2

Banco de Portugal Economic Studies

Volume IV

BANCO DE PORTUGAL
EUROSYSTEM

2
volume IV

Banco de Portugal
Economic Studies

Please address correspondence to
Banco de Portugal, Economics and Research Department
Av. Almirante Reis 71, 1150-012 Lisboa, Portugal
T +351 213 130 000 | estudos@bportugal.pt

BANCO DE PORTUGAL
EUROSYSTEM

Lisbon, 2018 · www.bportugal.pt

BANCO DE PORTUGAL ECONOMIC STUDIES | Volume IV - n.º 2 | Lisbon 2018 • Banco de Portugal Av. Almirante Reis, 71 | 1150-012 Lisboa • www.bportugal.pt • Edition Economics and Research Department • Design Communication and Museum Department | Publishing and Image Unit • ISSN 2183-5217 (online)

Content

Editorial

April 2018

The first essay, by Manuel Coutinho Pereira and Lara Wemans, is entitled "How long does it takes to enforce a debt in the Portuguese judicial system?". Business surveys in Portugal show that payment risks and the slowness of the justice system are among the most relevant framework costs faced by firms. It is essential for economic agents that the justice system quickly achieves payment of liabilities or at least an assessment of its impossibility.

The paper by Pereira and Wemans uses data from the years 2007 to 2016 covering almost 2.4 million cases to study the determinants and evolution of the length of enforcement lawsuits for collection of civil and commercial debts. Taking into account the cases for which information is available, they found that around 87% of the lawsuits are filed by private law entities, 12% by individuals and only 1% by public law entities (the latter use mostly administrative and tax courts). On average, 22% of the lawsuits were filed by mass litigants (companies that interpose more than 200 lawsuits per year), and this weight began to decrease in 2011. The cases are concentrated in Lisbon and Oporto. During the period under review, the proportion of lawsuits in the enforcement judgeships, specialized organic units, increased. These dealt with 15% of the cases between 2007 and 2009, with this percentage increasing to 90% in 2015 and 2016. The average value of a lawsuit is 21 000 euros but about two thirds are less than 5 000 euros. Only about 36% of the lawsuits ended with the payment of the debt. In 29% of the cases resolved the assets of the debtor were nonexistent or insufficient and in 27% of the cases the lawsuit ended with the plaintiff dropping it or some equivalent outcome. The remaining 9% correspond to more complex or unspecified situations.

The median duration, including data from pending cases, increased between 2007 and 2009, but it has declined since 2011. This decrease is probably related to judicial reforms in the area of debt enforcement. In particular for the reductions that took place after 2013, it will reflect policy measures that expedited the extinction of the lawsuit due to lack of activity of the plaintiff. Nevertheless, such a trend still holds when this latter factor is excluded. The data also point to lawsuits with higher value or lawsuits with procedural aspects (in Portuguese *apensos*) having a slower resolution. Finally, using the methodology of duration analysis, the authors estimated the hazard rates of lawsuit termination for the 2007-2016 global sample and for two sub periods: before 2014, and for 2014 and beyond. The hazard rates are increasing with the length of the case for the global sample and for the sub-sample before 2014. For the subsample thereafter the hazard rates are higher in the first five years, resulting in a substantially lower median duration of the lawsuits.

The econometric duration analysis was based on Cox models, where a set of variables influence the hazard rates. The most relevant results show that the lawsuits with private plaintiffs or mass litigants tend to be faster than those filed by collective persons in general. Lawsuits with higher value tend to have slower resolutions. By comparison with injunctions, lawsuits based on private documents and contracts tend to be more time-consuming. The existence of *apensos* has a negative effect on speed. Finally, some results point to the levels of congestion and complexity of litigation contributing in a quantitatively very relevant way to increase the duration of lawsuits. In general, the econometric results point to a shorter duration in the most recent years. These improvements are certainly associated with the various policy measures taken to improve the efficiency of the justice system. It remains to be seen whether this progress has deepened in the more recent years and whether it extends to other dimensions of the problem such as the proportion of the outstanding amounts that is recovered.

The second paper, by Sara Serra, is "Is the Phillips curve dead?-results for Portugal". The Phillips curve is a negative empirical relationship between inflation and unemployment or, alternatively, a positive relationship between inflation and the output gap. It was initially seen as a long-term economic relationship but work by Phelps and Friedman recast it as a short-term trade-off, dependent on inflation expectations. Over the years, estimates of Phillips curves have become less robust, but they are still important analytical tools relevant for policy makers and in particular for those in charge of conducting monetary policy. This implies that improving the robustness of Phillips curves estimates would be a valuable achievement. What are the reasons for their fragilities? What can we do to overcome them, to the extent that such a goal is feasible?

