

FRANCISCO JOSÉ VEIGA

Forecast Errors in Prices and Wages:
The Experience with Three Programme Countries

MANUEL CORREIA DE PINHO / MARIA MANUEL PINHO

Esforço Fiscal em Portugal:
Uma Avaliação no Período 1995-2015

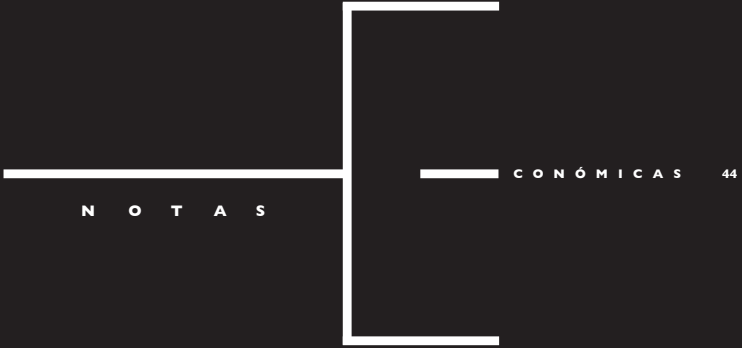
PATRÍCIA ANTUNES / PEDRO LOPES FERREIRA / LARA N. FERREIRA

A Utilização da Experiência de Escolha Discreta
na Valoração de Estados de Saúde

WEI-BIN ZHANG

A Small Open Economic Growth Model with
Imported Goods, Tourism, and Terms of Trade

(Página deixada propositadamente em branco)



N O T A S

CONÓMICAS 44

Propriedade

Faculdade de Economia
Universidade de Coimbra
Av. Dias da Silva, 165
3004-512 Coimbra

Apoios

Esta revista é financiada por Fundos Nacionais através da FCT – Fundação para a Ciência e a Tecnologia no âmbito do projeto PEst-OE/EGE/UI0031/2014

Número avulso

Portugal 7,5 €
Outros países 15 €

Condições de assinatura bienal

Portugal: 20 €
Outros países: 40 €

ISSN 0872-4733

ISSN Digital 2183-203X

DOI https://doi.org/10.14195/2183-203X_44

Depósito Legal 64932/93

Direcção Editorial (Editors)

Paulino Teixeira (Faculdade de Economia, U. Coimbra; pteixeira@fe.uc.pt)
Francisco Veiga (Escola de Economia e Gestão, U. Minho; fveiga@eeg.uminho.pt)
Joana Pais (ISEG, Universidade de Lisboa; jpais@iseg.utl.pt)
José Varejão (Faculdade de Economia, U. Porto; varejao@fep.up.pt)

Editores Associados (Associate Editors)

António Afonso (ISEG, Universidade de Lisboa; aafonso@iseg.utl.pt)
Óscar Afonso (Faculdade de Economia, U. Porto; oafonso@fep.up.pt)
João Ferreira do Amaral (ISEG, Universidade de Lisboa; famaral@iseg.utl.pt)
João Sousa Andrade (Faculdade de Economia, U. Coimbra; jasa@fe.uc.pt)
Manuela Arcanjo (ISEG, Universidade de Lisboa; marcanjo@iseg.utl.pt)
Miguel St. Aubyn (ISEG, Universidade de Lisboa; mstaubyn@iseg.utl.pt)
Rui Baptista (DEG-IST, U. Lisboa; rui.baptista@tecnico.ulisboa.pt)
Pedro Brinca (Nova School of Business and Economics; pedro.brinca@novasbe.pt)
Anabela Carneiro (Faculdade de Economia, U. Porto; anacar@fep.up.pt)
Filipe Coelho (Faculdade de Economia, U. Coimbra; fcoelho@fe.uc.pt)
João Paulo Costa (Faculdade de Economia, U. Coimbra; jpaulo@fe.uc.pt)
Rosa Branca Esteves (Escola de Economia e Gestão, U. Minho; rbranca@eeg.uminho.pt)
Pedro Lopes Ferreira (Faculdade de Economia, U. Coimbra; pedrof@fe.uc.pt)
José Rui Figueira (DEG-IST, U. Lisboa; figueira@tecnico.ulisboa.pt)
Francisco Louçã (ISEG, Universidade de Lisboa; flouc@iseg.utl.pt)
Jorge Braga de Macedo (Nova School of Business and Economics; jbmacedo@novasbe.pt)
José Pereirinha (ISEG, Universidade de Lisboa; pereirin@iseg.ulisboa.pt)
Esmeralda Ramalho (ISEG, Universidade de Lisboa; eramalho@iseg.ulisboa.pt)
Joaquim Ramalho (ISCTE, Instituto Universitário de Lisboa; Joaquim.Jose.Ramalho@iscte.pt)
Miguel Portela (Escola de Economia e Gestão, U. Minho; mangelo@eeg.uminho.pt)
Miguel Torres Preto (DEG-IST, U. Lisboa; miguel.preto@tecnico.ulisboa.pt)
Pedro Nogueira Ramos (Faculdade de Economia, U. Coimbra; pnramos@fe.uc.pt)
José Reis (Faculdade de Economia, U. Coimbra; jreis@fe.uc.pt)
Miguel Sousa (Faculdade de Economia, U. Porto; msousa@fep.up.pt)

(Página deixada propositadamente em branco)

SUMÁRIO

Artigos

- Francisco José Veiga
Forecast Errors in Prices and Wages:
The Experience with Three Programme Countries • 7
- Manuel Correia de Pinho / Maria Manuel Pinho
Esforço Fiscal em Portugal:
Uma Avaliação no Período 1995-2015 • 25
- Patrícia Antunes / Pedro Lopes Ferreira / Lara N. Ferreira
A Utilização da Experiência de Escolha
Discreta na Valoração de Estados de Saúde • 47
- Wei-Bin Zhang
A Small Open Economic Growth Model with
Imported Goods, Tourism, and Terms of Trade • 65

(Página deixada propositadamente em branco)

FORECAST ERRORS IN PRICES AND WAGES: THE EXPERIENCE WITH THREE PROGRAMME COUNTRIES

Francisco José Veiga

ABSTRACT

This paper evaluates the accuracy of price and wage forecasts made under the scope of the Greek, Portuguese and Irish Economic Adjustment Programmes (EAPs). Besides comparing the quality of forecasts for the three EAPs, it checks if they were less accurate than forecasts for other euro area countries, and compares the European Commission's forecasts with those of the IMF and OECD. Programme forecasts for Greece were, overall, the least accurate, but those made under the second Greek EAP seem to have gotten closer to actual values. The results also suggest that EC forecasts were not less accurate than those produced by the IMF or the OECD.

Keywords: Forecast errors; Prices and wages; Economic Adjustment Programmes; Greece; Portugal; Ireland.

JEL Classification: E37; E31; E6; J3.

Acknowledgments: Francisco Veiga is thankful for the very helpful comments of Ansgar Belke, Daniel Gros, Miguel Lebre de Freitas, and Antonis Adam.

Received for publication: November 16, 2016

Revision accepted for publication: February 6, 2017

1. INTRODUCTION

Economic outlooks were central pieces of the Economic Adjustment Programmes (EAPs) implemented in European Union countries affected by the sovereign debt crisis. In fact, those economic forecasts were used as a basis for the design of the policy measures Greece, Ireland, Portugal, and Cyprus were asked to adopt. Given this central role of economic forecasts, it is vital that they are of high quality. Unfortunately, forecasts are seldom as accurate as economists and policymakers would desire, and forecast errors can potentially lead to inappropriate policies which generate significant economic costs. That may have been the case in Programme countries such as Greece and Portugal, where austerity measures were especially harsh.

This paper evaluates the forecast performance of the Greek, Irish and Portuguese Programmes in the specific area of prices and wage developments, using data between 2009 and 2014. First, the forecasted behaviour of wages and prices is compared with the values that actually materialized, checking if forecast errors were larger for Greece (where the greatest forecast errors seem to have been committed)¹ than for Portugal and Ireland, and/or for the EU countries in general, and comparing EC forecast errors with those of the IMF and the OECD. Developments in prices are proxied by the annual percentage change in the Harmonized Index of Consumer Prices (HICP), while developments in wages are proxied by annual percentage changes in unit labour costs (ULC).²

The paper is structured as follows. Section 2 provides a graphical presentation of the forecasts and forecast errors, regarding prices and wages, for the Greek, Irish and Portuguese EAPs. Section 3 compares the forecast accuracy of the Programmes' forecasts using measures such as the mean error (ME), the mean absolute error (MAE), and the root mean squared error (RMSE). The following exercise, in Section 4, compares the current year, 1-year ahead and 2-years-ahead forecasts of the EC with those of other international agencies, such as the IMF and the OECD.³ In Section 5, econometric estimations are performed in order to check if error statistics for the 3 programme countries are higher than for other euro area countries. Section 6 concludes the paper.

2. PROGRAMME FORECASTS OF PRICE AND WAGE DEVELOPMENTS

This section presents the successive programmes' forecasts for price and wage developments, until April 2014, and compares them with actual values, taken from the February 2015 release of the European Commission's AMECO database.

¹ An earlier survey on the development of wages, prices and the real exchange rate in Greece vis-à-vis the projections was made in Pisany-Ferry, et al. (2013), p. 57.

² Wage developments were also evaluated using compensation of employees per capita. Since the results are similar to those obtained for ULC, they are not shown here in order to economize space.

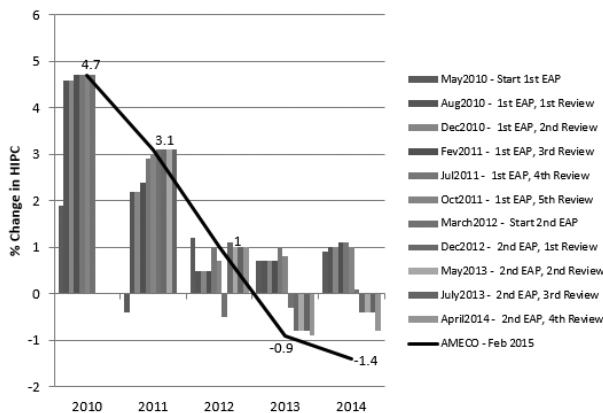
³ EC forecasts were obtained in the publication *European Economy* (several issues), IMF inflation forecasts are from the *World Economic Outlook* (WEO) database, and OECD forecasts are from the *Economic Outlook* database.

2.1 GREECE

The successive inflation forecasts included in the Greek EAPs are presented in Figure 1. A vertical bar is included for the initial forecasts and for those of each of the periodic reviews, while actual values are shown in a black line and with numbers next to it. The initial inflation forecasts (in May 2010) were generally far from the actual values. Forecasts for 2010 and 2011 were well below the values that materialized, while those of 2013 and 2014 were above them. Only the initial forecast for 2012 (1.2%) was close to the actual value (1%). Subsequent forecasts made in 2010 corrected the estimate for that year, but underestimated inflation in 2011 and 2012 and overestimated it in 2013 and 2014. In fact, practically all forecasts overestimated inflation for these latter years.

Inflation forecast errors for the current year, 1-year-ahead and 2-years-ahead are presented in Figure 2. These errors were calculated as the difference between each forecast and the corresponding value obtained from the February 2015 release of AMECO.⁴ As mentioned above, the initial programme forecasts (May 2010) were quite far from the final values. Except for the latter reviews, inflation tended to be underestimated in the forecasts for the current year and 1-year ahead.

Figure 1: Inflation forecasts for Greece



Sources: European Commission, *European Economy, Occasional Papers* and AMECO. Figures 2 to 8 are based on the same sources.

⁴ Contrary to Cabanillas and Terzi (2012), who calculate forecast errors using the ‘first available estimates’ for the current year forecast, and the ‘first settled estimates’ for the year-ahead forecasts, we preferred to use the most recent release of AMECO, which better reflects final values than the first available or settled estimates. The May2013, July 2013 2-years-ahead forecast errors and the April 2014 1-year and 2-years-ahead forecast errors are in fact forecast corrections, as they refer to 2015 and 2016 for which final values are not available yet. That is, in those cases the forecast error/correction was calculated as the difference between those programmes’ forecasts and the February 2015 AMECO forecasts. The same procedure will be used in the following figures for forecast errors in Greece and in other programme countries.

Figure 2: Inflation forecast errors for Greece

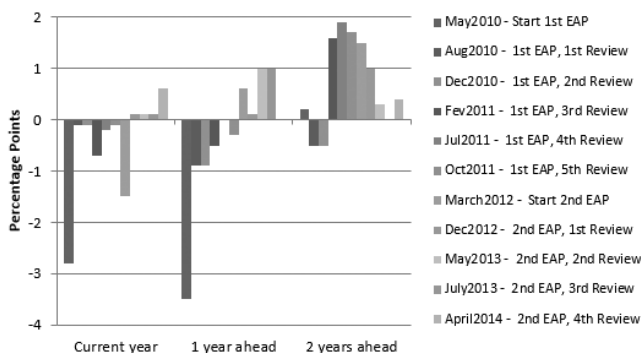
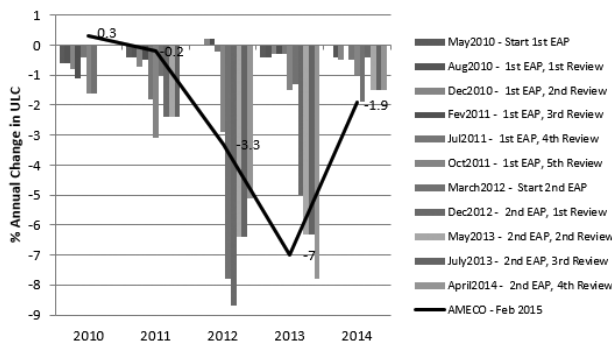


Figure 3 presents the forecasts and actual values of annual percentage changes in unit labour costs (ULC). As happened with inflation, the initial programme forecasts were below the actual values of 2010 and 2011 and above actual values in 2013 and 2014. Figure 4 shows that forecast errors for the current year were mostly negative (except for the last 3 reviews), while for 1-year-ahead and 2-years-ahead were mostly positive (again, with the exception of the last 3 reviews).

Figure 3: Unit Labour Costs forecasts for Greece

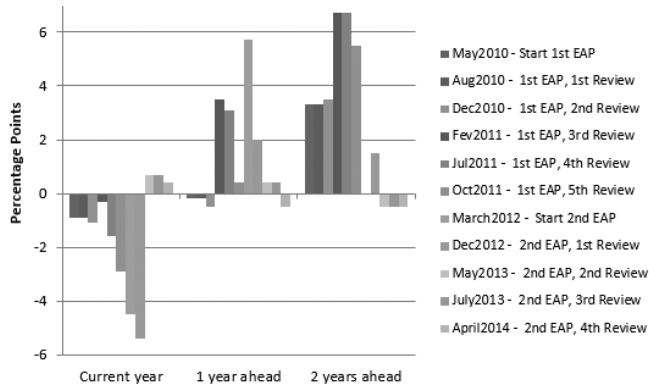


2.2 PORTUGAL

Figure 5 presents HIPC inflation forecast errors for Portugal. Current year forecast errors were generally quite small, except those of the 2014 reviews, which overestimated

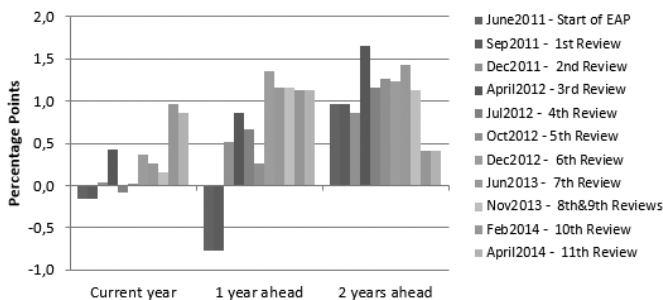
Forecast Errors in Prices and Wages: The Experience with Three Programme Countries

Figure 4: Unit Labour Costs forecast errors for Greece



the inflation rate by almost 1 percentage point. Except for the initial (June 2011) and first review (September 2011) forecasts, inflation forecasts for Portugal tended to overestimate inflation. In fact, that was always the case for 2-years-ahead forecasts. Nevertheless, inflation forecast errors for Portugal never reached 3 percentage points, as they did for Greece (see Figure 2).

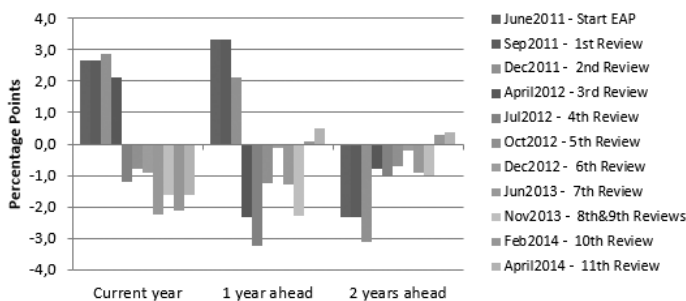
Figure 5: Inflation forecast errors for Portugal



Forecast errors for the annual changes in unit labour costs (ULC) are illustrated in Figure 6. Forecasts made in 2011 for the current year and 1-year-ahead overestimated the growth in ULC, while latter forecasts underestimated it. Two-years-ahead forecasts generally underestimated the change in ULC. Although errors are generally higher for ULC than for

inflation, the same happened for Greece, which had even higher forecast errors in ULC than Portugal, in some occasions exceeding 6 percentage points (see Figure 4).

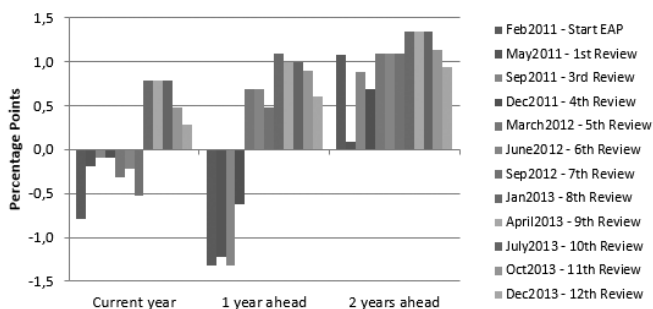
Figure 6: Unit Labour Costs forecast errors for Portugal



2.3 IRELAND

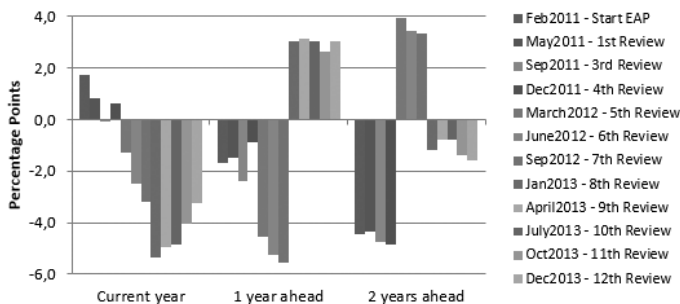
A pattern similar to that of inflation forecasts for Portugal is observed for Ireland in Figure 7, as initial current year and 1-year-ahead forecasts tended to underestimate inflation, while latter ones overestimated it. As for Portugal, 2-years-ahead forecasts always overestimated inflation, and the biggest forecast errors were around 1 percentage point.

Figure 7: Inflation forecast errors for Ireland



Forecast errors in changes in ULC for Ireland, shown in Figure 8, were generally higher than those for Portugal, and comparable to those for Greece. In fact, changes in ULC were sometimes underestimated by almost 6 percentage points.

Figure 8: Unit Labour Costs forecast errors for Ireland



3. FORECAST PERFORMANCE OF THE THREE ECONOMIC ADJUSTMENT PROGRAMMES

This section compares the forecast performance of the three EAPs (Greece, Portugal, and Ireland), using three summary statistics: the mean error (ME), the mean absolute error (MAE), and the root mean squared error (RMSE).⁵

3.1 FORECAST ERRORS AND SUMMARY STATISTICS

Following Cabanillas and Terzi (2012), and previous studies, we define the forecast errors and the summary statistics as explained below.

The forecast error is the difference between the forecast and the realization. That is:

$$\begin{aligned}
 e_{t,t} &= y_{t,t} - \mathcal{Y}_t, \text{ for the current year} \\
 e_{t+1,t} &= y_{t+1,t} - \mathcal{Y}_{t+1}, \text{ for the year ahead} \\
 e_{t+2,t} &= y_{t+2,t} - \mathcal{Y}_{t+2}, \text{ for two years ahead,}
 \end{aligned}$$

where $y_{t,t}$, $y_{t+1,t}$ and $y_{t+2,t}$ are the forecasts of y made at period t for t , $t+1$ and $t+2$, respectively; and y_t , y_{t+1} and y_{t+2} are the realizations of y in years t , $t+1$ and $t+2$, respectively.

The programmes' forecast accuracy is assessed using three commonly used summary statistics:

The mean error (ME), which is the average difference between the forecast and the outcome. Although it has the problem that negative errors may compensate positive ones, not being a good accuracy indicator, it may help signal biases in the forecasts.

⁵ These statistics were used by Artis (1996) to evaluate IMF-WEO forecasts, and by Keereman (1999), Melander et al. (2007), and Cabanillas and Terzi (2012) to evaluate the accuracy of the European Commission's forecasts. Although these studies only considered the current year and the year ahead forecasts, we also analyse the 2-years-ahead forecasts. See also Öller and Barot (2000), who analysed 1-year-ahead forecasts by the OECD and by national institutes of GDP growth and inflation in 13 European countries.

$$ME = \frac{1}{T} \sum_{t=1}^T e_{t,t}, \text{ for the current year}$$

$$ME = \frac{1}{T} \sum_{t=1}^T e_{t+1,t}, \text{ for the year ahead}$$

$$ME = \frac{1}{T} \sum_{t=1}^T e_{t+2,t}, \text{ for two years ahead.}$$

The mean absolute error (MAE), which is the average absolute difference between the forecast and the outturn. Since positive and negative errors no longer cancel each other out, it is a more accurate measure of the average forecast error than the ME.

$$MAE = \frac{1}{T} \sum_{t=1}^T |e_{t,t}|, \text{ for the current year}$$

$$MAE = \frac{1}{T} \sum_{t=1}^T |e_{t+1,t}|, \text{ for the year ahead}$$

$$MAE = \frac{1}{T} \sum_{t=1}^T |e_{t+2,t}|, \text{ for two years ahead.}$$

The root mean squared error (RMSE), which is a measure of the relative size of the forecast errors, gives greater weight to large than to small errors.

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T e_{t,t}^2}, \text{ for the current year}$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T e_{t+1,t}^2}, \text{ for the year ahead.}$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T e_{t+2,t}^2}, \text{ for two years ahead.}$$

3.2 PROGRAMMES` FORECAST ACCURACY

Table 1 shows the summary statistics for the inflation forecasts made in the EAPs of Greece, Portugal, and Ireland. As in the figures shown above, all available programme forecasts (initial and of all reviews) for each country were considered.

Regarding current year forecasts, the mean error for Greece is negative, indicating that forecasts, on average, underestimated inflation in that country. The positive ME in Portugal, which is close to the MAE, indicates that inflation forecasts for Portugal tended to overestimate actual inflation. For Ireland, the ME is much smaller than the MAE, which indicates that positive errors were in large part cancelled out by negative ones. The MAE is higher for Greece than for the other countries, although the difference to Ireland is not very big. But, the RMSE for Greece is twice as big as that for Ireland (the second largest), which is due to the large errors made in the initial inflation forecasts for Greece.

Table 1: Programme Inflation forecasts statistics

Country	Current Year			1-year-ahead			2-years-ahead		
	ME	MAE	RMSE	ME	MAE	RMSE	ME	MAE	RMSE
Greece	-0,42	0,58	1,00	-0,31	0,80	1,23	0,69	0,87	1,09
Portugal	0,24	0,32	0,44	0,61	0,89	0,94	1,04	1,04	1,11
Ireland	0,08	0,44	0,52	0,16	0,91	0,95	1,01	1,01	1,07

Sources: European Commission, *European Economy, Occasional Papers* (several issues) and AMECO.

Although the 1-year-ahead forecasts have a somewhat similar pattern (especially for the ME and RMSE), those of 2-years-ahead are quite different, with Portugal exhibiting the largest ME, MAE and RMSE. Thus, it is interesting to note that forecast accuracy for Portugal was the highest when considering the current year, but it was the lowest for 2-years-ahead. For 1-year-ahead, Greece has the lowest accuracy according to the RMSE (which penalizes the large errors of the initial programme forecasts), and the highest according to the MAE.

Table 2 presents the summary statistics for the forecasts of changes in unit labour costs (ULC). Regarding current year forecasts, the lowest accuracy was achieved for Ireland, according to any of the measures. The same applies to 1-year-ahead forecasts, according to the MAE and RMSE. As for 2-years-ahead forecasts, the lowest accuracy was for Greece, according to all statistics. Additionally, forecasts for Greece were always less accurate than those for Portugal.

Table 2: Programme ULC forecasts statistics

Country	Current Year			1-year-ahead			2-years-ahead		
	ME	MAE	RMSE	ME	MAE	RMSE	ME	MAE	RMSE
Greece	-1,44	1,76	2,42	1,28	1,54	2,32	2,64	2,91	3,78
Portugal	-0,02	1,89	2,01	-0,11	1,80	2,15	-1,07	1,19	1,51
Ireland	-2,18	2,72	3,23	-0,57	3,04	3,34	-1,13	2,90	3,29

Sources: European Commission, *European Economy, Occasional Papers* (several issues) and AMECO.

In sum, there is some indication that forecasts for Greece were less accurate than for Portugal and Ireland regarding inflation (for current year and 1-year-ahead forecasts). Regarding ULC forecasts, those for Greece were still less accurate than for Portugal, but more accurate than those for Ireland.

4. COMPARISON OF EC FORECAST ERRORS WITH THOSE OF THE IMF AND OECD

In order to check if the European Commission's forecasts were more or less accurate than those of the IMF and OECD, we show in this section the summary statistics of those

institutions' forecasts. Following Cabanillas and Terzi (2012), current-year forecasts are taken from the Commission's spring forecasts, whereas 1-year-ahead forecasts come from the Commission's autumn forecasts. The 2-years-ahead forecasts, not used in that study, are also taken from the autumn forecasts. The same procedure is used regarding the IMF forecasts, taken from its *World Economic Outlook*, and the OECD forecasts, taken from its *Economic Outlook*.⁶ The first forecast considered for each institution is that of autumn 2009 and the last is that of autumn 2014. We start by comparing forecasts for Greece, Portugal and Ireland, and then we show Ordinary least Squares (OLS) regression results for estimations including other euro area countries.

4.1 FORECASTS STATISTICS FOR GREECE, PORTUGAL AND IRELAND

The forecast summary statistics for the inflation forecasts of the EC, IMF and OECD, for Greece, Portugal and Ireland, are shown in Table 3. IMF forecasts were generally the least accurate, according to the MAE and RMSE, regardless of the country and time horizon considered.⁷ Regarding current year forecasts, the OECD performed somewhat better than the EC for Greece and Ireland, while their accuracy was practically the same for Portugal. The EC's 1-year-ahead forecasts were generally more accurate than those of the OECD for Greece, slightly better for Ireland, and about the same for Portugal (the EC has a lower MAE, but a higher RMSE). Regarding 2-years-ahead forecasts, the EC performed worse than the OECD for Greece and Ireland, but better for Portugal.

