



**FEBRABAN**

Federação Brasileira de Bancos

**FEBRABAN**

Diretoria de Regulação Prudencial,  
Riscos e Economia

**Fundação Getúlio Vargas**

Escola de Economia de São Paulo

**Indicadores macroeconômicos explicam  
e preveem a inadimplência? Um estudo  
dos dados brasileiros**

Dezembro/2015

**Emerson Fernandes Marçal  
Pedro Luiz Valls Pereira**

Este estudo foi realizado no âmbito do Convênio Febraban - Fundação Getúlio Vargas, que tem por objetivo estimular a produção de trabalhos e estudos na área de economia.

O conteúdo foi desenvolvido pelos autores de forma independente. As opiniões, hipóteses e conclusões / recomendações contidas neste material são de responsabilidade exclusiva dos mesmos, não refletindo, necessariamente, a visão da FEBRABAN.

# Indicadores macroeconômicos explicam e preveem a inadimplência? Um estudo dos dados brasileiros

**Emerson Fernandes Marçal**  
Escola de Economia de São Paulo  
emerson.marc@fgv.br

**Pedro Luiz Valls Pereira**  
Escola de Economia de São Paulo  
pedro.valls@fgv.br

## Resumo

O relatório avalia a contribuição incremental de indicadores macroeconômicos na previsão de inadimplência agregada. Os resultados indicam a importância do ciclo econômico e do ciclo monetário. A análise utilizou um modelo Vetor Autoregressivo (VAR) como processo gerador de dados. A série agregada de inadimplência utilizada está disponível desde janeiro de 2000, sendo feita uma análise desagregada nos níveis dos bancos privados (nacionais versus estrangeiros) e dos bancos públicos. Os indicadores macroeconômicos coletados foram a taxa básica Selic praticada no mercado, a produção industrial total, e o Índice de Preços ao Consumidor Amplo (IPCA).

**Palavras-chave:** Inadimplência; indicadores macroeconômicos; Causalidade de Granger; Estacionariedade; Modelos VAR.

## 1. INTRODUÇÃO

A análise dos níveis de inadimplência de uma economia é importante pois um aumento inesperado e forte da inadimplência pode gerar repercussões para o sistema financeiro como um todo, afetar o mecanismo de transmissão da política monetária e gerar repercussões negativas para a economia. Formuladores de política econômica bem como os gestores das carteiras dos bancos estão interessados em entender a relação que existem entre o ambiente macroeconômico e a evolução dos índices de inadimplência. O desenvolvimento de modelos de previsão da inadimplência é algo extremamente útil tais agentes na elaboração de cenários prováveis num horizonte mais longo.

O objetivo deste trabalho consiste em avaliar a influência de indicadores macroeconômicos nos indicadores de inadimplência no Brasil. Dois tipos de análise serão feitas. A primeira avalia se e como indicadores macroeconômicos - taxas de juros, ciclo econômico e inflação - contribuem na explicação das taxas de inadimplência da economia brasileira. O segundo tipo de análise realizada avalia se há e qual é o tamanho do ganho de poder preditivo das séries de inadimplência a partir de modelos que utilizam informações macroeconômicas.

O trabalho está dividido em sete seções incluindo esta introdução. A seção dois apresenta uma breve revisão da literatura. Na seção três a metodologia da análise feita no trabalho é discutida em detalhes. Na seção quatro descreve-se a base de dados. Na seção cinco os resultados são apresentados. A seção seis as limitações e possíveis extensões do trabalho são propostas. Por fim na seção sete as considerações finais são realizadas.

## 2. Motivação e Revisão da Literatura

Há uma literatura de teoria econômica extensa que ligam o ambiente financeiro ao ambiente macroeconômico sugerindo uma relação forte e de retroalimentação entre os mesmos. Um trabalho clássico nesta linha é dado por Bernanke et al. (1999). Neste trabalho os autores

desenvolvem um modelo em que há uma forte interação ciclo econômico e de crédito.

No plano de trabalhos empíricos há trabalhos que confirmam a relação entre ciclo econômico e inadimplência. Beck et al. (2013) utilizam dados de painel para um grupo de países mostram que outras variáveis que não apenas o ciclo econômico tais como taxa de câmbio e preços de ações podem influenciar o nível de inadimplência nos países. Já Claessens et al. (2012) analisam a relação entre ciclo econômico e ciclo financeiro para uma amostra grande de países. Os resultados dos autores sugerem que recessões estão associadas a crises de inadimplência mais fortes sendo o efeito do ciclo econômico assimétrico.

No caso brasileiro há trabalho que exploram microdados. Este trabalhos são de difícil replicação na medida que exige acesso a base de dados não públicas. Correa et al. (2014) utilizam microdados disponíveis a partir de informações do Banco Central do Brasil para avaliar a relação entre crédito e ciclo econômico. Os autores concluem que uma recessão aumenta a probabilidade de aumento da inadimplência e perda nas carteiras de empréstimos. O trabalho foca na análise de crédito ao consumidor e para aquisição de veículos cobrindo o período de 2008 a 2013. Neste trabalho os autores puderam explorar tanto a dimensão temporal quanto a dimensão de corte transversal na medida em que tiveram a acesso a dados de indivíduos. Tabak et al. (2008) estudam a relação existente entre concentração bancária e inadimplência controlada por fatores macroeconômicos. Os resultados sugerem que uma maior concentração bancária pode levar a menores níveis de inadimplência.

O trabalho de Linardi & Ferreira (2008) explora uma base de dados pública assim como o presente trabalho. Os autores estimam um modelo autoregressivo para avaliar a relação existente entre inadimplência e fatores macroeconômicos. Os resultados dos autores sugerem a existência de relação entre as variáveis. Os autores também propõe uma metodologia para avaliar cenários para a inadimplência, mas o poder preditivo dos modelos não é avaliado.

### **3. Metodologia**

Nesta parte do trabalho discute-se de forma mais detalhada a metodologia a ser utilizada. Dois tipos de análise são feitas. O primeiro tipo de análise foca o ajuste dentro da amostra. O objetivo neste caso é avaliar em que medida os indicadores macroeconômicos ajudam a explicar o comportamento dos índices de inadimplência e como se dá tais relações ao longo do tempo. A técnica utilizada para avaliar tal relação é dado por vetores autoregressivos. O segundo tipo de análise avalia-se se os indicadores macroeconômicos são capazes de antecipar e em que horizontes as variações da inadimplência no futuro. Trata-se de um exercício de previsão fora da amostra. A avaliação da qualidade preditiva é feita utilizando técnicas recentes de seleção de modelos de previsão. Utiliza-se o trabalho de Hansen et al. (2011).

#### **3.1. Análise dentro da amostra**

Na análise dentro da amostra, utiliza-se um modelo linear multivariado da classe Vetor Autoregressivo (VAR). Avaliamos em que medida este modelo é uma boa representação do processo gerador de dados. Para isto utiliza-se avanços recentes na área de seleção de modelos dado pelo algoritmo Autometrics desenvolvidos por Hendry & Doornik (2014). O algoritmo utiliza técnicas de seleção de modelos que obedecem a abordagem de Econometria da LSE desenvolvida por David Hendry ([Hendry (1995)]).

##### **3.1.1. Vetor Autoregressivo - VAR**

Optou-se por modelar os dados a partir de um Vetor Autoregressivo (VAR). Este modelo pode ser visto como sendo uma forma reduzida de algum modelo macroeconômico. Num VAR estima-se a relação existente entre as variáveis e o passado das mesmas. Não há relação contemporânea entre as variáveis:

$$Y_t = \mu + \Gamma_1 Y_{t-1} + \dots + \Gamma_p Y_{t-p} + \varepsilon_t \quad (1)$$

na qual  $\varepsilon_t$  é um vetor de erros normais e não serialmente correlacionados,  $\Omega$  a matriz de variância e covariância dos erros que é igual para todos os períodos de tempo e  $\theta = [\Gamma_1, \dots, \Gamma_{k-1}, \mu, \Omega]$  contém os parâmetros do modelo. O vetor  $Y_t$  conterá os indicadores de inadimplência e as variáveis macroeconômicas.

### 3.1.2. Testes de relevância das variáveis

Gourieroux & Jasiak (2001) discutem que restrições podem ser impostas ao VAR para que os processo podem ser vistos como não correlacionados para qualquer período temporal, sendo um forte indicador da inexistência de qualquer relação de causalidade entre as variáveis.

Caso todas as matrizes  $\Gamma_i$  com  $i$  variando de 1 a  $k$  (ordem do VAR) sejam diagonais e se a matriz de variância e covariância  $\Omega$  seja também diagonal, então a correlação temporal dos processo é zero para qualquer instante de tempo.

### 3.1.3. Causalidade de Granger

A restrição da diagonalidade da matriz  $\Gamma_i$  nada mais é que a inexistência de causalidade de Granger em qualquer direção. O conceito de causalidade de Granger [Granger (1969)] pode ser definido da seguinte forma:

Definição 1: Causalidade de Granger (Hendry (1995), pág, 176)

Sejam  $y$  e  $x$  duas variáveis cuja distribuição conjunta é dada por  $D_Z(y_t, x_t | Z_{t-1})$  em que  $Z = [y \ x]$  e  $Z_{t-1}$  representada toda a história passada até o instante  $t-1$ ,

Considerando a fatoração da distribuição conjunta dada por:

$$D_Z(y_t, x_t | Z_{t-1}) = D_Z(y_t | x_t, Z_{t-1}) D_x(x_t | Z_{t-1}),$$

Se  $D_x(x_t | Z_{t-1}) = D_x(x_t | X_{t-1})$  diz-se a variável  $y_{t-1}$  não causa no sentido de Granger  $x_t$ ,

Vale ressaltar que o conceito de causalidade de Granger está associado a precedência temporal e não como uma relação de causa e efeito. Principalmente em modelos em que os agentes olham para frente para tomar suas decisões (forward looking), a causalidade de Granger pode ocorrer na direção contrária a relação causal que implica causa e efeito.