Serra's analysis starts with a generic Phillips curve formulation where inflation in a given period is a function of inflationary expectations, past inflation levels, current and past import prices and the previous period measure of slack. Given this general specification, the next task becomes to estimate the curve. One first area of concern is finding the right operational variables. To begin with, what are the right import price and inflation measures that should be used in Phillips curves´ estimates? What are the best operational measures for inflation expectations? What is the best measure of the slack in the economy? In each case there are multiple alternatives, sometimes a large number. A second set of problems relates to finding the best specification for the Phillips curve model. Should import prices and inflation expectations really be in the equation? Are there significant non-linearities where, for example, the equation coefficients for extreme levels of slack measures are different from the coefficients for low values of slack measures?

In order to tackle these questions Serra uses data from the first quarter of 1996 up to the last quarter of 2017 and runs a large number of regressions (about 500 different models for each measure of inflation tested). Serra then evaluates the performance of each regression specification by looking at measures of fit and the mean squared errors of out-of-sample predictions over time.

Serra reaches the conclusion that the Harmonized Index of Consumer prices (HICP) excluding the volatile components related to food and energy prices is the best inflation measure for the purposes of Phillips curve estimation. Import prices only seem to have become relevant in the most recent years. Output gaps can be estimated using different methodologies (based on filters, productions functions, unemployment and underemployment rates), but the production function based estimates seem to provide the best Phillips curves estimates, along with two alternative measures of slack (a qualitative survey on labor as a limiting factor to production and the short-term unemployment rate).

After the financial crisis years, the estimates for the output gap coefficients decline, supporting the thesis of flattening of the standard (linear) Phillips curve. This is consistent with some evidence of non-linearities in the model.

Nevertheless, even when evaluated in real-time, Phillips curves estimates still represent an improvement over the forecasting performance of other models and as such they will keep being useful tools for policymakers and for the general understanding of macroeconomic phenomena.

The last paper, by Francisco Dias, Nuno Lourenço and António Rua, bears the title "Forecasting exports with targeted predictors". Forecasting key economic and financial variables is part of the way policymakers monitor developments in any economy. Advances on forecasting have drawn on the use of a multiplicity of large datasets.

The work of Dias, Lourenço and Rua deals with the possibilities generated by these contemporary data-rich environments. It starts from the list of Portugal’s main export destinations, which includes, by exports’ size, Spain, France, Germany, the UK, the US, and others. For each trading partner, a panel of variables is used that includes the main quantitative measures of economic activity (hard data), as well as qualitative assessments (soft data). The overall database includes 766 time series. These international datasets, as well as domestic data, are then used to forecast the year-on-year growth rate of the exports of goods on a monthly basis.

The forecasting relies on methodologies improving on traditional factor models by using targeted predictors. The use of factor analysis to extract the fundamental information in a database with a very large numbers of variables has become more frequent in forecasting. However, for the purposes of forecasting specific variables such as the Portuguese exports of goods, the basic factor extraction methodology may not work well. Even if the “factors” are good summaries of the original large data set, they were not built with the purpose of maximizing their predictive power vis-à-vis a specific variable. To address this problem the authors use a methodology, the LARS-EN algorithm that implicitly pre-selects a limited number of predictor variables, the most relevant, targeted for each specific forecasting problem.

The results obtained, for an out-of-sample period from January 2009 to December 2016, using soft indicators for these countries show that this methodology achieves forecasting gains up to 12-month ahead. The forecasting accuracy gains delivered by factor forecasts using targeted predictors are statistically significant against an autoregressive model with an optimized number of lags. In general, the best forecast performance is achieved with no more than 70 variables. However, in the particular case of nowcasting, the authors find that resorting solely to national data, with no pre-selection of predictors, yields greater forecasting accuracy. Hence, data from Portugal's main trading partners is more informative to produce h-step ahead forecasts. A final result is that pooling hard data with soft data does not seem to bring additional predictive power for the forecasting of exports of goods.

How long does it take to enforce a debt in the Portuguese judicial system?

