

LÍGIA APARECIDA FRANCISCO ALBANEZ

ANÁLISE DA INFLUÊNCIA DA TAXA DE JUROS E DO DESEMPREGO NA
INADIMPLÊNCIA

Monografia apresentada à Escola Politécnica
da Universidade de São Paulo, para
obtenção do título de MBA em Engenharia
Financeira.

Área de concentração:

Engenharia Financeira

Orientador: Prof. Dr. Claudio Garcia

SÃO PAULO

2015

MBA/EF
2015
A13a



Escola Politécnica - EPEL



31500023144

Catálogo-na-publicação

M 2015 N

Albarez, Lígia

Análise da influência da taxa de juros e do desemprego na inadimplência / L. Albarez -- São Paulo, 2015.

64 p.

Monografia (MBA em Engenharia Financeira) - Escola Politécnica da Universidade de São Paulo. PECE - Programa de Educação Continuada em Engenharia.

1.Séries Temporais 2.Modelos Estatísticos 3.Econometria 4.Desemprego 5.Juros Universidade de São Paulo. Escola Politécnica. PECE - Programa de Educação Continuada em Engenharia II.t.

[2752514]

AGRADECIMENTOS

Ao professor Dr. Claudio Garcia pelas orientações e atenção nesse período.

Ao Fernando Fernandes Neto por compartilhar sua experiência em modelos econométricos.

RESUMO

A inadimplência é um dos fatores que mais preocupa as instituições financeiras e empresas de um modo geral, independente do seu tamanho ou faturamento, se for considerado que impacta diretamente na receita. O conhecimento sobre quais variáveis podem afetar esse índice é interessante e pode ser utilizado nos modelos de previsões da inadimplência e do risco de crédito. O objetivo deste trabalho foi estudar a influência dos índices de desemprego e taxa de juros reais na inadimplência do consumidor, a partir da análise das séries históricas divulgadas pelos órgãos responsáveis, no período entre janeiro de 2004 a agosto de 2015. Foram utilizados modelos auto-regressivos (ARIMA) para as análises de aderência dos dados, relevância das variáveis independentes e erros dos modelos. Neste estudo foi observado que o modelo ARIMA(1,0,1) representa melhor os dados originais e que no modelo ARIMA(0,0,2) as variáveis independentes são relevantes e explicam a inadimplência. Entretanto observou-se que os modelos são mais adequados para serem aplicados nas previsões de curto prazo.

Palavras-chave: Inadimplência, Desemprego, Juros e ARIMA

ABSTRACT

The default is one of the factors that most concerns the financial institutions and companies in general, regardless of their size or billing if it is considered that directly impacts revenue. The knowledge about which variables may affect this rate is interesting and can be used in forecasting models of default and credit risk. The objective of this work was to study the influence of unemployment rates and real interest rates on consumer defaults, based on the analysis of historical data published by responsible authorities in the period from January 2004 to August 2015. Models autoregressive (ARIMA) were used to adjust the data of analysis, relevance of independent variables and models' errors. In this study it was observed that the ARIMA model (1,0,1) better represents the original data and that in the ARIMA model (0,0,2) the independent variables are relevant and explain the default. However it was observed that the models are more appropriate to be applied in short-term forecasts.

Keywords: Default, Unemployment, Interest rate e ARIMA

LISTA DE ILUSTRAÇÕES

Figura 1 - Oferta de crédito observada no período de 2011 a 2015.	14
Figura 2 - Demanda de crédito observada no período de 2011 a 2015.	15
Figura 3 - Curva de Phillips.	17
Figura 4 - Realizações de um processo estocástico.	20
Figura 5 - Correlograma da variável inadimplência dessazonalizada.	29
Figura 6 - Série de Juros Reais no período entre Janeiro/2004 a Agosto/2015.	31
Figura 7 - Série do IPCA acumulado 12 meses no período entre Janeiro/2004 a Agosto/2015.	32
Figura 8 - Taxa de desemprego original no período entre Janeiro/2004 a Agosto/2015.	33
Figura 9 - Série de desemprego original e série dessazonalizada no período entre Janeiro/2004 a Agosto/2015.	33
Figura 10 - Série de inadimplência original no período entre Janeiro/2004 a Agosto/2015.	34
Figura 11 - Série de inadimplência original e série dessazonalizada no período entre Janeiro/2004 a Agosto/2015.	35
Figura 12 - Variáveis observadas no período entre Janeiro/2004 a Agosto/2015.	36
Figura 13 - Análise de Componentes Principais, gerada através do software Gretl 2015d.	36
Figura 14 - Parâmetros estimados do Modelo ARIMA (1,0,0).	39
Figura 15 - Comparação do Modelo ARIMA (1,0,0) com a série temporal de inadimplência.	40
Figura 16 - Parâmetros estimados do Modelo ARIMA (0,0,1).	41
Figura 17 - Comparação do Modelo ARIMA (0,0,1) com a série temporal de inadimplência.	41
Figura 18 - Parâmetros estimados do Modelo ARIMA (0,0,2).	42
Figura 19 - Comparação do Modelo ARIMA (0,0,2) com a série temporal de inadimplência.	43
Figura 20 - Parâmetros estimados do Modelo ARIMA (1,0,1).	44
Figura 21 - Comparação do Modelo ARIMA (1,0,1) com a série temporal de inadimplência.	45
Figura 22 - Parâmetros estimados do Modelo ARIMA (1,0,0) com PCA.	46
Figura 23 - Comparação do Modelo ARIMA (1,0,0) com PCA em relação a série temporal de inadimplência.	47
Figura 24 - Parâmetros estimados do Modelo ARIMA (0,0,1) com PCA.	48
Figura 25 - Comparação do Modelo ARIMA (0,0,1) com PCA em relação a série temporal de inadimplência.	48
Figura 26 - Parâmetros estimados do Modelo ARIMA (0,0,2) com PCA.	49
Figura 27 - Comparação do Modelo ARIMA (0,0,2) com PCA em relação a série temporal de inadimplência.	50
Figura 28 - Parâmetros estimados do Modelo ARIMA (1,0,1) com PCA.	51
Figura 29 - Comparação do Modelo ARIMA (1,0,1) com PCA em relação a série temporal de inadimplência.	52
Figura 30 - Parâmetros estimados do Modelo ARIMA (1,0,1) sem variáveis regressoras.	53
Figura 31 - Comparação do Modelo ARIMA (1,0,1) sem variáveis regressoras em relação a série temporal de inadimplência.	54
Figura 32 - Parâmetros estimados do Modelo ARIMA (0,0,2) com atraso de 5 meses na variável desemprego.	55
Figura 33 - Parâmetros estimados do Modelo ARIMA (0,0,2) com atraso de 6 meses na variável desemprego.	56
Figura 34 - Comparação do Modelo ARIMA (0,0,2) com atrasos na variável desemprego em relação a série temporal de inadimplência.	57
Figura 35 - Previsões do Modelo ARIMA (1,0,1) para 3 períodos à frente.	59
Figura 36 - Previsões do Modelo ARIMA (0,0,2) para 3 períodos à frente com PCA.	60

Figura 37 - Previsões do Modelo ARIMA (0,0,2) para 3 períodos a frente com atrasos na variável desemprego.....	60
---	----

LISTA DE ABREVIATURAS E SIGLAS

AIC	Critério de Informação de Akaike
BIC	Critério de informação de Schwartz
IBGE	Instituto Brasileiro de Geografia e Estatística
IPCA	Índice Nacional de Preços ao Consumidor Amplo
IPTU	Imposto predial e territorial urbano
IPVA	Sistema Gerenciador de Séries Temporais
PCA	Análise de Componentes Principais
SFN	Sistema Financeiro Nacional
SGS	Serviço de Proteção ao Crédito
SQR	Soma Quadrática dos Resíduos

SUMÁRIO

1. Introdução	11
2. Referencial teórico	13
2.1 Revisão bibliográfica	13
2.2 Crédito	13
2.3 Inadimplência	15
2.4 Desemprego	17
2.5 Taxa de Juros	18
2.6 IPCA	18
2.7 Séries temporais	19
2.8 Análise Estatística de Séries Temporais	19
2.8.1 Gráficos	19
2.8.2 Processos Estocásticos	20
2.8.3 Estacionariedade	20
2.8.4 Sazonalidade	21
2.8.4.1 Filtro de Médias Móveis	22
2.8.5 Ruído Branco	22
2.8.6 Resíduo	22
2.8.7 Regressão Linear	23
2.8.8 Análise de Componentes Principais (PCA)	23
2.9.1 Modelo Autoregressivo (AR)	26
2.9.2 Modelo de Média Móvel (MA)	26
2.9.3 Modelo ARMA	27
2.9.4 Autocorrelação e Autocorrelação Parcial	27
2.9.5 Identificação da ordem do modelo	28
3. Metodologia	30
3.1 Obtenção dos dados	30
3.2 Dados utilizados para as análises	31
3.3.1 Juros Reais	31
3.3.2 IPCA	31
3.3.3 Desemprego	32
3.3.4 Inadimplência	34
3.3.5 Componentes principais	36
3.3 Softwares empregados nas análises	37
3.4 Critérios para a escolha das ordens do modelo	37
4. Simulações	39
4.1 Modelo ARIMA (1,0,0) com 3 variáveis	39
4.2 Modelo ARIMA (0,0,1) com 3 variáveis	40
4.3 Modelo ARIMA (0,0,2) com 3 variáveis	42
4.4 Modelo ARIMA (1,0,1) com 3 variáveis	43
4.5 Modelo ARIMA (1,0,0) com duas variáveis de componentes principais	45

4.6	Modelo ARIMA (0,0,1) com duas variáveis de componentes principais	47
4.7	Modelo ARIMA (0,0,2) com duas variáveis de componentes principais	49
4.8	Modelo ARIMA (1,0,1) com duas variáveis de componentes principais	50
4.9	Modelo ARIMA (1,0,1) sem variáveis regressoras	52
4.10	Modelo ARIMA (0,0,2) com atrasos na variável desemprego	54
4.11	Previsões 3 períodos à frente	58
5.	Conclusões	61
	Referências Bibliográficas	63

1. Introdução

Com as mudanças das políticas monetárias dos últimos anos, houve uma abertura da oferta de crédito no mercado, beneficiando algumas classes econômicas que historicamente não tinham acesso ao crédito.

As instituições financeiras têm um papel essencial na determinação das condições sob o qual o crédito é ofertado, seus objetivos são de ofertar crédito com menor risco de perda. Elas se baseiam em modelos matemáticos, dados históricos e variáveis econômicas tais como PIB, taxa de juros, inflação, renda, desemprego, inadimplência, entre outras.

O crédito sob o aspecto financeiro significa dispor recursos financeiros a um cliente para diversas finalidades, como pagamento de despesas, investimentos, aquisição de bens, entre outras.

No momento atual, a maioria das avaliações do risco de crédito estão relacionadas a fatores individuais do tomador, ou seja, renda mensal, idade, estado civil, quantidade de filhos, entre outros. Com o objetivo de melhorar a avaliação de risco, é provável que seja necessária a análise de fatores relacionados à economia monetária atual.

A taxa de inadimplência tornou-se um fator relevante na avaliação do risco de crédito. Segundo Houaiss (2001), a inadimplência pode ser definida como a falta de cumprimento de uma obrigação financeira até a data estabelecida e envolve questões econômicas, sociais e psicológicas.

O desemprego e a taxa de juros são fatores que podem estar relacionados ao não cumprimento das obrigações do tomador. Reinert (2001) caracteriza desemprego como sendo a não possibilidade do trabalho assalariado nas organizações de um modo geral.

Segundo Sobrinho (1981), a taxa de juros pode ser definida como a relação entre os juros pagos (ou recebidos) no final do período e o capital inicialmente tomado (ou aplicado) e esta pode ser classificada quanto ao regime de capitalização (simples e composta) e ao valor do capital inicial tomado como base de cálculo (nominal, efetiva e real).

Dentre as inúmeras variáveis que podem impactar diretamente ou indiretamente no aumento da inadimplência, o objetivo deste trabalho é estudar a possível influência ao longo do tempo da taxa de juros real e do desemprego nos índices de inadimplência divulgados.

Para a criação de um modelo estatístico que tenha a função de melhorar e auxiliar na assertividade das decisões das empresas, é necessária a utilização de séries históricas das variáveis que inicialmente possam impactar na situação a ser tratada.

Esse assunto torna-se interessante, pois estes índices (taxa de juros real e desemprego) são pouco utilizados ou considerados de baixa importância na avaliação do risco de crédito. Essas variáveis são ligadas à macroeconomia e não estão diretamente relacionadas com características individuais dos tomadores. A avaliação desses índices pode agregar qualidade na análise e minimizar as perdas por inadimplência.

De maneira a apresentar o trabalho, tem-se a seguinte estrutura:

No capítulo 2 tem-se a introdução do referencial teórico dos principais conceitos, necessidades e variáveis para a criação de um modelo de previsão dentro do escopo deste trabalho e como as áreas de finanças, economia, estatística e matemática se complementam nos estudos de previsão.

No capítulo 3 tem-se uma breve apresentação dos órgãos que disponibilizam as séries temporais dos índices utilizados, a metodologia para a transformação matemática da série de juros, as limitações das séries utilizadas e os *softwares* utilizados nas avaliações e os critérios de escolha do modelo utilizado.

No capítulo 4 são realizadas as simulações dos modelos e abordados os resultados obtidos após comparação entre eles.

Na conclusão são avaliados e discutidos se os resultados obtidos no trabalho são compatíveis com o objetivo inicial.

2. Referencial teórico

As principais áreas de pesquisa desse trabalho são:

- Gestão de risco de crédito (risco de não pagamento pelo tomador);
- Finanças (índices econômicos);
- Estatística e Matemática (estatística e análise de séries temporais);
- Economia (efeitos das variáveis macroeconômicas como juros e desemprego).

2.1 Revisão bibliográfica

As principais leituras para realização deste trabalho foram:

O trabalho de Rossi e Neves (2014), onde foi possível entender as premissas estatísticas de análise de séries temporais. Os autores se baseiam em conceito de autores como Box e Jenkins.

O livro de Barth (2004), em sua publicação sobre modelos de previsões de inadimplência, onde ele demonstra a construção de modelos baseados em diversas técnicas e cita as variáveis utilizadas nesses modelos.

O trabalho de Mankiw (2009) foi utilizado para conceituar as variáveis econômicas estudadas.

2.2 Crédito

Crédito é um termo que traduz confiança e deriva da expressão “crer”, acreditar em algo, ou em alguém.

Castro (2009), define sob o aspecto financeiro que “crédito é todo ato de vontade ou disposição de alguém ceder, temporariamente, parte do seu patrimônio a terceiro, com a expectativa que a parcela volte a sua posse integralmente após decorrido um tempo estimulado”. Por meio da definição do autor, observam-se dois itens fundamentais nesse conceito, a confiança no pagamento e o tempo entre a concessão de seu patrimônio até o pagamento da dívida.

Pode-se dividir o crédito basicamente em bancário, comercial, para pessoa física e para pessoa jurídica. Nesse trabalho, é abordada a oferta de crédito para pessoa física.

A oferta de crédito é indispensável para que haja desenvolvimento econômico do país, porém é necessário que exista a demanda de crédito pelos consumidores, que sinaliza a confiança econômica em cumprir com as obrigações de uma operação de crédito.

O cenário de crescimento econômico observado nos últimos anos, o aumento da oferta de crédito e a alta empregabilidade são alguns dos fatores que estão influenciando os consumidores a irem às compras, porém, levam aqueles que têm pouco controle sob suas despesas à possível situação de inadimplência.

O Banco Central do Brasil, através do SGS (Sistema Gerenciador de Séries Temporais), disponibiliza indicadores trimestrais de oferta e demanda de crédito do mercado, conforme mostrado nas Figuras 1 e 2. Pode-se observar que em alguns momentos da série tem-se uma boa correlação entre os dados (2011), mas em outros o mercado se comporta de maneira oposta (2012).

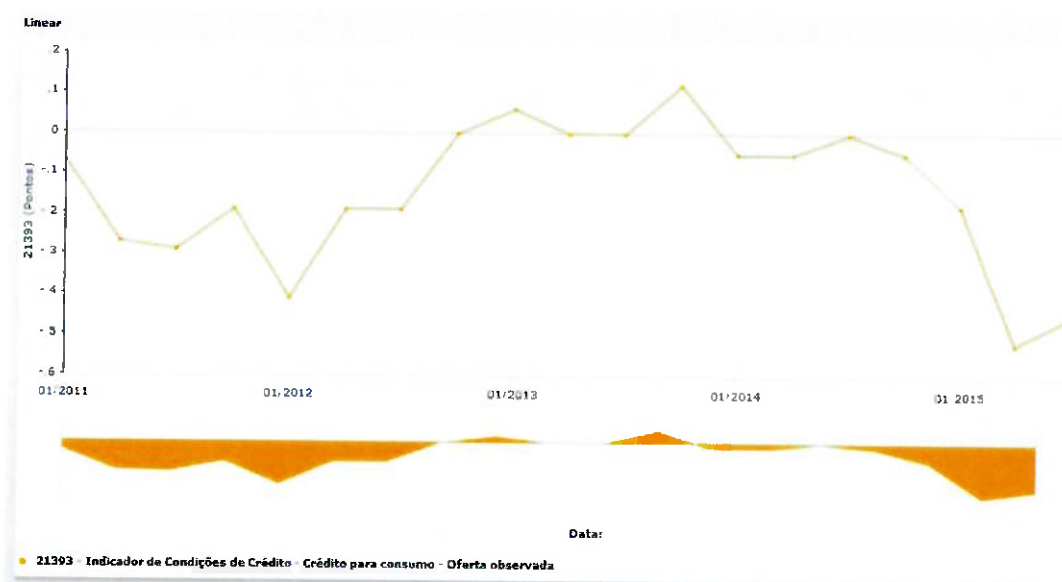


Figura 1 - Oferta de crédito observada no período de 2011 a 2015.

