

TAXA DE JURO DOS NOVOS EMPRÉSTIMOS PARA AQUISIÇÃO DE HABITAÇÃO

PREVISÃO DE UMA SÉRIE NÃO SAZONAL COM BASE EM MODELOS UNIVARIADOS

Ricardo Martinho

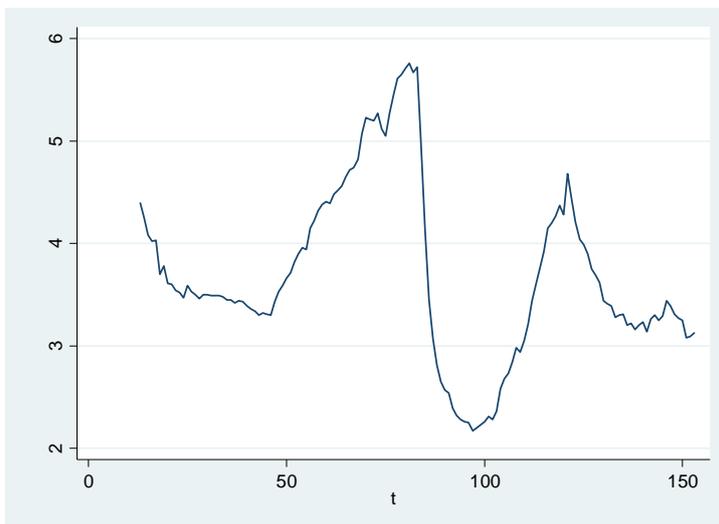
09/01/2015

Definição do problema, descrição dos dados e breve revisão da literatura

Neste trabalho pretende-se modelar o comportamento das taxas de juro sobre as novas operações de empréstimos para aquisição de habitação com o objectivo de obter uma previsão para a evolução futura da taxa de juro média, num horizonte temporal de 12 meses, com base em modelos univariados, determinísticos e estocásticos. Este pode ser um tema de interesse para diversas entidades, tais como: bancos, supervisores bancários, Ministério das Finanças, investidores no mercado de imobiliário, entre outros. Os dados utilizados estão disponíveis desde Janeiro de 2003 e seguem uma metodologia comum a outros países da área do euro, que pondera os empréstimos individuais por montante e ignora diferenças de maturidade/prazo contratual. Em particular, neste trabalho utilizaram-se 141 observações mensais, correspondentes ao período entre Janeiro de 2003 e Setembro de 2014, sendo que para estimação dos modelos apenas foram tidos em conta dados até Dezembro de 2013. Existe uma extensa literatura sobre modelos empíricos para previsão de taxas de juro de empréstimos à habitação e/ou de outros segmentos de crédito (empresas, consumo, automóvel, outros fins). Em todo o caso, as abordagens mais comuns tendem a aplicar modelos multivariados mais complexos, conjugando relações de cointegração entre taxas juro, volumes de crédito e um conjunto de variáveis macroeconómicas tidas como relevantes com dinâmicas de curto prazo, incluindo mecanismos correctores de erro.

Caracterização da série a modelar

Gráfico 1. Y_t : Taxa de juro média dos novos empréstimos para aquisição de habitação (Janeiro de 2003 - Setembro de 2014).

