do Brasil (BCB) e apresenta-se mensalmente em forma de índice da quantidade total de unidades produzidas mensalmente. Os momentos de desaceleração da produção industrial brasileira, representados pelas barras vermelhas, ocorrerem sempre antes dos momentos de recessão da atividade industrial. Notar que existem três momentos de desaceleração que não se revertem em recessões, que foram as desacelerações de 1993, 2004, 2005/06. Ou seja, nem toda desaceleração se traduz em recessão. Por isso, apesar dos momentos de desaceleração já servirem como um bons indicadores de recessão, eles ainda apresentam falsas sinalizações, justificando assim um índice que agregue outros indicadores antecedentes, com intuito de não dar falsos alarmes de recessões.

Se for realizada uma comparação entre as Figuras 30 e 31, é possível notar semelhanças entre as desacelerações americanas e brasileiras, mostrando um forte de grau de conectividade entre as duas economias.

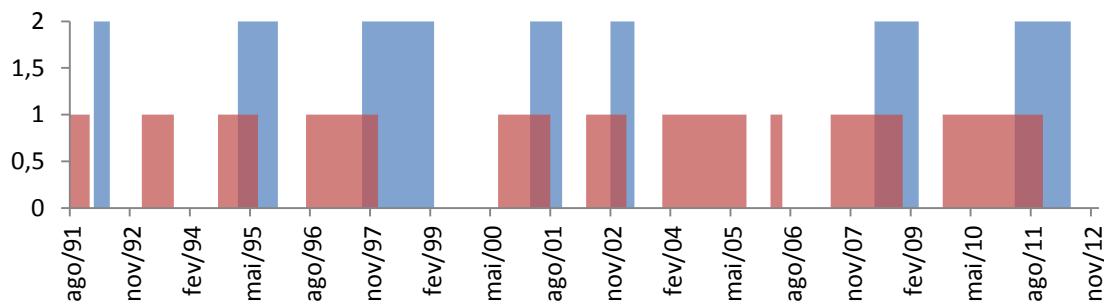


Figura 31 - Momentos de desaceleração da produção industrial brasileira.

6.4 Análise discriminante: previsões

Antes de apresentar os resultados da análise discriminante é necessário uma breve explicação sobre o termo previsão. Esse termo pode ser entendido de duas maneiras diferentes: previsões dentro da amostra ou fora da amostra. A primeira utiliza toda a informação disponível na amostra para elaborar o modelo e então analisa o seu poder preditivo em relação às observações dentro da própria amostra. A segunda utiliza informações de uma parte da amostra para estimar o modelo e então realiza as previsões para o restante das observações. A primeira metodologia, como considera toda a informação disponível para estimar o modelo, apresenta maior potência. Porém, as previsões fora da amostra possuem maior aplicação prática, pois aproximam-se melhor das restrições de informações que observamos na vida real. As duas metodologias serão aplicadas aqui.

Primeiramente serão apresentados os resultados para o método que utilizada toda a informação disponível (subseção 6.4.1). Posteriormente serão apresentados os resultados do método que utiliza informações de uma parte da amostra para estimar o modelo e a outra para realizar as previsões (subseção 6.4.2). Os dois tipos de método utilizaram as mesmas variáveis explicativas com suas respectivas defasagens contidas na Tabela 16, porém, com valores dos coeficientes distintos. Os dois métodos também utilizaram como target variable a série de produção industrial brasileira. Como já afirmado no final da seção 6.2 e na seção 6.3, devido as defasagens aplicadas nos indicadores antecedentes e aos cálculos das taxas de crescimento, perderam-se treze meses de análise, que começaria em janeiro de 1991 e passará a começar em fevereiro de 1992, cobrindo assim sete recessões, sendo seis completas e uma incompleta. A primeira recessão, de 1991, que na Tabela 10 está identificada com uma duração de nove meses, passa a durar apenas quatro meses. As outras seis recessões serão previstas de forma completa, de acordo como são mostradas na Tabela 10.

6.4.1 Análise discriminante com toda informação disponível

A amostra com toda informação disponível vai de fevereiro de 1992 até dezembro de 2012 e possui duzentas e cinquenta e uma observações. Destas, setenta e uma são de meses recessivos (28%) e cento e oitenta de meses não-recessivos (72%). Tomando como base as defasagens de cada variável dispostas na Tabela 16, a função discriminante genérica ficará da seguinte forma:

$$Z_t = a + \beta_1 X_{1_{t-11}} + \beta_2 X_{2_{t-12}} + \beta_3 X_{3_{t-10}} + \beta_4 X_{4_{t-6}} + \beta_5 X_{5_{t-6}} + \beta_6 X_{6_{t-6}} + \beta_7 X_{7_{t-7}} + \beta_8 X_{8_{t-6}} + \beta_9 X_{9_{t-6}} \quad (10)$$

em que,

$t - m$ é a antecedência da variável em relação a data t , em que t representa a data no formato mês/ano e m representa o número de meses antecedentes a essa data; com $0 < m \leq 12$;

Z_t é o escore discriminante para a data t ;

a é a constante da função discriminante;

β_n é o coeficiente discriminante;

X_n são os valores das variáveis independentes (tais variáveis estão dispostas na Tabela 16 e 17)

A Tabela 17 mostra os valores dos coeficientes discriminantes estimados para cada variável antecedente utilizada no modelo bem como o valor estimado da constante discriminante. É também mostrado a significância das variáveis a partir da estatística Wilks'Lambda⁴¹.

⁴¹ O teste de Wilks' Lambda é a estatística padrão usada para denotar a significância estatística do poder discriminatório da função discriminante em questão. O valor do teste pode variar de 0,0 a 1,0 considerando-se que, quanto mais próximo de 1,0 menor será o poder discriminatório da função, sendo o contrário verdadeiro (VIEIRA et al., 2009).

Tabela 17 - Valores dos coeficientes discriminantes e da constante discriminante.

Variável	Coeficiente Discriminante	Valor do Coeficiente Discriminante	Wikis'Lambda	Significância
Ibovespa	β_1	0,005	0,937	0,000
Exportações de Caminhões	β_2	0,022	0,995	0,247
Dívida Mobiliária	β_3	0,145	0,972	0,008
Pontos Bovespa	β_4	0,129	0,724	0,000
Produção de Materiais Elétricos	β_5	0,372	0,806	0,000
Produção de Carvão	β_6	0,500	0,896	0,000
Taxa de Câmbio	β_7	0,476	0,982	0,035
Desaceleração Industrial - EUA	β_8	0,727	0,857	0,000
Desaceleração Industrial - Brasil	β_9	0,741	0,820	0,000
Constante	a	-1,723		

Fonte: Elaboração própria.

O software estatístico SPSS® selecionou uma variável não significativa, que foi a variável referente as exportações de caminhões para fazer parte do modelo. A razão disto não foi devido a falta de outras variáveis para compor o algoritmo e sim as possíveis correlações entre as variáveis. Como pode ser observado, todas as variáveis apresentaram relação direta com a *target variable* (coeficientes discriminantes positivos). Das nove variáveis, as que indicam os momentos de desaceleração industrial no Brasil e nos Estados Unidos foram as que receberam maior peso. Porém, isto não significa um poder discriminante maior, visto que esse poder depende dos valores das variáveis independentes. O software listou as variáveis com melhor poder discriminatório: Pontos Bovespa β_4 , Desaceleração Industrial - EUA β_2 , Produção de Materiais Elétricos β_5 , Desaceleração Industrial - Brasil β_9 , Ibovespa β_1 , Produção de Carvão β_6 , Exportações de Caminhões β_2 , Taxa de Câmbio β_7 , Dívida Mobiliária β_3 .

Com a constante e os coeficientes discriminantes estimados, pode-se calcular o escore discriminante Z_t para cada mês. Calculados os escores, resta defini-los se eles representam uma recessão ou não. A regra de classificação para determinar se um mês é ou não de recessão, como já apresentado na metodologia, pode ser de dois tipos: através do ponto de corte Z_c ou através da distância euclidiana quadrática em relação aos centróides de cada grupo. Esses dois tipos de regra de classificação usam os valores dos centróides de cada grupo (recessão e não-recessão) em suas equações. Os valores calculados pelo *software* para os centróides dos grupos zero e um são, respectivamente, -1,790 e 0,706.

A seguir serão apresentados os resultados da previsão com toda informação disponível, primeiramente através da regra de classificação do ponto de corte e posteriormente através da regra de classificação através da distância euclidiana quadrática. A intenção é verificar qual metodologia de classificação se adéqua melhor aos dados desta previsão e então escolher qual delas será utilizada para classificar os dados não-amostrais da subseção 6.4.2.

6.4.1.1 Regra de classificação com o ponto de corte para grupos com tamanhos distintos

Quando os grupos têm tamanhos distintos, como no caso deste estudo, uma maneira de calcular o ponto de corte é atribuindo peso aos centróides em função do número de elementos de cada amostra, através da fórmula (06), mostrada na metodologia:

$$Z_c = \frac{N_0 Z_1 + N_1 Z_0}{N_0 + N_1} \quad (06)$$

em que,

Z_c = valor do escore crítico.

N_0 = número de observações do grupo zero.

N_1 = número de observações do grupo um.

Z_0 = centróide do grupo zero.

Z_1 = centróide do grupo um.

Valores de Z_t maiores do que $-1,084$ serão classificados no grupo dos uns e consequentemente, valores menores que $-1,084$ serão classificados no grupo dos zeros. Dada essa regra de classificação, 87,5% das observações foram previstas corretamente, como podem ser vistas na Tabela 18 abaixo:

Tabela 18 - Acertos e Erros das Previsões

	Ocorrência	Previsão		Total
		0	1	
Amostra	0	57	14	71
	1	8	172	180
	0	80%	20%	100%
	1	5%	95%	100%

Fonte: Elaboração própria.

A função discriminante conseguiu acertar 80% das recessões e 95% das não-recessões e daí resulta o acerto médio de 87,5%. Ela deixou de apontar quatorze casos que eram recessões dos setenta e um possíveis, ou seja, não apontou 20% das recessões (erro Tipo I). Dos cento e oitenta casos de não-recessão, a função errou apenas em oito casos, indicando pois, oito falsos alarmes da discriminante (erro Tipo II). Essa previsão apresentou um QPS⁴² no valor de 0,17.

A Figura 32 abaixo mostra os alarmes de recessões e as respectivas recessões para os dados deste estudo. As barras pretas representam os alarmes enquanto as barras azuis representam as recessões.

⁴² Quadratic Probability Score. Seu valor varia entre 0 e 2 e quanto menor mais eficiente é a previsão (ver o final da seção 5.1 para maiores detalhes).

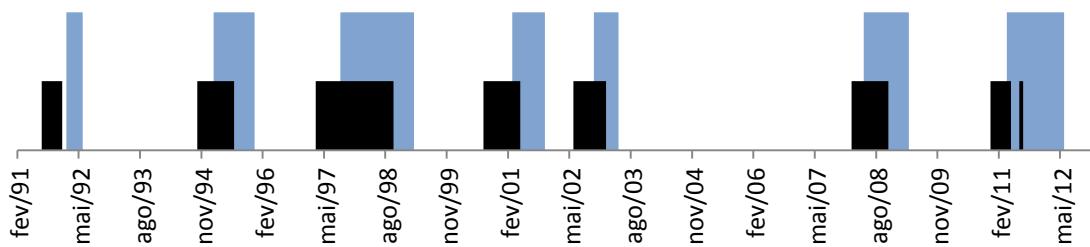


Figura 32 - Alarmes (barras pretas) e recessões (barras azuis), utilizando a regra de classificação através da fórmula (06).

Como pode ser observado, todos os momentos iniciais das recessões foram previstos pelo algoritmo. Dos oito erros Tipo II cometidos, nenhum deles representou um falso alarme preditivo, eles apenas inflaram o período recessivo, adicionando à previsão de um a dois meses antes do seu início e/ou um ou dois meses no final do período recessivo. Dos quatorze erros Tipo I nenhum deles atrapalhou de forma acentuada a previsão dos momentos iniciais das recessões. A Tabela 19 abaixo lista em quantos meses o algoritmo conseguiu antecipar cada recessão:

Tabela 19 - Duração dos alarmes antes de cada recessão

Recessão	Período do alarme	Duração (meses)
3	Ago/1991 - Dez/1991	5
4	Out/1994 - Jan/1995	4
5	Mar/1997 - Ago/1997	6
6	Ago/2000 - Fev/2001	7
7	Jun/2002 - Out/2002	5
8	Fev/2008 - Abr/2008	3
9	Dez/2010 - Mar/2011	4

Fonte: Elaboração própria.

As recessões que não foram alarmadas com pelo menos seis meses de antecedência é porque erros do Tipo I ocorreram nos momentos iniciais de cada previsão. Neste sentido, foram detectados nove erros do Tipo I que reduziram o espaço temporal das previsões do início das recessões. Os outros cinco erros do Tipo I ocorreram nos períodos finais dos alarmes, encurtando-os

ou tornando o alarme descontínuo perto do seu final, que foi o caso ocorrido na última recessão.

Nesta última recessão, a barra preta mais fina não representa nem um erro Tipo I nem Tipo II. Ela também não representa um falso alarme preditivo pois não é um erro Tipo II. Ela representa apenas um acerto de previsão como todos os outros. O que ocorreu foi que erros do Tipo I apareceram entre a barra preta mais grossa e a barra preta mais fina, por isso o alarme apresentou essa descontinuidade. O critério que anteriormente já foi estabelecido (seção 5.2) é que alarmes que ocorram dentro de períodos recessivos não são alarmes por si só, apenas fazem parte do alarme inicial.

6.4.1.2 Regra de classificação através da distância euclidiana quadrática

Esse tipo de método como regra de classificação consiste apenas em comparar os escores de cada observação com suas distâncias euclidianas quadráticas em relação aos centróides. A regra de classificação para cada observação será feita escolhendo-se a menor distância euclidiana quadrática em relação às centróides de cada grupo. A regra de decisão, como já mostrada na metodologia, será dada pela equação (07):

$$Classificação = \begin{cases} 0 & \text{se } (Z_t - Z_0)^2 < (Z_t - Z_1)^2 \\ 1 & \text{se } (Z_t - Z_1)^2 < (Z_t - Z_0)^2 \end{cases} \quad (07)$$

em que,

Z_t = escore discriminante para a data t ;

Z_0 = centróide do grupo zero;

Z_1 = centróide do grupo um;

$(Z_t - Z_0)^2$ = distância euclidiana quadrática para o grupo das recessões;

$(Z_t - Z_1)^2$ = distância euclidiana quadrática para o grupo das não-recessões.

Dada essa regra de classificação, 94,5% das observações foram previstas corretamente, como podem ser vistas na Tabela 20 abaixo:

Tabela 20 - Acertos e Erros das Previsões

	Ocorrência	Previsão		Total
		0	1	
Amostra	0	68	3	71
	1	13	167	180
Total	0	96%	4%	100%
	1	7%	93%	100%

Fonte: Elaboração própria.

A função discriminante conseguiu acertar 96% das recessões e 93% das não-recessões e daí resulta o acerto médio de 94,5%. Ela deixou de classificar três casos que eram recessões dos setenta e um possíveis, ou seja, não apontou 4% das recessões (erro Tipo I). Dos cento e oitenta casos de não-recessão, a função errou em treze casos, indicando pois, treze falsos alarmes da discriminante (erro Tipo II). Essa previsão apresentou um QPS no valor de 0,14.