Table 4 shows the summary statistics for EC and OECD forecasts of annual percentage changes in unit labour costs.⁸ The EC forecasts for Greece were more accurate than the OECD's when considering the current year and the year ahead, but the opposite happened for 2-years-ahead. Regarding forecasts of changes in ULC for Ireland, the OECD was always more accurate than the EC. A comparison is not possible for Portugal, as there were no ULC forecasts available for this country in the OECD's *Economic Outlook*.

In sum, it is not possible to state that the European Commission's forecasts were always the most accurate. But, it is safe to argue that the EC's inflation forecasts for the 3 programme countries considered were always more accurate than those of the IMF.⁹ Additionally, the

⁶ The Spring forecasts of the EC were generally released in May, those of the IMF in April, and those of the OECD in June. The Fall forecasts of the EC were generally released in November, those of the IMF in October, and those of the OECD in December. Although the release dates were not exactly the same, they were close enough to make EC, IMF and OECD forecasts comparable. Nevertheless, the fact that the IMF is the first and the OECD is the last institution to present forecasts, may be a disadvantage for the former and an advantage for the latter.

⁷ As mentioned in the previous footnote, the IMF is always the first of these three institutions to publish its forecasts. Using slightly more recent information and knowing the IMF forecasts may help the EC and the OECD produce better forecasts.

⁸ There are no forecasts for unit labour costs or compensation of employees in the IMF's WEO.

⁹ This may seem strange, as both the EC and the IMF were part of the Troika administering the Greek EAP, and we would expect their forecasts to be aligned. In practice, however, Programme forecasts seem to be more aligned with EC forecasts than with those of the IMF.

Table 3: Inflation forecast statistics

	Current Year (Spring Forecasts)			1-year-ahead (Fall Forecasts)			2-years-ahead (Fall Forecasts)		
	ME	MAE	RMSE	ME	MAE	RMSE	ME	MAE	RMSE
European Commission									
Greece	-0,60	0,86	1,05	0,16	0,55	0,73	0,75	0,93	1,15
Portugal	0,10	0,32	0,35	0,21	0,63	0,79	0,46	1,01	1,19
Ireland	0,19	0,35	0,41	-0,18	0,69	0,76	0,30	0,75	0,89
International Monetary Fund (World Economic Outlook)									
Greece	-1,02	1,22	1,59	-0,88	1,02	1,46	-0,06	1,05	1,26
Portugal	-0,09	0,61	0,69	-0,44	0,93	1,19	-0,28	1,29	1,46
Ireland	-0,11	0,43	0,52	-0,13	1,09	1,18	0,07	0,82	0,84
OECD (Economic Outlook)									
Greece	-0,41	0,46	0,77	-0,63	0,92	1,31	-0,25	0,75	0,89
Portugal	-0,19	0,33	0,35	-0,21	0,66	0,75	-0,40	1,21	1,45
Ireland	0,17	0,17	0,24	0,15	0,71	0,77	-0,05	0,56	0,58

Sources: European Commission, *European Economy, Economic Forecast* and AMECO; International Monetary Fund, *World Economic Outlook*; and OECD, *Economic Outlook*.

Table 4: Unit labour costs forecast statistics

	Current Year (Spring Forecasts)			1-year-ahead (Fall Forecasts)			2-years-ahead (Fall Forecasts)		
	ME	MAE	RMSE	ME	MAE	RMSE	ME	MAE	RMSE
European Commission									
Greece	-0,96	1,30	2,20	0,60	0,76	1,07	1,71	1,73	2,55
Portugal	0,47	1,25	1,51	0,67	1,01	1,22	0,22	1,03	1,41
Ireland	0,20	3,08	3,49	-1,68	2,36	2,96	-0,71	1,82	2,49
OECD (Economic Outlook)									
Greece	2,45	2,45	3,54	1,31	2,50	3,35	2,74	3,66	4,42
Ireland	-1,76	1,76	2,24	-0,60	1,90	2,47	-1,10	1,70	2,01

Sources: European Commission, *European Economy, Economic Forecast* and AMECO; and OECD, *Economic Outlook*.

EC forecasts for inflation and ULC do not seem to be, on average, less accurate than those of the OECD.

4.2 COMPARISON OF FORECAST ERRORS FOR PROGRAMME COUNTRIES WITH THOSE FOR OTHER EURO AREA COUNTRIES

In order to check if the forecast accuracy for the three programme countries analysed above was different from that for other 14 euro area countries,¹⁰ we estimated individual OLS cross-country regressions for each variable (inflation and ULC), institution (EC, IMF and OECD), and summary statistic (ME, MAE and RMSE). In each regression, the dependent variable is a summary statistic, and the explanatory variables are the dummy variables for the three programme countries and a constant. The estimated model is the following:

$$FS_i = \alpha + \beta \text{Greece}_i + \gamma \text{Portugal}_i + \delta \text{Ireland}_i + \varepsilon_i,$$

where FS_i is a forecast statistic (ME, MAE or RMSE) for country i ; Greece_i , Portugal_i , and Ireland_i are dummy variables that take the value of one for the respective countries and zero otherwise; ε_i is a white noise error term; and α , β , γ , and δ are parameters to be estimated.

The results of the regressions for inflation forecast statistics are presented in Table 5.¹¹ Regarding current year forecasts, the dummy variable for Greece is positive and statistically significant for the MAE and RMSE, which indicates that EC inflation forecasts for Greece were less accurate than for the remaining euro area countries. The same applies to the IMF and the OECD. There is also some indication that IMF and OECD forecasts for Ireland were also less accurate than for the other euro area countries (but more accurate than those for Greece). Finally, the regression for the MAE, indicates that OECD inflation forecasts for Portugal may have been more accurate than for the other euro area countries.

A different picture emerges for the 1-year-ahead forecasts, as EC forecasts for the three programme countries seem to have been more accurate than for the remaining euro area countries. This is not true for IMF forecasts, since there is a positive and statistically significant coefficient for Greece in the RMSE regression, and the dummy variables for Portugal and Ireland are not statistically significant. As for OECD year ahead forecasts, there is weak evidence of higher errors for Greece (for the RMSE) and lower ones for Portugal and Ireland, than for the remaining euro area countries.

Regarding the 2-years-ahead forecasts of the EC, there is no robust evidence of different forecast accuracy for the programme countries. For the IMF, there is indication of higher errors for Ireland and lower ones for Portugal. Finally, the OECD 2-years-ahead forecasts seem to have been more accurate for Greece and Portugal.

¹⁰ Only the first 17 members of the euro area were considered. We did not include Latvia and Lithuania, as they were not euro area members during the entire period under analysis.

¹¹ As the ME is not a good measure of accuracy, we will only refer to it when it clearly indicates biased forecasts. Thus, most of the comments in this section refer to the results for the MAE and the RMSE. The R-Squared of each estimation is also reported in Table 5.

Table 6 presents the results of regressions for the forecasts of changes in unit labour costs (ULC) by the European Commission and the OECD. Current year ULC forecasts for the programme countries seem to have been less accurate than for the other euro area countries. In fact, all dummy variables for programme countries are positive and statistically significant in the regressions for the MAE and the RMSE. The same applies to 1-year-ahead forecasts of the EC for Portugal and of the OECD for Greece and Ireland. Regarding 2-years-ahead forecasts, they seem to have been less accurate for Greece and Portugal than for the remaining euro area countries. Thus, overall, ULC forecasts for the 3 programme countries tended to be less accurate than for the other euro area members.

Table 5: Regressions for inflation forecasts

	Current Year (Spring Forecasts)			1-year-ahead (Fall Forecasts)			2-years-ahead (Fall Forecasts)		
	ME	MAE	RMSE	ME	MAE	RMSE	ME	MAE	RMSE
European Commission									
Greece	-0.593*** (-11.71)	0.471*** (8.023)	0.610*** (8.883)	0.0162 (0.283)	-0.258*** (-5.976)	-0.263*** (-4.406)	0.338*** (3.838)	0.0707 (0.910)	0.0701 (0.762)
Portugal	0.198*** (3.900)	-0.0369 (-0.628)	-0.0290 (-0.422)	-0.331*** (-5.794)	-0.121** (-2.810)	-0.232*** (-3.898)	-0.104 (-1.181)	-0.110 (-1.416)	-0.192* (-2.083)
Ireland	0.108* (2.122)	-0.0719 (-1.224)	-0.0959 (-1.397)	0.0593 (1.036)	-0.181*** (-4.196)	-0.202*** (-3.387)	0.0468 (0.531)	0.158* (2.029)	0.105 (1.141)
R-squared	0.508	0.299	0.343	0.183	0.255	0.202	0.095	0.043	0.040
International Monetary Fund (World Economic Outlook)									
Greece	-1.007*** (-13.22)	0.737*** (13.22)	1.003*** (11.63)	-0.660*** (-7.392)	0.0798 (0.864)	0.409*** (4.673)	-0.0355 (-0.254)	0.0251 (0.369)	0.0723 (1.024)
Portugal	-0.0892 (-1.171)	-0.0524 (-0.940)	-0.0736 (-0.854)	0.0931 (1.042)	0.143 (1.552)	0.125 (1.428)	0.0937 (0.671)	-0.204** (-2.989)	-0.353*** (-5.002)
Ireland	-0.0706 (-0.927)	0.1333** (2.386)	0.103 (1.190)	-0.221** (-2.469)	-0.0162 (-0.175)	0.130 (1.484)	-0.256* (-1.836)	0.272*** (3.993)	0.265*** (3.748)
R-squared	0.521	0.530	0.463	0.276	0.019	0.132	0.025	0.142	0.211
OECD (Economic Outlook)									
Greece	a-0.410*** (-8.017)	0.218*** (8.797)	0.488*** (16.15)	-0.295** (-2.637)	-0.0223 (-0.251)	0.221* (2.049)	0.00560 (0.0413)	-0.403** (-2.803)	-0.496** (-2.852)
Portugal	0.178*** (3.478)	-0.0694** (-2.797)	-0.0426 (-1.408)	0.478*** (4.279)	-0.233** (-2.617)	-0.314** (-2.914)	0.211 (1.556)	-0.596*** (-4.150)	-0.800*** (-4.606)
Ireland	-0.185*** (-3.606)	0.0889*** (3.585)	0.0695** (2.299)	0.120 (1.078)	-0.281*** (-3.160)	-0.330** (-3.067)	-0.144 (-1.062)	0.0591 (0.411)	0.0737 (0.424)
R-squared	0.446	0.465	0.702	0.197	0.121	0.166	0.032	0.175	0.199

Sources: European Commission, *European Economy, Economic Forecast* and AMECO; IMF, *World Economic Outlook*; and OECD, *Economic Outlook*.

Notes: OLS models estimated for a sample of 17 euro area countries, from autumn 2009 to autumn 2014. The estimated coefficient for the constant is not shown. T-statistics are in parentheses. Significance level at which the null hypothesis is rejected: ***, 1%; **, 5%, and *, 10%.

Table 6: Regressions for forecasts of annual changes in Unit Labour Costs (ULC)

	Current Year (Spring Forecasts)			1-year-ahead (Fall Forecasts)			2-years-ahead (Fall Forecasts)		
	ME	MAE	RMSE	ME	MAE	RMSE	ME	MAE	RMSE
European Commission									
Greece	-0.799*** (-6.392)	0.488** (2.838)	1.186*** (5.605)	0.688** (2.983)	-0.169 (-0.995)	-0.220 (-0.922)	1.670*** (5.864)	0.708*** (3.985)	1.150*** (4.888)
Portugal	0.359** (2.870)	2.276*** (13.25)	2.470*** (11.68)	-1.594*** (-6.909)	1.424*** (8.382)	1.673*** (7.001)	-0.749** (-2.628)	0.807*** (4.540)	1.084*** (4.608)
Ireland	0.633*** (5.063)	0.437** (2.546)	0.489** (2.313)	0.756*** (3.276)	0.0796 (0.469)	-0.0701 (-0.293)	0.186 (0.654)	0.0173 (0.0971)	0.00743 (0.0316)
R-squared	0.332	0.531	0.501	0.310	0.312	0.242	0.214	0.177	0.210
OECD (Economic Outlook)									
Greece	2.751*** (14.44)	1.577*** (15.81)	2.473*** (17.58)	1.824*** (6.143)	1.365*** (7.218)	1.895*** (7.952)	3.645*** (9.720)	2.111*** (9.642)	2.591*** (10.34)
Ireland	-1.456*** (-7.640)	0.870*** (8.819)	1.177*** (8.365)	-0.0880 (-0.296)	0.768*** (4.060)	1.016*** (4.264)	-0.199 (-0.530)	0.147 (0.672)	0.177 (0.707)
R-squared	0.700	0.715	0.745	0.238	0.345	0.385	0.440	0.431	0.466

Sources: European Commission, *European Economy, Economic Forecast* and AMECO; and OECD, *Economic Outlook*.

Notes: OLS models estimated for a sample of 17 euro area countries (EC) or 14 euro area countries (OECD), from autumn 2009 to autumn 2014. The estimated coefficient for the constant is not shown. T-statistics are in parentheses. Significance level at which the null hypothesis is rejected: ***, 1%; **, 5%, and *, 10%.

4.3 COMPARISON OF FORECAST ERRORS COMBINING THE THREE INSTITUTIONS

A final econometric analysis consisted on estimating regressions for a sample which combines the forecast statistics of the 3 institutions. In order to check if the EC's forecasts were in general more or less accurate than those of the IMF and OECD, dummy variables for those institutions were added to the list of explanatory variables. The results are presented in Table 7. In the regressions for current year inflation forecasts, the dummy variable for the IMF is positive, while that for the OECD is negative. Since both are statistically significant, this indicates that EC inflation forecasts were, on average, more accurate than those of the IMF, but less accurate than those of the OECD. As for ULC, the OECD dummy variable is not statistically significant, indicating that there was no difference in accuracy relative to EC forecasts. MAE regressions indicate that EC inflation forecasts 1 and 2-years-ahead may have been more accurate than those of the IMF and OECD. For ULC, there is no difference in 1-year-ahead forecasts, but the EC seems to have been more accurate than the OECD in 2-years-ahead forecasts (in the MAE estimation).

The results for the country dummy variables indicate that inflation current year forecasts were, on average (and considering the 3 institutions), less accurate for Greece than for the remaining euro area countries. As for ULC forecasts (by the EC and OECD), they were less accurate for the 3 programme countries, especially for Greece and Portugal, than for the others. For 1-year-ahead forecasts, the MAE regression indicates that greater than average accuracy may have been achieved in inflation forecasts for Ireland, while there was lower than average accuracy in the 1-year-ahead ULC forecasts for Portugal. Finally, concerning 2-years-ahead, inflation forecasts for Portugal were more accurate and those for Ireland were less accurate than average. ULC forecasts were less accurate for Greece than for the other countries.

Table 7: Regressions forecasts combining the three institutions

	Current Year (Spring Forecasts)			1-year-ahead (Fall Forecasts)			2-years-ahead (Fall Forecasts)			
	ME	MAE	RMSE	ME	MAE	RMSE	ME	MAE	RMSE	
Inflation	Greece	-0.671*** (-4.640)	0.476*** (3.835)	0.701*** (5.332)	-0.315* (-1.853)	-0.0661 (-0.744)	0.123 (0.750)	0.103 (0.875)	-0.100 (-0.807)	-0.115 (-0.741)
	Portugal	0.0948 (1.278)	-0.0521* (-1.816)	-0.0479 (-1.157)	0.0781 (0.389)	-0.0694 (-0.736)	-0.140 (-1.291)	0.0667 (0.624)	-0.301** (-2.450)	-0.446*** (-2.955)
	Ireland	-0.0499 (-0.630)	0.0507 (0.886)	0.0260 (0.432)	-0.0154 (-0.172)	-0.159** (-2.293)	-0.133 (-1.204)	-0.118 (-1.223)	0.165** (2.478)	0.150* (1.977)
	IMF	-0.0603 (-0.806)	0.118* (1.824)	0.179** (2.047)	-0.400*** (-4.380)	0.182** (2.249)	0.140 (1.568)	-0.460*** (-3.559)	0.164* (2.014)	0.110 (1.198)
	OECD	-0.000449 (-0.00761)	-0.154*** (-2.966)	-0.159** (-2.677)	-0.443*** (-4.290)	0.136* (1.800)	0.103 (1.062)	-0.678*** (-5.510)	0.229* (1.826)	0.222 (1.467)
	R-squared	0.430	0.507	0.524	0.389	0.133	0.090	0.390	0.170	0.159
Unit Labour Costs	Greece	0.972 (0.731)	1.033** (2.353)	1.830*** (3.572)	1.249*** (3.030)	0.596 (1.023)	0.833 (1.047)	2.651*** (3.617)	1.408** (2.609)	1.869*** (3.296)
	Portugal	-0.553 (-0.734)	1.579*** (2.924)	1.823*** (3.549)	-0.848 (-1.558)	1.093*** (3.622)	1.341*** (4.018)	-0.480* (-1.777)	0.475 (1.555)	0.629 (1.574)
	Ireland	0.686*** (4.656)	0.428** (2.679)	0.489** (2.514)	0.838*** (3.946)	0.107 (0.677)	-0.0250 (-0.113)	0.264 (1.014)	0.0402 (0.245)	0.0239 (0.111)
	OECD	-0.0246 (-0.0880)	0.0466 (0.247)	0.0486 (0.210)	-0.253 (-0.773)	0.257 (1.126)	0.262 (0.865)	-0.772* (-1.904)	0.582** (2.347)	0.463 (1.548)
	R-squared	0.169	0.469	0.523	0.230	0.252	0.224	0.380	0.373	0.344

Notes: OLS models estimated for a sample of 17 euro area countries, from autumn 2009 to autumn 2014. There are 49 observations for inflation and 31 for ULC. The estimated coefficient for the constant is not shown. T-statistics are in parenthesis. Significance level at which the null hypothesis is rejected: ***, 1%; **, 5%, and *, 10%.

Sources: European Commission, *European Economy, Economic Forecast* and AMECO; IMF, *World Economic Outlook*; and OECD, *Economic Outlook*.

5. CONCLUSIONS

The analysis presented in this study leads us to conclude that programme forecasts for Greece were less accurate than for Portugal and Ireland regarding inflation. Initial inflation forecasts (in May 2010) for Greece, for the current year and the year ahead, were far below from the actual values. The same, although to a smaller degree, happened with the initial forecasts of the second adjustment programme for Greece. Forecasts for unit labour costs (ULC) were less accurate for Ireland regarding the current year and year ahead, and less accurate for Greece for 2-years-ahead. Comparing to ULC programme forecasts for Portugal, forecasts for Greece were always less accurate. Although programme forecasts for Greece were, overall, the least accurate of the 3 programmes' forecasts, those made in more recent reviews of the Greek EAP seem to have gotten closer to actual values.

The forecasts errors in programme countries may in part have been due to greater difficulty to forecast in more volatile environments. In fact, IMF and OECD forecasts for these three programme countries also had relatively large errors, with IMF inflation forecasts being less accurate than those of the European Commission for any of the three countries.¹² Additionally, current year inflation and ULC forecasts for Greece were generally less accurate than those for the other euro area countries, regardless if they were made by the EC, IMF, or OECD. Actually, ULC forecasts for Greece, Ireland and Portugal were always less accurate than for the other euro area members.

Overall, EC forecast do not seem to be less accurate than those of the IMF and OECD. In fact, regarding inflation, they are on average more accurate than those of the IMF.

REFERENCES

- Artis, M. (1996) How accurate are the IMF's short-term economic forecasts? Another examination of the World Economic Outlook, IMF Working Paper, WP/96/89, August.
- Cabanillas, L. G.; Terzi, A. (2012) The accuracy of the European Commission's forecasts re-examined. Economic Papers No. 476, European Commission, Directorate-General for Economic and Financial Affairs, Brussels, Belgium.
- European Commission (2009-2014) European Economy, Economic Forecast, European Commission, Directorate-General for Economic and Financial Affairs, Brussels, Belgium. Forecasts consulted: Spring 2009, Autumn 2009, Spring 2010, Autumn 2010, Spring 2011, Autumn 2011, Spring 2012, Autumn 2012, Winter 2013, Spring 2013, Autumn 2013, Winter 2014, Spring 2014, Autumn 2014.
- European Commission (2009-2015) European Economy, Occasional Papers, European Commission, Directorate-General for Economic and Financial Affairs, Brussels, Belgium. Numbers consulted: Greek EAPs 61, 68, 72, 77, 82, 87, 94, 123, 148, 159, 192; Irish EAP 76, 78, 84, 88, 93, 96, 115, 127, 131, 154, 162, 167, 195, 210; Portuguese EAP 79, 83, 89, 95, 111, 117, 124, 153, 164, 171, 191, 202, 208.

¹² Again, it is worth noting that the EC and, especially, the OECD could have taken the IMF forecasts into account when computing their own. Furthermore, they may use slightly more recent data, which also helps producing more reliable forecasts.

- International Monetary Fund (2009-2014) World Economic Outlook Database, IMF, Washington, DC. Issues: October 2009, April 2010, October 2010, April 2011, September 2011, April 2012, October 2012, April 2013, October 2013, April 2014, October 2014.
- Keereman, F. (1999) The track record of the Commission forecasts, Economic Papers No. 137, European Commission, Directorate-General for Economic and Financial Affairs, Brussels, Belgium.
- Melander, A.; Sismanidis, G.; Grenouilleau, D. (2007) The track record of the Commission's forecasts: An update, Economic Papers No. 291, European Commission, Directorate-General for Economic and Financial Affairs, Brussels, Belgium.
- OECD (2009-2014) Economic Outlook Database. OECD, Paris. Issues: December 2009, June 2010, December 2010, June 2011, December 2011, June 2012, December 2012, June 2013, December 2013, May 2014, November 2014
- Öller; L.E.; Barot, B. (2000) The accuracy of European growth and inflation forecasts, *International Journal of Forecasting*, 16 (3), 293–315.
- Pisany-Ferry, J. Sapir, A.; Wolff, G. (2013) *EU-IMF Assistance to Euro-Area Countries: An Early Assessment*, Bruegel Blueprint 19, Brussels.

(Página deixada propositadamente em branco)

ESFORÇO FISCAL EM PORTUGAL: UMA AVALIAÇÃO NO PERÍODO 1995-2015

Manuel Correia de Pinho
Maria Manuel Pinho

RESUMO

O nível de fiscalidade praticada por um país tem implicações nos planos económico e social, podendo ser aferido por recurso aos conceitos de receita fiscal, de carga fiscal e de esforço fiscal. O esforço fiscal é expresso por um índice que relaciona a carga fiscal (rácio entre a receita fiscal e o PIB) com uma de três medidas alternativas: o PIB *per capita*, uma carga fiscal padrão e a capacidade fiscal da economia. Com base em dados para o período 1995-2015 e para os 28 Estados-membros da União Europeia, procede-se à estimação dos índices de esforço fiscal através de metodologias diferentes. Em três dos quatro índices calculados, os resultados sugerem que o esforço fiscal em Portugal seguiu uma tendência de redução entre 1996 e 2010 e de aumento no período de implementação do programa de ajustamento económico e financeiro de 2011-2014, sem que, no entanto, tenham sido atingidos os valores observados em 1996. No quarto índice calculado, os resultados apontam no sentido de uma tendência de aumento do esforço fiscal em Portugal desde 1996. A evidência empírica sugere ainda que o esforço fiscal em Portugal tem sido inferior à média dos países da União Europeia. Palavras-chave: carga fiscal, capacidade fiscal, fronteira fiscal estocástica, esforço fiscal.

ABSTRACT

The tax level in a country has an impact in both the economic and the social spheres and can be assessed using the concepts of tax revenue, tax burden and tax effort. The tax effort is expressed by an index relating the tax burden (ratio between the tax revenue and GDP) with one of three alternative measures: GDP *per capita*, benchmark tax burden and tax capacity. Using data for the 1995-2015 period and the 28 European Union Member States, we estimate tax effort indexes on the basis of different methodologies. For three of the four computed indices, the results suggest that the tax effort in Portugal followed a downward trend between 1996 and 2010 and an upward one in the period of the implementation of the 2011-2014 economic and financial adjustment program, without, however, having reached the figures observed in 1996. For the fourth index, the results point towards a trend of increase in the tax effort in Portugal since 1996. The empirical evidence also suggests that the tax effort in Portugal has remained below the EU Member States average. Keywords: tax burden, tax capacity, stochastic tax frontier, tax effort

JEL Classification: C51, E62, H21

Received for publication: November 14, 2016

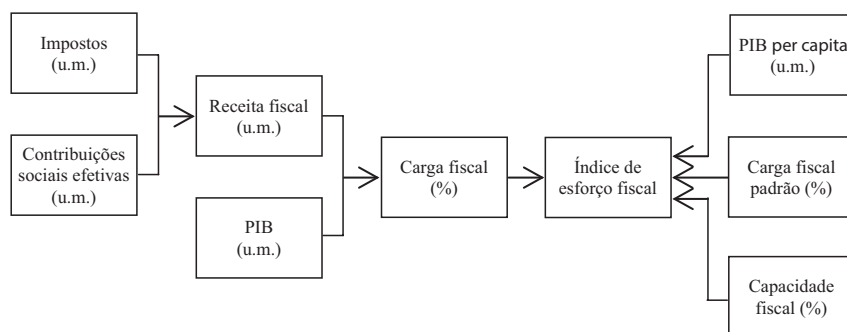
Revision accepted for publication: February 3, 2017

1. INTRODUÇÃO

O nível de fiscalidade praticado por um país tem implicações nos planos económico e social. Em primeira instância, o nível de fiscalidade oferece uma perceção da medida em que as sociedades apreciam a maior ou menor provisão pública de bens. Porém, o nível de fiscalidade é também um fator determinante da competitividade externa da economia e pode ter um efeito de *crowding out* sobre os recursos afetos à atividade económica das famílias e das empresas. Importa, ainda, sublinhar que o nível de fiscalidade é uma variável que condiciona a definição da política económica.

O nível de fiscalidade é, aqui, aferido por recurso aos conceitos de receita fiscal, carga fiscal¹ e esforço fiscal. A receita fiscal corresponde aos montantes arrecadados com impostos e contribuições sociais efetivas e é expressa em unidades monetárias. A carga fiscal corresponde ao quociente entre a receita fiscal e o Produto Interno Bruto (PIB) a preços de mercado e é expressa em percentagem. O esforço fiscal corresponde à relativização da carga fiscal pelo PIB per capita ou a uma comparação entre a carga fiscal e uma carga fiscal virtual resultante de um padrão vigente num conjunto de países e/ou num conjunto de períodos (carga fiscal padrão) ou, ainda, a uma comparação entre a carga fiscal e o máximo de receita fiscal cobrável por um país em função das suas características económicas, sociais e institucionais (capacidade fiscal) (Pessino e Fenochietto, 2010: 65). Em qualquer dos casos, o esforço fiscal é expresso por um índice que é determinado de acordo com um dos três modos esquematizados na Figura 1.