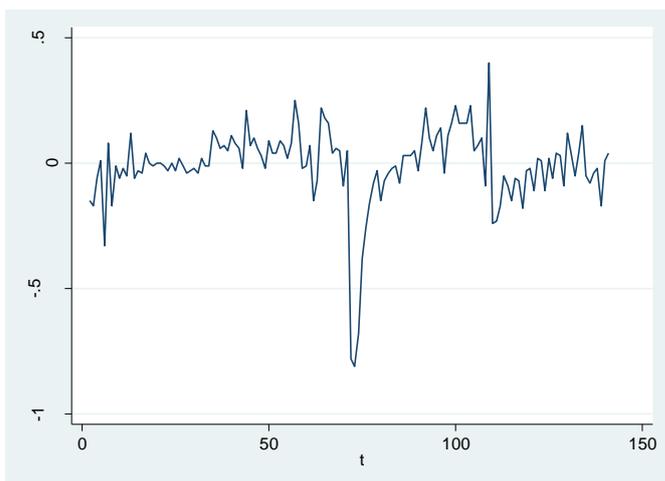


A análise gráfica sugere uma série não estacionária em média, sem estrutura sazonal, sendo possível sinalizar momentos de alteração abrupta no comportamento série. O correlograma da série original (ver Anexo 1, Fig. 1) revela uma elevada persistência, com um coeficiente de autocorrelação no primeiro lag próximo de 1 e com um decaimento lento para zero. O

resultado do teste de Dickey-Fuller para raiz unitária não permite rejeitar a hipótese nula (de existência de uma raiz unitária) para os níveis de significância convencionais, quer na versão simples, quer na versão com tendência (*Anexo, Fig 2*).

Neste contexto, procedeu-se à aplicação de uma diferença simples com objectivo de resolver o problema da elevada persistência da série. A série obtida (*Gráfico 2*) pode ser interpretada como a variação da taxa de juro dos empréstimos para aquisição de habitação (em pontos percentuais). Neste caso, os testes Dickey-Fuller rejeitam a hipótese de raiz unitária (*anexo, Fig. 4*). O correlograma da série transformada X_t é apresentado na Fig. 6, em anexo. A FAC apresenta um decaimento amortecido para zero, sugerindo um modelo univariado do tipo ARIMA que incorpore uma componente autoregressiva. Por sua vez, a FACP apresenta uma queda acentuada no 2º lag e uma queda para zero no 3º lag, o que sugere um AR(1) ou AR(2). Nos desfasamentos associados aos meses homólogos não existem correlações significativas, o que é consistente com a ausência de um comportamento sazonal.

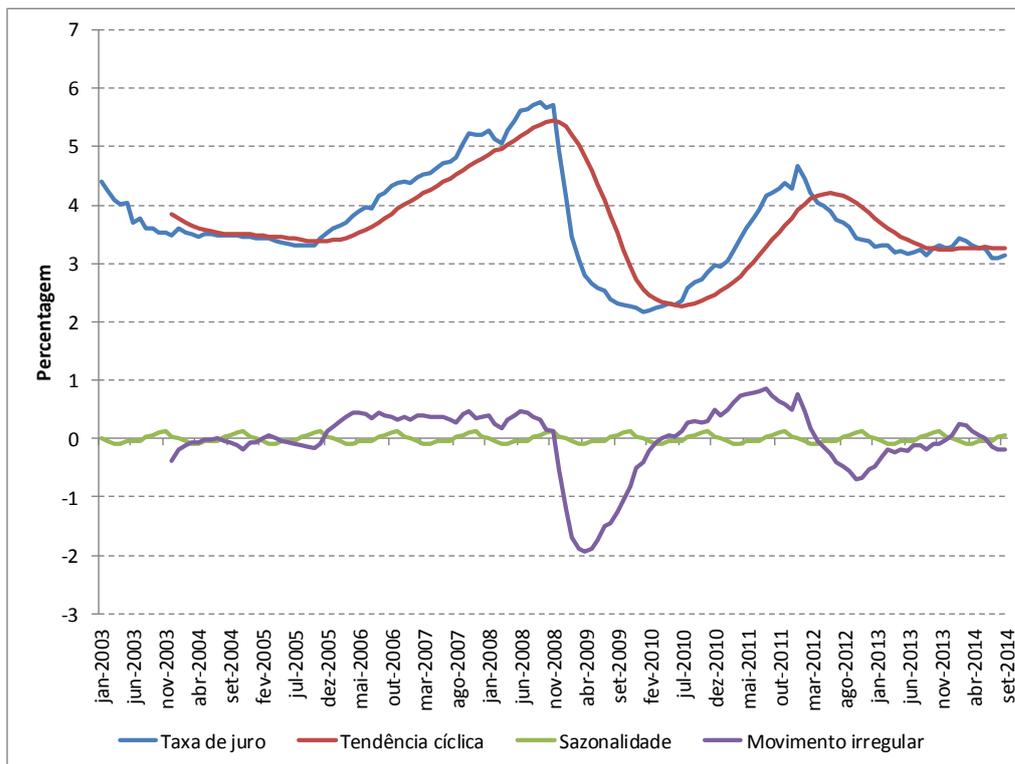
Gráfico 2. $X_t=(1-B)Y_t$. Variação da taxa de juro dos empréstimos para aquisição de habitação (em pontos percentuais)