Manuel Coutinho Pereira
Banco de Portugal

Lara Wemans
Banco de Portugal

Abstract
This paper applies duration analysis to a dataset covering all enforcement cases related to commercial and civil debts, dealt with by first instance courts in Portugal between 2007 and 2016. Evidence points to a strong reduction in the duration of cases since the beginning of the current decade. At that time, the median duration was above five years and it declined to around two and a half years by the end of the period under analysis. The probabilistic profile of case resolution changed significantly, with the hazard rate of resolution of pending cases being nowadays higher in their initial stage. Case duration is influenced notably by its complexity, proxied by a higher claim value, the existence of procedural aspects (*apensos*), and also by the overall enforcement litigation faced by the *comarca*. (JEL: K40, H11, C41)

Introduction

The ability to ensure the fulfilment of an obligation through the judicial system is essential for the regular functioning of the market, and the mass use of credit and deferred payment increased its relevance. The efficiency of debt enforcement should be a relevant part of the framework costs that companies consider in their investment decisions, thus affecting potential economic growth. Recent evidence based on surveys to companies shows that, on the one hand, payment risk is particularly high in Portugal (European Payment Report 2017) and, on the other hand, justice system delays are a very important issue for Portuguese companies (Business Cost of Context Survey from Statistics Portugal 2015).

From the perspective of companies, effectiveness of debt enforcement is related with the ability of the judicial system to quickly promote the fulfilment

Acknowledgements: The authors thank Direção-Geral de Política da Justiça for providing the data on enforcement cases and for valuable clarifications and Carmo Noronha and Marta Veloso for the treatment and help in the interpretation of data on the consultations made to the dataset of banking accounts. The authors are also grateful for the discussion with participants in a seminar at the Economics and Research Department and the comments made by Ana Gouveia, João Amador, Jorge Correia da Cunha, Manuela Espadaneira, Miguel Gouveia and Nuno Garoupa. The opinions expressed in the article are those of the authors and do not necessarily coincide with those of Banco de Portugal or the Eurosystem. Any errors and omissions are the sole responsibility of the authors.

E-mail: manuel.coutinho.pereira@bportugal.pt; lara.wemans@bportugal.pt

of the obligation or, if that is not possible (for instance, due to lack of assets to seize), to declare the debt as uncollectible. As a result, the effectiveness of the justice system in this area closely relates not only with the ability to ensure the fulfilment of the obligation, when the debtor has that capacity, but also to the duration of cases.

In broad terms, an enforcement case is a lawsuit by which the plaintiff demands the fulfilment of an obligation, generally a debt payment. Debt enforcement is based on a document that creates or recognizes such an obligation - the enforceable title (*título executivo*) - which can be generated inside the judicial system (as a court decision or an injunction procedure), or outside it (as a cheque or a promissory note). The pledging plays a major role in these cases. This is the judicial seizure of debtor's assets, giving the plaintiff the right to be paid before any other creditor without previous guarantees over those assets. Moreover, following an enforcement case, several procedural aspects (*apensos*) can be raised by the defendant or third parties that must be resolved so that the enforcement case can be closed. Therefore, the duration and effectiveness of enforcement lawsuits closely relates both with the ability of the judicial system to identify and pledge assets and the rights provided to the debtor and other creditors (Leiras 2015).

The enforcement procedure in Portugal was constantly changing over the last decades, with significant modifications in the role of the different parties involved (Pimenta 2012). The 2003 reform meant a clear paradigm shift, with the establishment of enforcement officers in order to reduce the reliance on the ability of the plaintiff to identify the assets to seize. This change promoted the de-jurisdictionalisation of debt enforcement as the enforcement officer was given the powers to coordinate the process even if the judge maintained a controlling role (Gomes 2007). The reform implementation generated clear blockages, as the clarification of the powers of different players[1] and the establishment of instruments for its proper functioning were only gradually accomplished (Lourenço 2017).

More recently, are also worth mentioning some changes set out by the new civil procedure code in the enforcement area. On the one hand, the transitory rules created by the decree-law no. 4/2013 of January 11 were made permanent, broadening the rules to close cases and defining a 6-month deadline for the closure of cases due to inactivity of the plaintiff. On the other hand, debtor's rights were enhanced, with a narrowing of the set of private documents that can be used as enforceable titles and the reduction of the claim value beyond which a greater intervention by the judge is required for cases based on non-judicial enforceable titles (de Freitas 2014).

1. In this regard, changes introduced by the decree-law no. 226/2008 of November 20, which transferred some responsibilities from judges and judicial clerks to enforcement officers, were particularly relevant (Araújo 2009).