A Figura 33 abaixo mostra os alarmes de recessões e as respectivas recessões para os dados deste estudo. As barras pretas representam os alarmes enquanto as barras azuis representam as recessões.

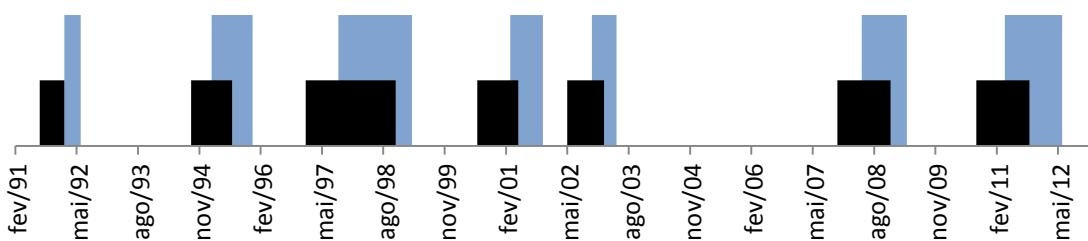


Figura 33 - Alarmes (barras pretas) e recessões (barras azuis), utilizando a regra de classificação através da fórmula (07).

Como pode ser observado, todos os momentos iniciais das recessões conseguiram ser previstos pelo algoritmo. Dos treze erros Tipo II cometidos, nenhum deles representou um falso alarme preditivo, eles apenas inflaram o período recessivo, adicionando à previsão de um a dois meses antes do seu início e/ou um ou dois meses no final do período recessivo. Dos três erros Tipo I nenhum deles atrapalhou de forma acentuada a previsão dos momentos

iniciais das recessões. A Tabela 21 abaixo lista em quantos meses o algoritmo conseguiu antecipar cada recessão:

Tabela 21 - Duração dos alarmes antes dos inícios de cada recessão

Recessão	Período do alarme	Duração (meses)
3	Ago/1991 - Jan/1992	6
4	Set/1994 - Jan/1995	5
5	Jan/1997 - Ago/1997	8
6	Jul/2000 - Fev/2001	8
7	Mai/2002 - Out/2002	6
8	Nov/2007 - Abr/2008	6
9	Set/2010 - Mar/2011	7

Fonte: Elaboração própria.

O único alarme que cometeu um erro Tipo I no início de sua recessão foi o representativo da crise número quatro. As outras sete recessões todos os alarmes apresentaram uma duração de pelo menos seis meses. Os alarmes que duram mais que seis meses refletem os erros Tipo II ocorridos nos momentos iniciais da previsão.

Com tais resultados é possível concluir que a metodologia apresenta aqui através das distâncias euclidianas quadráticas obteve um maior sucesso preditivo. Além de obter um percentual maior de acertos obteve também um QPS menor e tem a característica de antecipar com maior antecedência as recessões. Portanto, ela será adotada para ser utilizada na próxima subseção, a qual será realizada a previsão mais importante deste estudo, a previsão com dados não-amostrais.

Por fim, cabe aplicar o algoritmo desenvolvido nesta seção para prever o que ocorrerá na indústria até Agosto de 2013. Apesar desta análise está se efetuando em abril de 2013, as séries de produção de carvão, de material elétrico para veículos e a da própria produção industrial só estão disponíveis até o mês de fevereiro. Portanto, devido à demora de divulgação dos dados, o horizonte temporal de previsão de uma recessão pode ser diminuído em um ou

dois meses. A Tabela 22 abaixo mostra as previsões até o mês de agosto de 2013.

Tabela 22 - Previsão para 2013

Data	Prod. Industrial	Prod. Industrial Filtrada	Dummy	Escore Discriminante (Z_t)	$(Z_t - Z_0)^2$	$(Z_t - Z_1)^2$	Previsão
Jan/2013	128,76	126,85	1	0,12	3,66	0,33	1
Fev/2013	125,79	126,94	1	0,01	3,27	0,47	1
Mar/2013				0,58	5,65	0,01	1
Abr/2013				-0,6	2,97	0,59	1
Mai/2013				2,12	15,35	2,02	1
Jun/2013				1,26	9,33	0,31	1
Jul/2013				0,78	6,62	0,00	1
Ago/2013				-0,13	2,75	0,70	1

Fonte: Elaboração própria.

Como pode ser observado, apesar do número índice que mede o nível de produção industrial ter caído de 128,76 para 125,79, uma queda de 2,3%, o valor filtrado do índice teve um pequeno aumento de 0,07%. Ou seja, por pouco não aconteceu uma variação negativa na taxa de crescimento, o que levaria a um novo momento de recessão. Apesar do valor da produção industrial de dezembro de 2013 não está na Tabela 22, ele foi de 125,95 e seu valor filtrado foi de 126,71. Assim, houve uma variação positiva de 0,11% entre os valores filtrados de dezembro e janeiro. Dessa forma, a dummy recebeu o valor de um, indicando uma não-recessão, assim também ocorrido para o mês de fevereiro. Os demais meses ainda não possuem os dados divulgados. A função discriminante acertou para os dois casos possíveis no ano de 2013. Notar que no mês de fevereiro, o qual quase representou o início de uma recessão, o escore apresentou-se bastante baixo. O mês de fevereiro não foi previsto como recessão já que os valor do seu escore estava mais próximo do centróide do grupo dos uns do que do centróide do grupo dos zeros. O mês de abril a previsão indicou como não-recessivo, apesar de seu escore se encontrar com baixo valor. A partir de maio a previsão indica que a produção industrial vá se distanciando de outra possível recessão até julho, já que em

agosto há outra queda do escore, mas sem ser acusado como uma recessão. De qualquer maneira, é um alerta que a produção industrial do ano de 2013 pode não ser de grandes crescimentos.

6.4.2 Análise Discriminante para Dados Fora da Amostra

Esta subseção tem como objetivo principal verificar o poder do algoritmo para uma amostra de teste, composta por observações não-amostrais. Para alcançar esse objetivo a amostra total dos dados foi dividida em duas partes: uma amostra de desenvolvimento e uma amostra de teste. A amostra total vai de fevereiro de 1992 até dezembro de 2012 e possui duzentas e cinquenta e uma observações. Destas, cento e setenta e nove (71%) foram usadas para a amostra de desenvolvimento e setenta e duas (29%) para a amostra de teste. Ou seja, de fevereiro de 1992 até dezembro de 2006 é o período para a amostra de desenvolvimento e de janeiro de 2007 até dezembro de 2012 é o período para amostra de teste. Dentro da amostra de desenvolvimento ocorreram cinco recessões, com duração total de quarenta e seis meses. Na amostra de teste ocorreram duas recessões, com duração total de vinte e cinco meses.

6.4.2.1 Amostra de desenvolvimento

Antes de efetuar as análises na amostra de teste, é necessário discutir os resultados da amostra de desenvolvimento. A Tabela 23 mostra os valores dos coeficientes discriminantes estimados para cada variável antecedente utilizada no modelo bem como o valor estimado da constante discriminante. É também mostrado a significância das variáveis a partir da estatística Wikis'Lambda.

Tabela 23 - Valores dos coeficientes discriminantes e da constante discriminante

Variável	Coeficiente Discriminante	Valor do Coeficiente Discriminante	Wikis'Lambda	Significância
Ibovespa	β_1	0,008	0,910	0,000
Exportações de Caminhões	β_2	0,011	0,999	0,720
Dívida Mobiliária	β_3	0,090	0,980	0,059
Pontos Bovespa	β_4	0,162	0,690	
Produção de Materiais Elétricos	β_5	0,344	0,829	0,000
Produção de Carvão	β_6	0,383	0,861	0,000
Taxa de Câmbio	β_7	0,302	0,984	0,094
Desaceleração Industrial - EUA	β_8	1,219	0,779	0,000
Desaceleração Industrial - Brasil	β_9	0,388	0,865	0,000
Constante	α	-1,494		

Fonte: Elaboração própria.

O software selecionou uma variável não significativa, que foi as exportações de caminhões para fazer parte do modelo. A razão disto não foi falta de outras variáveis para compor o algoritmo e sim possíveis correlações entre as variáveis.

Como pode ser observado, todas as variáveis apresentam relação direta com a *target variable*. Os coeficientes calculados diferiram um pouco da análise anterior. O caso que mais se destacou foi a Desaceleração Industrial dos EUA que quase dobrou de valor (comparar com a Tabela 17). O segundo maior coeficiente é o da Desaceleração Industrial do Brasil, representando nesta análise quase a metade do coeficiente calculado na análise anterior. Como já afirmado, coeficientes maiores não significam um poder discriminante maior, visto que esse poder depende dos valores das variáveis independentes.

O SPSS® listou as variáveis com melhor poder discriminatório: Pontos Bovespa β_4 , Desaceleração Industrial EUA β_8 , Produção de Materiais Elétricos β_5 , Produção de Carvão β_6 , Desaceleração Industrial Brasil β_9 ,

Ibovespa β_1 , Dívida Mobiliária β_3 , Taxa de Câmbio β_7 e Exportações de Caminhões β_2 .

Com os coeficientes discriminantes e a constante estimados, pode-se calcular o escore discriminante Z_t para cada mês. Calculados os escores, resta defini-los se eles representam uma recessão ou não. A regra discriminante para determinar se um mês é ou não de recessão baseou-se na regra de classificação da distância euclidiana quadrática. Para a construção da regra é necessário ter os valores dos centróides, que foram: -2,145 para o grupo das recessões e de 0,742 para o grupo das não-recessões.

Dada a regra de classificação baseada na distância euclidiana quadrática, 94% das observações da amostra de desenvolvimento foram previstas corretamente, como pode ser visto na Tabela 24 abaixo:

Tabela 24 - Acertos e erros das previsões

Ocorrência	Previsão		Total
	0	1	
Amostra de Desenvolvimento	0	44	46
	1	10	133
	0	96%	4%
	1	8%	92%
			100%

Fonte: Elaboração própria.

A função discriminante conseguiu acertar 96% das recessões e 92% das não-recessões e daí resulta o acerto médio de 94%. Ela deixou de classificar dois casos que eram recessões dos quarenta e seis possíveis, ou seja, não apontou 4% das recessões (erro Tipo I). Dos cento e trinta e três casos de não-recessão, a função errou em dez casos, indicando pois, dez *falsos alarmes da discriminante* (erro Tipo II). Essa previsão apresentou um QPS no valor de 0,13.

A Figura 34 abaixo mostra os alarmes de recessões e as respectivas recessões para os dados deste estudo. As barras pretas representam os alarmes enquanto as barras azuis representam as recessões.

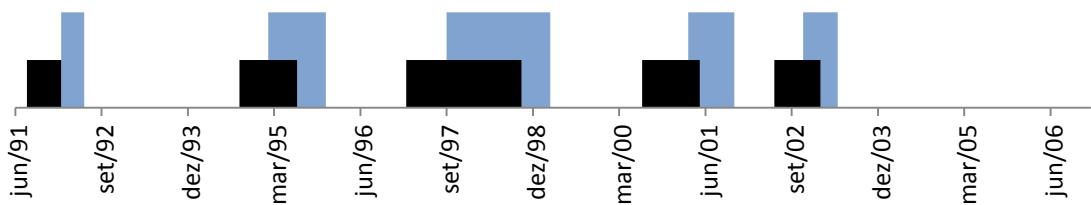


Figura 34 - Alarmes (barras pretas) e recessões (barras azuis), utilizando a regra de classificação através da fórmula (07).

Como pode ser observado, todos os momentos iniciais das recessões conseguiram ser previstos pelo algoritmo. Dos dez erros Tipo II cometidos, nenhum deles representou um *falso alarme preditivo*, eles apenas inflaram o período recessivo, adicionando à previsão de um a dois meses antes do seu início e/ou um ou dois meses no final do período recessivo. Dos dois erros Tipo I nenhum deles atrapalhou de forma acentuada a previsão dos momentos iniciais das recessões. A Tabela 25 abaixo lista em quantos meses o algoritmo conseguiu antecipar cada recessão:

Tabela 25 - Duração dos alarmes antes dos inícios de cada recessão

Recessão	Período do alarme	Duração (meses)
3	Ago/1991 - Jan/1992	6
4	Set/1994 - Jan/1995	5
5	Fev/1997 - Ago/1997	7
6	Jul/2000 - Fev/2001	8
7	Jun/2002 - Out/2002	5

Fonte: Elaboração própria.

Apenas em dois alarmes ocorreu erro Tipo I no momento inicial da previsão. Foram os alarmes da recessão Quatro e Sete. Nas outras três recessões todos os alarmes apresentaram uma duração de pelo menos seis meses. Os alarmes que duram mais que seis meses refletem os erros Tipo II ocorridos nos momentos iniciais da previsão. Dessa forma, dos dez erros Tipo II cometidos, três apresentam-se nos momentos iniciais da previsão, um deles na recessão Cinco e dois na recessão Seis. Os outros sete erros Tipo II

encontram-se nos finais dos alarmes. Desta forma, como não existiu um erro do Tipo II ocorrendo longe das recessões, não houve *falso alarme preditivo*.

6.4.2.2 Amostra de teste

Utilizando o algoritmo construído com os dados da amostra de desenvolvimento, testou-se a capacidade preditiva do modelo na amostra de teste. Dada a regra de classificação baseada na distância euclidiana quadrática, 86% das observações foram previstas corretamente, como pode ser visto na Tabela 26 abaixo:

Tabela 26 - Acertos e erros das previsões

	Ocorrência	Previsão		Total
		0	1	
Amostra de Teste	0	19	6	25
	1	2	45	47
		76%	24%	100%
		4%	96%	100%

Fonte: Elaboração própria.

A função discriminante conseguiu acertar 76% das recessões e 96% das não-recessões e daí resulta o acerto médio de 86%. Ela deixou de classificar seis casos que eram recessões dos vinte e cinco possíveis, ou seja, não apontou 24% das recessões (erro Tipo I). Dos quarenta e sete casos de não-recessão, a função errou em dois, indicando pois, dois *falsos alarmes da discriminante* (erro Tipo II). Essa previsão apresentou um QPS no valor de 0,22.

A Figura 35 abaixo mostra os alarmes de recessões e as respectivas recessões para os dados deste estudo. As barras pretas representam os alarmes enquanto as barras azuis representam as recessões.

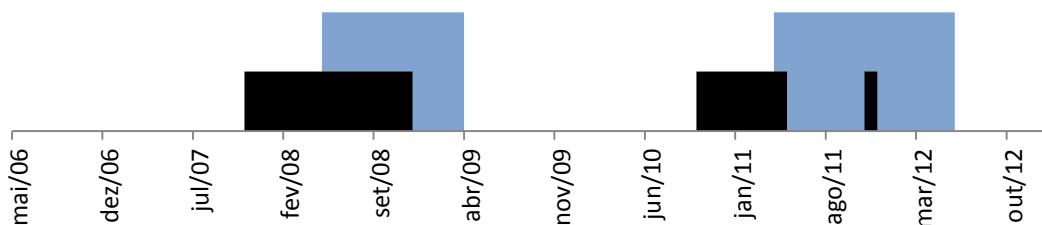


Figura 35 - Alarmes (barras pretas) e recessões (barras azuis), utilizando a regra de classificação através da fórmula (07)

Como pode ser observado, os dois momentos iniciais das recessões conseguiram ser previstos pelo algoritmo. Dos dois erros Tipo II cometidos, nenhum deles representou um *falso alarme preditivo*. A última barra preta não é considerada um *falso alarme preditivo*, já que ela não é um erro Tipo II. Ela representa apenas um acerto de previsão como todos os outros e por estar localizado dentro de uma recessão ele não representa um alarme. O que aconteceu foi que os seis erros do Tipo I apareceram concentrados entre a barra preta mais grossa e a barra preta mais fina, por isso o alarme apresentou essa descontinuidade. Dessa forma, dos seis erros Tipo I nenhum deles atrapalhou a previsão dos momentos iniciais das recessões. A Tabela 27 abaixo lista em quantos meses o algoritmo conseguiu antecipar cada recessão:

Tabela 27 - Duração dos alarmes antes dos inícios de cada recessão

Recessão	Período do alarme	Duração (meses)
8	Dez/2007 - Abr/2008	6
9	Jul/2010 - Mar/2011	6

Fonte: Elaboração própria.