### 3.2. Análise fora da amostra

Um bom ajuste dentro da amostra não garante necessariamente que o mesmo modelo conseguirá ter desempenho bom na antecipação de valores futuros. O campo da previsão em Economia é um dos mais desafiadores. As variáveis econômicas como parte de um sistema complexo, são de difícil antecipação. Alterações de toda ordem como crises, mudanças estruturais, erros de medida, instabilidade do processo gerador de dados podem fazer com que um modelo bem fundamentado não tenha desempenho preditivo adequado. Há também o fato de muitas variáveis alvo de previsão em Economia serem financeiras, o que torna sua previsibilidade extremamente lucrativa. Num ambiente de mercados eficientes informacionalmente as oportunidades de lucro econômico devem ser baixas ou inexistentes com base em informação pública. Logo a previsibilidade deve ser baixa ou inexistente neste contexto.

#### 3.2.1. Avaliação do poder preditivo do modelo

Na confecção de uma previsão econômica o analista deve construir algum modelo e a partir do mesmo construir sua previsão. A precisão da previsão pode ser avaliada sobre diversos critérios. Um primeiro critério diz respeito ao viés. Um conjunto de previsões do analista que apresente erros sistemáticos numa ou outra direção não deve ser utilizado em princípio. Outro critério diz respeito aos desvios com relação ao valor observado. Se dois



conjuntos de previsão acertam na média o valor da variável escolhe-se aquele com menor dispersão.

Desta forma métricas para avaliação dos erros de previsão foram criadas na literatura. Dois exemplos bem populares são o erro quadrático médio (EQM) e erro absoluto médio (EAM) definidos a seguir:

$$EQM = \sum_{\tau=1}^N (\tilde{y}_{t+\tau+h}^{t+\tau-1} - y_{t+\tau+h})^2 \quad (2)$$

$$EAM = \sum_{\tau=1}^N |\tilde{y}_{t+\tau+h}^{t+\tau-1} - y_{t+\tau+h}| \quad (3)$$

na qual  $\tilde{y}_{t+\tau+h}^{t+\tau-1}$  representa a previsão realizada para a variável  $y$  para  $t + \tau + h$  condicionada na informação disponível até  $t + \tau - 1$ ,  $h$  representa o número de passos a frente ao qual se relaciona a previsão.

Uma forma de selecionar entre diversos modelos de previsão consiste em calcular tais medidas e ranquear o modelo a partir das mesmas, optando pelo modelo de melhor desempenho.

### 3.2.2. O exercício de pseudo tempo real

Um analista que deseja construir um modelo de previsão pode simular um exercício de tempo real para avaliar se um conjunto de modelos tem bom poder preditivo. O analista deve escolher uma janela de avaliação e uma janela de estimação. Na janela de estimação os modelos são estimados, as variáveis escolhidas e as previsões geradas para os dados da janela de avaliação. A escolha dos modelos não pode conter nenhuma informação que não estava disponível ao analista no momento em que ele teria gerado a previsão para que o exercício seja válido. Modelos construídos a partir dos dados da janela de estimação podem ser então testados em termos de poder preditivo na janela de avaliação. Caso haja uma boa performance é possível que num exercício real venha a ter boa performance também. A seleção de modelos preditivos é um processo computacionalmente intensivo, na medida que a projeção tem de ser feita a todo instante de forma a tentar simular a projeção que teria sido feita em tempo real.

### 3.2.3. Escolha do benchmark

A escolha do benchmark é algo importante. Um possível critério diz respeito a construção de modelos ingênuos que seriam facilmente vencidos por modelos mais sofisticados. Um exemplo é o passeio aleatório no qual supõe-se que a melhor previsão para  $h$  períodos a frente é exatamente o valor da última observação. O modelo de passeio aleatório é amplamente utilizado como benchmark em Finanças e tem se mostrado muitas vezes um adversário “duro” a ser batido. No caso de taxa de câmbio para horizontes curtos é muito difícil vencer o passeio aleatório como mostra o trabalho seminal de Meese & Rogoff (1983).

Outro benchmark utilizado para dados com persistência é um modelo autoregressivo de ordem 1. Castle et al. (2014) mostram que é muito difícil batê-lo em exercícios de previsão de inflação.

### 3.3. O conjunto de confiança para (seleção de) modelos (model confidence set - MCS)

O conjunto de confiança para (seleção de ) de modelos (MCS) é uma técnica de seleção de modelos de previsão desenvolvidos por Hansen et al. (2011). Ele consiste num algoritmo que ranquea um conjunto de previsões feitas por um conjunto de modelos. O conjunto  $M^*$  contém o(s) melhor(es) modelo(s) escolhidos a partir de um conjunto inicial. O conjunto  $M^0$  é o conjunto que contém o(s) melhor(es) modelo(s) definido a partir de um critério de qualidade preditiva.

Definição 2: O conjunto que contém os elementos superiores é definido por:

$$M^* = \{i \in M^0 : E(d_{\tau}^{i,j}) \leq 0 \text{ para todo } j \in M^0\}$$

Defina-se  $M^{\dagger}$  para ser o complemento de  $M^*$ , ou seja,  $M^{\dagger} = \{i \in M^0 : E(d_{\tau}^{i,j}) > 0 \text{ para todo } j \in M^0\}$  na qual  $g(e_{\tau}^i)$  é alguma função de perda e

$$d_{\tau}^{i,j} = g(e_{\tau}^i) - g(e_{\tau}^j) \quad (4)$$

$$e_{\tau}^i = \tilde{y}_{t+\tau+h}^{t+\tau-1} - y_{t+\tau+h} \quad (5)$$

O MCS seleciona o(s) modelo(s) utilizando um teste de equivalência,  $\delta_M$ , e uma regra de eliminação,  $\rho_M$ . A regra de equivalência é aplicada ao conjunto  $M = M^0$ . Se a regra de equivalência é rejeita ao nível de confiança selecionado, então há, com alta probabilidade, um grupo de modelos inferiores em termos de poder preditivo que deve ser eliminado do conjunto de bons modelos. Neste caso uma regra de eliminação,  $\rho_M$  é utilizada para retirar o(s) modelo(s) com baixo poder preditivo do conjunto de bons modelos. Feito isto, utiliza-se a regra de equivalência novamente. O procedimento é repetido até a hipótese de equivalência preditiva no conjunto analisado,  $\delta_M$ , não é rejeitada. O conjunto de modelos do último passo ( $\widetilde{M}_F$ ) é selecionado e deve conter o(s) melhor(es) modelo(s) a um determinado nível de significância.

A hipótese nula do teste de equivalência é dada por:

$$H_M^0: E(d_{\tau}^{i,j}) = 0 \text{ para todo } i, j \in M \quad (6)$$

na qual  $M \subset M^0$ .

A hipótese alternativa é dada por:

$$H_M^1: E(d_{\tau}^{i,j}) \neq 0 \text{ para todo } i, j \in M \quad (7)$$

Um ponto importante a ressaltar é que pode existir modelos melhores fora do conjunto inicial de modelos “candidatos”  $M^0$ , O objetivo é ranquear um determinado conjunto de modelos para obter  $M^*$ .

A hipótese nula pode ser testada a partir da seguinte estatística:<sup>1</sup>

$$T_D = \sum_{i \in M} t_i^2 \quad (8)$$

na qual  $t_i = \frac{\bar{d}_i}{\sqrt{VAR(\bar{d}_i)}}$  e  $\bar{d}_i = \frac{1}{M} \sum_{j \in M} \bar{d}_{ij}$

<sup>1</sup> Existem outras possíveis escolhas.

A estatística de teste dada por (8) têm distribuição estatística não padrão mas que pode ser simulada usando técnicas de reamostragem (bootstrap). A regra de eliminação é dada por:

$$\rho_M = \operatorname{argmax}_i(t_i) \quad (9)$$

### 3.3.1. O algoritmo

O algoritmo do MCS tem os seguintes passos:

- (i) Inicializa-se o procedimento definindo o conjunto inicial de modelos a serem analisados  $M = M^0$ ;
- (ii) Testa-se  $H_M^0$  usando  $\delta_M$  ao nível de significância  $\alpha$ ;
- (iii) Se  $H_M^0$  não for rejeitada então o procedimento termina o conjunto final é  $\widehat{M}_{1-\alpha}^* = M$ , Caso contrário utiliza-se a regra de eliminação  $\rho_M$  para excluir um objeto de  $M$  e volta-se ao passo (i),

Os autores sugerem que o MCS têm as seguintes propriedades estatísticas:

- (i)  $\lim_{n \rightarrow \infty} (M^* \subset \widehat{M}_{1-\alpha}^*) > 1 - \alpha$  e;
- (ii)  $\lim_{n \rightarrow \infty} P(i^\dagger \in \widehat{M}_{1-\alpha}^*) = 0$  para todo  $i^\dagger \in M^\dagger$

### 3.3.2. Ranqueando os modelos: p-valores do MCS

A regra de eliminação,  $\rho_M$ , define uma sequência de conjuntos aleatórios,  $M_0 = M_1 \supset M_2 \supset \dots \supset M_{m_0}$ , nos quais  $M_i = \{\rho_i, \dots, \rho_{m_0}\}$  e  $m_0$  é o o número de elementos em  $M_0$ ,  $\rho_{m_0}$  é o primeiro elemento a ser eliminado,  $\rho_{m_1}$  é o segundo a ser eliminado, e assim por diante. No final apenas um modelo sobrevive. Define-se o p-valor do modelo final em 1. Armazena-se os p-valores dos modelos eliminados se eles forem maiores do que o p-valor do modelo eliminado previamente. Se o p-valor da eliminação corrente é armazenado.

