

Previendo Inflação usando Ridge Regressions

Luiz Renato Lima,
Fundação Getulio Vargas

Claudia F. Rodrigues,
Fundação Getulio Vargas

October 3, 2006

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Abstract

Este estudo tem como objetivo analisar o desempenho de vários modelos econométricos ao prever inflação. Iniciamos o trabalho utilizando como base de comparação para todos os modelos a tradicional curva de Phillips que usa a taxa de desemprego como variável explicativa para diferenças de preço. Dentre os modelos analisados temos univariados e bivariados, sendo estes últimos uma curva de Phillips alternativa já que apenas substitui a variável desemprego por outra variável macroeconômica. Além destes modelos também comparamos o desempenho de previsão de modelos que usam como covariadas uma combinação das previsões dos modelos anteriores (univariados e bivariados).

O resultado deste estudo aponta a combinação de modelos por "ridge regression" como uma técnica - dentre as analisadas para combinação de previsões - de menor erro de previsão sempre; sendo alcançado pela combinação da média em apenas um dos casos analisados. No entanto, a combinação de previsões não apresentou melhor resultado que algumas das covariadas testadas em modelos bivariados.

Keywords: previsão de inflação, curva de Phillips, "ridge regression".

1 Introdução

A economia brasileira vem experimentando, ao longo dos últimos anos, uma relativa estabilidade de preços. Isto foi conquistado inicialmente com a introdução do Plano Real, em junho de 1994, e ratificado com a adoção do regime de metas de inflação, em meados de 1999.

Em janeiro de 1999 houve uma acentuada desvalorização cambial, esta desvalorização resultou no abandono da âncora cambial. Em junho do mesmo ano nova âncora foi estabelecida para o real. Dessa vez uma âncora nominal definida com base em um regime de metas de inflação. Por este regime, o governo define a meta e atribui ao Banco Central a responsabilidade e independência operacional para conduzir a política monetária de forma a alcançar a meta estabelecida. Para isso é necessário que autoridades monetárias adotem uma atitude de visão futura e tome atitudes para evitar ou deter situações antecipadas como indesejáveis, conhecendo a defasagem entre decisões políticas e seus efeitos sobre a produção e preços. Nas palavras do ex-presidente do Banco Central americano, Alan Greenspan, "Implícito em qualquer atitude (ou falta de) está uma expectativa de como o futuro irá se desdobrar, isto é, uma previsão". De fato, os bancos centrais atuam antecipando a inflação, e tomando decisões baseadas nestas previsões, condicionais a diferentes trajetórias da taxa de juros, e ao desenvolvimento de outras variáveis exógenas que compõem o cenário macroeconômico. Como órgão deliberativo máximo do sistema financeiro nacional - o Conselho Monetário Nacional - utiliza como indicador de inflação um índice reportado pelo IBGE mensalmente o IPCA.

Neste estudo faremos uma análise e comparação do grau de previsibilidade de vários modelos econométricos capazes de auxiliar o processo de previsão da inflação brasileira. Tomamos como referência principal um estudo de Stock, Watson 1999 sobre a previsibilidade da inflação americana. Replicaremos essa experiência utilizando dados brasileiros e variáveis mais adequadas a explicação da inflação nacional. Utilizamos aqui dados mensais entre 1995 e 2005, período este escolhido por ser o de maior estabilidade da variável chave. Também consideramos um período mais curto, a partir de junho de 1999, pois nessa data o governo alterou a forma de condução da política monetária, e queremos verificar se a previsibilidade dos modelos se comporta de forma diferente a partir de então.

Os modelos econométricos utilizados neste estudo tiveram sua previsibilidade comparada a previsibilidade de um modelo que utiliza o desvio do desemprego de sua taxa natural como principal explicação para inflação, ou seja, o tradicional modelo da curva de Phillips. Não queremos com isso impor nenhum modelo como padrão, apenas escolhemos a curva de Phillips por ser tradicional na área macroeconômica e a utilizaremos como um bechmark para os demais modelos.

Neste estudo a inflação brasileira é representada pela série mensal do IPCA; também realizamos o mesmo exercício considerando a série INPC, ambas reportadas pelo IBGE e utilizadas nacionalmente como índices de inflação. Seleccionamos modelos dentre as categorias ARMA e ADL para explicar esses índices. Nosso objetivo, na primeira parte do estudo, é comparar modelos de acordo

com seus erros de previsão. A segunda parte do estudo lida com outros tipos de modelos - regressões resultantes das combinações de previsões dos modelos bivariados - é nesta parte que analisamos o tipo de regressão que deu nome ao artigo.

Este documento encontra-se dividido da seguinte forma. A seção dois descreve os modelos ADL e ARMA utilizados para previsão de inflação e analisa os resultados empíricos de performance destes modelos comparada a performance do modelo de "benchmark". A seção três apenas complementa o mesmo estudo incluindo novas covariadas também utilizadas para previsão das mesmas séries de inflação - estas novas covariadas foram identificadas em estudos de "leading indicators" para a inflação brasileira (papers de discussão do Banco Central nro 7). A seção quatro apresenta possíveis combinações dos resultados de previsão dos modelos estudados na seção 2 - ADL e ARMA - com a intenção de gerar uma previsão mais acurada e também analisa os resultados obtidos com essas combinações; esta seção descreve o método de combinação de previsões "ridge regression" e o motivo do seu uso quando se trata de combinações de regressores multicolineares; a seção cinco conclui o estudo.

2 Previsão de Inflação: Modelos Estruturais

O estudo de modelos de séries temporais tem como principal objetivo encontrar equações que se ajustem bem aos dados observados. Isto porque, equações que bem representam o processo gerador de dados, podem ser utilizadas para prever o comportamento futuro de séries de tempo.

Nesta seção abordaremos modelos ditos ADL (autoregressive distributed lag). Estes são modelos que utilizam defasagens da variável explicada e defasagens de uma variável explicativa como regressores da equação.