Como pode ser observado, as recessões foram antecipadas pelo algoritmo em exatamente seis meses, não ocorrendo erros do Tipo I nem II nos momentos iniciais da previsão. Se tal algoritmo fosse utilizado para prever as recessões datadas pelo CODACE, as recessões teriam sido antecipadas em 11 e 7 meses, respectivamente (Ver Tabelas 6 e comparar com os inícios dos alarmes mostrados na Tabela 27).

Cabe ainda nesta subseção plotar os valores assumidos por Z_t , valores estes que formam o índice antecedente. A Figura 36 mostra os valores do índice antecedente, formado pelos nove indicadores antecedentes ponderados por seus respectivos escores discriminantes. Os valores do índice estão plotados tanto para os dados da amostra de desenvolvimento como para a amostra de teste. O eixo vertical do gráfico mede os valores deste índice. Como a maioria dos indicadores antecedentes apresentam-se em taxas de crescimento, o valor do índice antecedente comporta-se de forma similar. As áreas hachuradas representam as recessões.

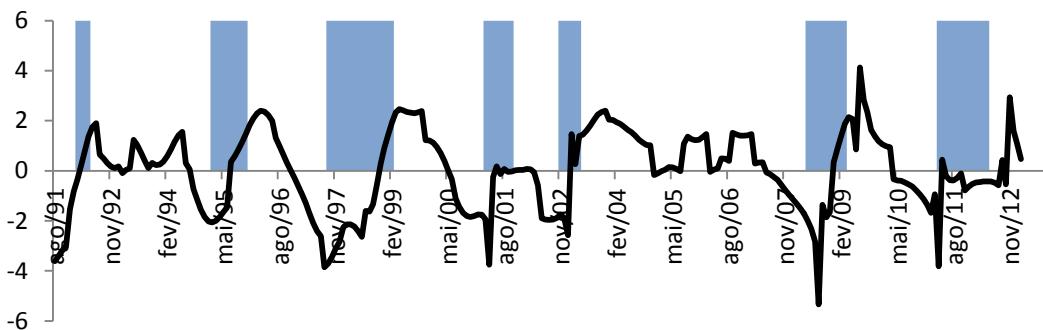


Figura 36 - Índice antecedente das recessões brasileiras.

Como pode ser observado, os valores de Z_t antes de todas as recessões tornam-se negativos. Se a regra de classificação fosse dada pelo ponto de corte, seria possível traçar uma linha horizontal indicando um limiar crítico (como na Figura 5) e assim determinar quando o alarme começaria a soar. Porém, como a regra de classificação é dada pela distância euclidiana quadrática, o alarme só pode ser observado através das Figuras 34 e 35 anteriormente apresentadas. Mas, mesmo assim, só o fato dos valores de Z_t apresentarem-se negativos já é um forte indicativo que este mês será classificado pela regra de classificação como recessivo. Notar que a suavização pretendida com o índice antecedente foi alcançada, visto que ele indicou de forma clara todas as recessões, fato este que não ocorreu analisando-se os indicadores de forma individual.

Por fim, cabe aplicar o algoritmo desenvolvido nesta seção para prever o que ocorrerá na indústria até agosto de 2013. Apesar desta análise está se efetuando em abril de 2013, as séries de produção de carvão, a de material elétrico para veículos e a da própria produção industrial só estão disponíveis

até o mês de fevereiro. Portanto, devido a demora de divulgação dos dados, o horizonte temporal de previsão de uma recessão pode ser diminuído em um ou dois meses. A Tabela 28 abaixo apresenta as previsões do algoritmo até agosto de 2013.

Tabela 28 - Previsão para 2013

Data	Prod. Industrial	Prod. Industrial Filtrada	Dummy	Escore Discriminante (Z_t)	$(Z_t - Z_0)^2$	$(Z_t - Z_1)^2$	Pre vis ão
Jan/2013	128,76	126,85	1	-0,47	2,79	1,47	1
Fev/2013	125,79	126,94	1	-0,57	2,45	1,74	1
Mar/2013				-0,42	6,61	0,09	1
Abr/2013				-0,54	2,57	1,64	1
Mai/2013				2,93	25,77	4,7	1
Jun/2013				1,59	13,95	0,72	1
Jul/2013				1,08	10,43	0,11	1
Ago/2013				0,46	6,82	0,07	1

Fonte: Elaboração própria.

Como pode ser observado, apesar do número índice que mede o nível de produção industrial ter caído de 128,76 para 125,79, uma queda de 2,3%, o valor filtrado do índice teve um pequeno aumento de 0,07%. Ou seja, por pouco não aconteceu uma variação negativa na taxa de crescimento, o que levaria a um novo momento de recessão. Apesar da produção industrial de dezembro de 2013 não está na Tabela 28, ele foi de 125,95 e seu valor filtrado foi de 126,71. Assim, houve uma variação positiva de 0,11% entre os valores filtrados de dezembro e janeiro. Dessa forma, a *dummy* recebeu o valor de um, indicando uma não recessão, assim também ocorrido para o mês de fevereiro. Os demais meses ainda não possuem os dados divulgados. A função discriminante acertou para os dois casos possíveis no ano de 2013. Notar que no mês de fevereiro, o qual quase representou o início de uma recessão, o escore apresentou-se bastante baixo, o menor dos oito escores mostrados na Tabela 28. O mês de fevereiro não foi previsto como recessão já que os valores do seu escore estavam mais próximos do centróide do grupo dos uns do que do centróide do grupo dos zeros. Os meses de março e abril a previsão

indicou como não recessivos, apesar de seus escores se encontrarem com baixos valores. A partir de maio a previsão indica que a produção industrial vá se distanciando de outra possível recessão. Em agosto o escore volta a cair, porém, sem indicativo de recessão. Essa queda pode indicar que para o ano de 2013 o crescimento da produção industrial não seja muito significativo.

6.5 Recessões brasileiras como eventos extremos

Continuando a discussão apresentada no Capítulo 4, intitulado "Eventos Extremos", esta seção tem como objetivo fazer testes estatísticos na série de produção industrial brasileira para verificar se seu comportamento pode ser inserido no arcabouço teórico dos eventos extremos. Essa verificação se dará de três formas. A primeira delas irá realizar testes de normalidade, linearidade e estacionariedade na série de produção industrial brasileira. O intuito principal é verificar se o comportamento dos dados segue ou não uma distribuição normal.

A segunda forma, inserida na subseção 6.5.1 vai analisar a evolução ao longo do tempo do expoente de Hurst e a terceira forma, subseção 6.5.2, irá analisar a evolução do componente auto-regressivo da série da produção industrial. A intenção do estudo do expoente de Hurst é verificar se as observações são ou não dependentes. Da mesma forma, a análise da evolução do componente auto-regressivo estuda este grau de dependência entre as observações, inserida na teoria da desaceleração crítica, que será mais a frente explicada.

A Figura 37 abaixo mostra a evolução da produção industrial brasileira⁴³ que aqui será analisada:

⁴³ Desde janeiro de 1975 até janeiro de 2013. Os dados estão dessazonalizados e transformados em índice com a média de 2002 = 100.

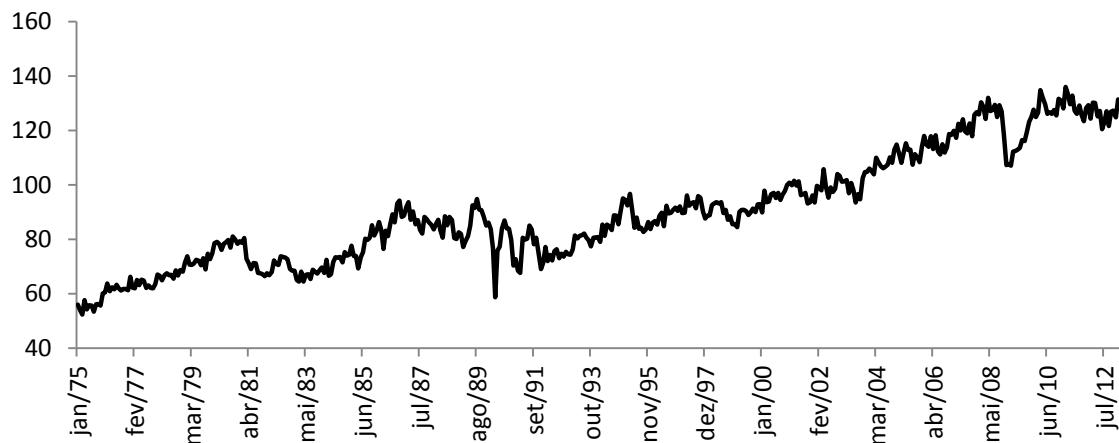


Figura 37 - Produção industrial brasileira.

Para verificar se estes dados se distribuem de forma normal, alguns procedimentos foram realizados. Primeiramente, o histograma da série pode ser visualizado na Figura 38 abaixo e comparado com a curva da distribuição normal.

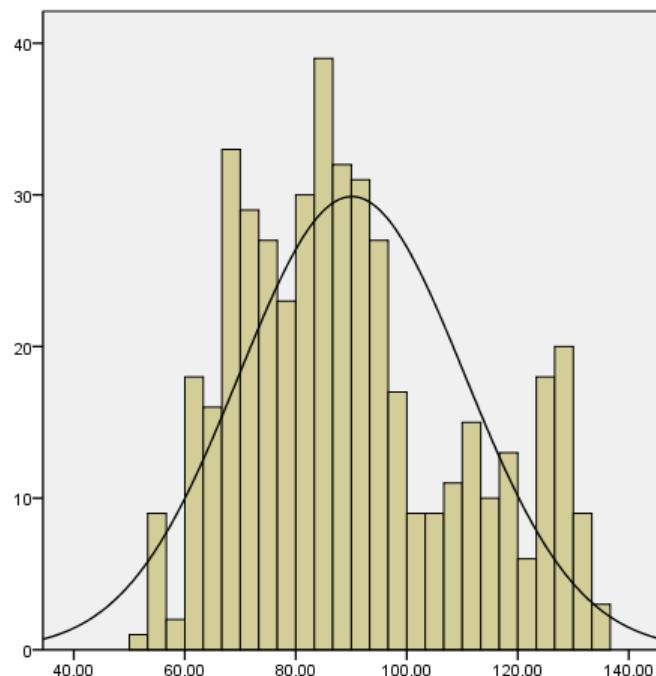


Figura 38 - Distribuição de frequência da série de produção industrial.

Valores de assimetria e curtose devem ser zero quando a distribuição da variável é normal. A produção industrial apresentou uma assimetria no valor de 0,482, indicando que existe uma acumulação de escores do lado esquerdo da distribuição. O valor da curtose foi de -0,687, indicando que a distribuição é

menos pontuda e mais horizontal, provocando "caudas mais pesadas". Foram realizados dois testes estatísticos para verificar se a produção industrial brasileira seguia uma distribuição normal. São os testes Kolmogorov-Smirnov (K-S) e Shapiro-Wilk (S-W) que testam a hipótese de normalidade dos dados. As hipóteses empregadas nos testes de normalidade são as seguintes:

$$H_0: \text{os dados exibem distribuição normal}$$

$$H_1: \text{os dados não exibem distribuição normal}$$

Com valores de 0,08 para o teste K-S e de 0,95 para o teste S-W, ambos rejeitaram a hipótese nula de que os dados exibem distribuição normal à 1% de significância estatística. Dado que a série de produção industrial brasileira não apresenta uma distribuição normal dos dados, provavelmente os dados também não são independentes e identicamente distribuídos. Esse fato foi confirmado através do teste BDS (BROCK et al., 1996) pelo qual a hipótese nula da série ser i.i.d. foi rejeitada, assumindo então a hipótese que os dados são não-lineares⁴⁴. Esse fato pode ter sido ocasionado devido a não-estacionariedade da série, verificada com o teste Dickey-Fuller Aumentado (DFA)⁴⁵.

Portanto, a série de produção industrial brasileira não apresenta uma distribuição normal dos dados, que também não são i.i.d., além da série não ser estacionária. Estes fatos apoiam as críticas dos econofísicos aos métodos de previsão econométricos tradicionais para prever recessões.

6.5.1 Análise do Exponente de Hurst

Esse teste foi elaborado por Hurst (1951) para testar se determinada série de tempo era aleatória, buscando mensurar se a série apresentava persistência em se manter acima ou abaixo de sua média. A essência do teste é medir a flutuação da série em torno da média ao longo do tempo. O alcance

⁴⁴ Em todas as dimensões testadas (2,3,4,5) a hipótese nula foi rejeitada.

⁴⁵ Foram feitos testes de estacionariedade com constante, com constante e tendência e sem constante e sem tendência. Dado que o teste assume a presença de raiz unitária e a hipótese nula foi rejeitada em todos os tipos de testes realizados, chega-se a conclusão que a série não é estacionária em nível.

R_n dessa flutuação depende do tamanho da amostra, e dado que a série seja aleatória, o alcance se mostra proporcional a raiz quadrada do tamanho da amostra. A teoria afirma que, por causa disso, em séries constituídas de informações independentes, o expoente de Hurst tende a ser 0,5. Dado que o alcance depende do tamanho da amostra, com o intuito de tornar a análise adimensional, reescalona-se o alcance R_n dividindo-o pelo desvio padrão S_n da série, daí a designação análise R/S ("rescaled range" ou "*range over standard deviation*"). Sendo a série $\{y_t\}$, o alcance da série é definido como:

$$R_{(n)} = \max \sum_{n=1}^i (y_t - \bar{y}) - \min \sum_{n=1}^i (y_t - \bar{y}) \quad (11)$$

que nada mais é que o intervalo normalizado de variação da série de tempo, isto é, o somatório das diferenças entre os valores máximos e mínimos normalizados das observações y_t em termos da sua média \bar{y} . Dividindo o termo R_n por S_n , Hurst (1951) chegou a seguinte equação:

$$E = \left[\frac{R_{(n)}}{S_{(n)}} \right] = Cn^H \quad (12)$$

em que C é uma constante e H é o expoente de Hurst. Este expoente pode ser interpretado como a probabilidade de no próximo período a série se manter na mesma direção. O motivo pelo qual o expoente de Hurst ajuda a prever a ocorrência de eventos extremos é relativamente simples. O teorema do limite central afirma que variações independentes em dados tendem a se cancelar com o crescimento da amostra. O teorema afirma que em séries constituídas de observações independentes o expoente de Hurst tende a ser 0,5. Quando o expoente difere de 0,5 significa que a série possui algum tipo de memória de longo prazo. Se for inferior, indica um comportamento de reversão à média, chamado de anti-persistência (Mandelbrot e Hudson, 2004). Se for superior, existe persistência na série. Expoentes maiores significam que ao menos parte das variações observadas não estão se cancelando, ou seja, não estão sendo

independentemente distribuídas. Segundo Mandelbrot e Hudson, 2004, a maior parte das séries financeiras apresentam um expoente de Hurst próximo a 0,75.