Fonte: <https://www3.bcb.gov.br/sgspub/>

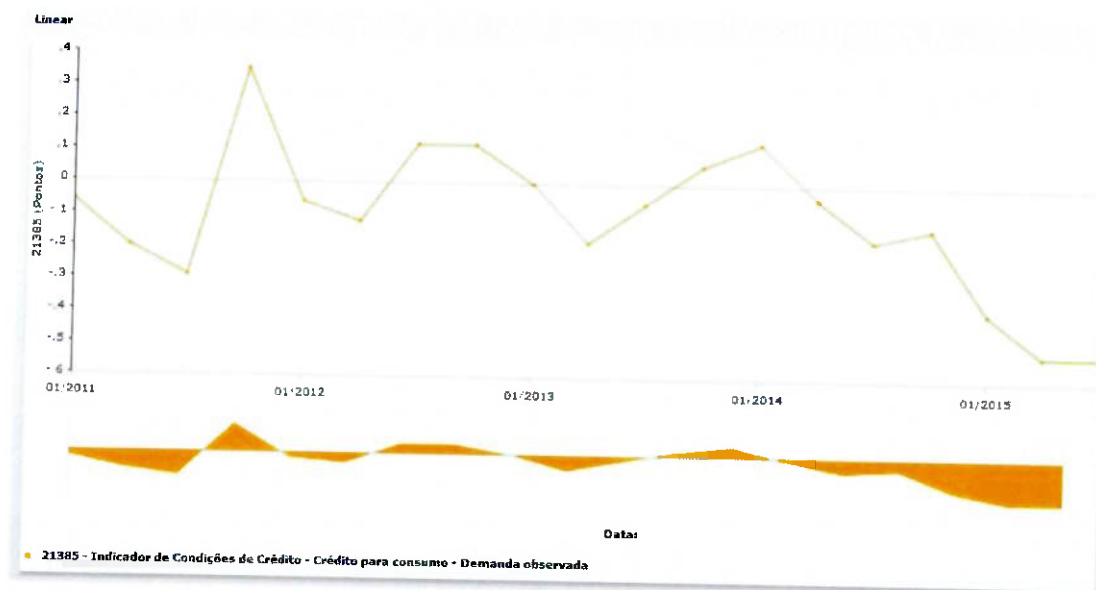


Figura 2 - Demanda de crédito observada no período de 2011 a 2015.

Fonte: <https://www3.bcb.gov.br/sgspub/>

Segundo economistas do SPC (Serviço de Proteção ao Crédito), uma possível alta taxa de juros pode contribuir para a diminuição da demanda de crédito, podendo ocasionar uma desaceleração da economia do país.

2.3 Inadimplência

O não cumprimento de uma obrigação financeira é caracterizado como uma inadimplência. A teoria que justifica a inadimplência dos consumidores, envolve fatores econômicos e psicológicos. Neste trabalho, são tratados apenas os fatores econômicos, devido a não existirem índices que demonstrem os fatores psicológicos das pessoas.

Um estudo da empresa Serasa Experian realizado no ano de 2014, revela que os índices de inadimplência são diferentes quando se analisam regiões do país e faixa etária dos consumidores. A inadimplência é uma das consequências de um baixo poder aquisitivo e um controle financeiro deficitário.

As empresas precisam conhecer os fatores que ocasionam a inadimplência para, a partir daí, utilizar práticas preventivas, podendo diminuir esses índices. A criação de um modelo de previsão pode auxiliar nas decisões, através de variáveis

individuais do consumidor e variáveis econômicas, disponibilizadas no mercado financeiro.

Outra característica da inadimplência é a sazonalidade, ou seja, existem épocas do ano em que o índice aumenta. Um exemplo, nos primeiros meses do ano, devido ao reflexo das compras de Natal e o pagamento de impostos como IPVA (Imposto sobre a propriedade de Veículos Automotores), IPTU (Imposto Predial e Territorial Urbano) e matrícula escolar, existe uma tendência do consumidor ficar inadimplente, mas em algumas situações, esse desequilíbrio financeiro pode refletir somente nos meses subsequentes. A inadimplência do consumidor, também afeta diretamente as empresas, pois elas deixam de receber dos clientes, conseqüentemente suas receitas diminuem e a probabilidade delas não conseguirem cumprir todos seus compromissos financeiros aumenta.

Barth (2004) afirma que é preciso estabelecer um critério para interpretar o resultado de uma operação de crédito. Quando o cliente será considerado um mau pagador? Cada empresa atua conforme sua política de crédito e possui o seu conceito de cliente inadimplente. Em algumas visões, o pagamento de contratos em dia significa sucesso na operação. Já em outras se o cliente atrasar três dias, poderá haver dúvidas em como classificá-lo. E, se houver atraso de um mês, mas houver cobrança de juros sem questionamentos, a operação pode ser lucrativa para a instituição. Tais dúvidas sobre a classificação demandam outros estudos, como o risco de crédito. No mercado de crédito, é natural que algumas perdas ocorram, mas a experiência da empresa na sua atividade e estabelecimento de critérios do que será considerado inadimplência é pré-requisito para a determinação do modelo.

Conforme acordo da Basileia II, para toda operação, existe uma perda esperada já prevista pelas instituições. Segundo Gerbelli (2015), quando maior for a inadimplência do consumidor junto às instituições, elas podem aumentar as taxas de juros das operações para compensar as perdas.

A inadimplência pode ocorrer devido ao consumo excessivo das famílias ou uma consequência do aumento da inflação, onde as famílias perdem o poder de compra devido à queda na renda e não conseguem cumprir com suas obrigações.

A prioridade dos consumidores é o pagamento das dívidas que estão relacionadas às necessidades básicas, como aluguel, alimentação, energia e água.

2.4 Desemprego

Reinert (2001), caracteriza desemprego como a não possibilidade do trabalho assalariado nas organizações.

Segundo Mankiw (2009), “perder o emprego pode ser um dos eventos econômicos mais desoladores na vida de uma pessoa. A maioria das pessoas conta com os ganhos do seu trabalho para manter o padrão de vida, e muitas obtêm não só a renda mas também um sentimento de realização pessoal. A perda do emprego significa um padrão de vida mais baixo no presente e uma ansiedade com o futuro e perda da autoestima”.

Um certo grau de desemprego é inevitável. Em uma economia complexa, algumas causas podem contribuir para o aumento do desemprego, como o desenvolvimento tecnológico, a globalização e os modernos métodos de gestão. Os índices de desemprego revelam quão bem está a situação econômica do país.

Segundo Mankiw (2009), no ano de 1958, o economista A. W. Phillips publicou um artigo que mostrava a existência de uma correlação negativa entre a taxa de desemprego e a taxa de inflação. Ou seja, Phillips provou que anos com baixo desemprego tendem a apresentar a inflação elevada e anos com desemprego elevado tendem a apresentar baixa inflação.

Quando a taxa de desemprego é baixa, a demanda aumenta, pois as pessoas empregadas consomem mais e a oferta não acompanha esse crescimento, com isso existe uma tendência de aumento dos preços e consequentemente da inflação, conforme Figura 3. Já quando o desemprego é alto, o mercado diminui o consumo e a oferta fica maior que a demanda e a tendência é que os preços não se elevem.

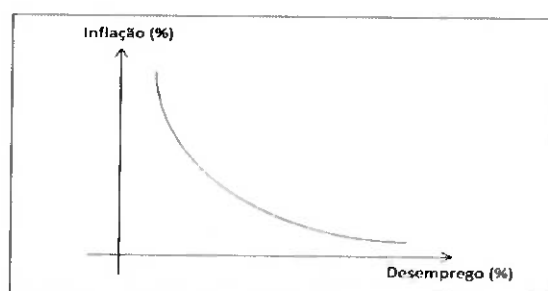


Figura 3 - Curva de Phillips.

Fonte: Google Imagens

2.5 Taxa de Juros

No mercado financeiro, existe uma confusão no que se refere ao conceito de taxa de juros nominal, efetiva e real. O desconhecimento generalizado desses conceitos, dificulta o fechamento de negócios e o entendimento das análises que envolvem os juros.

A taxa de juros que o banco paga é chamada taxa de juros nominal.

Mankiw (2009) define que a taxa de juros real é a diferença entre a taxa de juros nominal e a taxa de inflação. Como exemplo, a taxa de juros nominal, informa a velocidade a que cresce o número de dólares em sua conta bancária com o passar do tempo. A taxa de juros real nos diz a velocidade a que cresce o poder aquisitivo de sua conta bancária com o passar do tempo.

As taxas de juros nominal e real nem sempre se movem juntas. No fim da década de 1970, as taxas de juros nominais estavam elevadas. Mas como a inflação era muito alta, as taxas de juros reais eram baixas. Em alguns anos, as taxas de juros reais foram negativas porque a inflação corroía as poupanças mais rapidamente do que o pagamento de juros nominais as aumentavam. Na década de 1990, as taxas de juros nominais estavam baixas. Mas como a inflação também estava baixa, as taxas de juros reais estavam relativamente altas.

2.6 IPCA

Segundo Yassumura (2014), o IPCA (Índice Nacional de Preços ao Consumidor Amplo) é medido pelo IBGE desde 1979 e reflete o custo de vida para famílias com renda de 1 a 40 salários mínimos. A pesquisa é feita em 11 regiões metropolitanas e o período de coleta estende-se do dia 01 a 30 do mês de referência e mede a variação dos preços no comércio de uma cesta de produtos tais como: alimentação, bebidas, artigos de residência, comunicação, despesas pessoais, educação, habitação, saúde, cuidados pessoais, transportes e vestuário para o consumidor final. O IPCA é utilizado como índice oficial do governo para as metas de inflação.

Quando o IPCA sobe, significa que alguns dos itens da cesta sofrerão reajuste de preço para cima, então se terá inflação no período. Caso o IPCA no mês seguinte seja menor do que o do anterior, não significa que houve redução dos preços, somente que os preços subiram menos do que no mês anterior. Somente se o IPCA for negativo é que se terá deflação, ou seja, que os preços terão diminuído. Isso não acontece na série observada neste trabalho.

A inflação impacta nos juros, pois o governo pode aumentar os juros com o objetivo de diminuir a inflação.

2.7 Séries temporais

Segundo Stevenson (1981), o conceito de série temporal ou histórica, é uma sequência de observações organizadas cronologicamente, tais como vendas anuais de produtos, valores semestrais de produção, dados mensais de desemprego, preços diários de ações na bolsa de valores, registros de temperaturas diárias, entre outros. Na maioria das séries, o período entre as observações é constante, mas esse não é um aspecto necessário para se ter uma série histórica.

As características dos métodos de séries temporais é que estes utilizam os dados observados do passado para realizar previsões normalmente a curto prazo das variáveis.

Segundo Rossi (2014), os métodos de séries históricas em sua maioria necessitam de muitas observações, principalmente na fase de análise que precede a modelagem, sendo que quanto maior for o número de observações maior a probabilidade de assertividade no modelo estatístico.

De maneira geral, uma série histórica é uma variação aleatória imprevisível ao longo do tempo, ou seja, um processo estocástico.

2.8 Análise Estatística de Séries Temporais

2.8.1 Gráficos

Demonstrar o gráfico da série observada é o primeiro passo para realizar a modelagem dos dados. Pelo gráfico, é possível identificar intuitivamente comportamentos relativos a tendências, variabilidade dos dados, pontos anômalos e a possibilidade de erros.

2.8.2 Processos Estocásticos

Um conceito menos intuitivo, é o processo estocástico. Monteiro (2006) define processo estocástico como um fenômeno que evolui ao longo do tempo, de acordo com as leis probabilísticas. Segundo esse conceito, uma série histórica é descrita por uma variável aleatória (X_t) e sua distribuição de probabilidade.

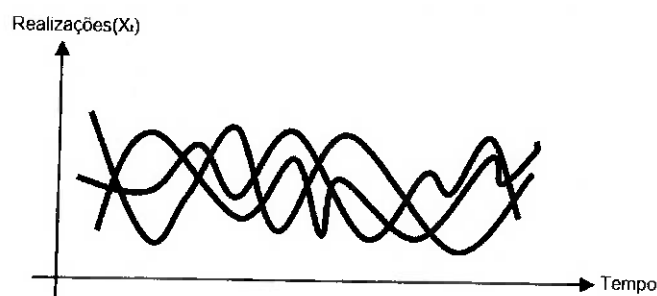


Figura 4 - Realizações de um processo estocástico.

Fonte: Autor.

Os processos estocásticos podem ser estacionários ou não estacionários. No sentido restrito, ele é estacionário quando a distribuição de probabilidade de uma realização for a mesma para qualquer instante do tempo, e não estacionário quando existe a variação da distribuição de probabilidade.

Matematicamente:

$Prob(X_t) = Prob(X_{t+k})$, para qualquer k é a condição restrita de estacionariedade.

2.8.3 Estacionariedade

Rossi (2014) cita que a condição restrita de estacionariedade é difícil de ser verificada e, normalmente se trabalha com condições mais simples, relativas aos

momentos da distribuição. A média e a variância são condições relevantes ao primeiro e segundo momentos do processo estocástico.

$$\text{Média: } \mu_x(t) = E[X_t]$$

$$\text{Variância: } \sigma_x^2(t) = E[X_t - \mu_x(t)]^2$$

em que $E[\]$ é o operador de valor esperado.

Diz-se que um processo estocástico é fracamente estacionário ou de segunda ordem se tanto o primeiro como o segundo momentos forem independentes de t . Tem-se:

Estacionariedade na média (condição 1)

$$E(X_t) = E(X_{t+k})$$

Estacionariedade na variância (condição 2)

$$\sigma_{x_t}^2 = \sigma_{x_{t+k}}^2$$

A estacionariedade de segunda ordem requer que ambas as condições se verifiquem. Quando a distribuição de probabilidade da variável X_t é normal, as condições acima definem um processo estritamente estacionário.

No mercado financeiro é comum as séries temporais não serem estacionárias originalmente, pois o ambiente financeiro é sujeito a efeitos externos da economia. Idealmente, os resíduos de uma série temporal devem seguir a distribuição normal.

2.8.4 Sazonalidade

Segundo Feijó (2008), as séries econômicas tipicamente afetadas por fenômenos sazonais são aquelas relativas à produção, vendas, estoque, despesas de consumos das famílias, despesa de consumo da administração pública, renda, ocupação, exportações e importações de bens e serviços.

Entende-se por eventos sazonais aqueles que se repetem com maior ou menor periodicidade todos os anos. Por exemplo, no fim do ano, com o recebimento do décimo terceiro salário e as festas de fim de ano, é comum que haja um aumento nas vendas no comércio varejista e nas despesas das famílias. Outro exemplo, são os

ajustes de salários e consequentemente o aumento da renda no período pós dissídio das empresas.

2.8.4.1 Filtro de Médias Móveis

É muito comum fazer ajustes sazonais dos dados para utilizá-los no modelo.

O filtro de médias móveis é uma técnica de tratamento de dados para eliminar a sazonalidade de uma série temporal. O método pode ser usado para eliminar flutuações de uma maneira geral. É utilizado quando se quer observar movimentos no tempo de uma variável de uma forma mais assertiva.

Spiegel (1993) afirma que mediante o emprego de médias móveis de ordens apropriadas, podem ser eliminadas as variações cíclicas e irregulares, conservando-se apenas o movimento de tendência.

Segundo Rossi (2014), o método consiste na estimativa de tendência polinomial a um conjunto de observações no tempo. Por definição, tem-se:

$$X_t = T_t + e_t$$

em que X_t é a série histórica e T_t a tendência que se quer determinar pela média móvel.

2.8.5 Ruído Branco

O ruído branco é definido como um processo estocástico estacionário que não se pode explicar. É um sinal sem memória e não correlacionado com o tempo.

Uma série temporal é chamada de ruído branco se for uma sequência de variáveis aleatórias independentes e identicamente distribuídas com média zero e variância finita.

2.8.6 Resíduo

É a diferença do valor original da série em relação ao valor previsto pelo modelo. Dependendo do modelo utilizado, o resíduo pode ser um ruído branco.

2.8.7 Regressão Linear

Segundo Stevenson (1981), o modelo de regressão tem o objetivo de estimar uma equação matemática que visa explicar o comportamento entre uma variável dependente (Y) e uma ou mais variáveis independentes (X). Na regressão pressupõe-se alguma relação de causa e efeito entre as variáveis.

Quando a função f é do tipo $f(X) = a + bX$, tem-se o modelo de regressão linear simples, a e b são valores que se determinam com base nos dados amostrais, a é a cota da reta em $x=0$ e b é o coeficiente angular. A variável X é a variável independente e Y a variável que deve ser predita e é igual a $f(X)$.

Para estimar uma variável com base em um conjunto de variáveis, utiliza-se a regressão múltipla.

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_kX_k$$

onde: Y é a variável dependente;

a : corresponde a um coeficiente fixo, isto é, um valor de base a partir do qual começa Y ;

b_k : são os pesos que o modelo atribui a cada variável explicativa;

X_k : são as variáveis independentes.

A obtenção dos pesos das variáveis independentes é realizada normalmente através do método dos mínimos quadrados.

2.8.8 Análise de Componentes Principais (PCA)

Barroso (2003) conceitua a análise de componentes principais (PCA) como sendo uma técnica estatística multivariada, que transforma linearmente um conjunto de n variáveis em um conjunto com um número menor k de variáveis não correlacionadas, que explica uma parcela substancial das informações do conjunto original. As variáveis n originais (X_1, X_2, \dots, X_n) são transformadas em p variáveis (Y_1, Y_2, \dots, Y_p), denominadas componentes principais, de modo que Y_1 é aquela que explica a maior parte da variação dos dados, Y_2 explica a segunda maior parte e assim por diante.

Artes (2003) descreve que os principais objetivos da análise de componentes principais são: reduzir a dimensionalidade dos dados, gerar novas variáveis que consigam expressar de modo satisfatório a informação contida no conjunto original dos dados, eliminar, quando for possível, algumas variáveis originais, caso elas contribuam com pouca informação.

A análise de componentes principais depende somente da matriz de covariância (σ) e da matriz de correlação (ρ) de (X_1, X_2, \dots, X_t) .

Neste trabalho, escolheu-se essa técnica pela possibilidade de obter uma nova série temporal com as características das variáveis independentes estudadas.

Nesta seção, não será demonstrado o rigor matemático da técnica devido ao fato dos *softwares* empregados fornecerem as informações necessárias para o uso nos modelos.

Em resumo, seja x o vetor das n variáveis originais $x^T = (X_1, X_2, \dots, X_n)$, com $\text{Cov}(x) = \sigma$. Considere p combinações lineares de X_1, X_2, \dots, X_t .

$$Y_1 = l_1^T x = l_{11}X_1 + l_{12}X_2 + \dots + l_{1n}X_n$$

$$Y_2 = l_2^T x = l_{21}X_1 + l_{22}X_2 + \dots + l_{2n}X_n$$

$$\vdots$$

$$Y_n = l_n^T x = l_{n1}X_1 + l_{n2}X_2 + \dots + l_{nn}X_n$$

Então:

$$\text{Var}(Y_i) = l_i^T \text{Var}(x) l_i = l_i^T \sigma l_i$$

$$\text{Cov}(Y_i, Y_j) = \text{Cov}(l_i^T x, l_j^T x) = l_i^T \sigma l_j$$

As componentes principais são combinações lineares Y_1, \dots, Y_n não correlacionadas cujas variâncias são as maiores possíveis.