Os p-valores do MCS são importantes pois tornam mais simples para o analista avaliar um determinado objeto  $\widehat{M}_{1-\alpha}^*$ ,

### 3.4. O pseudo exercício de tempo real

Os dados coletados permitem criar várias variantes de modelos para prever inadimplência. A amostra é dividida em dois pedaços. A primeira parte é utilizada para estimar os modelos e construir as previsões, enquanto que a segunda parte é utilizada para avaliar o desempenho preditivo sobre vários horizontes. Neste exercício procura-se simular uma previsão em tempo real. Utiliza-se o conjunto de informação mais próximo possível ao disponível aos agentes no momento em que construíam a previsão. Em outras palavras, os modelos são estimados e revistos a cada ponto no tempo de forma a incorporar o ganho de informação que decorre do passar do tempo. Cada modelo gera previsões até 1 ano à frente. Há portanto doze grupos de previsões pois os dados são mensais.

Um problema com a simulação de um exercício em tempo real tem a ver com as fontes de dados. Muitas das séries são descontinuadas ou tem metodologia revista ou sofrem revisão por nova informação não processada. Desta forma, o conjunto de informação não é exatamente o valor disponível aos agentes, mas muito próximo. O exercício utiliza um conjunto de informação ligeiramente superior ao disponível para os agentes no instante que realizaram as projeções em tempo real. Isto pode gerar um pequeno viés na direção de construirmos melhores modelos preditivos. Devido a pesquisa da base de dados originais ser extremamente custosa em termos de tempo e recursos, um eventual refinamento do exercício é deixado como extensão futura.

Recentemente Cusinato et al. (2013) analisam o efeito que revisões do dado divulgado pelo IBGE para o Produto Interno Bruto na série trimestral tem na estimação do produto potencial brasileiro. Os autores sugerem que entre a estimativa feita a partir da divulgação inicial e do dado final revisado a estimativa pode variar em cerca de 0.7% em termos anualizados, o que é algo

bem substancial. Não é de conhecimento dos autores outros trabalhos que façam o mesmo para a estimação do produto potencial numa frequência mais alta a partir dos dados de produção industrial utilizados neste trabalho.

### **3.5. Construindo Hiato do Produto**

O hiato do produto potencial foi estimado utilizando o filtro Hodrick and Prescott (1997). Este filtro permite ao analista separar ciclo de tendência, A tendência é interpretada como produto potencial enquanto o termo cíclico é interpretado como um desvio do pleno emprego. Existem outras possibilidades na literatura como Baxter & King (1999) ou um modelo de espaço-estado de forma similar a discutida em Harvey (1990) e Durbin & Koopman (2001) e uma aplicação desta para dados brasileiros em Valls Pereira (1986).

## **4. Descrição da Base de Dados**

Coletaram-se os dados de inadimplência no Banco Central do Brasil. Recentemente a instituição promoveu uma revisão de metodologia das séries de inadimplência descontinuando grande parte das séries existentes e não realizando a reconstrução das séries novas para períodos anteriores. Desta forma, grande parte dos indicadores de inadimplência em níveis de desagregação importantes tão como pessoa física e jurídica, modalidades de crédito, entre outras estão disponíveis apenas a partir de 2010, o que limita a análise. Seria interessante, caso seja possível tecnicamente, que o Banco Central continuasse a divulgar as séries de inadimplência na metodologia antiga também.

Dada a restrição de dados optamos por utilizar a série agregada de inadimplência que está disponível desde janeiro de 2000 e a única desagregação disponível desde 2000 é dada por setor privado (bancos nacionais e estrangeiros) e setor público.

Os indicadores macroeconômicos coletados foram a taxa básica Selic praticada no mercado, a produção industrial total, e o Índice de Preços ao Consumidor Amplo (IPCA).

Tabela 1: Descrição das Variáveis				
Séries	Fonte	Amostra	Frequência	Transformação
Inadimplência Privada Total	Banco Central do Brasil	2000(1) a 2015(9)	Mensal	Nenhuma
Inadimplência Privada Nacional	Banco Central do Brasil	2000(1) a 2015(9)	Mensal	Nenhuma
Inadimplência Privada Estrangeira	Banco Central do Brasil	2000(1) a 2015(9)	Mensal	Nenhuma
Inadimplência Pública	Banco Central do Brasil	2000(1) a 2015(9)	Mensal	Nenhuma
Produção Industrial	IBGE	1975(1) a 2015(7)	Mensal	Encadeamento
Taxa Selic	Banco Central do Brasil	1980(1) a 2015(7)	Mensal	Nenhuma
Índice de Preços ao Consumidor Amplo	IBGE	1980(1) a 2015(7)	Mensal	Nenhuma

A Figura 3 mostra a evolução do ciclo extraído para a amostra completa. O primeiro gráfico da Figura mostra a estimativa do ciclo não ajustada sazonalmente o que dificulta a visualização do ciclo. Já no segundo gráfico apresenta-se a medida de ciclo ajustada sazonalmente o que permite visualizar o ciclo de forma mais nítida.

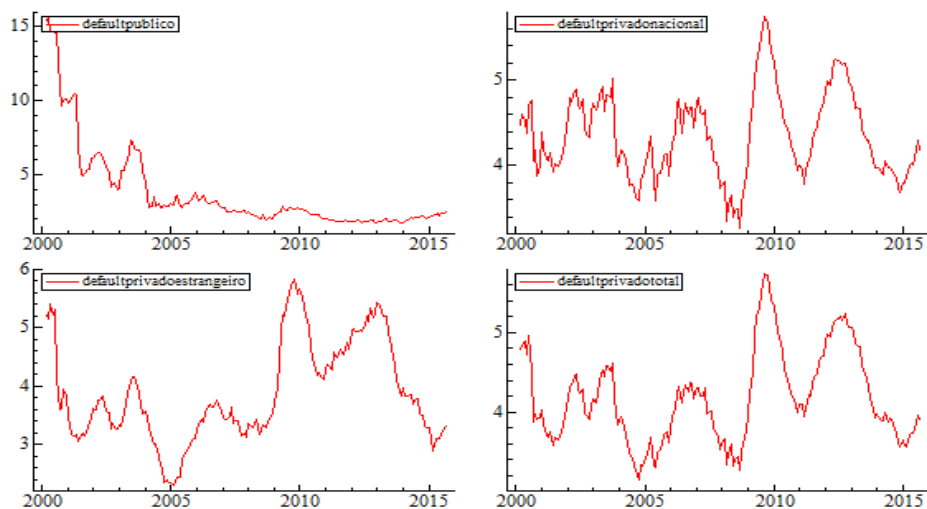


Figura 1: Inadimplência

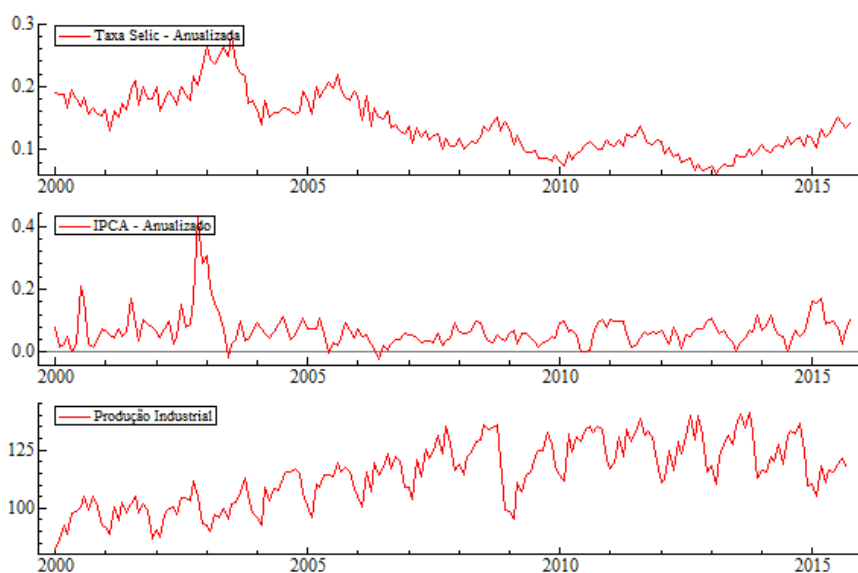


Figura 2: Indicadores Macroeconômicos



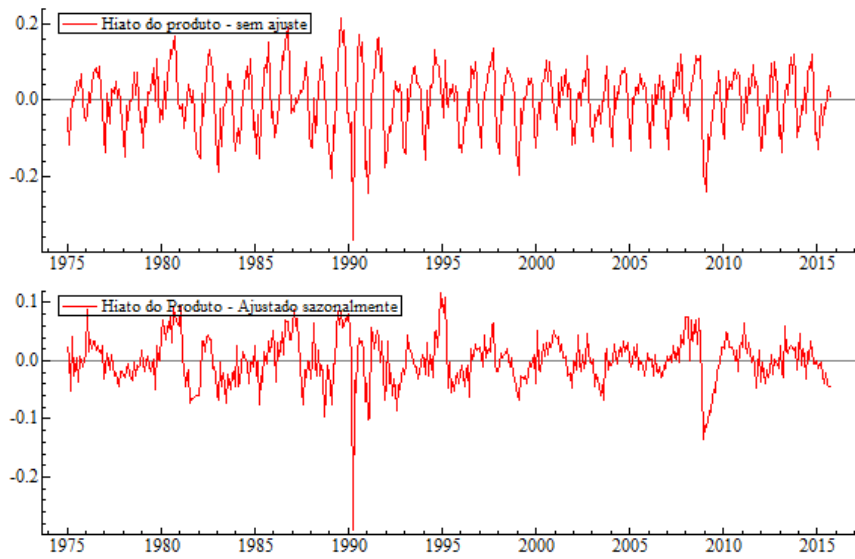


Figura 3: Hiato do Produto estimado com amostra completa

## **5. Resultados**

Nesta seção apresentamos os resultados dos modelos estimados para entender os movimentos da inadimplência no Brasil. Dois tipos de exercícios são apresentados. No primeiro, a análise é feita dentro da amostra e procuramos demonstrar a ligação existente entre inadimplência e variáveis macroeconômicas, em particular ao ciclo monetário e ao ciclo real econômico. Na segunda parte reportamos os resultados de um exercício de simulação em tempo real para avaliar a capacidade de preditiva de indicadores macroeconômicos sobre inadimplência.