O expoente de Hurst que aqui será analisado é ao longo do tempo. A literatura⁴⁶ indica que o expoente de Hurst consegue captar a força de dependência de longo prazo da série. Portanto, a série foi filtrada⁴⁷ para obter sua tendência de longo prazo e depois analisou-se os resíduos da série. A Figura 39 abaixo mostra a evolução do expoente de Hurst para os resíduos da tendência de longo prazo da série de produção industrial brasileira⁴⁸.

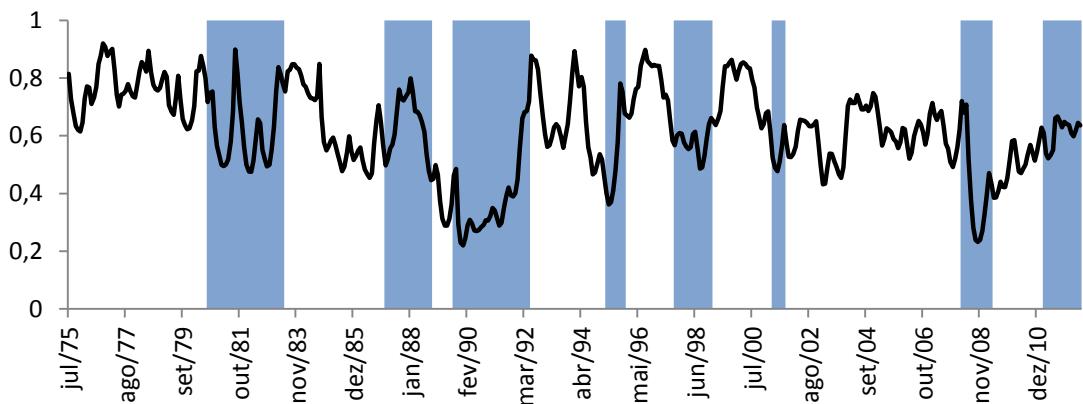


Figura 39 - Evolução do expoente de Hurst para os resíduos da tendência de longo prazo da produção industrial brasileira.

Como pode ser observado, antes de todas as recessões o expoente de Hurst apresentou-se elevado (maior do que 0,75), por vezes apresentando os momentos de queda até antes de começar as recessões. Existe portanto, uma memória de longo prazo na série, indicando que as mudanças que ocorrem na série da produção industrial brasileira não são meramente aleatórias. Este maior grau de dependência dos dados antes dos momentos recessivos será explicado na próxima subseção.

⁴⁶ Ver Ihlen (2012).

⁴⁷ O filtro utilizado foi o mesmo aplicado em todas as séries de tempo até agora já estudadas neste trabalho.

⁴⁸ A metodologia para o cálculo do expoente de Hurst ao longo do tempo seguiu os passos propostos por Ihlen (2012), através do método MFDAF (*Multifractal Detrended Fluctuation Analysis*). São cinco etapas mostradas em seu artigo, com as respectivas rotinas para serem aplicadas no software matemático Matlab®.

6.5.2 Evolução do Componente Auto-Regressivo de Primeira Ordem

Eventos extremos geralmente são precedidos por uma desaceleração crítica⁴⁹, na qual sistemas operando perto de seus *turning points* tornam-se crescentemente mais lentos a perturbações externas (BUENO, 2012). Scherffer et al. (2009) mostram que isso ocorre porque *loops* de *feedback* positivos começam a compensar o efeito estabilizador dos *loops* de *feedback* negativos dominantes em sistemas menos vulneráveis a choques, indicando que os sistemas estão menos resilientes.

Ives (1995) propôs um método para identificar esta desaceleração crítica através do aumento da auto-correlação de curto prazo dos resíduos⁵⁰ da série, através do coeficiente *AR*(1), ou seja, através do parâmetro auto-regressivo de primeira ordem, que determina quão rapidamente o efeito de um choque se dissipia na série. A literatura afirma que sistemas operando próximos a *turning points* tornam-se crescentemente mais lentos na recuperação a perturbações externas e assim apresenta altos valores para o parâmetro auto-regressivo ou valores crescentes deste parâmetro. Para realizar este teste, os resíduos são aqui caracterizados como os desvios da série de produção industrial em relação a sua tendência de longo prazo, que é medida através do filtro polinomial. Dados os resíduos, a seguinte equação foi estimada por MQO, sem constante:

$$e_t = \rho e_{t-1} + \nu_t \quad (13)$$

em que,

⁴⁹ Os teóricos de sistemas dinâmicos relacionam o fenômeno da desaceleração crítica à dificuldade em que determinado sistema tem de se recuperar do efeito de uma perturbação, mesmo pequena. Um exemplo é imaginar uma bola no fundo de um recipiente em que as paredes sejam íngremes. Se a bola for afastada um pouco do fundo, a inclinação da parede do recipiente fará com que ela volte ao seu estado de repouso rapidamente. Mas, se a mesma bola for colocada em um recipiente com paredes menos íngremes, ela poderá levar algum tempo até voltar ao seu estado de repouso. Essa última situação é que os teóricos denominam de desaceleração crítica. Este é um sinal de alerta precoce de que o sistema está se aproximando de uma zona de perigo (CASTI, 2012).

⁵⁰ Os resíduos aqui são caracterizados como os desvios da série de produção industrial em relação a sua tendência de longo prazo, que é medida através do filtro polinomial.

e_t são os resíduos da série de produção industrial após filtrada pela suavização polinomial;

e_{t-1} representa os resíduos defasados em um mês;

ρ é um parâmetro que determina as propriedades de correlação de e_t ; seu intervalo está entre: $-1 \leq \rho \leq 1$

ν_t representa os termos de erro para essa estimativa.

A equação (13) foi estimada para a série de produção industrial de janeiro de 1975 até dezembro de 2012. Foi escolhida a série completa de produção industrial para poder efetuar uma melhor comparação entre as recessões, apesar do fato deste estudo focar nas previsões das recessões a partir de 1991. Para obter a evolução do parâmetro ρ para cada mês foram realizadas 322 regressões através de uma janela deslizante com início em fevereiro de 1975, tendo a primeira regressão com apenas quinze observações, para obter-se o ρ de abril de 1976.

As figuras abaixo apresentam as evoluções do parâmetro ρ momentos antes de todas as recessões da atividade industrial captadas pelo método do filtro polinomial. As áreas hachuradas representam as recessões. É importante observar que o valor do parâmetro ρ não é central para o estudo da desaceleração crítica, interessando a análise de sua trajetória de crescimento ou decrescimento.

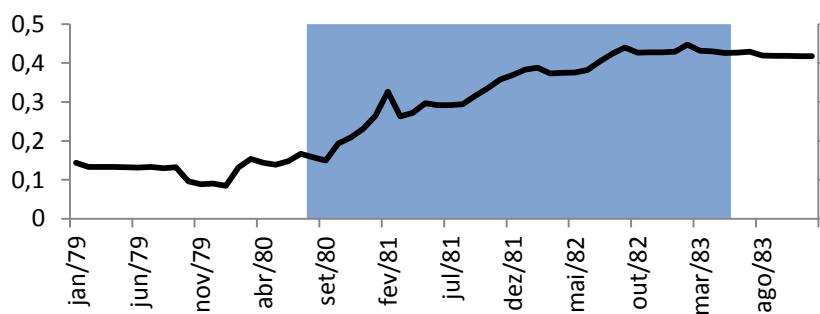


Figura 40 - Evolução do AR(1) antes da recessão de 1980.

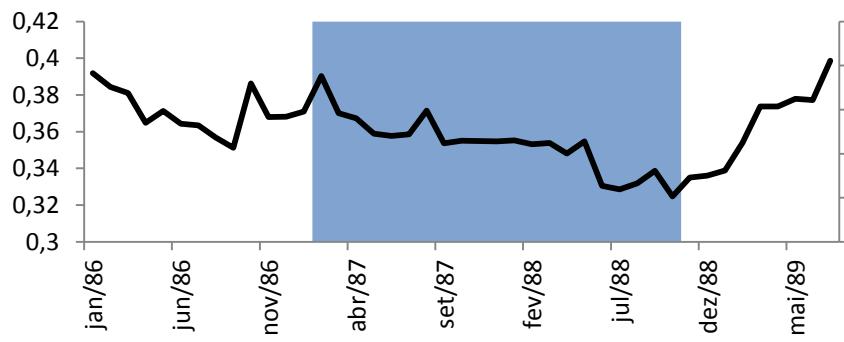


Figura 41 - Evolução do AR(1) antes da recessão de 1987.

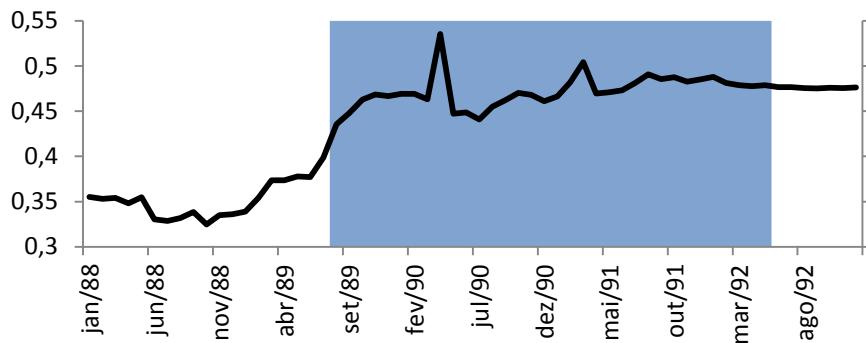


Figura 42 - Evolução do AR(1) antes da recessão de 1989.

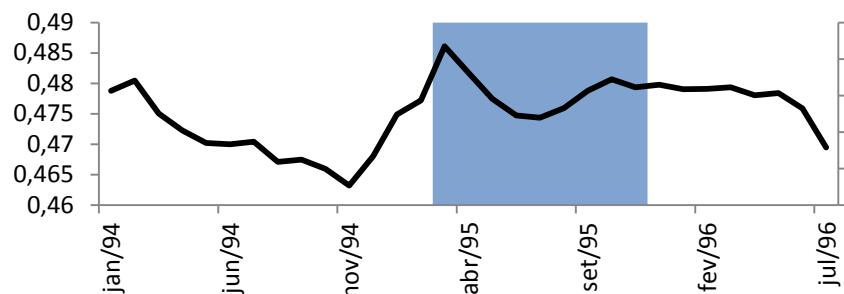


Figura 43 - Evolução do AR(1) antes da recessão de 1995.

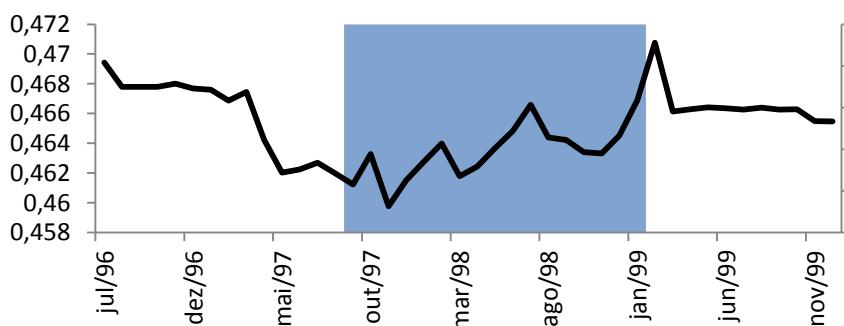


Figura 44 - Evolução do AR(1) antes da recessão de 1997.

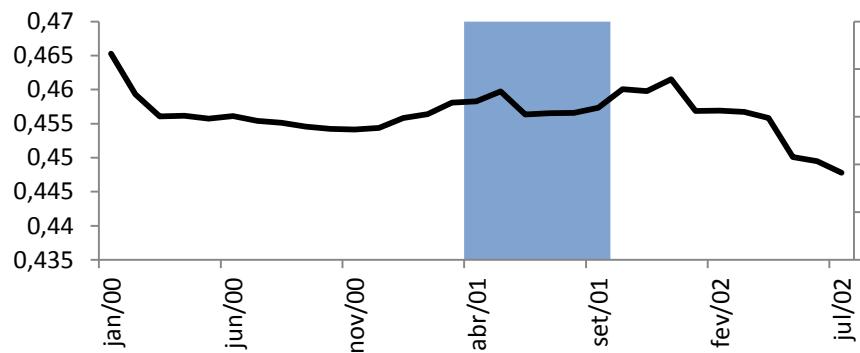


Figura 45 - Evolução do AR(1) antes da recessão de 2001.

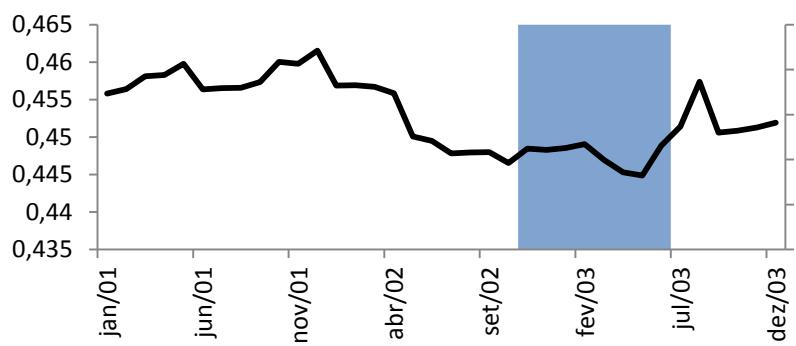


Figura 46 - Evolução do AR(1) antes da recessão de 2003.

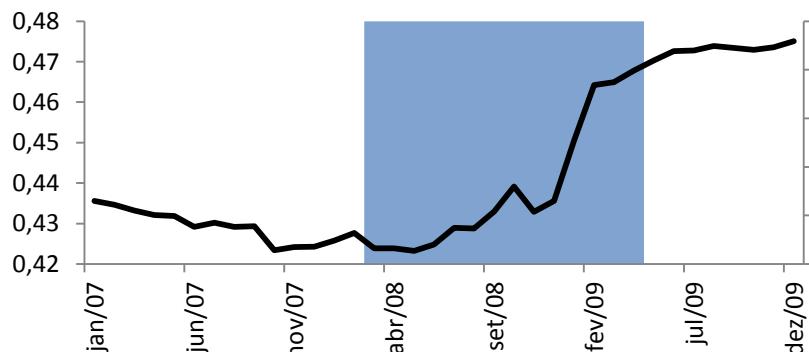


Figura 47 - Evolução do AR(1) antes da recessão de 2008.

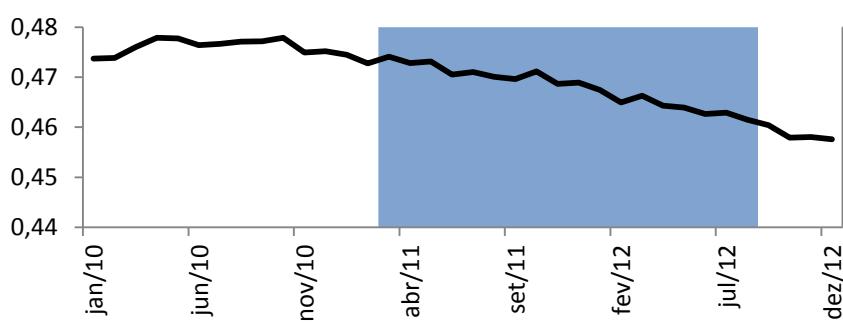


Figura 48 - Evolução do AR(1) antes da recessão de 2011.