A primeira componente principal é a combinação linear $l_1^T x$ que maximiza $\text{Var}(l_1^T x)$ sujeita a restrição $l_1^T l_1 = 1$. Se essa restrição não é imposta, pode-se fazer a variância de Y_1 tão grande quanto se queira.

A segunda componente principal é a combinação linear $l_2^T x$ que maximiza $Var(l_2^T x)$ sujeita a restrição $l_2^T l_2 = 1$ e $Cov(l_1^T x, l_2^T x) = 0$.

A n -ésima componente principal é a combinação linear $l_n^T x$ que maximiza $Var(l_n^T x)$ sujeita a restrição $l_n^T l_n = 1$ e $Cov(l_i^T x, l_j^T x) = 0$ para qualquer $j < n$.

Estas restrições garantem que a soma das variâncias das variáveis originais seja igual a soma das variâncias das componentes principais e que estas sejam não correlacionadas.

Seja σ a matriz de covariância associada ao vetor de variáveis aleatórias x . Sejam $(\lambda_1, \alpha_1), \dots, (\lambda_n, \alpha_n)$ os autovalores e autovetores ortogonais padronizados, associados a σ , ordenados de modo que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. A n -ésima componente principal é dada por:

$$Y_n = \alpha_n^T x = \alpha_{n1}X_1 + \alpha_{n2}X_2 + \dots + \alpha_{np}X_p, i = 1, 2, 3, \dots, p$$

com esta escolha,

$$Var(Y_i) = \alpha_i^T \sigma \alpha_i = \lambda_i, i = 1, 2, 3, \dots, p$$

$$Cov(Y_i, Y_j) = \alpha_i^T \sigma \alpha_j = 0, i \neq j$$

O autovalor representa o percentual da quantidade de variância total que está associado ao componente. O autovetor associado ao autovalor calculado, corresponde à correlação entre as componentes principais e as variáveis, e a variância de cada elemento individual do autovetor.

A soma dos autovalores fornece a variância total que corresponde ao número de variáveis consideradas.

BARROSO (2003) descreve que primeiro autovalor corresponde ao maior percentual variabilidade máxima. O segundo autovalor corresponde ao segundo maior percentual e assim por diante. Uma vez calculados dos autovalores e autovetores pode-se calcular as componentes principais.

Na seção 3.3.5 é apresentada a análise de componentes principais gerada pelo *software* Gretl para o uso neste trabalho.

2.9 Previsões em séries temporais

2.9.1 Modelo Autoregressivo (AR)

Os modelos AR foram criados com a ideia que a presente observação da série Y_t , pode ser explicada como uma função das observações passadas $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$, onde p determina o número de passos entre as observações passadas e a previsão da próxima observação.

$$Y_t = c + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} + a_t$$

Um modelo autoregressivo é simplesmente uma regressão linear do valor da série sobre um ou mais dos valores anteriores da série. Por isso, estes modelos podem ser criados usando a técnica dos mínimos quadrados e têm uma interpretação fácil.

2.9.2 Modelo de Média Móvel (MA)

O termo média móvel provém do fato que Y_t seja construído a partir de uma soma ponderada dos valores mais recentes de a_t (ruído branco).

A estrutura de médias moveis é expressa por:

$$Y_t = c + \phi_0 a_t + \phi_1 a_{t-1} + \dots + \phi_p a_{t-p}$$

onde:

c : é uma constante;

ϕ_i : são parâmetros da estrutura, $i = 1, \dots, p$ (a ordem da estrutura);

a_t : é o ruído branco .

O modelo MA possui pouca memória, pois suas previsões se aproximam rapidamente da média da série temporal.

Os modelos autoregressivo $AR(p)$ e média móvel $MA(q)$ podem ser combinados, gerando-se um modelo misto ARMA (p, q) .

2.9.3 Modelo ARMA

Este modelo apresenta simultaneamente, tanto uma estrutura autoregressiva como uma estrutura de média móvel.

$$Y_t = c + \sum_{i=1}^p \theta_i Y_{t-i} + \sum_{i=1}^q \phi_i a_{t-i} + a_t$$

onde:

c : é uma constante;

$\sum_{i=1}^p \theta_i Y_{t-i}$: são as observações anteriores (AR);

$\sum_{i=1}^q \phi_i a_{t-i}$: são os erros de previsões anteriores (MA);

a_t : é o ruído branco.

É interessante notar que os modelos AR(p), MA(q) e ARMA (p, q) pressupõe que a série analisada é estacionária (média e variância constantes ao longo do tempo). Caso a série não seja estacionária, uma alternativa é a utilização do modelo ARIMA (p, d, q), onde d corresponde ao número de diferenciações necessárias para que a série se torne estacionária.

2.9.4 Autocorrelação e Autocorrelação Parcial

Algumas medidas de variação, dispersão e dependência são importantes para a definição da autocorrelação e da autocorrelação parcial.

A covariância é um valor usado para descrever a relação linear entre duas variáveis.

A variância e desvio padrão são medidas de dispersão que mostram o quão distante cada valor da série está do valor central (médio).

Conhecendo-se o valor dos desvios padrão e da covariância, pode-se obter a medida de correlação entre as variáveis. A correlação pode variar no intervalo $[-1, 1]$, onde 1 indica uma correlação linear positiva perfeita e -1 indica uma correlação linear negativa perfeita. É raramente observada uma correlação de 1 ou -1 entre variáveis.

Uma correlação de coeficiente zero significa que os dados não têm nenhuma relação linear entre eles.

Por definição, a função de autocorrelação (FAC) mede o quanto uma série histórica é correlacionada consigo mesma na escala $(-1, 1)$ para uma defasagem h .

A função de autocorrelação parcial (FACP) é a correlação entre duas observações distantes da série, eliminando-se a dependência dos termos intermediários.

2.9.5 Identificação da ordem do modelo

A Tabela 1 resume como usar as funções de autocorrelação e de autocorrelação parcial para identificar a ordem do modelo de uma série temporal.

Tabela 1 – Resumo do comportamento das funções de autocorrelação e autocorrelação parcial para cada tipo de modelo.

Modelo	(FAC)	(FACP)
AR(p)	Decai exponencialmente	Truncada na defasagem p
MA(q)	Truncada na defasagem p	Decai exponencialmente
ARMA(p,q)	Decai exponencialmente	Decai exponencialmente

Fonte: Adaptado Autor.

Ao analisar o correlograma da série de inadimplência dessazonalizada mostrada na Figura 5, ele sugere um modelo AR(1) devido à função de autocorrelação decair exponencialmente e a autocorrelação parcial mostrar um pico na defasagem 1.

Esse modelo pode ser utilizado como ponto de partida para as simulações. A experiência do analista em modelos econométricos pode auxiliar na definição do melhor modelo.

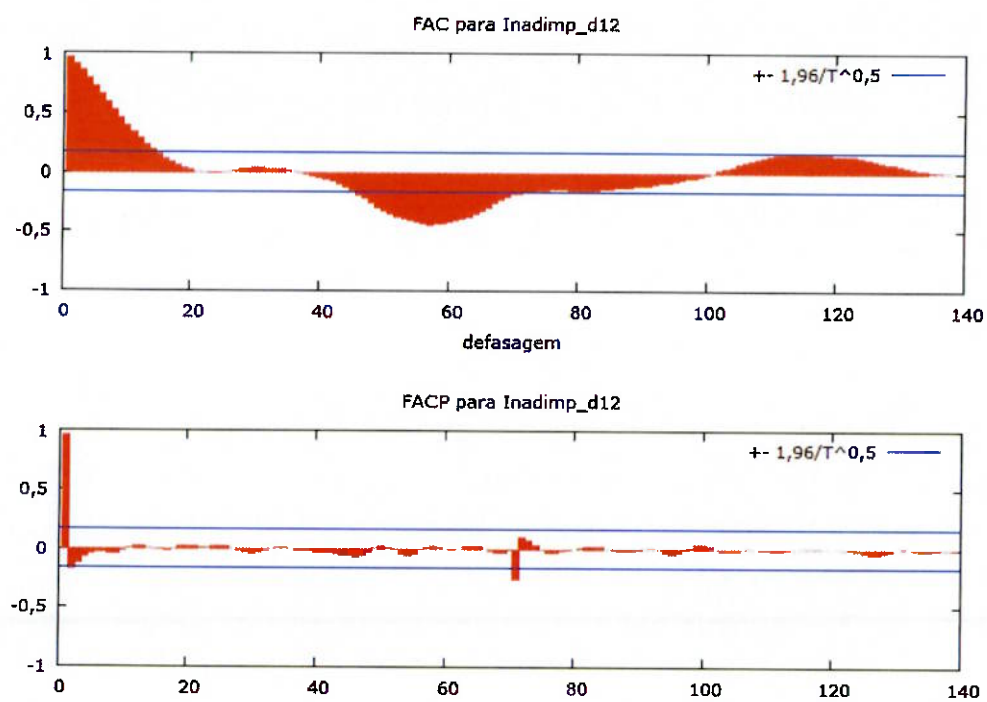


Figura 5 – Correlograma da variável inadimplência dessazonalizada.

Fonte: Autor, gerada através do *software* Gretl 2015d.

3. Metodologia

3.1 Obtenção dos dados

As bases de dados utilizadas neste trabalho são séries temporais dos índices disponibilizados pelos órgãos Ipea, Banco Central do Brasil e IBGE e pela empresa BMF&Bovespa.

O Instituto de Pesquisa Econômica Aplicada (Ipea) é uma fundação pública federal vinculada ao Ministério do Planejamento, Orçamento e Gestão. Suas atividades de pesquisa fornecem suporte técnico e institucional às ações governamentais para a formulação e reformulação de políticas públicas e programas de desenvolvimento brasileiros.

O Banco Central do Brasil é uma autarquia federal, vinculada ao Ministério da Fazenda, que tem por missão assegurar a estabilidade do poder de compra da moeda e um sistema financeiro sólido e eficiente. Entre as suas atividades principais destacam-se: a condução das políticas monetária, cambial, de crédito e de relações financeiras com o exterior; a regulação e a supervisão do Sistema Financeiro Nacional (SFN) e a administração do sistema de pagamentos e do meio circulante.

O Instituto Brasileiro de Geografia e Estatística - IBGE é o principal provedor de dados e informações do País, que atendem às necessidades dos mais diversos segmentos da sociedade civil, bem como dos órgãos das esferas governamentais federal, estadual e municipal.

A BMF&BOVESPA é uma companhia que administra mercados organizados de títulos, valores mobiliários e contratos derivativos, além de prestar serviços de registro, compensação e liquidação, atuando principalmente como contraparte central garantidora da liquidação financeira das operações realizadas em seus ambientes. A Bolsa oferece ampla gama de produtos e serviços, tais como: negociação de ações, títulos de renda fixa, câmbio pronto e contratos derivativos referenciados em ações, ativos financeiros, índices, taxas, mercadorias, moedas, entre outros; listagem de empresas e outros emissores de valores mobiliários; depositária de ativos; empréstimo de títulos e licença de *softwares*.

3.2 Dados utilizados para as análises

3.3.1 Juros Reais

A série histórica de juros reais mostrada na Figura 6 contém 140 observações e foram obtidas a partir da taxa referencial do Swap de DI contra a taxa Pré de 360 dias ajustada pela expectativa da inflação (IPCA) para os próximos 12 meses. A taxa reflete melhor a percepção dos juros reais que os agentes econômicos têm para o futuro, já que o rendimento futuro está diretamente ligado à perspectiva da inflação e não da inflação passada.

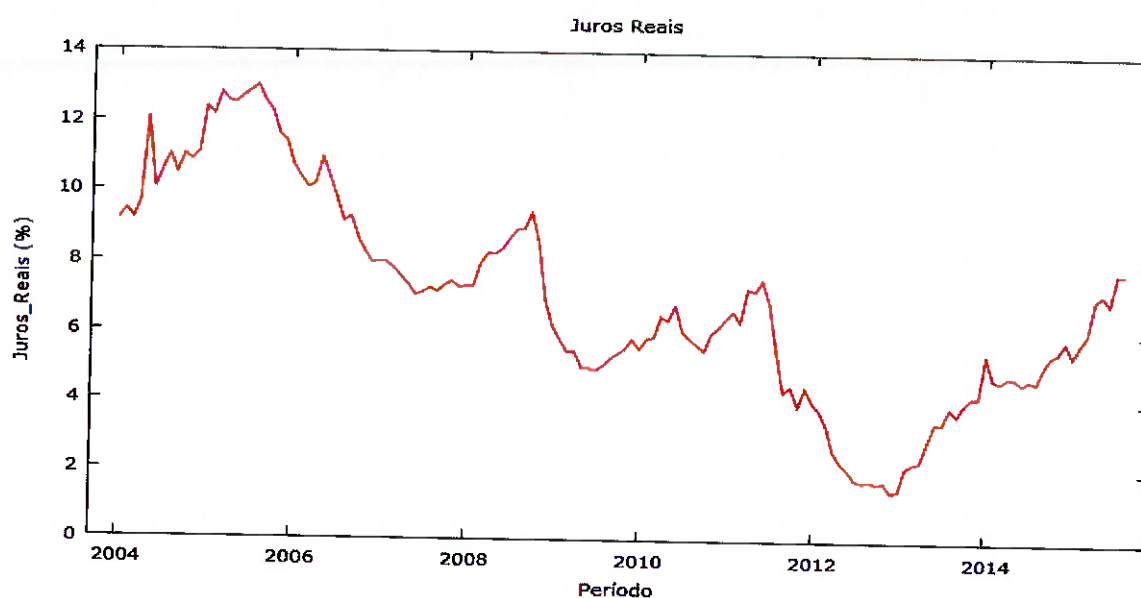


Figura 6 - Série de Juros Reais no período entre Janeiro/2004 a Agosto/2015.

Fonte: Autor, gerada através do *software* Gretl 2015d.

3.3.2 IPCA

O IPCA torna-se relevante na análise de impacto dos juros na inadimplência. Será utilizada a série histórica do período entre janeiro de 2004 e agosto de 2015, conforme mostrado na Figura 7, que é disponibilizada no *site* do IBGE.

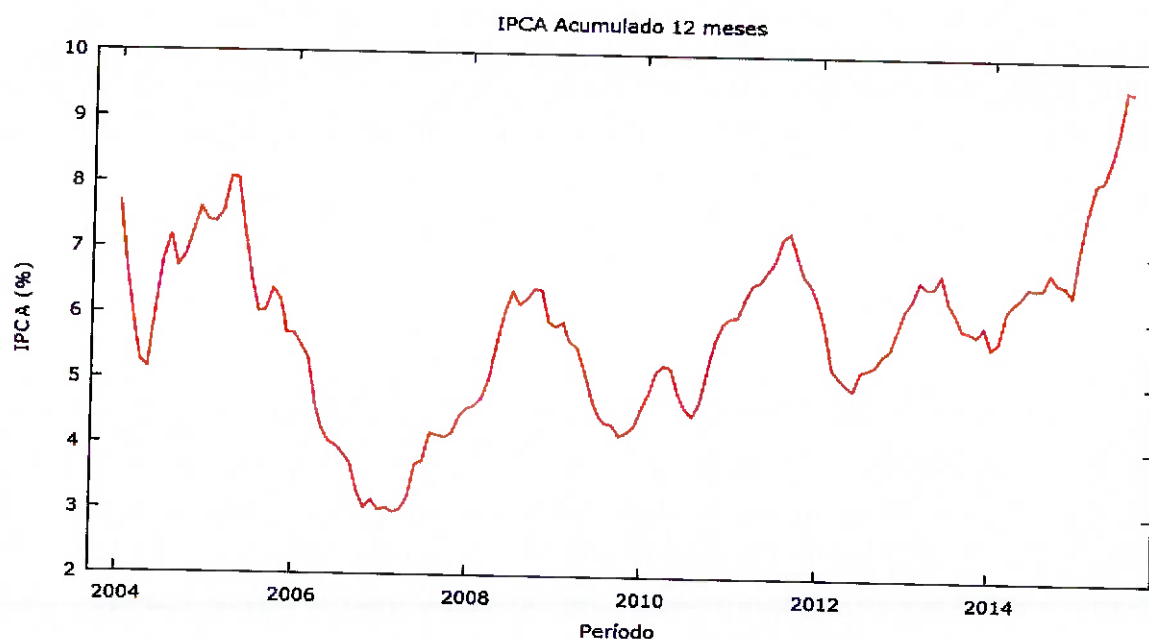


Figura 7 - Série do IPCA acumulado 12 meses no período entre Janeiro/2004 a Agosto/2015.

Fonte: Autor, gerada através do *software* Gretl 2015d.

3.3.3 Desemprego

A série histórica de desemprego contém 140 observações com periodicidade mensal. Foram obtidas no *site* de séries temporais do banco central (SGS), com referência ao IBGE. Os valores estão relacionados ao desemprego aberto, que é a relação entre o número de pessoas desocupadas que já trabalharam no mínimo uma vez e o número de pessoas economicamente ativas.

Durante a análise gráfica da série, foram observados alguns movimentos repetitivos na série que podem prejudicar a construção do modelo. Um deles foi uma queda no índice nos meses de novembro e dezembro, que provavelmente ocorre pela oferta de vagas temporárias para o Natal. Foi necessário um tratamento nos dados retirando a sazonalidade da série através do pacote de análise X12-ARIMA do *software* Gretl. As Figuras 8 e 9 apresentam respectivamente a série original dos dados (desemprego) e um comparativo entre a série original e a série dessazonalizada (desemprego_12).