### **5.1. Testes dentro da amostra**

Primeiro são feitos testes de raízes unitárias para determinar o ordem de integração das variáveis. Isto é feito para que todas as variáveis que entram na estimação do VAR tenham a mesma ordem de integração. A seguir são estimados VAR para os diversos modelos propostos para inadimplência e variáveis macroeconômicas onde é utilizado a seleção de modelos Autometrics. A seguir testes de Causalidade são apresentados.

#### **5.1.1. Testes de Estacionaridade**

Inicialmente avaliamos a ordem de integração das variáveis pois isto é importante para a análise que se segue na medida que traz implicações importantes para a análise estatística. A presença ou não de processos não estacionários causa grandes alterações na análise de séries de tempo. Uma discussão detalhada é feita por exemplo em Hamilton (1994).

A Tabela 2 mostra os resultados do teste desenvolvidos por Dickey e Fuller na versão aumentada para diversas séries em nível. Os resultados sugerem que a hipótese nula de raiz unitária pode ser rejeitada para a maioria das séries com exceção da taxa Selic, produção Industrial e inadimplência privada estrangeira.

Variáveis Incluídas	Sazonalidade	Defasagens	Estatística t	Amostra
Inadimplência Total	Sim	3	-3,493	2000(1) a 2015(9)
Inadimplência Privada Nacional	Sim	3	-3,408	2000(1) a 2015(9)
Inadimplência Privada Estrangeira	Sim	2	-2,376	2000(1) a 2015(9)
Inadimplência Pública	Sim	2	-5,037	2000(1) a 2015(9)
Selic	Não	3	-1,882	2000(1) a 2015(9)
Produção Industrial	Sim	12	-1,693	1975(1) a 2015(9)
Selic Real (deflacionado IPCA)	Sim	0	-5,675	2000(1) a 2015(9)

Nota: Valores Críticos: 5%  $\geq -2,882$  e 1%  $\geq -3,47$ .

### 5.1.2. Análise dos Modelos Estimados

Estimaram-se modelos vetoriais autoregressivos (VAR) para modelar inadimplência total e seus determinantes macroeconômicos. Os modelos estimados estão descritos abaixo. O ponto de partida foi um VAR com doze defasagens, dummies sazonais e correção para outliers. Os modelos finais são reportados a seguir e a estratégia de simplificação seguiu um abordagem geral para o específico nos moldes proposta pela abordagem da LSE popularizada pelos trabalhos de David Hendry. Utilizou o algoritmo de simplificação Autometrics desenvolvidos por Hendry & Doornik (2014). Utiliza-se a amostra completa que compreende o período de janeiro de 2000 a setembro de 2015.

Com base nesta estratégia conseguiu-se obter modelos que apresentam bons resultados em termos de especificação e podem ser vistos com uma aproximação boa do processo gerador dos dados.

Modelo	Variáveis Incluídas	Defasagens	Sazonalidade	Correção para Outliers
	Inadimplência, Variação da Selic, hiato do produto	12	Sim	Sim
	Inadimplência, hiato do produto, juros reais	12	Sim	Sim
	Inadimplência, juros reais, hiato do produto, variação da Selic	12	Sim	Sim

Em todas as especificações avaliadas há uma relação forte entre ciclo econômico e nível de inadimplência. O hiato do produto é significativo na equação de inadimplência em todas as especificações. Já a variável juros, seja na forma da Selic ou da Selic deflacionadas (juros real) parece não causar influência direta na inadimplência. Contudo isto não implica que é irrelevante para explicar o que ocorre com a inadimplência. Existe uma dinâmica importante entre juros nominal e-ou real com o ciclo econômico. De tal sorte que fatores que influenciam os juros acabam afetando o ciclo econômico que por sua vez acaba gerando efeito na inadimplência. A evolução do nível de inadimplência também influencia as variáveis macroeconômicas.

Vale também ressaltar que foram necessários a adição de dummies pontuais para controlar outliers aditivos e de inovação<sup>2</sup>. O algoritmo de seleção do Autometrics sugere que a estrutura estimada apresenta pontos de instabilidade em: 2002 (11), 2002(12) , 2003(2), 2003(10), 2003(11), 2004(2), 2005(4), 2005(5), 2005(6), 2008(12) e 2009(4). Os pontos de 2002 estão associados a instabilidade ocasionada pelo período eleitoral de 2002. O ponto em 2008 está provavelmente associado aos efeitos da crise internacional sobre o Brasil. O ponto de 2013(11) talvez esteja associado ao lançamento do Lei do crédito consignado. Os demais pontos não possuem clara interpretação mas apresentam significância estatística.

<sup>2</sup> Ver Nielsen (2004) para discussão e definição.

Todos os modelos estão bem especificados uma vez que passam por todos os testes de especificação e falta de especificação.

Tabela 4: Estimativa do Modelo 1: default privado total, variação da Selic e hiato do produto

$$\begin{aligned}
 \text{Defaultprivtotal}_t = & \quad 1.069 \text{Defaultprivtotal}_{t-1} \quad - 0.1136 \text{Defaultprivtotal}_{t-7} \\
 & \quad (0.02) \quad - (0.0176) \\
 & + \frac{9.36}{(7.03)} \text{Dlnselic}_{t-1} \quad - \frac{1.276}{(0.17)} \text{Hiato\_do\_Produto}_{t-1} \quad - \frac{0.1618}{(0.297)} \text{Hiato\_do\_Produto}_{t-3} \\
 & + \frac{0.8544}{(0.23)} \text{Hiato\_do\_Produto}_{t-4} \quad + \frac{0.4008}{(0.159)} \text{Hiato\_do\_Produto}_{t-10} \quad - \frac{0.07617}{(0.0314)} \text{CSeasonal}_{t-1} \\
 & - \frac{0.2017}{(0.051)} \text{CSeasonal}_{t-2} \quad + \frac{0.1521}{(0.0395)} \text{CSeasonal}_{t-3} \quad + \frac{0.1046}{(0.0411)} \text{CSeasonal}_{t-4} \\
 & + \frac{0.05777}{(0.0323)} \text{CSeasonal}_{t-6} \quad + \frac{0.08823}{(0.0288)} \text{CSeasonal}_{t-7} \quad + \frac{0.01186}{(0.034)} \text{CSeasonal}_{t-8} \\
 & + \frac{0.1696}{(0.0289)} \text{CSeasonal}_{t-9} \quad + \frac{0.1346}{(0.031)} \text{CSeasonal}_{t-10} \quad + \frac{0.1562}{(0.0594)} \text{DI: 2003(2)}_t \\
 & - \frac{0.3799}{(0.0682)} \text{DI: 2005(5)}_t \quad - \frac{0.235}{(0.0685)} \text{DI: 2005(6)}_t \quad + \frac{0.3445}{(0.0827)} \text{I: 2002(12)}_t \\
 & - \frac{0.3845}{(0.0865)} \text{I: 2003(11)}_t \quad + \frac{0.1471}{(0.084)} \text{I: 2008(12)}_t \quad + \frac{0.1835}{(0.0588)}
 \end{aligned}$$

$$\begin{aligned}
 \text{Dlnselic}_t = & \quad - \frac{0.0002078}{(0.000173)} \text{Defaultprivtotal}_{t-1} \quad + \frac{4.86e-5}{(0.000152)} \text{Defaultprivtotal}_{t-7} \\
 & - \frac{0.454}{(0.0606)} \text{Dlnselic}_{t-1} \quad + \frac{0.3455}{(0.0794)} \text{Dlnselic}_{t-3} \quad - \frac{0.0006022}{(0.0146)} \text{Hiato\_do\_Produto}_{t-1} \\
 & + \frac{0.005569}{(0.0256)} \text{Hiato\_do\_Produto}_{t-3} \quad - \frac{0.0004785}{(0.00198)} \text{Hiato\_do\_Produto}_{t-4} \quad - \frac{0.04591}{(0.0137)} \text{Hiato\_do\_Produto}_{t-10} \\
 & - \frac{0.001473}{(0.000271)} \text{CSeasonal}_{t-1} \quad + \frac{0.001238}{(0.000439)} \text{CSeasonal}_{t-2} \quad + \frac{0.0005494}{(0.00034)} \text{CSeasonal}_{t-3} \\
 & + \frac{0.01752}{(0.000271)} \text{CSeasonal}_{t-4} \quad + \frac{0.001594}{(0.000278)} \text{CSeasonal}_{t-6} \quad + \frac{0.0005082}{(0.000248)} \text{CSeasonal}_{t-7} \\
 & - \frac{0.0003908}{(0.000293)} \text{CSeasonal}_{t-8} \quad - \frac{0.0006096}{(0.000249)} \text{CSeasonal}_{t-9} \quad - \frac{0.001042}{(0.000267)} \text{CSeasonal}_{t-10} \\
 & + \frac{0.01951}{(0.000512)} \text{DI: 2003(2)}_t \quad + \frac{0.000397}{(0.000588)} \text{DI: 2005(5)}_t \quad + \frac{0.0009989}{(0.000591)} \text{DI: 2005(6)}_t \\
 & + \frac{0.001167}{(0.000713)} \text{I: 2002(12)}_t \quad - \frac{0.001061}{(0.000746)} \text{I: 2003(11)}_t \quad - \frac{0.0006703}{(0.000724)} \text{I: 2008(12)}_t \\
 & + \frac{0.000669}{(0.000506)}
 \end{aligned}$$









Tabela 6: Estimativas do Modelo 3: inadimplência privada total, variação da selic, hiato do produto e juros real