Figura 1: Modos de determinação do esforço fiscal



A carga fiscal padrão e a capacidade fiscal são estimadas por modelos econométricos que produzem valores dependentes quer da estrutura dos modelos quer das variáveis explicativas utilizadas determinando, em consequência, índices de esforço fiscal com valores diferentes. A literatura exhibe divergências, por vezes acentuadas, nos resultados obtidos pelas diversas investigações. A procura de robustez para as análises efetuadas com base nesses índices

¹ Em Pereira *et al.* (2012), Santos (2013) e Pereira (2014), este conceito é designado por nível de fiscalidade.

estimula, por isso, o aprofundamento dos estudos com diferentes quadros metodológicos, bases de dados e especificações.

Cabe anotar que os índices de esforço fiscal não têm um significado concreto, mas são indicadores de ordenação no espaço e de evolução no tempo. No entanto, estes indicadores devem ser usados com ponderação em comparações entre países ou períodos de tempo porque não refletem a qualidade dos bens de provisão pública que os impostos proporcionam. Acresce, ainda, que a medida das grandezas subjacentes pode incorporar alterações concetuais e metodológicas no espaço e no tempo.

Este estudo pretende analisar a evolução do esforço fiscal em Portugal entre 1995 e 2015. Em particular, procura-se avaliar se o aumento da carga fiscal decorrente da aplicação do plano de ajustamento económico e financeiro relativo ao período 2011-2014 conduziu a um aprofundamento do esforço fiscal. Paralelamente, comparar-se-á o esforço fiscal observado em Portugal com o esforço fiscal registado nos países da União Europeia.

Com este propósito, o trabalho está estruturado em três secções. Na secção 2, são revisitadas as abordagens e aplicações encontradas na literatura e salientados os contributos relativos aos fatores determinantes do esforço fiscal, as medidas de esforço fiscal e os resultados obtidos em diversos estudos que abrangem a economia portuguesa. Na secção 3, desenvolve-se um exercício de aplicação empírica visando situar a evolução do esforço fiscal em Portugal no quadro dos países da União Europeia. Na secção 4, sistematizam-se as conclusões.

2. LITERATURA SOBRE O ESFORÇO FISCAL: ABORDAGENS E APLICAÇÕES

A literatura sobre o esforço fiscal destaca os fatores que o determinam e identifica os conceitos usados para o medir. Nesta secção, são revisitados os principais contributos que a literatura regista a este propósito e são apresentadas as posições de Portugal em ordenações do esforço fiscal extraídas da literatura.

2.1 DETERMINANTES DO ESFORÇO FISCAL

A literatura afirma que o esforço fiscal de um país depende da sua capacidade fiscal e, por esta via, das suas características económicas, sociais e institucionais (Pessino e Fenochietto, 2010). Assim, a identificação dos fatores determinantes da capacidade fiscal é uma questão central na investigação sobre o esforço fiscal.

O nível de desenvolvimento económico de um país é amplamente acolhido como determinante do esforço fiscal e a variável que, habitualmente, o representa é o PIB *per capita* (Lotz e Morss, 1967; Gupta, 2007; Pessino e Fenochietto, 2010; Fenochietto e Pessino, 2013; Santos, 2013; Pereira, 2014)². À medida que o PIB *per capita* se eleva, o cidadão médio dispõe de uma proporção maior do seu rendimento para pagar

² Lotz e Morss (1967) recorrem ao PNB *per capita* para representar o nível de desenvolvimento económico, enquanto Gupta (2007), Pessino e Fenochietto (2010), Fenochietto e Pessino (2013), Santos (2013) e Pereira (2014) usam o PIB *per capita* expresso em paridades de poder de compra.

impostos e conforma-se mais com as normas fiscais. Por outro lado, de acordo com a lei de Wagner, a procura de serviços públicos é elástica em relação ao rendimento, sendo de pressupor a existência de uma relação positiva entre o PIB *per capita* e a capacidade fiscal.

O grau de abertura da economia ao exterior é, também, identificado pela literatura como determinante da capacidade fiscal (Lotz e Morss, 1967; Gupta, 2007; Pessino e Fenochietto, 2010; Fenochietto e Pessino, 2013; Santos, 2013; Pereira, 2014). A literatura salienta a presunção de que existe uma relação positiva entre o grau de abertura da economia ao exterior e a capacidade fiscal, o que é explicado, principalmente, pela circunstância de as transações com o exterior implicarem formalidades que facilitam a identificação da base tributária e dificultam a evasão fiscal. Contudo, Gupta (2007) concluiu que o efeito pode ser ambíguo na medida em que a liberalização do comércio externo pode ser conseguida através de redução da tributação.

O grau de desigualdade na distribuição do rendimento é outro fator relevado pela literatura como sendo determinante da capacidade fiscal, no sentido em que uma distribuição do rendimento mais equitativa facilita a arrecadação de receita fiscal, na medida em que tal situação favorece a conformidade dos cidadãos com as normas fiscais (Pessino e Fenochietto, 2010; Fenochietto e Pessino, 2013).

De acordo com a literatura, a corrupção tem um impacto negativo significativo sobre a capacidade fiscal (Tanzi e Davodi, 2000; Bird e Martinez-Vazquez, 2008; Pessino e Fenochietto, 2010; Fenochietto e Pessino, 2013). Este impacto resulta de a corrupção tender a informalizar a economia e, conseqüentemente, a reduzir a base de incidência fiscal. Em particular, nos países menos desenvolvidos, este efeito pode ser potenciado pelo desvio de impostos arrecadados em benefício dos agentes tributários.

A literatura sugere que o grau de intervenção do Estado na economia determina positivamente a capacidade fiscal. Como salientam Varsano *et al.* (1998), espera-se que a provisão pelo Estado de educação, saúde e proteção social liberte a sociedade de aplicar os seus recursos na obtenção destes bens no mercado e, por essa via, aumente a sua capacidade para pagar impostos.

A inflação é referida pela literatura como um fator que determina negativamente a capacidade fiscal (Varsano *et al.*, 1998; Pessino e Fenochietto, 2010; Fenochietto e Pessino, 2013). O mecanismo através do qual se concretiza o efeito da inflação sobre a capacidade fiscal é conhecido por efeito Tanzi e significa que “no caso dos impostos usualmente especificados em termos nominais e arrecadados desfasadamente, um aumento da inflação tipicamente reduz a receita fiscal real” (Romer, 2006: 542).

A composição setorial do produto é, de acordo com Gupta (2007), um fator determinante da capacidade fiscal visto que há setores onde é mais difícil controlar a evasão fiscal (por exemplo, a agricultura de subsistência). No mesmo sentido, inclinam-se Pessino e Fenochietto (2010), Fenochietto e Pessino (2013) e Pereira (2014).

O nível de educação é apontado em Pessino e Fenochietto (2010) e em Fenochietto e Pessino (2013) como um fator determinante da capacidade fiscal dado que populações com maior nível de escolaridade compreendem melhor a necessidade de pagar impostos, aceitam melhor a função redistributiva do Estado e revelam maior conformidade com as normas fiscais.

2.2 MEDIDAS DO ESFORÇO FISCAL

A literatura regista medidas de esforço fiscal constituídas por índices de três tipos, consoante estejam baseadas em: (i) indicadores da relativização da carga fiscal; (ii) na comparação entre a carga fiscal e uma carga fiscal padrão; e (iii) na comparação entre a carga fiscal e a capacidade fiscal.

No âmbito das medidas do primeiro tipo, situam-se os índices de Frank e de Bird descritos em Santos (2013) e Pereira (2014). O índice de Frank, apresentado em 1959 por Henry Frank, mede o esforço fiscal através da expressão:

$$F = \frac{T/Y}{Y_p} \times 100 , \quad (1)$$

onde T representa a receita fiscal, Y representa o Produto Nacional Bruto (PNB) e Y_p representa o PNB *per capita*.

O índice de Bird, apresentado em 1964 por Richard Bird, mede o esforço fiscal através da expressão:

$$B = \frac{T/(Y-T)}{Y_p} \times 100 , \quad (2)$$

onde os símbolos mantêm o significado previamente referido. O índice de Bird difere do índice de Frank por definir a carga fiscal em função do PNB deduzido da receita fiscal.

As medidas do segundo tipo – medidas que comparam a carga fiscal com uma carga fiscal padrão – derivam de modelos econométricos que definem a carga fiscal padrão por ajustamento, em regra linear, a variáveis representativas dos fatores considerados determinantes da capacidade fiscal. São exemplos desta abordagem os modelos apresentados em Lotz e Morss (1967), em Piancastelli (2001), em Gupta (2007), em Bird e Martinez-Vazquez (2008) e em Le *et al.* (2012).

Lotz e Morss (1967) estimaram, com base em médias de três anos extraídas de dados do período 1961-1966 para 72 países em vias de desenvolvimento, uma relação linear entre a carga fiscal (receita fiscal relativizada pelo PNB) e o PNB *per capita*. A comparação entre a carga fiscal e a carga fiscal assim estimada – correspondente a uma carga fiscal padrão – originou uma ordenação do esforço fiscal desenvolvido pelos países incluídos na base de dados. Adicionalmente, Lotz e Morss (1967) aprofundaram a sua análise considerando a relevância que o comércio externo pode ter na definição da carga fiscal padrão. Assim, estimaram, também, uma relação linear que adicionou à especificação inicial uma variável representativa do grau de abertura da economia ao exterior.

Piancastelli (2001) usou uma base de dados constituída por informação sobre 75 países, desenvolvidos e em vias de desenvolvimento, no período 1985-1995, para estimar índices

de esforço fiscal. Esta investigação suportou-se em análises de regressão linear da carga fiscal, incluindo apenas as receitas fiscais do governo central, sobre o PIB *per capita*, o grau de abertura da economia ao exterior e o produto do setor agrícola relativizado pelo PIB. Piancastelli (2001) estimou modelos baseados em dados seccionais (constituídos pelas médias de 11 anos observadas em cada país) e em dados em painel. Os índices de esforço fiscal são representados pelo quociente entre a carga fiscal e a carga fiscal estimada (designada no texto original por carga fiscal prevista).

Usando uma base de dados de um painel de 105 países em vias de desenvolvimento observados ao longo de 25 anos entre 1980 e 2004, Gupta (2007) investigou o desempenho das receitas fiscais dos respetivos governos centrais. O autor estimou as relações lineares entre a receita do governo central (não incluindo transferências) relativizada pelo PIB e variáveis representativas dos fatores determinantes daquele desempenho como o PIB *per capita* e, relativizados pelo PIB, o produto da agricultura, o montante das importações, o montante da ajuda externa e o montante da dívida pública. Os índices de esforço fiscal são representados pelo quociente entre os valores efetivo e estimado. Dado que a investigação incidiu sobre países em vias de desenvolvimento, o artigo não facultou uma medida do esforço fiscal em Portugal.

Bird e Martínez-Vázquez (2008) investigaram o efeito da participação dos cidadãos e das instituições e da responsabilização dos governos através de um índice de *voice and accountability*³ e da perceção da corrupção sobre a fiscalidade de um país. Esta investigação usou um modelo de regressão linear da carga fiscal sobre o PIB *per capita*, a taxa de crescimento da população, o grau de abertura da economia ao exterior, a parcela das atividades não agrícolas no PIB, o índice de *voice and accountability* e o índice de perceção da corrupção. Os dados usados foram os valores médios das variáveis relevantes no período 1990-1999 em países da América Latina. A investigação incidiu também sobre a exploração do modelo com aplicação a 25 países de elevado rendimento (entre os quais Portugal) para dados de 1998 e de 2000. Dado que a investigação de Bird e Martínez-Vázquez (2008) visava principalmente a avaliação do impacto da *voice and accountability* e da corrupção sobre a carga fiscal, o artigo não disponibiliza índices de esforço fiscal.

Le *et al.* (2012) estimaram uma carga fiscal padrão – que os autores designaram de *taxable capacity* – para um conjunto de 107 países desenvolvidos e em vias de desenvolvimento com base em dados anuais para o período 1994-2009. As variáveis determinantes da carga fiscal padrão são o PIB *per capita* a preços constantes, a taxa de crescimento da população em idade ativa (15-64 anos) (ou, em alternativa, o índice de dependência da população), o grau de abertura da economia ao exterior, o peso do valor acrescentado pela agricultura no PIB e um índice da qualidade institucional e de governança. A estimação da carga fiscal padrão usou uma regressão sobre dados em painel. O esforço fiscal foi medido por um índice constituído pela relação entre a carga fiscal e a carga fiscal padrão.

As medidas do terceiro tipo – medidas que comparam a carga fiscal com a capacidade fiscal – derivam de modelos econométricos que definem a capacidade fiscal como a receita

³ Os índices de *voice and accountability* são índices compostos que medem a extensão em que os cidadãos de um país estão capacitados para participar na seleção dos seus governos bem como a extensão em que existe liberdade de expressão, liberdade de associação e comunicação social livre.

máxima que um país pode obter dadas as suas características económicas, sociais e institucionais. A questão crucial da construção destes modelos centra-se na avaliação da capacidade fiscal e um modo de a abordar é o de definir a capacidade fiscal através de um procedimento análogo ao usado para estimar fronteiras de produção estocásticas⁴. Os *inputs* usados são variáveis representativas dos fatores determinantes da capacidade fiscal.

Os modelos adotados em Varsano *et al.* (1998), em Pessino e Fenochietto (2010) e em Fenochietto e Pessino (2013) usam esta metodologia para estimar índices de esforço fiscal. Varsano *et al.* (1998) estimaram a capacidade fiscal em função da população total, do PIB *per capita*, da taxa de inflação e da proporção da população em idade de trabalhar na população total. A base de dados usada integrou 27 países desenvolvidos e em vias de desenvolvimento e reportou-se ao ano de 1991. Pessino e Fenochietto (2010) adotaram um procedimento idêntico para estimarem a capacidade fiscal a partir de informação relativa a um painel de 96 países, desenvolvidos e em vias de desenvolvimento, ao longo dos anos 1991 a 2006⁵. Os *inputs* usados foram o PIB *per capita*, o grau de abertura da economia ao exterior, a despesa pública em educação relativizada pelo PIB, a taxa de inflação, o índice de Gini, o peso da agricultura no PIB e um índice de perceção da corrupção. Fenochietto e Pessino (2013) usaram o modelo econométrico adotado em Pessino e Fenochietto (2010) para estenderem a análise a 113 países e ao período 1991-2012.

A Tabela 1 mostra a posição do esforço fiscal em Portugal registada nos artigos apresentados.

3. APLICAÇÃO EMPÍRICA

Nesta secção, estima-se o esforço fiscal observado nos países da União Europeia ao longo do período 1995-2015, através das três metodologias alternativas identificadas na literatura relevante: a carga fiscal relativizada pelo PIB *per capita*, a carga fiscal comparada com uma carga fiscal padrão e a carga fiscal comparada com a capacidade fiscal.

Em concreto, a aplicação empírica desenvolve-se através do cálculo de índices que expressam a carga fiscal relativizada pelo PIB *per capita* e da estimação de modelos econométricos definidores da carga fiscal padrão e da capacidade fiscal com vista à comparação da carga fiscal com estes dois agregados.

⁴ A estimação de fronteiras de produção estocásticas é tratada, por exemplo, em Aigner *et al.* (1977) e em Jondrow *et al.* (1982).

⁵ Por indisponibilidade de informação, em 43 dos 96 países a carga fiscal considera apenas as receitas fiscais do governo central.

Tabela 1: Esforço fiscal em Portugal

Modelo	Período da base de dados	Ordenação abrangendo				
		Todos os países		Países da União Europeia		
		Número de países	Posição de Portugal	Número de países	Posição de Portugal	
Santos (2013)	2010	15	Índice de Frank	3º	3º	
Índice de Bird			7º	7º		
Modelo de Lotz-Morss			11º	11º		
Pereira (2014)	2011	17	Modelo de Lotz-Morss	11º	17	11º
Lotz e Morss (1967)	1961-1966 (a)	72	Versão inicial	28º	14	5º
Versão aprofundada			24º		7º	
Piancastelli (2001) / Dados em painel	1985-1995	75		30º	17	10º
Le <i>et al.</i> (2012)	1994-2009	107		40º	28	14º
Pessino e Fenochietto (2010)	1991-2006	96	Modelo truncado	37º	25	17º
Modelo truncado normal com heterogeneidade			30º		18º	
Fenochietto e Pessino (2013) (b)	1991-2012	96	Modelo normal	45º		19º
Modelo truncado normal			46º	27	19º	
Modelo truncado normal com heterogeneidade			46º		19º	
Fenochietto e Pessino (2013) (c)	1991-2012	113	Modelo normal	50º		19º
Modelo truncado normal			50º	27	19º	
Modelo truncado normal com heterogeneidade			52º		19º	
Modelo Mundlack (efeitos aleatórios)			72º		27º	

Notas: (a) Médias de três anos; (b) Modelo com 96 países; (c) Modelo com 113 países. A posição de Portugal corresponde à ordenação decrescente do índice de esforço fiscal.

3.1 BASE DE DADOS

A base de dados é constituída pelas variáveis descritas na Tabela 2, cujas fontes estão indicadas no Apêndice 1.

Tabela 2: Estatísticas descritivas

Variável	Unidade	Número de observações	Mínimo	Média	Máximo	Desvio-padrão
Carga fiscal (<i>cf</i>)	% do PIB	582	23,9 (Irlanda, 2015)	35,6	50,6 (Dinamarca, 2014)	6,0
PIB a preços constantes <i>per capita</i> (<i>pibpc</i>)	10 ³ euros	574	2,8 (Bulgária, 1999)	22,6	83,7 (Luxemburgo, 2007)	15,1
PIB a preços correntes (<i>pibn</i>)	10 ⁹ euros	588	2,8 (Malta, 1995)	402,1	3 032,8 (Alemanha, 2015)	641,7
Exportações de bens e serviços (<i>expt</i>)	10 ⁹ euros	588	1,4 (Letónia, 1995)	147,3	1 418,8 (Alemanha, 2015)	213,1
Importações de bens e serviços (<i>impt</i>)	10 ⁹ euros	588	1,6 (Letónia, 1995)	142,0	1 189,3 (Alemanha, 2015)	198,9
Grau de abertura da economia ao exterior (<i>gabe</i>)	% do PIB	588	37,1 (Estónia, 1995)	108,6	438,2 (Alemanha, 2015)	60,5
Indicador de desigualdade na distribuição do rendimento (<i>dren</i>)	-	444	20,0 (Dinamarca, 1995, 1997)	29,5	38,9 (Letónia, 2006)	4,1
Índice de perceção de corrupção (<i>corp</i>) (a)	-	539	26,0 (Roménia, 2002)	63,5	100,0 (Dinamarca, 1998, 1999) (Finlândia, 2000)	19,1
Despesa pública corrente primária (<i>dcpf</i>)	% do PIB	582	22,3 (Bulgária, 1997)	37,1	52,7 (Suécia, 1995)	6,1
Hiato do PIB (<i>pibgap</i>)	%	565	-14,7 (Grécia, 2013)	-0,2	14,5 (Estónia, 2007)	3,2

Nota: (a) O valor do mínimo (máximo) do índice corresponde à perceção de um grau de maior (menor) corrupção.

A base de dados abrange os 28 países da União Europeia (UE28) e informação relativa ao período de 1995 a 2015. Existem 436 observações com informação completa; no Apêndice 2, registam-se as contagens de valores omissos por país e por variável.

3.2 ANÁLISE EMPÍRICA

A análise empírica foi desenvolvida em torno do conceito de esforço fiscal expresso por três versões de índices: (i) carga fiscal relativizada pelo PIB *per capita*, (ii) carga fiscal comparada com uma carga fiscal virtual representada por um padrão vigente num conjunto de países e/ou num conjunto de períodos (carga fiscal padrão) e (iii) carga fiscal comparada com um máximo de receita fiscal cobrável por um país em função das suas características económicas, sociais e institucionais (capacidade fiscal).

CARGA FISCAL RELATIVIZADA PELO PIB *PER CAPITA*

A análise empírica iniciou-se com a determinação do índice de esforço fiscal medido pela relação entre a carga fiscal (*cf*) e o PIB *per capita* (*pibpc*). A Tabela 3 regista os valores do índice para os anos de 1995, 1998, 2007, 2010 e 2015⁶.

De acordo com este conceito, o esforço fiscal em Portugal aumentou de 2010 para 2015, situando o país na 12^a posição com esforço fiscal mais elevado entre os 28 países da União Europeia. Ainda assim, o índice observado em Portugal em 2015 fica aquém do registado em 1995. A Tabela 3 indicia, também, que, no período anterior a 2010, o esforço fiscal desenvolvido em Portugal seguiu uma tendência de redução.

CARGA FISCAL COMPARADA COM CARGA FISCAL PADRÃO

A observação gráfica das relações entre a carga fiscal e as variáveis *pibpc*, *gabe*, *dren*, *corp*, *dcp* e *pibgap* nos países da UE28 (Figura 2) indica que, tendencialmente, a carga fiscal (i) aumenta com o aumento do PIB *per capita*, (ii) diminui com o aumento do grau de abertura da economia ao exterior, (iii) é mais elevada quando a distribuição do rendimento é menos desigual, (iv) é mais elevada quando a corrupção percebida é menor, (v) aumenta com o aumento do rácio da despesa pública corrente primária (relativamente ao PIB) e (vi) não se altera em função dos desvios do produto efetivo face ao produto potencial.

A análise empírica desenvolvida com vista à definição de uma carga fiscal padrão conduziu à seleção de uma relação linear entre a carga fiscal (*cf*) e o PIB *per capita* (*pibpc*), o grau de abertura da economia ao exterior (*gabe*), o indicador de desigualdade na distribuição do rendimento (*dren*), o índice de perceção de corrupção (*corp*), a despesa pública corrente primária (*dcp*) e o hiato do PIB (*pibgap*). A estimação desta relação é feita através do método 2SLS, com recurso aos valores desfasados em um ano das variáveis *dcp* e *pibgap* em virtude

⁶ Os anos de 1998, 2007 e 2010 correspondem aos períodos que antecederam a constituição da União Económica e Monetária, o início da crise financeira internacional e o período de ajustamento da economia portuguesa.

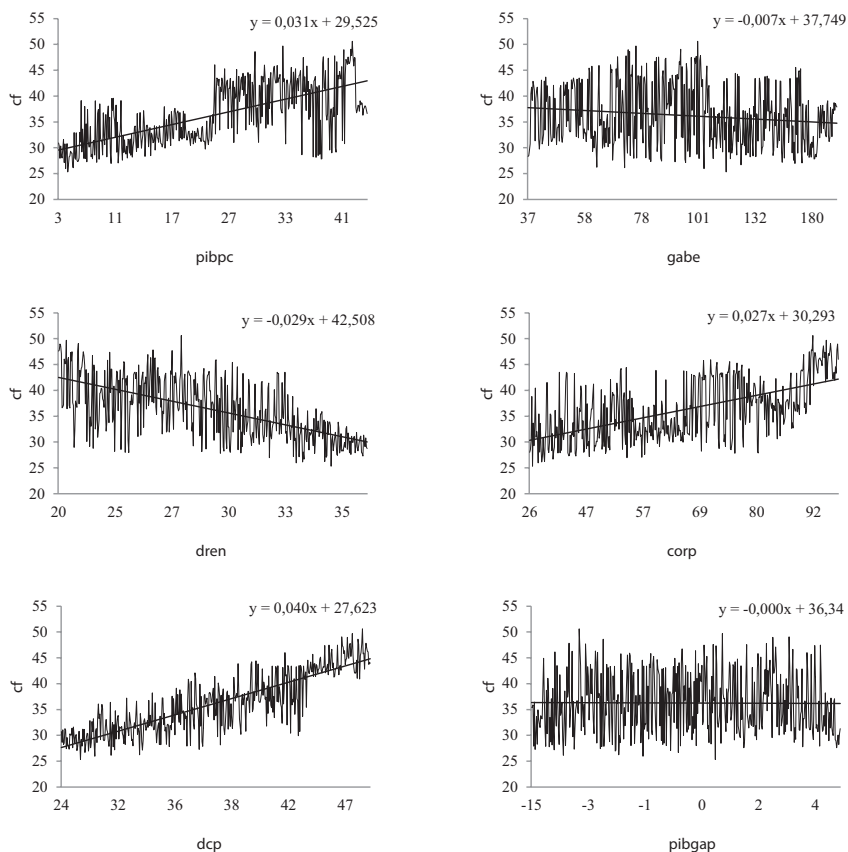
Tabela 3: Índice de esforço fiscal na UE 28 (Carga fiscal/PIB per capita a preços constantes de 2010)

País	1995	1998	2007	2010	2015
Alemanha	1,48	1,43	1,17	1,15	1,14
Áustria	1,52	1,48	1,15	1,17	1,22
Bélgica	1,61	1,56	1,26	1,28	1,31
Bulgária	-	9,30	6,45	4,99	5,09
Chipre	1,37	1,33	1,48	1,37	1,57
Croácia	-	-	3,31	3,44	3,58
Dinamarca	1,28	1,21	1,03	1,06	1,05
Eslováquia	5,99	4,70	2,45	2,26	2,28
Eslovénia	3,36	2,88	1,99	2,10	2,05
Espanha	1,75	1,66	1,48	1,35	1,47
Estónia	6,54	4,85	2,35	3,02	2,55
Finlândia	1,85	1,62	1,12	1,17	1,29
França	1,62	1,57	1,35	1,37	1,46
Grécia	1,89	1,86	1,40	1,58	2,14
Hungria	6,02	5,16	3,84	3,82	3,56
Irlanda	-	1,13	0,75	0,76	0,48
Itália	1,56	1,57	1,44	1,55	1,69
Letónia	8,03	6,76	2,77	3,29	2,72
Lituânia	6,91	6,55	3,07	3,14	2,51
Luxemburgo	0,67	0,64	0,44	0,48	0,47
Malta	-	-	2,12	1,96	1,74
Países Baixos	1,30	1,14	0,93	0,95	0,98
Polónia	7,32	5,99	4,07	3,34	2,98
Portugal	2,14	1,97	1,85	1,79	2,08
Reino Unido	1,35	1,31	1,12	1,18	1,10
República Checa	3,42	3,07	2,26	2,18	2,14
Roménia	7,42	8,00	4,75	4,15	3,89
Suécia	1,62	1,60	1,13	1,12	1,06

de se ter concluído pela endogeneidade destes dois regressores, e utilizando dados agrupados, tendo resultado na equação (3)⁷.

$$\begin{aligned}
 cf_i = & 18,328 + 0,103pibpc_i - 0,012gabe_i - 0,277dren_i \\
 & \begin{matrix} (7,678) & (7,724) & (-4,979) & (-6,770) \\ [0,000] & [0,000] & [0,000] & [0,000] \end{matrix} \\
 & -0,014 corp_i + 0,679dcp_i + 0,204pibgap_i + u_i . \\
 & \begin{matrix} (-1,261) & (19,701) & (3,743) \\ [0,208] & [0,000] & [0,000] \end{matrix}
 \end{aligned}
 \tag{3}$$

Figura 2: Relações entre a carga fiscal e as variáveis explicativas selecionadas



⁷ Os valores entre parêntesis curvos (retos) correspondem às estatísticas *t* (*p-value*).