O ADL utilizado como "benchmark" para os demais é uma representação econométrica da curva de Phillips convencional; isto é, é uma relação entre desvios do desemprego de sua taxa natural e inflação passada. Podemos escrevê-lo da seguinte forma:

$$\pi_{t+1} = \phi + \beta(L)u_t + \gamma(L)\pi_t + e_{t+1} \quad (1)$$

onde:

π_t	inflação no período t
u_t	desemprego no período t
$\beta(L), \gamma(L)$	polinômios no operador de defasagem L

Uma particularidade desta equação (1) é que consideramos a inflação em t dependente de inflações passadas e de taxas de desemprego passadas, e não do desvio dessas taxas de uma taxa natural ($u_t - \bar{u}_t$). Podemos explicar esta simplificação se pensarmos que a taxa de desemprego natural é constante ao longo dos anos ($\bar{u}_t = \bar{u}$) e com isso seu papel pode ser representado pelo intercepto de (1).

Os demais modelos ADL foram criados a partir de uma reespecificação da curva de Phillips tradicional utilizando outra variável de atividade agregada como explicativa para a inflação. Isto é, temos modelos econométricos mais genéricos representados por:

$$\pi_{t+1} = \phi + \beta(L)x_t + \gamma(L)\pi_t + e_{t+1} \quad (2)$$

onde:

x_t indicador de atividade real agregada da economia

A questão a identificar nestes modelos é de quantas defasagens são necessárias para explicar a dependência entre as variáveis.

As equações dadas por (1) e (2) definem genericamente a forma dos ADLs utilizados mas não determinam um único modelo pois não especificam o número de defasagens a ser utilizado pelos regressores. Na realidade este número depende do regressor em questão e foi determinado segundo um critério de informação. Um critério de informação é uma medida que pondera número de regressores (no caso defasagens da covariada) e erro quadrático médio do modelo. A intenção ao analisarmos um critério de informação é escolhermos um modelo que possua um número mínimo de regressores e um erro de previsão também baixo.

Podemos aprovar o uso de um critério de informação para escolha de número de regressores se pensarmos que inerente a cada modelo existe um erro de previsão; este erro é a diferença entre o verdadeiro valor futuro (observado posteriormente) e a previsão hoje deste valor futuro. O erro de previsão pode ser dividido em duas partes: a primeira está relacionada aos estimadores calculados para a amostra (ie. quão bem foram estimados), e a segunda é proveniente de choques idiossincráticos. Quando fazemos uma previsão dentro-da-amostra nossa estimativa dos parâmetros são os valores corretos para a amostra, logo este erro de estimação não é de interesse, mas quando fazemos uma previsão fora-da-amostra este erro é importante. Esta subdivisão do erro nos ajuda a perceber que ao incluirmos novos regressores ao modelo podemos explicar melhor os choques (ie, diminuimos o erro médio), mas por outro lado criamos o problema da estimação de novos parâmetros - uma tarefa na qual se incorre em novos erros. A idéia por trás do critério de informação é atribuir um valor numérico a cada modelo compensando os que possuem simultaneamente o menor número de regressores (ie os que terão menor erro no cálculo de estimadores) e o menor erro quadrático médio (ie menor erro de previsão). Utilizamos neste estudo o SIC (Schwartz Information Criteria) como critério de informação.

As variáveis dependentes utilizada para os ADLs foram IPCA, INPC. As covariadas alternativas para atividade econômica agregada foram: produção industrial, renda média pessoal dos trabalhadores, horas trabalhadas, volume de comércio com o exterior (exportações líquidas), folha de pagamento da indústria, taxa de utilização da capacidade instalada, e taxa de desemprego. Todas estas séries encontram-se descritas no apêndice deste texto.

As séries utilizadas nos modelos foram testadas quanto a estacionariedade. As dependentes foram detectadas como estacionárias (no período amostral considerado) pelo teste de raiz unitária, e não sofreram qualquer transformação em seus dados para o cálculo da regressão. Já as covariadas quando identificadas como não estacionário (pelo teste de raiz unitária), foram transformadas tirando-se uma diferença da série (quando esta possuía raiz unitária em modelo com tendência temporal e intercepto) ou aplicando-se o filtro de Hodrick-Prescott (HP) (quando a raiz unitária só foi verificada em modelo sem tendência temporal, mas com intercepto). Este filtro é um algoritmo para escolha de valores mais suaves para uma série de tempo. O parâmetro lambda deste algoritmo define o grau de suavização da série. Quanto maior, mais suave a série resultante. Para o estudo de séries estacionárias não usamos esta suavização mas sim a série de resíduos resultante da diferença entre a série original e a suavizada. Os dados filtrados por HP mostram menor flutuação que dados em primeira diferença.

A análise empírica foi feita examinando o desempenho de previsão de cada modelo ADL e de um modelo ARMA (univariado) comparado ao desempenho de previsão do modelo tradicional (ADL com covariada desemprego) numa simulação de previsão "pseudo" fora-da-amostra. Procedemos da seguinte forma: primeiramente identificamos o melhor modelo para cada covariada (número ótimo de defasagens) a partir do critério de informação SIC; em seguida, calculamos diversas estatísticas para resumir o desempenho da previsão "pseudo" fora-da-amostra de cada modelo: uma dessas estatísticas foi o erro quadrático médio de previsão reportado na tabela de resultados (tabela 1). Para o cálculo deste erro utilizamos o procedimento "pseudo" fora-da-amostra que consiste em quebrar os dados observados em dois conjuntos, um para ser usado na estimação dos parâmetros da regressão e outro para ser usado confrontando valores determinados pelo modelo estimado, a eles. Assim, supondo que estamos no ano 2000:1 fizemos uma estimação dos parâmetros do modelo utilizando dados observados até essa data e a partir desses parâmetros fizemos uma previsão 12 meses a frente (isto é, até 2001:1). Os valores previstos foram então comparados aos observados no período (2000:1 à 2001:1) e o erro quadrático médio computado. Fizemos isso para n períodos de 12 meses, (sendo $n=1,2,3, 4$) utilizando o método OLS recursivo. O valor de n não pode ser muito grande senão perdemos muitas das observações usadas para estimar os parâmetros, e com isso teremos estimadores pobres e erro de previsão alto demais.