As regards procedural changes, it is important to highlight, in 2013, the streamlining of the procedure to freeze bank accounts. The need for an intervention by the judge was abolished and the procedure to identify the banks in which the debtor has deposits or bank accounts was simplified[2] (Leiras 2015). Moreover, in 2014 a new procedure was created - PEPEX (*Procedimento Extrajudicial Pré-Executivo*) - which allows the holder of an enforceable title to obtain information regarding the feasibility of an enforcement case before presenting it, namely by gathering information available on several datasets about the assets held by the debtor.

The constant legislative changes in enforcement area and the considerable investment in the streamlining of procedures show the high priority given by the Portuguese legislator to the effectiveness of debt enforcement. This paper focuses on a particular aspect of enforcement effectiveness: its duration. In particular, duration analysis methods are applied to a dataset covering enforcement cases related to commercial and civil debts dealt with by first instance courts in Portugal between 2007 and 2016, containing more than 2 million observations. This methodology allows an estimation of the probability that an enforcement case is resolved depending on the time elapsed since it entered the system and to identify the characteristics impacting that duration.

The vast majority of studies of court efficiency are based on court level data. The main findings indicate high heterogeneity in efficiency, with a lack of a clear link between the human resources allocated and the duration of cases (Voigt 2016). Evidence with Portuguese data confirms these conclusions (Pereira and Wemans 2017). There is also evidence of a large impact of judge characteristics on the number of cases resolved and duration of cases (Christensen and Szmer 2012).

The implementation of duration analysis to judicial cases requires case level information and the studies that apply this methodology are relatively scarce,[3] probably due to the difficulty in gathering datasets with that detail. One of the main benefits in using case level information as opposed to more aggregate data is the availability of information about characteristics capturing case complexity, such as normative complexity cited in the decision (Vita 2012), claim value or existence of multiple parties on either the plaintiff or the defendant sides (Bielen *et al.* 2017). In addition, these datasets allow the

2. Between September 2013 and December 2016, more than one million requests to access this dataset, managed by Banco de Portugal, were granted to judicial clerks and enforcement officers, in order to identify the financial institutions legally authorized to receive deposits in which the debtor held bank accounts or deposits.

3. There have been applications to administrative cases in Italy (Vita 2012) and medical malpractice in the same country (Grembi and Garoupa 2013) and in the UK (Fenn and Rickman 2014), to compensation of damages in commercial relationships in Slovenia (Grajzl and Zajc 2017) and to cases focused on contracts in the construction sector in Belgium (Bielen *et al.* 2017).

estimation of the impact on duration of some procedural events, such as the availability of an expert's report or the holding of court hearings (Fenn and Rickman 2014, Grajzl and Zajc 2017 and Bielen *et al.* 2017). Another element covered by this type of studies is the characteristics of the different parties involved (Heise 2000). This may have a special impact for debt enforcement, as the resources put forward to resolve the case by companies vis-a-vis individuals would be distinct, and this could influence case duration.

In contrast to the analysis presented in this study, other papers that apply duration models to judicial cases are based on relatively small samples. The authors are not aware of any paper applying duration analysis to enforcement cases. Moreover, despite the importance of this area in the reform agenda for the judicial sector in the past years, particularly during the Economic and Financial Assistance Programme (European Commission 2014), quantitative analyses of the effectiveness of enforcement in Portugal are relatively scarce. In this respect, the analyses in Correia and Videira (2015, 2016) stressing the improvement of performance indicators in this area between 2011 and 2014 are noteworthy.

This paper is organized as follows. The first part describes the dataset, characterizing enforcement cases in Portugal, including type of litigant, enforceable title, court where cases were resolved, claim value and incidence of specific procedural aspects. The second and third sections focus on duration and profile of case resolution considering all cases, also the pending ones, overcoming the limitations of duration measures based only on resolved cases (a discussion of these shortcomings is made by Pereira and Wemans 2017). The fourth section addresses the role of enforcement cases' characteristics as determinants of duration. Finally, one makes some concluding remarks.

Main characteristics of enforcement cases related with the payment of civil and commercial debts

The dataset used in this paper includes 2,351,768 observations corresponding to all enforcements related to civil and commercial debts (including those regarding provision of services) which were dealt with in first instance judicial courts in Portugal between 2007 and 2016.[4] The dataset includes around 75% of all resolved enforcement cases, mainly excluding the ones related to the payment of judicial fees, fines, penalties and insurance premia.

4. Several observations were excluded from the dataset. Firstly, those with a duration shorter than one day, probably corresponding to case re-openings. Secondly, special enforcement cases regarding the payment of alimony. Thirdly, those for which the *comarca* where the case was resolved was not compatible with the territorial organization in place at that time. Finally, those with zero claim value and with missing information on the enforceable title. These exclusions totalled 10,890 observations.