O termo independente e os coeficientes associados às variáveis explicativas são estatisticamente significativos a um nível de significância de 1%, com exceção do coeficiente associado à variável *corp* que exige 21% para o nível de significância estatística. O coeficiente de determinação associado a esta estimação é de 0,831.

Os sinais dos coeficientes sugerem que a carga fiscal padrão mais elevada está associada a: (i) níveis mais elevados do PIB *per capita*; (ii) menores níveis do grau de abertura da economia ao exterior; (iii) menor desigualdade na distribuição do rendimento; (iv) maiores níveis de corrupção percebida (não obstante o menor nível de confiança associado à estimativa deste coeficiente); (v) maior peso da despesa pública corrente primária relativamente ao PIB; e (vi) variações positivas do hiato do produto efetivo face ao produto potencial. Os sinais dos coeficientes estimados conformam-se com a observação gráfica representada na Figura 2, exceto quanto ao hiato do produto (e ao nível da corrupção percebida que se revelou estatisticamente não significativo).

A literatura regista vantagens na estimação de relações lineares baseada em dados estruturados em painel relativamente a dados agrupados. Em primeiro lugar, a modelação com dados em painel permite identificar heterogeneidades individuais, isto é, características próprias de uma certa entidade (no caso vertente, países) ou de um determinado período de tempo.

Uma segunda vantagem potencial do uso de dados em painel por confronto com dados seccionais ou temporais é que os dados em painel podem oferecer informação mais completa sobre cada entidade porque a mesma entidade é observada repetidamente. Como é sublinhado em Verbeek (2004: 342), os dados em painel são apropriados para modelizar a forma como uma certa entidade se comporta em diferentes períodos do tempo.

Uma terceira vantagem é destacada por Verbeek (2004: 343) ao anotar que, variando as variáveis explicativas em duas dimensões na estimação com dados em painel (e não apenas numa dimensão como acontece com dados seccionais e com séries temporais), os estimadores são frequentemente mais precisos. Verbeek acrescenta que “mesmo com amostras de idêntica dimensão, o uso de uma base de dados estruturados em painel produz estimadores mais eficientes do que uma série de dados seccionais em que diferentes entidades são extraídas em cada período”.

Por último, Verbeek (2004: 344) destaca outra vantagem do uso de dados em painel: permite reduzir problemas de identificação decorrentes da presença de regressores endógenos, de erros de medida ou de omissão de variáveis explicativas.

No caso particular aqui investigado, a estimação baseada em dados estruturados em painel tem a vantagem de permitir captar os efeitos decorrentes das alterações do enquadramento na União Europeia. Com efeito, ao longo dos 21 anos do período em análise, o quadro institucional e regulamentar europeu, relevante do ponto de vista da definição dos regimes orçamentais dos Estados-membros, sofreu alterações que interessa ter em consideração. Estas alterações são comuns a todos os Estados-membros mas, naturalmente, alteram aquele quadro ao longo dos anos. Paralelamente, a crise económica e financeira internacional desencadeada no final da década de 2000 condicionou o desempenho fiscal dos países. Para segregar os efeitos destas alterações, é apropriado usar a estimação com efeitos fixos anuais.

Neste contexto, a análise empírica prosseguiu com a exploração da relação linear entre a carga fiscal (*cf*) e o PIB *per capita* (*pibpc*), o grau de abertura da economia ao exterior (*gbe*), o indicador de desigualdade na distribuição do rendimento (*dren*), o índice de perceção de

corrupção (*corp*), a despesa pública corrente primária (*dcp*) e o hiato do PIB (*pibgap*) no quadro de uma modelação baseada em dados estruturados em painel e com efeitos fixos temporais.

A aplicação de um teste F (para avaliar a hipótese nula de os efeitos fixos serem redundantes face à especificação alternativa com dados agrupados), de um teste de Breusch-Pagan para efeitos aleatórios (para avaliar a hipótese nula de os efeitos aleatórios serem redundantes face à especificação alternativa com dados agrupados) e do teste de Hausman (para avaliar a hipótese nula de os efeitos aleatórios estarem correlacionados com as variáveis explicativas) sugere que a especificação com efeitos fixos é a modelação mais apropriada. A opção por efeitos fixos temporais decorreu do facto de ser esta a especificação que capta melhor o poder explicativo das variáveis explicativas seleccionadas.

A equação (4) apresenta os resultados obtidos, por 2SLS (com recurso aos valores desfasados em um ano das variáveis *dcp* e *pibgap*, em virtude de se ter concluído pela endogeneidade destes dois regressores), com efeitos fixos temporais.

$$\begin{aligned}
 cf_{i,t} = & 12,241 + 0,088pibpc_{i,t} - 0,005gabe_{i,t} - 0,193dren_{i,t} \\
 & \quad \begin{matrix} (5,296) & (6,955) & (-1,850) & (-4,928) \\ [0,000] & [0,000] & [0,065] & [0,000] \end{matrix} \\
 & - 0,022corp_{i,t} + 0,777dcp_{i,t} + 0,194pibgap_{i,t} + u_{i,t}. \\
 & \quad \begin{matrix} (-2,224) & (24,005) & (3,197) \\ [0,027] & [0,000] & [0,002] \end{matrix}
 \end{aligned} \tag{4}$$

O termo independente e os coeficientes associados às variáveis dependentes são estatisticamente significativos, exceto o coeficiente associado à variável *gabe* que, contudo, é estatisticamente significativo a um nível de significância de 7%. O coeficiente de determinação é de 0,870.

Os sinais dos coeficientes associados às variáveis explicativas, consistentes com os obtidos na estimação com dados agrupados, sugerem que, tendencialmente, a carga fiscal padrão (i) aumenta à medida que o PIB *per capita* aumenta, (ii) diminui quando o grau de abertura da economia ao exterior aumenta, (iii) é menor quando ocorre maior desigualdade na distribuição do rendimento, (iv) é maior quando o nível de corrupção percebida é maior, (v) aumenta com o aumento do peso da despesa pública corrente primária no PIB e (vi) aumenta à medida que o hiato do produto efetivo face ao produto potencial aumenta.

Sublinha-se, em particular, o resultado da estimação para o sinal do coeficiente associado ao índice de percepção da corrupção (*corp*) na medida em que não se conforma com a evidência apresentada pela literatura.

CARGA FISCAL COMPARADA COM CAPACIDADE FISCAL

O terceiro conceito de esforço fiscal objeto de análise empírica é expresso por um índice que compara a carga fiscal com uma capacidade fiscal aferida por um máximo de receita fiscal cobrável em função de condições económicas, sociais e institucionais. O modelo de fronteira estocástica de Aigner *et al.* (1977) é a plataforma econométrica padrão que sustenta

a estimação da capacidade fiscal no âmbito da literatura sobre modelos de fronteira fiscal estocástica.

A capacidade fiscal, para dados em painel, é estimada a partir da definição de uma fronteira fiscal estocástica no sentido adotado em Pessino e Fenochietto (2010) e em Fenochietto e Pessino (2013), cujo modelo pode ser descrito assim:

$$y_{i,t} = \alpha + \beta^T x_{i,t} + v_{i,t} - u_{i,t} , \quad (5)$$

onde:

$y_{i,t}$ representa o logaritmo da carga fiscal para o país i no período t ;

$x_{i,t}$ é o vetor que representa as variáveis determinantes da carga fiscal para o país i no período t ;

β^T é um vetor de parâmetros desconhecidos (a estimar);

$v_{i,t}$ é o ruído estatístico (termo de perturbação ou erro); é uma variável aleatória representativa das variáveis independentes que explicam a variável dependente mas que não foram explicitamente consideradas bem como os erros de medida e a incorreção das formas funcionais, podendo ser positiva ou negativa;

$u_{i,t}$ representa a ineficiência, isto é, a insuficiência na obtenção de receita fiscal relativamente ao nível máximo de receita fiscal que podia ser obtida dadas as condições económicas, sociais e institucionais; é uma variável aleatória não negativa associada aos fatores específicos de cada país que contribuem para que esse país não atinja a sua capacidade fiscal no período t .

Habitualmente, é assumido que (i) v_{it} tem uma distribuição simétrica, tal como a distribuição normal e (ii) v_{it} e u_{it} são estatisticamente independentes, sendo $u_{it} > 0$ e podendo v_{it} assumir qualquer valor.

A análise visa medir a eficiência, $ef_{i,t}$, através do seguinte conceito de esforço fiscal:

$$ef_{i,t} = \frac{\text{carga fiscal}}{\text{carga fiscal na fronteira}} . \quad (6)$$

Na aplicação desta metodologia ao esforço fiscal, a ineficiência não deve ser entendida em sentido estrito. A capacidade fiscal não usada pode resultar de preferências sociais favoráveis a uma provisão menos expressiva de bens públicos e, portanto, de uma escolha política sobre o nível de fiscalidade mas também pode resultar de ineficiências na coleta dos impostos.

A capacidade fiscal foi, assim, estimada pelo método da máxima verosimilhança⁸, com as variáveis explicativas PIB *per capita* (*pibpc*), grau de abertura da economia ao exterior (*gabe*), indicador de desigualdade na distribuição do rendimento (*dren*), índice de perceção

⁸ A estimação foi efetuada por recurso ao *software* FRONTIER *version* 4.1. O procedimento usado na estimação e o significado dos parâmetros daí resultantes são descritos em Coelli (1996).

de corrupção (*corp*) e hiato do PIB (*pibgap*), ensaiando três especificações em que se assume, relativamente à variável associada à ineficiência, que:

- no modelo A, $u_{i,t} = u_i e^{\eta(t-T)}$ (η é um parâmetro a ser estimado e T representa o último período do painel) e u_i tem uma distribuição seminormal⁹;
- no modelo B, $u_{i,t} = u_i e^{\eta(t-T)}$ e u_i tem uma distribuição normal $N(\mu, \sigma^2)$ truncada¹⁰ (μ e σ^2 são parâmetros a serem estimados); e,
- no modelo C, $u_{i,t} = u_i e^{\eta(t-T)}$ e u_i tem uma distribuição normal $N(m, \sigma^2)$ truncada, sendo $m_{it} = \delta_0 + \delta_1 \times dcp_{it}$ (δ_0 e δ_1 são parâmetros a serem estimados).

Nestes modelos, é assumido que a despesa pública corrente primária (relativamente ao PIB) – representada pela variável *dcp* – não constitui uma determinante direta da capacidade fiscal mas que no modelo C influencia a medida em que a carga fiscal se aproxima daquela capacidade.

A Tabela 4 apresenta as estimativas de máxima verosimilhança obtidas nas condições descritas que, em conjunto, permitem definir a capacidade fiscal.

Os resultados das estimações dos modelos A e B revelam que os coeficientes associados ao grau de abertura da economia ao exterior (*gabe*), ao índice de perceção da corrupção (*corp*) e ao hiato do PIB (*pibgap*) não são estatisticamente significativos a um nível de significância de 10%. Acresce que, nestes modelos, o sinal do coeficiente associado ao indicador do nível de desenvolvimento económico (*pibpc*) não se conforma com o efeito sugerido pela literatura, nem com a representação gráfica da relação entre *cf* e *pibpc* nem com os resultados estimados para a carga fiscal padrão.

Contudo, a estimação do modelo C mostra que os coeficientes associados às variáveis explicativas são significativos a um nível de significância de 1%, com exceção do coeficiente associado ao índice de perceção da corrupção que não é estatisticamente significativo. Os sinais associados aos coeficientes do PIB *per capita* (*pibpc*) e do indicador de desigualdade na distribuição do rendimento (*dren*) mostram conformidade com a literatura, com a representação gráfica e com os resultados das estimações da carga fiscal padrão.

⁹ A distribuição seminormal (*half-normal*) é uma distribuição $Y = |X|$ em que $X \sim N(0, \sigma^2)$.

¹⁰ A distribuição normal truncada é uma distribuição da variável aleatória normalmente distribuída cujo domínio é finitamente limitado à esquerda ou à direita. No caso vertente, a distribuição normal é truncada à esquerda por zero.

Tabela 4: Coeficientes das variáveis explicativas e parâmetros da fronteira fiscal estocástica (estimativas da capacidade fiscal pelo método da máxima verosimilhança)

	Modelo A	Modelo B	Modelo C
Coefficientes			
constante	56,184 (23,689) [0,000]	52,368 (35,141) [0,000]	53,356 (31,560) [0,000]
pibpc	-0,090 (-2,577) [0,001]	-0,044 (-1,395) [0,163]	0,107 (8,188) [0,000]
gabe	0,000 (0,035) [0,972]	-0,001 (-0,156) [0,876]	-0,013 (-5,393) [0,000]
dren	-0,152 (-3,820) [0,000]	-0,148 (-3,995) [0,000]	-0,275 (-7,302) [0,000]
corp	-0,009 (-0,618) [0,537]	0,009 (0,711) [0,477]	-0,014 (-1,362) [0,173]
pibgap	-0,001 (-0,052) [0,958]	-0,020 (-0,829) [0,407]	0,388 (7,923) [0,000]
Parâmetros			
δ_0	-	-	35,408 (22,661) [0,000]
δ_1	-	-	-0,693 (-20,382) [0,000]
σ^2	206,515 (3,374) [0,001]	28,207 (25,303) [0,000]	5,826 (13,923) [0,000]
$\gamma = \sigma_{\epsilon_1}^2 / (\sigma_{\epsilon_1}^2 + \sigma_{\epsilon_2}^2)$	0,992 (401,077) [0,000]	0,942 (165,927) [0,000]	0,572 (2,231) [0,026]
μ	-	10,308 (8,718) [0,000]	-
η	0,005 (4,007) [0,000]	0,004 (2,711) [0,007]	-
Função de verosimilhança	-802,399	-809,618	-996,107

3.3 RESULTADOS

A Tabela 5 tem por base os índices de esforço fiscal calculados com base nas metodologias previamente apresentadas e sintetiza os resultados obtidos no quadro dos países da União Europeia em 1996, 1998, 2007, 2010 e 2015.

Tabela 5: Esforço fiscal em Portugal, no quadro dos países da União Europeia

Índices de esforço fiscal (a)		1996 (b)	1998	2007	2010	2015
Esforço fiscal medido pelo índice “Carga fiscal/PIB <i>per capita</i> ”	Portugal	2,12	1,97	1,85	1,79	2,08
	Média	3,33	3,09	2,09	2,04	1,99
	N.º países	25	26	28	28	28
Esforço fiscal medido pelo índice “Carga fiscal/Carga fiscal padrão” – dados agrupados (equação 3)	Portugal	0,99	0,98	0,93	0,81	0,98
	Média	1,02	1,05	1,02	0,94	1,02
	N.º países	11	13	27	28	27
Esforço fiscal medido pelo índice “Carga fiscal/Carga fiscal padrão” – dados em painel (equação 4)	Portugal	0,96	0,92	0,90	0,85	0,95
	Média	1,00	1,00	1,00	1,00	1,00
	N.º países	11	13	27	28	27
Esforço fiscal medido pelo índice “Carga fiscal/Capacidade fiscal” – modelo C (c)	Portugal	0,70	0,70	0,75	0,76	0,80
	Média	0,82	0,80	0,76	0,79	0,81
	N.º países	11	13	27	28	27

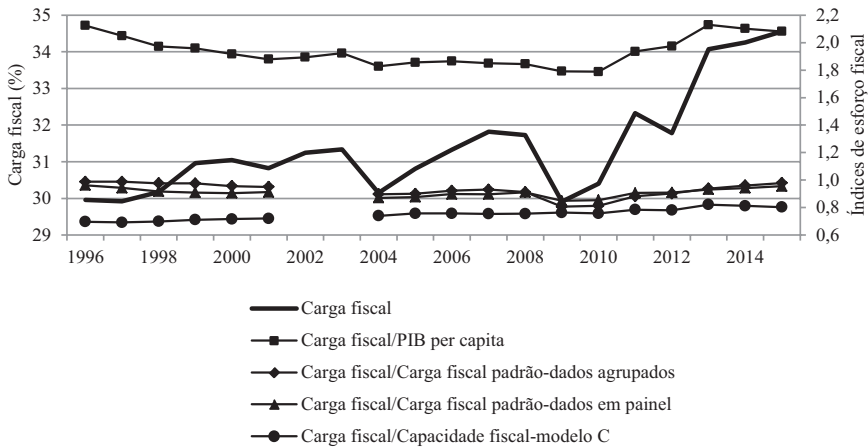
Notas: (a) Para cada modelo e ano, o quadro regista o índice do esforço fiscal em Portugal (primeira linha), a média do índice aplicado aos países analisados (segunda linha) e o número de países analisados (terceira linha). (b) Nos modelos baseados na carga fiscal padrão, o recurso a variáveis instrumentais implica a perda da estimativa para o ano de 1995. Nestas circunstâncias, optou-se por assumir como ano de partida nesta tabela o ano de 1996. (c) Para valor do índice, é tomada a medida de eficiência técnica apresentada em Coelli (1996: 7).

Estes resultados não são suficientemente robustos para permitirem ilações definitivas. Não obstante, no quadro da comparação com países da União Europeia, os resultados obtidos sugerem que o esforço fiscal em Portugal tem sido inferior à média dos países europeus considerados em cada modelo e ano, com exceção para o ano de 2015 na comparação com o índice “Carga fiscal/PIB *per capita*”.

No quadro da evolução temporal, os resultados obtidos através dos índices que comparam a carga fiscal com o PIB *per capita* e com uma carga fiscal padrão indicam que o esforço fiscal em Portugal diminuiu de 1996 para 1998 (período de preparação para a entrada no euro), de 1998 para 2007 (primeiros anos de integração monetária) e de 2007 para 2010 (período que se seguiu ao início da crise financeira internacional). Estes resultados sugerem que o esforço fiscal em Portugal aumentou de 2010 para 2015 (período de ajustamento da economia portuguesa), sem que, no entanto, tenham sido atingidos os valores observados em 1996. Diferentemente, os resultados obtidos através do índice que compara a carga fiscal com a capacidade fiscal sugerem que o esforço fiscal em Portugal tem vindo a aumentar desde 1996.

A Figura 3 sintetiza os resultados obtidos para Portugal no período 1996-2015.

Figura 3: Evolução da carga fiscal e dos índices de esforço fiscal em Portugal (1996-2015)



4. CONCLUSÕES

A relevância que o nível de fiscalidade assume nos planos económico e social motivou a realização deste estudo. O nível de fiscalidade foi avaliado por índices de esforço fiscal que relacionam a carga fiscal com o PIB *per capita*, com uma carga fiscal padrão ou com a capacidade fiscal. A carga fiscal padrão e a capacidade fiscal são estimadas por modelos econométricos inspirados na literatura.

Os resultados obtidos comparam a situação portuguesa com a situação em países da União Europeia bem como a sua evolução entre momentos considerados relevantes no período entre 1996 e 2015. Estes resultados sugerem que, no período em análise, Portugal suportou um esforço fiscal inferior à média dos países europeus considerados.

Os resultados obtidos através dos índices que comparam a carga fiscal com o PIB *per capita* e com uma carga fiscal padrão sugerem que o esforço fiscal em Portugal seguiu uma tendência de redução entre 1996 e 2010 e de aumento no período de implementação do programa de ajustamento económico e financeiro de 2011-2014. Em qualquer caso, resulta que o recente ajustamento da economia portuguesa foi acompanhado por um aumento do esforço fiscal, sem que tenha sido atingido o valor observado em 1996. No entanto, os resultados obtidos através do índice que compara a carga fiscal com a capacidade fiscal apontam no sentido de uma tendência de aumento do esforço fiscal em Portugal desde 1996.

REFERÊNCIAS BIBLIOGRÁFICAS

- Aigner, D.; Lovell, C. A. K.; Schmidt, P. (1977) Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics*, 6, 21-37.
- Bird, R. M.; Martinez-Vazquez, J. (2008) Tax effort in developing and high income countries: The impact of corruption, voice and accountability, Georgia State University Economics Faculty Publications, Paper 27.
- Coelli, T. (1996) A Guide to FRONTIER version 4.1: A computer program for stochastic frontier production and cost function estimation, Centre for Efficiency and Productivity Analysis, Working Paper 96/07.
- Fenochietto, R.; Pessino, C. (2013) Understanding countries' tax effort, IMF Working Paper 13/244.
- Gupta, A. S. (2007) Determinants of tax revenues efforts in developing countries, IMF Working Paper 07/184.
- Jondrow, J. ; Lovell, C. A. K.; Materov, I. S.; Schmidt, P. (1982) On the estimation of technical inefficiency in the stochastic frontier production function model, *Journal of Econometrics*, 19, 233-238.
- Le, T. M.; Moreno-Dodson, B.; Bayraktar, N. (2012) Tax capacity and tax effort: Extended cross-country analysis from 1994 to 2009, World Bank Working Paper 6252.
- Lotz, J. R.; Morss, E. R. (1967) Measuring 'tax effort' in development countries, Staff Papers-International Monetary Fund, 14 (3), 478-491.
- Pereira, M. F. (2014) *Fiscalidade*, Coimbra, Edições Almedina.
- Pereira, P. T.; Afonso, A.; Arcanjo, M.; Santos, J. C. G. (2012) *Economia e Finanças Públicas*, Lisboa, Escolar Editora.
- Pessino, C. ; Fenochietto, R. (2010) Determining countries' tax effort, *Hacienda Pública Española/Revista de Economía Pública, Instituto de Estudios Fiscales*, 195 (4), 65-87.
- Piancastelli, M. (2001) Measuring the tax effort of developing countries: Cross country panel data analysis, 1985/95, *Texto para Discussão* n° 818, IPEA – Instituto de Pesquisa Económica Aplicada (Rio de Janeiro).
- Romer, D. (2006) *Advanced Macroeconomics*, New York, McGraw-Hill Irwin.
- Santos, J. A. (2013) *Teoria Fiscal*, Lisboa, Instituto Superior de Ciências Sociais e Políticas.
- Tanzi, V.; Davodi, H. R. (2000) Corruption, growth and public finances, IMF Working Paper 00/182.
- Varsano, R. ; Pessoa, E. P.; Silva, N. C.; Afonso, J. R. R.; Raimundo, J. C. M. (1998) Uma análise da carga tributária do Brasil, *Texto para Discussão* n° 583, IPEA-Instituto de Pesquisa Económica Aplicada (Rio de Janeiro).
- Verbeek, M. (2004) *A Guide to Modern Econometrics*, Chicester, John Wiley & Sons.

APÊNDICE I: DESCRIÇÃO DA BASE DE DADOS

A carga fiscal foi extraída da série da AMECO “Total tax burden excluding imputed social security contributions: total economy-ESA 2010 (UTAT)” e está expressa em percentagem do PIB (Código: *cf*).

O PIB *per capita* foi extraído da série do Eurostat “Gross domestic product at market prices (Chain linked volumes (2010), euro *per capita* (nama_10_pc)” e está expresso em mil euros por indivíduo (Código: *pihpc*).

O PIB a preços correntes foi extraído da série da AMECO “Gross domestic product at current market prices (UVGD)” e está expresso em mil milhões de ecu/euros (Código: *pihn*).

As exportações de bens e serviços foram extraídas da série da AMECO “Exports of goods and services at current market prices (National accounts) (UXGS)” e estão expressas em mil milhões de ecu/euros (Código: *expt*).

As importações de bens e serviços foram extraídas da série da AMECO “Imports of goods and services at current market prices (National accounts) (UMGS)” e estão expressas em mil milhões de ecu/euros (Código: *impt*).

O grau de abertura da economia ao exterior foi calculado pela expressão “[Exportações de bens e serviços+Importações de bens e serviços]/PIB a preços correntes]” e está expresso em percentagem (Código: *gabe*).

O indicador de desigualdade na distribuição do rendimento foi extraído da série do Eurostat “Gini coefficient of equivalised disposable income (Source: SILC) (ilc_di12)” e está expressa em percentagem variando entre 0 (igual distribuição) e 100 (completa concentração) (Código: *dren*).

O índice de perceção da corrupção é produzido pela *Transparency International*. Para os anos de 1995 a 2013, o índice foi extraído de “DICE Database (2014), Corruption Perceptions Index, 1995-2013”, Ifo Institute, Munich, disponível *on line* em <http://www.cesifo-group.de/DICE/fb/pXBucnP5>; para os anos 2014 e 2015, o índice foi extraído do sítio <http://www.transparency.org/research/cpi/overview>. Este índice varia entre 0 e 10 para o período 1995-2011 e entre 0 e 100 para os anos seguintes, onde 0 significa que o país é percebido como muito corrupto e 10 ou 100 significa que o país é percebido como pouco corrupto. Para tornar a série temporalmente consistente, os índices de 1995-2011 foram reescalados para 0-100 por multiplicação por 10 do valor original (Código: *corp*).

A despesa pública corrente primária foi extraída da série da AMECO “Total current expenditure excluding interest: general government:-ESA 2010 (UUCGI)” e está expressa em percentagem do PIB (Código: *dcp*).

O hiato do PIB foi extraído da série da AMECO “Gap between actual and potential gross domestic gross product at 2010 reference levels (AVGDGP)” e está expresso em percentagem do PIB potencial (Código: *pihgap*).

APÊNDICE 2: CONTAGEM DE VALORES OMISSOS NA BASE DE DADOS

Países	Variáveis					
	cf	pibpc	dren	corp	dcp	pibgap
Alemanha			3			
Áustria			1			
Bélgica			1			
Bulgária		1	8	3		
Chipre			9	8		3
Croácia	6	5	15	4	6	6
Dinamarca			4			
Eslováquia			10	3		2
Eslovénia			7	4		3
Espanha			1			
Estónia			6	3		
Finlândia			2			
França			1			
Grécia			1			
Hungria			7			3
Irlanda		3	2			
Itália			2			
Letónia			9	3		2
Lituânia			8	4		2
Luxemburgo			1	3		
Malta		5	9	9		
Países Baixos			2			
Polónia			8	1		
Portugal			2			
Reino Unido			2			
República Checa			9	2		2
Roménia			9	2		
Suécia			5			

Nota: As variáveis *pibn*, *expt*, *impt* e *gabe* dispõem de informação completa.

A UTILIZAÇÃO DA EXPERIÊNCIA DE ESCOLHA DISCRETA NA VALORAÇÃO DE ESTADOS DE SAÚDE

Patrícia Antunes
Pedro Lopes Ferreira
Lara N. Ferreira

RESUMO

A Experiência de Escolha Discreta ou Discrete Choice Experiment (DCE) é uma técnica de eliciação de preferências que recentemente começou a ser utilizada na economia da saúde. Este estudo tem por objetivo identificar e descrever as etapas da construção do desenho desta técnica e avaliar a sua aplicação na valoração de estados de saúde em artigos de economia da saúde.

Foram selecionados estudos com a aplicação da DCE na valoração de estados de saúde. Todos os nove estudos avaliados explicitam as cinco etapas principais: identificação dos atributos e respetivos níveis, desenho experimental, recolha e análise de dados.

Conclui-se que as etapas de construção da DCE são seguidas na generalidade.

Palavras-chave: *discrete choice experiment*; eliciação de preferências; modelo de escolha discreta.

ABSTRACT

Discrete Choice Experiment (DCE) is a technique for preferences elicitation that recently started to be used in health economics.

