

Previsão contemporânea das exportações portuguesas de turismo

Sónia Cabral
Banco de Portugal

Cláudia Duarte
Banco de Portugal

Outubro de 2016

Resumo

Dada a crescente importância da monitorização contínua da atividade económica, técnicas que permitam o aproveitamento da divulgação atempada de dados em frequência alta desempenham um papel fundamental na previsão de curto prazo. Este artigo compara duas abordagens uni-equacionais, nomeadamente os tradicionais modelos de ligação (*bridge models*) e as mais recentes regressões *Mixed Data Sampling* (MIDAS), para produzir uma previsão contemporânea (*nowcasting*) das exportações trimestrais de turismo em Portugal. São consideradas diferentes especificações de modelos de ligação e MIDAS, bem como combinações de previsões num exercício recursivo em pseudo tempo real. Os resultados sugerem que devem ser utilizados indicadores de curto prazo para prever as exportações de turismo. As regressões MIDAS tendem a superar as equações de ligação, especialmente quando menos informação do trimestre corrente está disponível. Os melhores resultados são sempre obtidos a partir de uma combinação de previsões de uma especificação MIDAS com dinâmica auto-regressiva. (JEL: C53, F47, Z39)

Introdução

O turismo é o setor mais importante do comércio internacional de serviços em Portugal e tem dado um contributo significativo para o excedente da balança de transações internacionais de serviços nas últimas duas décadas (Gráfico 1). Ainda que a importância das exportações de outros serviços tenha aumentado progressivamente ao longo do tempo, as exportações nominais de turismo ainda representavam mais de 45 por cento do total de exportações de serviços e mais de 15 por cento do total de exportações portuguesas de bens e serviços em 2015. Adicionalmente, as exportações portuguesas de turismo têm crescido de forma acentuada nos últimos anos, aumentando cerca de 50 por cento de 2010 a 2015. Neste

Agradecimentos: Os autores agradecem ao Instituto Nacional de Estatística (INE - <http://ine.pt>) o acesso à série das exportações trimestrais de turismo a preços constantes utilizada neste artigo. Os autores agradecem igualmente a António Rua os seus comentários e sugestões, e os esclarecimentos de Carla Ferreira e Ana Mouta sobre os dados. Eventuais erros e omissões são da exclusiva responsabilidade dos autores. As opiniões expressas no artigo são da responsabilidade dos autores, não coincidindo necessariamente com as do Banco de Portugal ou do Eurosistema.

E-mail: scabral@bportugal.pt; cfduarte@bportugal.pt

contexto, as exportações nominais de turismo representavam 6.3 por cento do PIB em 2015 e o excedente da balança de viagens e turismo ascendeu a mais de 4 por cento do PIB em 2015, o valor mais elevado das últimas duas décadas.

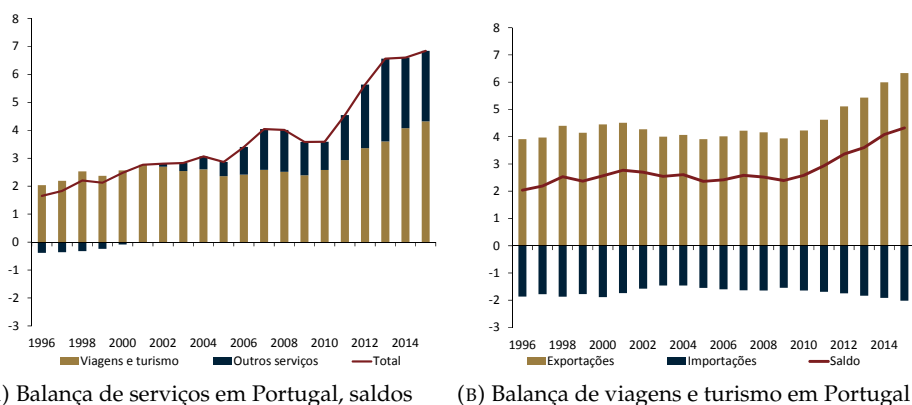


GRÁFICO 1: Exportações portuguesas de turismo

Notas: Gráficos em percentagem do PIB nominal.

Fontes: Instituto Nacional de Estatística (INE) e Banco de Portugal.

Comparando com outros países da União Europeia (UE), a importância económica do setor do turismo em Portugal é também evidente (Gráfico 2). O rácio das receitas de turismo internacional relativamente ao PIB aumentou de 4.8 por cento na média dos anos 1995-2000 para 6.3 por cento no período 2009-2014. Este rácio do PIB corresponde a mais do dobro da média da UE e apenas é superado por outros seis países da UE, a maioria deles com economias tipicamente associadas a exportações significativas de turismo.

Como o turismo contribui significativamente para o crescimento da economia portuguesa, previsões precisas sobre a procura de turismo são particularmente importantes. O cálculo de previsões atempadas normalmente requer a identificação de variáveis que, não só acrescentam informação útil, mas também são publicadas mais cedo. O objetivo deste artigo é utilizar indicadores mensais de curto prazo para a previsão contemporânea (*nowcasting*) do crescimento real das exportações de turismo das contas nacionais trimestrais portuguesas. O princípio básico assenta na utilização de informação que é publicada mais cedo e com maior frequência do que a variável de interesse a fim de obter projeções antes de esta ser observada.

Considerando que estamos interessados em prever uma variável trimestral numa base mensal, a expressão “previsão contemporânea” (*nowcasting*) refere-se, em geral, às projeções mensais do trimestre corrente e, portanto, para cada trimestre, há pelo menos 3 projeções distintas, uma efetuada em cada mês do trimestre. Neste artigo, definimos “previsão contemporânea” como as projeções de um trimestre desde o primeiro mês

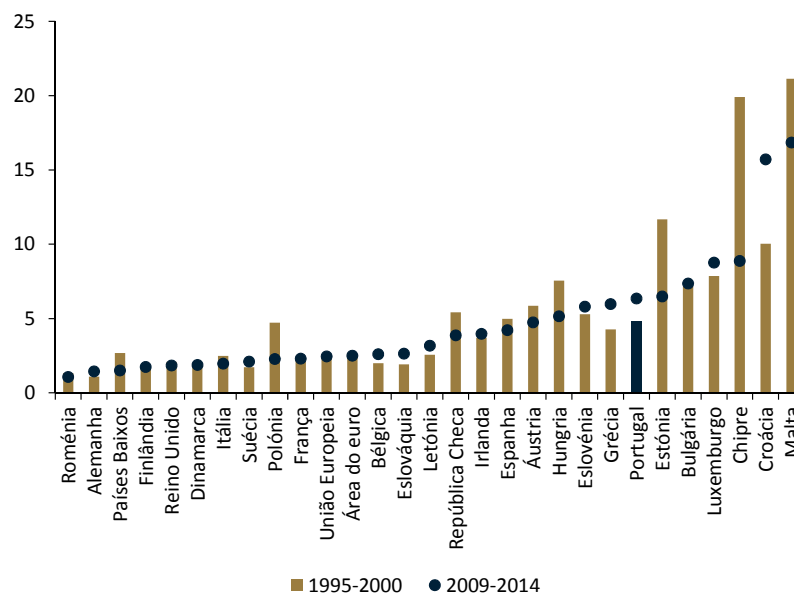


GRÁFICO 2: Receitas de turismo internacional

Notas: Receitas de turismo internacional, em percentagem do PIB (dólares norte-americanos, preços correntes). As receitas de turismo internacional incluem os gastos efetuados por visitantes internacionais, incluindo pagamentos a companhias aéreas nacionais para o transporte internacional. Estas receitas incluem qualquer outro pré-pagamento feito por bens ou serviços recebidos no país de destino.

Fonte: *The World Bank - World Development Indicators (WDI)*.

desse trimestre até ao momento em que os números oficiais são divulgados. Dado que as contas nacionais trimestrais portuguesas estão normalmente disponíveis 60 dias após o final do trimestre de referência, neste artigo são produzidas 5 projeções distintas para cada trimestre. A amostra cobre o período abrangido entre Outubro de 2000 e Março de 2016 e são utilizados modelos de forma reduzida numa abordagem puramente de séries temporais para prever as exportações de turismo, num exercício recursivo em pseudo tempo real, que replica o padrão de publicação dos indicadores de maior frequência verificado em tempo real. O artigo vai para além dos tradicionais modelos de ligação (*bridge models*) e também aplica as regressões *Mixed Data Sampling* (MIDAS) propostas por Ghysels *et al.* (2007). Em comparação com outros modelos de frequência mista, a abordagem MIDAS é apelativa porque é uma técnica uni-equacional simples, flexível e parcimoniosa. Tanto quanto sabemos, esta é a primeira aplicação de regressões MIDAS à previsão de curto prazo das exportações trimestrais de turismo.

Os resultados obtidos mostram que, em geral, a utilização de indicadores de curto prazo para a previsão contemporânea das exportações de turismo

é útil, uma vez que proporciona projeções mais precisas do que as de um modelo de referência univariado, nos períodos considerados. Os modelos MIDAS tendem a apresentar melhores resultados do que as tradicionais equações de ligação, especialmente quando há menos dados disponíveis sobre os indicadores no trimestre corrente. Combinar previsões é uma boa estratégia independentemente do modelo de frequência mista considerado, pois permite melhorar face à referência univariada e face aos modelos com apenas um indicador. Considerando todos os modelos e combinações de projeções, o melhor resultado em qualquer período é sempre obtido a partir de uma combinação de previsões de um modelo MIDAS com dinâmica auto-regressiva. Em suma, apesar da amostra utilizada para previsão ser relativamente curta, este artigo oferece evidência robusta de uma melhoria na precisão de previsão através da combinação de projeções de modelos MIDAS.

O artigo está organizado da seguinte forma. A secção 2 discute alguns dos estudos que enquadram este trabalho e destaca as suas principais contribuições para a literatura. A secção 3 apresenta resumidamente as equações de ligação e as regressões MIDAS. A secção 4 descreve as variáveis e o exercício empírico de previsão contemporânea das exportações portuguesas de turismo. A secção 5 discute os resultados do exercício realizado. Finalmente, a secção 6 apresenta algumas conclusões.