$Defaultprivtotal_t =$	$1.072$ $(0.0161)$	$Defaultprivtotal_{t-1}$	$-0.1201$ $(0.018)$	$Defaultprivtotal_{t-7}$	
$+ 8.201$ $(7.86)$	$Dlnselic_{t-1}$	$+ 1.43$ $(7.38)$	$Dlnselic_{t-3}$	$- 4.438$ $(7.51)$	$Dlnselic_{t-5}$
$+ 9.399$ $(8.63)$	$Dlnselic_{t-9}$	$- 8.328$ $(6.01)$	$Dlnselic_{t-11}$	$- 0.663$ $(0.152)$	$Hiato\_do\_Produto_{t-1}$
$+ 0.6191$ $(0.222)$	$Hiato\_do\_Produto_{t-8}$	$- 0.004911$ $(0.245)$	$Hiato\_do\_Produto_{t-9}$	$+ 0.4221$ $(0.157)$	$Hiato\_do\_Produto_{t-10}$
$- 0.5077$ $(1.93)$	$lnjurosreal_{t-1}$	$- 0.8028$ $(1.85)$	$lnjurosreal_{t-12}$	$- 0.149$ $(0.0321)$	$CSeasonal_{t-2}$
$+ 0.001232$ $(0.0335)$	$CSeasonal_{t-4}$	$- 0.2028$ $(0.0351)$	$CSeasonal_{t-5}$	$- 0.002004$ $(0.0335)$	$CSeasonal_{t-6}$
$+ 0.03732$ $(0.0443)$	$CSeasonal_{t-7}$	$+ 0.1574$ $(0.0329)$	$CSeasonal_{t-9}$	$+ 0.08252$ $(0.0403)$	$CSeasonal_{t-10}$
$- 0.3017$ $(0.091)$	$DI: 2002(11)_t$	$+ 0.02391$ $(0.0843)$	$DI: 2003(10)_t$	$+ 0.03472$ $(0.0588)$	$DI: 2004(2)_t$
$- 0.02286$ $(0.0836)$	$DI: 2005(4)_t$	$+ 0.326$ $(0.123)$	$I: 2002(11)_t$	$- 0.3854$ $(0.126)$	$I: 2003(11)_t$
$- 0.5115$ $(0.12)$	$I: 2005(5)_t$	$+ 0.1621$ $(0.0821)$	$I: 2008(12)_t$	$+ 0.2092$ $(0.0723)$	

$Dlnselic_t =$	$-0.0001633$ $(0.00013)$	$Defaultprivtotal_{t-1}$	$+ 9.827e - 5$ $(0.000145)$	$Defaultprivtotal_{t-7}$	
$- 0.4908$ $(0.0633)$	$Dlnselic_{t-1}$	$+ 0.3694$ $(0.0595)$	$Dlnselic_{t-3}$	$+ 0.08857$ $(0.0605)$	$Dlnselic_{t-5}$
$+ 0.2806$ $(0.0696)$	$Dlnselic_{t-9}$	$- 0.3311$ $(0.0485)$	$Dlnselic_{t-11}$	$+ 0.002636$ $(0.00122)$	$Hiato\_do\_Produto_{t-1}$
$+ 0.005414$ $(0.00179)$	$Hiato\_do\_Produto_{t-8}$	$- 0.00388$ $(0.00198)$	$Hiato\_do\_Produto_{t-9}$	$- 0.004081$ $(0.00127)$	$Hiato\_do\_Produto_{t-10}$
$- 0.05664$ $(0.0155)$	$lnjurosreal_{t-1}$	$- 0.001074$ $(0.0149)$	$lnjurosreal_{t-12}$	$+ 0.0005029$ $(0.000259)$	$CSeasonal_{t-2}$
$+ 0.001289$ $(0.00027)$	$CSeasonal_{t-4}$	$+ 0.0003597$ $(0.000196)$	$CSeasonal_{t-5}$	$+ 0.001749$ $(0.00027)$	$CSeasonal_{t-6}$
$+ 0.001123$ $(0.000358)$	$CSeasonal_{t-7}$	$- 0.000196$ $(0.000265)$	$CSeasonal_{t-9}$	$- 0.000612$ $(0.000325)$	$CSeasonal_{t-10}$
$- 0.0001302$ $(0.000733)$	$DI: 2002(11)_t$	$- 0.001993$ $(0.00068)$	$DI: 2003(10)_t$	$- 0.001692$ $(0.000474)$	$DI: 2004(2)_t$

$+ \frac{0.0009124}{(0.000674)} DI: 2005(4)_t$	$+ \frac{0.001428}{(0.000993)} I: 2002(11)_t$	$- \frac{0.002141}{(0.00102)} I: 2003(11)_t$
$+ \frac{0.0006891}{(0.000967)} I: 2005(5)_t$	$- \frac{2.61 e - 005}{(0.000663)} I: 2008(12)_t$	$+ \frac{0.0014}{(0.000583)}$

$Hiato\_do\_Produto_t =$	$- \frac{0.009001}{(0.00471)} Defaultprivtotal_{t-1}$	$+ \frac{0.01626}{(0.00525)} Defaultprivtotal_{t-7}$
$- \frac{9.71}{(2.29)} Dlnselic_{t-1}$	$+ \frac{9.043}{(2.15)} Dlnselic_{t-3}$	$- \frac{8.32}{(2.19)} Dlnselic_{t-5}$
$+ \frac{11.54}{(2.52)} Dlnselic_{t-9}$	$- \frac{1.468}{(1.76)} Dlnselic_{t-11}$	$+ \frac{0.6624}{(0.0443)} Hiato\_do\_Produto_{t-1}$
$+ \frac{0.1757}{(0.0647)} Hiato\_do\_Produto_{t-8}$	$- \frac{0.3892}{(0.0716)} Hiato\_do\_Produto_{t-9}$	$+ \frac{0.3194}{(0.0459)} Hiato\_do\_Produto_{t-10}$
$- \frac{1.614}{(0.563)} lnjurosreal_{t-1}$	$+ \frac{1.99}{(0.539)} lnjurosreal_{t-12}$	$+ \frac{0.08269}{(0.00937)} CSeasonal_{t-2}$
$+ \frac{0.1063}{(0.00978)} CSeasonal_{t-4}$	$+ \frac{0.006228}{(0.0103)} CSeasonal_{t-5}$	$+ \frac{0.0893}{(0.00979)} CSeasonal_{t-6}$
$+ \frac{0.1074}{(0.013)} CSeasonal_{t-7}$	$+ \frac{0.09792}{(0.00961)} CSeasonal_{t-9}$	$+ \frac{0.02443}{(0.0118)} CSeasonal_{t-10}$
$+ \frac{0.07214}{(0.0266)} DI: 2002(11)_t$	$- \frac{0.02053}{(0.0246)} DI: 2003(10)_t$	$- \frac{0.02398}{(0.0172)} DI: 2004(2)_t$
$+ \frac{0.08005}{(0.0244)} DI: 2005(4)_t$	$- \frac{0.04949}{(0.036)} I: 2002(11)_t$	$+ \frac{0.03894}{(0.0368)} I: 2003(11)_t$
$+ \frac{0.08558}{(0.035)} I: 2005(5)_t$	$- \frac{0.1381}{(0.024)} I: 2008(12)_t$	$- \frac{0.03319}{(0.0211)}$

$lnjurosreal_t =$	$- \frac{0.0001049}{(0.000458)} Defaultprivtotal_{t-1}$	$- \frac{0.0006258}{(0.00051)} Defaultprivtotal_{t-7}$
$+ \frac{0.4594}{(0.223)} Dlnselic_{t-1}$	$+ \frac{0.1404}{(0.209)} Dlnselic_{t-3}$	$+ \frac{0.6271}{(0.213)} Dlnselic_{t-5}$
$+ \frac{0.5024}{(0.245)} Dlnselic_{t-9}$	$- \frac{0.1429}{(0.171)} Dlnselic_{t-11}$	$- \frac{0.003477}{(0.0043)} Hiato\_do\_Produto_{t-1}$
$+ \frac{3.898e - 5}{(0.00629)} Hiato\_do\_Produto_{t-8}$	$- \frac{0.001828}{(0.00696)} Hiato\_do\_Produto_{t-9}$	$- \frac{0.002642}{(0.00446)} Hiato\_do\_Produto_{t-10}$
$+ \frac{0.7139}{(0.0547)} lnjurosreal_{t-1}$	$+ \frac{0.1108}{(0.0524)} lnjurosreal_{t-12}$	$+ \frac{0.002271}{(0.00951)} CSeasonal_{t-2}$
$+ \frac{0.002574}{(0.000951)} CSeasonal_{t-4}$	$+ \frac{0.002591}{(0.000997)} CSeasonal_{t-5}$	$+ \frac{0.002113}{(0.000951)} CSeasonal_{t-6}$
$+ \frac{0.001354}{(0.00126)} CSeasonal_{t-7}$	$- \frac{0.0006061}{(0.000934)} CSeasonal_{t-9}$	$+ \frac{0.0006846}{(0.00114)} CSeasonal_{t-10}$

$- \frac{0.004886}{(0.00258)} DI: 2002(11)_t$	$+ \frac{0.005454}{(0.00239)} DI: 2003(10)_t$	$- \frac{0.001284}{(0.00167)} DI: 2004(2)_t$
$+ \frac{0.004227}{(0.00237)} DI: 2005(4)_t$	$- \frac{0.01236}{(0.00349)} I: 2002(11)_t$	$+ \frac{0.009088}{(0.00358)} I: 2003(11)_t$
$- \frac{0.002535}{(0.00341)} I: 2005(5)_t$	$- \frac{0.002818}{(0.00233)} I: 2008(12)_t$	$+ \frac{0.003968}{(0.00205)}$

<i>Vector AR1 – 7 test</i>	$F(112,459) = 1.3439$	[0.0192]
<i>Vector Normality test</i>	$Chi^2(8) = 14.948$	[0.0602]
<i>Vector Hetero test</i>	$F(140,516) = 0.9569$	[0.6177]
<i>Vector RESET test</i>	$F(32,499) = 1.0563$	[0.3862]

### 5.1.3. Causalidade de Granger

Os resultados dos testes de causalidade de Granger entre as variáveis nos diversos modelos são apresentados nas Tabelas a seguir. Para todos os modelos, obtém-se evidência de causalidade de Granger de inadimplência para os indicadores macroeconômicos e também no sentido contrário nos testes conjuntos. Isto não implica necessariamente que todos os indicadores macroeconômicos causem ou sejam causados no sentido de Granger pela inadimplência, mas que pelo um indicador causa ou é causado no sentido de Granger pela inadimplência.