As duration analysis is the focus of this paper, dates when the cases were filed and resolved are the key variables. The dataset includes 661,898 cases that started before 2007. However, these are only a fraction of the cases brought in those years (the ones not resolved until the end of 2006) and consequently the dataset is left-truncated. In addition, there are 616,073 cases without a resolution date: those pending at the end of 2016 (right-censored observations).

Information about the plaintiff is available for around 80% of resolved cases, but only for around 17% of cases pending at the end of 2016. Companies filed the vast majority of enforcement cases (87%), followed by individuals (12%). Finally, public entities represent a residual share (1%) as enforcement cases brought by the state are generally dealt with by tax and administrative courts. Available information also allows an identification of mass litigants[5] for around 70% of resolved cases and 15% of those pending. Considering resolved cases, the weight of mass litigants between 2007 and 2016 is approximately 22% but this percentage has been falling since 2011 (Figure 1A). As regards the economic activity sector, information is available for only 43% of all cases brought by companies. In this subsample, the most important sectors are financial and insurance services as well as information and communication services, each corresponding to around 30% of cases. Also relevant are the commercial sector (13%), manufacturing (7%) and real estate (5%).[6]

Moreover, the dataset includes information about the *comarca* where the case was resolved or pending at the end of 2016[7] and the cases resolved in judgeships specialised in enforcement cases - *juízos de execução*. There is a clear concentration of enforcement cases in the two major *comarcas* (Lisboa and Porto), although this has dropped significantly up to 2012, recovering in the following years, particularly in 2015, the first complete year after the implementation of the new judicial map which increased the territorial scope of these *comarcas*. As regards specialisation, between 2007 and 2009 only around 15% of cases were resolved in enforcement judgeships. This

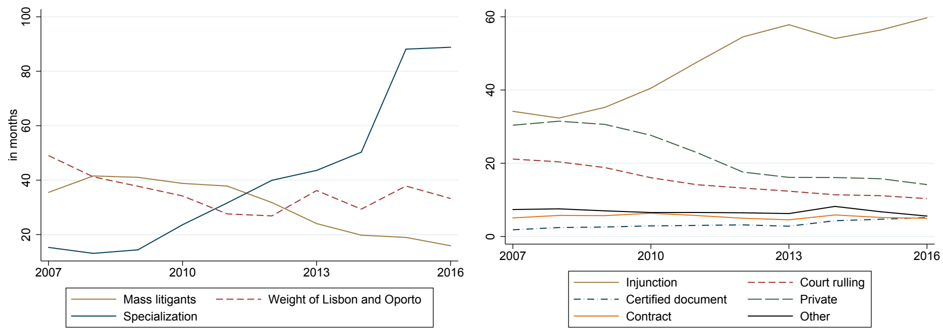
5. Companies that filed more than 200 cases, protective measures, proceedings or enforcements each year and which have to pay higher judicial fees, according to the *Regulamento das Custas Processuais* and to *Portaria* no 200/2011 of May 20. It is important to note that this variable, along with the one regarding procedural aspects which will be presented ahead, were especially compiled for this study and therefore were not subject to the same consolidation and validation procedures performed for the other variables that are included in official statistics.

6. Taking into account that the activity sector is only available for around half of the cases filed by companies, the actual percentages can differ substantially from those presented here.

7. A small fraction of cases do not have such information about the *comarca* (7,127), because they were dealt with in courts which cover several *comarcas* as labour and family courts.

proportion increased steadily up to 2014 and it jumped to 90% in the last two sample years, with the implementation of the new judicial map (Figure 1A).[8]

Regarding case characteristics, there is information about the type of enforceable title and, for 98% of the observations, claim value. Among enforceable titles[9], the injunction procedure is the most common, underlying around 30% of resolved cases in the beginning of the sample period and almost 60% in the more recent years (Figure 1B). The increasing role of this procedure should be related with legislative changes that broadened its scope, as discussed in Pereira and Wemans (2015), and was accompanied by a reduction in the relevance of private documents (from 30% to 14%) and court sentences (from 20% to 10%). Authenticated documents, contracts and other titles cover around 5% of cases each (Figure 1B).



(A) Characteristics of enforcement cases (B) Enforceable title

FIGURE 1: Main trends in enforcement by resolution year (%)

Notes: Note that the territorial scope of Lisboa and Porto *Comarcas* was enlarged in September 2014, with the implementation of the new judicial map.
Sources: DGPJ and authors' calculations.