Decomposição das forças componentes

A decomposição da série em tendência cíclica, componente sazonal e movimentos irregulares foi efectuada pelo método aditivo dada a aparente ausência de sazonalidade e facto dos movimentos irregulares não variarem com nível da série. A tendência cíclica foi obtida através de uma média móvel não centrada a 12 meses. No período em análise, destaque-se a existência de alterações significativas na tendência da série (*Gráfico 3*). Dada a simplicidade da metodologia subjacente, a queda abrupta no final de 2008, não sequência dos sucessivos cortes na taxa de juro de referência do Banco Central Europeu, é capturada com algum desfasamento pela tendência cíclica, o que se reflecte em variações significativas na componente de movimentos irregulares. Refira-se ainda que a componente sazonal é virtualmente inexistente.

Gráfico 3. Modelo aditivo: $Y_t=f(TC_t, S_t, E_t)= TC_t+S_t+E_t$



Método de alisamento exponencial (Holt)

Dado que a série não apresenta carácter sazonal foi aplicado o método de Holt, definido pelas equações de actualização:

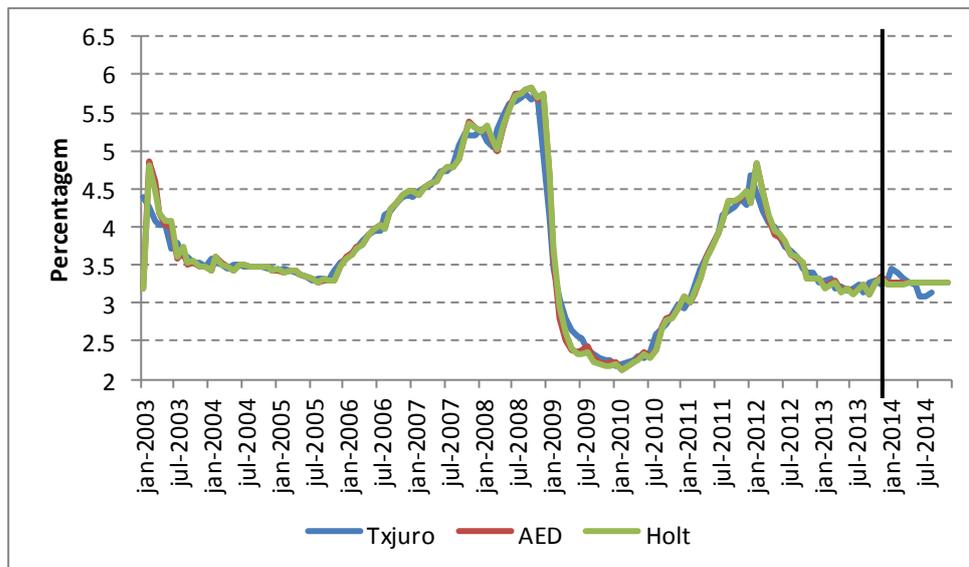
$$a(t)=\alpha Y_t+(1-\alpha)[a(t-1)+b(t-1)], 0<\alpha<1$$

$$b(t)=\beta[a(t)-a(t-1)]+(1-\beta)b(t-1), 0<\beta<1$$

$$P_{t+h}=a(t)+b(t) \times h$$

A estimação que teve em consideração 132 observações até dez-2013, resultando em $\alpha=1$, $\beta=0.3236$ e com $RMSE=0.18386$. Alternativamente, o método de alisamento exponencial duplo, um caso especial do método de Holt, resultou em $\alpha=0.6802$ e num $RMSE$ ligeiramente superior 0.18646. O gráfico 4 mostra as previsões *in-sample* obtidas com base nos dois métodos até dez-2013, sendo que a partir de Jan-2014 os valores dizem respeito a previsões *out-of-sample*, um passo à frente. É interessante verificar que a previsão para os meses de 2014 é virtualmente constante, o que é justificado pelo facto da tendência da série original ser praticamente nula nos últimos meses utilizados na estimação.