Em todos os modelos a variável hiato do produto Granger causa inadimplência e também é Granger causada ao nível de pelo 5% de significância, ou seja, existe uma relação robusta entre ciclo econômico e inadimplência por todos os modelos.

Já o efeito da taxa de juros seja em termos nominais avaliada através da taxa selic, seja em termos reais avaliada pela taxa selic deflacionada pelo IPCA não parece ser direto. A evidência de uma relação direta entre juros e inadimplência não é forte. A hipótese nula não é rejeitada aos níveis de significância tradicionais. Isto não quer dizer que juros seja uma variável irrelevante para explicar inadimplência. O efeito que juros tem sobre o hiato

do produto e vice-versa é pronunciado e estatisticamente robusto. Desta forma, o canal existe, mas a ação ocorre de forma indireta mediada pelo ciclo econômico.

Tabela 7: Causalidade de Granger - Modelo 1					
Variável(is)	Causa(m) no sentido de Granger	Variável(is)	Estatística	Distribuição	p-valor
2,3	→	1	103,02	$\chi^2(6)$	[0,0000]
2	→	1	2,2567	$\chi^2(2)$	[0,3236]
3	→	1	102,50	$\chi^2(4)$	[0,0000]
1	→	2,3	13,438	$\chi^2(4)$	[0,0093]
1	→	2	2,1304	$\chi^2(2)$	[0,3447]
1	→	3	9,7748	$\chi^2(2)$	[0,0075]
2	→	3	7,3027	$\chi^2(2)$	[0,0260]
3	→	2	29,217	$\chi^2(4)$	[0,0000]

Nota: Variável 1 corresponde a Default Total; 2 a Dlnselic e 3 a Hiato do Produto.

Tabela 8: Causalidade de Granger - Modelo 2					
Variável(is)	Causa(m) no sentido de Granger	Variável(is)	Estatística	Distribuição	p-valor
2,3	→	1	72.737	$\chi^2(4)$	[0.0000]
2	→	1	0.49749	$\chi^2(1)$	[0.4806]
3	→	1	68.931	$\chi^2(3)$	[0.0000]
1	→	2,3	31.462	$\chi^2(6)$	[0.0000]
1	→	2	11.644	$\chi^2(3)$	[0.0087]
1	→	3	19.956	$\chi^2(3)$	[0.0002]
2	→	3	2.9412	$\chi^2(1)$	[0.0863]
3	→	2	18.306	$\chi^2(3)$	[0.0004]

Notas: Variável 1 corresponde a Default Total; 2 a Dlnselic e 3 a Hiato do Produto.

Tabela 9: Causalidade de Granger - Modelo 3					
Variável(is)	Causa(m) no sentido de Granger	Variável(is)	Estatística	Distribuição	p-valor
2,3,4	→	1	14.124	$\chi^2(6)$	[0.0283]
2	→	1	0.19219	$\chi^2(1)$	[0.6611]
3	→	1	2.3773	$\chi^2(2)$	[0.3046]
4	→	1	12.901	$\chi^2(3)$	[0.0049]
1	→	2, 3, 4	16.134	$\chi^2(6)$	[0.0131]
1	→	2	7.7112	$\chi^2(2)$	[0.0212]
1	→	3	3.0309	$\chi^2(2)$	[0.2197]
1	→	4	5.6911	$\chi^2(2)$	[0.0581]
2	→	3	16.289	$\chi^2(1)$	[0.0001]
2	→	4	4.0504	$\chi^2(1)$	[0.0442]
3	→	2	2.0710	$\chi^2(2)$	[0.3550]
3	→	4	11.643	$\chi^2(2)$	[0.0030]

4	→	2	7.1776	$\chi^2(3)$	[0.0664]
4	→	3	7.6223	$\chi^2(3)$	[0.0545]
Notas: Variável 1 corresponde a default total; 2 a Lnjuros real; 3 a Dlnselic e 4 a Hiato do Produto.					

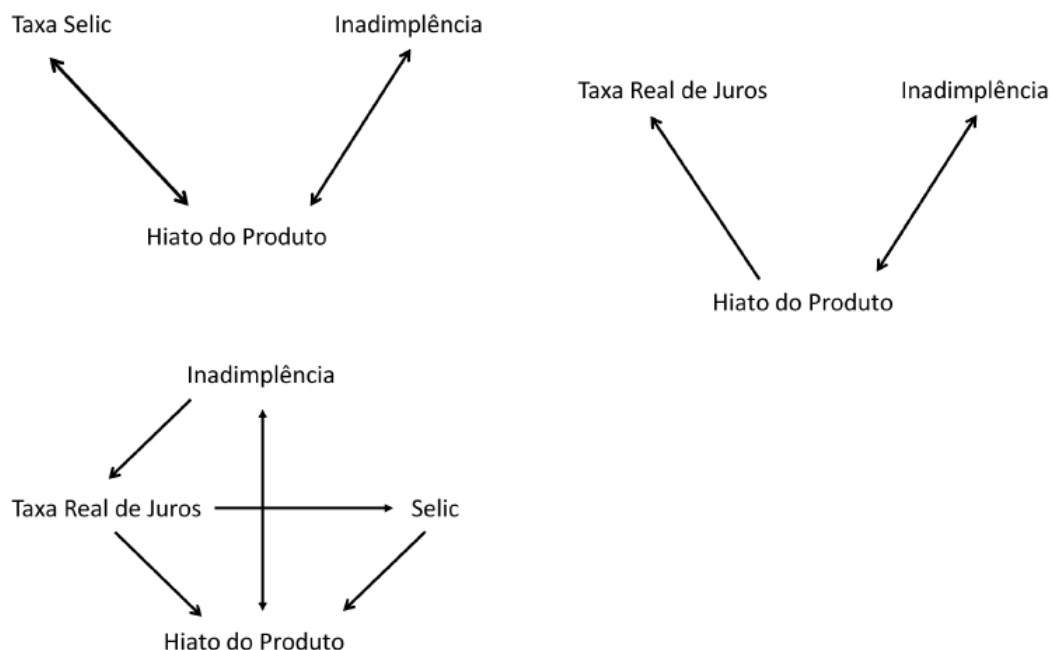


Figura 4: Resumo dos testes de Causalidade de Granger: 5% de significância

## 5.2. Avaliação do Poder Preditivo

O ajuste de um modelo dentro da amostra não necessariamente se converte em bom poder preditivo fora da amostra. Alterações inesperadas, instabilidades na estrutura, incerteza na estimação dos parâmetros entre outros fatores podem levar modelos cujas variáveis tenham um bom ajuste dentro da amostra, gerem um desempenho preditivo baixo. Nesta seção deseja-se avaliar se as variáveis macroeconômicas ajudam a melhorar o poder preditivo dos modelos e em que horizontes de tempo. A janela de previsão contempla o período a partir de janeiro de 2008 até o final da amostra.

### 5.2.1. Analisando a linha temporal

A fonte primária dos dados de inadimplência analisados é o Banco Central do Brasil. A instituição informa que a divulgação destes indicadores é feita na publicação “Nota para imprensa” com a defasagem de pelo menos quatro semanas (um mês) com relação ao último dia útil do mês que se refere o indicador. A primeira divulgação do índice de produção industrial é feita com dois meses de defasagens. O índice de inflação medido pelo IPCA é feito cerca de uma semana após o fechamento do mês. A taxa selic é conhecida ao final do dia útil. Desta forma, a produção industrial é a série que mais restringe a elaboração da previsão. Todos os dados estarão na melhor das hipóteses um mês a frente do dado de produção industrial.

Por exemplo, no último dia de novembro, temos disponíveis os dados de inadimplência de outubro, a taxa selic e IPCA de novembro, e a produção industrial de setembro. Com isto o analista em tempo real é capaz de gerar previsões para outubro, novembro, dezembro, janeiro do ano seguinte e assim por diante. Note que a previsão para outubro para inadimplência só é factível quando este dado já está disponível. Desta forma a previsão um passo a frente não é útil.

Na confecção da projeção dois passos a frente - final de dezembro no exemplo - seria possível utilizar os dados de inflação, juros e inadimplência disponíveis até novembro em tempo real, mas opta-se no exercício deste trabalho utilizar valores projetados. A razão para deve-se a uma atitude conservadora na avaliação dos modelos. Como estamos trabalhando com os dados revistos e não com a primeira divulgação dos dados de produção, optamos por descartar a informação disponível dos demais indicadores para tentar controlar um viés de melhor desempenho do nosso procedimento frente ao factível em tempo real. Mas na execução em tempo real, não há razão para descartar tal informação. Como os dados de produção industrial são defasados de dois períodos, as previsões feitas em tempo real com os modelos para dois passos a frente, na prática, equivalem a exercício de nowcast, enquanto a previsão 3 passos a frente equivale a previsão para um mês a frente e assim sucessivamente. Esta nomenclatura será utilizada daqui para

frente. Um exercício de nowcast consiste em utilizar um modelo para prever uma variável no instante atual, mas não divulgada ainda, a partir de indicadores disponíveis que tragam informação sobre a mesma.

Na avaliação de um modelo de previsão, algumas questões tem de ser avaliadas. Uma forma de avaliar consiste em escolher um benchmark bem ingênuo que qualquer modelo razoável deveria ser capaz de batê-lo. Iremos escolher como benchmark o passeio aleatório. Neste modelo, a melhor previsão é dada pela não alteração da variável no horizonte de previsão.

### **5.2.2. Resultados do exercício de simulação**

Nesta parte do trabalho apresentamos os resultados do ranqueamento do modelos estimados. Como conclusão geral, o desempenho preditivo dos indicadores macroeconômicos é melhor para as séries de inadimplência do setor privado. O desempenho preditivo dos modelos na série pública é bem inferior.