Outra estatística apresentada na tabela 1 avalia se a covariada alternativa (x_t) realiza uma contribuição útil para a previsão da inflação quando comparada a previsão da covariada tradicional (desemprego). Para essa estatística utilizamos uma regressão que pondera a previsão destas duas covariadas ao explicar a inflação. Esta regressão foi uma estimação OLS do seguinte modelo:

$$\pi_t = \lambda f_t^x + (1 - \lambda) f_t^u + e_t \quad (3)$$

onde:

- π_t inflação no período t
- f_t^x previsão realizada pela covariada x
- f_t^u previsão realizada pela covariada desemprego
- e_t erro de previsão associado a combinação de previsões

Pela equação (3) $\lambda = 0$ implica que as previsões do modelo que utiliza a covariada x não trazem melhor informação para previsibilidade da inflação que a variável convencional (desemprego); por outro lado, $\lambda = 1$ implica que as previsões obtidas com a covariada x contribuem significativamente para previsão da inflação. Os resultados obtidos encontram-se também na tabela 1 (coluna peso).

Sobre as tabelas deste estudo, cabe salientar a estrutura comum usada por essas para a apresentação dos resultados. Primeiramente descrevemos o nome da variável utilizada como dependente na regressão e em seguida a covariada que foi utilizada para comparação dos demais modelos. Repare que o MSE reportado é sempre relativo, ou seja, é o MSE do modelo i em relação ao MSE do modelo de "benchmark". Neste estudo utilizamos sempre o desemprego como covariada para "benchmark". A coluna variável indica a covariada utilizada no modelo i , cada linha da tabela apresenta resultados de um modelo específico. A coluna Transf possui siglas que indicam se a covariada sofreu transformações antes de ser utilizada na estimação dos parâmetros; a legenda utilizada foi: LN: log da série original, DLN: log e diferença da série original, DLV: diferença tirada em nível da série original; DT: aplicado o filtro HP na série em log; LV: série utilizada sem transformação.

A coluna peso, desvio padrão e p-valor estão relacionadas a estimação do parâmetro (λ) apresentado a cima. O número total de observações diz respeito ao tamanho da amostra antes da divisão dos dados; o número de observações utilizadas para estimação considera o volume de dados considerado para estimação dos parâmetros e o número de observações fora da amostra indica o número de observações utilizadas para se confrontar com valores obtidos do modelo estimado. Este número será sempre um múltiplo de doze porque estamos querendo estimadores com melhor previsão no período de um ano.

Os resultados obtidos considerando a inflação medida pelo IPCA no período pós Plano Real, mais precisamente entre 1995 e 2005 estão reportados na tabela 1. Todas as séries utilizadas são mensais. O significado das covariadas listadas abaixo está descrito no apêndice deste documento.

Tabela 1

Variável Dependente: IPCA
 Covariada benchmark: desemprego
 Amostra completa: 1995 a 2005
 Fora da amostra: 2003:12 a 2005:11

Variável	Transf	MSE Relativo	Peso	DP	P-Valor
pi_bcons	DLN	1,008149	0,322471601	1,39640597	0,819411314
rm_nsp	LN	1,000469	0,497080145	0,76901687	0,524428893
venda_nom	DLN	0,967597	1,768535625	1,1968376	0,153058997
htrab	DLN	0,845669	0,893873144	0,45092561	0,059515726 *
folhapagto	DLN	1,005032	0,474659832	0,53996124	0,388459219
utilcapacinst	DT	1,03411	0,280450008	0,51093263	0,588364939
exportliq	DLV	0,965527	2,377366337	1,43221996	0,110505859
Arma		1,511207	-0,50249699	0,46101115	0,287006631
Covariada benchmark		RMSE			
desemprego	DLN	0,235030878			

Nro total de observações: 131

Nro de observações fora da amostra: 24 = (2 períodos de 12 meses)

Podemos observar que 'horas trabalhadas' (htrab) apresenta melhor previsão para o IPCA que desemprego, considerando um horizonte de previsão de dois períodos de doze meses. Somente esta covariada realiza previsões com peso superior (e significativo a 10%) ao das previsões realizadas por desemprego. O modelo ARMA bem como as demais covariadas não tiveram um bom desempenho relativo.

Em toda a análise reportamos resultados obtidos quando o período de dois anos é colocado fora-da-amostra. No entanto, fizemos testes para outros valores de n (número de anos fora-da-amostra) indicados anteriormente e o resultado apresentado para dois anos é preservado para 'horas trabalhadas'.

Os resultados obtidos considerando a inflação medida pelo INPC no período pós Plano Real estão reportados na tabela 2. As mesmas covariadas utilizadas anteriormente para o IPCA são reutilizadas aqui.

Tabela 2

Variável Dependente: INPC
 Covariada benchmark: Desemprego
 Amostra completa: 1995 a 2005
 Fora da amostra: 2003:12 a 2005:11

Variável	Transf	MSE Relativo	Peso	DP	P-Valor
pi_bcons	DLN	1,035909	-1,509244419	1,968480495	0,451055292
rm_nsp	LN	1,116282	-0,455041555	0,869765928	0,60585534
venda_nom	DLN	1,016149	0,081547094	0,991566007	0,935166831
htrab	DLN	0,8124	0,937033485	0,426769191	0,038469858 *
folhapagto	DLN	1,037898	0,198999662	0,663165237	0,766814915
utilcapacinst	DT	1,060879	0,328881078	0,287944588	0,26513574
exportliq	DLV	1,046277	-2,514337978	2,262909244	0,278002321
Arma		1,432464	-1,031764323	0,600982206	0,099455492

Covariada benchmark RMSE
 desemprego DLN 0,311553508

Nro total de observações: 131

Nro de observações fora da amostra: 24 = (2 períodos de 12 meses)

Também para INPC 'horas trabalhadas' é a única covariada com previsão melhor que desemprego e significativa a 5%. O modelo ARMA também neste caso não faz uma previsão melhor que desemprego.