This study aimed to identify and describe the stages of design construction of this technique and to assess their implementation in health economics papers that address the elicitation of health state values.

Published research studies were selected implementing DCE in the valuation of health status. All nine evaluated studies explain the five main: identification of attributes, identification of levels of attributes, experimental design, data collection and analysis.

The analysis performed evidenced that the DCE building steps are in general followed.

Keywords: Discrete choice experiment; elicitation of preferences; discrete choice model.

JEL Classification: I10; I12; I18; D71; D81

Received for publication: September 6, 2016

Revision accepted for publication: February 6, 2017

1. INTRODUÇÃO

A Experiência de Escolha Discreta (*Discrete Choice Experiment* ou DCE) é, segundo Earl e Kemp (1999) e Lancsar e Louviere (2008), uma técnica utilizada desde a década de 60 que assenta na psicologia matemática e cuja aplicação se iniciou originalmente na área do *marketing*, como uma forma de avaliar o comportamento do consumidor e prever as vendas de novos produtos. Desde então, a DCE tem sido utilizada para elicitare as preferências dos indivíduos em outras áreas académicas como a economia dos transportes, a economia do ambiente e, mais recentemente, a economia da saúde. Dada a contínua evolução tecnológica a que se assiste no sector da saúde e ao eterno problema dos recursos escassos associados à procura crescente de serviços, Lancsar e Louviere (2008) defendem que a utilização deste método de elicitação é particularmente relevante pelo facto de as preferências dos indivíduos poderem constituir um papel fundamental para a determinação das prioridades na aplicação dos recursos em saúde.

Não obstante a existência de outras técnicas ou escalas utilizadas para elicitação de preferências em saúde, como a escala visual analógica (VAS), o *standard gamble* (SG) e o *time trade-off* (TTO), Rowen et al. (2015) defendem que os valores elicitados por cada uma destas técnicas geram pouco consenso. Segundo estes autores e Brazier et al. (2012), o TTO e o SG têm sido também criticados pela sua complexidade e morosidade na aplicação aos respondentes e por excluírem grupos particulares, como crianças, idosos ou indivíduos de outras culturas. Por outro lado, Brazier et al. (2007) afirmam que a VAS é uma técnica que nem sequer é estritamente baseada em preferências na medida em que não implica uma escolha entre opções.

As limitações destas técnicas têm conduzido a um interesse crescente por outras, tais como a estimação de magnitude, o compromisso na pessoa e a DCE. A DCE tem vindo a obter cada vez mais destaque em estudos para elicitação de preferências em saúde. Segundo Kjær (2005), a utilização desta técnica justifica-se ainda pela sua simplicidade de aplicação e pela fácil compreensão do método.

Este artigo tem por objetivos identificar e descrever as etapas da construção do desenho do método e avaliar a sua aplicação em artigos publicados na área da economia da saúde e que se dedicam à valoração de estados de saúde.

2. EXPERIÊNCIA DE ESCOLHA DISCRETA

A escolha discreta ocorre quando um indivíduo se encontra face a uma opção entre um conjunto de alternativas finitas, mutuamente exclusivas e que contemplam todas as possibilidades.

A DCE é habitualmente implementada com recurso a questionários que definem o bem ou serviço em análise, em função de atributos, considerados como os mais relevantes, aos quais o indivíduo associa uma utilidade. O objetivo desta técnica é observar a resposta do indivíduo a alterações nesses atributos através de níveis definidos. Cada combinação, construída com os vários atributos e níveis associados (cf. Figura 1), vai constituir um cenário cujo objetivo final é permitir que o indivíduo se sinta apto a fazer trocas entre cenários de

**A utilização da experiência
 de escolha discreta na
 valoração de estados de saúde**

acordo com o que mais valoriza. Assim, pelo que defendem Kjær (2005) e Lancsar e Louviere (2008), torna-se possível identificar o nível que influencia a escolha de cada atributo, isto é, estimar a taxa marginal de substituição dos atributos. O atributo custo (e.g., custos de transporte, salários, impostos) desempenha um papel extremamente importante na DCE na medida em que permite que o método proceda à elicitación das preferências relativamente à disposição para pagar dos indivíduos. Isto não significa que os respondentes sejam diretamente questionados sobre a sua disponibilidade para pagar, mas é-lhes pedido que efetuem trocas monetárias por melhorias nos atributos.

Figura 1: Exemplo de aplicação de um par de cenários para escolha discreta.

	Cenário 1	Cenário 2
Modelos de organização dos cuidados primários	Inscrição associada a um médico de família (MF)	Inscrição associada a uma equipa de família (médico e enfermeiro de família)
Possibilidade de escolha do prestador de cuidados	Com possibilidade de escolha do MF	Sem possibilidade de escolha do MF
Pagamento (por consulta)	50€	25€
Tempo de espera para uma consulta de vigilância	4 dias	2 dias
Escolha a opção	<input type="checkbox"/>	<input type="checkbox"/>
→ Atributo	→ Alternativa/Cenário	→ Nível

Fonte: Adaptado de Hjelmgren & Anell (2007).

Earl e Kemp (1999) argumentam que a DCE tem por base as teorias da escolha probabilística e da utilidade aleatória, sendo um método compatível com a teoria do consumo e com a economia neoclássica. A teoria da utilidade aleatória permite ao investigador proceder à elicitación de preferências dos indivíduos para vários bens ou serviços e estimar modelos de preferências. Já a teoria da escolha probabilística subentende alguma incerteza no processo de escolha dos indivíduos, implicando que não seja possível prever, com perfeição, as suas escolhas individuais. De forma a analisar os resultados da DCE, Lancsar e Louviere defendem que se proceda à convergência do perfil de cada indivíduo respondente num único valor de utilidade, atribuindo pesos a cada um dos atributos e constituindo uma combinação linear.

As escolhas na DCE analisam-se de acordo com a teoria da utilidade aleatória, isto é, para um indivíduo i condicionado à escolha j , a sua utilidade U pode ser desagregada numa componente explicável V_{ij} e numa componente não explicável, aleatória, ε_{ij} :

$$U_{ij} = V_{ij} + \varepsilon_{ij}, \quad j = 1, \dots, J. \quad (1)$$

Esta última componente não explicável pode ser devida a atributos não observáveis, variações não observáveis nas preferências ou resultado de erros de medição. Por outro lado, a componente explicável é uma função dos atributos do bem ou serviço (vetor X_{ij} dos atributos) e das características dos indivíduos (vetor Z_i das características dos indivíduos):

$$V_{ij} = X'_{ij}\beta + Z'_i\lambda. \quad (2)$$

Nesta equação, β e λ representam os coeficientes a ser estimados. Assume-se que um indivíduo escolhe uma determinada opção 1, se e só se, a sua utilidade é maior do que qualquer outra utilidade de uma opção num conjunto de J alternativas. Considerando que Y_i é uma variável aleatória que denota o resultado da escolha e assumindo uma distribuição de probabilidade conjunta para e_i a probabilidade P da utilidade a ser maximizada ao escolher a opção 1 é dada por:

$$\begin{aligned} P(Y_i = 1) &= P(U_{i1} > U_{ij}) = P(V_{i1} + \varepsilon_{i1} > V_{ij} + \varepsilon_{ij}) \\ &= P(V_{i1} - V_{ij} > \varepsilon_{ij} - \varepsilon_{i1}), \forall j \neq 1. \end{aligned} \quad (3)$$

Note-se que a DCE considera a escolha em termos relativos, comparando escolhas intra-indivíduo e não inter-indivíduo. Para este último nível de análise, deverão ser usados modelos de escolha discreta, nomeadamente os que controlam a heterogeneidade decorrente da perceção inerente a cada indivíduo, tal como o modelo desenvolvido por Rebelo e Pereira (2014).

Não é prudente, no entanto, aplicar a DCE sem se compreender na totalidade o método, pelo que se torna fundamental estudar a teoria, o desenho da experiência e os resultados que se podem obter de forma a construir um modelo válido nos estudos de investigação.

2.1 DESENHO

Cada desenho de uma DCE tem em consideração o contexto específico da investigação em curso, dividindo-se por fases distintas com o intuito de simplificar a utilização do método. Os autores defendem desenhos semelhantes ainda que possuam ligeiras discrepâncias entre si. Assim, Kjær (2005) afirma que existe algum consenso em se dividir o desenho de uma DCE em cinco grandes fases principais: (1) identificação dos atributos, (2) identificação dos níveis dos atributos, (3) desenho experimental, (4) recolha e (5) análise de dados. Para o sucesso da implementação do método é fundamental que os investigadores tenham plena noção do que está a ser estudado e que os respondentes tenham conhecimento do bem ou serviço em análise e sejam devidamente informados de forma a possibilitar escolhas assertivas.

Passaremos, de seguida, a apresentar estas cinco fases.

Na primeira fase, os atributos são características qualitativas ou quantitativas de um bem ou serviço, pelo que a sua escolha deve ter em consideração a abrangência e mensurabilidade. Segundo Kjær (2005) e Lancsar e Louviere (2008), para a identificação dos atributos mais adequados, há várias fontes de informação possíveis designadamente a literatura existente, grupos focais, entrevistas a pessoas chave (e.g., decisores) e peritos. O objetivo principal na definição e identificação dos atributos é que estes sejam relevantes para os decisores e que, simultaneamente, se revistam de significado para os respondentes. Embora não exista um limite máximo para o número de atributos, Kjær (2005) sugere não se exceder os oito atributos. No entanto, na identificação dos atributos há vários aspetos a ter em consideração, nomeadamente se estes são irrelevantes ou não, isto é, até que ponto a sua exclusão ou inclusão interfere com os resultados obtidos e se os atributos são, entre si, mutuamente dependentes ou possuem relação de causalidade. Estas últimas características são particularmente importantes porque, se existirem, podem influenciar os indivíduos induzindo alguma diminuição da utilidade ou mesmo a transferência de utilidade entre atributos.

Na segunda fase, a cada um dos atributos identificados na etapa anterior são associados níveis considerados razoáveis e relevantes. O intervalo de variação destes níveis deve ser tal que possibilite aos respondentes a realização de trocas entre as várias combinações de atributos. Além disto, o número de níveis de um atributo influencia a sua importância, isto é, quantos mais níveis forem determinados para um atributo maior será a importância dada pelo respondente (em circunstâncias de intervalos de variação iguais). No entanto, quanto maior for o número e níveis dos atributos, mais complexo será construir o desenho da DCE, dado o número de combinações possíveis envolvidas. A inclusão de um atributo relacionado com o custo pode ser, ainda segundo Kjær (2005) e Lancsar e Louviere (2008), útil em alguns estudos nomeadamente quando se pretende efetuar análises económicas de custo-benefício.

Na terceira fase procede-se à construção de várias combinações de atributos e conjuntos de escolhas. O objetivo principal é criar uma DCE de tal forma que o número de alternativas seja o menor possível, sem colocar em causa a capacidade de o respondente inferir utilidade sobre todas as alternativas apresentadas. Este desenho vai permitir a estimação de uma matriz resultante de todas as possíveis combinações dos níveis de atributos das distintas alternativas (conjunto de escolhas). O número de alternativas aumenta significativamente à medida que aumenta o número de atributos e o número de níveis associados. Assim, partindo da hipótese que todos os atributos têm o mesmo número de itens, o número máximo de alternativas é dado pelo número de alternativas, elevado ao número de atributos.

Nas circunstâncias em que todas as alternativas possíveis são apresentadas aos respondentes está-se perante um desenho fatorial completo. Este desenho, para além de garantir a independência dos efeitos relevantes dos atributos, permite estimá-los, bem como aos efeitos resultantes da interação entre atributos. No entanto, por questões de praticabilidade, o desenho fatorial completo apenas pode ser aplicado a pequenas experiências, com um número muito limitado de atributos e de níveis. Em investigações com maior número de atributos e de níveis será, no entanto, necessário reduzir a dimensão do desenho da DCE e, nessas circunstâncias, o desenho fatorial fracionário constitui naturalmente a opção mais adequada. Nesta vertente mais reduzida utiliza-se um subconjunto ou fração que garante as propriedades mais significativas do desenho fatorial completo. Apesar de pressupor alguma

perda de informação estatística, o desenho fatorial fracionário permite, tanto quanto possível, estimar eficientemente os efeitos dos atributos. Ainda que se considere o desenho fatorial completo pouco praticável em estudos já com alguma dimensão, é sempre possível lidar com este tipo de desenho através de versões mais pequenas resultantes do modelo original. Com isto possibilita-se que sejam criados vários pontos de desenho sem aumentar o número do conjunto de escolhas para os respondentes (e.g., um desenho fatorial completo com 256 combinações pode ser reduzido a 16 versões de 16 conjuntos de escolhas). À semelhança do que ocorre com o desenho fatorial completo, a versão fracionária também permite obter versões mais reduzidas. Importa assinalar que, apesar de no desenho fatorial completo se utilizarem todas as alternativas possíveis, ainda assim, a eficiência do desenho de uma DCE necessita, para ser garantida, do correto emparelhamento das alternativas na formação do conjunto de escolhas. Deste modo, de forma a assegurar a otimização do desenho, é necessário que as seguintes propriedades se verifiquem: equilíbrio dos níveis (cada nível de um atributo surge com a mesma frequência nas alternativas incluídas no desenho); ortogonalidade (a diferença nos níveis de cada atributo varia de forma independente ao longo dos conjuntos de escolhas); justaposição mínima (os níveis não se repetem a si mesmo num conjunto de escolhas); e equilíbrio de utilidade (as utilidades das alternativas são semelhantes em cada conjunto de escolhas). Os tipos de desenhos referidos podem ser obtidos através de classificações manuais ou com recurso a *software* específico.

A quarta fase preocupa-se com a recolha de dados. Neste âmbito, os estudos realizados sobre o desenho da DCE recomendam a inclusão de uma alternativa adicional no conjunto de escolhas, particularmente em circunstâncias em que é irrealista considerar que o bem ou serviço é consumido com toda a certeza. Esta alternativa é considerada uma não escolha e impede que o indivíduo seja forçado a escolher algo que não é importante para si (i.e., que possui utilidade negativa). Com a inclusão desta alternativa o respondente pode “optar por sair”, escolher “nenhuma”, “não participar” ou manter a situação existente (*status quo*). A limitação desta alternativa reside, em situações de desenhos com alguma complexidade, no facto do indivíduo poder escolher a resposta mais fácil, isto é, aplicar uma heurística para evitar fazer uma escolha considerada difícil. Outra perspetiva relativa à inclusão de uma alternativa adicional é a possibilidade de esta refletir uma situação de indiferença para o indivíduo. A introdução destes tipos de alternativas deverá ser efetuada com ponderação e adaptar-se a cada estudo na medida em que acarretam alguns problemas do foro estatístico, nomeadamente no que se refere ao tratamento dos dados econométricos. Lancsar e Louviere (2006) referem que em alguns estudos, quando os indivíduos revelaram preferências lexicográficas ou manifestaram comportamentos irracionais, os investigadores tomaram a decisão de retirar do conjunto de dados da DCE essas respostas, com o intuito de garantir a consistência e a racionalidade da experiência. No entanto, os mesmos autores consideram que a decisão dos investigadores de eliminar dados que não correspondem às expectativas de um comportamento aceitável, impondo as suas próprias preferências, é uma atitude paradoxal e no mínimo paternalista. Estes autores apontam, por isso, um conjunto de fatores a ter em consideração quando se pondera a eliminação desses dados. Em primeiro lugar, mencionam a importância de utilizar abordagens alternativas à teoria do consumo. O que é considerado irracional para a abordagem clássica de preferências poderá ser encarado como racional em abordagens alternativas à teoria do consumo, como é o caso da teoria do arrependimento.

Por exemplo, se um indivíduo face a duas alternativas – A e B – escolhe A e se entretanto o conjunto de alternativas de escolha for alargado e passar a incluir mais uma opção – C – a informação que C contém poderá induzir o indivíduo a alterar a sua escolha para B em vez de A. De acordo com a abordagem clássica, este comportamento é irracional mas para a teoria do arrependimento tal já não sucede. O segundo fator apontado pelos autores como relevante é o desenho da investigação. Em circunstâncias de estudos de investigação com desenho deficiente, as preferências podem ser encaradas como irracionais mas serem, na verdade, racionais. Essa percepção poderá dever-se a problemas de desenho e implementação da DCE, por exemplo, quando os indivíduos são influenciados por atributos que não estão previstos no modelo; quando o uso de legendas nas diversas alternativas de escolha incentiva os indivíduos a fazerem inferências sobre atributos que não estão referenciados; e também pelo potencial de utilização de heurísticas por parte dos indivíduos quando se encontram a realizar o processo de escolha. Tal poderá igualmente acontecer devido ao facto de um ou mais atributos poderem não ter qualquer importância para os indivíduos. A eliminação destes indivíduos, nas diferentes circunstâncias apresentadas, fará com que a DCE apresente resultados robustos nos parâmetros dos atributos mas dificilmente os seus resultados poderão ser generalizados para a população em geral porque representam apenas os indivíduos que não foram excluídos do estudo.

Antes de os indivíduos responderem ao questionário é necessário que, por parte dos investigadores, seja efetuada uma explicação exaustiva no que se refere ao contexto e objetivos do estudo, à descrição dos atributos e à contextualização de cada um dos cenários (conjuntos de escolhas). Para o sucesso da recolha de dados é fundamental garantir que o respondente compreende o que se espera dele.

Nesta fase, os investigadores podem também considerar pertinente proceder à validação do desenho recorrendo a testes de racionalidade tornando-se, para tal, necessário incluir conjuntos de escolhas adicionais para testar os diversos axiomas, tais como monotonicidade, transitividade, convexidade e continuidade, como defendem Lancsar e Louviere (2006). É igualmente possível incluir perguntas de acompanhamento que possibilitam a compreensão por parte do investigador das razões que levaram o indivíduo a responder de determinada forma. Por fim, é importante a inclusão de algumas perguntas relativas a dados sociodemográficos (e.g., idade, género, habilitações literárias, rendimento, profissão). A recolha de dados pode ser efetuada com recurso a entrevistas, questionários ou combinação destes métodos.

No que respeita à amostra, esta deve refletir a população para a qual os resultados vão ser generalizados e a quem os custos de oportunidade importam. A dimensão da amostra está muitas vezes sujeita ao orçamento da investigação e a outros eventuais constrangimentos, mas deverá sempre permitir a estimação de modelos com garantia de fiabilidade. A dimensão da amostra vai depender do número de perguntas que cada indivíduo responde, do tamanho da população, da taxa de resposta que se espera obter no estudo e da potência estatística que se pretende obter do modelo.

A quinta e última etapa inclui a escolha do modelo de probabilidade e a inserção de dados, ambos com recurso a *software* apropriado, e também a especificação do modelo estatístico.

Segundo Kjær (2005), há vários modelos disponíveis para a estimação da DCE cuja escolha depende, entre outros fatores, de novo, do desenho escolhido para o método. No caso da DCE ser dicotómica (resposta ‘sim/não’ para uma determinada alternativa)

ou apresentar um conjunto de escolhas constituído por duas alternativas são utilizados modelos binomiais de escolha discreta (*logit* ou *probit*). Quando o conjunto de escolhas é constituído por três ou mais alternativas, é apropriada a utilização de modelos *logit* ou *probit* multinomiais, *mixed logit*, *nested logit* e valor extremo com heterocedasticidade. Em geral, este autor defende ainda que a estimação dos modelos de escolha discreta se baseie no método da máxima verosimilhança, embora outros métodos possam também ser considerados. Outros modelos com desenvolvimento mais recente e que fazem uso dos pontos de corte ou de fronteira endógenos com vista à distinção entre os níveis dos atributos poderão abrir caminho a outras opções de investigação em escolhas discretas ordinais, permitindo a categorização ou determinação de uma ordem de preferências ou respostas do indivíduo (Hensher et al., 2015).

Para além do *software* estatístico mais clássico como o SPSS ou o STATA, outros programas mais específicos, como o Lindep ou o Ngene, são também utilizados. A etapa de análise de dados pressupõe ainda a interpretação e a discussão dos resultados obtidos.

2. MÉTODOS E RESULTADOS

De forma a atingir os objetivos propostos procedeu-se a uma pesquisa sistemática na base de dados da PubMed em novembro de 2015, utilizando como palavras-chave '*Discrete Choice Experiment*' e '*Preferences*', tendo resultado em 482 artigos. Destes, selecionaram-se todos os que utilizaram a DCE como técnica de valoração de estados de saúde, restando nove estudos para análise. A opção por artigos relacionados com a aferição de valor dos estados de saúde deveu-se à importância de que se reveste a determinação das preferências da população no contexto de avaliação económica, particularmente no papel que desempenham no apoio à tomada de decisão e na definição de prioridades na alocação de recursos em saúde. Em cada um dos estudos selecionados foram identificadas e avaliadas as etapas de aplicação da DCE, tendo por base as fases descritas anteriormente.

Apresentam-se, de seguida, os nove artigos seleccionados sobre a valoração de estados de saúde:

O Estudo 1, de Xie et al. (2014), desenvolvido no Canadá, teve por objetivo comparar a DCE binária (escolha entre duas alternativas) e a DCE ternária (método que inclui três alternativas) no que concerne à sua viabilidade e fiabilidade para a valoração dos estados de saúde obtidos através do instrumento genérico de medição em saúde EQ-5D-5L. Os atributos e níveis escolhidos foram obtidos através das cinco dimensões deste instrumento de medição (mobilidade, cuidados pessoais, atividades habituais, dor/mal-estar e ansiedade/depressão) e dos cinco níveis associados a cada dimensão. No desenho da DCE, e a fim de selecionar os estados de saúde, foi utilizado um modelo Bayesiano eficiente que permitiu aleatoriamente gerar e emparelhar 48 estados de saúde. Este processo foi repetido 1.000 vezes com o objetivo de gerar 1.000 conjuntos de 24 pares e 16 conjuntos de escolhas. Todos os pares e conjuntos de escolhas foram manualmente examinados com o intuito de garantir que não existiam alternativas dominantes. Foram igualmente divididos em dois grupos de forma a manter o equilíbrio em termos de gravidade entre estados de saúde dentro de cada par/conjunto de escolhas.

Os participantes deste estudo foram constituídos por uma amostra de conveniência de profissionais da Universidade McMaster e por indivíduos resultantes de uma seleção aleatória de números de telefone pertencentes à cidade de Hamilton da província de Ontário. Os critérios para inclusão destes participantes consistiam em ter idade igual ou superior a 18 anos, compreensão da língua inglesa e vontade em participar numa entrevista direta com recurso a um questionário aplicado por computador.

Na fase inicial da entrevista foram colocadas algumas perguntas demográficas. Após descreverem o próprio estado de saúde com recurso ao EQ-5D-5L, cada participante executou um grupo aleatório com 12 tarefas de DCE binária e oito tarefas de DCE ternária.

A fim de determinar a validade das respostas na DCE binária, foi apresentada uma tarefa adicional para escolha entre alternativas em que um estado de saúde era dominante em relação ao outro. Para avaliar a fiabilidade teste-reteste, os participantes executaram pela segunda vez três tarefas de DCE binária escolhidas aleatoriamente do grupo original de 12 pares, equivalendo a 25% das tarefas de escolha.

No que concerne à DCE ternária, depois de finalizadas as oito tarefas, os participantes completaram duas tarefas adicionais aleatoriamente selecionadas do grupo original de oito conjuntos de escolhas. Neste estudo a viabilidade de execução das tarefas foi avaliada através da análise do tempo despendido em cada uma e do nível de dificuldade percebido pelo participante. A fiabilidade teste-reteste foi avaliada com recurso ao coeficiente de correlação intraclassas. Os dados da DCE binária e ternária foram respetivamente analisados com recurso a um modelo econométrico de regressão *probit* e de regressão logística ordenada. Os resultados do estudo foram apresentados e discutidos.

O Estudo 2, de Robinson et al. (2015), foi desenvolvido em Inglaterra com o objetivo de conceber um desenho de DCE envolvendo risco que permitisse elicitar valores de utilidade para os estados de saúde, ancorados ao estado de saúde normal e à morte, e desenvolver uma estrutura para elicitar valores para estados considerados piores que morte, assim como para estados melhores que morte. Nele, os investigadores pretenderam também estimar modelos alternativos de escolha envolvendo risco na DCE. Os atributos considerados neste estudo foram os resultados ou os estados de saúde do EQ-5D e o risco (ou probabilidade) de ter um dos resultados. Não obstante a utilização do instrumento de medição em saúde EQ-5D, este estudo não pretendeu estimar o sistema de valores para este sistema descritivo. Assim, no que concerne aos níveis, foram imputados quatro níveis ao atributo resultados/estados de saúde do EQ-5D (21121, 22222, 22323 e morte imediata) e nove níveis ao atributo risco (1%, 5%, 10%, 20%, 30%, 40%, 50%, 70% e 90%). Os autores do estudo optaram por um desenho fatorial completo, isto é, por incluir todas as combinações devido à incerteza associada ao desenho ótimo de modelos multiplicativos e para evitar limitar o âmbito da análise por se utilizar um desenho fatorial fracionário.

Para a recolha de dados, recrutaram-se, em 2011-2012, 60 estudantes de duas universidades de Londres e de Exeter, organizados em pequenos grupos constituídos, em média, por oito a nove participantes. No âmbito da participação, foi pago a cada respondente um valor monetário. A cada um dos grupos constituídos foi efetuada uma pequena introdução para apresentar os objetivos do estudo e alguns estados de saúde do EQ-5D. Foi, de seguida, pedido aos participantes que ordenassem esses estados de saúde ao longo da escala. Após esta fase inicial apresentaram-se, de forma aleatória, 15 perguntas de DCE e três perguntas

de SG modificado. Foi ainda pedido aos participantes que respondessem a quatro perguntas finais desenhadas para elicitare atitudes face ao risco através de lotarias monetárias. Neste estudo, os dados referentes à componente DCE foram analisados com recurso a um modelo econométrico *probit* ordenado. Os resultados foram apresentados e discutidos.

O Estudo 3, de Norman et al. (2013a), recorrendo à DCE, teve por objetivo testar a fiabilidade da estimação de valores de utilidade para a versão mais atual do instrumento de medição em saúde EQ-5D-5L. O estudo incluiu 973 respondentes que integravam uma base de dados *online* já existente e que foram submetidos à aplicação desta técnica de elicitacão de preferências através de um questionário aplicado por computador.

A DCE foi delineada a partir de um desenho fatorial fracionário composto por 200 conjuntos de escolha. Esses conjuntos foram agrupados em 20 blocos, cada um com 10 conjuntos de escolha. Os autores sugerem que este número de escolhas é facilmente gerido pelos respondentes, sem perda de qualidade nos dados. Cada conjunto de escolhas compreendia três opções: dois perfis de saúde e a possibilidade de morte imediata. O perfil de saúde das opções foi desenhado tendo por base as cinco dimensões do EQ-5D-5L, os cinco níveis de cada dimensão e a duração de vida (2, 4, 8 e 16 anos). O questionário, para além de incluir as opções de escolha sobre os perfis de saúde, integrava ainda perguntas demográficas e uma avaliação sobre a dificuldade de execução da DCE. No início do questionário foi disponibilizada aos respondentes uma explicação exaustiva sobre a tarefa a desenvolver. Para além disso, em todas as etapas de resposta, esteve sempre presente um botão de ajuda. Não obstante a inclusão de três opções em cada conjunto de escolhas, a resposta referente à “morte imediata” não foi analisada para efeitos do estudo publicado, avaliando-se apenas a preferência relativa sobre as opções dos perfis de saúde. A análise econométrica foi efetuada tendo por base um modelo *probit* com efeitos aleatórios.