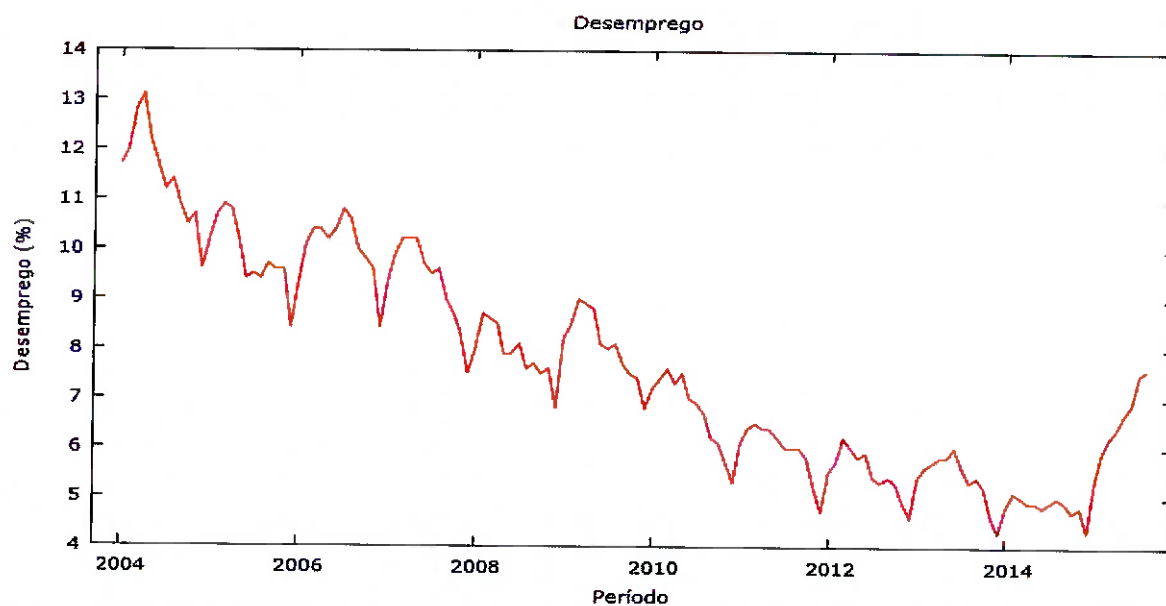


Figura 8 - Taxa de desemprego original no período entre Janeiro/2004 a Agosto/2015.

Fonte: Autor, gerada através do software Gretl 2015d.

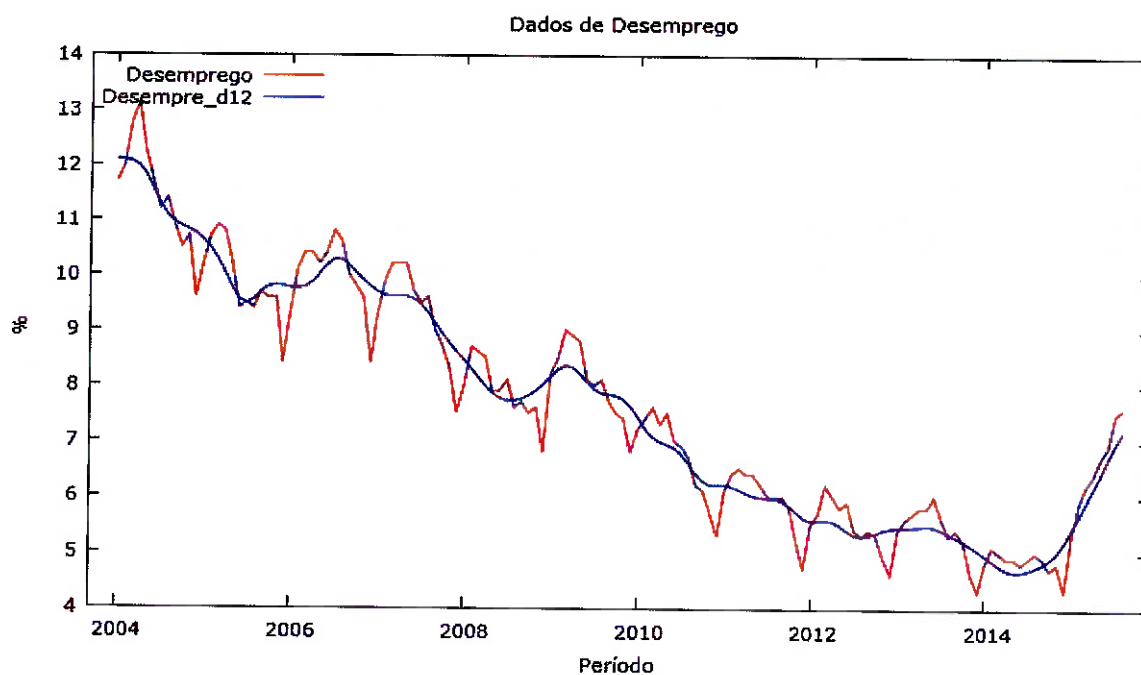


Figura 9 - Série de desemprego original e série dessazonalizada no período entre Janeiro/2004 a Agosto/2015.

Fonte: Autor, gerada através do software Gretl 2015d.

3.3.4 Inadimplência

A série histórica de inadimplência contém 140 observações e foram obtidas através do indicador inadimplência em t-3 do IpeaData. Percebe-se na Figura 10 uma sazonalidade nos dados. Há uma queda na inadimplência nos meses finais do ano, onde possivelmente com o recebimento do décimo terceiro salário as pessoas quitam suas dívidas, mas no início do ano, existe um aumento no índice, provavelmente devido aos gastos extras de início de ano e reflexo das compras de Natal e por fim no período entre junho e agosto, percebe-se novamente uma queda no índice. Nesse período, uma grande quantidade de empresas ajusta os salários dos funcionários, conforme o dissídio dos sindicatos de cada classe e as pessoas solicitam férias e recebem uma renda extra, que pode ser utilizada para quitar dívidas.

As justificativas citadas acima, são apenas prováveis motivos da sazonalidade da série. Foi necessário um tratamento para retirar a sazonalidade nos dados da série através do pacote de análise X12-ARIMA do *software* Gretl, cujo resultado é indicado na Figura 11.

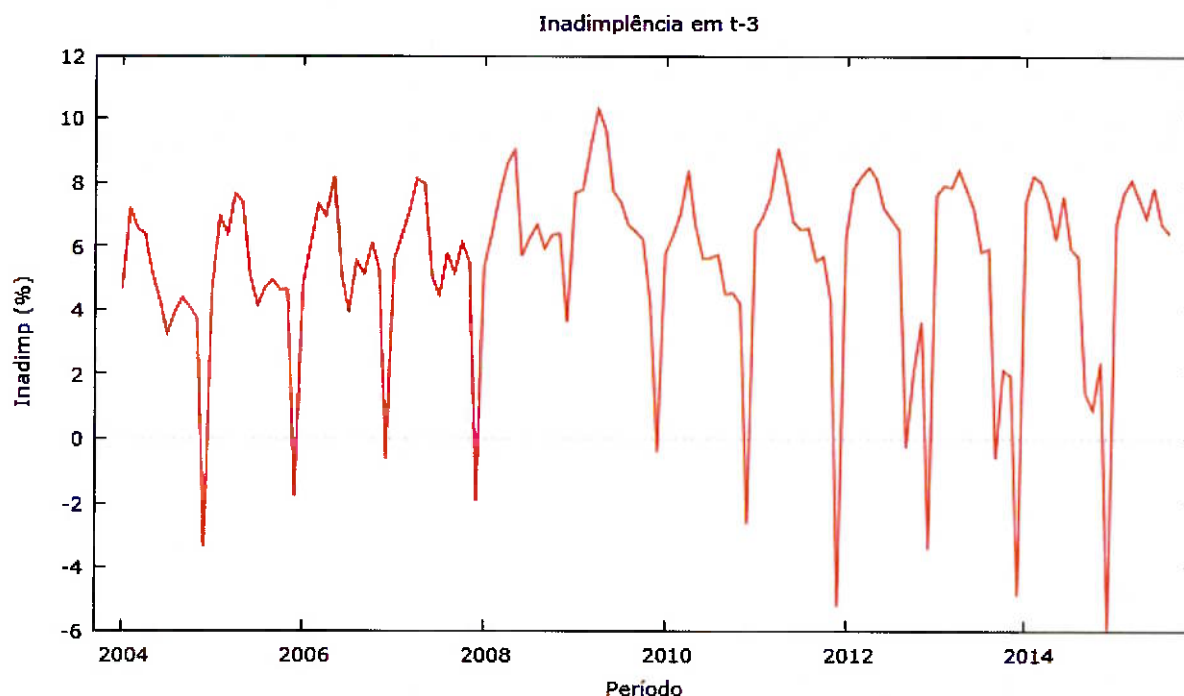


Figura 10 - Série de inadimplência original no período entre Janeiro/2004 a Agosto/2015.

Fonte: Autor, gerada através do *software* Gretl 2015d.

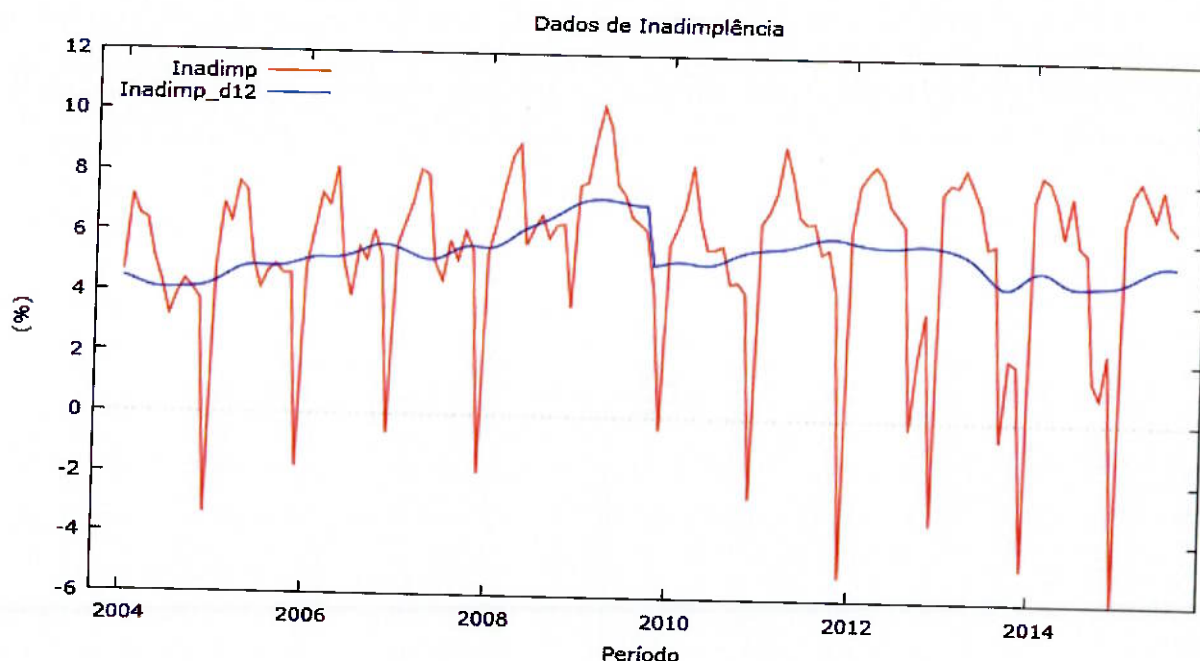


Figura 11 - Série de inadimplência original e série dessazonalizada no período entre Janeiro/2004 a Agosto/2015.

Fonte: Autor, gerada através do software Gretl 2015d.

Mesmo após dessazonalizar a série de inadimplência, observa-se uma irregularidade no mês de novembro de 2009.

Adotou-se o uso de uma variável *dummy* no conjunto de dados, como uma forma de introduzir uma característica qualitativa no modelo econométrico.

A Figura 12 apresenta em um único gráfico, as quatro séries e a variável *dummy* que serão utilizadas na construção e análise do melhor modelo.

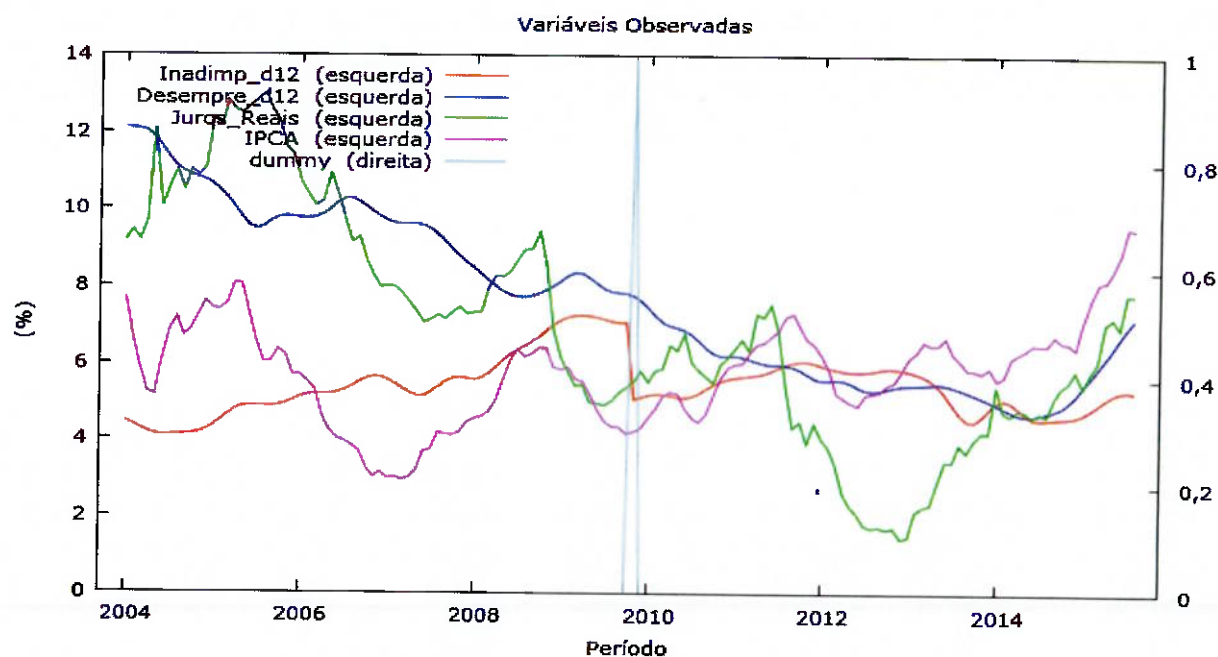


Figura 12 - Variáveis observadas no período entre Janeiro/2004 a Agosto/2015.

Fonte: Autor, gerada através do software Gretl 2015d

3.3.5 Componentes principais

A Figura 13 apresenta a análise de componentes principais gerada pelo software Gretl com as variáveis desemprego, juros reais e IPCA. A componente 1 representa 60,95% das características dos dados e a componente 2 representa 34,94%. Nas simulações, serão utilizadas as componentes 1 e 2, pois juntas representam 95,89% dos dados originais. A componente 3 será desconsiderada. Foi realizada a mesma análise no software Minitab17 e os resultados foram os mesmos.

Análise de Componentes Principais			
Análise de Autovalores da Matriz de Correlação			
Componente	Autovalor	Proporção	Acumulada
1	1,8283	0,6094	0,6094
2	1,0481	0,3494	0,9588
3	0,1236	0,0412	1,0000
Autovetores (cargas dos componentes)			
	PC1	PC2	PC3
Desemprego_d12	-0,713	0,103	0,694
IPCA	0,100	-0,964	0,246
Juros_Reais	-0,694	-0,245	-0,677

Figura 13 - Análise de Componentes Principais, gerada através do software Gretl 2015d.

Fonte: Autor.

3.3 Softwares empregados nas análises

Foram usados os *softwares* Gretl (GNU Regression and Econometric Tool Library com X-12 ARIMA), Microsoft Excel e Minitab 17.

Constatou-se uma facilidade no uso do software Geral em relação a outros softwares como o Matlab por exemplo, para construção dos modelos, devido sua interface gráfica amigável e as funções pré-programadas do software. O software é de código aberto e está disponível para download na internet.

Entretanto, como as funções do software estão disponíveis para uso, apenas é necessário entrar com os dados no sistema. Com isso algumas análises de parâmetros estimados e equações podem se tornar complexas no entendimento. A ferramenta deve ser utilizada como apoio nas análises de séries temporais.

3.4 Critérios para a escolha das ordens do modelo

As ordens p e q de um modelo ARMA são desconhecidas e devem ser obtidas empiricamente. Somente o uso das funções de autocorrelação e autocorrelação parcial não é suficiente para determinar as ordens dos modelos. Os critérios de informação traduzem a qualidade de ajuste de um modelo e devem ser utilizados para a escolha do modelo.

Uma medida estatística útil para a análise do modelo é a SQR (soma dos quadrados dos resíduos). Ela representa a parte que não é explicada pelo modelo. O menor valor representa o melhor ajuste do modelo. A SQR é definida como:

$$SQR = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

onde Y_i : valor original da série, \hat{Y}_i : valor previsto pelo modelo e n o número de observações.

O log da função de verossimilhança pode ser utilizado para comparar o ajuste de diferentes coeficientes do modelo. Com o objetivo de maximizar o log, um valor mais alto representa um modelo com melhor ajuste.

O critério de Informação de Akaike (AIC) é definido como:

$$AIC_{p=} - 2 \log(L_p) + 2[(p + 1) + 1]$$

onde L_p é a função de máxima verossimilhança do modelo e p é o número de variáveis explicativas consideradas no modelo.

O critério de Informação de Schwarz (SIC ou BIC) é definido como:

$$SIC_{p=} - 2 \log(L_p) + [(p + 1) + 1] \log(n)$$

onde L_p é a função de máxima verossimilhança do modelo, p é o número de variáveis explicativas e n o número de observações consideradas no modelo.

Os dois critérios aumentam conforme a soma dos quadrados dos resíduos aumenta. Ambos penalizam modelos com muitas variáveis e são preferíveis valores menores de AIC e SIC.

4. Simulações

4.1 Modelo ARIMA (1,0,0) com 3 variáveis

Na simulação, o uso das variáveis independentes como regressoras, tem o objetivo de verificar sua relevância nos dados originais. Conforme mostrado nas Figuras 14 e 15, observou-se que o modelo converge, o valor da medida SQR é baixa e igual 3,46, mas devido aos valores do *p-valor* das variáveis explicativas serem altos, neste modelo, estas variáveis não contribuem com a explicação. Assim, se essas variáveis forem retiradas do modelo, o resultado das previsões não será modificado.

Modelo 1: ARMAX, usando as observações 2004:01-2015:08 (T = 140)
Estimado usando o filtro de Kalman (Máxima verossimilhança exata)
Variável dependente: Inadimp_d12
Erros padrão baseados na Hessiana

	coeficiente	erro padrão	z	p-valor	
const	4,85403	1,06596	4,554	5,27e-06	***
phi_1	0,981789	0,0140505	69,88	0,0000	***
Desemprego_d12	0,0172978	0,0955235	0,1811	0,8563	
Juros_Reais	-0,0263934	0,0234714	-1,124	0,2608	
IPCA	0,0546757	0,0377155	1,450	0,1471	
dummy	-1,02187	0,102688	-9,951	2,49e-023	***
Média var. dependente	5,408136	D.P. var. dependente	0,768289		
Média de inovações	0,005368	D.P. das inovações	0,143839		
Log da verossimilhança	71,15628	Critério de Akaike	-128,3126		
Critério de Schwarz	-107,7211	Critério Hannan-Quinn	-119,9448		

Figura 14 - Parâmetros estimados do Modelo ARIMA (1,0,0).