Das nove recessões aqui listadas, é possível notar que em seis o valor do parâmetro ρ apresentou uma trajetória de crescimento antes das recessões. Esta trajetória de crescimento pode ser evidente como nas recessões de 1989 e 1995 ou mais suaves, como nas recessões de 1980 e 2001. Antes da recessão de 1987 o valor do ρ aumenta, porém de uma forma mal comportada, com movimentos bruscos. Na recessão de 2008 o valor de ρ começa uma trajetória suave de crescimento cinco meses antes do início da recessão, mas que aumenta de forma acentuada dentro da recessão. Três recessões não apresentaram aumento do ρ meses antes de seu início, que foram as recessões de 1997, 2003 e 2012. A recessão de 2003 não foi captada pelo filtro polinomial na série de produção industrial que começa em 1975 (como pode ser observado na Figura 18), mas como foi captada na série de produção industrial que começa em 1991 e também foi captada pela datação do CODACE para a atividade agregada, tornou-se necessária aqui sua identificação e estudo da evolução do ρ . Quanto a recessão de 1997, pode-se afirmar que o ρ passa a aumentar apenas dentro da recessão. A recessão de 2012 o ρ não apresenta sinais de aumento nem antes nem dentro da recessão. Os motivos de nem todas as recessões terem apresentado aumento no ρ precisam ser melhor estudados, porém, existe um indicativo de que isto seja devido a baixa intensidade dos momentos recessivos, a citar a recessão de 2012, que em níveis de intensidade apresenta-se menos severa que a recessão de 2008.

Fazendo uma análise macro do parâmetro ρ é possível observar outros aspectos. A Figura 49 abaixo mostra esse comportamento macro, que nada mais é que a junção de todos os gráficos anteriormente mostrados em um só.

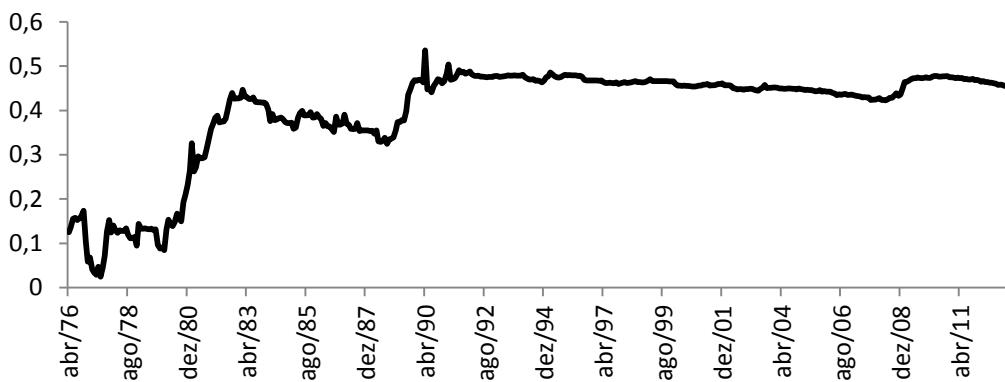


Figura 49 - Evolução do AR(1) da série de produção industrial em relação a sua tendência de longo prazo entre os anos de 1975 e 2012.

É possível observar que durante a recessão de 1980 (que dura até 1983, ver Tabela 7) o valor ρ apresenta uma trajetória de crescimento que começa cinco meses antes do início da recessão e que só acaba em maio de 1983, justamente no mês final da recessão.

Outra trajetória visível de crescimento do ρ se dá na recessão de 1989. Essa trajetória de crescimento antecede a recessão de 1989 em nove meses. O valor do ρ continua a crescer dentro da recessão, cessando este crescimento dez meses após o início da recessão.

Em termos de meses, essas duas recessões (1980 e 1989) foram as mais longas, ambas com 34 meses, e talvez por isso terem se destacado no comportamento macro da evolução do ρ . É possível notar também este fato na recessão de 2008, na qual o valor de ρ passa a aumentar cinco meses antes do seu início e continua aumentando dentro de toda a recessão. Desta forma, pode-se concluir que antes de recessões mais severas o ρ tende a aumentar de forma mais significativa, aumento esse que se dá a maior parte dentro das recessões.

6.6 Recessões na indústria de bens de capital

Esta seção tem como objetivo construir um algoritmo de previsão para as recessões da indústria de bens de capital brasileira. A intenção é comparar a metodologia de previsão apresentada neste estudo com a metodologia de previsão apresentada por Chauvet e Morais (2009). Os resultados do trabalho de Chauvet e Morais já foram expostos no Capítulo 3 deste estudo. Os autores realizaram previsões de curto e longo prazo, com dados amostrais e não-amostrais.

Será aqui realizado o procedimento de previsão de curto prazo, com dados amostrais e não-amostrais, devendo-se deixar claro que para este procedimento não foi realizado um aprofundamento investigativo sobre os melhores indicadores a serem utilizados, apenas utilizou-se dos mesmos indicadores encontrados por Chauvet e Morais, já que a principal intenção é uma comparação de metodologias. Os autores dessazonalizaram as séries utilizadas em seu modelo e o mesmo foi feito aqui.

A série de tempo considerada como a *target variable* será a série de produção industrial de bens de capital⁵¹. A periodicidade vai de janeiro de 1991 até janeiro de 2013 e seu início em 1991 justifica-se pois foi o mesmo utilizado por Chauvet e Morais. Os valores da série foram filtrados através do *gaussian kernel filter*, valores estes que podem ser observados através da Figura 50 abaixo, na qual as áreas hachuradas representam os momentos das recessões. Complementando a Figura 50, os momentos das recessões são listados na Tabela 29, que mostra os meses dos picos e dos vales de todas as recessões.

⁵¹ Dessazonalizada e transformada em índice com a média de 2002 = 100.

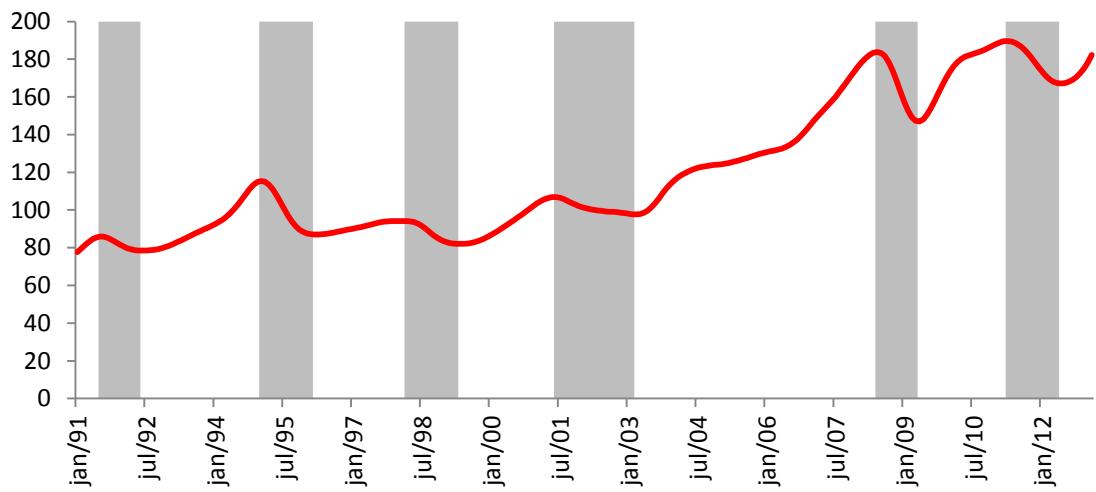


Figura 50 - Valores filtrados da produção de bens de capital desde 1991 até 2012 e datação de seus *business cycles*.

Tabela 29 - Recessões da Indústria de bens de capital desde 1991 até 2012 datadas pelo filtro polinomial

Recessão	Pico	Vale
1	1991:07	1992:06
2	1995:01	1996:03
3	1998:03	1999:05
4	2001:06	2003:03
5	2008:06	2009:05
6	2011:04	2012:06

Fonte: Elaboração própria.

Como pode ser observado, foram identificados seis momentos recessivos. Dos 266 meses filtrados, 81 representaram meses recessivos (30%) e 185 representaram meses não-recessivos (70%). Estes momentos foram bastante semelhantes aos encontrados por Chauvet e Moraes, mostrados na Tabela 3 da seção 3.2. Apenas a recessão seis não foi captada pelos autores, dado que seus dados iam até o ano de 2009.

O procedimento metodológico de previsão foi o mesmo aplicado neste estudo para prever as recessões da produção industrial, ou seja, primeiramente transformou-se a *target variable* em uma variável dicotômica, com os zeros representando as recessões e os valores unitários os momentos de expansão. Em seguida, colheu-se na literatura os indicadores antecedentes

e calculou-se as defasagens ótimas. A seleção final dos indicadores se deu através do método *stepwise* realizado através da análise discriminante. A subseção seguinte entrará nos detalhes da previsão.

6.6.1 Análise de previsão no curto prazo

Para a previsão de curto prazo, Chauvet e Morais utilizaram cinco indicadores antecedentes: produção de artefatos têxteis (-2); produção de produtos químicos inorgânicos (-1, -2, -3, -4); produção de tratores, máquinas e equipamentos agrícolas (-3,-5); produção de material elétrico para veículos (-1) e a produção de carrocerias e reboques (-2). Os números entre parênteses representam as defasagens dos indicadores. Além destes, os autores também utilizaram a variável da produção industrial de bens de capital de forma defasada, não indicando, porém, as defasagens aplicadas.

Para a previsão que aqui será construída adotou-se os seguintes indicadores antecedentes: produção de carvão mineral (-7), produção de produtos químicos inorgânicos (-6); produção de material elétrico para veículos (-6); e a produção de carrocerias e reboques (-6). Também foi incluído como indicador antecedente a própria série de produção de bens de capital defasada em seis meses⁵². Como a menor defasagem dos indicadores é de seis meses, o modelo que aqui será construído se limita à previsão das recessões com seis meses de antecedência. As durações dos alarmes que fujam a este número de meses é porque apresentaram erros do Tipo I ou II nos momentos iniciais da previsão.

Apesar de Chauvet e Morais não terem utilizado a produção de carvão mineral em seu modelo de curto prazo⁵³, esta variável foi utilizada no modelo de longo prazo dos autores. Seu uso no modelo que aqui será construído justificou-se pois apresentou um bom nível de correlação com a *target variable*. O procedimento para escolher as defasagens ótimas foi o mesmo relatado na seção 6.3.

⁵² A transformação utilizada nesta série foi a mesma efetuada para as séries dos momentos de desaceleração das indústrias americana e brasileira.

⁵³ Como relatado no Capítulo 3, só foram utilizados seis indicadores antecedentes em todo o trabalho de Chauvet e Morais. Destes seis, cinco estão presentes no modelo de curto prazo, no qual apenas o indicador que representa a produção de carvão mineral não fez parte do modelo, sendo incluído apenas no modelo de longo prazo, como será visto mais a frente.

A previsão será efetuada utilizando uma parte da amostra (amostra de desenvolvimento) para estimar o modelo e então realizar as previsões para o restante das observações (amostra de teste). Dos 266 meses antes relatados, doze foram retirados devido as defasagens aplicadas nos indicadores antecedentes e mais outro mês foi retirado devido a que todos os indicadores foram utilizados na forma de taxa de crescimento, totalizando assim treze meses a menos. Assim, a *target variable* teve que se apresentar só a partir de fevereiro de 1992, indo até fevereiro de 2013, totalizando 253 meses. Destes, 177 meses (70%) fizeram parte da amostra de desenvolvimento e os 76 meses restantes (30%) fizeram parte da amostra de teste.

Dado isso, o *software* estatístico SPSS® calculou os coeficientes discriminantes para os seis indicadores antecedentes bem como suas significâncias através da estatística Wikis'Lambda. Calculou também o valor da constante discriminante. Tais resultados podem ser analisados através da Tabela 30 abaixo:

Tabela 30 - Valores dos Coeficientes Discriminantes e da Constante Discriminante

Variável	Coeficiente Discriminante	Valor do Coeficiente Discriminante	Wikis'Lambda	Significância
Produção de Carvão Mineral	β_1	0,482	0,721	0,000
Produção de Produtos Químicos	β_2	1,105	0,541	0,000
Produção de Material Elétrico	β_3	0,226	0,680	0,000
Produção de Carrocerias	β_4	0,191	0,881	0,000
Desaceleração da Produção de Bens de Capital	β_5	2,594	0,723	0,000
Constante	α	0,071		

Fonte: Elaboração própria.

Como pode ser observado, todas as variáveis apresentam relação direta com a *target variable* (valores positivos dos coeficientes discriminantes) e todas

apresentaram-se estatisticamente significativas. Dada a *target variable* e os indicadores antecedentes já citados, realizou-se a análise discriminante. Através da regra de classificação da distância euclidiana quadrática o modelo acertou 91,5% dos casos na amostra de desenvolvimento. Os resultados podem ser vistos na Tabela 31 abaixo:

Tabela 31 - Acertos e erros das previsões

	Ocorrência	Previsão		Total
		0	1	
Amostra de Desenvolvimento	0	47	7	54
	1	5	118	123
	0	87%	13%	100%
	1	4%	96%	100%

Fonte: Elaboração própria.

A função discriminante conseguiu acertar 87% das recessões e 96% das não-recessões e daí resulta o acerto médio de 91,5%. Ela deixou de classificar sete casos que eram recessões dos cinquenta e quatro possíveis, ou seja, não apontou 13% das recessões (erro Tipo I). Dos cento e vinte e três casos de não-recessão, a função errou em cinco casos, indicando pois, cinco *falsos alarmes da discriminante* (erro Tipo II). Essa previsão apresentou um QPS no valor de 0,13.

Utilizando o algoritmo construído com os dados da amostra de desenvolvimento, testou-se a capacidade preditiva do modelo na amostra de teste. Dada a regra de classificação baseada na distância euclidiana quadrática, 74,5% das observações foram previstas corretamente, como pode ser visto na Tabela 32 abaixo:

Tabela 32 - Acertos e erros das previsões

	Ocorrência	Previsão		Total
		0	1	
Amostra de Teste	0	17	8	25
	1	10	41	51
	0	68%	32%	100%
	1	19%	81%	100%

Fonte: Elaboração própria.

A função discriminante conseguiu acertar 68% das recessões e 81% das não-recessões e daí resulta o acerto médio de 74,5%. Ela deixou de classificar oito casos que eram recessões dos vinte e cinco possíveis, ou seja, não apontou 32% das recessões (erro Tipo I). Dos cinquenta e um casos de não-recessão, a função errou em dez, indicando pois, dez *falsos alarmes da discriminante* (erro Tipo II). Essa previsão apresentou um QPS no valor de 0,46.

A Figura 51 abaixo mostra os alarmes de recessões e as respectivas recessões tanto para a amostra de desenvolvimento como para a amostra de teste. As barras pretas representam os alarmes enquanto as barras azuis representam as recessões. As quatro primeiras recessões fazem parte da amostra de desenvolvimento e as duas últimas fazem parte da amostra de teste.

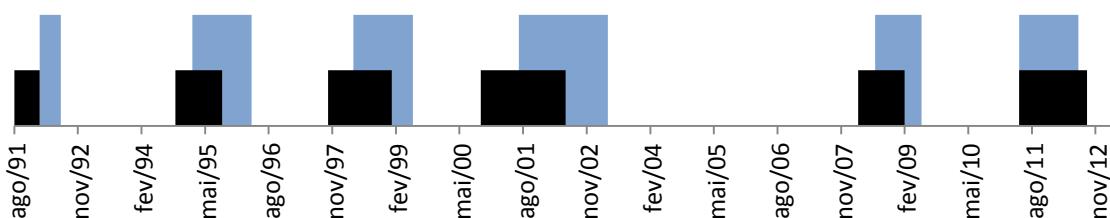


Figura 51 - Alarmes (barras pretas) e recessões (barras azuis) para a indústria de produção de bens de capital.