Average claim value, at constant 2011 prices (private consumer deflator), is 21 thousand euros and its distribution is highly asymmetric - around two thirds of enforcement cases have a claim value lower than 5 thousand euros (Figure 2). Average claim value is considerably higher for companies in the

8. One could identify resolved cases in enforcement judgeships in the *comarcas* of Lisboa, Porto, Guimarães and Oeiras already in January 2007 and Vila Nova de Gaia from May of that year onwards. With the implementation of *comarcas-piloto*, one could find enforcement cases resolved in this type of judgeships also in Grande Lisboa-Noroeste and Baixo Vouga. Finally, with the implementation of the new judicial map, in 2014, 16 out of 23 *comarcas* have enforcement judgeships.

9. Different enforceable titles were aggregated in relatively homogeneous categories with a similar treatment by the judicial system. Appendix A in the Portuguese version of this paper details the construction of these categories.

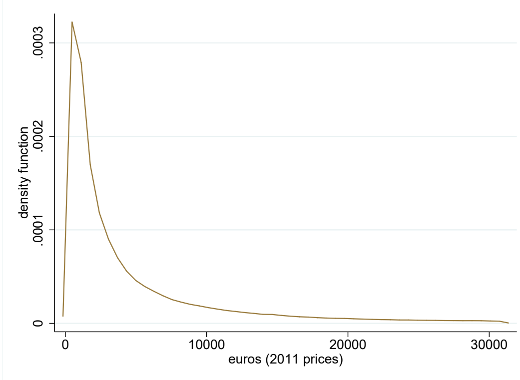


FIGURE 2: Distribution of claim value, excluding the upper decile

Sources: DGPJ and authors' calculations.

financial sector (around 40 thousand euros) and lower for companies in the information and communication sector (2 thousand euros).

Procedural aspects (*apensos*) were identified in around 5% of the observations.[10] The most common type of *apensos*, identified in around 3% of enforcement cases, are oppositions by the defendant and *embargos de terceiro* by other creditors.[11] In addition, around 2% of the cases have creditors' claims[12] and 0.4% have other types of *apensos*. Note that resolved cases with *apensos* have, on average, a considerably high claim value (74.8 thousand euros).

Finally, there is also information about the type of closure, spread over 39 categories, several of those with a very residual importance in the case of enforcement. Therefore, these categories were aggregated into four groups, with the aim of measuring the importance of successful enforcement cases - cases that ended in the fulfilment of the obligation - and of the different reasons the remainder cases did not succeed. Indeed, only around 36% of

10. For cases pending at the end of 2016, the identification of procedural aspects was made including information available up to 15 January 2018, which can lead to an underestimation of their incidence, as *apensos* can still be filed for such cases. However, in the following sections only *apensos* filed until the end of 2016 were considered, as only those might have a direct influence on observed duration. Moreover, only *apensos* recorded in the same court as the related enforcement case were considered, a procedure that may also lead to an underestimation of their incidence, particularly for cases resolved in the months that followed the implementation of the new judicial map, in September 2014.

11. *Embargos de terceiro* can be filed by any asset holder whose ownership of the asset is not compatible with the scope or the implementation of the seizure (see Gomes 2007, pp. 74).

12. This procedure allows other creditors with secured claims over the seized assets to intervene in the case, even if their debt is not overdue or they do not yet have an enforceable title (Leiras 2015).

enforcement cases ended with the fulfilment of the obligation,[13] 29% were closed due to a lack of assets, while 27% ended by reasons attributable to the plaintiff, as withdrawals or inactivity. Note that around 9% of the observations do not fit with in the mentioned categories or have an unspecified type of closure (appendix B in the Portuguese version of this paper presents a list of the types of closure included in each category). In this respect, not only the speed at which the system promotes the payment of debts, but also the rules governing the closure of cases may have a significant impact on the duration of enforcement cases.

Duration of enforcement cases

Statistics available concerning case duration in Portugal focus on the duration of the cases resolved in a given year, an indicator that does not correctly express what happened in that year, as most of the cases started in previous years. However, using information about the dates the cases were filed and resolved, it is possible to use duration analysis (see, for instance, Kiefer 1988 and, for the implementation in *Stata*, Cleves *et al.* 2010) to obtain the survival function of enforcement cases, which depicts how the probability of a case remaining pending evolves as time since case filing goes by.[14] In this approach, the function is estimated taking into account all the information regarding each time period, namely all pending cases in the beginning of the period and new cases filed, even if they remain pending at the end of the period.