Das duas tabelas percebemos que a previsão para série IPCA é mais precisa que a previsão para INPC em razão do menor erro quadrático médio ser reportado para a primeira.

Fazendo o mesmo exercício apenas alterando o período para o que segue a fixação de metas de inflação (1999:8), mas ainda utilizando dois anos como período fora-da-amostra, temos os seguintes resultados reportados pelas tabela 3 e tabela 4.

Tabela 3

Variável Dependente: IPCA
 Covariada benchmark: desemprego
 Amostra completa: 1999 a 2005
 Fora da amostra: 2003:12 a 2005:11

Variável	Transf	MSE Relativo	Peso		P-Valor
pi_bcons	DLN	1,195418	-1,118695008	0,84241422	0,197219306
rm_nsp	DT	1,405077	0,237645365	0,18391816	0,2091462
venda_nom	DT	1,25121	0,027678875	0,40380758	0,945944463
htrab	DLN	0,885	0,920558618	0,50184488	0,079572368 *
folhapagto	DLN	1,038535	-0,034342594	0,87632852	0,969077748
utilcapacinst	LN	1,45243	-0,561751073	0,28860906	0,063922026
exportliq	DT	1,857045	-0,121951883	0,24962116	0,629787622
Arma		1,004813	-1,390837404	6,28621137	0,826848649
Covariada benchmark		RMSE			
desemprego	DLN	0,261807925			

Nro total de observações: 76

Nro de observações fora da amostra: 24 = (2 períodos de 12 meses)

Tabela 4

Variável Dependente: INPC
 Covariada benchmark: desemprego
 Amostra completa: 1999 a 2005
 Fora da amostra: 2003:12 a 2005:11

Variável	Transf	MSE Relativo	Peso		P-Valor
pi_bcons	DLN	1,23228	-1,509026062	0,641466407	0,027578208
rm_nsp	DT	0,901518	0,613474149	0,254653234	0,024400214 *
venda_nom	DT	1,632266	-0,27527503	0,365353815	0,458824812
htrab	DLN	0,923565	0,939319657	0,621128174	0,144082572
folhapagto	DLN	1,180502	-0,076011684	0,411824943	0,85518191
utilcapacinst	LN	1,9212	-0,634502374	0,111783142	8,86E-06
exportliq	DT	1,959563	-0,333300236	0,314143305	0,299712999
Arma		1,063722	-1,503862062	1,663895644	0,37546056
Covariada benchmark		RMSE			
desemprego	DLN	0,381522206			

Nro total de observações: 76

Nro de observações fora da amostra: 24 = (2 períodos de 12 meses)

Podemos observar que o IPCA continua tendo 'horas trabalhadas' como variável com melhor informação para previsão que a covariada desemprego. No entanto, para o INPC 'horas trabalhadas' apesar de ter um MSE relativo menor não é significativo a 10% quando se compara suas previsões a do modelo "benchmark". Neste caso `rm_nsp` apresentou melhor poder de previsão.

3 Outros Indicadores de Atividade Econômica

Podemos estender a análise da seção anterior a novos modelos bivariados, que apenas possuem outra covariada x indicadora da atividade econômica. A mesma metodologia para identificar a performance de previsão adotada na seção anterior também foi considerada para essas novas covariadas.

Ao todo 44 séries foram utilizadas e tiveram seu desempenho quanto a previsão de inflação comparado ao desempenho de previsão de desemprego. Todas as séries estão listadas e descritas no apêndice no final do documento.

Repare que as covariadas que representam taxa de câmbio, salários e os modelos ARMA não ajudaram na previsibilidade de inflação reportada pelo IPCA (tabela 5); já as covariadas de taxa de juros, moeda e crédito apresentaram performance de previsão melhor que o "benchmark", bem como covariadas de atividade real da economia, como produção industrial e consumo.

Para o INPC observamos que as taxas de câmbio apresentaram melhor poder de previsão que a taxa de desemprego (tabela 6); no entanto, essa constatação não pôde ser observada quando um número maior de observações foi colocado fora-da-amostra. Neste caso, as covariadas que representam produção industrial apresentaram melhor comportamento previsor de inflação em todos os casos testados para INPC.

Um ponto importante a salientar é que as tabelas 5 e 6 são idênticas as tabelas 1 e 2 apenas sendo uma versão estendidas destas últimas para novas covariadas. As covariadas selecionadas para as tabelas 1 e 2 apenas foram escolhidas por uma intuição do que poderia explicar a inflação; percebemos nesta análise ampliada que outras covariadas podem ser mais interessantes para explicar cada índice de inflação separadamente.

Tabela 5

Variável Dependente: IPCA
 Covariada benchmark: desemprego
 Amostra completa: 1995 a 2005
 Fora da amostra: 2003:12 a 2005:11