Como pode ser observado, todas as recessões conseguiram ser previstas, com exceção da última, na qual o alarme só começou a soar exatamente no mês que a recessão se iniciou. A Tabela 33 abaixo lista em quantos meses o algoritmo conseguiu antecipar cada recessão.

Tabela 33 - Duração dos alarmes antes dos inícios de cada recessão para a Indústria de Bens de Capital

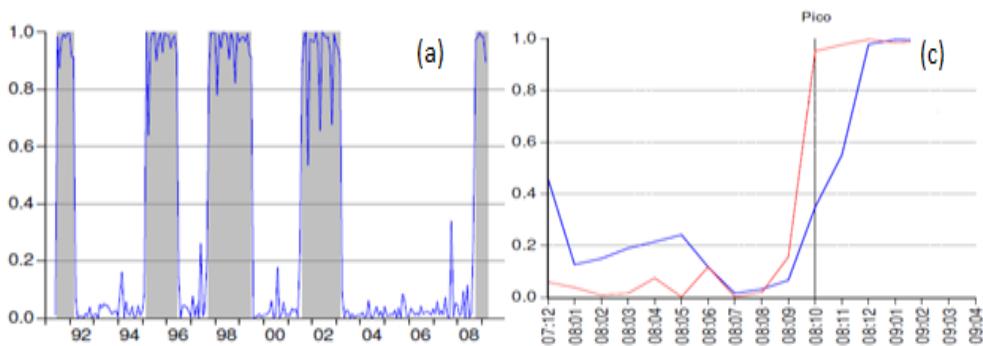
Recessão	Período do alarme	Duração (meses)
1	Ago/1991 - Jan/1992	6
2	Out/1994 - Jan/1995	4
3	Out/1997 - Mar/1998	6
4	Out/2000 - Jun/2001	9
5	Mar/2008 - Jun/2008	4
6	n/d	n/d

Fonte: Elaboração própria.

Observando a Tabela 33, podem ser identificados os erros Tipo I e II que influenciaram a previsão nos momentos iniciais das recessões. Primeiramente em relação as recessões da amostra de desenvolvimento (recessões de 1 a 4), apenas nas recessões dois e quatro houveram erros nos momentos iniciais. Na recessão de número dois apareceram dois erros do Tipo I no momento inicial da previsão, encurtando o alarme que era para ser de seis meses, passando a ser de quatro meses. Na recessão de número quatro apareceram três erros do Tipo II no início da previsão, ampliando o alarme para nove meses.

Em relação à amostra de teste, a recessão de número cinco apresentou dois erros do Tipo I no início do alarme, fazendo com que seu comprimento se resumisse a quatro meses. A recessão de número seis não conseguiu ser prevista pelo algoritmo. O motivo, como já relatado no começo desta seção, é que os indicadores antecedentes não foram construídos especificamente para a metodologia aqui adotada, apenas utilizou-se dos mesmo indicadores propostos por Chauvet e Morais, que utilizaram de outra metodologia de previsão, como já relatado no Capítulo 3.

A comparação do modelo aqui construído com o modelo de previsão de Chauvet e Morais pode ser analisada observando-se a Figura 7 do Capítulo 3, especificamente os gráficos (a) e (c), que representam as previsões de curto prazo para dados amostrais e não-amostrais, respectivamente. Segue os gráficos:



No gráfico (a) é possível observar que nenhuma recessão foi antecipada, dado que a probabilidade de recessão (linha azul) só subiu acima dos 50% quando as recessões (áreas hachuradas) já haviam se iniciado. No gráfico (c), que mostra a previsão da crise de 2008 (linha vertical) com dados não-amostrais, a previsão do modelo probit autoregressivo conseguiu antecipá-la em dois meses (a linha azul se refere a probabilidade de um probit simples e a linha vermelha a probabilidade do probit autoregressivo).

Os autores ainda realizaram previsões de longo prazo, gráficos (b) e (d), que apresentaram um maior sucesso que as previsões de curto prazo. Cabe, em trabalho futuro, uma comparação também com este modelo de longo prazo. O que se pode concluir é que a metodologia apresentada neste estudo mostrou-se promissora para prever as recessões da indústria de produção de bens de capital no curto prazo. Cabe também, no mesmo trabalho futuro, uma investigação mais detalhada acerca dos melhores indicadores antecedentes a serem utilizados de acordo com as especificidades da metodologia aqui apresentada. Espera-se que, por exemplo, a previsão da recessão de 2011 consiga também ser prevista pelo algoritmo, assim como aconteceu com a recessão de 2008 para os dados da indústria geral anteriormente apresentados na seção 6.5.

7 CONCLUSÃO

A previsão dos momentos recessivos é de suma importância para formuladores de política econômica (i.e. políticas fiscais e monetárias), visto o problema das defasagens internas e externas de tais políticas. Sendo a ocorrência de recessões motivada por diversas causas, o estudo de suas previsões diverge em termos teóricos e metodológicos. Os fundamentos teóricos do porque recessões ocorrem não foi objeto principal de estudo deste trabalho, que tratou de analisar de forma pragmática como, nos dias de hoje, uma recessão é definida e quais os meios disponíveis para tentar prevê-las. A teoria que deu sustentação a hipótese aqui formulada foi a de tratar recessões como sendo eventos extremos. Ou seja, a suposição que economias comportam-se de forma anormal antes das recessões encontra sustentação na teoria dos eventos extremos, na qual, em uma de suas suposições, afirma que, como as séries a serem previstas não são independentes e identicamente distribuídas, existe um processo de desaceleração crítica, pelo qual as observações da série, em certos momentos, são dependentes umas das outras. A tendência é que sistemas operando próximos a *turning points* tornam-se crescentemente mais lentos na recuperação a perturbações externas, ocasionando assim o processo de desaceleração crítica, gerado por perdas temporárias de sustentabilidade do sistema econômico.

A série utilizada para se estudar as recessões brasileiras foi a da produção industrial (*target variable*), que se mostrou como uma boa *proxy* dos momentos de recessão quando comparada com as recessões datadas pelo CODACE. O método para datar as recessões, utilizando um filtro na série de produção industrial mostrou-se eficiente não só em relação as datações brasileiras do CODACE, mas também com as datações de recessões americanas do NBER. Assim, por este método alternativo, desde 1991, no Brasil, foram identificadas seis recessões completas, todas elas historicamente conhecidas.

A metodologia de previsão tratou de identificar séries temporais que apresentassem precocemente os movimentos da *target variable*, chamadas de indicadores antecedentes. O fato de haver co-movimento entre a *target variable*

e os indicadores antecedentes não implica uma relação de causalidade, mas que estes são mais sensíveis a responder antes às variações de conjuntura.

No modelo de previsão aqui construído, das seis recessões completas, os dados da série temporal da produção industrial referentes as duas últimas recessões (de 2008 e 2011) não foram utilizados para a construção do algoritmo preditivo. Eles serviram de teste para verificar se o modelo era eficiente na previsão de dados não-amostrais, o que foi verificado, dado que estas duas últimas recessões foram previstas com seis meses de antecedência.

Segundo a teoria de eventos extremos, tais previsões foram possíveis porque a série da produção industrial brasileira não se comporta de forma aleatória. Existe uma persistência nos seus dados antes das recessões, medida pelo expoente de Hurst. Existe também uma desaceleração crítica da série antes das recessões, medida pelo componente auto-regressivo de primeira ordem. Dessa forma, o estudo de recessões na economia brasileira pode ser analisado sobre alguns aspectos da teoria dos eventos extremos, na qual os *econofísicos* estão inseridos.

Comparou-se o modelo aqui construído, um modelo heurístico, com um modelo de previsão de recessões elaborado por Chauvet e Morais (2009), um modelo estatístico-econométrico. O modelo aqui construído adequou-se ao formato do modelo de Chauvet e Morais. Ou seja, procurou-se testar a previsão de recessões na mesma *target variable* (produção de bens de capital) com seus respectivos indicadores antecedentes, diferenciando apenas a metodologia de previsão. Os resultados indicaram uma superioridade do modelo aqui proposto, dado que, para a mesma previsão, a metodologia de Chauvet e Morais só conseguiu antecipar uma das cinco recessões ocorridas na indústria de bens de capital, ocorridas desde 1991. Já o modelo aqui proposto, das mesmas cinco recessões identificadas por Chauvet e Morais, todas conseguiram ser previstas com quatro a nove meses de antecedência.

Também foram realizadas previsões para as recessões Norte-americanas, que obtiveram sucesso na previsão com dados não-amostrais, antecipando as duas últimas recessões americanas (2001 e 2008), datadas pelo NBER, em 13 e 8 meses, respectivamente.

Dado isso, apesar de ser a primeira vez que a metodologia aqui proposta está sendo analisada, os resultados mostraram-se promissores. Evidente que melhorias precisam ser feitas, como dar um maior respaldo teórico a respeito das variáveis a serem utilizadas como indicadores antecedentes. Para trabalhos posteriores fica a sugestão de testar novos indicadores antecedentes, a partir de séries de tempo fornecidas pela base de dados da FGV ou da Serasa Experian, que fornecem dados importantes para a previsão de recessões. A sugestão principal é de encontrar novas séries de indicadores antecedentes que consigam ampliar o horizonte de previsão das recessões brasileiras.

REFERÊNCIAS BIBLIOGRÁFICAS

ALBEVERIO, S.; ALBEVERIO, V.; KANTZ, H. **Extreme events in nature and society**. Nova York: Springer, 2005.

ALTMAN, E.I. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. **The Journal of Finance**, v. 23, No. 4, September, 1968.

ANDERSON, P. Complexity theory and organization science. **Organization science**, v.10, n. 3, p. 216 – 232, Mai./Jun. 1999.

ARTHUR, W. B. Complexity and the Economy. **Science**, v. 284, p. 107 – 109, Abr./1999.

BROCK, W.; DECHERT, W.; SCHEINKMAN, J.; LE BARON, B. A teste for independence based on the correlation dimension. **Econometric Reviews**. V.15, 1996.

BRY, G.; BOSCHAN, C. Cyclical analysis of time series: selected procedures and computer programs. **NBER Technical Paper** 20, 1971.

BUENO, Newton Paulo. Assessing the resilience of small socio-ecological systems from polarity dominance of their feedback structures. **System Dynamics Review**, Cambridge, p. 351-360. 10 nov. 2012.

BURNS, A. F.; MITCHELL, W.C. **Measuring Business Cycles**. The National Bureau of Economic Research, 1946. Disponível em <<http://www.nber.org/books/burn46-1>>.

CARVALHO, F. J. C.; HERMANNY, F. P. Ciclos e previsão cíclica: o debate teórico e um modelo de indicadores antecedentes para a economia brasileira. **Análise Econômica**, Porto Alegre, Ano 21, n. 39, Mar, 2003.

CASTI, J.. **O colapso de tudo**. Rio de Janeiro: Editora Intrínseca, 2012.

CHAUVET, M.; MORAIS, I. A. de M. Indicadores antecedentes para a indústria de bens de capital. **Confederação Nacional da Indústria**, 2009.

CHAUVET, M.; SILVA, J. A. B. da. **Indicadores antecedentes de recessões brasileiras**. In: XXVI ENCONTRO BRASILEIRO DE ECONOMETRIA. João Pessoa, 10, 11, 12 dez. 2004.

CHAUVET, M. An economic characterization of business cycle dynamics with factor structure and regime switching. **International Economic Review**, 1998.

CHEN, P.; KRAUSS, A. Spearman correlation coefficient. In M. Lewis-Beck, A. Bryman, & T. Liao (Eds.). **Encyclopedia of social science research methods**. (pp. 1052-1054). Thousand Oaks, CA: SAGE Publications, 2002.

CODACE. **Cronologia Mensal dos Ciclos de Negócios Brasileiros e Datação de séries conjunturais com algoritmos tradicionais.** Disponível em: <<http://portalibre.fgv.br/main.jsp?lumChannelId=4028808126B9BC4C0126BEA1755C6C93>>. Acesso em: 28 jun. 2013.

CONTADOR, R. C. Ciclos econômicos e indicadores de atividade. Rio de Janeiro, **INPES/IPEA**, 1977.

CONTADOR, R. C.; FERRAZ, C. Previsão com indicadores antecedentes. Rio de Janeiro: **Silcon**, 2001.

CORRAR, L.J.; PAULO, E. FILHO, J.M.D. **Análise multivariada para os cursos de Administração, Ciências Contábeis e Economia.** São Paulo: Atlas, 2011.

CYSNE, R.P.; SIMONSEN, M.H. **Macroeconomia.** São Paulo: Atlas, 2009.
 DAKOS, V.; SCHEFFER, M. van Nes E.H.; BROVKIN V.; PETOUKHOV, V.; HELD, H. Slowing down as an early warning signal for abrupt climate change. **Proceedings of the National Academy of Science**, USA, 105(38), 14308-14312, 2008.

DE VRIES, C.G. The simple economics of bank fragility. **Journal of Banking & Finance**, v. 29, n. 04, p. 803 – 825, Abr./2005.

DIEBOLD, F.X. The Past, Present and Future of Macroeconomic Forecasting. **Journal of Economic Perspectives**. Vol. 12, 1998.

DIEBOLD, F.X. The Past, Present, and Future of Macroeconomic Forecasting. **Journal of Economic Perspectives**, v. 12, p. 175-192, 1997.

DIEBOLD, F.X.; RUDEBUSCH, G.D. Scoring the Leading Indicators. **The Journal of Business**, Vol. 62, No. 3, Jul., 1989.

ESTRELLA, A.; MISHKIN, F. S. Predicting U.S. recessions: financial variables as leading indicators. **Review of Economics and Statistics**, n. 80, p. 45-61, 1999.

FOLHA ONLINE (São Paulo). **Indicadores econômicos dos EUA registram queda em fevereiro.** Publicada em: 20 mar. 2006. Disponível em: <<http://www1.folha.uol.com.br/folha/dinheiro/ult91u106123.shtml>>. Acesso em: 18 abr. 2013.

FONSECA, E.L. **O estudo das propriedades multifractais de séries temporais financeiras.** 2012. 86 f. Dissertação (Mestrado em Modelagem de Sistemas Complexos) – Universidade de São Paulo, São Paulo, 2012.

GLERIA, I.; MATSUSHITA, R.; SILVA, S. Sistemas complexos, criticalidade e leis de potência. **Revista Brasileira de Ensino de Física**, v. 26, n.2, p. 99 – 108, 2004.

GUJARATI, D. **Econometria Básica**. Rio de Janeiro: Elsevier, 2006.

HAIR, J. F. et al. **Análise Multivariada de Dados**. 5. Ed. Porto Alegre: Bookman, 2005.

HAIR, J. F. et al. **Multivariate data analysis**. Nova Jersey: Prentice Hall, 1984.

HAMILTON, J. D. Calling recessions in real time. **International Journal of Forecasting**, Elsevier, vol. 27(4), pages 1006-1026, 2011.

HOBSBAWM, E. **Era dos Extremos: o Breve Século XX**. Cia das Letras, 1994.

HSIEH, D.; FUNG, W. The risk in hedge fund strategies: theory and evidence from trend followers. **Review of Financial Studies**, v. 14, p. 313-341, 2001.