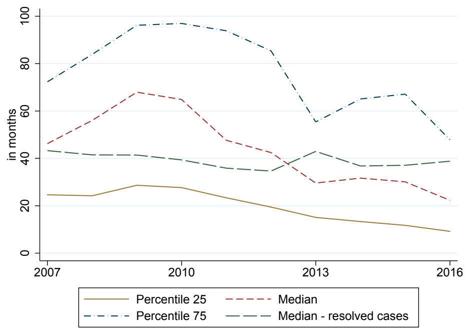
The annual median case duration - the time until which half of cases are expected to be resolved - estimated according to these two indicators presents a very different pattern in the period under analysis (Figure 3A). Whereas the median duration of resolved cases remained stable around 40 months, the median estimated from the survival function increased significantly between 2007 and 2009, presenting thereafter a decreasing trend, with particularly significant reductions in 2011, 2013 and 2016. In 2013 the decline was very steep in longer durations, while the time needed to resolve 25% of the cases had a constant decrease since 2010, from 25 to only 9 months. This reduction in 2013 may be related with the following measures - see decree-law no. 4/2013 of January 11 (Correia and Videira 2015). Firstly, special units were established to reduce case backlog in the courts of Lisboa, Porto, Maia, Vila Nova de Gaia and Oeiras. Secondly, new measures were introduced to facilitate the closure of cases due to inactivity of the plaintiff. Finally, for cases filed before

13. This type of closure includes not only situations in which effective payment existed, but also in which fulfilment of the obligation is expected, such as when payments occur in instalments.

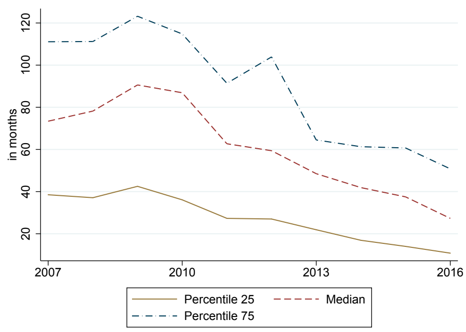
14. Formally, the survival function is given by $S(t) = Pr(T > t)$, T being the random variable which represents the time elapsed until the case is resolved.

September 2003, the closure resulting from a lack of seizable assets was also simplified.

As mentioned, significant changes were introduced to the rules governing the closure of cases due to inactivity of the plaintiff. Consequently, it is important to analyse whether the reduction in duration was actually accomplished by a quicker response of the judicial system or simply reflected a swifter closure of cases without giving a proper response to the plaintiff, namely the fulfilment of the obligation or the recognition that it could not be fulfilled due to a lack of sufficient assets. An analysis of the duration of enforcement cases that only considers as resolved the ones ending in the fulfilment or non-fulfilment of the obligation, and in which the cases ending by reasons attributable to the plaintiff are treated as censored[15], also points to a strong decline in average duration between 2010 and 2015 (Figure 3B). As will be discussed ahead, after controlling for the impact of the change in observable characteristics of enforcement cases, evidence of shorter durations for the most recent period remains.



(A) All types of closure



(B) Type of closure: effective resolution

FIGURE 3: Developments in the duration of enforcement cases 2007-2016, comparing different measures

Note: Percentiles driven from survival functions estimated with complete information about the enforcement cases dealt with by the system in each year. For comparison, figure A also shows the median of resolved cases in each year.
Sources: DGPJ and authors' calculations.

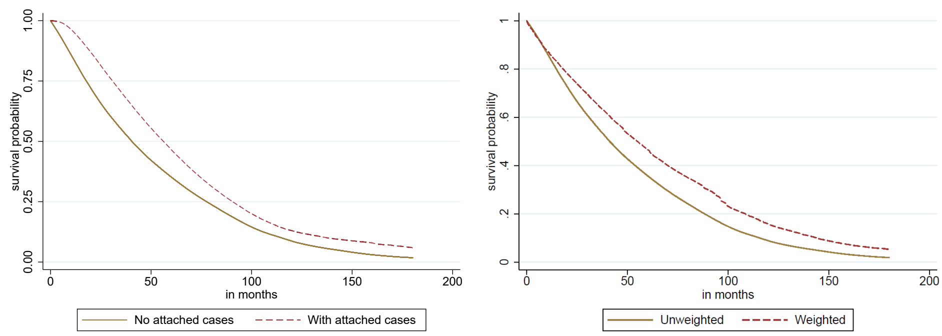
The decline in the duration of cases might be associated with changes in the judicial system, either through a streamlining of procedures, availability of new resources or the implementation of new rules for case closure. However

15. This analysis excludes resolved cases for which the type of closure is not identified. The closure of cases due to reasons attributable to the plaintiff should reflect a cost-benefit analysis, taking into account that the system was not able to respond to the claim up to that time. As this event prevented case closure by the system, these observations can be regarded as censored.

it can also relate to a change in the quantity or the average complexity of cases filed. In this respect, it is important to understand how duration varies with case characteristics.