Fonte: Autor, gerada através do software Gretl 2015d

Conforme parâmetros estimados pelo software Gretl, temos a equação do modelo estimado ARIMA (1,0,0).

$$\widehat{InadPrev}_t = 4,85403 + 0,981789Inadimp_d12_{t-1} + 0,0172978Desemprego_d12_{t-1} \\ - 0,0263934JurosReais_{t-1} + 0,0546757IPCA_{t-1} - 1,02187dummy_{t-1} \\ + at$$

onde: *at* é um erro associado ao modelo.

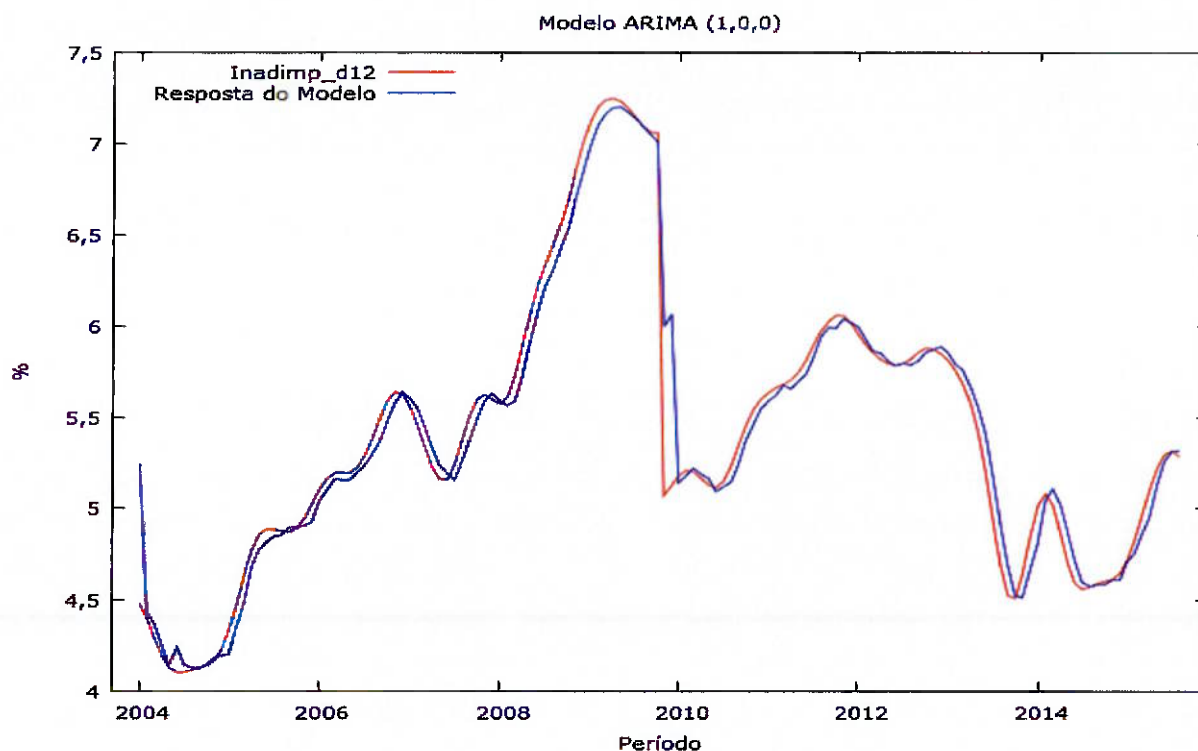


Figura 15 - Comparação do Modelo ARIMA (1,0,0) com a série temporal de inadimplência.

Fonte: Autor, gerada através do software Gretl 2015d.

4.2 Modelo ARIMA (0,0,1) com 3 variáveis

Similar ao modelo anterior, são incluídas no modelo as variáveis independentes como regressoras. Nas Figuras 16 e 17 é possível verificar que o modelo converge, o valor da medida SQR é superior ao modelo ARIMA(1,0,0) e igual a 18,27, o que impacta na auto-validação do modelo. As variáveis Juros Reais e IPCA são relevantes devido ao *p-valor* ser menor que 5%, mas o desemprego neste modelo não é significativo.

Modelo 2: ARMAX, usando as observações 2004:01-2015:08 (T = 140)
 Estimado usando X-12-ARIMA (Máxima verossimilhança exata)
 Variável dependente: Inadimp_d12

	coeficiente	erro padrão	z	p-valor	
const	6,83906	0,393232	17,39	9,50e-068	***
theta_1	0,999979	0,0396035	25,25	1,14e-140	***
Desempre_d12	-0,0329562	0,0437536	-0,7532	0,4513	
Juros_Reais	-0,0570728	0,0273288	-2,088	0,0368	**
IPCA	-0,135002	0,0451271	-2,992	0,0028	***
dummy	-1,12465	0,0816435	-13,78	3,60e-043	***
Média var. dependente	5,408136	D.P. var. dependente		0,768289	
Média de inovações	0,000665	D.P. das inovações		0,354599	
Log da verossimilhança	-55,97681	Critério de Akaike		125,9536	
Critério de Schwarz	146,5451	Critério Hannan-Quinn		134,3214	

Figura 16 - Parâmetros estimados do Modelo ARIMA (0,0,1).

Fonte: Autor, gerada através do software Gretl 2015d

Conforme parâmetros estimados pelo software Gretl, a equação do modelo estimado ARIMA (0,0,1) pode ser representada por:

$$\widehat{InadPrev}_t = 6,83906 + a_t + 0,999979a_{t-1} - 0,0329562Desempre_{d12t-1} \\ - 0,0570728JurosReais_{t-1} - 0,135002IPCA_{t-1} - 1,12465dummy_{t-1}$$

onde: a_t é um erro associado ao modelo.

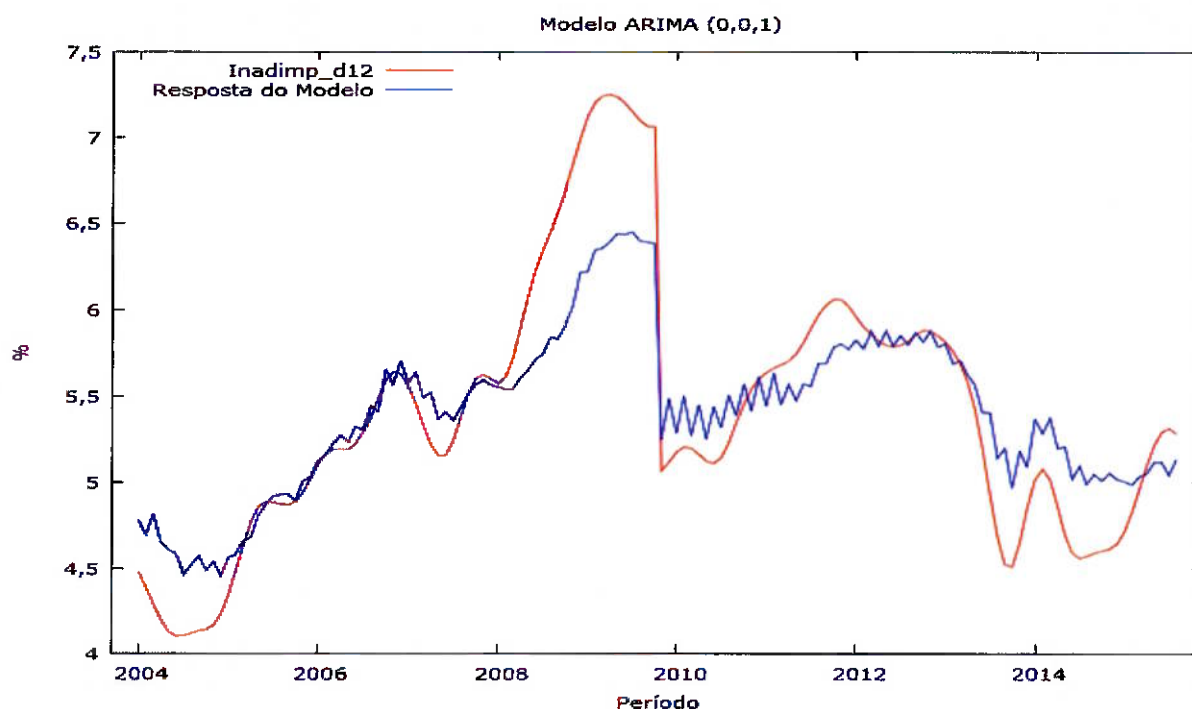


Figura 17 - Comparação do Modelo ARIMA (0,0,1) com a série temporal de inadimplência.

Fonte: Autor, gerada através do software Gretl 2015d.

4.3 Modelo ARIMA (0,0,2) com 3 variáveis

Para o modelo ARIMA (0,0,2) com as variáveis regressoras, observou-se uma melhor aderência dos dados ao modelo, se comparado com o modelo ARIMA (0,0,1). O desemprego continua não sendo relevante no modelo, mas os critérios de AIC e Schwarz são os menores obtidos nas simulações realizadas. O valor de SQR é igual a 7,16, que representa uma queda nos erros de previsão. As Figuras 18 e 19 demonstram os parâmetros do modelo e o gráfico da auto-validação do modelo.

Modelo 3: ARMAX, usando as observações 2004:01-2015:08 (T = 140)
 Estimado usando X-12-ARIMA (Máxima verossimilhança exata)
 Variável dependente: Inadimp_d12

	coeficiente	erro padrão	z	p-valor	
const	6,42999	0,334214	19,24	1,74e-082	***
theta_1	1,73309	0,0447668	38,71	0,0000	***
theta_2	0,860128	0,0449413	19,14	1,20e-081	***
Desemprego_d12	-0,0446054	0,0351203	-1,270	0,2041	
Juros_Reais	-0,0397496	0,0142140	-2,797	0,0052	***
IPCA	-0,0697316	0,0336834	-2,070	0,0384	**
dummy	-0,996821	0,0415635	-23,98	4,17e-127	***
Média var. dependente	5,408136	D.P. var. dependente		0,768289	
Média de inovações	0,001509	D.P. das inovações		0,219024	
Log da verossimilhança	11,58992	Critério de Akaike		-7,179850	
Critério de Schwarz	16,35329	Critério Hannan-Quinn		2,383314	

Figura 18 - Parâmetros estimados do Modelo ARIMA (0,0,2).

Fonte: Autor, gerada através do software Gretl 2015d.

Conforme parâmetros estimados pelo software Gretl, temos a equação do modelo estimado ARIMA (0,0,2).

$$\begin{aligned}
 \widehat{InadPrev}_t = & 6,4299 + a_t + 1,73309a_{t-1} + 0,860128a_{t-2} \\
 & - 0,0446054(\text{Desemprego}_{d12,t-2} - \text{Desemprego}_{d12,t-1}) \\
 & - 0,0397496(\text{JurosReais}_{t-2} - \text{JurosReais}_{t-1}) - 0,0697316(\text{IPCA}_{t-2} \\
 & - \text{IPCA}_{t-1}) - 0,996821(\text{dummy}_{t-2} - \text{dummy}_{t-1})
 \end{aligned}$$

onde: a_t é um erro associado ao modelo.

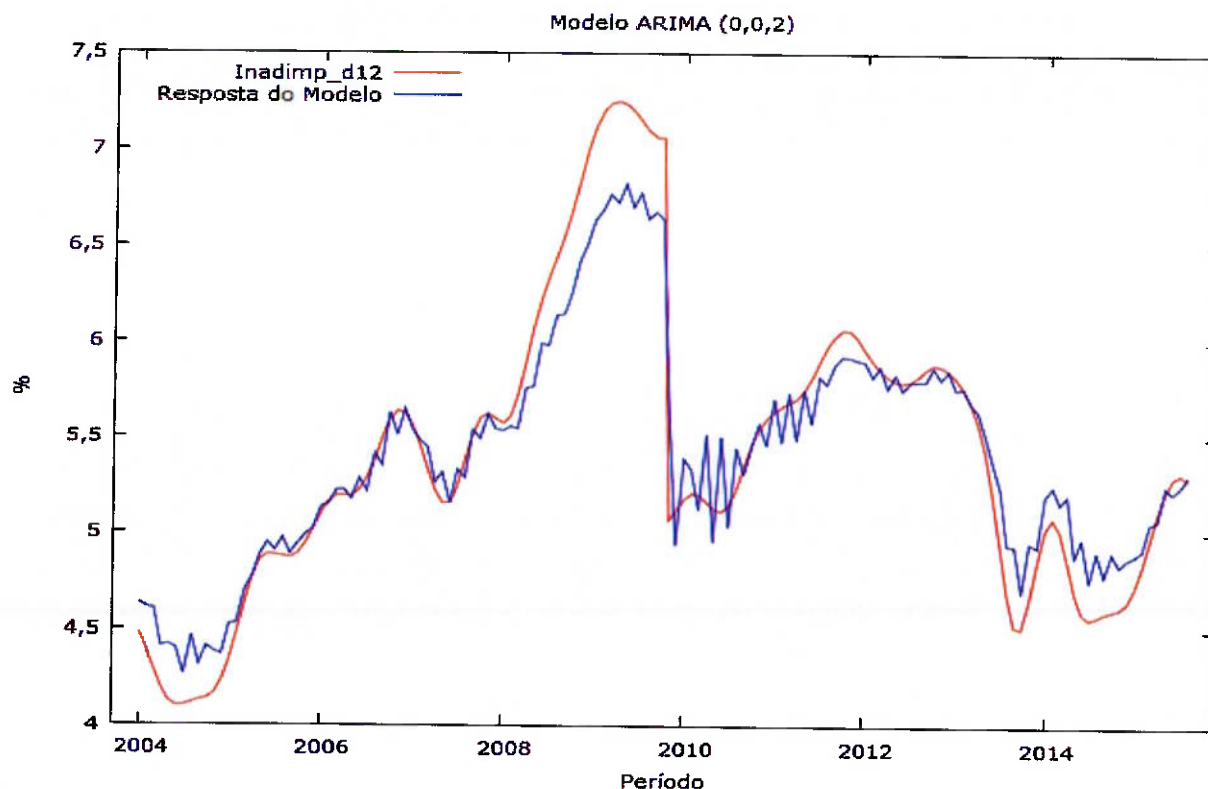


Figura 19 - Comparação do Modelo ARIMA (0,0,2) com a série temporal de inadimplência.

Fonte: Autor, gerada através do software Gretl 2015d.

4.4 Modelo ARIMA (1,0,1) com 3 variáveis

O modelo ARIMA(1,0,1) apresentou melhor auto-validação com os dados originais, mas conforme as Figuras 20 e 21, as variáveis regressoras não impactam nas previsões, devido ao alto valor do *p-valor* de cada variável. O valor de SQR é igual a 1,76, a previsão fica próxima dos valores originais, mas as variáveis não contribuem com a explicação do modelo. Este é o melhor modelo de auto-validação dos dados originais.

Modelo 4: ARMAX, usando as observações 2004:01-2015:08 (T = 140)
 Estimado usando X-12-ARIMA (Máxima verossimilhança exata)
 Variável dependente: Inadimp_d12

	coeficiente	erro padrão	z	p-valor	
const	5,41940	0,946940	5,723	1,05e-08	***
phi_1	0,966221	0,0191837	50,37	0,0000	***
theta_1	0,999999	0,0394969	25,32	2,00e-141	***
Desempre_d12	-0,0177771	0,102348	-0,1737	0,8621	
Juros_Reais	-0,00759122	0,00519661	-1,461	0,1441	
IPCA	0,00720384	0,0182742	0,3942	0,6934	
dummy	-1,02282	0,0134426	-76,09	0,0000	***
Média var. dependente	5,408136	D.P. var. dependente	0,768289		
Média de inovações	0,004564	D.P. das inovações	0,093990		
Log da verossimilhança	127,8855	Critério de Akaike	-239,7710		
Critério de Schwarz	-216,2378	Critério Hannan-Quinn	-230,2078		

Figura 20 - Parâmetros estimados do Modelo ARIMA (1,0,1).

Fonte: Autor, gerada através do software Gretl 2015d.

Conforme parâmetros estimados pelo software Gretl, temos a equação do modelo estimado ARIMA (1,0,1).

$$\widehat{InadPrev}_t = 5,4194 + 0,966221Inadimp_d12_{t-1} + a_t + 0,999999a_{t-1} \\ - 0,0177771Desempre_{d12}_{t-1} - 0,00759122JurosReais_{t-1} \\ + 0,00720384IPCA_{t-1} - 1,02282dummy_{t-1}$$

onde: a_t é um erro associado ao modelo.

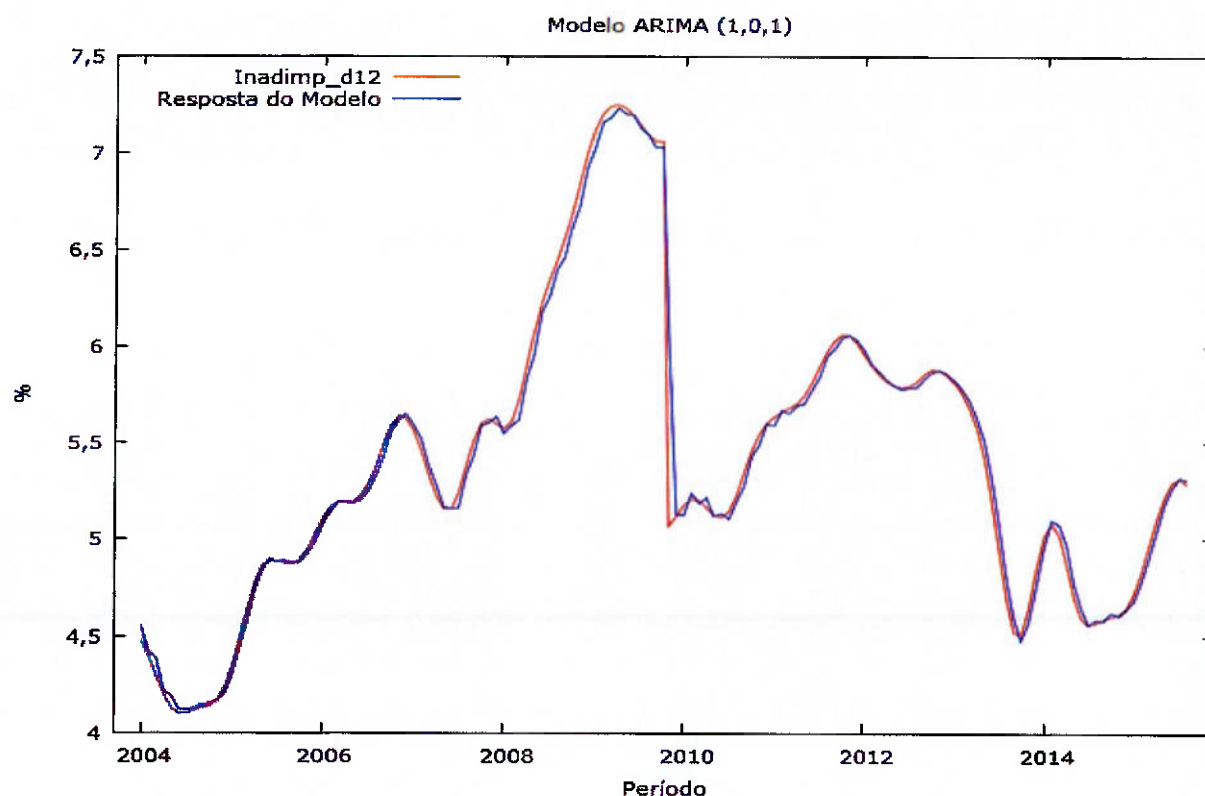


Figura 21 - Comparação do Modelo ARIMA (1,0,1) com a série temporal de inadimplência.