HURST, E. Long term storage capacity of reservoirs. **Transactions of the American society of civil engineers**. V.116, p.770-7999, 1951.

IHLEN, V. Introduction to multifractal detrended fluctuation analysis in matlab. **Frontiers in Physiology** ;3:141, 2012.

ISSLER, J. V.; HOLLAUER, G. Construção de Indicadores Antecedentes para a Atividade Industrial Brasileira e Comparação de Metodologias. **Texto para discussão do IPEA**, TD 1191, IPEA, 2006.

ISSLER, J. V.; NOTINI, H. H.; RODRIGUES, C. F. Um indicador coincidente e antecedente para a atividade econômica brasileira. **Ensaios Econômicos**, nº 295, FGV, Rio de Janeiro, 2009.

ISSLER, J. V.; NOTINI, H.H; RODRIGUES, C. F.; SOARES, A. F. Constructing Coincident Indices of Economic Activity for the Latin American Economy. **Revista Brasileira de Economia**, v. 67 n. 1 / p. 67–96 Jan-Mar, 2013.

IVES, A. R.. Measuring resilience in stochastic-systems. **Ecological Monograph**, Washington Dc, p. 217-233. 1 maio, 1995.

JUNIOR, A. M. C. C. **Indicadores Antecedentes De Atividade Industrial No Brasil**. Dissertação de mestrado da Fundação Getúlio Vargas (FGV). Biblioteca Virtual da FGV, 2008.

KAMINSKY, G. L.; LIZONDO, J. S.; REINHART, C. M. Leading Indicators of Currency Crises. **International Monetary Fund**, Staff Paper 45, p. 1-48, 1998.

KEILIS-BOROK, V.; SOLOVIEV, A.A.; INTRILIGATOR, M. D.; WINBERG, F. E. Pattern of Macroeconomic Indicators Preceding the End of an American Economic Recessions. **Pattern Recognition**, 2008.

KEILIS-BOROK, V.; STOCK, J.H.; SOLOVIEV, A.; MIKHALEV, P. Pre-recession pattern of six economic indicators in the USA. **Journal of Forecasting**, v. 19, p. 65 – 80, 2000.

KEILIS-BOROK, V.I; SOLOVIEV, A.A.; ALLÈGRE, C.B.; SOBOLEVSKII, A.N.; INTRILIGATOR, M.D. Patterns of macroeconomic indicators preceding the unemployment rise in Western Europe and the USA. **Pattern Recognition**, v. 38, 2005.

KAHNEMAN, D. **Rápido e Devagar - Duas Formas de Pensar**. Rio de Janeiro: Editora Objetiva, 2011.

KOEDIJK, K. Schafgans, M. ; deVries, C. The tail index of exchange rate returns. **Journal of International Economics**, 29: 93-108, 1990.

KOSSOBOKOV, V.G.; SOLOVIEV, A. A. Prediction of extreme events: Fundamentals and prerequisites of verification. **Russian Journal of Earth Sciences**, v. 10, 2008.

LAHIRI, K., MOORE, G.H. The leading indicator approach in leading economic indicators. **New approaches and forecasting methods**. Cambridge University Press, 1991.

LIMA, I. C.; MORO, S.; JUNIOR, F. G. J. Ciclos e previsão cíclica: um modelo de indicadores antecedentes para a economia brasileira. **Encontro Nacional da ANPEC**, 2006.

LIRA, S.A. **Análise de correlação: abordagem teórica e de construção dos coeficientes com aplicações**. 2004. 196 f. Dissertação (Mestrado em Métodos Numéricos em Engenharia) – Universidade Federal do Paraná, Curitiba, 2004.

MANDELBROT, B.; HUDSON, R.L. **The (mis) behavior of markets: a fractal view of financial turbulence**. Nova York: Basic Books, 2004.

MANDELBROT, Benoit. The variation of certain speculative prices. **Journal Of Business**, Chicago, p. 394-419. 01 out. 1963a.

MANDELBROT, Benoit. New methods in statistical economics. **Journal Of Political Economy**, Cambridge, p. 421-440. 01 fev. 1963b.

MANTEGNA, R. N.; STANLEY, H. E. **An Introduction to Econophysics: Correlations and Complexity in Finance**. Cambridge: Cambridge University Press, 1999.

MARSH, L.C.; CORMIER D.R. **Spline Regressions Models**. California: Sage Publications, Inc, 2002.

MEYERS, R.A. **Complex Systems in Finance and Econometrics**. Volume 1. Nova York: Ed. Springer, 2009.

MOORE, G. H. New developments in leading indicators in leading economic indicators. **New approaches and forecasting methods**. Cambridge University Press, 1991.

NBER. **Business Cycle Expansions and Contractions**. Disponível em: <<http://www.nber.com/cycles.html>>. Acesso em: 28 jul. 2013.

NORUSIS, M. J. **The SPSS Guide to Data Analysis for Release 4**. Chicago: SPSS Inc., 1990.

PAULO, E.; DIAS FILHO, J.M (Coord.). Análise Multivariada. São Paulo: Atlas, 2011.

PICCHETTI, P., TOLEDO, C. Estimating and interpreting a common stochastic component for the Brazilian industrial production index. **Revista Brasileira de Economia**, v. 56, p. 107-120, 2002.

REINHART, C.M.; ROGOFF, K.S. **This time is different**. Nova Jersey: Princeton University Press, 2009.

SCHEFFER, M.; BASCOMPTE, J.; BROCK, W.; BROVKIN, V.; Carpenter, S.; Dakos, V.; Held, H.; Van Ness, E.; Rietkerk, M. and G. Sugihara. Early-warning signals for critical transitions. **Nature**, 461 (3): 53-59, 2009.

SIMONSEN, M.H.; CYSNE, R.P. **Macroeconomia**. São Paulo: Atlas, 2009.

STOCK, J.H.; WATSON, M. W. A probability model of the coincident economic indicators. **NERBER Working Paper**, nº 2.772, 1988.

STOCK, J.H.; WATSON, M. W. New indexes of coincident and leading economic indicators. **NERBER Macroeconomics Annual**, p. 351-395, 1989.

STOCK, J.H.; WATSON, M. W. A procedure for predicting recessions with leading indicators: econometric issues and recent experience. In: STOCK, J., WATSON, M. (eds.). New research on business cycles, indicators and forecasting. **Chicago: University of Chicago Press**, 1993a.

STOCK, J.H.; WATSON, M. W. New research on business cycles, indicators and forecasting. **Chicago: University of Chicago Press**, 1993b.

STOCK, J.H.; WATSON, M. W. Forecasting using principal components from a large number of predictors. **Journal of the American statistical association**, v.97, nº 460, p. 1167-1179, dec., 2002.

TALEB, N. **A lógica do cisne negro – o impacto do altamente improvável**. Rio de Janeiro: BestSeller, 2007.

TCB. The Conference Board Leading Economic Index (Lei) for the United States. **U.S. Business Cycle Indicators** (United States Of America). Washington, 2013. Disponível em: <http://www.conferenceboard.org/pdf_free/press/TechnicalPDF_4850_1371717637.pdf>. Acesso em: 20 jun. 2013.

TOLEDO, C.; PICCHETTI, P. Estimating and interpreting a common stochastic component for the Brazilian industrial production index. **Revista Brasileira de Economia**, São Paulo, v. 56, p.107-120, 2002.

VIEIRA, E. de O.; PRATES, H. T.; PEREIRA, J. R. B.; SILVA, G. F. da; DUARTE, F. V.; NERES, P. M. Avaliação da contaminação do carbofurano nos solos do Distrito de Irrigação do Gorutuba. **Rev. Brasileira de Engenharia Agrícola Ambiental**, vol.13 no.3, Campina Grande, Mai/Jun, 2009.

ZALIAPIN, I.; KEILIS-BOROK, V. Premonitory spreading of seismicity over the faults' network in southern California: Precursor Accord. **Journal Of Geophysical Research**, Los Angeles. 10 out. 2002. Disponível em: <<http://wolfweb.unr.edu/~zal/pubs/Accord.pdf>>. Acesso em: 20 jun. 2013.

ZARNOWITZ, V.; MOORE, G.H. Sequential Signals of Recessions and Recoveries. **Journal Of Business**, Chicago, p. 22-37. 5 abr. 1982.

ZARNOWITZ, V.; OZYILDIRIM, A. Time series decomposition and measurement of business cycles, trends and growth cycles. **Journal of Monetary Economics**. Volume 53, páginas 1717-1739, 2006. Disponível em: <<http://www.nber.org/papers/w8736.pdf>>. Acesso em: 20 jun. 2013.

APÊNDICE

DATAÇÃO DOS MOMENTOS RECESSIVOS DA PRODUÇÃO INDUSTRIAL AMERICANA

Com o intuito de prever as recessões americanas, o mesmo procedimento de previsão aplicado em toda dissertação foi aqui utilizado. Datou-se as recessões na série da produção industrial americana (*target variable*), fornecida pelo *Board of Governors of the Federal Reserve System* (BGRFS). A série contém 653 observações, compreendendo o período entre janeiro de 1959 e maio de 2013. A Figura 23 abaixo mostra os valores filtrados desta série, na qual as áreas hachuradas representam os momentos de recessão. Complementando as informações da referida figura, a Tabela 34 apresenta os *turning points* da datação.

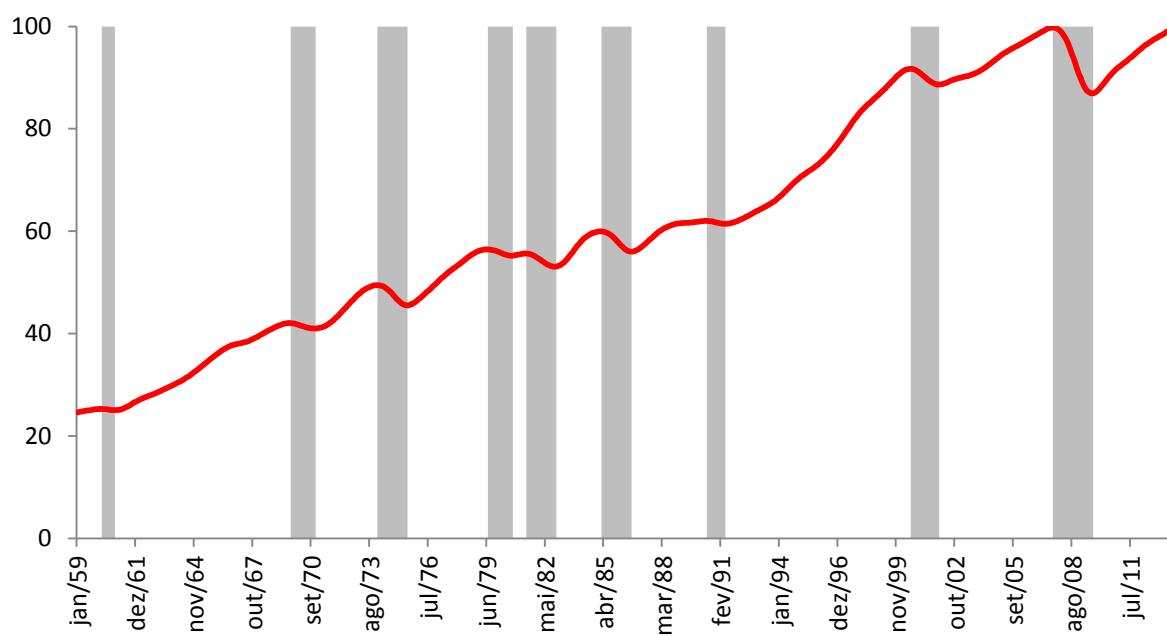


Figura 52 - Valores filtrados da produção industrial americana (linha vermelha) de 1959 a 2013 e a datação de seus *business cycles* (áreas hachuradas).

Tabela 34 - *Turning Points* dos *business cycles* da produção industrial americana
datados pelo filtro polinomial - 1959 a 2013

Recessão	Picos	Vales	Meses em recessão
1	Março/1960	Novembro/1960	8
2	Agosto/1969	Novembro/1970	15
3	Dezembro/1973	Junho/1975	18
4	Junho/1979	Setembro/1980	15
5	Maio/1981	Novembro/1982	18
n/d	Fevereiro/1985	Agosto/1986	18
6	Maio/1990	Abril/1991	11
7	Julho/2000	Dezembro/2001	17
8	Agosto/2007	Agosto/2009	24
Duração total de meses			144

Fonte: Elaboração própria com dados do BGRFS.

Os números de cada recessão são vinculados às recessões americanas datadas pelo NBER, mostradas na Tabela 13. As datações são bastante semelhantes aquelas já apresentadas anteriormente através da Figura 22 e da Tabela 14 e consequentemente, também semelhantes às recessões do NBER.

As datações inseridas na Tabela 14 foram realizadas com os dados da produção industrial fornecidos pelo BGRFS, mas que foram colhidos no *site* do NBER, no *link* que disponibilizava as planilhas de todas as séries utilizadas na datação da recessão de 2008. Neste *link* encontra-se a série do índice da produção industrial americana de 1959 a 2010. Observando a Figura 22, pode ser notado que não houve nenhum caso de uma recessão na atividade industrial que também não fosse uma recessão na atividade agregada. Porém, quando aplicada a mesma metodologia de datação na série que vai de 1959 até 2013, fornecida pelo mesmo BGRFS (colhida através do depositário *Economagic*), surge uma nova recessão que se inicia em 1985, que pode ser observada na Figura 23 e na Tabela 16 acima. Devido a este fato, esta recessão também será alvo do modelo de previsão.

Com o intuito de justificar as previsões que aqui serão feitas, faz-se importante analisar os momentos de desaceleração da série de produção industrial aqui apresentada. A Figura 24 abaixo mostra estes momentos de

desaceleração através das barras vermelhas. As barras azuis indicam os momentos recessivos da produção industrial.

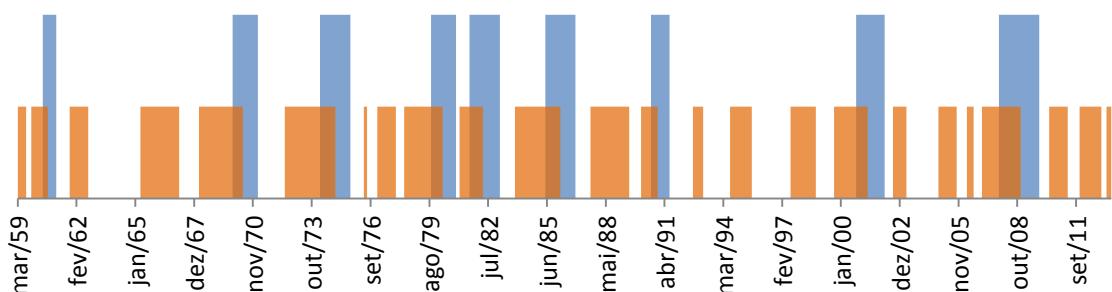


Figura 53 - Momentos de desaceleração da produção industrial americana - 1959 a 2012.

Como pode ser observado, antes de toda recessão existe um momento de desaceleração. Porém, existem diversos momentos de desaceleração que não se traduzem em recessões. Este fato justifica a adoção de um modelo de previsão de recessões que leve em conta outras variáveis, não apenas os momentos de desaceleração da série. Uma questão que deve ser abordada é: por que não prever os momentos de desaceleração ao invés dos momentos de recessão? A resposta se encontra exatamente no fato de que diversas desacelerações não se traduzem em recessões e deste modo o comportamento anormal dos indicadores antecedentes não é verificado. Desde 1959 a economia americana passou por vinte e quatro momentos de desaceleração e apenas nove se reverteram em recessão.