Variável	Transf	MSE Relativo	Peso		P-Valor
Consumo e Venda					
consgas	DLN	0,968435	0,90353379	0,634742	0,168026
consoleocomb	DLN	0,948996	3,65575733	1,753828	0,048415 *
consoleodiesel	DLN	0,965395	2,97157575	2,237159	0,197116
spc	LN	1,083704	-0,2607319	1,080946	0,811532
vendareal	DLN	0,953388	1,10346524	0,689405	0,123112
Moeda e Crédito					
mcm0	DLN	0,825507	1,28087799	0,524599	0,022728 *
mcm1	DLN	0,792325	1,23300372	0,490563	0,019409 *
mcm2	DLN	0,972994	5,59575721	2,088981	0,013411 *
mcibvsp12	LV	0,88647	2,54041439	0,815056	0,00485 *
mcrdv12	DT	1,027182	0,09876417	0,731327	0,893749
mcvcdp12	DT	0,96223	3,56031018	2,249425	0,12713
mcvcdv12	DT	0,99471	0,66980313	0,900686	0,464613
Prod Industrial					
pi_bcons	DLN	1,008149	0,3224716	1,396406	0,819411
pi_bcap	DT	1,091197	-0,327814	0,732746	0,658787
pi_bconsdur	DLN	0,947512	4,94291524	2,498885	0,060024 *
pi_bconsndur	LN	1,164892	-1,0968388	1,231345	0,382279
pi_bint	DT	0,644549	1,6063674	0,430461	0,001093 *
pi_extmin	DT	0,900106	4,08913042	1,896148	0,041734 *
pi_fumo	DT	0,859047	2,51288789	1,334329	0,072373
pi_indtransf	DT	0,949336	8,17748361	2,645403	0,005154 *
pi_insumos	DLN	0,916793	0,68981852	0,294841	0,02835 *
pi_leite	DLN	0,991297	1,34940116	2,752984	0,628664
pi_madeira	DLN	0,986553	0,57368009	0,529573	0,289906
pi_maquequip	DLN	1,013628	0,43809968	0,500552	0,390496
pi_metal	DLN	0,993279	0,62685785	1,004467	0,538723
pi_oleocomb	DT	0,976397	2,08876508	2,449902	0,402677
pi_transporte	DLN	0,934818	1,56064851	0,933366	0,108063
prodpetrol	DLN	4,143583	0,01382321	0,014064	0,3359
Salário					
sal_indn	DLN	1,015243	0,31795496	1,197036	0,792899
sal_indr	DLN	0,997765	0,52010609	1,104435	0,64213
Taxas de Câmbio					
tccompra	DLN	1,239111	0,39496101	0,222392	0,088972
tcvenda	DLN	1,130067	0,42041552	0,258644	0,117691
Taxas de Juros					
tjcdB	DT	0,979572	2,54847061	2,435583	0,306272
tjcdi	DLN	0,919147	3,34241956	1,401949	0,025756 *
tjpoup	DLN	0,914775	2,88690222	1,316055	0,038634 *
tjselic	DLN	0,920912	3,42542273	1,430419	0,025173 *
tjtr	DLN	0,934228	1,67777975	0,942665	0,088322
Arma		1,511207	-0,502497	0,461011	0,287007
Covariada benchmark					
desemprego	DLN	RMSE	0,235030878		

Nro total de observações: 131

Nro de observações fora da amostra: 24 = (2 períodos de 12 meses)

Tabela 6

Variável Dependente: INPC
 Covariada benchmark: desemprego
 Amostra completa: 1995 a 2005
 Fora da amostra: 2003:12 a 2005:11

Variável	Transf	MSE Relativo	Peso		P-Valor
Consumo e Venda					
consgas	DLN	1,036108	-0,18942335	1,353061	0,889882
consoleocomb	DLN	1,002496	0,42454781	1,452831	0,772737
consoleodiesel	DLN	1,033628	-1,61394754	2,033514	0,435498
spc	LN	1,325906	-1,92751256	0,757766	0,018152
vendareal	DLN	0,988903	0,62575213	0,606898	0,313231
Moeda e Crédito					
mcm0	DLN	0,991852	0,54951851	0,545699	0,324412
mcm1	DLN	0,865891	1,02584444	0,665133	0,136645
mcm2	DLN	0,983998	0,89736568	0,982178	0,370381
mcibvsp12	LV	1,21863	-0,58981207	0,724081	0,423675
mcrdv12	DT	1,048662	-0,32753966	1,111451	0,77087
mcvcdp12	DT	1,030258	-1,42332666	2,106042	0,505888
mcvcdv12	DT	1,038841	-0,92002812	1,89897	0,632622
Prod Industrial					
pi_bcons	DLN	1,035909	-1,50924442	1,96848	0,451055
pi_bcap	DT	0,936771	2,208702	0,936106	0,027166 *
pi_bconsdur	DLN	1,037429	-1,53539129	1,88244	0,423074
pi_bconsndur	LN	0,828615	1,71345155	1,077124	0,125315
pi_bint	DT	0,737452	1,16022352	0,505873	0,03129 *
pi_extmin	DT	1,096038	-0,06114451	0,739685	0,934835
pi_fumo	DT	0,827201	2,13687122	1,106223	0,065811 *
pi_indtransf	DT	0,989111	1,68787916	3,14358	0,596473
pi_insumos	DLN	0,89124	0,73633421	0,345834	0,044168 *
pi_leite	DLN	1,031185	-2,28300165	2,653599	0,39848
pi_madeira	DLN	1,119327	0,14302326	0,355463	0,691136
pi_maquequip	DLN	1,033895	0,15544062	0,524658	0,769682
pi_metal	DLN	1,031862	-0,1831776	0,954749	0,849536
pi_oleocomb	DT	1,05284	-0,99370191	1,976057	0,619836
pi_transporte	DLN	0,968146	0,99312064	0,919121	0,291115
prodpetrol	DLN	1,045387	-0,26290368	0,735127	0,723881
Salário					
sal_indn	DLN	1,061518	-0,38536637	1,401606	0,785812
sal_indr	DLN	1,042441	-0,1065851	1,549479	0,945753
Taxas de Câmbio					
tccompra	DLN	0,742074	0,65348723	0,261097	0,019869 *
tcvenda	DLN	0,696008	0,77691909	0,343769	0,033598 *
Taxas de Juros					
tjcd	DT	1,036519	-1,97598717	2,834354	0,492692
tjcdi	DLN	0,993206	0,76624332	1,893644	0,689481
tjpoup	DLN	0,957642	1,45085634	1,523023	0,350689
tjselic	DLN	0,99711	0,62410059	1,990929	0,756747
tjtr	DLN	0,977943	0,94496698	1,22839	0,44956
Arma		1,432464	-1,03176432	0,600982	0,099455
Covariada benchmark					
desemprego	DLN			RMSE	0,311553508

Nro total de observações: 131

Nro de observações fora da amostra: 24 = (2 períodos de 12 meses)

4 Previsão de Inflação: Modelos Multivariados

O modelo bivariado da seção anterior poderia ser estendido para um caso com vários regressores, no entanto, isso acarretaria overfitting e conseqüentemente um pior desempenho do modelo para previsões futuras.