Fonte: Autor, gerada através do *software* Gretl 2015d.

4.5 Modelo ARIMA (1,0,0) com duas variáveis de componentes principais

Nas próximas simulações, as variáveis regressoras serão as duas componentes principais geradas pelo *software* Gretl, que representam juntas 95,88% das características das variáveis.

O modelo converge, o valor de SQR é igual a 3,22, próximo ao valor do mesmo modelo quando são utilizadas as variáveis Juros Reais, desemprego e IPCA. As variáveis independentes, representadas pelas componentes principais não contribuem para a explicação do modelo, conforme mostrado nas Figuras 22 e 23.

Modelo 1: ARMAX, usando as observações 2004:01-2015:08 (T = 140)
 Estimado usando o filtro de Kalman (Máxima verossimilhança exata)
 Variável dependente: Inadimp_d12
 Erros padrão baseados na Hessiana

	coeficiente	erro padrão	z	p-valor	
const	5,18549	0,474950	10,92	9,46e-028 ***	
phi_1	0,979619	0,0140277	69,83	0,0000 ***	
PC1	0,112789	0,0917490	1,229	0,2190	
PC2	0,0601178	0,0516237	1,165	0,2442	
dummy	-1,01998	0,102956	-9,907	3,88e-023 ***	
Média var. dependente	5,408136	D.P. var. dependente	0,768289		
Média de inovações	0,004045	D.P. das inovações	0,144116		
Log da verossimilhança	70,94280	Critério de Akaike	-129,8856		
Critério de Schwarz	-112,2357	Critério Hannan-Quinn	-122,7132		

Figura 22 - Parâmetros estimados do Modelo ARIMA (1,0,0) com PCA.

Fonte: Autor, gerada através do software Gretl 2015d.

Conforme parâmetros estimados pelo software Gretl, temos a equação do modelo estimado ARIMA (1,0,0) com componentes principais.

$$\widehat{InadPrev}_t = 5,18549 + 0,979619Inadimp_d12_{t-1} + 0,112789PC1_{t-1} + 0,061178PC2_{t-1} - 1,01998dummy_{t-1} + at$$

onde: at é um erro associado ao modelo.

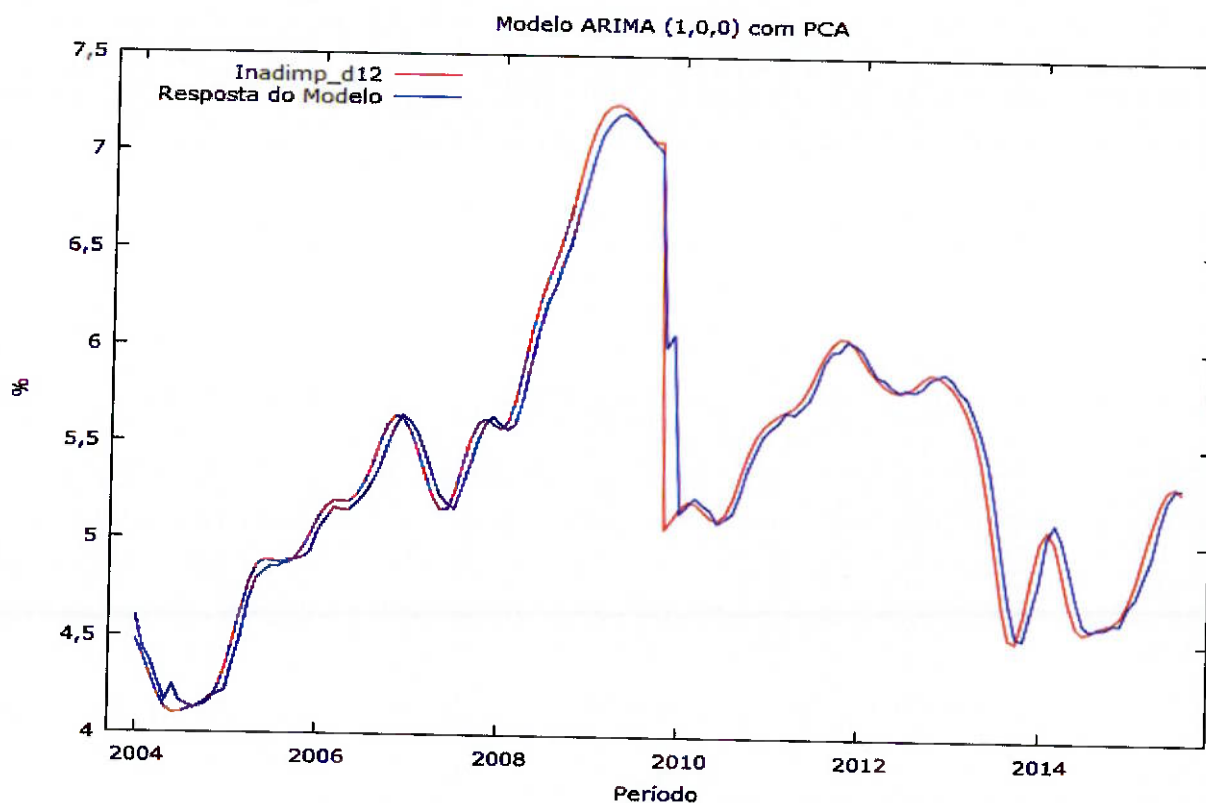


Figura 23 - Comparação do Modelo ARIMA (1,0,0) com PCA em relação a série temporal de inadimplência.

Fonte: Autor, gerada através do software Gretl 2015d.

4.6 Modelo ARIMA (0,0,1) com duas variáveis de componentes principais

O modelo converge e as componentes principais conseguem explicar o valor da inadimplência, devido ao valor do *p-valor* das variáveis ser menor que 5%. O valor de SQR é igual a 18,25, o maior valor encontrado nas simulações. Conforme demonstrado nas Figuras 24 e 25, no período entre dezembro de 2008 a novembro de 2009, o erro do modelo é alto, onde é possível associar a consequência da crise econômica de 2008 a esse erro.

Modelo 2: ARMAX, usando as observações 2004:01-2015:08 (T = 140)
 Estimado usando o filtro de Kalman (Máxima verossimilhança exata)
 Variável dependente: Inadimp_d12
 Erros padrão baseados na Hessiana

	coeficiente	erro padrão	z	p-valor	
const	5,41653	0,0597423	90,66	0,0000	***
theta_1	1,00000	0,0191201	52,30	0,0000	***
PC1	0,150073	0,0425426	3,528	0,0004	***
PC2	-0,208363	0,0554227	-3,760	0,0002	***
dummy	-1,11550	0,0638203	-17,48	2,08e-068	***
Média var. dependente	5,408136	D.P. var. dependente		0,768289	
Média de inovações	-0,002726	D.P. das inovações		0,354636	
Log da verossimilhança	-55,99294	Critério de Akaike		123,9859	
Critério de Schwarz	141,6357	Critério Hannan-Quinn		131,1583	

Figura 24 - Parâmetros estimados do Modelo ARIMA (0,0,1) com PCA.

Fonte: Autor, gerada através do software Gretl 2015d.

Conforme parâmetros estimados pelo software Gretl, a equação do modelo estimado ARIMA (0,0,1) com componentes principais pode ser representada por:

$$\widehat{InadPrev}_t = 5,41653 + a_t + a_{t-1} + 0,150073PC1_{t-1} - 0,208363PC2_{t-1} - 1,1155dummy_{t-1}$$

onde: a_t é um erro associado ao modelo.

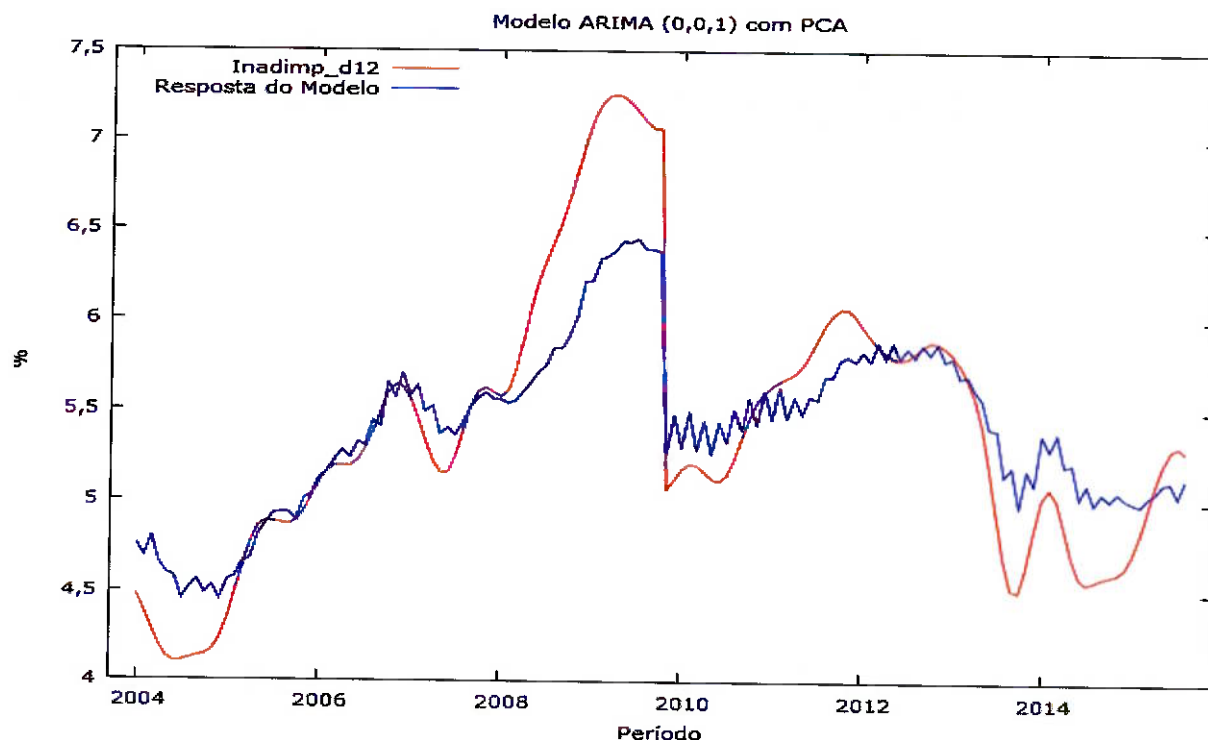


Figura 25 - Comparação do Modelo ARIMA (0,0,1) com PCA em relação a série temporal de inadimplência.

Fonte: Autor, gerada através do software Gretl 2015d.

4.7 Modelo ARIMA (0,0,2) com duas variáveis de componentes principais

O modelo converge e as componentes principais conseguem explicar o valor da inadimplência, devido ao valor do *p*-valor das variáveis ser menor que 5%. O valor de SQR é igual a 7,18, inferior ao encontrado no modelo anterior (ARIMA(0,0,1)). Isso representa um menor erro nas previsões, se comparado com o modelo anterior. Conforme demonstrado nas Figuras 26 e 27, os critérios de AIC e BIC são menores, isso significa que o modelo não foi penalizado pela inclusão das variáveis.

Modelo 3: ARMAX, usando as observações 2004:01-2015:08 (T = 140)
Estimado usando o filtro de Kalman (Máxima verossimilhança exata)
Variável dependente: Inadimp_d12
Erros padrão baseados na Hessiana

	coeficiente	erro padrão	z	p-valor	
const	5,41339	0,0661062	81,89	0,0000	***
theta_1	1,73365	0,0483767	35,84	2,99e-281	***
theta_2	0,860795	0,0389929	22,08	5,41e-108	***
PC1	0,136067	0,0388740	3,500	0,0005	***
PC2	-0,109609	0,0539139	-2,033	0,0420	**
dummy	-0,997248	0,0422960	-23,58	6,51e-123	***
Média var. dependente	5,408136	D.P. var. dependente	0,768289		
Média de inovações	-0,004083	D.P. das inovações	0,219030		
Log da verossimilhança	11,58181	Critério de Akaike	-9,163614		
Critério de Schwarz	11,42788	Critério Hannan-Quinn	-0,795845		

Figura 26 - Parâmetros estimados do Modelo ARIMA (0,0,2) com PCA.

Fonte: Autor, gerada através do software Gretl 2015d.

Conforme parâmetros estimados pelo software Gretl, a equação do modelo estimado ARIMA (0,0,2) com componentes principais é descrita por:

$$\widehat{InadPrev}_t = 5,41339 + a_t + 1,73365a_{t-1} + 0,860795a_{t-2} + 0,136067(PC1_{t-2} - PC1_{t-1}) - 0,109609(PC2_{t-2} - PC2_{t-1}) - 0,997248(dummy_{t-2} - dummy_{t-1})$$

onde: a_t é um erro associado ao modelo.

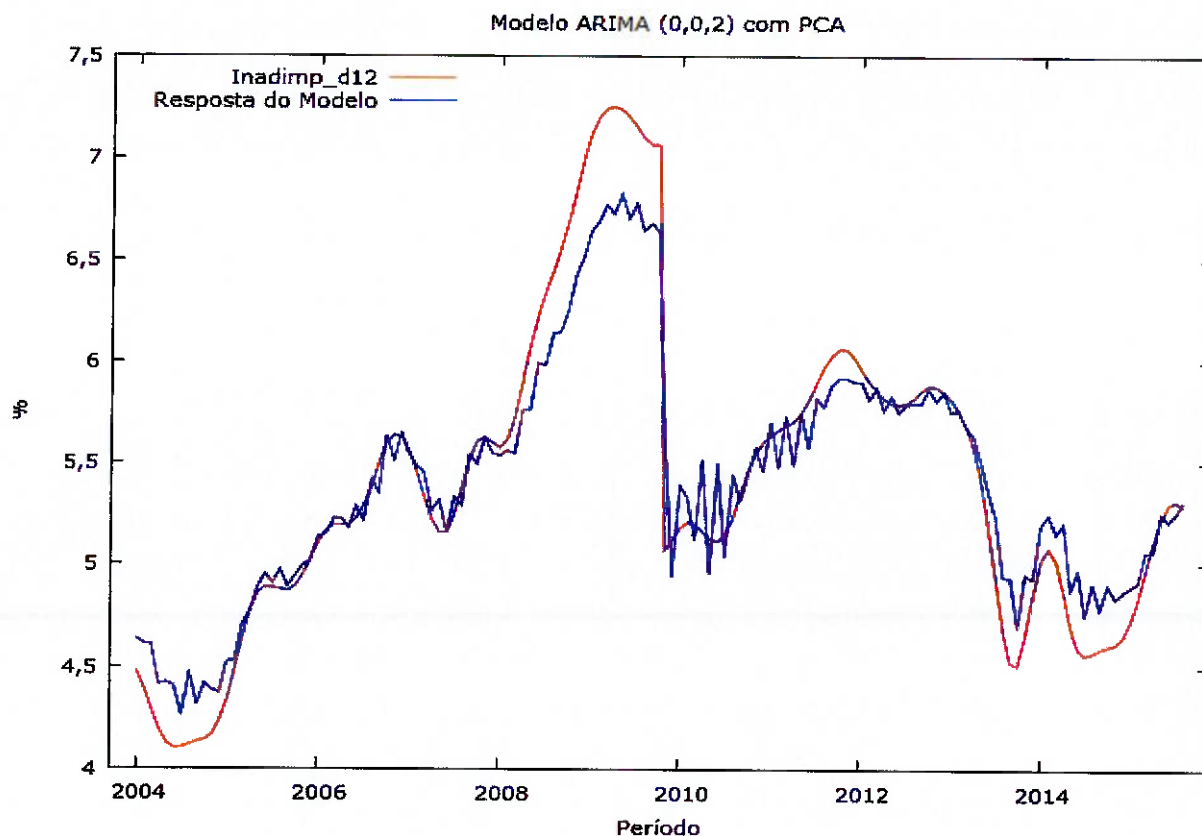


Figura 27 - Comparação do Modelo ARIMA (0,0,2) com PCA em relação a série temporal de inadimplência.

Fonte: Autor, gerada através do software Gretl 2015d.

4.8 Modelo ARIMA (1,0,1) com duas variáveis de componentes principais

O modelo apresenta ótima representação dos dados originais mas, conforme Figuras 28 e 29, as componentes regressoras não impactam nas previsões. O valor de SQR é baixo e igual a 1,79, a previsão fica próxima dos valores originais, mas as variáveis não explicam o modelo. O ARIMA (1,0,1) desta simulação é similar ao da simulação da seção 4.4.

Modelo 4: ARMAX, usando as observações 2004:01-2015:08 (T = 140)
 Estimado usando o filtro de Kalman (Máxima verossimilhança exata)
 Variável dependente: Inadimp_d12
 Erros padrão baseados na Hessiana

	coeficiente	erro padrão	z	p-valor	
const	5,26770	0,404478	13,02	9,00e-039	***
phi_1	0,966643	0,0190365	50,78	0,0000	***
theta_1	1,00000	0,0194605	51,39	0,0000	***
PC1	0,0346527	0,0216825	1,598	0,1100	
PC2	0,00656785	0,0248537	0,2643	0,7916	
dummy	-1,02283	0,0134477	-76,06	0,0000	***
Média var. dependente	5,408136	D.P. var. dependente	0,768289		
Média de inovações	-0,000458	D.P. das inovações	0,093986		
Log da verossimilhança	127,8841	Critério de Akaike	-241,7682		
Critério de Schwarz	-221,1767	Critério Hannan-Quinn	-233,4005		

Figura 28 - Parâmetros estimados do Modelo ARIMA (1,0,1) com PCA.