Gráfico 4. Séries obtidas pelos métodos de alisamento exponencial.



Modelos estocásticos ARMA/ARIMA

De acordo com a caracterização da desenvolvida anteriormente foi estimado inicialmente um modelo ARIMA (2,1,0) com os seguintes resultados:

Quadro 1. Resultados da estimação do modelo proposto pela análise do correlograma da série transformada. ARIMA (2,1,0)

ARIMA regression

| | | | |
|---------------------------|---------------|---|--------|
| Sample: 2 - 132 | Number of obs | = | 131 |
| | Wald chi2(2) | = | 151.06 |
| Log likelihood = 83.22802 | Prob > chi2 | = | 0.0000 |

| D.txjuro | OPG | | | | |
|----------|-----------|-----------|-------|-------|----------------------|
| | Coef. | Std. Err. | z | P> z | [95% Conf. Interval] |
| txjuro | | | | | |
| _cons | -.0120197 | .0399274 | -0.30 | 0.763 | -.0902759 .0662365 |
| ARMA | | | | | |
| ar | | | | | |
| L1. | .4633843 | .0384639 | 12.05 | 0.000 | .3879964 .5387722 |
| L2. | .2062638 | .0815459 | 2.53 | 0.011 | .0464367 .3660908 |
| /sigma | .1279473 | .0037485 | 34.13 | 0.000 | .1206003 .1352943 |

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at zero.

Os parâmetros associados aos termos autoregressivos são estatisticamente significativos para um nível de significância de 5% em ambos os casos. Adicionalmente, a constante é não significativa, consistente com uma variação média da taxa de juro próxima de zero no período considerado. O correlograma dos resíduos (Quadro 2) confirma que estamos na presença de um ruído branco, com o *p-value* de 0.63 (63%) para o teste de significação global da FAC dos resíduos, e, portanto, na presença de um modelo com as propriedades estatísticas desejáveis.

Quadro 2. Correlograma dos resíduos do modelo ARIMA (2,1,0)