Uma abordagem alternativa para construir previsões multivariadas, seria tratar como série de dados as previsões dos modelos bivariados (vistos na seção anterior) e combiná-las de alguma forma a obter uma previsão única. Um tipo de combinação poderia ser a média aritmética ou a mediana das previsões dos modelos bivariados, ou uma nova regressão da variável inflação sobre suas próprias previsões adquiridas com os modelos bivariados.

Considerando que temos n modelos de previsão de inflação - cada um originado a partir de uma covariada i ($i = 1, \dots, n$) - temos a seguinte equação de regressão:

$$\pi_{t+1} = \phi_i + \beta(L)_i x_{it} + \gamma(L)_i \pi_t + e_{it+1}$$

Esta expressão é a mesma utilizada na seção de modelos bivariados; apenas o índice i na equação a cima indica apenas que estamos analisando o modelo cujo regressor é dado por x_i .

A previsão deste modelo é indicada por:

$$f_{it} = \hat{\pi}_{t+1} = \hat{\phi}_i + \hat{\beta}_i(L) x_{it} + \hat{\gamma}_i(L) \pi_t$$

onde os coeficientes são estimados usando os dados disponíveis até a data t , e f_{it} representa a previsão do modelo que utiliza covariada i .

A partir das previsões (f_{it}) de cada modelo podemos realizar uma nova previsão combinando os diversos f_{it} . Assim, a equação abaixo indica uma forma de combinação genérica:

$$f_t = \sum w_{it} * f_{it}; \text{ somatório em } i; i = 1, \dots, n$$

Substituímos o termo de ponderação de três formas diferentes: primeiramente consideramos a previsão de cada modelo como igualmente satisfatória, ie. atribuímos mesmo peso a cada uma delas (neste caso $w_{it} = 1/n$). A previsão estimada foi uma média amostral das previsões dos n modelos. Em seguida usamos a mediana da previsão dos modelos estimados para determinar a mediana amostral da previsão e utilizamos a mediana como melhor previsão para inflação. A terceira forma utilizada para determinar os pesos foi a que deu nome a este estudo; esta forma associa a cada

ponderador um valor estimado por uma regressão chamada "ridge regression". O motivo do uso deste tipo de regressão (e não de uma regressão por mínimos quadrados) foi em razão das variáveis independentes apresentarem alto grau de multicolinearidade - já que todas são previsões de uma mesma variável dependente. A seguir descreveremos algumas características da "ridge regression" que validam seu uso neste experimento. O peso associado a cada previsão atribuído pelo estimador "ridge regression" é dado pela seguinte fórmula:

$$\hat{w}_{t,RR} = (cI_n + F'F)^{-1}((F'Y) + c/n)$$

onde:

- $\hat{w}_{t,RR}$ coeficiente estimado por "ridge regression"
- F matriz com as previsões (fit) dos melhores modelos (regressores da "ridge regression")
- Y vetor com a série de inflação observada
- n número de observações
- c $k * Traco(1/n * F'F)$
- k regula o grau de encolhimento; $k = 0 : \hat{w}_{RR} = \hat{w}_{ols}; k \rightarrow \infty : \hat{w}_{RR} \rightarrow 1/n$

O método de regressão por mínimos quadrados (OLS) é o procedimento estatístico mais frequentemente aplicado na estimação de parâmetros de um modelo genérico. Portanto, é bem documentado que este método é extremamente não confiável para a estimação de parâmetros quando as variáveis independentes são dependentes (ie. problema de multicolinearidade). Isto significa que, com dados multicolineares alguns coeficientes estimados por OLS podem ser estatisticamente insignificantes e/ou podem ter sinal contrário ao esperado.

Trabalhando no campo de engenharia Hoerl e Kennard (2000, 1970) perceberam que esse problema de colinearidade ocorria frequentemente o que os levou a propor um estimador alternativo que, apesar de viesado, pode ter um menor erro quadrático médio que o reportado por OLS. O estimador GRR (Generalized Ridge Regression) e uma versão simplificada deste - ORR (Ordinary Ridge Regression) - foram propostos por Hoerl e Kennard (2000).

A estimação por "ridge regression" é similar a estimação por mínimos quadrados ordinários, exceto por um pequeno valor constante que é adicionado a diagonal principal da matriz de variância-covariância dos estimadores. O objetivo desta alteração nos dados é criar um novo conjunto de informação com um menor grau de multicolinearidade.

Obviamente, um aspecto importante da "ridge regression" é determinar o melhor valor da constante que deve ser adicionado a diagonal principal da matriz de variância-covariância para otimizar a previsão (k). A forma típica usada para determinar este valor é um procedimento iterativo, no qual a cada momento se verifica o erro produzido para um determinado valor da constante. O melhor valor desta constante é o que está associado ao menor erro quadrático médio

da previsão. Usar valores altos sucessivos tipicamente diminui o erro quadrático médio até um certo ponto, quando a partir de então valores mais altos começam a aumentar o erro quadrático médio.

É importante perceber que a equação resultante da "ridge regression" é uma estimativa viesada e não reflete os parâmetros populacionais. Por isso a "ridge regression" é de pouco uso em modelos teóricos (Darlington, 1978), quando estamos interessados em uma interpretação sobre os valores dos parâmetros. A principal vantagem da "ridge regression" está na previsão, e como este é o propósito neste estudo, ela aparece como uma ferramenta particularmente útil. Além desta característica a "ridge regression" pode ser valiosa porque não necessita de amostras grandes como a regressão por mínimos quadrados, isto é, podemos usar amostras menores sem sacrificar o poder de previsão do modelo por "ridge regression".