Fonte: Autor, gerada através do software Gretl 2015d.

Conforme parâmetros estimados pelo software Gretl, temos a equação do modelo estimado ARIMA (1,0,1) com componentes principais

$$\widehat{InadPrev}_t = 5,2677 + 0,966643Inadimp_d12_{t-1} + a_t + a_{t-1} + 0,0346527PC1_{t-1} - 0,00656785PC2_{t-1} - 1,02283dummy_{t-1}$$

onde: a_t é um erro associado ao modelo.

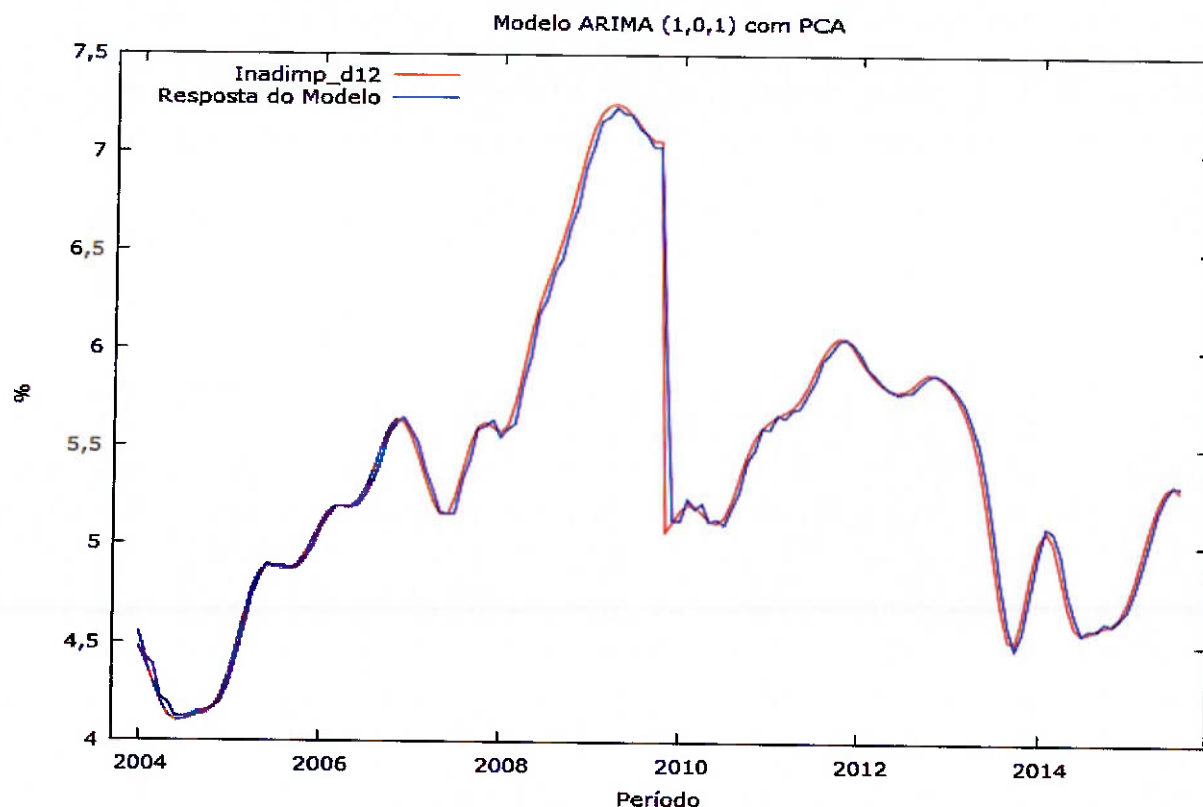


Figura 29 - Comparação do Modelo ARIMA (1,0,1) com PCA em relação a série temporal de inadimplência.

Fonte: Autor, gerada através do software Gretl 2015d.

4.9 Modelo ARIMA (1,0,1) sem variáveis regressoras

Nas seções 4.4 e 4.8 foram demonstrados os modelos ARIMA (1,0,1) com a utilização de variáveis regressoras e verificou-se que para os modelos, as variáveis não são relevantes. Conforme mostrado nas Figuras 30 e 31, as variáveis regressoras podem ser retiradas do modelo que sua auto-validação permanece praticamente a mesma, o valor de SQR para este modelo é igual a 1,87 que significa um baixo erro e o valor próximo aos obtidos nas seções 4.4. e 4.8 .

Modelo 5: ARMAX, usando as observações 2004:01-2015:08 (T = 140)
 Estimado usando o filtro de Kalman (Máxima verossimilhança exata)
 Variável dependente: Inadimp_d12
 Erros padrão baseados na Hessiana

	coeficiente	erro padrão	z	p-valor	
const	5,25365	0,417324	12,59	2,43e-036	***
phi_1	0,967489	0,0188464	51,34	0,0000	***
theta_1	1,00000	0,0230220	43,44	0,0000	***
dummy	-1,00727	0,00817619	-123,2	0,0000	***
Média var. dependente	5,408136	D.P. var. dependente	0,768289		
Média de inovações	-0,000361	D.P. das inovações	0,094832		
Log da verossimilhança	126,6174	Critério de Akaike	-243,2347		
Critério de Schwarz	-228,5265	Critério Hannan-Quinn	-237,2578		

Figura 30 - Parâmetros estimados do Modelo ARIMA (1,0,1) sem variáveis regressoras.

Fonte: Autor, gerada através do software Gretl 2015d.

Conforme parâmetros estimados pelo software Gretl, temos a equação do modelo estimado ARIMA (1,0,1) sem componentes regressores.

$$\widehat{InadPrev}_t = 5,25365 + 0,967489Inadimp_d12_{t-1} + a_t + a_{t-1} - 1,00727dummy_{t-1}$$

onde: a_t é um erro associado ao modelo.

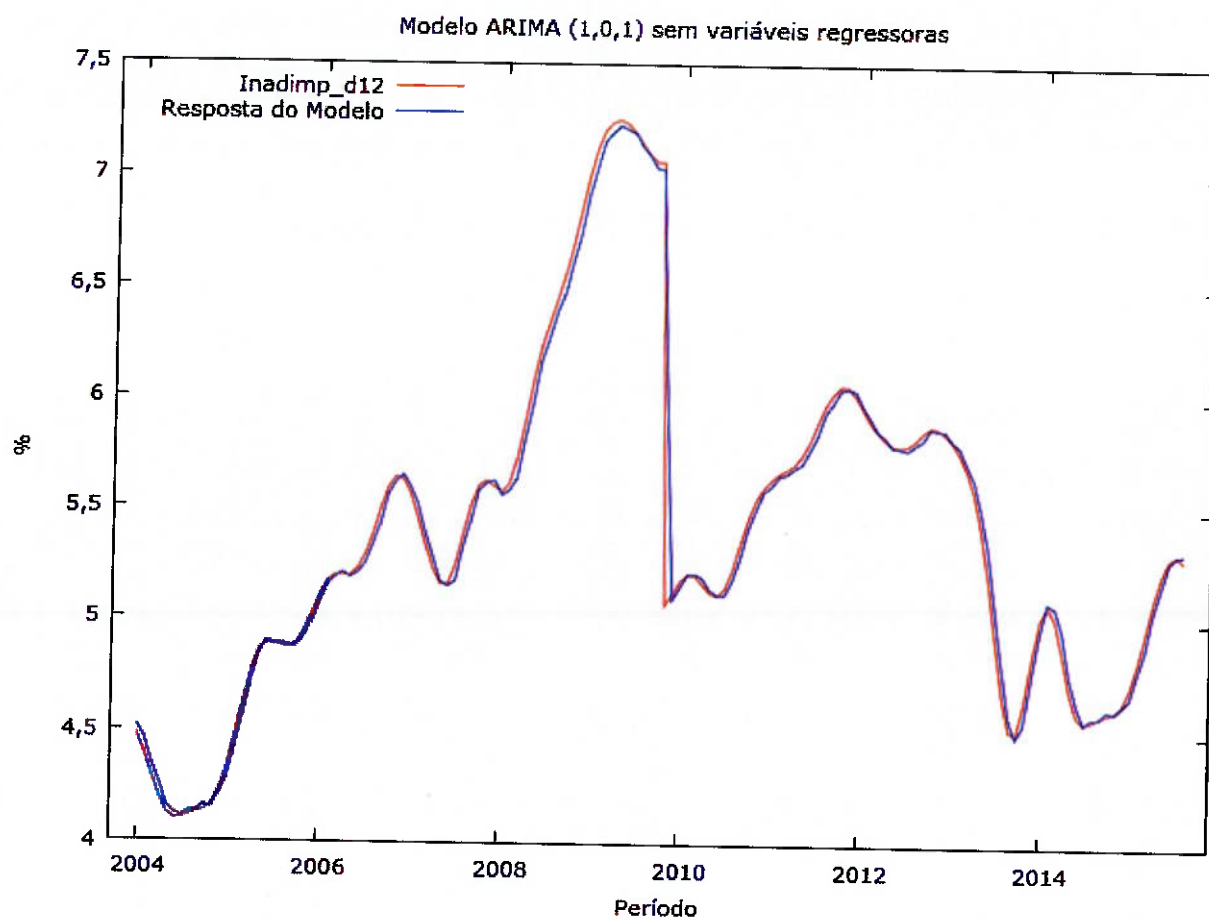


Figura 31 - Comparação do Modelo ARIMA (1,0,1) sem variáveis regressoras em relação a série temporal de inadimplência.

Fonte: Autor, gerada através do *software* Gretl 2015d.

4.10 Modelo ARIMA (0,0,2) com atrasos na variável desemprego

Uma hipótese de simulação do modelo é utilizar a série de desemprego com 5 e 6 meses de atraso como variável regressora em conjunto com as demais variáveis independentes. Supõem-se que a grande maioria das pessoas que perdem o emprego, recebem o benefício do seguro-desemprego concedido pelo governo federal. Esse benefício consiste no pagamento de 5 parcelas calculadas conforme regras estipuladas em relação ao salário individual. A inadimplência pode se refletir no mês posterior ao final do recebimento do seguro desemprego ou seja, 6 meses após a demissão.

Para o modelo ARIMA (0,0,2) com o atraso de 5 meses na série de desemprego, as variáveis desemprego e IPCA não são relevantes conforme mostrado na Figura 32,

o valor da soma quadrática dos resíduos (SQR) é igual a 7,63 e está próximo aos valores encontrados nas seções 4.3 e 4.7.

Modelo 27: ARMAX, usando as observações 2004:01-2015:08 (T = 140)
Estimado usando o filtro de Kalman (Máxima verossimilhança exata)
Variável dependente: Inadimp_d12
Erros padrão baseados na Hessiana

	coeficiente	erro padrão	z	p-valor	
const	5,97144	0,352007	16,96	1,52e-064	***
theta_1	1,76187	0,0688016	25,61	1,24e-144	***
theta_2	0,882950	0,0579457	15,24	1,99e-052	***
Juros_Reais	-0,0482708	0,0126231	-3,824	0,0001	***
IPCA	-0,0568590	0,0391188	-1,453	0,1461	
dummy	-1,00914	0,0419749	-24,04	1,02e-127	***
Desempre_t5	0,0127898	0,0206427	0,6196	0,5355	
Média var. dependente	5,408136	D.P. var. dependente		0,768289	
Média de inovações	-0,005470	D.P. das inovações		0,219680	
Log da verossimilhança	10,97582	Critério de Akaike		-5,951640	
Critério de Schwarz	17,58150	Critério Hannan-Quinn		3,611524	

Figura 32 - Parâmetros estimados do Modelo ARIMA (0,0,2) com atraso de 5 meses na variável desemprego.

Fonte: Autor, gerada através do software Gretl 2015d.

Conforme parâmetros estimados pelo software Gretl, temos a equação do modelo estimado ARIMA (0,0,2) com atraso de 5 meses na variável desemprego.

$$\begin{aligned} \widehat{InadPrev}_t = & 5,97144 + a_t + 1,76187a_{t-1} + 0,882950a_{t-2} - 0,0482708(JurosReais_{t-2} \\ & - JurosReais_{t-1}) - 0,0568590(IPCA_{t-2} - IPCA_{t-1}) \\ & - 1,00914(dummy_{t-2} - dummy_{t-1}) + 0,0127898(Desempre_{t-2} \\ & - Desempre_{t-1}) \end{aligned}$$

onde: a_t é um erro associado ao modelo.

O mesmo modelo com a série de desemprego atrasada em 6 meses, todas as variáveis são relevantes conforme mostrado na Figura 33, é o valor de SQR é igual a 6,74, os valores dos critérios de informação AIC e BIC podem ser considerados bons.

Modelo 28: ARMAX, usando as observações 2004:01–2015:08 (T = 140)
 Estimado usando o filtro de Kalman (Máxima verossimilhança exata)
 Variável dependente: Inadimp_d12
 Erros padrão baseados na Hessiana

	coeficiente	erro padrão	z	p-valor	
const	6,36777	0,217768	29,24	5,83e-188	***
theta_1	1,88669	0,0402698	46,85	0,0000	***
theta_2	1,00000	0,0420218	23,80	3,57e-125	***
Juros_Reais	-0,0370767	0,00808239	-4,587	4,49e-06	***
IPCA	-0,0548659	0,0280286	-1,958	0,0503	*
dummy	-1,03570	0,0233297	-44,39	0,0000	***
Desempre_t6	-0,0484029	0,0105907	-4,570	4,87e-06	***
Média var. dependente	5,408136	D.P. var. dependente	0,768289		
Média de inovações	-0,005088	D.P. das inovações	0,208003		
Log da verossimilhança	15,81101	Critério de Akaike	-15,62203		
Critério de Schwarz	7,911111	Critério Hannan-Quinn	-6,058864		

Figura 33 - Parâmetros estimados do Modelo ARIMA (0,0,2) com atraso de 6 meses na variável desemprego.

Fonte: Autor, gerada através do software Gretl 2015d.

Conforme parâmetros estimados pelo software Gretl, temos a equação do modelo estimado ARIMA (0,0,2) com atraso de 6 meses na variável desemprego.

$$\begin{aligned} \widehat{InadPrev}_t = & 6,36777 + a_t + 1,88669a_{t-1} + a_{t-2} - 0,0370767(JurosReais_{t-2} \\ & - JurosReais_{t-1}) - 0,0548659(IPCA_{t-2} - IPCA_{t-1}) \\ & - 1,03570(dummy_{t-2} - dummy_{t-1}) - 0,0484029(Desempre_{t-2} \\ & - Desempre_{t-1}) \end{aligned}$$

onde: a_t é um erro associado ao modelo.

Na Figura 34 é mostrada a auto-validação do modelo ARIMA (0,0,2) com atrasos na variável desemprego, observa-se uma boa aderência dos dados originais ao valores do modelo.

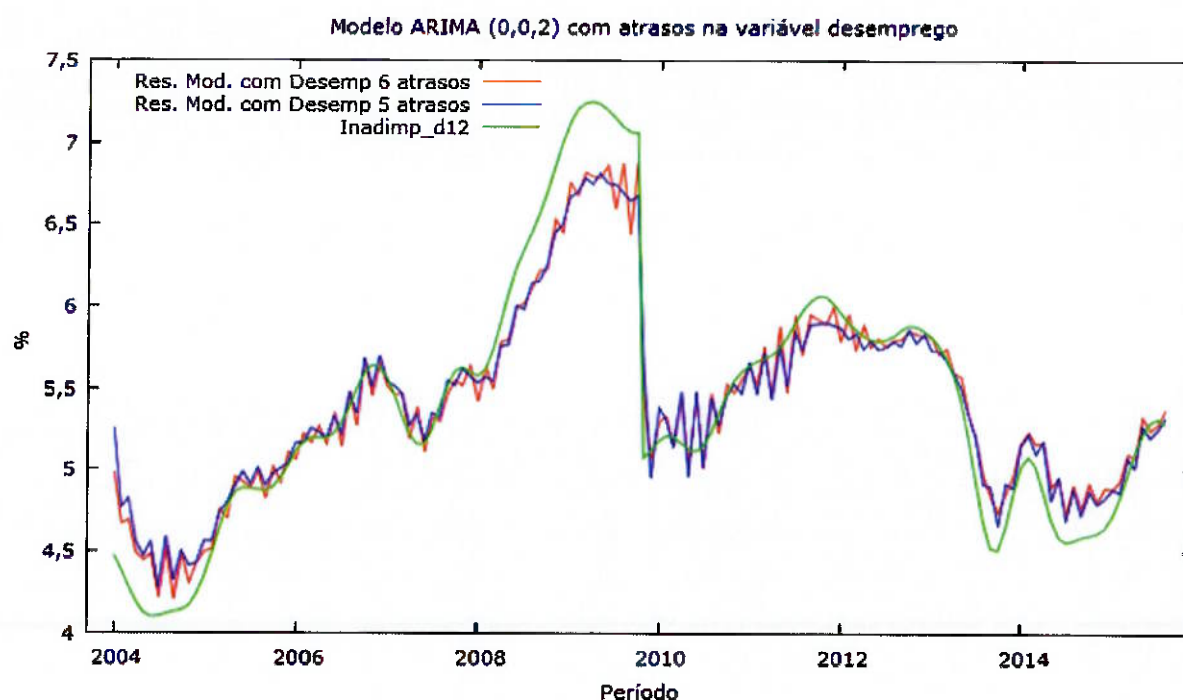


Figura 34 - Comparação do Modelo ARIMA (0,0,2) com atrasos na variável desemprego em relação a série temporal de inadimplência.

Fonte: Autor, gerada através do software Gretl 2015d.

A Tabela 2 resume a auto-validação dos modelos, observando os valores dos três últimos meses da série observada. Os modelos ARIMA (1,0,1) apresentaram uma maior assertividade dos valores em relação a série original.

Tabela 2 – Auto-validação dos modelos.