. corrgram erro_txjuro_arima_is_1

| LAG | AC | PAC | Q | Prob>Q | -1 | 0 | 1 | -1 | 0 | 1 |
|-----|---------|---------|--------|--------|--------|-------------------|-------------------|----|---|---|
| | | | | | Prob>Q | [Autocorrelation] | [Partial Autocor] | | | |
| 1 | 0.0015 | 0.0014 | .00029 | 0.9864 | | | | | | |
| 2 | 0.0017 | 0.0017 | .0007 | 0.9996 | | | | | | |
| 3 | 0.0158 | 0.0161 | .03451 | 0.9983 | | | | | | |
| 4 | 0.0239 | 0.0242 | .11271 | 0.9985 | | | | | | |
| 5 | -0.0273 | -0.0284 | .21592 | 0.9989 | | | | | | |
| 6 | -0.0175 | -0.0170 | .2588 | 0.9997 | | | | | | |
| 7 | -0.0157 | -0.0150 | .29333 | 0.9999 | | | | | | |
| 8 | -0.0081 | -0.0074 | .30273 | 1.0000 | | | | | | |
| 9 | -0.0988 | -0.1003 | 1.6968 | 0.9954 | | | | | | |
| 10 | 0.0574 | 0.0693 | 2.1713 | 0.9949 | | | | | | |
| 11 | 0.0408 | 0.0348 | 2.413 | 0.9964 | | | | | | |
| 12 | -0.0632 | -0.0670 | 2.9974 | 0.9956 | | | | | | |
| 13 | 0.1371 | 0.1559 | 5.773 | 0.9540 | | | | | | |
| 14 | 0.0386 | 0.0134 | 5.9951 | 0.9666 | | | | | | |
| 15 | -0.1569 | -0.1849 | 9.6903 | 0.8388 | | | | | | |
| 16 | -0.1061 | -0.0816 | 11.397 | 0.7843 | | | | | | |
| 17 | 0.0242 | 0.0354 | 11.486 | 0.8302 | | | | | | |
| 18 | 0.0118 | -0.0107 | 11.508 | 0.8716 | | | | | | |
| 19 | -0.1544 | -0.1378 | 15.215 | 0.7088 | | | | | | |
| 20 | -0.1056 | -0.0927 | 16.965 | 0.6553 | | | | | | |
| 21 | 0.0953 | 0.0994 | 18.405 | 0.6233 | | | | | | |
| 22 | -0.0300 | -0.0091 | 18.548 | 0.6730 | | | | | | |
| 23 | -0.0431 | -0.0969 | 18.848 | 0.7100 | | | | | | |
| 24 | -0.0149 | -0.0652 | 18.884 | 0.7581 | | | | | | |
| 25 | 0.0537 | 0.1196 | 19.358 | 0.7797 | | | | | | |
| 26 | -0.1018 | -0.1757 | 21.077 | 0.7380 | | | | | | |
| 27 | -0.0215 | -0.0633 | 21.154 | 0.7790 | | | | | | |
| 28 | -0.0915 | -0.1179 | 22.572 | 0.7542 | | | | | | |
| 29 | -0.0869 | -0.0212 | 23.86 | 0.7358 | | | | | | |
| 30 | -0.0429 | -0.0564 | 24.178 | 0.7639 | | | | | | |
| 31 | 0.0434 | -0.0467 | 24.505 | 0.7894 | | | | | | |
| 32 | -0.0839 | -0.1009 | 25.745 | 0.7748 | | | | | | |
| 33 | -0.0444 | 0.0174 | 26.095 | 0.7979 | | | | | | |
| 34 | -0.0058 | -0.0955 | 26.101 | 0.8318 | | | | | | |
| 35 | 0.0406 | -0.0991 | 26.4 | 0.8520 | | | | | | |
| 36 | -0.0595 | -0.0714 | 27.049 | 0.8593 | | | | | | |
| 37 | -0.1106 | -0.1472 | 29.317 | 0.8119 | | | | | | |
| 38 | 0.1818 | 0.2135 | 35.506 | 0.5853 | | | | | | |
| 39 | 0.0666 | -0.0029 | 36.346 | 0.5916 | | | | | | |
| 40 | -0.0240 | -0.0335 | 36.457 | 0.6305 | | | | | | |

Adicionalmente, foram estimados diversos modelos na vizinhança do proposto com base na análise do correlograma, isto é, com +1/-1 termo autoregressivo ou de médias móveis. O quadro resumo da qualidade de ajustamento dos diversos modelos é apresentado abaixo. O modelo sugerido pela análise do correlograma – ARIMA (2,1,0) é o que revela melhor qualidade de ajustamento, quer à luz dos critérios de informação de Akaike e Schwartz, quer de acordo com o erro quadrático médio. Os modelos com +1 parâmetro na especificação apresentam um valor virtualmente igual da função de verosimilhança, mas introdução de um parâmetro adicional é penalizada pelos critérios de informação, não justificando a sua adopção. O modelo ARIMA (0,1,0), usualmente conhecido como *random-walk*, é o que apresenta de longe uma pior qualidade de ajustamento. Refira-se ainda que no teste à significância global da FAC dos resíduos, a hipótese nula de estarmos na presença de um ruído branco apenas é rejeitada no ARIMA (0,1,0).