Literatura relacionada

A modelação e previsão da procura de turismo tem sido uma importante área de investigação nas últimas décadas, tendo surgido na literatura um conjunto de novos métodos e técnicas (ver Song e Li (2008) para um resumo dos estudos publicados após 2000). Uma edição especial do *International Journal of Forecasting* fornece uma descrição útil e detalhada dos desenvolvimentos recentes na previsão de turismo (ver Song e Hyndman (2011) para uma introdução a esta edição especial e aos trabalhos aí incluídos). A revisão da vasta literatura empírica sobre modelação e previsão de turismo está fora do âmbito deste artigo. Em alternativa, esta secção oferece uma lista não exaustiva de referências em diferentes vertentes da literatura, que estão diretamente relacionadas com o nosso estudo e fornecem um enquadramento da nossa análise, com um foco especial na economia portuguesa.

Em geral, a literatura sobre previsão de turismo ainda é dominada por dois métodos principais: modelos de séries temporais não-causais e abordagens econométricas causais. O nosso estudo enquadra-se na primeira categoria. Athanasopoulos *et al.* (2011) concluem que abordagens puras de séries temporais produzem previsões da procura de turismo mais precisas do que os métodos alternativos. No entanto, esta conclusão não é unânime na literatura, pois há vários resultados discordantes, especialmente quando

se utilizam modelos causais mais sofisticados (ver, por exemplo, Song *et al.* (2011)).

No contexto dos modelos de séries temporais, este trabalho contribui para uma análise global da atividade turística em Portugal. Dada a relevância do setor do turismo, existem diversos estudos nesta área, cujo foco é a economia portuguesa. No entanto, a literatura sobre o comércio internacional português de serviços de turismo ainda é limitada quando comparada com o grande número de estudos sobre o comércio internacional português de mercadorias. Algumas exceções são Daniel e Ramos (2002) que realizam uma análise econométrica do número de turistas oriundos de cinco origens distintas, utilizando métodos de cointegração e modelos com mecanismo corretor do erro, e Teixeira e Fernandes (2012, 2014) que usam modelos de redes neurais artificiais para prever as receitas de turismo e as dormidas na hotelaria em Portugal. Utilizando dados mensais sobre as dormidas nos estabelecimentos hoteleiros, Gouveia e Rodrigues (2005) aplicam um método não paramétrico para identificar os ciclos de crescimento do turismo, concluindo que existe um desfasamento entre os ciclos da procura de turismo e os ciclos económicos, e Rodrigues e Gouveia (2004) usam um modelo auto-regressivo periódico parcimonioso, mostrando a sua superioridade no desempenho preditivo em comparação com outros modelos. Andraz *et al.* (2009) utilizam um modelo de índice de difusão para prever a procura turística no Algarve oriunda do Reino Unido e confirmam o seu melhor desempenho preditivo. Mais recentemente, Serra *et al.* (2014) usam modelos dinâmicos de dados de painel para modelar a procura internacional de turismo em sete regiões turísticas portuguesas, encontrando um comportamento heterogéneo por região.

No nosso caso, a análise centra-se na previsão contemporânea da evolução das exportações trimestrais de turismo, em termos reais. Desta forma, este artigo está também relacionado com uma vasta literatura empírica sobre previsão de curto prazo de variáveis económicas e, em particular, sobre a previsão de componentes do PIB do lado da procura em Portugal. Neste contexto, as equações de ligação são uma das técnicas mais usadas para lidar com bases de dados de frequência mista. Normalmente, estas equações ligam variáveis mensais e trimestrais que têm uma correlação significativa e a escolha dos regressores tende a ter em conta o facto estarem disponíveis mais cedo (ver, por exemplo, Baffigi *et al.* (2004)). Esteves e Rua (2012) fornecem uma descrição geral da metodologia do exercício de previsão de curto prazo do Banco de Portugal, onde os modelos de ligação são a ferramenta de modelação preferencial. Outras aplicações de modelos de ligação em exercícios de previsão de curto prazo da economia portuguesa incluem Cardoso e Duarte (2006) para as exportações de bens, Maria e Serra (2008) para o investimento e Esteves (2009) para o consumo privado.

Para além da abordagem tradicional de modelos de ligação, também são consideradas regressões MIDAS. Ao aplicar esta técnica, este artigo contribui para uma vertente recente da literatura empírica que utiliza modelos

MIDAS para lidar com diferentes frequências de amostragem e a divulgação assíncrona de dados. Inspiradas nos modelos de desfasamento distribuído, as regressões MIDAS são muito flexíveis, podendo considerar diferentes frequências, diferentes polinómios de agregação e diferentes horizontes de previsão (para um breve resumo dos principais tópicos relacionados com a modelação MIDAS, ver Andreou *et al.* 2011). Recentemente, Duarte *et al.* (2016) utilizam modelos MIDAS para a previsão de curto prazo do consumo privado trimestral em Portugal. Tanto quanto sabemos, a abordagem MIDAS não foi ainda aplicada à previsão de curto prazo das exportações trimestrais de turismo e este artigo visa preencher essa lacuna.

Equações de ligação e regressões MIDAS

A existência de informação atempada sobre o estado da economia é fundamental para a análise e elaboração de políticas. No entanto, algumas estatísticas oficiais importantes, como as contas nacionais, só estão disponíveis numa base trimestral e com desfasamentos relevantes na divulgação. Por exemplo, a estimativa rápida para o PIB português está disponível 45 dias após o final do trimestre, enquanto os principais agregados do lado da despesa estão disponíveis 60 dias após o final do trimestre de referência. Neste contexto, técnicas que incorporam dados de frequência mista são ferramentas úteis para tirar proveito do grande número de indicadores de curto prazo relevantes, permitindo uma avaliação atempada da situação económica corrente. Na literatura têm sido propostos diversos tipos de ferramentas econométricas para combinar dados com diferentes frequências temporais e explorar divulgações mais atempadas dos dados de maior frequência, de modo a melhorar a precisão das previsões; Foroni e Marcellino (2014) descrevem sucintamente as principais abordagens. Esta secção centra-se em duas abordagens econométricas específicas, que utilizam dados de frequência mista de uma forma simples e apelativa: equações de ligação e regressões MIDAS (ver, por exemplo, Schumacher (2016) para uma comparação recente destes modelos).

Equações de ligação

As equações de ligação são uma das técnicas mais utilizadas para ligar dados com diferentes frequências temporais. Normalmente, as séries com maior frequência temporal são, em primeiro lugar, agregadas para a (menor) frequência da variável dependente e, em seguida, incluídas em modelos de previsão tradicionais. Estes modelos têm sido amplamente utilizados na literatura, especialmente por instituições nacionais e internacionais para a previsão do crescimento do PIB (e.g., Baffigi *et al.* 2004, Diron 2008, Barhoumi *et al.* 2012 e Bulligan *et al.* 2015).

Considerando y_t numa frequência trimestral (intervalo de referência) como variável dependente, a especificação de uma equação de ligação simples com um único indicador e com termos auto-regressivos é dada por:

$$y_{t+h} = \beta_0 + \beta(L)x_t^Q + \gamma(L)y_t + \varepsilon_{t+h}, \quad (1)$$

em que o preditor x_t^Q é uma variável trimestral obtida através da agregação da sua contraparte em frequência alta $x_t^{(m)}$, m vezes mais frequente (por exemplo, para dados mensais m é igual 3), h é o horizonte trimestral, e ε_{t+h} é o habitual termo de erro i.i.d. O polinómio de desfasamento trimestral $\beta(L)$ de ordem k é definido como $\beta(L) = \sum_{i=0}^k \beta_{i+1}L^i$, com $Lx_t^Q = x_{t-1}^Q$. Da mesma forma, $\gamma(L)$ é um polinómio de ordem p no operador de desfasamento definido como $\gamma(L) = \sum_{i=1}^p \gamma_i L^i$, em que p é o número de termos auto-regressivos, e $Ly_t = y_{t-1}$. A equação 1 pode ser facilmente alargada a um formato multivariado, através da simples introdução de regressores adicionais, sendo que cada um pode ter um polinómio $\beta(L)$ distinto.

Dependendo dos prazos de publicação dos dados, os indicadores com maior frequência podem ter que ser prolongados com estimativas, antes de serem agregados temporalmente e incluídos no modelo de ligação. Considerando dados trimestrais e mensais, as estimativas para as observações mensais em falta, obtidas a partir de modelos univariados simples, são justapostas aos dados mensais observados, sendo as séries mensais transformadas em séries trimestrais e, em seguida, utilizadas para a previsão no modelo de ligação trimestral.

Regressões MIDAS

Esta secção apresenta uma descrição geral das regressões MIDAS utilizadas neste artigo. Armesto *et al.* (2010) fornecem uma introdução simples e intuitiva a este assunto e discussões mais abrangentes sobre regressões MIDAS no contexto da previsão de curto prazo podem ser encontradas em Andreou *et al.* (2011), Foroni e Marcellino (2014), Schumacher (2016) e nas referências aí incluídas. Finalmente, uma edição recente do *Journal of Econometrics* (Ghysels e Marcellino 2016) discute em detalhe vários métodos econométricos que permitem a utilização de dados de frequência mista.

As regressões MIDAS, introduzidas por Ghysels *et al.* (2004), são uma ferramenta de previsão direta para vários horizontes, inspirada nos modelos de desfasamento distribuído. Adicionalmente, como discutido em Duarte (2014), um termo auto-regressivo pode facilmente ser adicionado à equação MIDAS. Considere-se novamente y_t numa frequência trimestral e $x_t^{(m)}$ disponível m vezes mais frequentemente. Uma regressão MIDAS simples com termos auto-regressivos tem a seguinte forma:

$$y_{t+h} = \beta_0 + \beta_1 B(L^{1/m}; \theta)x_t^{(m)} + \gamma(L)y_t + \varepsilon_{t+h}, \quad (2)$$

na qual h é o horizonte trimestral, $B(L^{1/m}; \theta) = \sum_{j=0}^{jmax} B(j; \theta)L^{j/m}$ é um polinómio de dimensão $jmax$ no operador $L^{1/m}$, $B(j; \theta)$ representa o esquema de ponderação utilizado para a agregação, que se assume ser normalizado para 1, $L^{j/m}x_t^{(m)} = x_{t-j/m}^{(m)}$, e ε_{t+h} é o habitual termo de erro i.i.d.