O Estudo 4, de Bansback et al. (2014), pretendeu analisar a fiabilidade da utilização de um desenho de DCE com a inclusão de um atributo de tempo (DCE_{TTO}) para estimar os valores de utilidade dos estados de saúde obtidos através do EQ-5D-5L. Os atributos escolhidos para a DCE resultaram das cinco dimensões do sistema descritivo do EQ-5D-5L e de um atributo adicional respeitante à duração de vida. Os níveis escolhidos para os atributos deste sistema descritivo decorreram dos cinco níveis presentes neste instrumento genérico de medição em saúde (sem alterações, alterações ligeiras, moderadas, graves ou extremas). No que concerne ao atributo da duração de vida foram escolhidos três níveis: 1, 5 ou 10 anos, seguidos de morte. A combinação dos estados de saúde do EQ-5D-5L com três níveis de duração de tempo permitiu obter 9.375 cenários e, consequentemente, 87.900 possíveis pares de escolha. Apesar do instrumento de medição EQ-5D-5L poder gerar estados de saúde pouco plausíveis (e.g., problemas extremos de mobilidade, sem problemas com as atividades habituais), os investigadores optaram por utilizar um desenho não restrito, aceitando a possibilidade de existência de combinações inverosímeis, mas não identificando nenhuma.

A recolha de dados foi elaborada e organizada por uma empresa de sondagens e os respondentes foram recrutados através de correio eletrónico, a partir de um painel comercial já existente na Internet. A seleção de respondentes foi efetuada seguindo as quotas da distribuição por idade e sexo da população do Reino Unido. Os dados foram obtidos com recurso a um questionário aplicado por computador. Depois de responder às perguntas demográficas e sobre o estado de saúde percebido, cada respondente completou 15

tarefas de DCE_{TTO} através de três módulos experimentais, cada um com cinco tarefas. Para a análise econométrica dos dados assumiu-se um modelo *logit* condicional. Os resultados do estudo foram apresentados e discutidos.

O Estudo 5, de Craig et al. (2013), foi o primeiro estudo com o objetivo de estimar os valores de utilidade da população norte-americana para os estados de saúde descritos pelo instrumento genérico de medição SF-6D, utilizando para tal as técnicas DCE e SG. Os atributos escolhidos para a construção da DCE seguiram as dimensões do instrumento de medição SF-6D (função física, limitação no desempenho, função social, dor, saúde mental e vitalidade) e os respetivos níveis (quatro a seis, consoante a dimensão). O sistema descritivo do instrumento genérico de medição em saúde SF-6D tem 18.000 possíveis estados de saúde, permitindo assim a construção de 161.991.000 possíveis pares. Neste estudo, os investigadores optaram naturalmente por um desenho parcial do tipo *pivot* com um cenário base adicionado por 12 pares com alterações efetuadas nos níveis dos atributos de forma a permitir a concentração e diminuir o cansaço por parte do respondente.

Para a recolha de dados, recrutaram-se participantes através de um painel de indivíduos na Internet. O questionário, administrado por computador, consistiu em perguntas iniciais de cariz sociodemográfico, seguido por tarefas relacionadas com a DCE. Para efetuar a análise dos dados recorreu-se a um modelo econométrico do tipo *probit*. Os resultados foram apresentados e discutidos.

O Estudo 6, de Bansback et al. (2012), teve como objetivo avaliar a aplicação da DCE associada à duração de tempo (DCE_{TTO}) na obtenção de resultados na escala de utilidade em saúde e compará-lo com os resultados obtidos com a técnica iterativa TTO. Na construção deste modelo utilizaram-se os estados de saúde de acordo com o modelo descritivo do EQ-5D-3L. Neste contexto, o desenho foi desenvolvido tendo por base seis atributos: cinco correspondentes às dimensões do EQ-5D e um atributo relativo ao tempo de vida associado ao estado de saúde descrito. Aos cinco atributos relacionados com o sistema descritivo do EQ-5D foram associados três níveis de intensidade (sem problemas, alguns problemas e problemas extremos) e, para o atributo tempo, escolheram-se quatro níveis (1, 4, 7 e 10 anos). No desenho do TTO considerou-se uma matriz quase ortogonal de 36 estados. A estes acrescentaram-se mais 12 estados permitindo a comparação da técnica TTO com estudos prévios de avaliação. Quanto ao desenho da DCE_{TTO} obtiveram-se 972 combinações de atributos e níveis, considerando as 243 combinações de estados de saúde possíveis gerados pelo EQ-5D-3L e os quatro níveis de tempo do atributo anos de vida, o que gerou o número de 471.906 pares de escolha potenciais. Os conjuntos de escolha resultantes da combinação dos atributos e níveis foram gerados através do *software* SAS. Com o intuito de testar a consistência lógica, foi incluído um cenário em que uma das opções foi considerada dominante. Considerando que a saúde é tipicamente definida como o produto do valor do estado de saúde e dos anos de vida, neste estudo, assumiu-se que a técnica DCE_{TTO} adotaria um desenho multiplicativo. Inicialmente, os investigadores construíram um desenho fatorial fracionário, mas, devido às combinações implausíveis entre atributos e níveis, optaram por gerar 1.000 desenhos adicionais baseados em matrizes quase ortogonais e considerando todas as possíveis combinações. A escolha recaiu num desenho com menor correlação entre os

efeitos especificados, menor número de perfis de escolha dominados que foram depois manualmente alterados para se tornarem não dominados, maior número de estados de saúde nos cenários também incluídos no TTO para efeitos de comparação em estudos posteriores e maior eficiência. O desenho final incluiu 117 perfis de saúde do EQ-5D ao longo de 144 pares de escolha, previamente testados com dados simulados para assegurar que um modelo poderia ser estimado.

Antes de se iniciar a recolha de dados, aplicou-se um teste piloto a fim de garantir a compreensão dos respondentes, a obtenção de respostas fiáveis e a melhoria do desenho do questionário. Para a recolha de dados foi aplicado um questionário *online* a uma amostra da população canadiana, maior de idade e de língua inglesa, contendo um conjunto de tarefas com as técnicas TTO e DCE_{TTO}. A amostra foi recrutada através de uma empresa de sondagens que procedeu também à recolha dos perfis sociodemográficos dos respondentes. O questionário iniciava-se pedindo a estes para descreverem o seu estado de saúde atual utilizando o EQ-5D. Na segunda fase do questionário os participantes foram aleatoriamente alocados a um conjunto de tarefas TTO ou DCE_{TTO}, precedidas por uma descrição explicativa e vídeo de demonstração. Em ambas as tarefas, os atributos surgiram segundo uma ordem aleatória exceto no que concerne à duração no tempo (anos de vida) que na DCE_{TTO} surgiu sempre no final do cenário. No final do exercício, os respondentes foram questionados quanto à dificuldade na compreensão e execução das tarefas. Os resultados do estudo foram apresentados e discutidos.

O Estudo 7, de Ratcliffe et al. (2011), difere dos restantes analisados porque incide sobre a valoração dos estados de saúde em adolescentes e não em adultos. O interesse reside no facto de cada vez mais surgir evidência que sugere que há heterogeneidade significativa nos valores atribuídos por adolescentes e adultos a estados de saúde idênticos. A técnica utilizada para a elicitación de preferências neste estudo é, uma vez mais, a DCE a par com o TTO e o SG. A introdução destas três técnicas de elicitación de preferências teve por objetivo a sua comparação no que concerne ao grau de compreensão, facilidade de utilização e tarefas finalizadas. O instrumento de medição em saúde utilizado neste estudo foi o Child Utility 9D (CHU9D), desenvolvido pela Universidade de Sheffield. Trata-se de um instrumento genérico de medição em saúde, baseado em preferências, cujo sistema descritivo é constituído por nove dimensões (preocupação, tristeza, dor, cansaço, irritação, trabalho escolar, sono, rotina diária e capacidade para aderir a atividades), cada uma com cinco níveis. O algoritmo utilizado é baseado na população adulta inglesa com recurso ao SG. Na medida em que os investigadores consideram que a valoração dos estados de saúde obtida através de adultos poderá não refletir o valor atribuído por crianças e adolescentes, analisaram outras técnicas a fim de elicitar valores diretamente a partir de jovens.

A amostra foi constituída por 16 alunos de uma escola de Adelaide, no Sul da Austrália, com idades compreendidas entre os 11 e os 13 anos, que participaram numa entrevista direta onde lhes foi solicitado que indicassem a melhor e a pior combinação de atributos e níveis de uma série de estados de saúde definidos pelo CHU9D num desenho DCE. Os participantes foram também aleatoriamente alocados a tarefas convencionais de TTO e SG para as quais foi ainda pedido que indicassem o nível de dificuldade na sua execução. No que concerne ao desenho da DCE, o estudo não indica explicitamente a metodologia

seguida, nem informa sobre o modelo econométrico utilizado. No entanto, os resultados foram apresentados e discutidos.

O Estudo 8, de Shah et al. (2015), tendo por base o debate sobre se os ganhos em saúde devem ter pesos distintos para diferentes grupos etários de doentes, pretendeu determinar se a preferência da população em geral de Inglaterra e do País de Gales é consistente com a decisão de dar prioridade a tratamentos para prolongamento da esperança de vida em doentes terminais relativamente a outro tipo de tratamentos, considerando que os primeiros não obedecem aos critérios de reembolso utilizados em outros procedimentos. Para este estudo foram escolhidos quatro atributos de acordo com os critérios do *National Institute for Health and Care Excellence* (NICE) na definição de políticas de saúde em situações terminais: esperança de vida com e sem tratamento e qualidade de vida com e sem tratamento. Foi potencialmente considerado incluir outros atributos como a idade do doente, mas a decisão final foi de não o fazer de forma a restringir a complexidade das tarefas de escolha, diminuindo a possibilidade de ocorrência de decisões lexicográficas ou heurísticas por parte dos respondentes. No que concerne aos níveis atribuídos a cada atributo, foram imputados cinco ao atributo “esperança de vida sem tratamento” (3, 12, 24, 36 e 60 meses), dois níveis ao atributo “qualidade de vida sem tratamento” (50 e 100%), seis ao atributo “esperança de vida com tratamento” (0, 1, 2, 3, 6 e 12 meses) e três ao atributo “qualidade de vida com tratamento” (0, 25 e 50%). Neste estudo foi efetuado um desenho fatorial completo com as combinações dos atributos e respetivos níveis, resultando em 180 possíveis cenários. No entanto, alguns destes cenários mostraram-se implausíveis (e.g., soma da qualidade de vida com e sem tratamento superior a 100%). No desenho foi, também, incluída uma restrição de que a soma da esperança de vida e da qualidade de vida com tratamento seria maior que zero. Com recurso ao *software* STATA construíram-se 80 conjuntos de pares de escolha, a partir dos 110 cenários obtidos com as restrições incluídas. Os investigadores optaram por organizar os 80 conjuntos em oito blocos de dez escolhas, tendo estes conjuntos sido classificados em um de 13 tipos, um deles era um cenário dominante sobre os outros. Os investigadores asseguraram-se que todos os oito blocos continham entre um a dois deste tipo de conjunto de escolhas dominantes. Os conjuntos de escolhas foram incluídos num questionário de autopreenchimento disponibilizado na Internet. O questionário iniciava com uma introdução sobre o impacto que diferentes doenças e tratamentos têm na esperança de vida e, seguidamente, foi pedido aos participantes que escolhessem um de dois doentes para tratamento, assumindo que o serviço de saúde tinha recursos apenas para tratar um deles e que não existiam terapêuticas alternativas. Não se disponibilizou uma opção de resposta de indiferença ou “*status quo*” e as 10 escolhas de DCE foram apresentadas de forma aleatória aos respondentes. Depois de completadas estas tarefas, foram apresentadas duas tarefas adicionais com o intuito de examinar o impacto da introdução de informação relacionada com o tempo respeitante ao conhecimento da doença por parte dos respondentes. Por fim, foi pedido aos participantes que respondessem a perguntas pessoais e sobre a sua saúde.

A amostra deste estudo foi constituída por indivíduos adultos da população em geral de Inglaterra e do País de Gales que integravam um painel de uma empresa de sondagens. A abordagem utilizada garantiu a representatividade da população em geral. Excluíram-

se os indivíduos que tinham recentemente respondido a questionários relacionados com questões de saúde. Os participantes foram recompensados através de um esquema de conquista de pontos usados para troca de *vouchers* de oferta ou donativos para caridade. Antes de iniciar o estudo, realizou-se um teste piloto, de forma a garantir a compreensão dos métodos apresentados e a resposta às perguntas. Os dados foram modelados utilizando uma estrutura de maximização de utilidade aleatória com recurso ao *software* STATA. O modelo econométrico utilizado foi o *logit* condicional. O artigo incluiu a apresentação dos resultados e a respetiva discussão.

Por último, o Estudo 9, de Norman et al. (2013b), pretendeu estimar os valores de utilidade para o instrumento de medição em saúde SF-6D, através da técnica DCE. Para o efeito, foram recrutados de uma base de dados *online* já existente 1.017 respondentes, representativos da população australiana. Tal como no estudo anterior, os atributos escolhidos para a construção da DCE seguiram as dimensões do instrumento de medição e os respetivos níveis, tendo para além disso sido incluído um atributo de cariz quantitativo relativo à sobrevivência ou duração de vida (1, 2, 4, 8, 12, 16 e 20 anos). A combinação dos atributos e níveis possibilitou a criação de um desenho fatorial completo com 126.000 estados de saúde, reduzidos a 180 conjuntos de escolha. Estes 180 conjuntos foram aleatoriamente divididos em 12 blocos, constituídos por 15 conjuntos de escolha, destinados aos respondentes. À semelhança do estudo 3, cada conjunto de escolha possuía três opções de resposta: duas baseadas nos perfis de saúde, resultantes da combinação de atributos e níveis das dimensões do SF-6D bem como da duração de vida, e uma respeitante à “morte imediata”. Uma vez mais, o estudo publicado não contemplou a análise desta última opção concentrando-se nos resultados das preferências relativas entre os perfis de saúde (“não morte”). Esta decisão pretendeu evitar o aparecimento de preferências lexicográficas, relacionadas com a existência da opção “morte imediata” e a não violação dos pressupostos da teoria da utilidade aleatória. Para além das opções sobre perfis de saúde, os respondentes foram questionados com perguntas demográficas e sobre a dificuldade da execução do processo de escolha. O respondente teve ainda a possibilidade de sugerir contributos ou emitir a sua opinião através de espaço destinado a texto livre. Todo este processo de resposta foi operacionalizado através de um questionário aplicado por computador, previamente submetido e aprovado por uma comissão de ética. Os respondentes receberam um incentivo monetário para participar no estudo e, para uma melhor compreensão da tarefa a executar, disponibilizou-se uma explicação detalhada no início do questionário. Adicionalmente, foi proporcionado um botão de ajuda que acompanhou o respondente durante todo o processo de resposta.

Para efetuar a análise de dados os autores recorreram a um modelo econométrico *probit*, com efeitos aleatórios em que os níveis das dimensões do SF-6D foram tratados como *dummies*. O estudo testou ainda um modelo *logit* com efeitos aleatórios. Para efeitos de análise econométrica utilizou-se o *software* STATA.

Após esta breve descrição dos nove estudos, e face às distintas etapas que o desenho da DCE deve seguir, importa apresentar os resultados da avaliação do seguimento dessas fases em cada um dos estudos analisados. O quadro 1 apresenta os resultados obtidos dessa avaliação.

**A utilização da experiência
 de escolha discreta na
 valoração de estados de saúde**

Quadro 1: Resultados da avaliação das fases do desenho da DCE nos estudos em análise

Estudo	Fase 1 Identificação dos Atributos	Fase 2 Identificação dos Níveis	Fase 3 Desenho Experimental	Fase 4 Recolha de Dados	Fase 5 Análise de Dados
Estudo 1 Xie et al. (2014)	5 atributos (EQ-5D)	5 níveis do EQ-5D	Desenho fatorial fracionário com 1.000 conjuntos de 24 pares resultando em 16 conjuntos de escolha	Profissionais de uma universidade e indivíduos selecionados aleatoriamente a partir de números de telefone	Regressão <i>probit</i> e regressão logística ordenada
Estudo 2 Robinson et al. (2015)	5 atributos (EQ-5D) + 1 atributo (risco)	4 a 9 níveis	Desenho fatorial completo	60 estudantes	Regressão <i>probit</i> ordenado
Estudo 3 Norman et al. (2013a)	5 atributos (EQ-5D) +1 atributo (sobrevivência)	5 níveis do EQ-5D + 2, 4, 8 e 16 anos	Desenho fatorial fracionário constituído por 200 conjuntos de escolha, agrupados em 20 blocos de 10 conjuntos de escolha, aleatoriamente atribuídos aos respondentes	Recrutados a partir de uma base de dados <i>online</i> já existente	Regressão <i>probit</i> , com efeitos aleatórios
Estudo 4 Bansback et al. (2014)	5 atributos (EQ-5D) + 1 atributo (sobrevivência)	5 níveis do EQ-5D + 1,5 e 10 anos	Desenho fatorial completo	Recrutados a partir de uma base de dados <i>online</i>	Regressão <i>logit</i> condicional
Estudo 5 Craig et al. (2013)	6 atributos (SF-6D)	4 a 6 níveis	Desenho parcial do tipo <i>pivot</i> (cenário base adicionado de 12 pares)	Recrutados a partir de um painel <i>online</i> .	Regressão <i>probit</i>
Estudo 6 Bansback et al. (2012)	5 atributos (EQ-5D) + 1 atributo (sobrevivência)	3 níveis EQ- 5D + 1, 4, 7 e 10 anos	Desenho fatorial fracionário incluindo 117 perfis de saúde do EQ-5D ao longo de 144 pares de escolha	Amostra representativa da população canadiana	Modelo não especificado
Estudo 7 Ratcliffe et al. (2011)	9 atributos (CHU9D)	5 níveis	Não explicitado	Alunos jovens	Modelo não especificado
Estudo 8 Shah et al. (2015)	4 atributos (NICE)	2 a 5 níveis	Desenho fatorial completo, com 110 cenários em 80 conjuntos	Amostra representativa da população em geral	Regressão <i>logit</i> condicional
Estudo 9 Norman et al. (2013b)	6 atributos (SF-6D) + 1 atributo (sobrevivência)	4 a 6 níveis + 1, 2, 4, 8, 12, 16 ou 20 anos	Desenho fatorial completo com 180 conjuntos de escolha aleatoriamente divididos em 12 blocos	Recrutados a partir de uma base de dados <i>online</i>	Regressão <i>probit</i> , com efeitos aleatórios

4. DISCUSSÃO

A análise dos resultados obtidos nestes nove estudos sobre a avaliação da aplicação das etapas da DCE permite verificar que, de uma forma geral, todos os casos analisados explicitam as cinco etapas definidas: identificação dos atributos, identificação dos níveis dos atributos, desenho experimental, recolha e análise de dados. No entanto, verifica-se a existência de insuficiências em algumas destas etapas de aplicação da DCE.

Assim, no que respeita aos atributos, e não obstante o seu número se encontrar intrinsecamente associado ao contexto da investigação, os estudos analisados, à exceção do estudo de Ratcliffe et al. (2011), não excederam o número de oito em teoria recomendado por Kjær (2005). O número de atributos escolhidos para cada um dos estudos variou de quatro em Shah et al. (2015) a nove em Ratcliffe et al. (2011). De uma forma geral, pode-se dizer que o número de atributos escolhido na maior parte dos estudos foi fortemente influenciado pelo sistema descritivo do instrumento de medição em saúde utilizado para elicitar os valores de utilidade associados aos estados de saúde por si descritos. Por esse motivo, apenas um estudo (Shah et al. 2015) referiu ter em conta o contexto local de definição de políticas de saúde na escolha dos atributos. A maior parte dos estudos (67%) escolheu atributos qualitativos e quantitativos.

Quanto ao número de níveis escolhidos para os atributos verificou-se, na maior parte dos estudos, uma grande heterogeneidade no número de níveis atribuído, tendo variado de dois (mínimo) a nove (máximo). Apenas os estudos de Xie et al. (2014) e de Ratcliffe et al. (2011) optaram por atribuir o mesmo número de níveis (cinco) a todos os atributos. Tal como foi referido anteriormente, os níveis deverão ter um intervalo de variação tal que permita aos respondentes a realização de trocas entre as várias combinações. Assim, torna-se possível que um maior número de níveis permita por parte do respondente a atribuição de maior importância a um atributo. No entanto, compreende-se que os estudos optem por reduzir o número de níveis dos atributos de forma a não tornar o desenho da DCE muito complexo. Uma vez mais, os níveis definidos para cada um dos atributos qualitativos da maior parte dos estudos obedeceram ao número de níveis já existente em cada um dos instrumentos de medição em saúde utilizados.

No que se refere à etapa do desenho experimental, encontrou-se uma grande diversidade na escolha. Quatro estudos optaram por um desenho fatorial completo (estudos 2, 4, 8 e 9), enquanto outros passaram para subdivisões ou para um desenho fatorial fracionário. Apenas o estudo de Ratcliffe et al. (2011) foi inespecífico quanto ao desenho experimental realizado.

Para a recolha de dados, seis dos nove estudos optaram pela utilização de um questionário disponível *online* e apenas três recorreram a entrevistas diretas. Os instrumentos utilizados em seis estudos continham também perguntas de carácter sociodemográfico. No que respeita a esclarecimentos prévios aos respondentes, quase metade dos estudos (45%) não refere ter prestado qualquer esclarecimento sobre os atributos e seu significado. Nenhum estudo incluiu outras opções de escolha como “optar por sair” ou “situação atual” que poderiam facilitar uma melhor correspondência à realidade do respondente. Relativamente a eventuais testes de racionalidade, realizados para validação do modelo, apenas três dos nove estudos (os estudos 1, 6 e 8) referiram explicitamente a inclusão de uma questão dominante. Será, no entanto, difícil não encontrar num estudo escolhas irracionais ou aparentemente irracionais

que podem, entre outros aspetos, ser resultado de uma falha do próprio investigador dada a possível dificuldade em compreender algumas das escolhas dos indivíduos. Foi precisamente para evitar a ocorrência de preferências lexicográficas que dois estudos não analisaram a opção “morte imediata”, nomeadamente, os de Norman et al. (2013a, 2013b). Considerando que a inclusão deste cenário foi efetuada com a finalidade de permitir a valoração de todos os perfis de saúde ao longo da escala de utilidade, mesmo que a decisão final dos investigadores passe por eliminar esse tipo de respostas do conjunto de dados da DCE, os dados poderão ser sempre analisados na perspetiva que inclui esses respondentes e na perspetiva que os exclui.

Por fim, a maior parte dos estudos referiu explicitamente os modelos econométricos escolhidos para análise dos dados. Todos apresentaram os resultados e efetuaram a respetiva discussão.

5. CONCLUSÃO

A análise efetuada permitiu concluir que, nos estudos selecionados, as diversas etapas de construção da DCE são seguidas na generalidade. No entanto, pelo facto da maior parte dos estudos analisados optar por acompanhar o sistema descritivo dos instrumentos de medição em saúde (e.g. EQ-5D, SF-6D) na seleção dos atributos, esta análise estatística torna-se eventualmente menos rica. Para o sucesso de aplicação deste método contribui sem dúvida um adequado esclarecimento do modelo aos participantes envolvidos e essa componente não foi conseguida em todos os estudos. É, ainda, de referir que nesta análise foram apenas contemplados nove estudos e todos eles relacionados com a valoração de estados de saúde. Por esse motivo, a investigação futura deverá alargar este tipo de avaliação a mais estudos e com outro objeto de análise. Na verdade, antevê-se que a utilização da DCE continue a crescer particularmente em estudos na área da saúde. A integração recente desta técnica no protocolo de valoração do EQ-5D-5L descrito em Oppe et al. (2014), será um fator determinante na sua crescente utilização.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bansback, N.; Brazier, J.; Tsuchiya, A.; Anis, A. (2012) Using a discrete choice experiment to estimate health state utility values, *Journal of Health Economics*, 31(1), 306-313.
- Bansback, N.; Hole, A.R.; Mulhern, B.; Tsuchiya, A. (2014) Testing a discrete choice experiment including duration to value health states for large descriptive systems: Addressing design and sampling issues, *Social Science & Medicine*, 114, 38-48.
- Brazier, J.; Ratcliffe, J.; Salomon, J.A. ; Tsuchiya, A. (2007) *Measuring and Valuing Health Benefits for Economic Evaluation*, Oxford, Oxford University Press.
- Brazier, J.; Rowen, D.; Yang, Y.; Tsuchiya, A. (2012) Comparison of health state utility values derived using time trade-off, rank and discrete choice data anchored on the full health-dead scale, *European Journal of Health Economics*, 13(5), 575-587.
- Craig, B.; Pickard, A.S.; Stolk, E.; Brazier, J. E. (2013) US valuation of the SF-6D, *Medical Decision Making*, 33(6), 793-803.

- Earl, P.E.; Kemp, S. (1999) *The Elgar Companion to Consumer Research and Economic Psychology*, Cheltenham, Northampton, MA, USA, Edward Elgar.
- Hensher, D.; Rose, J.M.; Green, W.H. (2015) *Applied Choice Analysis*, Cambridge, Cambridge University Press.
- Hjelmgren, J.; Anell, A. (2007) Population preferences and choice of primary care models: A discrete choice experiment in Sweden, *Health Policy*, 83(2-3), 314-322.
- Kjær, T. (2005) A review of the discrete choice experiment, with emphasis on its application in health care, Health Economics Papers 1, University of Southern Denmark.
- Lancsar, E.; Louviere, J. (2006) Deleting “irrational” responses from discrete choice experiments: A case of investigating or imposing preferences?, *Health Economics* 15(8), 797-811.
- Lancsar, E.; Louviere, J. (2008) Conducting discrete choice experiments to inform healthcare decision making: a user’s guide, *Pharmacoeconomics*, 26(8), 661-677.
- Norman, R.; Cronin, P.; Viney, R. (2013a) A pilot discrete choice experiment to explore preferences for EQ-5D-5L health states, *Applied Health Economics and Health Policy*, 11(3), 287-298.
- Norman, R.; Viney, R.; Brazier, J.; Burgess, L.; Cronin, P.; King, M.; Ratcliffe, J.; Street, D. (2013b) Valuing SF-6D health states using a discrete choice experiment, *Medical Decision Making*, 34(6), 773-786.
- Oppe, M.; Devlin, N.; van Hout, B.; Krabbe, P.; de Charro, F. (2014) A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol, *Value in Health*, 17(4), 445-453.
- Ratcliffe, J.; Couzner, L.; Flynn, T.; Sawyer, M.; Stevens, K.; Brazier, J.; Burgess, P.L. (2011) Valuing child health utility 9D health states with a young adolescent sample: A feasibility study to compare best-worst scaling discrete-choice experiment, standard gamble and time trade-off methods’, *Applied Health Economics and Health Policy*, 9(1), 15-27.
- Rebelo, L.P.; Pereira, N.S. (2014) Assessing health endowment, access and choice determinants: Impact on retired europeans’ (in)activity and quality of life, *Social Indicators Research*, 119, 1411-1446.
- Robinson, A.; Spencer, A.; Moffat, P. (2015) A framework for estimating health state utility values within a discrete choice experiment: Modeling risky choices, *Medical Decision Making*, 35(3), 341-350.
- Rowen, D.; Brazier, J.; van Hout, B. (2015) A comparison of methods for converting DCE values onto the full health-dead QALY scale, *Medical Decision Making*, 35(3), 328-340.
- Shah, K.K.; Tsuchiya, A.; Wailoo, A.J. (2015) Valuing health state at the end of life: A stated preference discrete choice experiment, *Social Science & Medicine*, 124, 48-56.
- Xie, F.; Pullenayegum, E.; Gaebel, K.; Oppe, M.; Krabbe, P.F. (2014) Eliciting preferences to the EQ-5D-5L health states: Discrete choice experiment or multiprofile case of best-worst scaling?, *European Journal of Health Economics*, 15(3), 281-288.