Quadro 3. Comparação da qualidade de ajustamento in-sample

| | AIC | BIC | loglikelihood | MSE | RMSE |
|---------------|----------|----------|---------------|----------|----------|
| ARIMA (2,1,0) | -158.456 | -146.955 | 83.22802 | 0.016424 | 0.128157 |
| ARIMA (1,1,0) | -154.74 | -146.114 | 80.36997 | 0.017161 | 0.130999 |
| ARIMA (2,1,1) | -156.456 | -142.08 | 83.22803 | 0.016424 | 0.128158 |
| ARIMA (3,1,0) | -156.456 | -142.08 | 83.22803 | 0.016424 | 0.128158 |
| ARIMA (0,1,0) | -101.837 | -96.0862 | 52.91827 | 0.026101 | 0.161557 |

A detecção de *outliers* foi efectuada recorrendo ao *software* TRAMO, disponível no Eviews. Foi detectado um *outlier* do tipo aditivo (AO) em Junho de 2003 e dois *outliers* do tipo *temporary change* (TC) em Novembro de 2008 e Janeiro de 2012. No primeiro caso, observa-se uma queda mensal da taxa de juro significativa entre dois períodos de relativa estabilização. No caso dos TCs as datas identificadas correspondem ao início das duas fases mais notórias de diminuição da taxa de juro. O modelo escolhido pelo TRAMO foi o ARIMA (1,1,1) com $\phi=-0.81$ e $\theta=-0.21$, ambos significativos para um nível de significância de 5%, o que representa uma diferença significativa face aos parâmetros positivos do ARIMA (2,1,0) escolhido sem correcção de *outliers*.

Os modelos de intervenção/detecção de *outliers* são úteis para melhorar a precisão das estimativas após uma intervenção. No entanto, têm pouca utilidade na previsão *out-of-sample*, na medida em que não é possível estimar os parâmetros do modelo de intervenção antes da mesma ocorrer.

Previsão out-of-sample

Muitas vezes, mais do que escolher um modelo com boa qualidade de ajustamento e propriedades desejáveis do ponto de vista estatístico (e.g. ausência de autocorrelação nos resíduos) é importante escolher um modelo que funcione bem fora da amostra. Uma necessidade de muitas entidades é a de obter, no início do ano, uma projecção para a evolução de uma variável até ao final do ano, i.e. previsões $P_t(1)$ até $P_t(12)$. Neste contexto, foram estimadas versões sucessivas dos 5 modelos ARIMA identificados anteriormente, a cada 12 meses, para incorporar as “novas” observações disponíveis. A primeira ronda foi estimada

com observações até dez-2007¹, a partir da qual se efectuaram previsões *out-of-sample* de a 1 até 12 passos à frente. As previsões obtidas foram confrontadas com os valores reais ao longo de 2008, tendo sido calculados os respectivos erros de previsão. O processo foi depois repetido para observações até dez-2009, dez-2010,... dez-2013². O modelo com melhor capacidade de previsão *out-of sample* no período em análise é o ARIMA (2,1,1), enquanto que o modelo com melhor ajustamento *in-sample*, identificado anteriormente, é o ARIMA(2,1,0). Note-se, contudo, a heterogeneidade observada ao longo do tempo, com o modelo mais simples (ARIMA 0,1,0) a apresentar a melhor performance para prever o ano de 2014 (Janeiro a Setembro, apenas). Focando agora a análise no período completo, concluiu-se, através da ferramenta de optimização do Excel, que não existe nenhuma combinação de dois ou mais dos cinco modelos apresentados, com pesos inversamente proporcionais aos respectivos erros quadráticos médios (ou mesmo com pesos iguais), que produza melhores resultados que o modelo ARIMA (2,1,1). Isto não quer, contudo, dizer que não exista uma infinidade de combinações com pesos totalmente arbitrários, e com pouco significado económico, capaz de produzir melhores resultados. Refira-se ainda que, um exercício análogo, avaliado à luz do Erro Absoluto Médio, produziu exactamente o mesmo resultado para o conjunto do período (2008-2014).

Quadro 4. Comparação das previsões *out-of-sample*. De 1 a 12 passos à frente. Erro quadrático médio.