Se bem que a ordem do polinómio $B(L^{1/m}; \theta)$ possa ser potencialmente infinita, algumas restrições têm de ser impostas por motivos de exequibilidade. Numa regressão MIDAS, os coeficientes $B(L^{1/m}; \theta)$ são captados por uma função polinomial conhecida $B(j; \theta)$, que depende apenas de um conjunto reduzido de parâmetros, resumidos no vetor θ . Os modelos MIDAS são, portanto, parcimoniosos, o que é uma das principais características desta técnica.

A literatura sugere algumas alternativas para a função polinomial de agregação; ver, nomeadamente, Ghysels *et al.* (2007). O polinómio mais utilizado é o polinómio de defasamento exponencial de Almon:

$$B(k; \theta_1, \theta_2) := \frac{e^{(\theta_1 k + \theta_2 k^2)}}{\sum_{k=1}^K e^{(\theta_1 k + \theta_2 k^2)}}, \quad (3)$$

em que $f(q, \theta_1, \theta_2) = (q^{\theta_1 - 1}(1 - q)^{\theta_2 - 1}\Gamma(\theta_1 + \theta_2))/(\Gamma(\theta_1)\Gamma(\theta_2))$ e $\Gamma(\theta) = \int_0^\infty e^{-k} k^{\theta - 1} dk$. Uma vez que o polinómio exponencial de Almon tem uma especificação funcional não-linear, as regressões MIDAS têm que ser estimadas utilizando métodos não-lineares, nomeadamente os mínimos quadrados não-lineares.

Uma variante MIDAS discutida por Chen e Ghysels (2011) é o MIDAS multiplicativo (M-MIDAS) que está mais perto dos esquemas de agregação tradicionais. Em vez de agregar todos os defasamentos da variável com maior frequência num único agregado, os modelos MIDAS multiplicativos incluem agregados com m elementos de frequência alta, assim como os seus defasamentos, i.e.,

$$y_{t+h} = \beta_0 + \sum_{i=1}^p \beta_i x_{t-i+1}^{mult} + \gamma(L)y_t + \varepsilon_{t+h}, \quad (4)$$

em que $x_t^{mult} = \sum_{j=0}^{m-1} B(j; \theta)L^{j/m}x_t^{(m)}$.

Uma abordagem MIDAS distinta é o MIDAS irrestrito (U-MIDAS, do inglês *unrestricted MIDAS*) proposto por Foroni *et al.* (2015):

$$\begin{aligned} y_{t+h} &= \beta_0 + B_u(L^{1/m})x_t^{(m)} + \gamma(L)y_t + \varepsilon_{t+h} \\ &= \beta_0 + \sum_{j=0}^J \beta_{j+1} L^{j/m} x_t^{(m)} + \gamma(L)y_t + \varepsilon_{t+h} \\ &= \beta_0 + \beta_1 x_t^{(m)} + \beta_2 x_{t-1/m}^{(m)} + \dots + \beta_{J+1} x_{t-J/m}^{(m)} + \gamma(L)y_t + \varepsilon_{t+h}. \end{aligned} \quad (5)$$

A regressão U-MIDAS não recorre a funções polinomiais de desfasamento distribuído, tendo, por isso, a vantagem de poder ser estimada por OLS. No entanto, dada a proliferação de parâmetros, os modelos U-MIDAS são mais indicados para lidar com dados mensais do que com dados semanais ou diários, uma vez que grandes diferenças de frequências de amostragem entre as variáveis consideradas são muito penalizadas em termos de parcimónia.

Finalmente, Clements e Galvão (2008) sugerem uma forma alternativa de introduzir dinâmica auto-regressiva em regressões MIDAS. Os autores interpretam a dinâmica de y_t como um fator comum, assumindo como hipótese que y_{t+h} e $x_t^{(m)}$ partilham a mesma dinâmica auto-regressiva. Considere-se uma regressão MIDAS simples em que o termo de erro pode ser representado por um modelo auto-regressivo de ordem 1. O modelo MIDAS de fator comum (CF-MIDAS, do inglês *common factor* MIDAS) pode ser escrito como:

$$(1 - \gamma L)y_t = \beta_0(1 - \gamma) + \beta_1(1 - \gamma L)B(L^{1/m}; \theta)x_t^{(m)} + \varepsilon_t. \quad (6)$$

Embora o trabalho inicial de Clements e Galvão (2008) considere apenas um único termo auto-regressivo, é possível aplicar esta técnica a casos com mais termos auto-regressivos.

Em resumo, os modelos MIDAS usam esquemas de ponderação mais flexíveis do que os modelos tradicionais de frequência baixa e tendem a ser mais parcimoniosos. Os modelos MIDAS também permitem incorporar facilmente as divulgações atempadas de dados de maior frequência. Na equação 2, assume-se que todas as observações de maior frequência de $x_t^{(m)}$ ao longo do período de referência (menor frequência) são conhecidas. Considerando dados trimestrais e mensais, isso significa que já estão disponíveis os três meses de informação sobre o indicador de curto prazo para o trimestre de interesse. Se em vez de um trimestre completo, por exemplo, estiver disponível apenas o primeiro mês, então a regressão MIDAS pode ser escrita como:

$$y_{t+h} = \beta_0 + \beta_1 B(L^{1/3}; \theta)x_{t-2/3}^{(3)} + \gamma(L)y_t + \varepsilon_{t+h}. \quad (7)$$

Adicionalmente, as regressões MIDAS podem também incluir mais do que um indicador de frequência alta, e, em alguns casos, sem necessidade de estimar muitos parâmetros adicionais. Além disso, podem considerar-se diferentes polinómios $B(L^{1/m}; \theta)$ para cada regressor.

Dados e descrição do exercício

Dados

A variável dependente são as exportações portuguesas de turismo das contas nacionais trimestrais a preços constantes e ajustadas de sazonalidade e efeitos

de calendário. Ao longo deste artigo, as exportações de turismo referem-se ao conceito do Sistema de Contas Nacionais de despesa de consumo final das famílias em turismo por parte de visitantes não residentes em Portugal, não incluindo o consumo intermédio de turismo associado a viagens de negócios de não-residentes.¹

Quatro tipos de variáveis de curto prazo relacionados com as exportações de turismo são publicados mensalmente e, portanto, foram a base dos quatro indicadores individuais incluídos neste exercício de previsão contemporânea das exportações trimestrais de turismo.

Em primeiro lugar, este artigo utiliza as exportações nominais (créditos) da balança de viagens e turismo da Balança de Pagamentos (BoP) portuguesa, deflacionadas com o Índice Harmonizado de Preços no Consumidor (IHPC) total.²

Em segundo lugar, consideramos as transações com cartões emitidos no exterior efetuadas em terminais localizados em Portugal (ATM/POS). Estas operações incluem quer levantamentos de dinheiro em caixas automáticos (ATM, do inglês *Automated Teller Machines*), quer transações em terminais de pagamento automático (POS, do inglês *Points of Sale*) e estão disponíveis desde Setembro de 2000. Os valores das transações mensais ATM/POS foram deflacionados utilizando o IHPC total.³

Em terceiro lugar, outro indicador é o número de dormidas de não-residentes em estabelecimentos hoteleiros em Portugal. Para levar em conta potenciais efeitos de qualidade, o número de dormidas em cada tipo de estabelecimento hoteleiro foi ponderado pelos respetivos proveitos totais médios no ano anterior. Foram considerados cinco tipos individuais distintos de estabelecimentos hoteleiros (hotéis, pousadas, hotéis-apartamentos, aldeamentos turísticos, apartamentos turísticos) e uma categoria residual (incluindo pensões, estalagens e motéis).⁴

Finalmente, foi calculado um índice composto de confiança dos consumidores em alguns dos principais países de origem dos turistas –

1. Os dados detalhados sobre as exportações de turismo foram gentilmente cedidos pelo Instituto Nacional de Estatística (INE - <http://ine.pt/>).

2. Dois outros deflatores foram também testados para deflacionar os dados da BoP. Em primeiro lugar, foi utilizado o IHPC para o agregado de serviços. Em segundo lugar, um deflator composto foi construído através da ponderação de várias componentes de preços pelo seu peso nas despesas de turistas em Portugal. Optámos por usar o IHPC total, dado que teve o melhor desempenho, mas os resultados não mudam qualitativamente com os dois deflatores alternativos.

3. Tal como para as exportações nominais de turismo da balança de viagens e turismo, foram também considerados dois deflatores alternativos para as transações ATM/POS (ver nota de rodapé 2 para mais detalhes), tendo os resultados permanecido, de uma forma geral, praticamente inalterados.

4. Foram igualmente testados os dados brutos do total de dormidas de não-residentes, mas o seu desempenho de previsão não foi melhor.

Espanha, Reino Unido, França, Alemanha, Itália e Países Baixos. Os inquéritos de opinião são particularmente interessantes por estarem disponíveis mais cedo, sendo os primeiros dados mensais relativos ao trimestre corrente a serem divulgados. O indicador mensal de confiança dos consumidores de cada país, publicado pela Direção-Geral dos Assuntos Económicos e Financeiros (DG ECFIN) da Comissão Europeia, foi ponderado pela importância relativa desse país nas dormidas em Portugal de não-residentes no ano anterior.⁵

Quando necessário, as séries mensais foram corrigidas de sazonalidade e ajustadas de efeitos de calendário. Aplicou-se o mesmo procedimento usado pelo Instituto Nacional de Estatística para o ajustamento sazonal das estatísticas mensais oficiais, ou seja, o X-13 ARIMA com ajustamento de efeitos de calendário através do programa JDemetra+ fornecido pelo Eurostat. A amostra começa em Outubro de 2000, que corresponde ao primeiro mês do primeiro trimestre para o qual as transações ATM/POS estão disponíveis, e termina em Março de 2016. Com exceção do indicador de confiança, as séries originais foram transformadas em taxas de variação homóloga. No caso do indicador de confiança, foram utilizadas diferenças absolutas em relação ao mesmo período do ano anterior.