Período Observado da Inadimplência	Junho /15	Julho /15	Agosto /15
Valores da Série	5,29978	5,319683	5,288046
Auto-validação (140 Observações) - Variáveis Juros Reais, Desemprego e IPCA como regressores			
ARIMA(1,0,0)	5,261426	5,31685	5,321551
ARIMA(0,0,1)	5,124011	5,047234	5,141199
ARIMA(0,0,2)	5,220635	5,256145	5,313331
ARIMA(1,0,1)	5,270589	5,322537	5,312328
Auto-validação (140 Observações) - Duas componentes principais como regressores			
ARIMA(1,0,0)	5,245097	5,30594	5,31072
ARIMA(0,0,1)	5,124732	5,041512	5,140299
ARIMA(0,0,2)	5,218805	5,256465	5,314847
ARIMA(1,0,1)	5,27118	5,323185	5,312667
Auto-validação (140 Observações) – Sem variáveis regressoras			
ARIMA(1,0,1)	5,283972	5,313975	5,323204

Auto-validação (140 Observações) – Atrasos na variável desemprego			
ARIMA(0,0,2) 5 meses de atraso	5,202968	5,257072	5,327994
ARIMA(0,0,2) 6 meses de atraso	5,249030	5,288642	5,378236

Fonte: Autor, gerada através do *software* Microsoft Excel com os dados obtidos das simulações no Gretl.

4.11 Previsões 3 períodos à frente

A quantidade de observações de uma série temporal é importante para a qualidade do modelo. Não é recomendado utilizar séries temporais com um número de observações menor que 50. O período de observação das séries analisadas neste trabalho, foi escolhido a partir dos efeitos das mudanças políticas nos últimos anos e devido à limitação dos dados publicados.

As séries temporais com 140 observações, que são utilizadas neste trabalho são limitadas, pois alguns índices podem precisar de um tempo maior para se explicar. Com isso, não é possível fazer previsões a longo prazo, por exemplo 15 meses. Para previsões a longo prazo um conjunto com mais observações deveriam ser analisados.

Com o objetivo de testar a capacidade de previsão do modelo, foram utilizados os 137 primeiros pontos da série original para realizar simulações dos últimos 3 períodos conforme mostrado na Tabela 3.

Tabela 3 – Valores das previsões de 3 períodos para as observações dos meses de junho, julho e agosto de 2015.

Período Observado da Inadimplência	Junho /15	Julho /15	Agosto/15
Valores da Série	5,29978	5,319683	5,288046
Previsão Modelo (137 Observações - 3 períodos a frente) - Variáveis Juros Reais, Desemprego e IPCA como regressores			
ARIMA(1,0,0)	5,260181	5,276576	5,278772
ARIMA(0,0,1)	5,093021	4,792542	4,790764
ARIMA(0,0,2)	5,216567	5,100636	5,109961
ARIMA(1,0,1)	5,27056	5,264373	5,261325
Previsão Modelo (137 Observações - 3 períodos a frente) - Duas componentes principais como regressores			
ARIMA(1,0,0)	5,243983	5,249103	5,241331
ARIMA(0,0,1)	5,093257	4,789296	4,786662
ARIMA(0,0,2)	5,215176	5,098546	5,110578

ARIMA(1,0,1)	5,271366	5,266832	5,265023
Previsão Modelo (137 Observações - 3 períodos a frente) – Sem variáveis regressoras			
ARIMA(1,0,1)	5,284655	5,283794	5,282962
Previsão Modelo (137 Observações - 3 períodos a frente) – Atrasos na variável desemprego			
ARIMA(0,0,2) 5 meses de atraso	5,199583	5,071071	5,110864
ARIMA(0,0,2) 6 meses de atraso	5,249186	5,193201	5,271742

Fonte: Autor, gerada através do *software* Microsoft Excel.

A melhor assertividade das previsões, foram as realizadas pelo modelo ARIMA (1,0,1) mostrado na Figura 35, nesse modelo as variáveis independentes desemprego, juros e IPCA não tem influência conforme visto nas seções 4.4, 4.8 e 4.9.

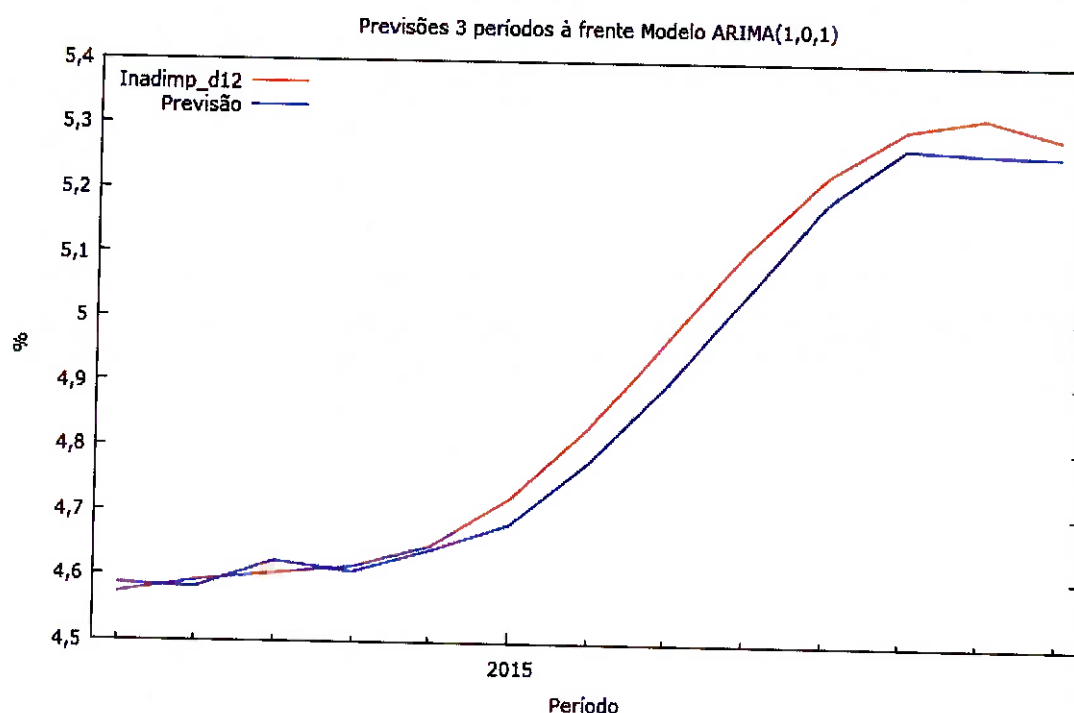


Figura 35 - Previsões do Modelo ARIMA (1,0,1) para 3 períodos à frente.

Fonte: Autor, gerada através do *software* Gretl 2015d.

O modelo ARIMA (0,0,2) com as componentes principais, realiza a previsão com o valor de SQR igual a 7,16. O modelo não é tão assertivo como o ARIMA (1,0,1). As variáveis independentes explicam a variável dependente, mas a previsão deste modelo não é boa, conforme mostrado na Figura 36.

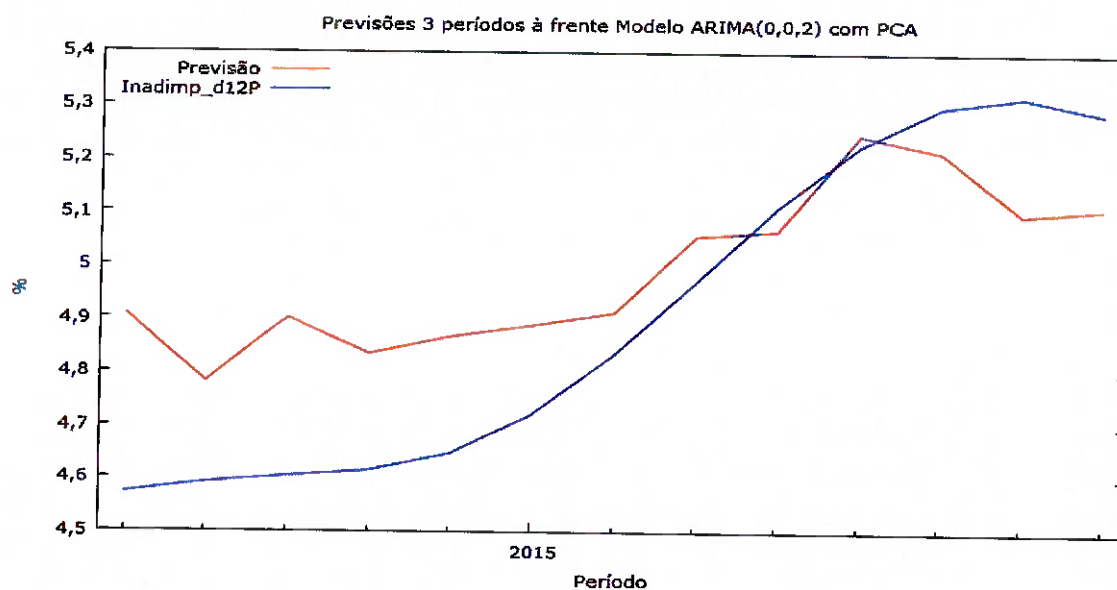


Figura 36 - Previsões do Modelo ARIMA (0,0,2) para 3 períodos à frente com PCA.

Fonte: Autor, gerada através do software Gretl 2015d.

O modelo ARIMA (0,0,2) com atrasos de 5 e 6 meses na variável desemprego, realiza as previsões com o valor de SQR respectivamente igual a 7,63 e 6,74. As variáveis independentes explicam a variável dependente no modelo com atraso de 6 meses, mas a previsão mostrada na figura 37 deste modelo não é boa.

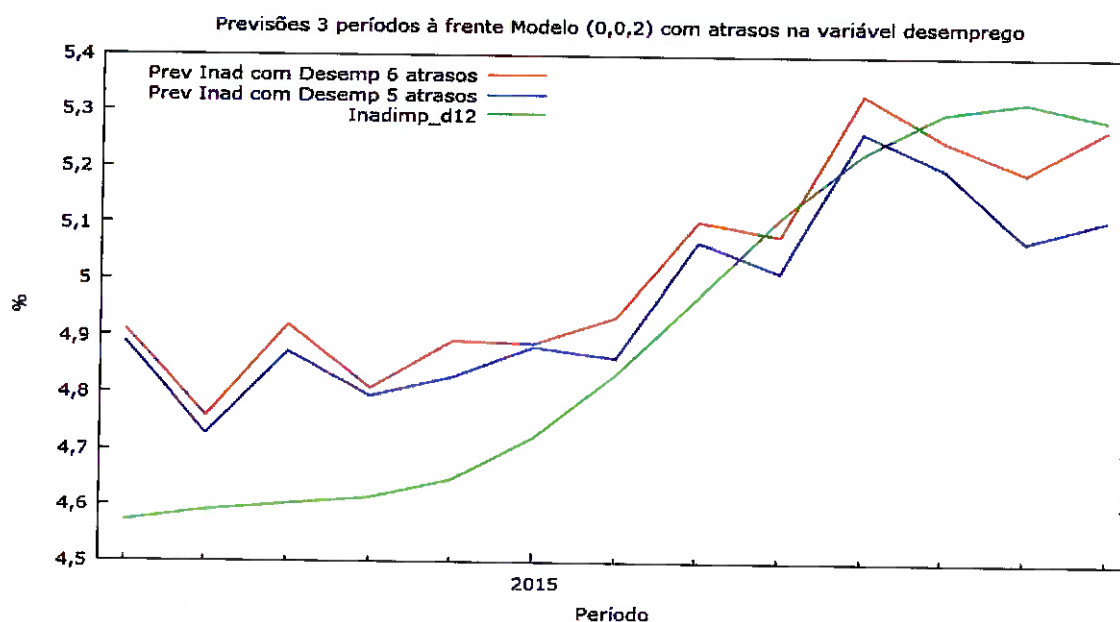


Figura 37 - Previsões do Modelo ARIMA (0,0,2) para 3 períodos a frente com atrasos na variável desemprego.

Fonte: Autor, gerada através do software Gretl 2015d.

5. Conclusões

Ao longo do trabalho foram aplicadas técnicas estatísticas nas séries temporais utilizadas para transformar os dados originais em dados adequados para análise, ou seja, sem alta sazonalidade.

Foram mostrados alguns modelos auto-regressivos com a utilização de variáveis regressoras. A inclusão da variável IPCA enriqueceu a análise, pelo fato da inflação estar relacionada com os juros reais.

Como sugestão após análise da função de autocorrelação e autocorrelação parcial, o modelo ARIMA(1,0,0) foi simulado utilizando as 3 variáveis independentes e as componentes principais como regressoras e apresentou uma boa aderência aos dados, mas não se observou relevância nas variáveis em ambos os casos.

Em relação aos modelos ARIMA(0,0,1) e ARIMA(0,0,2) com as 3 variáveis utilizadas como regressoras, as variáveis independentes Juros Reais e IPCA são relevantes, já quando aplicada a técnica estatística de análise de componentes principais, todas as variáveis foram relevantes para os dois modelos.

O modelo que possui a melhor assertividade em relação aos dados originais é o modelo ARIMA(1,0,1), mas as variáveis estudadas nesse trabalho não são relevantes, mesmo utilizando a técnica de componentes principais e podem ser excluídas, pois não impactam significativamente na aderência do modelo.

No modelo ARIMA(0,0,2) com a utilização das componentes principais ou da série de desemprego com atraso de 6 meses, as variáveis são relevantes e explicam os dados.

Devido à limitação da quantidade de observações das séries temporais observadas, conclui-se que não é possível realizar previsões de longo prazo. Foram realizadas as previsões 3 períodos à frente utilizando os 137 primeiros pontos das séries e o melhor modelo de previsão foi o ARIMA(1,0,1).

Baseado em dados sem respaldo científico, acredita-se que a inadimplência possa estar relacionada ao descontrole financeiro das pessoas, pois a situação

econômica do país, na grande parte do período observado é de crescimento e os índices de desemprego apresentam queda.

Em relação ao período estudado notou-se uma queda no índice de desemprego no período entre janeiro de 2004 a dezembro de 2014, o que pode justificar uma baixa correlação entre as variáveis. Já a partir de janeiro de 2015 o desemprego aumentou e provavelmente o índice de inadimplência será afetado nos períodos posteriores aos analisados neste trabalho. Outro provável problema pode ser a forma da coleta dos dados de inadimplência, pois não existe uma classificação de motivo de inadimplência.

As taxas de juros impactam na inadimplência em vários modelos, pois seu crescimento impacta diretamente no endividamento das pessoas.

Pode-se concluir que após as análises realizadas, a influência das variáveis de desemprego e juros reais sobre a inadimplência foram baixas, pois observou-se que em alguns modelos as variáveis independentes são relevantes e em outros não. As características econômicas do período podem ter influenciado o resultado.

Um período de tempo maior e a inclusão de outras variáveis econômicas e individuais, provavelmente poderão explicar melhor o índice de inadimplência. Entretanto, o objetivo inicial deste trabalho foi apenas a análise das variáveis desemprego e juros reais.

Referências Bibliográficas

- ALENCAR, B. **Análise Multivariada de Dados, dissertação de mestrado**. Belo Horizonte, Universidade Católica de Minas Gerais, 2005. 101p.
- ANIBAL, C. **Trabalhos para Discussão 192, inadimplência no setor bancário brasileiro**, no sítio <http://www.bcb.gov.br/pec/wps/port/wps192.pdf> visitado em 10 de outubro de 2015.
- Banco Central do Brasil**, no sítio <https://www3.bcb.gov.br/sgspub/>, visitado nos meses de agosto, setembro, outubro, novembro de 2015.
- BARROSO, Lúcia P. e ARTES, R. **Análise Multivariada, notas de aula**. São Paulo, 2003. 159p.
- BARTH, Nelson L. **Inadimplência Construção de Modelos de Previsão**, 1ª edição, São Paulo - Editora Nobel, 2004.
- BMF&Bovespa**, no sítio <http://www.bmfbovespa.com.br/>, visitado nos meses de agosto, setembro, outubro, novembro de 2015.
- CASTRO, Neto J. L. **Análise de Risco e Crédito**, 1ª edição, Curitiba - Editora IESDE Brasil, 2006.
- Equipe Docente. **Estimação de Modelos ARIMA** Acessado em 24 de janeiro de 2016. Disponível em: [http://portal.uned.es/ESTIMACIÓN DE MODELOS ARIMA.pdf](http://portal.uned.es/ESTIMACIÓN%20DE%20MODELOS%20ARIMA.pdf)
- FEIJÓ, Carmen. **Contabilidade Social**, 3ª edição, Rio de Janeiro - Editora Elsevier 2008.
- GERBELLI, Luiz G. **Juro para o consumidor é maior desde 2009**. *Jornal O Estado de São Paulo*, São Paulo, p.B5, 15 set. 2015.
- HOUAISS, Antônio. **Dicionário Houaiss da Língua Portuguesa**. Rio de Janeiro: Objetiva, 2001.
- IBGE**, no sítio <http://www.ibge.gov.br/>, visitado nos meses de agosto, setembro, outubro, novembro de 2015.
- IpeaData**, no sítio <http://www.ipeadata.gov.br/>, visitado nos meses de agosto, setembro, outubro, novembro de 2015.
- MANKIW, Gregory N. **Introdução a economia**, 5. reimpr. 1ª edição, São Paulo - Editora Cengage Learning, 2009.
- MONTEIRO, Luiz H. A. **Sistemas Dinâmicos**, 2ª edição, São Paulo - Editora Livraria da Física, 2006.
- MORRETTIN, A. Pedro, TOLOI, M. C. **Análise de Séries Temporais**, 2ª edição, São Paulo - Editora Edgar Blucher, 2006.
- PASTORE, José. **Proteções em tempo de desemprego**. *Jornal O Estado de São Paulo*, São Paulo, p.A7, 22 set. 2015.
- REINERT, J. N. **Desemprego: causas, consequências e possíveis soluções**. *Revista de Ciências da Administração*, v. 3, n. 5, p. 45-48, 2001.

ROSSI, José. W., NEVES, Cesar. **Econometria e séries temporais com Aplicações a dados da economia brasileira**, 1ª edição, Rio de Janeiro - Editora LTC, 2014.

Serasa Experian, no sítio <http://www.serasaexperian.com.br/estudo-inadimplencia/>, visitado em 12 de outubro de 2015.

SOBRINHO, J. D. **Taxa de Juros: Nominal, efetiva ou real?**. Revista de Administração de Empresas, p. 77-82, 1981.

SPIEGEL, Murray R. **Estatística**, 3ª edição, São Paulo - Editora Makron Books, 1993.

STEVENSON, William J. **Estatística aplicada a administração**, 1ª edição, São Paulo - Editora Harper & Row do Brasil, 1981.

TEIXEIRA, Flavia. D. **Influência de Variáveis Macroeconômicas na Inadimplência do Sistema Financeiro da Habitação**. Revista de Finanças Aplicadas, p. 1-16, 2010. Acessado em 05 de novembro de 2015. Disponível em: <http://www.financasaplicadas.net/ojs/index.php/financasaplicadas/article/viewFile/31/36>