Neste estudo utilizamos a "ridge regression" para a estimação dos parâmetros da regressão que combina as previsões de inflação, por percebermos que as covariadas desse modelo são altamente correlacionadas. No entanto, não podemos concluir que os estimadores de "ridge regression" são sempre melhores que outras combinações de previsão como a média, por nem sempre apresentarem resultado superior - em alguns testes foram equiparados a média, algumas vezes, um pouco inferior.

Testamos a estimação por "ridge regression" para os seguintes valores do parâmetro $k = 0, .25, .5, 1, 10, 100, 1000, 10000$; os resultados mais precisos (menor MSE) estão reportados abaixo nas *tabelas 7* (IPCA, pós-Plano Real) e *8* (INPC, pós-Plano Real). Para esses resultados utilizamos todas as 44 covariadas disponíveis e combinamos a previsão de todas elas.

A combinação utilizando estimadores de "ridge regression" apresenta melhor previsão (menor MSE relativo) e significativa a 5% que as demais combinações, tanto para a IPCA quanto para a o INPC. Fizemos o mesmo exercício considerando períodos mais curtos fora da amostra como 3 meses, e 6 meses e ainda assim obtivemos a "ridge regression" como melhor combinação de previsões.

No entanto, a combinação de previsões não foi melhor que algumas covariadas para explicar inflação; no caso, π_bint superou a "ridge regression" para o IPCA e covariadas de câmbio e novamente π_bint superaram as combinações no que diz respeito a explicação do INPC.

Tabela 7

Combinação das previsões de todas as covariadas

Variável Dependente: IPCA

Covariada benchmark: desemprego

Amostra completa: 1995 a 2005

Fora da amostra: 2003:12 a 2005:11

Variável	MSE Relativo		Peso		P-Valor
Combinação Média	0,901085	3,74304	1,296632	0,008324	
Combinação Mediana	0,943969	4,288654	1,614846	0,014124	
Combinação Ridge Regression	0,8207	1,603375	0,698099	0,031073	K min = 0.25

Tabela 8

Combinação das previsões de todas as covariadas

Variável Dependente: INPC

Covariada benchmark: desemprego

Amostra completa: 1995 a 2005

Fora da amostra: 2003:12 a 2005:11

Variável	MSE Relativo	Peso	P-Valor	
Combinação Média	0,933717	3,30214	2,093437	0,12836473
Combinação Mediana	1,002148	0,372053	2,273525	0,87143949
Combinação Ridge Regression	0,82439	3,212175	1,295091	0,02088208 K min = 0.25

Para o período pós fixação de metas de inflação a "ridge regression" apresentou resultados ainda mais precisos de previsão dentre as combinações testadas para ambas as séries de inflação. Os resultados estão listados nas tabelas abaixo:

Tabela 9

Combinação das previsões de todas as covariadas

Variável Dependente: IPCA

Covariada benchmark: desemprego

Amostra completa: 1999 a 2005

Fora da amostra: 2003:12 a 2005:11

Variável	MSE Relativo	Peso	P-Valor	
Combinação Média	0,848185	2,528985	1,09277	0,026643
Combinação Mediana	0,90476	4,382912	1,358055	0,002713
Combinação Ridge Regression	0,678739	0,868811	0,310098	0,008226 K min = 0.25

Tabela 10

Combinação das previsões de todas as covariadas

Variável Dependente: INPC

Covariada benchmark: desemprego

Amostra completa: 1999 a 2005

Fora da amostra: 2003:12 a 2005:11

Variável	MSE Relativo	Peso	P-Valor	
Combinação Média	0,73774	4,182746	0,824812	1,29E-05
Combinação Mediana	0,788062	4,638409	0,969292	3,06E-05
Combinação Ridge Regression	0,737944	4,183865	0,826178	1,32E-05 K min = 10000

Não incluímos neste documento as tabelas que possuem análise de modelos bivariados a partir de 1999, mas estas também foram analisadas e pudemos concluir que a combinação de previsões perde para alguns poucos modelos bivariados como verificamos anteriormente para o período pós-Plano Real.

5 Conclusão

Alguns pontos precisam ser ressaltados. Primeiramente, como estamos interessados em previsão estamos comparando modelos com melhor performance fora-da-amostra. Para evitar "overfitting" e conseqüentemente previsões pobres em tempo real utilizamos o critério de informação de Schwartz que penaliza cada novo regressor adicionado a equação. Os modelos selecionados foram os que apresentaram melhor relação custo benefício entre número de regressores e erro amostral.

Em seguida, devemos lembrar que consideramos apenas modelos lineares de regressão, isto é, modelos em que a inflação e a covariada possuem relação linear; no caso desta relação ser não linear estes resultados subestimam uma melhoria da previsão que poderia ser obtida neste último caso quando comparada a previsão da curva de Phillips.

Após a análise empírica podemos constatar que a previsão da curva de Phillips convencional nem sempre foi o melhor previsor para a inflação; outras covariadas são igualmente úteis na análise de previsibilidade e apresentaram menor erro de previsão que a covariada desemprego para o período analisado. Destacamos a covariada *htrab* como a mais precisa em todos os testes mesmo quando consideramos 12, 36, 48, 50 observações fora-da-amostra no período pós plano real. É importante destacar que ao variarmos o número de observações fora-da-amostra, diferentes covariadas apresentaram desempenho de previsão melhor que desemprego para o horizonte de um ano, mas somente *htrab* manteve comportamento constante.

Além disso, a combinação de previsões de todas as covariadas nos permite concluir que combinações utilizando "ridge regression" tem performance melhor ou igual as demais técnicas de combinação mencionadas no estudo. E uma performance sempre melhor quando foi testada previsão para um horizonte temporal mais curto como trimestral e semestral.