Descrição do exercício

O objetivo deste artigo é a previsão contemporânea do crescimento trimestral das exportações reais de turismo em Portugal usando quatro indicadores mensais distintos. Para tal, foi implementado um exercício direto para vários horizontes, recursivo e em pseudo tempo real com as seguintes características.

Todos os modelos de ligação e MIDAS foram recursivamente estimados com amostra crescente (*expanding window*) e selecionados usando o critério de informação Bayesiano (BIC). Partindo de uma primeira sub-amostra (do quarto trimestre de 2000 ao quarto trimestre de 2007) que foi utilizada para a especificação dos modelos dentro da amostra (*in-sample*), a amostra de estimação é sucessivamente expandida através da adição de uma nova observação.⁶ Conforme cada nova observação é adicionada à amostra, todos os modelos são re-estimados e, assim, os coeficientes podem variar ao longo do tempo. Relativamente ao exercício de previsão fora da amostra (*out-of-sample*), a amostra de avaliação começa no primeiro trimestre de 2008 e termina no primeiro trimestre 2016.

5. Foram igualmente utilizados os indicadores de confiança dos consumidores, quer para a UE, quer para a área do euro, publicados mensalmente pela DG ECFIN e os resultados foram qualitativamente semelhantes.

6. Também testámos uma amostra rolante (*rolling window*) e os principais resultados sobre as diferenças entre os modelos de ligação e MIDAS mantêm-se relativamente inalterados.

Foram utilizados diferentes desfasamentos (até 3 trimestres), também para os termos auto-regressivos. Os modelos MIDAS foram estimados usando o polinómio exponencial de Almon definido na equação 3.⁷ As equações de ligação e os diferentes modelos MIDAS descritos na secção 3 foram estimados com e sem termos auto-regressivos.

Uma seleção adequada dos preditores é crucial para a obtenção dos melhores resultados de previsão ao longo dos períodos considerados. Tendo em conta que, no nosso caso, o conjunto de informação contém um pequeno número de variáveis foram considerados quer modelos com apenas um indicador, quer modelos multivariados. Para além disso, foi testada uma estratégia alternativa que pode permitir melhorar a precisão das previsões: a combinação de previsões (*forecast pooling*). Diferentes métodos de combinação de previsões estão disponíveis na literatura, desde pesos simples e iguais (e constantes) até pesos baseados no desempenho. Como esquemas simples de agregação tendem a ter bons desempenhos, neste artigo são utilizadas duas técnicas de combinação distintas: a média simples, com pesos iguais, e a média ponderada pelo erro quadrático médio de previsão (MSFE) descontado, proposta por Stock e Watson (2004). Os pesos de Stock e Watson (2004) são os seguintes:

$$w_{it} = \frac{m_{it}^{-1}}{\sum_{i=1}^n m_{it}^{-1}} \quad m_{it} = \sum_{s=t_0}^T \delta^{T-s} (y_s - \hat{y}_s^i)^2, \quad (8)$$

em que \hat{y}^i são as previsões do modelo i e δ é o fator de desconto. Os pesos desta técnica de combinação dependem inversamente do desempenho de previsão de cada modelo no passado. Assim, quanto maior for o MSFE de uma previsão individual, menor será o peso que lhe está associado.⁸

Os dados utilizados correspondem à versão final, ou seja à última versão disponível quando a base de dados foi construída. No caso dos indicadores de confiança dos consumidores e das transações ATM/POS, a versão final é igual às várias versões sucessivamente divulgadas ao longo do período coberto pela amostra (dados em tempo real), uma vez que estas séries normalmente não são revistas. As revisões das exportações da BoP, das dormidas e das exportações trimestrais de turismo não são consideradas nesta análise, mas são geralmente relativamente pequenas em Portugal, pelo que o seu impacto deve ser reduzido.

A ausência de sincronia nos calendários de publicação das séries de maior frequência implica painéis desequilibrados, havendo diferentes padrões

7. O tradicional polinómio de desfasamento de Almon também foi testado como alternativa para a função polinomial. No entanto, o desempenho dos modelos não melhorou.

8. Em relação ao parâmetro de desconto, foram considerados valores alternativos, tendo a opção de não-desconto ($\delta = 1$) apresentado os melhores resultados.

de valores em falta no final da amostra (o chamado problema "ragged-edge"). Existe evidência na literatura de que considerar estes diferentes padrões pode ter um impacto considerável na precisão de previsão (ver, por exemplo, Giannone *et al.* (2008)). Por isso, neste artigo levamos em conta esta característica importante dos dados macroeconômicos em tempo real. Seguindo Foroni e Marcellino (2014) e Schumacher (2016), o nosso exercício em pseudo tempo real replica o padrão de publicação dos indicadores, conforme eles se tornam disponíveis em tempo real. Mais especificamente, o padrão de valores em falta é replicado sucessivamente em cada uma das sub-amostras, de acordo com um calendário de publicação estilizado, ou seja, para cada série observa-se o número de valores em falta no final da amostra e impõe-se o mesmo número de observações em falta em cada iteração.

Como discutido em Banbura *et al.* (2011), uma característica importante de um exercício de previsão contemporânea é que raramente se executa apenas uma única projeção para um determinado trimestre, mas sim uma sequência de projeções que são atualizadas conforme existem novos dados. Assim, considerando previsões de variáveis trimestrais numa base mensal, normalmente a previsão contemporânea refere-se às projeções mensais do trimestre corrente e há pelo menos 3 projeções distintas para esse trimestre (uma em cada mês do trimestre). No entanto, tendo em conta os desfasamentos na divulgação da variável de interesse, o número de projeções produzidas antes dos dados serem observados pode aumentar. Por exemplo, Banbura *et al.* (2013) produzem previsões contemporâneas para o PIB dos USA a partir do primeiro mês do trimestre corrente até o primeiro mês do trimestre seguinte, quando os dados oficiais são publicados.

O nosso exercício segue igualmente esta perspectiva mais ampla sobre previsão contemporânea. Assim, a partir do final do primeiro mês do trimestre t até ao final do segundo mês de trimestre $t + 1$, quando se observam os dados oficiais, é possível obter até 5 projeções diferentes para o trimestre t , dependendo dos preditores considerados e da informação intra-trimestral disponível para cada preditor. Dado que todos os indicadores mensais são tipicamente observados antes do dia de publicação das contas nacionais trimestrais portuguesas (relembramos que o desfasamento de publicação dos agregados do lado da despesa é de 60 dias após o final do trimestre de referência), nos meses da sua publicação, ou seja, Fevereiro, Maio, Agosto e Novembro, é possível obter uma estimativa antecipada para as exportações de turismo antes de o número oficial se tornar disponível.

Um exemplo simples pode ajudar a esclarecer a estrutura da base de dados no nosso exercício em pseudo tempo real. Suponha que estamos interessados em obter uma projeção do crescimento real das exportações portuguesas de turismo no primeiro trimestre de 2016 (Q_t). No final de Janeiro de 2016 (1º m Q_t), o indicador de confiança dos consumidores está disponível para Janeiro, mas não há informação do trimestre corrente para as outras variáveis: as transações ATM/POS estão disponíveis para Dezembro

de 2015 e quer as exportações da BoP, quer as dormidas estão disponíveis para Novembro de 2015. Um mês depois, no final de Fevereiro (2º m Qt), há dois meses de informação do trimestre corrente para o indicador de confiança dos consumidores, os dados das transações ATM/POS estão disponíveis para Janeiro, e ainda não há dados do trimestre em curso para as outras duas variáveis: tanto as exportações da BoP como as dormidas estão disponíveis para Dezembro de 2015. Novamente, um mês depois, no final de Março (3º m Qt) há três meses de dados do trimestre corrente para o indicador de confiança dos consumidores, dois meses de dados para as transações ATM/POS e a informação quer das exportações da BoP, quer das dormidas está disponível para Janeiro. Além disso, a partir desta data, os dados sobre as exportações trimestrais de turismo para o quarto trimestre de 2015 (Qt-1) podem também ser incluídos. No final de Abril 2016 (1º m Qt+1), tanto o indicador de confiança dos consumidores como as transações ATM/POS têm três meses de dados do trimestre de interesse e a informação para as exportações da BoP e para as dormidas está disponível até Fevereiro. Finalmente, no final de Maio (2º m Qt+1), existe informação completa para todos os meses do trimestre de referência para todas as variáveis.

Neste exemplo, as duas últimas projeções são realizadas em Abril e Maio de 2016 e referem-se ao trimestre anterior. Note-se que, em contraste com esta nossa perspetiva mais ampla sobre o termo “previsão contemporânea”, que nos permite simplificar a redação, em algumas aplicações as previsões para um trimestre realizadas nesse mesmo trimestre e no trimestre seguinte são denominadas como “previsões contemporâneas” (*nowcasts*) e “previsões para o passado” (*backcasts*), respetivamente (ver Banbura *et al.* (2011)).

Como a utilização de dados para todos os meses do trimestre e para todos os indicadores permite ter previsões para o crescimento das exportações portuguesas de turismo num determinado trimestre apenas alguns dias antes da divulgação dos dados oficiais, é essencial ter projeções que explorem a informação parcial intra-trimestral disponível mais atempadamente. No âmbito dos modelos de ligação, quando não estão disponíveis todos os meses do trimestre de cada preditor, são utilizadas estimativas para as observações mensais em falta, obtidas a partir de modelos univariados simples, como descrito na secção 3. Todos as previsões são calculadas de forma direta, isto é, não são utilizadas projeções da variável dependente para obter futuras projeções, o que implica diferentes modelos de ligação para cada horizonte trimestral. No âmbito das regressões MIDAS, as diferentes projeções para o trimestre de interesse são calculadas utilizando modelos distintos para cada conjunto de informação intra-trimestral, ou seja, uma nova regressão é usada sempre que nova informação (mensal e trimestral) é incluída.