#### **5.2.2.1. Inadimplência Total**

Na tabela abaixo apresenta-se os resultados dos modelos classificados pelo algoritmo MCS. Reportamos apenas os modelos finalistas que não puderam ser excluídos ao nível de 10%. O modelo com p-valor 1 é o melhor modelo em termos de erro quadrático médio ou erro absoluto médio, mas não é possível afirmar que o melhor desempenho frente aos demais que estão no conjunto final tenha acontecido por acaso.

No horizonte de nowcast e de um mês a frente é difícil bater o modelo de passeio aleatório. Este resultado é esperado na medida em que mudanças nas variáveis macroeconômicas impactam com defasagem sobre os indicadores de inadimplência. Na análise de seis meses a frente, os modelos com indicadores macroeconômico parecem ter o melhor desempenho colocando o modelo de passeio aleatório num nível próximo de ser eliminado do conjunto final. Os modelos com dois indicadores macroeconômicos (1 e 2) são o que



tem melhor desempenho frente o modelo com 3 indicadores (modelo 3) no horizonte de 12 meses. Os modelos que supõe estacionaridade da série de inadimplência tem melhor desempenho que os modelos que não fazem esta hipótese. Este resultado está em linha com o teste de raiz unitária, sugerindo que determinar a ordem de integração pode ser importante para previsão.

A nomenclatura abaixo seja a definida na Tabela 2. Os modelos M1-b, M2-b e M3-b contém as mesmas variáveis que os modelos M1, M2 e M3, com a diferença que a série de inadimplência é suposta não estacionária e logo modelada em primeira diferença nos modelos b e nível nos outros casos.

Tabela 10: Avaliação da qualidade das previsões do modelos estimados - Inadimplência Total

Nowcast			1 mês à frente			6 meses à frente			12 meses à frente		
Modelos	EQM	p-valor	Modelos	EQM	p-valor	Modelos	EQM	p-valor	Modelos	EQM	p-valor
RW-1	0.01779	1.0000	AR-12	0.04951	1.0000*	M2	0.35973	1.0000*	M1	0.81443	1.0000*
AR-12	0.02529	0.1611	RW-2	0.05120	0.9868*	M1	0.37714	0.9526*	AR-12	0.83128	0.9196*
M1-b	0.03088	0.1611	M1	0.05385	0.9868*	AR12	0.38475	0.9526*	RW-13	1.0774	0.2002*
M1	0.03212	0.1611	M1b	0.05329	0.9868*	M3	0.39321	0.4889*	M2	0.93265	0.1525*
			M2	0.05620	0.9656*	RW-7	0.45866	0.2418*			
			M2b	0.06074	0.8250*						
			M3	0.06298	0.8250*						
			M3-b	0.07142	0.3331*						
Modelos	EAM	p-valor	1 mês à frente			Modelos	EAM	p-valor	Modelos	EAM	p-valor
RW-1	0.10674	1.0000	M1	0.16690	1.0000*	M 2	0.43601	1.0000*	M-1	0.68310	1.0000*
AR-12	0.11736	0.4865	AR-12	0.16747	0.9787*	M 1	0.44491	0.7978*	AR-12	0.80300	0.1583*
M1	0.12373	0.4865	M2	0.18048	0.8974*	AR-12	0.47868	0.7629*	M3	0.78097	0.1209*
M1-b	0.1269	0.4222	RW-2	0.18132	0.8974*	M-3	0.47153	0.2319*	RW-13	0.86075	0.1209*
M2	0.13275	0.2986	M3	0.18712	0.8974*	RW-7	0.53105	0.1290*			
			M1-b	0.17069	0.9177*						
			M2-b	0.19032	0.6885*						
			M3-b	0.19674	0.6885*						

Notas: EAM é o Erro Absoluto Médio\*1000; EQM é o Erro Quadrático Médio \*1000 e Modelos que estão no Conjunto Final ao nível de 10% de significância.

### 5.2.2.2. Inadimplência Bancos Privados Nacionais

A tabela abaixo apresenta o desempenho dos indicadores macroeconômicos para um dos componentes da série de inadimplência total: a série de inadimplência dos bancos nacionais. Neste caso o desempenho dos modelos com indicadores macroeconômicos é bem superior para os horizontes de seis meses e um ano. Em nenhum destes casos, nem o passeio aleatório nem o modelo AR-12 estão no conjunto final. No horizonte de doze meses o modelo M1 está sozinho. No caso de um ano, apenas o modelo 1 é o finalista, eliminando todos os demais concorrentes. No curto prazo, contudo o desempenho continua sendo inferior ao passeio aleatório.

Tabela 11: Avaliação da qualidade das previsões do modelos estimados - Inadimplência Bancos Privados Nacionais											
Nowcast			1 mês à frente			6 meses à frente			12 meses à frente		
Modelos	EQM	p-valor	Modelos	EQM	p-valor	Modelos	EQM	p-valor	Modelos	EQM	p-valor
RW-1	0.01973	1.0000	RW-2	0.05389	1.0000	M1	0.22354	1.0000	M1	0.38421	1.0000
			M 1-b	0.0601	0.7279	M2	0.2962	0.1866			
			M1	0.06218	0.7279						
			AR-12	0.06722	0.5135						
			M2	0.0764	0.2788						
			M3	0.08758	0.1225						
Modelos	EAM	p-valor	Modelos	EAM	p-valor	Modelos	EAM	p-valor	Modelos	EAM	p-valor
RW-1	0.11239	1.0000	M1-b	0.18986	1.0000	M1	0.35894	1.0000	M1	0.45515	1.0000
			RW-2	0.19022	0.9858	M2	0.39947	0.2328			
			M 1	0.19272	0.9858	M3	0.44812	0.1134			
			AR-12	0.20221	0.9095						
			M2	0.21316	0.7612						
			M3	0.2275	0.4269						

Notas: EAM é o Erro Absoluto Médio\*1000; EQM é o Erro Quadrático Médio \*1000 e Modelos que estão no Conjunto Final ao nível de 10% de significância.

### 5.2.2.3. Inadimplência Bancos Privados Estrangeiros

O outro componente da inadimplência do setor privado é dado pelo carteira dos bancos estrangeiros. Neste caso os indicadores macroeconômicos brasileiros não tem um bom desempenho. A razão talvez seja a característica estocástica das séries. A série de inadimplência apresenta sinais claros de não estacionaridade, o que torna o passeio aleatório um oponente difícil de ser batido. Os modelos que supõe que a série de inadimplência é não estacionária são aqueles com maior poder preditivo, embora nenhum deles tenha desempenho superior ao passeio aleatório, mas estão no conjunto final para os horizontes de um, seis e doze meses.

Tabela 12: Avaliação da qualidade das previsões do modelos estimados - Inadimplência Bancos Privados Estrangeiros											
Nowcast			1 mês a frente			6 meses frente			12 meses a frente		
Modelos	EQM	p-valor	Modelos	EQM	p-valor	Modelos	EQM	p-valor	Modelos	EQM	p-valor
RW-1	0.03075	1.0000	RW-2	0.0781	1.0000	RW-7	0.58817	1.0000	RW-13	1.33957	1.0000
			AR-12	0.10004	0.5460	M 2-b	0.66678	0.6450	AR-12	1.51307	0.5804
			M 2-b	0.10436	0.5460	AR-12	0.66702	0.6450	M 2-b	1.91159	0.5804
			M 2	0.11875	0.5460	M 2	0.72851	0.5918	M 2	1.92637	0.5804
			M 1-b	0.13893	0.5460	M 1b	0.8914	0.2185	M 1	2.0181	0.5804
			M 1	0.15667	0.5350	M 1	0.93073	0.2185	M 3	2.05606	0.5804
			M 3-b	0.15147	0.2764	M 3	0.97248	0.2185	M 1-b	2.11029	0.5804
			M 3	0.1587	0.1374	M 3-b	1.00175	0.2158	M 3-b	2.70316	0.3542
Modelos	EAM	p-valor	Modelos	EAM	p-valor	Modelos	EAM	p-valor	Modelos	EAM	p-valor
RW-1	0.12652	1.0000	RW-2	0.19670	1.0000	RW-7	0.57512	1.0000	RW-13	0.98225	1.0000
						M 2b	0.61694	0.7016	M 2-b	1.02884	0.7793
						AR-12	0.62743	0.7016	AR-12	1.06999	0.7793
						M 2	0.64774	0.6474	M 2	1.08069	0.7793

						M 1b	0.68224	0.4714	M 3	1.09629	0.7793
						M 1	0.7018	0.4714	M 1	1.10647	0.7793
						M 3b	0.70681	0.4714	M 1-b	1.10084	0.6563
						M 3	0.71866	0.4714	M 3-b	1.1894	0.2540
Notas: EAM é o Erro Absoluto Médio*1000; EQM é o Erro Quadrático Médico *1000 e Modelos que estão no Conjunto Final ao nível de 10% de significância.											

#### 5.2.2.4. Inadimplência Público

Por fim a série de inadimplência dos bancos públicos tem claramente sinais de mudança estrutural e sua previsão a partir de indicadores macroeconômicos é claramente ineficaz. Em todos os casos o modelo de passeio aleatório mostrou-se superior sendo o único presente no conjunto final de modelos.

Tabela 13: Avaliação da qualidade das previsões do modelos estimados - Inadimplência Bancos Públicos											
Nowcast			1 mês à frente			6 meses à frente			12 meses à frente		
Modelos	EQM	p-valor	Modelos	EQM	p-valor	Modelos	EQM	p-valor	Modelos	EQM	p-valor
RW-1	0.00974	1.0000	RW-2	0.01756	1.0000	RW-7	0.06682	1.0000	RW-13	0.13804	1.0000
Modelos	EAM	p-valor	Modelos	EQM	p-valor	Modelos	EAM	p-valor	Modelos	EAM	p-valor
RW-1	0.07511	1.0000	RW-2	0.10670	1.0000	RW-7	0.19849	1.0000	RW-13	.29350	1.0000
Notas: EAM é o Erro Absoluto Médio*1000; EQM é o Erro Quadrático Médico *1000 e Modelos que estão no Conjunto Final ao nível de 10% de significância.											