Outra questão que deve ser abordada é: por que não prever os momentos de desaceleração que se revertem em recessão? A resposta é que, a priori, ainda não se tem uma metodologia de indicadores antecedentes que consigam distinguir esses dois tipos de desaceleração. O que se pode fazer é, através do método da tentativa e do erro, testar novos indicadores antecedentes que se comportem de forma anormal antes destes momentos de desaceleração. Por exemplo, criar-se-ia uma *dummy* na qual os zeros seriam os momentos de desaceleração pré-recessão e depois testar-se-ia o nível de correlação desta *dummy* com diversos indicadores antecedentes. Fica, porém, esta sugestão para trabalhos futuros.

SELEÇÃO DOS INDICADORES ANTECEDENTES

As séries candidatas a antecedentes foram as mesmas utilizadas pelo TCB. Porém, das dez séries utilizadas por este instituto, apenas cinco conseguiram ser colhidas pelo autor com a abrangência temporal necessária, através do depositário *Economagic*. Das seguintes dez séries utilizadas pelo TCB, as séries 1,2,4,6 e 9 serão utilizadas neste modelo:

- 1) a semana de trabalho média dos operários da indústria;
- 2) a média dos pedidos iniciais semanais de seguro-desemprego ;
- 3) novas encomendas de bens e materiais de consumo;
- 4) ISM índice de novas encomendas;
- 5) novas encomendas de bens de capital que não seja para defesa;
- 6) concessões de novas autorizações de construção;
- 7) índices de preços de títulos imobiliários;
- 8) oferta monetária (M2) ajustada à inflação;
- 9) diferencial da taxa de juros de longo e curto prazo;
- 10) índice de expectativas do consumidor.

Tais séries foram colhidas já dessazonalizadas pelo próprio depositário. Somando-se a estas séries, também utilizou-se três diferentes transformações na série da produção industrial que conseguissem captar seus momentos de desaceleração. A Tabela 17 abaixo mostra as defasagens ótimas de cada série candidata a fazer parte do algoritmo preditivo, bem como o nível de correlação de Spearman e a transformação aplicada na variável.

Tabela 35 - Defasagens e correlações das variáveis antecedentes com a variável *dummy* das recessões

Série	Defasagem ótima	Correlação	Transformação
Pedidos de seguro desemprego	-6	-0,626**	Taxa de crescimento (filtro)
Diferencial de juros	-9	0,473**	Valores filtrados
Índice de novas encomendas	-10	0,560**	Taxa de crescimento (filtro)
Horas trabalhadas	-6	0,596**	Taxa de crescimento (filtro)
Permissão de novas construções	-11	0,528**	Taxa de crescimento (filtro)
Desaceleração da prod. industrial americana (I)	-6	0,556**	Taxa de crescimento (filtro)
Desaceleração da prod. industrial americana (II)	-6	0,534**	Diferenças das taxas de crescimento do filtro ¹
Desaceleração da prod. industrial americana (III)	-7	0,416**	Dummy ²

Fonte: Elaboração própria.

** Correlação significativa ao nível de 1%. ¹ Essa diferença permite encontrar os momentos de desaceleração da série. ² Variável dicotômica na qual os zeros representam as desacelerações e os valores unitários as acelerações. Tais valores foram mostrados na Figura 24. Essa demarcação foi feita com os dados da Desaceleração (II).

O número de defasagens foi escolhido a partir do maior nível de correlação de Spearman entre a *target variable* e os indicadores antecedentes. A Tabela 18 abaixo mostra o resultado do segundo processo de escolha dos indicadores, o método *stepwise* de seleção, já dentro da análise discriminante, realizado através do software SPSS®.

As variáveis mostradas na tabela a seguir são as que farão parte do algoritmo preditivo. Ela mostra os valores dos coeficientes discriminantes estimados para cada variável antecedente utilizada no modelo bem como o valor estimado da constante discriminante. É também mostrado a significância das variáveis a partir da estatística Wikis'Lambda.

Tabela 36 - Valores dos coeficientes discriminantes e da constante discriminante

Variável	Coeficiente Discriminante	Valor do Coeficiente Discriminante	Wikis'Lambda	Significância
Pedidos de seguro desemprego	β_1	-0,524	0,544	0,000
Diferencial de juros	β_2	0,242	0,704	0,000
Desaceleração da prod. industrial americana (I)	β_3	1,200	0,724	0,000
Desaceleração da prod. industrial americana (II)	β_4	9,815	0,765	0,000
Constante	a	-0,398		

Fonte: Elaboração própria.

Como pode ser observado, todas as variáveis apresentaram-se significativas. Das oito variáveis testadas, apenas quatro fizeram parte do modelo. Destas quatro, apenas os pedidos de seguro desemprego não tem uma relação direta com a *target variable*, o que já era esperado. Serão estas variáveis, com estes respectivos coeficientes, que vão construir a função discriminante, que servirá para prever tanto os dados da amostra de desenvolvimento quanto da amostra de teste.

ANÁLISE DISCRIMINANTE: PREVISÃO

A previsão se dará em duas partes. Na primeira, amostra de desenvolvimento, será construído o algoritmo de previsão (os coeficientes deste algoritmo já foram adiantados na subseção anterior). Na segunda será testado esse algoritmo com dados não-amostrais, ou seja, a amostra de teste.

Na amostra de desenvolvimento os alvos de previsão são as recessões de número 2, 3, 4, 5, a recessão "n/d" e a recessão 6. Na amostra de teste os alvos de previsão serão as recessões de número 7 e 8. Os dados vão de fevereiro de 1968 a maio de 2013, com 544 observações. A amostra de desenvolvimento possui 378 observações (70%), indo de fevereiro de 1968 a julho de 1999. A amostra de teste possui 166 observações (30%), indo de agosto de 1999 até maio de 2013.

As previsões que aqui serão mostradas foram feitas a partir da equação 07, ou seja, através da regra de classificação da distância euclidiana quadrática. O centroide do grupo das recessões Z_0 foi calculado no valor de -1,996 e o centróide do grupo das não-recessões Z_1 foi calculado no valor de 0,670.

Amostra de Desenvolvimento

Dada a regra de classificação baseada na distância euclidiana quadrática, 88,5% das observações da amostra de desenvolvimento foram previstas corretamente, como pode ser visto na Tabela 19 abaixo:

Tabela 37 - Acertos e erros das previsões

	Ocorrência	Previsão		Total
		0	1	
Amostra de Desenvolvimento	0	83	12	95
	1	26	257	283
	0	87%	13%	100%
	1	10%	90%	100%

Fonte: Elaboração própria.

A função discriminante conseguiu acertar 87% das recessões e 90% das não-recessões e daí resulta o acerto médio de 88,5%. Ela deixou de classificar doze casos que eram recessões dos noventa e cinco possíveis, ou seja, não apontou 13% das recessões (erro Tipo I). Dos duzentos e oitenta e três de não-recessão, a função errou em vinte e seis casos, ou seja, apontou 10% das não-recessões como recessões, gerando falsos alarmes da discriminante (erro Tipo II). Essa previsão apresentou um QPS no valor de 0,20.

A Figura abaixo mostra os alarmes de recessões e as respectivas recessões para os dados da amostra de desenvolvimento. As barras pretas representam os alarmes enquanto as barras azuis representam as recessões.

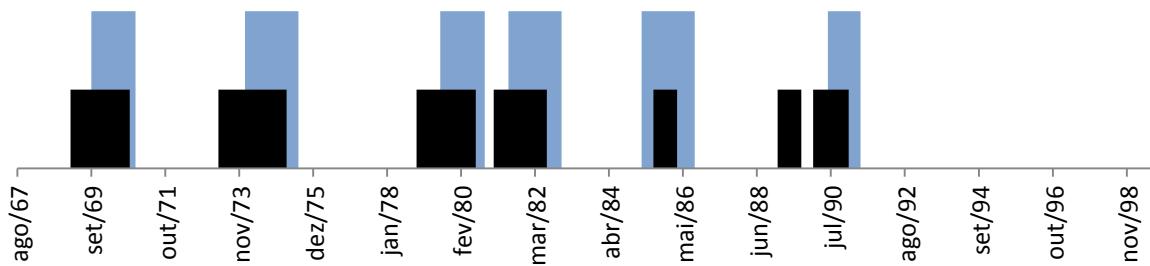


Figura 54 - Alarmes (barras pretas) e recessões (barras azuis) para a amostra de desenvolvimento.

Como pode ser observado, a única recessão que não conseguiu ser antecipada foi a recessão de 1985. Esta não antecipação pode residir no fato que é exatamente esta recessão que não se traduziu em uma recessão da atividade agregada datada pelo NBER.

Dos vinte e seis erros Tipo II cometidos, cinco representaram *falsos alarmes preditivos* e estão inseridos sexta barra preta da Figura (antes da recessão de 1990). Este alarme tem uma duração de oitos meses. Apenas os três últimos é que são considerados alarmes. O motivo é que a recessão em questão só inicia-se em junho de 1990 e, como já abordado na metodologia, alarmes com mais de 12 meses antecedendo uma recessão são considerados *falsos alarmes preditivos*. Os demais erros do Tipo II apenas inflaram o período recessivo, por vezes aumentando a abrangência temporal dos alarmes em antecipar os momentos iniciais das recessões.

Dos doze erros Tipo I nenhum deles atrapalhou de forma acentuada a previsão dos momentos iniciais de recessão. A Tabela abaixo lista em quantos meses o algoritmo conseguiu antecipar cada recessão:

Tabela 38 - Duração dos alarmes antes do início de cada recessão

Recessão	Período do alarme	Duração (meses)
2	Fev/1969 - Ago/69	7
3	Abr/1973 - Dez/73	9
4	Nov/1978 - Jun/1979	8
5	Jan/1981 - Mai/1981	5
n/d*	-	-
6	Jun/1989 - Mai/1990**	12

Fonte: Elaboração própria.

* Esta recessão foi apenas captada na produção industrial, não tendo referência com nenhuma recessão da atividade agregada datada pelo NBER. Ela é referente aos anos de 85/86 e não conseguiu ser antecipada pelo modelo de previsão. ** Os dois alarmes que antecedem a recessão 6 foram unificados, já que dizem respeito a mesma recessão. A duração de doze meses colocada na tabela levou em conta a descontinuidade do alarme, anteriormente explicada através da Figura 54.

Dado que a defasagem mínima dos indicadores é de seis meses, o alarme ideal deve anteceder cada recessão nos exatos seis meses. Caso difira disto, é porque existiram erros do Tipo I ou II da análise discriminante. Como pode ser visto na Tabela, apenas na recessão de número 5 é que o alarme durou menos do que o ideal, devido a um erro Tipo I em seu início. Os demais alarmes que duram mais que seis meses são ocasionados por erros do Tipo II em seus momentos iniciais.

Tabela 39 - Duração dos alarmes antes dos inícios das recessões da atividade agregada datadas pelo NBER

Recessão	Mês inicial do alarme	Mês inicial da recessão do NBER	Antecedência em relação às recessões do NBER
2	Fev/1969	Jan/1970	11
3	Abr/1973	Dez/1973	8
4	Nov/1978	Fev/1980	15
5	Jan/1891	Ago/1981	7
6	Jan/1989	Ago/1990	19

Fonte: Elaboração própria.

Cabe informar que esta análise de antecipação das recessões da atividade agregada pode ultrapassar os doze meses, já que, como não foi objetivo direto prever estas recessões, faz-se uma análise sem rigor, apenas observando-se os fatos. Como pode ser observado, todas as recessões incluídas na amostra de desenvolvimento foram bem antecipadas.

Amostra de Teste

Dada a regra de classificação baseada na distância euclidiana quadrática, 96% das observações da amostra de teste foram previstas corretamente, como pode ser visto na Tabela 22 abaixo:

Tabela 40 - Acertos e erros das previsões

Amostra de Teste	Ocorrência	Previsão		Total
		0	1	
0	38	3	41	
1	1	124	125	
0	93%	7%	100%	
1	1%	99%	100%	

Fonte: Elaboração própria.

A função discriminante conseguiu acertar 93% das recessões e 99% das não-recessões e daí resulta o acerto médio de 96%. Ela deixou de classificar três casos que eram recessões dos quarenta e um possíveis, ou seja, não apontou 7% das recessões (erro Tipo I). Dos cento e vinte e cinco casos de não-recessão, a função errou em um caso, ou seja, apontou 1% das não-

recessões como recessões, gerando um falso alarme da discriminante (erro Tipo II). Essa previsão apresentou um QPS no valor de 0,04.

A Figura abaixo mostra os alarmes de recessões e as respectivas recessões para os dados da amostra de desenvolvimento. As barras pretas representam os alarmes enquanto as barras azuis representam as recessões.

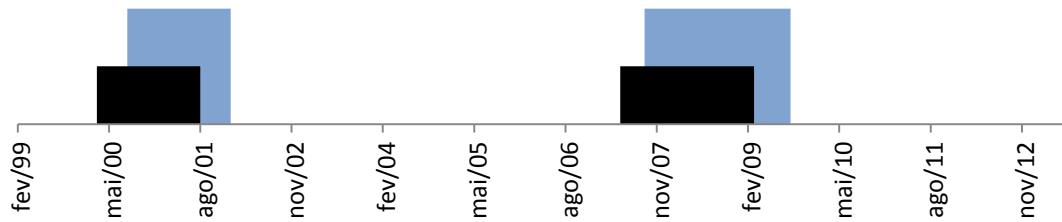


Figura 55 - Alarmes (barras pretas) e recessões (barras azuis) para a amostra de desenvolvimento.

Como pode ser observado, as duas recessões conseguiram ser antecipadas pelo algoritmo. Não houve falsos alarmes preditivos, justificando o poder de previsão do modelo. É importante ressaltar que no período da amostra de teste ocorreram oito períodos de desaceleração da produção, mas somente dois deles se reverteram em recessão. O fato do algoritmo não ter "confundido" estes momentos de desaceleração como os momentos recessivos se dá pelo papel dos indicadores antecedentes do diferencial de juros e dos pedidos de seguro-desemprego. Ou seja, caso eles não tivessem se comportado de maneira anormal antes destas duas recessões, apenas os indicadores de desaceleração não conseguiram captar as recessões.

A Tabela abaixo lista em quantos meses o algoritmo conseguiu antecipar cada recessão:

Tabela 41 - Duração dos alarmes antes do início de cada recessão

Recessão	Período do alarme	Duração (meses)
7	Mar/2000 - Jul/2000	5
8	Mai/2007 - Ago/2007	4

Fonte: Elaboração própria.

Como pode ser observado, as duas recessões conseguiram ser bem antecipadas, em cinco e quatro meses. O alarme não durou os seis meses ideais, pois os três erros do Tipo I desta análise se concentraram nos momentos iniciais dos alarmes. A Tabela 24 abaixo mostra em quanto tempo estes mesmos alarmes conseguiram antecipar as recessões da atividade agregada datadas pelo NBER.

Tabela 42 - Duração dos alarmes antes dos inícios das recessões da atividade agregada datadas pelo NBER

Recessão	Mês inicial do alarme	Mês inicial da recessão do NBER	Antecedência em relação às recessões do NBER
7	Mar/2000	Abr/2001	13
8	Mai/2007	Jan/2008	8

Fonte: Elaboração própria.

Como também pode ser observado, as duas últimas recessões americanas datadas pelo NBER foram bem antecipadas pelo algoritmo aqui construído.