Finalmente, para avaliar o desempenho de previsão dos vários modelos de ligação e MIDAS no período fora da amostra foi utilizada a raiz quadrada do erro quadrático médio de previsão (RMSE). Foi calculado o RMSE relativo para comparar o desempenho destas duas abordagens face a um modelo de

referência (*benchmark*) trimestral. O modelo de referência é um modelo auto-regressivo (AR) univariado, que é estimado de forma recursiva, com a ordem de defasamento (de 0 a 3 defasamentos) a ser determinada pelo BIC.

Principais resultados

Esta secção apresenta os resultados do exercício de previsão contemporânea em pseudo tempo real. Como, em média nos períodos de avaliação, os modelos MIDAS com termos AR superaram as regressões MIDAS que não os incluem, seguidamente vamos-nos centrar apenas nas especificações MIDAS com dinâmicas AR. Este resultado está em linha com outros estudos que mostram que os modelos MIDAS sem componente AR têm geralmente um desempenho pior do que as especificações MIDAS que a incluem (ver, por exemplo, Kuzin *et al.* (2011) e Duarte (2014)). Adicionalmente, as regressões CF-MIDAS tiveram o pior desempenho em termos de precisão de previsão, pelo que foram igualmente excluídas da análise.⁹

Em relação aos resultados das regressões com um único indicador, o Gráfico 3 fornece evidência sobre os desempenhos das diferentes classes de modelos de frequência mista. Os valores referem-se ao desempenho em termos de RMSE relativo, nos vários períodos de previsão, face ao modelo AR de referência. Um rácio inferior a 1 indica um ganho de previsão dos modelos de ligação e/ou MIDAS, enquanto um valor superior a 1 significa que o modelo univariado de referência supera os modelos alternativos. O Gráfico 3 mostra heurísticamente um dos factos estilizados desta literatura: a precisão de previsão deste tipo de modelos tende a aumentar à medida que o tempo passa e mais informação se torna disponível.

Começando com as equações de ligação, nos dois primeiros períodos de avaliação, que correspondem aos dois primeiros meses do trimestre de referência, nenhum indicador supera o modelo de referência. Porém, à medida que mais dados sobre o trimestre vão sendo observados, a incorporação das exportações da balança de viagens e turismo da BoP, das transações ATM/POS e, em menor medida, das dormidas leva a um RMSE menor do que o da referência univariada nos três períodos de avaliação seguintes. As exportações da balança de viagens e turismo da BoP são o melhor indicador em todos os casos, exceto um: a exceção são as transações ATM/POS no terceiro mês do trimestre de referência, quando há apenas um mês de dados de exportações da BoP, mas dois meses de informação de transações ATM/POS.

Por seu lado, as regressões MIDAS que apresentam previsões mais precisas superam sempre o modelo AR de referência, ao longo dos vários períodos de avaliação. Se analisarmos os diferentes indicadores de curto

9. Todos os resultados estão disponíveis mediante pedido aos autores.

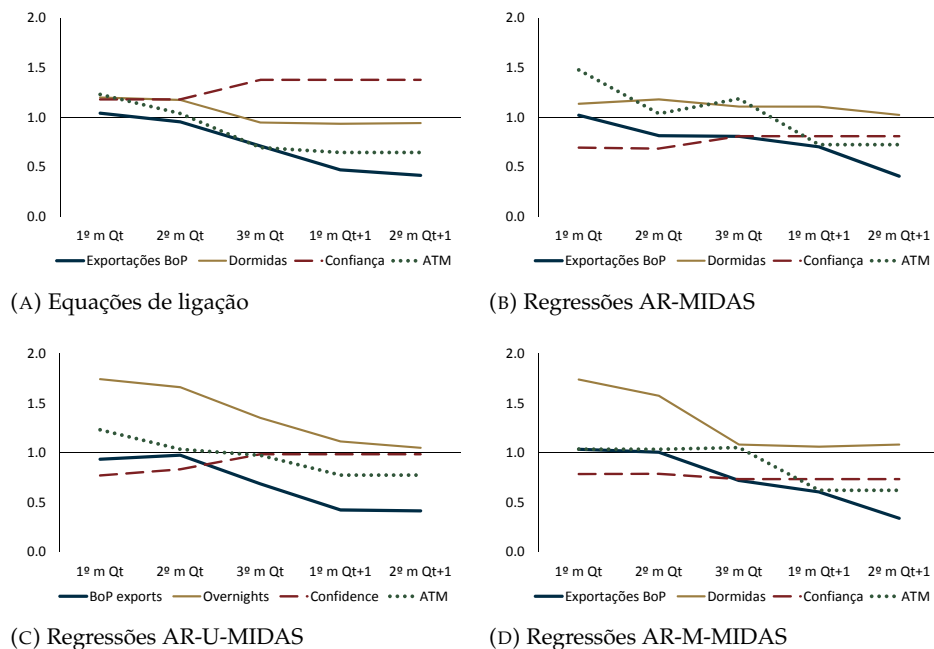


GRÁFICO 3: RMSE relativo dos modelos com indicadores individuais (modelo de referência = AR)

Notas: Ver secção 4 para uma descrição detalhada das variáveis e da informação utilizada em cada previsão. Rácios de RMSE relativamente a um modelo AR. Um rácio inferior a 1 indica um ganho de previsão dos modelos de ligação e/ou MIDAS, enquanto um valor superior a 1 significa que o modelo univariado de referência supera os modelos alternativos.

prazo, verificamos que existe um padrão comum às variantes MIDAS: nos dois primeiros meses do trimestre de referência, o indicador com melhor desempenho é o índice de confiança dos consumidores; as exportações da BoP têm o melhor desempenho a partir daí, à medida que se tornam gradualmente disponíveis mais dados sobre este indicador para o trimestre de referência. Adicionalmente, as dormidas tendem a estar associadas a maus desempenhos preditivos nos modelos MIDAS, piores do que o modelo AR de referência em todos os períodos.

A fim de examinar melhor as suas propriedades e aferir as suas diferenças e semelhanças, o Gráfico 4 apresenta os desempenhos mínimos e médios – do conjunto de todos os indicadores individuais – do RMSE relativo (face a um modelo AR de referência) dos vários modelos de frequência mista analisados. Em termos gerais, o modelo com melhor desempenho é sempre um modelo MIDAS, ou seja, a variante MIDAS com o menor RMSE supera sempre os modelos de ligação e isso é verdade tanto nos mínimos como nas médias. No

entanto, em ambos os casos, o modelo MIDAS com melhor desempenho não é sempre da mesma variante.

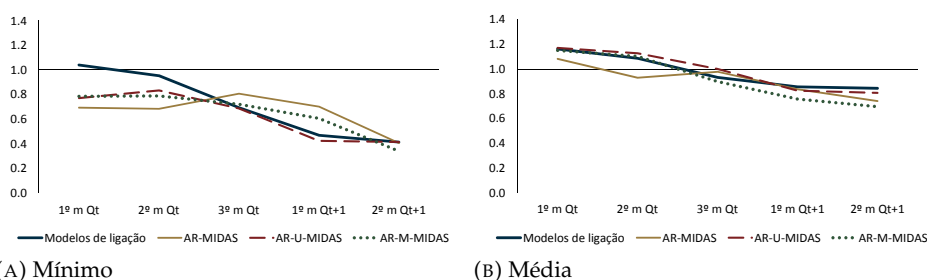


GRÁFICO 4: Mínimo e média do RMSE relativo dos modelos com indicadores individuais (modelo de referência = AR)

Notas: Ver secção 4 para uma descrição detalhada da informação utilizada em cada previsão. Mínimo e média dos rácios de RMSE relativo face a um modelo AR de referência, considerando todos os indicadores em cada classe de modelos. Um rácio inferior a 1 indica um ganho de previsão dos modelos de ligação e/ou MIDAS, enquanto um valor superior a 1 significa que o modelo univariado de referência supera os modelos alternativos.

Considerando apenas os valores mínimos do RMSE relativo, o melhor desempenho de previsão de um modelo MIDAS é sempre melhor do que o do modelo AR de referência, permitindo ganhos de cerca de 30 a 65 por cento ao longo dos vários períodos. De facto, em comparação com as equações de ligação, as regressões MIDAS parecem funcionar especialmente bem no curto prazo, i.e., quando menos informação do trimestre corrente está disponível. Pelo contrário, a falta de dados relativos às exportações da BoP para o trimestre corrente nos dois primeiros períodos de avaliação é crítica para o desempenho das equações de ligação, que nunca superam o modelo AR de referência. Nos últimos três períodos de avaliação, existem apenas ligeiras diferenças entre as regressões MIDAS com menor RMSE relativo e os modelos de ligação. Em ambos os casos, os ganhos de previsão relativamente à referência univariada aumentam de cerca de 30 para cerca de 60 por cento no último período.

Em relação à média dos desempenhos de previsão, é difícil superar o modelo AR de referência nos dois primeiros meses do trimestre (a única exceção é o modelo AR-MIDAS no segundo período). Adicionalmente, ainda que os melhores modelos MIDAS sejam (ligeiramente) melhores, em média, do que as equações de ligação, não existem diferenças substanciais entre os desempenhos médios destes modelos com um único indicador, ao longo de todos os períodos.

Combinação de previsões

Esta secção examina os resultados do exercício de combinação de previsões para cada classe de modelos. A combinação de projeções individuais recorrendo aos pesos descritos na equação 8 tem um melhor desempenho global do que os modelos multivariados e do que a média simples das projeções. Assim, esta secção analisa apenas os resultados das combinações de previsões usando os pesos de Stock e Watson (2004).

O Gráfico 5 apresenta os desempenhos mínimos e médios do RMSE relativo dos vários modelos (face a um modelo AR de referência), considerando todas as combinações possíveis dos quatro indicadores individuais para cada tipo de modelo considerado.¹⁰

Comparando os resultados incluídos nos Gráficos 4 e 5, torna-se mais claro que a combinação de previsões é uma boa estratégia, que tende a superar os modelos com uma única variável em cada período e para cada tipo de modelo. O resultado de que a combinação de previsões é mais estável do que a previsão com modelos individuais está de acordo com outros estudos na literatura sobre modelos MIDAS. Kuzin *et al.* (2013) concluem que a combinação de previsões supera os modelos com uma única variável na previsão do crescimento do PIB trimestral e Ghysels e Ozkan (2015) mostram que as combinações de previsões de modelos MIDAS fornecem ganhos em relação aos modelos tradicionais na previsão do orçamento federal anual dos EUA. Para além disso, Clements e Galvão (2008) indicam que combinações de previsões de modelos MIDAS são pelo menos tão boas como combinações de previsões de modelos de ligação e de outros modelos de frequência mista.