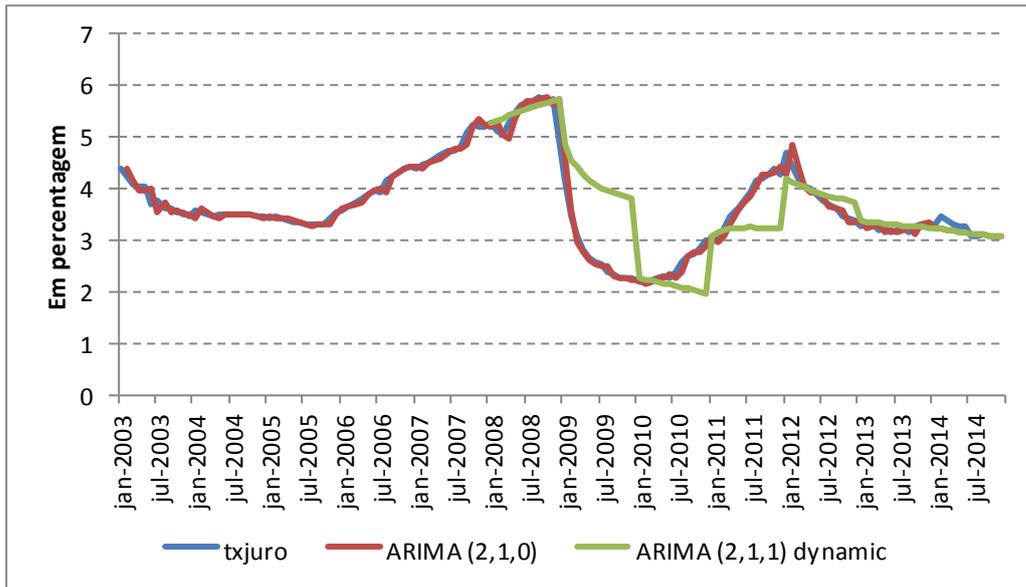
| RMSE | arima(2,1,0) dynamic | arima(1,1,0) dynamic | arima(2,1,1) dynamic | arima(3,1,0) dynamic | arima(0,1,0) dynamic |
|--------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| 2008 | 0.360 | 0.311 | 0.267 | 0.335 | 0.303 |
| 2009 | 1.542 | 2.096 | 1.441 | 1.454 | 2.334 |
| 2010 | 0.515 | 0.516 | 0.498 | 0.512 | 0.524 |
| 2011 | 0.721 | 0.792 | 0.719 | 0.762 | 1.035 |
| 2012 | 0.415 | 0.410 | 0.259 | 0.406 | 0.555 |
| 2013 | 0.082 | 0.093 | 0.090 | 0.090 | 0.112 |
| 2014 | 0.127 | 0.137 | 0.127 | 0.127 | 0.111 |
| (2008-2014) | 0.718 | 0.909 | 0.667 | 0.694 | 1.034 |

Por fim apresenta-se um gráfico comparativo entre a série original, a série com as previsões *in-sample* do modelo ARIMA (2,1,0) e a série com as previsões *out-of-sample* do modelo ARIMA(2,1,1). De facto, observa-se que o erro de projecção é significativamente elevado no caso do exercício de estimação de 1 até 12 passos à frente, e que o erro tende a aumentar de mês para mês, dentro do ano de projecção. A ausência de uma tendência definida e de uma estrutura sazonal marcada, reflecte-se na dificuldade de previsão da taxa de juro média dos empréstimos à habitação com base nos valores da própria série (modelos univariados), mesmo num horizonte temporal até 1 ano.

¹ Para que existisse um mínimo de 60 observações na primeira estimação. No limite este processo poderia ser repetido mês após mês, mesmo com o objectivo de prever 1 a 12 passos à frente. Os modelos de ajustamento exponencial não foram incluídos neste exercício pelo facto de o STATA não conseguir convergir para uma solução em algumas iterações do exercício.

² Para 2014 apenas foram calculados em erros de previsão até Setembro.