6 Apêndice

6.1 Descrição das séries utilizadas

Nesta seção apresentamos as séries utilizadas como covariadas nos modelos de regressão apresentados no texto. Todas estas séries foram otidas do site [ipeadata](http://ipeadata.gov.br); são mensais e representam o período de 1995 à 2005.

Taxas de Juros

tjcdB	Taxa de Juros: Rendimentos nominais das principais aplicações financeiras (CDB).
tjcdi	Taxa de Juros: Certificado de Depósito Interbancário (CDI).
tjlp	Taxa de Juros de Longo Prazo (TJLP).
tjpoup	Taxa de Juros: refere-se ao rendimento do primeiro dia útil do mês da poupança.
tjselic	Taxa de Juros: média do mês (Selic).
tjtr	Taxa Referencial de juros (TR).

Moeda e Crédito

mcm0	Estoque de moeda m0.
mcm1	Estoque de moeda m1.
mcm2	Estoque de moeda m2.
mcibvsp12	Variação mensal do Índice da Bolsa de Valores de São Paulo (Ibovespa).
mcrdv12	Depósito à vista. Resgate no mês.
mcvcdp12	Depósitos a prazo. Velocidade de circulação dos principais ativos financeiros
mcvcdv12	Depósitos à vista. Velocidade de circulação dos principais ativos financeiros.

Taxa de Câmbio

tccompra	Taxa de câmbio - R\$ / US\$ - comercial - compra - média
tcvenda	Taxa de câmbio - R\$ / US\$ - comercial - venda - fim período

Salário

rm_nsp	Rendimento médio - nominal - assalariados - trabalho principal - SP
folhapgto	Folha de pagamentos - indústria - índice (jan. 2003 = 100) - SP
sal_indn	Salário nominal - médio - indústria - índice (jan. 2003 = 100) - SP
sal_in dr	Salário real - médio - indústria - índice (jan. 2003 = 100) - SP
rm_rsp	Rendimento médio real dos assalariados no trabalho principal na RMSP

Índice de Preços

inpc	Inflação - INPC - (% a.m.)
igpm	Inflação - IGP-M - (% a.m.)
ipca	Inflação - IPCA - (% a.m.)

Produção Industrial

pi_bcap	Produção industrial - bens de capital - quantum - índice dessaz.
pi_bint	Produção industrial - bens intermediários - quantum - índice dessaz.
pi_bcons	Produção industrial - bens de consumo - quantum - índice dessaz.
pi_bconsdur	Produção industrial - bens de consumo duráveis - quantum - índice dessaz.
pi_bconsndur	Produção industrial - bens de consumo não duráveis - quantum - índice dessaz.
pi_ig	Produção industrial - indústria geral - quantum - índice dessaz.
pi_extmin	Produção industrial - extrativa mineral - quantum - índice dessaz.
pi_indtransf	Produção industrial - indústria de transformação - quantum - índice dessaz.
pi_fumo	Produção industrial - fumo - quantum - índice (média 2002 = 100)
pi_madeira	Produção industrial - madeira - quantum - índice (média 2002 = 100)
pi_metal	Produção industrial - metalurgia básica - quantum - índice (média 2002 = 100)
pi_maquequip	Produção industrial - máquinas e equipamentos - quantum - índice
pi_insumos	Produção industrial - insumos - construção civil - quantum
pi_transporte	Produção industrial - automóveis, caminhões, ônibus - quantum - nível 80
pi_oleocomb	Produção industrial - óleos combustíveis - quantum - nível 80
pi_leite	Produção industrial - leite beneficiado - quantum - nível 80
prodpetrol	Produção de petróleo - média - qde./dia - Barril(Mil) - ANP - Anp12_PDPET12
utilcapacinst	Utilização da capacidade instalada - indústria - (%) - CNI - Cni12_NUCAP12

Emprego

htrab	Horas trabalhadas - indústria - índice (média 1992 = 100) - CNI - Cni12_HTRAB12
desemprego	Taxa de desemprego - RMSP - (%) - Seade e Dieese/PED - SEADE12_TDTGSP12

Consumo e Venda

consgas	Consumo aparente - gasolina - média - qde./dia - Barril(Mil)
consoleocomb	Consumo aparente - óleo combustível - média - qde./dia - Barril(Mil)
consoleodiesel	Consumo aparente - óleo diesel - média - qde./dia - Barril(Mil)
vendanom	Vendas nominais - indústria - índice (média 1992 = 100)
vendareal	Vendas reais - indústria - índice (média 1992 = 100)
spc	SPC - número de consultas -

7 Bibliografia

References

- [1] Chauvet, M (2000) "Leading Indicators of Inflation for Brazil" working paper 7, Banco Central do Brasil
- [2] Enders, W. (2003) "Applied Econometric Time Series"
- [3] Granger, C. W and P. Newbold (1976), "Forecasting Economic Time Series", Academic Press: New York.
- [4] Hamilton, J, 1994, "Time Series Analysis", Princeton University Press
- [5] Hoerl, A.E. and Kennard, R. W. (1970a). "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, 12, 55-67.
- [6] Ibrahim, M and J. Zhang (2005) "A simulation study on SPSS ridge regression and ordinary least squares regression procedures for multicollinearity data", *Journal of Applied Statistics*, 2005, vol. 32, issue 6, pages 571-588
- [7] Minella, A e I Goldfajn e A. Fraga "Inflation Targeting in Emerging Market Economies" working paper 76, Banco Central do Brasil.
- [8] Silva, M e M. Portugal "A Recente Experiência Brasileira com Metas de Inflação: uma avaliação preliminar"
- [9] Stock, James and Mark W. Watson. "Forecasting Inflation," *Journal of Monetary Economics*, 1999, v44(2,Oct), 293-335
- [10] Werlang, S e A. Tombini e J. Bogdanski "Implementing Inflation Targeting in Brazil" working paper 1, Banco Central do Brasil.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)