Os RMSE relativos das combinações de previsões são sempre inferiores a 1 em todos os casos representados no Gráfico 5, o que implica que, não só o melhor modelo de cada classe tem um desempenho melhor do que o modelo AR de referência, mas também que, em média, é possível melhorar a precisão de previsão usando modelos de frequência mista. Tal como nos modelos com uma única variável, o modelo com melhor desempenho é sempre uma regressão MIDAS, tanto nos mínimos como nas médias. Ainda que os melhores resultados não sejam sempre obtidos a partir do mesmo tipo de modelo MIDAS, o modelo AR-M-MIDAS proporciona bons resultados de previsão ao longo dos vários períodos.

Para examinar com maior detalhe o desempenho das combinações de previsões dos vários modelos, o Quadro 1 compara os seus RMSE com os do modelo AR de referência. Das onze combinações possíveis apresentadas

10. O Apêndice A inclui os resultados detalhados da precisão de previsão das onze combinações possíveis para todos os modelos de frequência mista considerados: a primeira tabela apresenta os RMSE relativos de cada modelo face a um modelo AR de referência e a segunda tabela inclui os RMSE das diferentes variantes de modelos MIDAS em relação ao RMSE das equações de ligação.

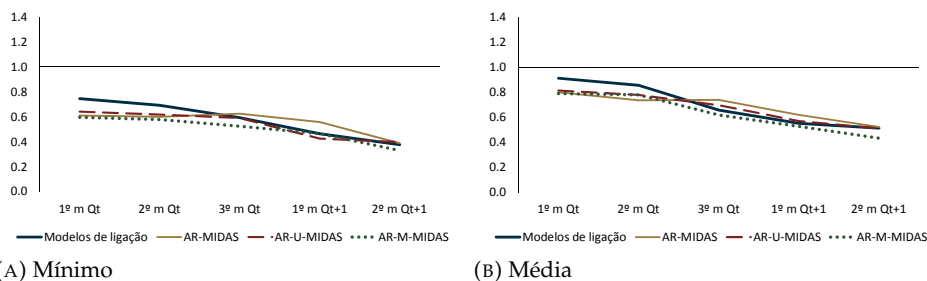


GRÁFICO 5: Mínimo e média do RMSE relativo das combinações de previsões (modelo de referência = AR)

Notas: Ver secção 4 para uma descrição detalhada da informação utilizada em cada previsão. Mínimo e média dos rácios de RMSE relativo face a um modelo AR de referência, considerando todas as possíveis combinações de indicadores em cada classe de modelos. Um rácio inferior a 1 indica um ganho de previsão dos modelos de ligação e/ou MIDAS, enquanto um valor superior a 1 significa que o modelo univariado de referência supera os modelos alternativos.

no Apêndice A, este quadro inclui as combinações com melhor desempenho em qualquer dos modelos de ligação/MIDAS em cada período de avaliação. Seguindo Foroni e Marcellino (2014), testamos a hipótese de igualdade de precisão no desempenho da previsão usando o teste de Diebold e Mariano (1995) modificado para amostras curtas por Harvey *et al.* (1997). Os casos em que, de acordo com este teste, a hipótese de igual precisão de previsão é rejeitada estão indicados no quadro por um ou mais *, dependendo do nível de significância estatística.

Os resultados mostram que a combinação de previsões, quer de modelos de ligação, quer de modelos MIDAS, tem, em geral, um bom desempenho: o modelo AR de referência é superado pela maioria das combinações em cada período e as diferenças em termos de RMSE são, na grande maioria dos casos, estatisticamente significativas. Por exemplo, nos últimos três períodos de avaliação, é possível obter uma projeção que é estatisticamente superior à do modelo univariado de referência em 96.4 por cento dos casos.

Em relação às melhores combinações para cada modelo e em cada período (os números a negrito no quadro), todas elas são estatisticamente superiores ao modelo AR de referência. Os bons desempenhos são partilhados por diferentes combinações, mas há uma característica comum a todos os modelos: o indicador de confiança dos consumidores faz sempre parte da combinação com melhor desempenho nos dois primeiros períodos, e as exportações da BoP estão sempre incluídas na melhor combinação dos últimos três momentos de projeção.

O mínimo RMSE relativo global para cada período de avaliação e considerando todos os modelos e combinações de indicadores (os números sombreados a escuro e escritos a branco no quadro) é sempre obtido a partir de

um modelo MIDAS: AR-M-MIDAS em quatro casos e AR-U-MIDAS em um dos casos. Não só as especificações MIDAS com melhor desempenho variam ao longo do tempo, como a melhor combinação de preditores também muda consoante os diferentes conjuntos de informação intra-trimestral. O modelo AR-M-MIDAS funciona particularmente bem nos três meses do trimestre de referência: produz os melhores resultados nos dois primeiros meses do trimestre de referência através da combinação de transações ATM/POS e confiança dos consumidores e, no último mês do trimestre, através da combinação dos quatro indicadores individuais. Nos últimos dois períodos de avaliação, quando mais dados do trimestre de referência já são observados, a combinação preferida conjuga as exportações da BoP e as dormidas, estando primeiro associada ao modelo AR-U-MIDAS e, no último período, ao modelo AR-M-MIDAS.

Outra forma de comparar os modelos alternativos de frequência mista é calcular os RMSEs dos vários modelos MIDAS em relação ao RMSE das equações de ligação para cada combinação de indicadores. As áreas sombreadas a claro no Quadro 1 representam os casos em que as previsões obtidas a partir de um modelo MIDAS são significativamente mais precisas do que as respectivas previsões das equações de ligação, com um nível de significância de, pelo menos, 10 por cento.

As combinações de previsões de modelos MIDAS mais úteis devem superar quer o modelo AR de referência, quer a equação de ligação concorrente (Schumacher 2016). No Quadro 1, estes são os casos em que os *s estão sombreados a claro. As melhorias estatisticamente significativas dos desempenhos preditivos de modelos MIDAS face a ambas as referências em simultâneo ocorrem, em particular, nos dois primeiros períodos: em cerca de 35 por cento dos casos para o AR-U-MIDAS e em mais de 70 por cento dos casos quer para o AR-MIDAS, quer para o AR-M-MIDAS. Considerando os cinco períodos de avaliação e as sete melhores combinações de indicadores, o AR-M-MIDAS é o modelo com o melhor desempenho global: produz projeções que são estatisticamente melhores que ambas as referências em cerca de 57 por cento dos casos. Em geral, e tendo em conta todos os períodos de avaliação, as combinações de previsões que incluem 3 e 4 indicadores tendem a ser mais fiáveis, no sentido em que tendem a superar ambas as referências mais frequentemente do que as combinações com menos indicadores.

	1º m Qt	2º m Qt	3º m Qt	1º m Qt+1	2º m Qt+1
Modelos de ligação					
Dormidas + Confiança	0.861	0.845	0.755 *	0.769 *	0.786
ATM + Confiança	0.754 *	0.727	0.623 ***	0.599 ***	0.599 ***
Exportações BoP + Confiança	0.745 *	0.691 **	0.694 **	0.493 ***	0.450 ***
Exportações BoP + Dormidas	1.046	0.981	0.672 ***	0.464 ***	0.376 ***
Exportações BoP + Dormidas + Confiança	0.805 *	0.762 **	0.633 **	0.473 ***	0.406 ***
Exportações BoP + Confiança + ATM	0.794 *	0.747 **	0.592 ***	0.483 ***	0.435 ***
Exportações BoP + Dormidas + Confiança + ATM	0.848 *	0.802 **	0.604 ***	0.488 ***	0.427 ***
AR-MIDAS					
Dormidas + Confiança	0.611 **	0.599 **	0.656 **	0.655 **	0.738 *
ATM + Confiança	0.609 **	0.604 **	0.657 **	0.614 **	0.614 **
Exportações BoP + Confiança	0.683 **	0.631 **	0.662 **	0.662 **	0.433 ***
Exportações BoP + Dormidas	0.995	0.822 **	0.770 **	0.574 ***	0.388 ***
Exportações BoP + Dormidas + Confiança	0.668 **	0.613 ***	0.624 ***	0.569 ***	0.422 ***
Exportações BoP + Confiança + ATM	0.657 **	0.624 ***	0.637 ***	0.561 ***	0.434 ***
Exportações BoP + Dormidas + Confiança + ATM	0.679 ***	0.637 ***	0.653 ***	0.557 ***	0.439 ***
AR-U-MIDAS					
Dormidas + Confiança	0.823	0.814	0.922	0.754 *	0.810
ATM + Confiança	0.715 *	0.666 *	0.662 ***	0.588 ***	0.588 ***
Exportações BoP + Confiança	0.655 **	0.668 **	0.631 ***	0.495 ***	0.407 ***
Exportações BoP + Dormidas	0.880	0.849	0.658 ***	0.425 ***	0.407 ***
Exportações BoP + Dormidas + Confiança	0.679 **	0.676 **	0.629 ***	0.518 ***	0.407 ***
Exportações BoP + Confiança + ATM	0.641 ***	0.617 ***	0.589 ***	0.467 ***	0.397 ***
Exportações BoP + Dormidas + Confiança + ATM	0.688 **	0.656 **	0.599 ***	0.503 ***	0.405 ***
AR-M-MIDAS					
Dormidas + Confiança	0.634 **	0.619 **	0.582 ***	0.577 ***	0.582 ***
ATM + Confiança	0.595 ***	0.577 **	0.544 ***	0.509 ***	0.509 ***
Exportações BoP + Confiança	0.633 **	0.620 **	0.615 **	0.593 ***	0.355 ***
Exportações BoP + Dormidas	1.088	1.022	0.647 ***	0.492 ***	0.329 ***
Exportações BoP + Dormidas + Confiança	0.632 ***	0.609 ***	0.550 ***	0.503 ***	0.338 ***
Exportações BoP + Confiança + ATM	0.632 ***	0.611 ***	0.538 ***	0.484 ***	0.359 ***
Exportações BoP + Dormidas + Confiança + ATM	0.659 ***	0.641 ***	0.524 ***	0.465 ***	0.358 ***

QUADRO 1. RMSE relativo das combinações de previsões face a um modelo de referência AR

Notas: Ver secção 4 para uma descrição detalhada das variáveis e da informação utilizada em cada previsão. Rácios de RMSE relativamente a um modelo AR de referência. Um rácio inferior a 1 indica um ganho de previsão dos modelos de ligação e/ou MIDAS. *, ** e *** indicam as previsões que são estatisticamente superiores às do modelo de referência com um nível de significância de 10, 5 e 1 por cento, respetivamente, de acordo com o teste de Diebold e Mariano (1995) modificado para amostras curtas por Harvey *et al.* (1997). Os valores a negrito representam o mínimo RMSE relativo dentro de cada modelo em cada período. Os números sombreados a escuro e escritos a branco representam o mínimo RMSE relativo para cada período de avaliação considerando todos os modelos e todas as combinações de indicadores. As áreas sombreadas a claro representam os casos em que as previsões obtidas a partir do modelo MIDAS são estatisticamente superiores à respetivas previsões da equação de ligação (com um nível de significância de, pelo menos, 10 por cento).