A SMALL OPEN ECONOMIC GROWTH MODEL WITH IMPORTED GOODS, TOURISM, AND TERMS OF TRADE

Wei-Bin Zhang

ABSTRACT

This paper constructs an economic growth model of a small open economy with tourism and imported goods in a perfectly competitive economy. The study focuses on the effects of changes in terms of trade, with a preference for imported goods, on the dynamic paths of trade balance and economic growth. The basic framework for modelling a national economy is based on the Solow-Uzawa neoclassical growth model with Zhang's alternative approach to household behaviour. We build a nonlinear dynamic model with interdependence between economic growth, economic structure, tourism, prices, wealth and income. We provide a computational process to follow the motion of the economic system. Simulation is used to carry out a comparative dynamic analysis of the terms of trade, the propensity to consume imported goods, the rate of interest, the price elasticity of tourism, and the total productivity of the service sector. The comparative dynamic analysis provides some insights into the complexity of the tourism economy.

Keywords: tourism, price elasticity of tourism, terms of trade, small open economy.

JEL Classification: Z23; O41; F43

Acknowledgements: The author is grateful for the constructive comments of the anonymous referee. The author is also grateful for the financial support from the Grants-in-Aid for Scientific Research (C), Project No. 25380246, Japan Society for the Promotion of Science.

Received for publication: August 26, 2016

Revision accepted for publication: January 5, 2017

1. INTRODUCTION

Tourism has become an important part of many economies (Andereck et al., 2005; Matarrita-Cascante, 2010; and Antonakakis, *et al.* 2015). Using data from the World Travel and Tourism Council, Chou (2013) describes the economic importance of tourism as follows: “The total impact of the industry is impressive. In 2011, it contributed to 9% of global GDP, a value of over US\$6 trillion, and accounted for 255 million jobs. Over the next ten years, this industry is expected to grow by an average of 4% annually. This will bring it to 10% of global GDP, or about US\$10 trillion. By 2022, it is anticipated that it will account for 328 million jobs, 1 in every 10 jobs on the planet.” There are extensive studies about the relationship between tourism spending and economic growth for both developed and developing economies (e.g., Sinclair and Stabler, 1997; Luzzi and Flückiger, 2003; Hazari and Sgro, 2004; and Hazari and Lin, 2011). Most literature on relations between tourism and economic growth is empirical, even though there are a few mathematical models (e.g., Corden and Neary, 1982; and Copeland, 1991, 2012). The lack of theoretical research in the field is partly explained by the fact that introducing tourism into economic growth theory is analytically not easy. Different from other goods, tourism converts non-traded goods into tradable ones. This special character of tourism makes it difficult to use traditional theoretical frameworks directly to analyse the dynamic interdependence between tourism and economic growth. This study builds a dynamic model of tourism and economic structural change in a small open economic growth framework by applying an alternative approach to economic growth theory.

National economies can be affected by tourism in different ways. Tourism may compete for resources such as labour, capital and housing, with other sectors of the economy. Tourism also brings in income for the development of other economic activities. A tourism boom may cause congestion in transport systems and other public facilities. Empirical studies show an opposite relationship between a tourism boom and economic development (e.g., Balaguer and Cantavella-Jorda, 2002; Dritsakis, 2004; Durbarry, 2004; Oh, 2005; and Kim *et al.* 2006). Hazari and Sgro (1995) studied the dynamic relationship between tourism, capital accumulation, and the terms of trade. An increase in the international demand for tourism leads to a positive effect on long-run economic growth. Chao *et al.* (2006) find that an expansion of tourism can lead to capital decumulation. As pointed out by Chao *et al.* (2009), theoretical research on tourism has been mainly static. In order to comprehensively study possible effects of tourism on national economic development and economic structure, it is necessary to build a dynamic general equilibrium framework (Dwyer *et al.* 2004, and

Blake *et al.* 2006). This study deals with interdependence between tourism and economic growth on the basis of Uzawa's two-sector growth model in the context of a small open economy. We examine the response of economic growth and trade balance in a small open two-sector economy with capital accumulation to changes in the price elasticity of tourism, the price of imported goods, and preference for imported goods.

An important issue in the literature of small open economies is the impact of disturbances such as global economic crisis, terms of trade and prices of inputs (e.g., Sachs, 1982; Svensson and Razin, 1983; and Matsuyama, 1987). A preference for foreign brands, for instance, could affect the economy's growth rate and structure. A widely discussed topic in the literature is the impact of a change in a country's terms of trade on the economy's growth rate (Mendoza, 1995; Kose, 2002; and Turnovsky and Chattopadhyay, 2003). This study examines effects of the preference for foreign goods on trade balance and long-run economic growth. We show that a stronger desire for foreign brands may hurt economic growth. To properly deal with interactions between growth, tourism, preference for foreign goods and trade balance, we build a genuine dynamic framework. As reviewed by Zeng and Zhu (2011), almost all the growth models in tourism economics are based on a small open economy (e.g., Obstfeld and Rogoff, 1996; Lane, 2001; Kollmann, 2001, 2002; Benigno and Benigno, 2003; and Galí and Monacelli, 2005). We follow this tradition. Our approach of introducing tourism into growth theory is influenced by Chao *et al.* (2006). A main difference between our approach and the model developed by Chao *et al.* (2006) is that this study is based on an alternative utility function proposed by Zhang (1993, 2005). This paper is an extension of Zhang's model (Zhang, 2012). The main difference between this paper and Zhang's model is that we now add the imported goods and terms of trade into Zhang's growth model with tourism. The rest of the paper is organised as follows: Section 2 defines the basic model; Section 3 provides a computational procedure to plot the motion of the economy and simulates the model; Section 4 examines effects of changes in some parameters on the economic system over time; Section 5 concludes the study. The appendix gives the proof of the main results in Section 3.

2. THE GROWTH MODEL WITH TOURISM

The model is a combination of the basic features of three well-known models, Solow's growth model, Uzawa's two-sector growth model (Uzawa, 1961), and the growth models with tourism.

Following Chao *et al.* (2009), we consider an economy that is small and open and produces two goods: an international industrial good and national services. National services are “tradable” in the sense that foreign tourists come to visit the country and consume services. This study emphasises changes in the propensity to save and to consume imported goods, and in the terms of trade. Following the model by Eicher *et al.* (2008), we include another good, called an imported good. This good is not produced by the economy, but it is consumed by the domestic consumers. The introduction of this good enables us to consider the impact of domestic households’ preference for goods which cannot be produced by the domestic economy. A stronger desire, for instance, for foreign luxury goods or special agricultural products could affect the domestic economic structure. There are two types of consumer, domestic households and foreign tourists. To simplify the analysis, we assume that domestic households consume the two goods and services, while foreign tourists consume only services. Tourism converts services into an exportable commodity. In our model, the economy freely imports goods, and freely exports goods. The price of the industrial good is unity. Capital depreciates at a constant exponential rate δ_k . We assume that the economy is too small to affect the world interest rate r^* and price of imported goods p_Z . We assume that r^* and p_Z are constant. The households hold wealth and land and receive income from wages, land rent, and interest payments on wealth. Land is only for residential and service use.

The production sector technologies are characterised by constant returns to scale. All markets are perfectly competitive and capital and labour are completely mobile between the two sectors. Capital is perfectly mobile in international markets and we neglect the possibility of emigration or/and immigration. We assume that the population N is constant and homogeneous. We use subscript indices, i and s , to denote respectively the industrial and service sectors. Let $K_j(t)$ and $N_j(t)$ stand for the capital stocks and labour force employed by sector j , $j = i, s$, at time t . We use $F_j(t)$ to represent the output level of sector j .

THE INDUSTRIAL SECTOR

The production function of the industrial sector is given by

$$F_i(t) = A_i K_i^{\alpha_i}(t) N_i^{\beta_i}(t), \quad \alpha_i, \beta_i > 0, \quad \alpha_i + \beta_i = 1, \quad (1)$$

where A_i , α_i , and β_i are parameters. Labour and capital earn their marginal products, and firms earn zero profits. The wage rate $w(t)$ is determined in the domestic labour market. The marginal conditions for the industrial sector are

$$r_\delta = \alpha_i A_i k_i^{-\beta_i}(t), \quad w(t) = \beta_i A_i k_i^{\alpha_i}(t), \quad (2)$$

where $k_i(t) \equiv K_i(t)/N_i(t)$ and $r_\delta \equiv r^* + \delta_k$. From (2) we solve

$$k_i = \left(\frac{\alpha_i A_i}{r_\delta} \right)^{1/\beta_i}, \quad w = \beta_i A_i k_i^{\alpha_i}. \quad (3)$$

We consider k_i and w as functions of r^* .

THE SERVICE SECTOR

Service production needs three inputs, capital $K_s(t)$, labour force $N_s(t)$, and land $L_s(t)$. The production function of the service sector is

$$F_s(t) = A_s K_s^{\alpha_s}(t) N_s^{\beta_s}(t) L_s^{\gamma_s}(t), \quad \alpha_s, \beta_s, \gamma_s > 0, \quad \alpha_s + \beta_s + \gamma_s = 1, \quad (4)$$

where A_s , α_s , β_s , and γ_s are parameters. We use $p(t)$ and $R(t)$ to stand respectively for the price of the service and the land rent. The marginal conditions for the service sector are

$$\begin{aligned} r_\delta &= \alpha_s A_s p(t) k_s^{\alpha_s-1}(t) l_s^{\gamma_s}(t), \quad w = \beta_s A_s p(t) k_s^{\alpha_s}(t) l_s^{\gamma_s}(t), \\ R(t) &= \gamma_s A_s p(t) k_s^{\alpha_s}(t) l_s^{\gamma_s-1}(t), \end{aligned} \quad (5)$$

where

$$k_s(t) \equiv \frac{K_s(t)}{N_s(t)}, \quad l_s(t) \equiv \frac{L_s(t)}{N_s(t)}.$$

From (5) we have

$$k_s = \frac{\alpha_s w}{\beta_s r_\delta}. \quad (6)$$

We can consider k_s as a function of r^* .

FULL EMPLOYMENT OF CAPITAL AND LABOUR

The total capital stock employed by the country $K(t)$ is employed by the two sectors. The full employment of labour and capital is represented by

$$K_i(t) + K_s(t) = K(t), \quad N_i(t) + N_s(t) = N.$$

We rewrite the above equations as

$$k_i N_i(t) + k_s N_s(t) = K(t), \quad N_i(t) + N_s(t) = N. \quad (7)$$

The capital intensity of the two sectors, k_i and k_s , are uniquely determined in (3) and (4) as functions of the fixed rate of interest. From (7), we have

$$N_i(t) = (K(t) - k_s N)k_0, \quad N_s(t) = (k_i N - K(t))k_0, \quad (8)$$

where $k_0 \equiv (k_i - k_s)^{-1}$. We require $k_0 \neq 0$. The labour distribution is a unique function of the total capital used by the country.

DEMAND FUNCTION OF FOREIGN TOURISTS

Following Schubert and Brida (2009), we use the following iso-elastic tourism demand function

$$D_T(t) = a(t)y_f^\phi(t)p^{-\varepsilon}(t), \quad (9)$$

where $y_f(t)$ stands for the disposable income of foreign countries, ϕ and ε are respectively the income and price elasticities of tourism demand. We consider $a(t)$ dependent on many conditions, such as environment (e.g., crime rates, pollutants and congestion) and infrastructure (airports and transport systems). We assume that tourists pay the same price for services as the domestic household does. We do not consider other features of the tourism industry which have important effects on pricing (e.g., Marin-Pantelescu and Tigu, 2010; Stabler, *et al.*, 2010).

BEHAVIOUR OF DOMESTIC HOUSEHOLDS

We now describe the behaviour of domestic households on the basis of the approach by Zhang (1993). The implications of this approach are similar to those in the Keynesian consumption function and models based on the permanent income hypothesis, which are empirically much more valid than

the approaches in the Solow model or the Ramsey model. The approach is discussed at length by Zhang (2005, 2008). Zhang (2005) has also examined in detail the relations between his approach and the Solow growth theory, the Ramsey growth theory, the permanent income hypothesis, and the Keynesian consumption function. It can be shown that the behaviour generated by the traditional approaches can also be observed in Zhang's approach if certain patterns of preference changes are specified. As it assumes a fixed proportion of disposable income (which is the current income in Zhang's approach) is saved, the Solow model does not take account of how changes in wealth affect saving behaviour. The traditional Keynesian consumption function lacks a rational microfoundation. The approach becomes less effective when one has to deal with multiple goods and saving in a consistent manner. The Ramsey approach is based on the assumption that the utility is additional for the same person over that person's life time. There are extensive studies on the limitations of this formation. A comprehensive survey of the literature is given by Frederick *et al.* (2002). Zhang proposes an alternative approach to household behaviour by taking account of the valid points and overcoming the shortcomings in the traditional approaches.

Let L denote the total land available for service production and residential use. Each household obtains income from land ownership, wealth and wages. To simplify the model, we assume that the land is equally owned by the population. The revenue from land is equally shared among the population. The total land revenue is $LR(t)$. The income from land per household $\bar{r}(t)$ is $\bar{r}(t) = LR(t)/N$. The household chooses the lot size, consumption levels of industrial goods, imported goods and services, and how much to save. The current income is

$$y(t) = r^* \bar{k}(t) + w + \bar{r}(t), \quad (10)$$

where $\bar{k}(t)$ is wealth held per household. We call $y(t)$ the current income in that it comes from consumers' wages and current earnings from ownership of wealth. The sums of income that consumers are using for consuming and saving are not necessarily equal to the current income because they can sell wealth to pay for current consumption, for instance, if the current income is not sufficient for the purpose. The total value of the wealth that a consumer can sell to purchase goods and to save is equal to $p_i(t)\bar{k}(t)$, with $p_i(t) = 1$ at any t , where $p_i(t)$ is the price of the industrial good. Here, we assume that selling and buying wealth can be done instantaneously without any transaction cost.

The disposable income at any point in time is

$$\hat{y}(t) = y(t) + \bar{k}(t). \quad (11)$$

Disposable income is used for saving and consumption. At time t the consumer has the total amount of income equalling \hat{y} to distribute between consuming and saving. In the growth literature, for instance in the Solow model, the saving comes out of the current income $y(t)$, while in this study the saving comes out of the disposable income, which is dependent both on the current income and wealth.

At each point in time, a consumer distributes the total available budget between lot size $l(t)$, consumption of services $c_s(t)$, industrial goods $c_i(t)$, imported goods $c_z(t)$, and saving $s(t)$. The budget constraint is

$$R(t)l(t) + p(t)c_s(t) + c_i(t) + p_z c_z(t) + s(t) = \hat{y}(t). \quad (12)$$

Equation (12) means that the consumption and saving exhaust the consumer's disposable income. We assume that utility level $U(t)$ of the household is dependent on $l(t)$, $c_s(t)$, $c_i(t)$ and $s(t)$ as follows:

$$U(t) = \theta l^{\eta_0}(t) c_s^{\gamma_0}(t) c_i^{\xi_0}(t) c_z^{\varsigma_0}(t) s^{\lambda_0}(t), \quad \eta_0, \gamma_0, \xi_0, \varsigma_0, \lambda_0 > 0,$$

in which η_0 , γ_0 , ξ_0 , ς_0 and λ_0 are the elasticities of utility with regard to lot size, services, industrial goods, imported good, and saving. We call η_0 , γ_0 , ξ_0 , ς_0 and λ_0 propensities to consume the lot size, to consume services, to consume industrial goods, to consume imported goods, and to hold wealth, respectively. The time horizon of the consumer's decisions is not explicitly mentioned in this utility function. As argued in Zhang (2005), the time horizon can be taken into account by making the propensities endogenous. Maximizing $U(t)$ subject to (12) yields

$$l(t) = \frac{\eta \hat{y}(t)}{R(t)}, \quad c_s(t) = \frac{\gamma \hat{y}(t)}{p(t)}, \quad c_i(t) = \xi \hat{y}(t), \quad c_z(t) = \frac{\varsigma \hat{y}(t)}{p_z}, \quad s(t) = \lambda \hat{y}(t), \quad (13)$$

where

$$\eta \equiv \rho \eta_0, \quad \gamma \equiv \rho \gamma_0, \quad \xi \equiv \rho \xi_0, \quad \varsigma \equiv \rho \varsigma_0, \quad \lambda \equiv \rho \lambda_0,$$

$$\rho \equiv \frac{1}{\eta_0 + \gamma_0 + \xi_0 + \varsigma_0 + \lambda_0}.$$

We assume that saving and transactions are conducted without any barriers. According to the definition of $s(t)$, the wealth accumulation of the household is

$$\dot{\bar{k}}(t) = s(t) - \bar{k}(t). \quad (14)$$

This equation states that the change in wealth equals saving minus dissaving.

FULL USE OF LAND

Land is used for the residential purposes and service production

$$l(t)N + L_s(t) = L. \quad (15)$$

BALANCE OF DEMAND AND SUPPLY FOR SERVICES

The equilibrium condition for services is

$$c_s(t)N + D_T(t) = F_s(t). \quad (16)$$

We define the GDP by: $Y(t) = F_i(t) + p(t)F_s(t)$.

TRADE BALANCE

Let $\bar{K}(t)$ stand for the total wealth owned by the country's population, that is, $\bar{K}(t) = \bar{k}(t)N$.

The capital owned by the population is not necessarily equal to the level of capital stocks employed by the country. We use $E(t)$ to denote the balance of trade. We have

$$E(t) = r^* (\bar{K}(t) - K(t)). \quad (17)$$

We have thus built the dynamic growth model with endogenous wealth, consumption, and tourism.

3. THE DYNAMICS OF THE NATIONAL ECONOMY

The appendix shows that the motion of the economic system is determined by a single differential equation. The following lemma shows how we can determine the motion of all the variables in the dynamic system.

LEMMA

The variables, k_i , k_s , and w are uniquely determined as functions of r^* by (3) and (4). The motion of the land rent is determined by

$$\dot{R}(t) = \bar{\Lambda}(R(t)), \quad (18)$$

in which $\bar{\Lambda}(t)$ is a function of $R(t)$ defined in the appendix. We determine all the other variables as functions of $R(t)$ as follows: $\bar{k}(t)$ by (A12) $\rightarrow \bar{K}(t) = \bar{k}(t)N \rightarrow K(t)$ by (A11) $\rightarrow p(t)$ by (A9) $\rightarrow K_i(t)$ and $K_s(t)$ by (A1) $\rightarrow N_i(t)$ and $N_s(t)$ by (A1) $\rightarrow D_T(t)$ by (8) $\rightarrow \hat{y}(t)$ by (A4) $\rightarrow l(t)$, $c_i(t)$, $c_s(t)$, $c_z(t)$, and $s(t)$ by (13) $\rightarrow L_s(t)$ by (15) $\rightarrow F_i(t)$ by (1) $\rightarrow F_s(t)$ by (4).

The lemma implies that the motion of economic system at any point in time can be uniquely described as functions of the land rent and the other exogenous variables (the rate of interest, land resource, technology, and preference). In Turnovsky's (1996) model for a small open economy domestic capital accumulation involves convex adjustment costs. The equilibrium growth rates of domestic capital and consumption are found to be largely independent. The equilibrium growth rate of domestic capital is determined by production conditions. The equilibrium growth rate of consumption is determined primarily by tastes. Turnovsky's model has a single production sector and the production function is explicitly dependent only on capital, as follows: $F(t) = aK(t)$ where $F(t)$ is the output of the sector, $K(t)$ is the capital employed by the domestic sector, and a is a parameter. In our model, the capital intensity (rather than the capital stock, as in the Turnovsky model) and the wage rate are determined by the internationally fixed interest rate and production conditions. This result derives from the assumption that capital is internationally freely mobile and the interest rate is fixed. Nevertheless, the total output levels, the capital stocks employed by the economy, and the economic production structure are determined not only by the production conditions and the international rate of interest, but also by tastes. This results from our modelling structure, which has economic structures as endogenous variables. As preference determines demand and demand affects the labour distribution, the variables just mentioned are dependent on the preference. Moreover, consumption is not only determined by preferences but it is also related to the rate of interest and the production conditions. This is because the rate of interest and the production conditions affect the disposable income.

Eicher *et al.* (2008) observe, “Previous authors have specified the borrowing cost to increase with the nation’s level of debt. This specification, together with a constant rate of time preference and inelastic labor supply, implies that terms of trade shocks have no dynamic effects. The only response is that consumption fully adjusts instantaneously, with the current account remained unchanged.” Although it is formed in an alternative framework, our model also predicts similar conclusions. To explain this, we first note that from the appendix we know that the disposable income $\hat{y}(t)$ is not affected by the price of imported goods p_Z . The price affects the system only through the equation, $p_Z c_Z(t) = \varsigma \hat{y}(t)$. From this equation we have

$$\frac{d c_Z(t)}{c_Z(t)} = - \frac{d p_Z}{p_Z}.$$

This implies that the rise in the price (terms of trade) has no effect on the other variables in the dynamic system, except for the purchase of imported goods. According to Harberger (1950) and Laursen and Metzler (1950) a deterioration in the terms of trade would reduce real income, thereby reducing saving and investment to cause a deterioration of the current account balance. Our analysis does not predict the Harberger-Laursen-Metzler effect. This result depends on the assumed Cobb-Douglas utility function. From the above equation we see that a deterioration in the terms of trade causes the same proportional reduction in the consumption of imported goods. Accordingly, there is no Harberger-Laursen-Metzler effect in our model. It should be noted that the Harberger-Laursen-Metzler effect might occur if we use other forms of utility function.

As it is difficult to explicitly interpret the analytical results, we simulate the model. We specify the parameters as follows:

$$\begin{aligned} r^* &= 0.05, \quad p_Z = 4, \quad \delta_k = 0.05, \quad N = 10, \quad L = 1, \quad A_i = 1.4, \quad A_s = 1.1, \\ \alpha_i &= 0.31, \alpha_s = 0.22, \quad \beta_s = 0.58, \quad \lambda_0 = 0.62, \quad \xi_0 = 0.15, \quad \gamma_0 = 0.07, \\ \eta_0 &= 0.07, \quad \varsigma_0 = 0.03, a = 1, \quad y_f = 5, \quad \phi = 1.8, \quad \varepsilon = 1.3. \end{aligned} \tag{19}$$

The rate of interest is fixed at 5 per cent and the population is 10. The land is unity. The propensity to save is 0.62. The propensity to consume is 0.10. The propensities to consume services, imported goods, and housing are respectively 0.07, 0.03, 0.07. The price of imported goods is 4. Some empirical studies show that income elasticity of tourism demand is well above unity

(Syriopoulos, 1995; Lanza *et al.*, 2003). According to Lanza *et al.* (2003), the price elasticity is in the range from 1.03 to 1.82 and income elasticities are in the range from 1.75 to 7.36. There are other studies on elasticities of tourism (e.g., Gafin-Múnos, 2007).

Following the lemma, we calculate the time-independent variables as follows:

$$k_i = 8.39, \quad k_s = 7.09, \quad w = 1.87. \tag{20}$$

We choose the initial condition with $R(0) = 12$. We plot the motion of the dynamic system in Figure 1. As the initial value of the land is fixed higher than its long-term equilibrium value, the land rent falls over time. As the price falls, demand of tourism rises. The lot size falls and the land used by the service sector rises. The consumption level of goods and services is reduced. The consumption level of imported goods fall as the disposable income falls. The output of the industrial sector rises, while the output of the service sector falls. The labour force is shifted from the service sector to the industrial sector. The capital stocks employed by the economy and the industrial sector are slightly higher. The capital stock employed by the service sector falls slightly.

Figure 1. The Motion of the National Economy

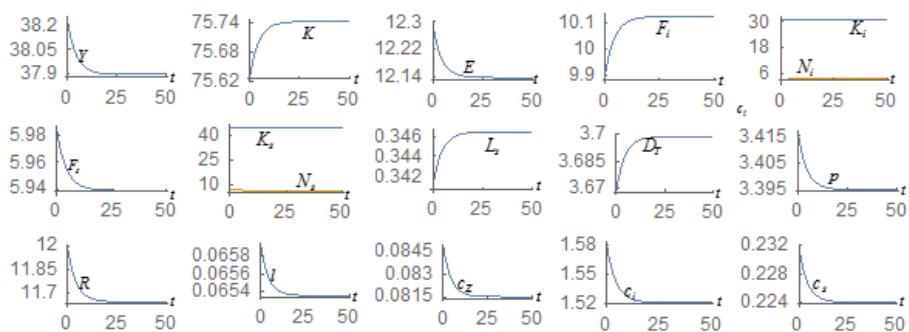


Figure 1 shows how the system approaches an equilibrium point. We calculate the equilibrium values of the variables as follows

$$\begin{aligned}
 p &= 3.4, \quad R = 11.64, \quad Y = 37.9, \quad E = 12.14, \quad K = 75.74, \quad \bar{K} = 67.38, \\
 DT &= 3.7, \quad N_i = 3.74, \quad N_s = 6.26, \quad K_i = 31.38, \quad K_s = 44.36, \\
 L_s &= 0.35, \quad F_i = 10.12, \quad F_s = 5.94, \quad c_i = 1.52, \quad c_s = 0.22, \\
 c_z &= 0.081, \quad l = 0.065.
 \end{aligned}$$

The eigenvalue at the equilibrium point is -0.23 . Accordingly the unique equilibrium point is stable. The stability is important as it ensures that we can effectively conduct comparative dynamic analysis.