#### 5.2.2.5. Limitações do Trabalho e Possíveis Extensões

O trabalho aqui realizado pode ser aprimorado em várias frentes. Um primeiro ponto diz respeito a uma melhor mensuração do ciclo econômico real

e monetário. Utilizou-se apenas a taxa selic e a produção industrial como proxy do ambiente econômico. Mesmo assim obteve uma evidência forte da importância destas em explicar a dinâmica da inadimplência privada no Brasil. Existe um conjunto de mais amplo com indicadores de emprego, produção, sondagem de atividade econômica, diversos tipos de taxas de juros e indicadores monetários que poderiam ser utilizados para uma melhor compreensão do ambiente macroeconômico.

No que tange ao modelo dentro da amostra, há a possibilidade de tentar avançar na direção de identificação de choques econômicos e buscar um melhor entendimento da dinâmica de reação da inadimplência a diferentes tipos de choque. Um exemplo clássico de estratégia de identificação pode ser encontrado em Blanchard & Quah (1988).

Existe também a possibilidade da utilização de modelos de espaço estado multivariados desenvolvidos por Harvey (1990) e Durbin & Koopman (2001). Este modelo também pode ser uma interessante opção na construção de previsões.

No que tange a previsão há uma ampla agenda de pesquisa que pode ser tocada. Uma revisão extensa é feita em Elliot & Timmermann (2008). Existem avanços importantes tanto na área de geração de previsões quanto nas técnicas de seleção de modelos.

Um primeiro refinamento está em linha com a busca de melhor mensuração do ciclo econômico. Modelos fatoriais no espírito proposto por Stock & Watson (2002a-b) são opções a serem tentadas.

Um das grandes razões pelas quais as previsões econômicas falham tem a ver com alterações não antecipadas das médias dos processos por Clements & Hendry (1998) e Clements & Hendry (2001). Modelos estatísticos com estruturas flexíveis podem ajudar a melhorar o poder preditivo na medida em que mudanças estruturais podem ser tratadas e incorporadas à elaboração de previsões.

Dentro deste espírito há a literatura de correção de viés. No caso de uma previsão apresentar um viés sistemático de erro numa determinada

direção, existem técnicas que permite tentar corrigir o viés pela adição de termos corretivos a previsão realizada. Ver sobre este ponto Issler & Lima (2009).

Atualmente, existe a possibilidade de confecção de uma ampla gama de modelos para previsão. Num texto seminal, Granger & Ramanathan (1984) propõe que uma combinação de previsões de modelos pode ter um poder preditivo melhor do que cada uma das previsões individuais. De lá para cá uma vasta literatura discute técnicas de combinação de previsão e os potenciais ganhos das mesmas veja Timmermann (2006).

Também há a possibilidade de aprimorar o trabalho no que tange a técnicas de seleção de modelos. O método proposto por Hansen et al. (2011) e utilizado neste trabalho é um claro avanço nesta área, mas mais pode ser feito. Por exemplo, é possível utilizar outras métricas com funções de perda que levem em conta assimetria. Tanto erro quadrático médio, quanto erro absoluto médio são métricas que utilizam simetria. Muitas vezes, os analistas dão valores diferentes a erros positivos e negativos. Por exemplo, um Banco Central que utiliza metas de inflação talvez esteja mais preocupado em subestimar do que superestimar a inflação no futuro próximo. Um banco comercial talvez esteja mais preocupado com a subestimação do nível de inadimplência.

Por fim um refinamento possível é feito em Linardi & Ferreira (2008). Consiste em utilizar o VAR estimado para elaborar um leque de previsões possíveis e atribuir as mesmas probabilidade de ocorrência. Com isto é possível construir cenários para valores máximos e mínimos de evolução da inadimplência e não apenas a previsão de valores médios. Desta forma é possível construir limites a partir dos quais há baixa probabilidade de serem ultrapassados pelo nível de inadimplência. Tal estratégia é importante para estabelecer valores para testes de stress da carteira e simulações de perda máxima.

## 6. CONCLUSÃO

Este trabalho teve por objetivo avaliar em que medida indicadores macroeconômicos contribuem para explicar a dinâmica da inadimplência brasileira. Utilizando dados de inadimplência total privada foi possível mostrar que indicadores macroeconômicos apresentam uma relação robusta com ciclo econômico e com o ciclo monetário. O processo gerador dos dados estimado foi um Vetor Autoregressivo (VAR) e o melhor modelo foi selecionado utilizando a abordagem geral para específico proposta pela LSE. Obteve-se modelos que satisfazem as condições para serem uma boa aproximação dos dados em termos de testes de especificação tais como normalidade, ausência de correlação e homocedasticidade nos resíduos, bem como ausência de evidência de mudança estrutural.

Os modelos sugerem que há uma relação robusta e direta entre ciclo econômico e inadimplência. O ciclo monetária também é importante mas sua influência se dá via impacto que gera o ciclo econômico. O efeito existe mas é essencialmente indireto.

Um exercício de simulação de previsão em tempo real também foi feito e sugere que as variáveis macroeconômicas tem um melhor desempenho preditivo em horizontes mais largos e para as séries de inadimplência privada. Para as séries de inadimplência pública o desempenho preditivo dos modelos com variáveis macroeconômicas é bem desapontador.

Por fim discute-se uma série de limitações e possíveis extensões desta trabalho, principalmente no que tange o de desenvolvimento de modelos com bom poder preditivo para inadimplência. A possibilidade de construção de previsões precisas é um importante tópico de pesquisa em Economia e avanços recentes vêm sendo realizados.

## Referências

Baxter, Marianne & King, Robert G. (1999) Measuring business cycles: Approximate band-pass filters for economic time series. *The Review of Economics and Statistics*, 81: 575-593.

Beck, Roland; Jakubik, Petr & Piloju, Anamaria (2013) Non-performing loans: What matters in addition to the economic cycle. European Central Bank working paper no 1515.

Bernanke, Ben S.; Gertler, Mark & Gilchrist, Simon (1999) The financial accelerator in a quantitative business cycle framework. Em *Handbook of macroeconomics*, 1:1341-1393, Editado por J. B. Taylor & M. Woodford, Elsevier Science: Amsterdam.

Blanchard, Olivier J. & Quah, Danny (1988) The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, 79: 655-673.

Castle, Jennifer; Hendry, David & Clements, Michael P. (2014) Robust Approaches to Forecasting. *Economics Series Working Papers 697*, University of Oxford, Department of Economics.

Claessens, Stijn ; Kose, M. Ayhan & Terrones, Marco E. (2012) How do business and financial cycles interact? *Journal of International Economics*, 87(1):178 - 190.

Clements, Michael P & Hendry, David F. (2001) *Forecasting non-stationary economic time series*. MIT Press: Boston.

Clements, Michael P & Hendry, David F. (1998) *Forecasting economic time series*. Cambridge University Press: Cambridge.

Correa, Arnildo da S.; Marins, Jaqueline T. M.; das Neves, Myrian B. E. & da Silva, Antonio C. M. (2014) Credit default and business cycles: An empirical investigation of Brazilian retail loans. *Revista Brasileira de Economia*, 68 (3): 337-362.



Cusinato, Rafael T.; Minella, André & Pôrto Júnior, Sabino da S. (2013) Output gap and GDP in Brazil: a real-time data analysis. *Empirical Economics*, 44 (3): 1113-1127.

Durbin, James & Koopman, Siem J. (2001) *Time Series Analysis by State Space Methods*. Oxford University Press: Oxford.

Elliott, Graham & Timmermann, Allan (2008) Economic forecasting. *Journal of Economic Literature*, 46 (1): 3-56.

Gourieroux, Christian & Jasiak, Joann (2001) *Financial Econometrics*. Princeton University Press: Princeton.

Granger, Clive W. J. (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3): 424-438.

Granger, Clive, W. & Ramanathan, Ramu (1984) Improved methods of combining forecasts. *Journal of Forecasting*, 3 (2): 197-204.

Hamilton, James D. (1994) *Time Series Analysis*. Princeton University Press: Princeton.

Hansen, Peter R.; Lunde, Asger & Nason, James M. (2011) The model confidence set. *Econometrica*, 79(2): 453-497.

Harvey, Andrew C. (1990) *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press: Cambridge.

Hendry, David F. & Doornik, Jurgen A. (2014) *Empirical Model Discovery and Theory Evaluation: Automatic Selection Methods in Econometrics*. MIT Press: Boston.

Hendry, David F. (1995) *Dynamic Econometrics*. Oxford University Press: Oxford.

Hodrick, Robert J. & Prescott, Edward C. (1997) Postwar u.s. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1): 1-16.

Issler, João V. & Lima, Luiz R. (2009) A panel data approach to economic forecasting: The bias-corrected average forecast. *Journal of Econometrics*, 152(2): 153-164.

Linardi, Fernando de M. & Ferreira, Mauro. S. (2008) Avaliação dos determinantes macroeconômicos da inadimplência bancária no Brasil. XXXVI Encontro Nacional de Economia, Salvador, Bahia.

Meese, Richard A. & Rogoff, Kenneth (1983) Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of international economics*, 14 (1): 3-24.

Nielsen, Heino B. (2004) Cointegration analysis in the presence of outliers. *The Econometrics Journal*, 7 (1): 249-271.

Stock. James H. & Matson, Mark W. (2002-a) Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97 (460): 1167-1179.

Stock. James H. & Matson, Mark W. (2002-b) Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20 (2): 147-162

Tabak, Benjamin, M.; Guerra, Solange, M.; Lima, Eduardo, J. A. & Chang, Eui J. (2008) The stability-concentration relationship in the Brazilian banking system. *Journal of International Financial Markets, Institutions and Money*, 18(4):388-397.

Timmermann, Alan (2006) Forecast combinations, Em G. Elliot, C. Granger & A. Timmermann (ed.) *Handbook of Economic Forecasting*, vol 1, n0 1: 135-196. Elsevier|: Amsterdam.

Valls Pereira, Pedro L. (1986) Estimação do hiato do produto via componentes não observados. *Brazilian Review of Econometrics*, 6 (2): 81-95.