Conclusões

As exportações de turismo são uma componente extremamente importante do comércio internacional de bens e serviços em Portugal. As previsões de curto prazo desta variável têm um papel relevante no acompanhamento da atividade económica e das contas externas portuguesas.

O objetivo deste artigo é a previsão contemporânea do crescimento real das exportações trimestrais de turismo, utilizando quatro indicadores mensais distintos num exercício recursivo em pseudo tempo real. Para tal são utilizadas duas abordagens uni-equacionais que lidam com dados de frequência mista: equações de ligação e regressões MIDAS. As equações de ligação são uma das técnicas mais utilizadas para ligar variáveis mensais e trimestrais. Nestes modelos, as variáveis em ambos os lados da equação estão na mesma frequência (baixa): no nosso caso, os indicadores mensais são agregados para os valores trimestrais correspondentes. Por seu lado, nas regressões MIDAS as observações da variável dependente na frequência baixa são ligadas diretamente às observações dos preditores com maior frequência, sem qualquer agregação temporal prévia. Neste artigo são avaliadas diferentes especificações de modelos de ligação e MIDAS com indicadores individuais, assim como combinações de previsões.

Os resultados obtidos sugerem que, como esperado, a utilização de indicadores de curto prazo em modelos de frequência mista contribui para aumentar a precisão de previsão quando comparada com uma referência univariada. Em geral, os modelos MIDAS tendem a ter um melhor desempenho do que os tradicionais modelos de ligação para a maioria dos preditores e períodos de avaliação, havendo maiores diferenças quando menos informação do trimestre corrente está disponível. As combinações de previsões, quer de equações de ligação, quer de regressões MIDAS, proporcionam sempre ganhos face a modelos com um único indicador. Com efeito, um resultado geral comum a todos os modelos de frequência mista considerados é que o modelo AR de referência pode sempre ser superado pela melhor combinação de projeções em cada período de avaliação, e que as diferenças em termos de RMSE relativo são estatisticamente significativas. Em suma, o melhor desempenho de previsão é sempre obtido a partir de uma combinação de projeções de uma variante MIDAS com dinâmica AR, o que sugere a utilização desta classe de modelos de frequência mista na previsão de curto prazo das exportações de turismo.

Referências

- Andraz, Jorge L.M., Pedro M.D.C.B. Gouveia, e Paulo M.M. Rodrigues (2009). "Modelling and forecasting the UK tourism growth cycle in Algarve." *Tourism Economics*, 15(2), 323–338.
- Andreou, Elena, Eric Ghysels, e Andros Kourtellos (2011). "Forecasting with Mixed-Frequency Data." In *The Oxford Handbook of Economic Forecasting*, edited by Michael P. Clements e David F. Hendry, chap. 8, pp. 225–267. Oxford University Press.
- Armesto, Michelle T., Kristie M. Engemann, e Michael T. Owyang (2010). "Forecasting with mixed frequencies." *Federal Reserve Bank of St. Louis Review*, (Nov/Dec), 521–536.
- Athanasopoulos, George, Rob J. Hyndman, Haiyan Song, e Doris C. Wu (2011). "The tourism forecasting competition." *International Journal of Forecasting*, 27(3), 822–844.
- Baffigi, Alberto, Roberto Golinelli, e Giuseppe Parigi (2004). "Bridge models to forecast the euro area GDP." *International Journal of Forecasting*, 20(3), 447–460.
- Banbura, Marta, Domenico Giannone, Michele Modugno, e Lucrezia Reichlin (2013). "Now-Casting and the Real-Time Data Flow." In *Handbook of Economic Forecasting*, vol. 2, Part A, edited by Graham Elliott e Allan Timmermann, chap. 4, pp. 195–237. Elsevier.
- Banbura, Marta, Domenico Giannone, e Lucrezia Reichlin (2011). "Nowcasting." In *The Oxford Handbook of Economic Forecasting*, edited by Michael P. Clements e David F. Hendry, chap. 7, pp. 193–224. Oxford University Press.
- Barhouni, Karim, Olivier Darné, Laurent Ferrara, e Bertrand Pluyaud (2012). "Monthly GDP forecasting using bridge models: Application for the French economy." *Bulletin of Economic Research*, 64, s53–s70.
- Bulligan, Guido, Massimiliano Marcellino, e Fabrizio Venditti (2015). "Forecasting economic activity with targeted predictors." *International Journal of Forecasting*, 31(1), 188–206.
- Cardoso, Fátima e Cláudia Duarte (2006). "The use of qualitative information for forecasting exports." *Banco de Portugal Economic Bulletin*, Winter, 67–74.
- Chen, Xilong e Eric Ghysels (2011). "News – Good or Bad – and Its Impact on Volatility Predictions over Multiple Horizons." *Review of Financial Studies*, 24(1), 46–81.
- Clements, M. P. e A. B. Galvão (2008). "Macroeconomic Forecasting With Mixed-Frequency Data: Forecasting Output Growth in the United States." *Journal of Business and Economic Statistics*, 26(4), 546–554.
- Daniel, Ana Cristina M. e Francisco F. R. Ramos (2002). "Modelling inbound international tourism demand to Portugal." *International Journal of Tourism Research*, 4(3), 193–209.
- Diebold, Francis X. e Roberto S. Mariano (1995). "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics*, 13(3), 253–63.

- Diron, Marie (2008). "Short-term forecasts of euro area real GDP growth: an assessment of real-time performance based on vintage data." *Journal of Forecasting*, 27(5), 371–390.
- Duarte, Cláudia (2014). "Autoregressive augmentation of MIDAS regressions." Working Paper 01/2014, Banco de Portugal.
- Duarte, Cláudia, Paulo M. M. Rodrigues, e António Rua (2016). "A Mixed Frequency Approach to Forecast Private Consumption with ATM/POS Data." *International Journal of Forecasting*, Forthcoming.
- Esteves, Paulo Soares (2009). "Are ATM/POS Data Relevant When Nowcasting Private Consumption?" Working Paper 25/2009, Banco de Portugal.
- Esteves, Paulo Soares e António Rua (2012). "Short-term forecasting for the Portuguese economy: A methodological overview." *Banco de Portugal Economic Bulletin*, Autumn, 137–149.
- Foroni, C. e M. Marcellino (2014). "A comparison of mixed frequency approaches for nowcasting Euro area macroeconomic aggregates." *International Journal of Forecasting*, 30, 554–568.
- Foroni, Claudia, Massimiliano Marcellino, e Christian Schumacher (2015). "Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(1), 57–82.
- Ghysels, E., P. Santa-Clara, e R. Valkanov (2004). "The MIDAS touch: Mixed data sampling regressions." Discussion paper, UNC and UCLA.
- Ghysels, Eric e Massimiliano Marcellino (2016). "The econometric analysis of mixed frequency data sampling." *Journal of Econometrics*, 193(2), 291 – 293.
- Ghysels, Eric e Nazire Ozkan (2015). "Real-time forecasting of the US federal government budget: A simple mixed frequency data regression approach." *International Journal of Forecasting*, 31(4), 1009 – 1020.
- Ghysels, Eric, Arthur Sinko, e Rossen Valkanov (2007). "MIDAS Regressions: Further Results and New Directions." *Econometric Reviews*, 26(1), 53–90.
- Giannone, Domenico, Lucrezia Reichlin, e David Small (2008). "Nowcasting: The real-time informational content of macroeconomic data." *Journal of Monetary Economics*, 55(4), 665–676.
- Gouveia, Pedro M.D.C.B. e Paulo M.M. Rodrigues (2005). "Dating and synchronizing tourism growth cycles." *Tourism Economics*, 11(4), 501–515.
- Harvey, David, Stephen Leybourne, e Paul Newbold (1997). "Testing the equality of prediction mean squared errors." *International Journal of Forecasting*, 13(2), 281 – 291.
- Kuzin, Vladimir, Massimiliano Marcellino, e Christian Schumacher (2011). "MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area." *International Journal of Forecasting*, 27(2), 529 – 542.
- Kuzin, Vladimir, Massimiliano Marcellino, e Christian Schumacher (2013). "Pooling versus model selection for nowcasting GDP with many predictors: Empirical evidence for six industrialized countries." *Journal of Applied*

- Econometrics*, 28(3), 392–411.
- Maria, José Ramos e Sara Serra (2008). "Forecasting investment: A fishing contest using survey data." Working Paper 18/2008, Banco de Portugal.
- Rodrigues, Paulo M.M. e Pedro M.D.C.B. Gouveia (2004). "An application of PAR models for tourism forecasting." *Tourism Economics*, 10(3), 281–303.
- Schumacher, Christian (2016). "A comparison of MIDAS and bridge equations." *International Journal of Forecasting*, 32(2), 257–270.
- Serra, Jaime, Antónia Correia, e Paulo M.M. Rodrigues (2014). "A comparative analysis of tourism destination demand in Portugal." *Journal of Destination Marketing and Management*, 2(4), 221–227.
- Song, Haiyan e Rob J. Hyndman (2011). "Tourism forecasting: An introduction." *International Journal of Forecasting*, 27(3), 817–821.
- Song, Haiyan e Gang Li (2008). "Tourism demand modelling and forecasting - A review of recent research." *Tourism Management*, 29(2), 203–220.
- Song, Haiyan, Gang Li, Stephen F. Witt, e George Athanasopoulos (2011). "Forecasting tourist arrivals using time-varying parameter structural time series models." *International Journal of Forecasting*, 27(3), 855–869.
- Stock, James H. e Mark W. Watson (2004). "Combination forecasts of output growth in a seven-country data set." *Journal of Forecasting*, 23(6), 405–430.
- Teixeira, João Paulo e Paula Odete Fernandes (2012). "Tourism Time Series Forecast - Different ANN Architectures with Time Index Input." *Procedia Technology*, 5, 445–454.
- Teixeira, João Paulo e Paula Odete Fernandes (2014). "Tourism time series forecast with artificial neural networks." *Tékhne*, 12(1-2), 26–36.