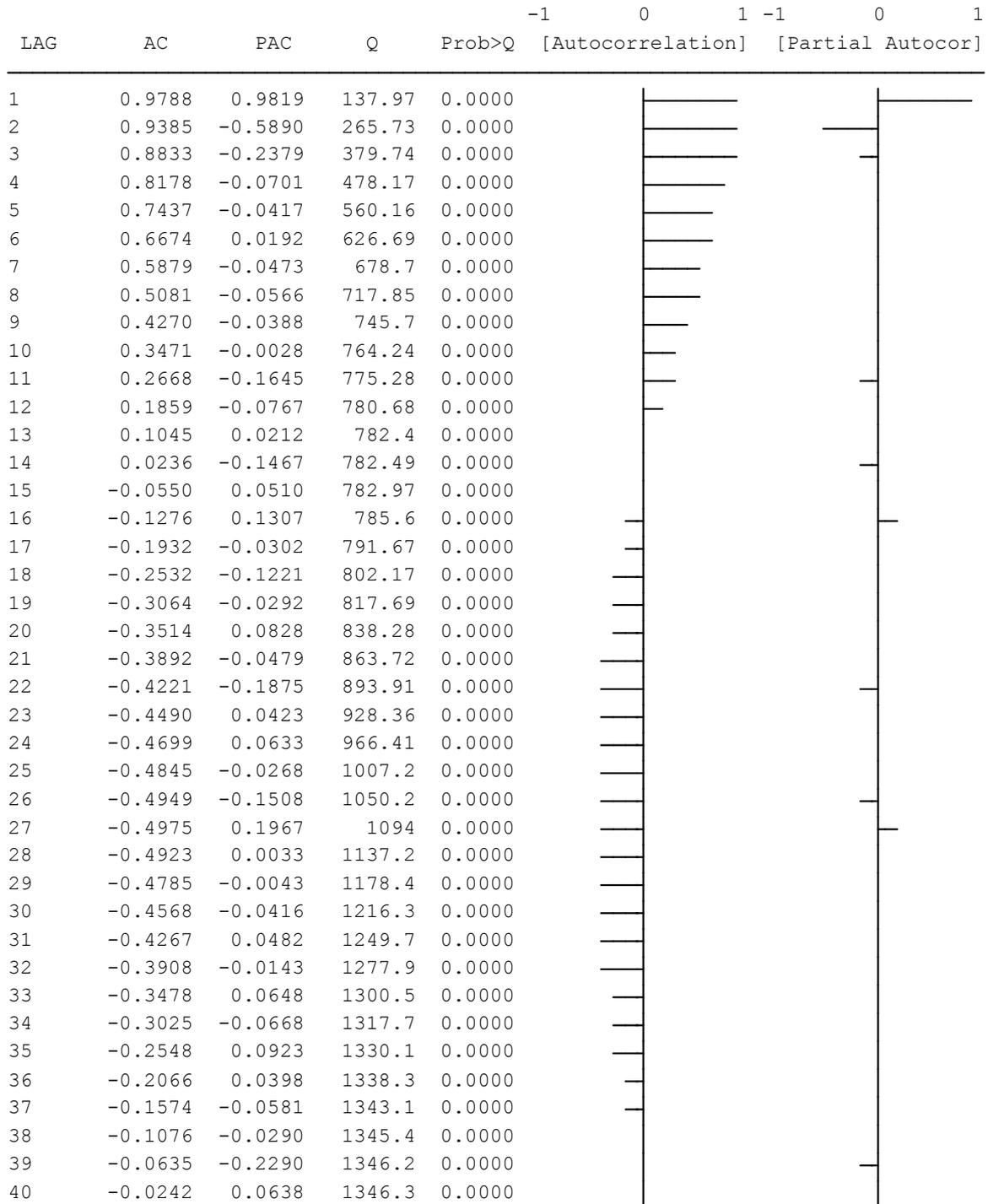
Gráfico 5. Série original, previsões in-sample e previsões out-of-sample, 1 a 12 passos à frente.



Anexo 1

Fig. 1 Correlograma da variável original: Y_t

. corrgram txjuro



.

Fig. 3 Correlograma da variável transformada: X_t

. corrgram d.txjuro