4. COMPARATIVE DYNAMIC ANALYSIS

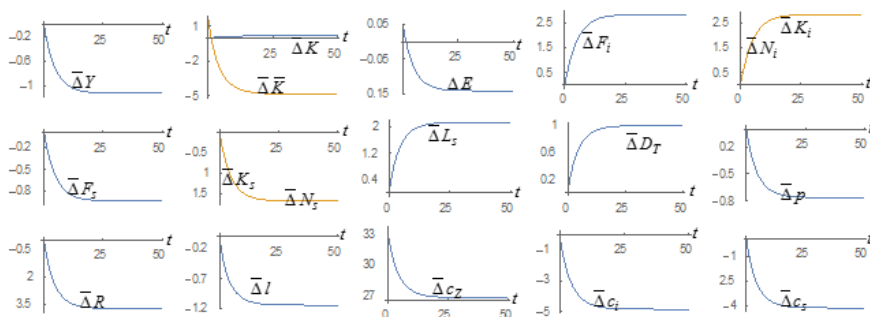
The previous section plotted the changes of the variables. This section examines how changes in some parameters affect the national economy. As we have shown how to simulate the motion of the system, the comparative dynamic analysis is straightforward. We introduce a variable $\bar{\Delta}x(t)$ to stand for the percentage change rate of the variable $x(t)$ due to the change in a parameter value.

A RISE IN THE PROPENSITY TO CONSUME IMPORTED GOODS

First, we examine the impact of the following change in the propensity to consume imported goods: $\zeta_0: 0.03 \Rightarrow 0.04$. We have $\bar{\Delta}k_i = \bar{\Delta}k_s = \bar{\Delta}w = 0$. The change has no impact on the capital intensity and the wage rate. As these variables are dependent only on the interest rate and technologies, the change in the preference has no impact on their values. The changes in the time-dependent variables are plotted in Figure 2. The demand for imported goods is increased. The economy spends more money on the consumption of foreign goods and less on industrial goods and services. The fall in domestic demand for the local services reduces the price of services. The fall in the price attracts more foreign tourists. As the fall of service consumption by the population dominates the rise by foreign tourists, the total service production is reduced. In association with the falls in service supply and price of services, some labour force is shifted from the service sector to the industrial sector.

Initially the rise in wealth is faster than the rise in the total capital employed by the country. The trade balance is improved, but soon the wealth falls and the total capital employed by the country rises slightly. The trade balance deteriorates. In association with the fall in the price of services, the industrial sector increases the scale of production. As a rise in the propensity to consume imported goods also implies a fall in the (relative) propensity to consume housing, the lot size is reduced, which results in a fall in the land rent. Some of the land is shifted from the housing market to the service sector. As the value produced by the service sector is lower than that produced by the industrial sector, GDP falls over time.

Figure 2. A Rise in the Propensity to Consume Imported Goods



A RISE IN THE INTEREST RATE

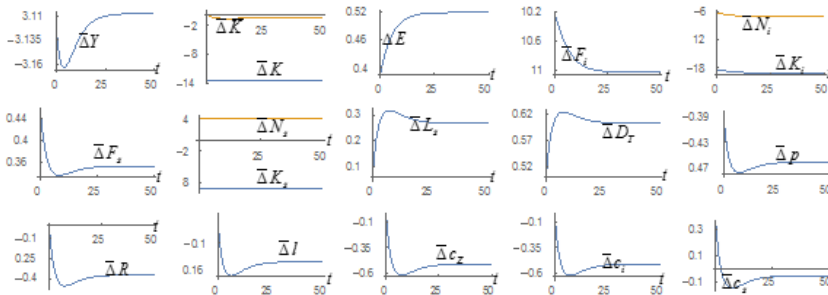
We now allow the rate of interest to be changed as follows: $r^* = 0.05 \Rightarrow 0.06$. The rise in capital cost causes the capital intensity and wage rate to fall as follows:

$$\bar{\Delta} k_i = \bar{\Delta} k_s = -12.9, \quad \bar{\Delta} w = -4.2.$$

The net result in the rise of the interest rate and fall in capital intensity leads to the fall in the wage rate. We plot the changes in the time-dependent variables in Figure 3. The two sectors use less capital in response to the rising cost of capital. The capital stock employed by the national economy is reduced. The net consequence of reduced wage income and increased rate of interest is a slight fall in the wealth. The price of services and the land rent both increase over time. There are more foreign visitors as the country's economic attractiveness improves. The output of the industrial sector falls in response to the

higher production cost, while the output level of the service sector is slightly increased as a net result of increased demand and production cost. Some of the labour force is shifted from the industrial sector to the service sector. The GDP falls over time, but the trade balance is improved. The lot size is slightly reduced. The household consumes more services and less industrial and imported goods.

Figure 3. An Increase in the Rate of Interest

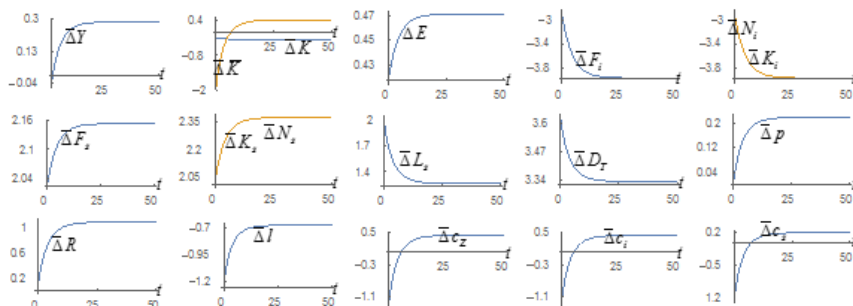


AN IMPROVEMENT IN GLOBAL ECONOMIC CONDITIONS

As the country is dependent on tourism, changes in global economic conditions should affect the national economy. We allow the foreign income to be enhanced as follows: $y_f = 5 \Rightarrow 5.1$. Figure 4 plots the changes in the time-dependent variables. The capital intensity and wage rate are not affected. More foreigners come to visit the country. The boom in tourism causes the price of services and output of the service sector to rise. The labour, capital, and land inputs employed by the service sector are all increased. More foreign visitors reduce the lot size and increase the land used by the service sector. The income of the household is initially reduced but increases later on. The consumption level of goods and services, wealth level, and consumption of imported goods are all reduced initially but increase in the long term. The trade balance is improved. An improvement in the global income reduces the living conditions and wealth of the domestic household in the short term but improves these variables in the long term. According to Harzri and Sgro (1995) an increase in the international tourism has a positive effect on long-run economic growth. Our result shows that this conclusion is true in the long term, but not necessarily true in the short term. Similarly, their model

shows that for a small open economy the growth in tourist consumption of services increases welfare. Our simulation demonstrates that this conclusion is valid in the long term, but not necessarily true in the short term. We get this “new insight” because our model explicitly shows transitional processes in the economic dynamics. As demonstrated by Chao *et al.* (2006), an expansion of tourism can result in capital decumulation in a two-sector dynamic model with a capital-generating externality. Our simulation demonstrates the same conclusion without any externality.

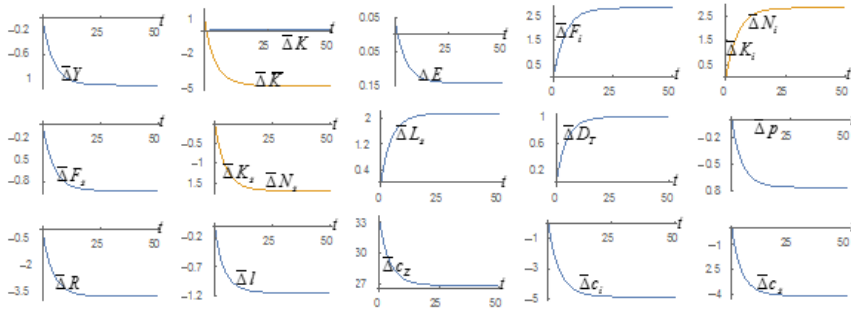
Figure 4. An Improvement in Economic Conditions of the Global Economy



A RISE IN THE PRICE ELASTICITY OF TOURISM

Many factors affect the price elasticity of tourism. For instance, fast development of other countries' tourist industries may make the home country more elastic. We now change the price elasticity as follows: $\varepsilon = 1.3 \Rightarrow 1.4$. The capital intensity and wage rate are not affected. The changes in the time-dependent variables are plotted in Figure 5. A rise in price elasticity reduces the tourist demand and the price of services. The national trade balance deteriorates as a consequence of the lower number of tourists. The output of the industrial sector is increased as some of labour force is shifted from the service sector to the industrial sector. The lot size rises, but the land input of the service sector is reduced. GDP rises initially but soon falls. The consumption levels of industrial goods, imported goods and services and wealth level rise initially but fall in the long term.

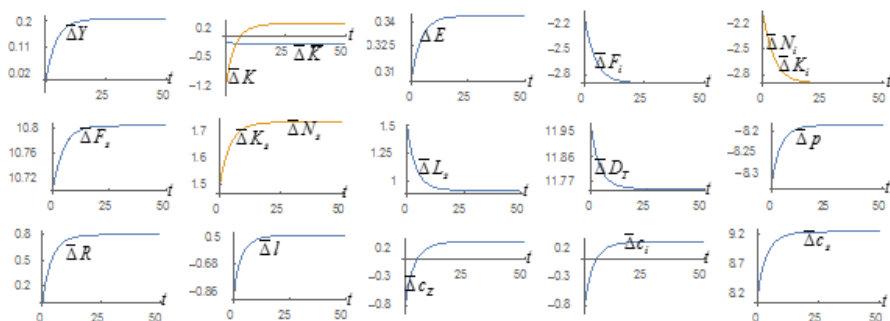
Figure 5. A Rise in the Price Elasticity of Tourism



A RISE IN TOTAL PRODUCTIVITY OF THE SERVICE SECTOR

We now change the total productivity of the service sector as follows: $A_s = 1.1 \Rightarrow 1.2$. There is no impact on the capital intensity and wage rate. The changes in the time-dependent variables are plotted in Figure 6. The output level of services is increased and the price of services falls. The net impact of the rise in the output of the service sector and the fall in the output of the industrial sector is that the GDP is almost not affected. Some of labour force is shifted from the industrial sector to the service sector. The lot size is reduced and the land input employed by the service sector is increased. The consumption level of services increases in association with the fall in the price. The consumption levels of industrial and imported goods are slightly affected. The trade balance is improved. Wealth falls initially but rises in the long term. As the price of services falls, there are more foreign visitors.

Figure 6. A Rise in Total Productivity of the Service Sector



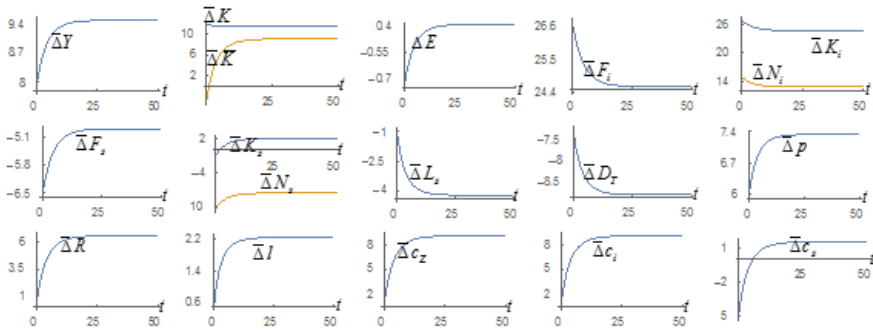
A RISE IN TOTAL PRODUCTIVITY OF THE INDUSTRIAL SECTOR

We now allow the total productivity of the industrial sector to be enhanced as follows: $A_i : 1.4 \Rightarrow 1.5$. The effects on the capital intensity and wage rate are as follows

$$\bar{\Delta} k_i = \bar{\Delta} k_s = \bar{\Delta} w = 10.52.$$

The changes in the time-dependent variables are plotted in Figure 7. It is interesting to note that Baumol (1967) builds a two-sector growth model of a closed economy. Baumol discusses different implications of technological changes in the manufacturing sector. A main conclusion from his model is that technological advances in the manufacturing sector will enhance service prices and the output of the service sector. Our model predicts a rise in the service price and a fall in the service sector's output. The conclusions arrived at by our model and Baumol's analysis differ because our model is for an open economy and Baumol's is for a closed economy.

Figure 7. A Rise in Total Productivity of the Industrial Sector



5. CONCLUSIONS

This paper has described an economic growth model of a small open economy with tourism and imported goods in a perfectly competitive economy. The study focuses on the effects of changes in the terms of trade, with a preference for imported goods, on the dynamic paths of trade balance and economic growth. The national economy consists of the two sectors. Following the traditional literature on small open economies, we assume that the rate of interest is fixed in the international market. The production side is the same as in neoclassical growth theory. We applied the utility function proposed by Zhang to determine saving and consumption. We simulated the model and demonstrated that the system has a unique stable equilibrium point. We examined the effects of changes on the terms of trade, the propensity to consume imported goods, the rate of interest, the price elasticity of tourism, and the total productivity of the service sector. The comparative dynamic analysis provides some important insights.

We may also get hints on policy implications from our analysis. For instance, we have shown that a stronger preference for imported goods harms national economic growth. It may thus be desirable for governments to weaken the preference for foreign goods or to tax them. The government could also improve safety, infrastructure elements and other amenities to reduce the price elasticity of

tourism demand. It should be noted that the model contains with many strict assumptions. We may extend and generalise the model in different directions. For instance, it is important to study the economic dynamics when utility and production functions take on other functional forms. We do not consider the possibility of domestic households travelling to other countries. Monetary issues such as exchange rates and inflation policies are important for understanding trade issues, and we also neglect tariffs and other taxes.

APPENDIX: PROVING THE LEMMA

We have shown that k_i , w , and k_s are determined as functions of r^* , which is fixed in the international market. From $K_j = k_j N_j$ and (8), we have

$$K_i = (K - k_s N)k_0 k_i, \quad K_s = (k_i N - K)k_0 k_s, \quad (\text{A1})$$

where we omit the time variable in the expressions. From (5), we solve

$$R = \frac{w_s N_s}{L_s}, \quad (\text{A2})$$

where we also use $l_s = L_s / N_s$ and $w_s \equiv w \gamma_s / \beta_s$. Insert (A2) in (15)

$$l N + \frac{w_s N_s}{R} = L. \quad (\text{A3})$$

From the definition of \hat{y} , we have

$$\hat{y} = (1 + r^*)\bar{k} + w + \frac{RL}{N}. \quad (\text{A4})$$

From (A4) and $l = \eta \hat{y} / R$ in (13)

$$l = \frac{(1 + r^*)\eta \bar{k} + \eta w}{R} + \frac{\eta L}{N}.$$

Insert this equation in (A3)

$$(1 + r^*)\eta \bar{k} + \eta w + \frac{w_s N_s}{N} = (1 - \eta)\frac{L}{N}R. \quad (\text{A5})$$

From $r^* + \delta_k = \alpha_s p F_s' / K_s$ and (16) we have

$$c_s N + D_T = \frac{r_\delta K_s}{\alpha_s p}. \quad (\text{A6})$$

Insert $c_s = \gamma \hat{y} / p$ in (A6)

$$\gamma \hat{y} N + p D_T = \frac{r_\delta K_s}{\alpha_s}.$$

Insert (A4) into this equation

$$\left((1 + r^*) \bar{k} + w + \frac{RL}{N} \right) \gamma N + a y_f^\phi p^{1-\varepsilon} = \frac{r_\delta K_s}{\alpha_s}, \quad (\text{A7})$$

where we also use (9). We now discuss two cases, $\varepsilon = 1$ and $\varepsilon \neq 1$, separately.

In $\varepsilon = 1$, (A7) becomes

$$\left((1 + r^*) \bar{k} + w + \frac{RL}{N} \right) \gamma N + a y_f^{-\phi} = \frac{r_\delta K_s}{\alpha_s}. \quad (\text{A8})$$

From (A5) and (A8), we delete R and get an equation which gives the relation between \bar{k} and K .

As discussed in the other case, we can determine a differential equation in terms of K .

We now examine $\varepsilon \neq 1$. From (5) we have

$$p = \frac{w R^{\gamma_s}}{\beta_s A_s k_s^{\alpha_s} w_s^{\gamma_s}}, \quad (\text{A9})$$

where we also use $l_s = w_s / R$ from (A2). Insert (A9) in (A7)

$$(1 + r^*) \gamma N \bar{k} + \gamma N w + \gamma RL + a y_f^\phi \left(\frac{w R^{\gamma_s}}{\beta_s A_s k_s^{\alpha_s} w_s^{\gamma_s}} \right)^{1-\varepsilon} = \frac{r_\delta K_s}{\alpha_s}. \quad (\text{A10})$$

Substituting $N_s = (k_i N - K) k_0$ from (8) into (A5) and $K_s = (k_i N - K) k_0 k_s$ from (A1) into (A10) respectively yields

$$\begin{aligned} (1 + r^*) \eta N \bar{k} - w_s k_0 K &= \bar{\Psi}, \\ (1 + r^*) \gamma N \bar{k} + \frac{r_\delta k_0 k_s K}{\alpha_s} &= \tilde{\Psi}, \end{aligned} \quad (\text{A11})$$

where

$$\bar{\Psi}(R(t)) = (1 - \eta)LR - \eta Nw - k_i w_s k_0 N,$$

$$\tilde{\Psi}(R(t)) \equiv \frac{r_\delta k_0 k_i k_s N}{\alpha_s} - \gamma Nw - \gamma RL - a y_f^\phi \left(\frac{wR^{\gamma_s}}{\beta_s A_s k_s^{\alpha_s} w_s^{\gamma_s}} \right)^{1-\varepsilon}.$$

From (A11), we solve

$$\bar{k} = \Psi(R) \equiv \frac{1}{(1 + r^*)N} \left(\tilde{\Psi} + \frac{r_\delta k_s \bar{\Psi}}{\alpha_s w_s} \right) \left(\gamma + \frac{\eta r_\delta k_s}{\alpha_s w_s} \right)^{-1}. \quad (\text{A12})$$

By the following procedure we can determine all the variables as functions of R : \bar{k} by (A12) $\rightarrow K$ by (A11) $\rightarrow p$ by (A9) $\rightarrow K_i$ and K_s by (A1) $\rightarrow N_i$ and N_s by (A1) $\rightarrow D_T$ by (8) $\rightarrow \hat{y}$ by (A4) $\rightarrow l, c, c_s, c_z, s$ by (13) $\rightarrow L_s$ by (15) $\rightarrow F_i$ by (1) $\rightarrow F_s$ by (4).

From this procedure and (14), we have

$$\dot{\bar{k}} = \Lambda(R) \equiv \lambda \hat{y} - \bar{k}. \quad (\text{A13})$$

Taking derivatives of (A12) with respect to time yields

$$\dot{\bar{k}} = \frac{d\Psi}{dR} \dot{R}. \quad (\text{A14})$$

We do not provide the expression of $d\Psi/dR$ because it is tedious. Solve (A13) and (A14)

$$\dot{R} = \bar{\Lambda}(R) \equiv \Lambda \left(\frac{d\Psi}{dR} \right)^{-1}. \quad (\text{A15})$$

We have thus proved the lemma.

REFERENCES

- Andereck, K.L.; Valentine, K.M.; Knopf, R.C.; Vogt, C.A. (2005) Residents' perceptions of community tourism impacts, *Annals of Tourism Research*, 32 (4), 1056-1076.
- Antonakakis, N.; Dragouni, M.; Filis, G. (2015) How strong is the linkage between tourism and economic growth in Europe?, *Economic Modelling*, 44, 142-155.
- Balaguer, L.; Cantavella-Jorda, M. (2002) Tourism as a long-run economic growth factor: The Spanish case, *Applied Economics*, 34, 877-884.
- Baumol, W.J. (1967) Macroeconomics of unbalanced growth: The anatomy of urban crisis, *American Economic Review*, 57(3), 415-426.
- Benigno, G.; Benigno, P. (2003) Price stability in open economies, *Review of Economic Studies*, 70, 743-764.
- Blake, A.; Sinclair, M.T.; Campos, J.A. (2006) Tourism productivity: Evidence from the United Kingdom, *Annals of Tourism Research*, 33, 1099-1120.
- Chao, C.C.; Hazari, B.R., Laffargue, Y.P.; Yu, E. S.H. (2006) Tourism, Dutch disease and welfare in an open dynamic economy, *Japanese Economic Review*, 57, 501-515.
- Chao, C.C.; Hazari, B.R.; Laffargue, Y.P.; Yu, E. S.H. (2009) A dynamic model of tourism, employment, and welfare: The case of Hong Kong, *Pacific Economic Review*, 14, 232-245.
- Chou, M.C. (2013) Does tourism development promote economic growth in transition countries? A Panel Data Analysis, *Economic Modelling*, 33, 226-232.
- Copeland, B. R. (1991) Tourism, welfare and de-industrialization in a small open economy, *Economica*, 58, 515-529.
- Copeland, B.R. (2012) Tourism and welfare-enhancing export subsidies, *Japanese Economic Review*, 63, 232-243.
- Corden, W.M.; Neary, J.P. (1982) Booming sector and de-industrialization in a small open economy, *Economic Journal*, 92, 825-848.
- Dritsakis, N. (2004) Tourism as a long-run economic growth factor: An empirical investigation for Greece using causality analysis, *Tourism Economics*, 10, 305-316.
- Drubarry, R. (2004) Tourism and economic growth: The case of Mauritius, *Tourism Economics*, 10, 389-401.

- Dwyer, L.; Forsyth, P.; Spurr, R. (2004) Evaluating tourism's economic effects: New and old approaches, *Tourism Management*, 25, 307-317.
- Eicher, T.S.; Schubert, S.F.; Turnovsky, S.J. (2008) Dynamic effects of terms of trade shocks: The impact on debt and growth, *Journal of International Money and Finance*, 27, 876-896.
- Frederick, S.; Loewenstein, G.; O'Donoghue, T. (2002) Time discounting and time preference: A critical review, *Journal of Economic Literature*, 11, 317-336.
- Gali, J.; Monacelli, T. (2005) Monetary policy and exchange rate volatility in a small open economy, *Review of Economic Studies*, 72, 707-734.
- Gafin-Múños, T. (2007) German demand for tourism in Spain, *Tourism Management*, 28, 12-22.
- Harberger, A.C. (1950) Currency depreciation, income, and the balance of trade, *Journal of Political Economy*, 58, 47-60.
- Hazari, B.R.; Lin, J.J. (2011) Tourism, terms of trade and welfare to the poor, *Theoretical Economics Letters*, 1, 28-32.
- Hazari, B.R.; Sgro, P.M. (1995) Tourism and growth in a dynamic model of trade, *Journal of International Trade and Economic Development*, 4, 243-252.
- Hazari, B.R.; Sgro, P.M. (2004) *Tourism, Trade and National Welfare*, Amsterdam, Elsevier.
- Ikedo, S. (2001) Weakly non-separable preferences and the Harberger-Laursen-Metzler effect, *Canadian Journal of Economics*, 34, 290-307.
- Kim, H.J.; Chen, M.H.; Jang, S. (2006) Tourism expansion and economic development: The case of Taiwan, *Tourism Management*, 27, 925-933.
- Kollmann, R. (2001) The Exchange rate in a dynamic-optimizing business cycle model with nominal rigidities: A quantitative Investigation, *Journal of International Economics*, 55, 243-262.
- Kollmann, R. (2002) Monetary policy rules in the open economy: Effects on welfare and business cycles, *Journal of Monetary Economics*, 49, 899-1015.
- Kose, M.A. (2002) Explaining business cycles in small open economies: How much do world prices matter?, *Journal of International Economics*, 56, 299-327.
- Lane, P.R. (2001) The new open economy macroeconomics: A survey, *Journal of International Economics*, 54, 235-266.
- Lanza, A.; Temple, P.; Urga, G. (2003) The implications of tourism specialisation in the long run: An econometric analysis for 13 OECD economies, *Tourism Management*, 24, 315-321.

- Laursen, S.; Metzler, L.A. (1950) Flexible exchange rates and the theory of employment, *Review of Economics and Statistics*, 32, 281-299.
- Luzzi, G.F.; Flückiger, Y. (2003) Tourism and international trade: Introduction, *Pacific Economic Review*, 8, 239-243.
- Mansoorian, A. (1993) Habit persistence and the Harberger-Laursen-Metzler effect in an infinite horizon model, *Journal of International Economics*, 34, 153-166.
- Marin-Pantelescu, A.; Tigu, G. (2010) Features of the travel and tourism industry which may affect pricing, *Journal of Environmental Management and Tourism*, 1, 8-11.
- Matarrita-Cascante, D. (2010) Beyond growth: Reaching tourism-led development, *Annals of Tourism Research*, 32(4), 1141-1163.
- Matsuyama, K. (1987) Current account dynamics in a finite horizon model, *Journal of International Economics*, 23, 299-313.
- Mendoza, E.G. (1995) The terms of trade, the real exchange rate, and economic fluctuations, *International Economic Review*, 36, 101-137.
- Obstfeld, M. (1982) Aggregate spending and the terms of trade: Is there a Harberger-Laursen-Metzler effect?, *Quarterly Journal of Economics* 97, 251-270.
- Obstfeld, M.; Rogoff, K. (1996) *Foundations of International Macroeconomics*, Mass., Cambridge: MIT Press.
- Oh, C.O. (2005) The contribution of tourism development to economic growth in the Korean economy, *Tourism Management*, 26, 39-44.
- Sachs, J. (1982) The current account in the macroeconomic adjustment process, *Scandinavian Journal of Economics*, 84, 147-159.
- Schubert, S. F.; Brida, J.G. (2009) A dynamic model of economic growth in a small tourism driven economy, Munich Personal RePEc Archive.
- Svensson, L.E.O.; Razin, A. (1983) The terms of trade and the current account: The Harberger-Laursen-Metzler effect, *Journal of Political Economy*, 91, 97-125.
- Sinclair, M.T.; Stabler, M. (1997) *The Economics of Tourism*, London: Routledge.
- Solow, R. (1956) A contribution to the theory of growth, *Quarterly Journal of Economics*, 70, 65-94.
- Stabler, M.J., Papatheodorou, A.; Sinclair, M.T. (2010) *The Economics of Tourism*, London: Routledge.

- Syriopoulos, T.C. (1995) A Dynamic model of demand for Mediterranean tourism, *International Review of Applied Economics*, 9, 318-336.
- Turnovsky, S.J. (1996) Fiscal policy, growth, and macroeconomic performance in a small open economy, *Journal of International Economics*, 40, 41-66.
- Turnovsky, S.J.; Chattopadhyay, P. (2003) Volatility and growth in developing economies: Some numerical results and empirical evidence, *Journal of International Economics*, 59, 267-295.
- Uzawa, H. (1961) On a two-sector model of economic growth, *Review of Economic Studies*, 29, 47-70.
- Zeng, D.Z.; Zhu, X.W. (2011) Tourism and industrial agglomeration, *Japanese Economic Review*, 62, 537-561.
- Zhang, W.B. (2005) *Economic Growth Theory*, London: Ashgate.
- Zhang, W.B. (2008) *International Trade Theory: Capital, Knowledge, Economic Structure, Money and Prices over Time and Space*, Berlin: Springer.
- Zhang, W.B. (2012) Tourism and economic structure in a small-open growth model, *Journal of Environmental Management and Tourism*, 3(2), 76-92.

(Página deixada propositadamente em branco)

I
IMPRESSA DA UNIVERSIDADE DE COIMBRA
COIMBRA UNIVERSITY PRESS
U

• U



C •