Apêndice: Resultados detalhados da combinação de previsões

	1° m Qt	2° m Qt	3° m Qt	1° m Qt+1	2° m Qt+1
Modelos de ligação					
Dormidas + ATM	1.144	1.056	0.731***	0.679***	0.679***
Dormidas + Confiança	0.861	0.845	0.755*	0.769*	0.786
ATM + Confiança	0.754*	0.727	0.623***	0.599***	0.599***
Exportações BoP + Confiança	0.745*	0.691**	0.694**	0.493***	0.450***
Exportações BoP + Dormidas	1.046	0.981	0.672***	0.464***	0.376***
Exportações BoP + ATM	1.103	0.984	0.619***	0.481***	0.421***
Dormidas + Confiança + ATM	0.822	0.801*	0.652***	0.624***	0.627***
Exportações BoP + Dormidas + Confiança	0.805*	0.762**	0.633**	0.473***	0.406***
Exportações BoP + Confiança + ATM	0.794*	0.747**	0.592***	0.483***	0.435***
Exportações BoP + Dormidas + ATM	1.094	0.995	0.638***	0.492***	0.417***
Exportações BoP + Dormidas + Confiança + ATM	0.848*	0.802**	0.604***	0.488***	0.427***
AR-MIDAS					
Dormidas + ATM	1.196	1.186	1.127	0.813***	0.758**
Dormidas + Confiança	0.611**	0.599**	0.656**	0.655**	0.738*
ATM + Confiança	0.609**	0.604**	0.657**	0.614**	0.614**
Exportações BoP + Confiança	0.683**	0.631**	0.662**	0.662**	0.433***
Exportações BoP + Dormidas	0.995	0.822**	0.770**	0.574***	0.388***
Exportações BoP + ATM	1.022	0.857**	0.807**	0.561***	0.421***
Dormidas + Confiança + ATM	0.624***	0.623***	0.689***	0.642***	0.649**
Exportações BoP + Dormidas + Confiança	0.668**	0.613***	0.624***	0.569***	0.422***
Exportações BoP + Confiança + ATM	0.657**	0.624***	0.637***	0.561***	0.434***
Exportações BoP + Dormidas + ATM	1.033	0.881**	0.830**	0.583***	0.426***
Exportações BoP + Dormidas + Confiança + ATM	0.679***	0.637***	0.653***	0.557***	0.439***
AR-U-MIDAS					
Dormidas + ATM	1.313	1.195	0.910	0.800**	0.728***
Dormidas + Confiança	0.823	0.814	0.922	0.754*	0.810
ATM + Confiança	0.715*	0.666*	0.662***	0.588***	0.588***
Exportações BoP + Confiança	0.655**	0.668**	0.631***	0.495***	0.407***
Exportações BoP + Dormidas	0.880	0.849	0.658***	0.425***	0.407***
Exportações BoP + ATM	0.846*	0.815**	0.673***	0.511***	0.415***
Dormidas + Confiança + ATM	0.794	0.751*	0.712**	0.628***	0.613***
Exportações BoP + Dormidas + Confiança	0.679**	0.676**	0.629***	0.518***	0.407***
Exportações BoP + Confiança + ATM	0.641***	0.617***	0.589***	0.467***	0.397***
Exportações BoP + Dormidas + ATM	0.899	0.835*	0.658***	0.553***	0.417***
Exportações BoP + Dormidas + Confiança + ATM	0.688**	0.656**	0.599***	0.503***	0.405***
AR-M-MIDAS					
Dormidas + ATM	1.154	1.226	0.945	0.662***	0.668***
Dormidas + Confiança	0.634**	0.619**	0.582***	0.577***	0.582***
ATM + Confiança	0.595***	0.577**	0.544***	0.509***	0.509***
Exportações BoP + Confiança	0.633**	0.620**	0.615**	0.593***	0.355***
Exportações BoP + Dormidas	1.088	1.022	0.647***	0.492***	0.329***
Exportações BoP + ATM	0.978	0.981	0.649***	0.489***	0.359***
Dormidas + Confiança + ATM	0.620***	0.613***	0.545***	0.516***	0.520***
Exportações BoP + Dormidas + Confiança	0.632***	0.609***	0.550***	0.503***	0.338***
Exportações BoP + Confiança + ATM	0.632***	0.611***	0.538***	0.484***	0.359***
Exportações BoP + Dormidas + ATM	1.044	1.035	0.647***	0.483***	0.363***
Exportações BoP + Dormidas + Confiança + ATM	0.659***	0.641***	0.524***	0.465***	0.358***

QUADRO A.1. RMSE relativo das combinações de previsões face a um modelo de referência AR

Notas: Ver secção 4 para uma descrição detalhada das variáveis e da informação utilizada em cada previsão. Rácios de RMSE relativamente a um modelo AR de referência. Um rácio inferior a 1 indica um ganho de previsão dos modelos de ligação e/ou MIDAS. *, ** e *** indicam as previsões que são estatisticamente superiores às do modelo de referência com um nível de significância de 10, 5 e 1 por cento, respetivamente, de acordo com o teste de Diebold e Mariano (1995) modificado para amostras curtas por Harvey *et al.* (1997).

	1° m Qt	2° m Qt	3° m Qt	1° m Qt+1	2° m Qt+1
AR-MIDAS					
Dormidas + ATM	1.045	1.123	1.542	1.197	1.116
Dormidas + Confiança	0.709***	0.708***	0.869	0.852*	0.938
ATM + Confiança	0.808***	0.830**	1.056	1.026	1.026
Exportações BoP + Confiança	0.917	0.912	0.954	1.343	0.964
Exportações BoP + Dormidas	0.951	0.837	1.146	1.237	1.032
Exportações BoP + ATM	0.927	0.871	1.304	1.165	0.999
Dormidas + Confiança + ATM	0.760***	0.778***	1.057	1.029	1.035
Exportações BoP + Dormidas + Confiança	0.829**	0.804**	0.986	1.203	1.041
Exportações BoP + Confiança + ATM	0.828**	0.835*	1.076	1.162	0.999
Exportações BoP + Dormidas + ATM	0.945	0.886	1.301	1.185	1.020
Exportações BoP + Dormidas + Confiança + ATM	0.801**	0.794**	1.081	1.141	1.029
AR-U-MIDAS					
Dormidas + ATM	1.147	1.132	1.245	1.179	1.072
Dormidas + Confiança	0.956	0.963	1.222	0.980	1.030
ATM + Confiança	0.948	0.916	1.064	0.982	0.982
Exportações BoP + Confiança	0.879	0.967	0.909	1.004	0.905*
Exportações BoP + Dormidas	0.841	0.865	0.979	0.916	1.084
Exportações BoP + ATM	0.767*	0.829	1.088	1.061	0.987
Dormidas + Confiança + ATM	0.966	0.938	1.091	1.006	0.978
Exportações BoP + Dormidas + Confiança	0.843*	0.887	0.992	1.095	1.002
Exportações BoP + Confiança + ATM	0.807**	0.827*	0.995	0.966	0.912
Exportações BoP + Dormidas + ATM	0.822*	0.839	1.031	1.125	1.000
Exportações BoP + Dormidas + Confiança + ATM	0.812**	0.819*	0.992	1.031	0.948
AR-M-MIDAS					
Dormidas + ATM	1.008	1.161	1.294	0.974	0.984
Dormidas + Confiança	0.736***	0.733***	0.771***	0.750**	0.740
ATM + Confiança	0.789***	0.793***	0.874	0.850**	0.850**
Exportações BoP + Confiança	0.850**	0.897	0.887	1.202	0.788***
Exportações BoP + Dormidas	1.040	1.042	0.963	1.060	0.876
Exportações BoP + ATM	0.887	0.998	1.049	1.016	0.853**
Dormidas + Confiança + ATM	0.754***	0.765***	0.836*	0.827**	0.829**
Exportações BoP + Dormidas + Confiança	0.785***	0.800***	0.868*	1.062	0.833**
Exportações BoP + Confiança + ATM	0.797***	0.819***	0.908	1.003	0.827***
Exportações BoP + Dormidas + ATM	0.955	1.040	1.014	0.982	0.870
Exportações BoP + Dormidas + Confiança + ATM	0.777***	0.800***	0.867	0.953	0.838**

QUADRO A.2. RMSE relativo das combinações de previsões face a modelos de ligação de referência

Notas: Ver secção 4 para uma descrição detalhada das variáveis e da informação utilizada em cada previsão. Rácios de RMSE relativamente aos modelos de ligação. Um rácio inferior a 1 indica um ganho de previsão dos modelos MIDAS. *, ** e *** indicam as previsões que são estatisticamente superiores às do modelo de referência com um nível de significância de 10, 5 e 1 por cento, respetivamente, de acordo com o teste de Diebold e Mariano (1995) modificado para amostras curtas por Harvey *et al.* (1997).