As regards the enforceable title, there is no clear distinction among survival functions estimated for injunctions, court sentences, private documents or other titles. However, duration is much longer for cases with authenticated documents as the underlying enforceable title - which also have a very high average claim value (134 thousand euros) - and for enforcement cases based on contracts. The incidence of *apensos* attached to the case considerably increases duration (Figure 4A), as expected given that the case can only be resolved after the resolution of these *apensos*.

Cases featuring higher debts take, on average, longer to be solved, as illustrated by the clear distinction between survival functions weighted by claim value and unweighted (Figure 4B). This result can be driven by a positive association between complexity and claim value as these cases may demand the fulfilment of a higher number of procedural steps. In addition, the connection between procedural rules and claim value, as well as the relationship between the amounts at issue and the resources allocated by the defendant to prevent enforcement could explain such result.



(A) Procedural aspects

(B) Weighted by claim value

FIGURE 4: Survival functions

Note: The figures depict Kaplan-Meier survival functions for durations up to 15 years that cover the time span of the vast majority of cases. The dashed line in figure 4B is weighted by claim value in real terms.
Sources: DGPJ and authors' calculations.

Regarding the characteristics of the litigant and the type of closure, it is only feasible to look at the duration of resolved cases.[16] By type of litigant,

16. As information regarding pending cases on these variables is virtually unavailable, survival functions would be clearly biased to shorter durations.

case resolution is, on average, faster for individuals, lasting 45 months, which compares to 49 months for companies. By activity sector, financial and insurance companies present higher durations (61 months), while information and communications, trade and manufacturing sectors have a duration close to that for all companies and real estate activities have a significantly lower average duration (32 months). The enforcement cases filed by mass litigants took on average 6 months less than those from other companies. Regarding the type of closure, the fastest cases are those that end with the fulfilment of the obligation (39 months), followed by those that end because of insufficient assets (48 months) and finally those closed by reasons attributable to the plaintiff (60 months).

Probabilistic profile of resolution of enforcement cases

The time profile of case resolution can be studied through the hazard function which is an alternative way to characterize the duration distribution along with the survival function presented above. The hazard function measures the instantaneous rate of resolution of a pending case, in probabilistic terms, reported to a given moment after start.[17] Figure 5 shows this function for the entire period under analysis. As described in the introduction, recent years have been marked by profound changes, both in the enforcement procedure with the entry into force of the 2013 Civil Procedure Code and the availability of electronic platforms to support the activity of enforcement officers, and organization with the implementation of the new judicial map in the 2014 judicial year.[18]

Figures 6A and 6B show the hazard functions, respectively, for the periods before and after the beginning of 2014 judicial year. Taking the latter as the break point of the sample, the second figure will capture approximately the current profile of case resolution's probabilistic intensity, already reflecting the various changes that have been affecting enforcement procedure. The figures also show the time elapsed until the median (red line) and the 90th percentile (blue line) of case resolution are reached, giving an indication of duration intervals during which most cases are resolved.

17. More specifically, the hazard function measures the probability of case resolution to occur during an infinitesimal interval of time, conditional on the case having remained pending until that moment, divided by the amplitude of the interval. The hazard function ($h(t)$) relates to the survival function through the expression $h(t) = -dlnS(t)/dt$.

18. In this article references to judicial years always take the period from 1 September to 31 August of the following year and not the official judicial year which sometimes coincided with the calendar year. In fact, the main policy measures took effect after the end of the judicial holidays, not at the beginning of the year, signalling that period as the most relevant reference for the Portuguese judicial system.

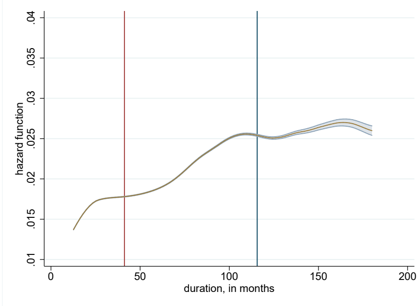
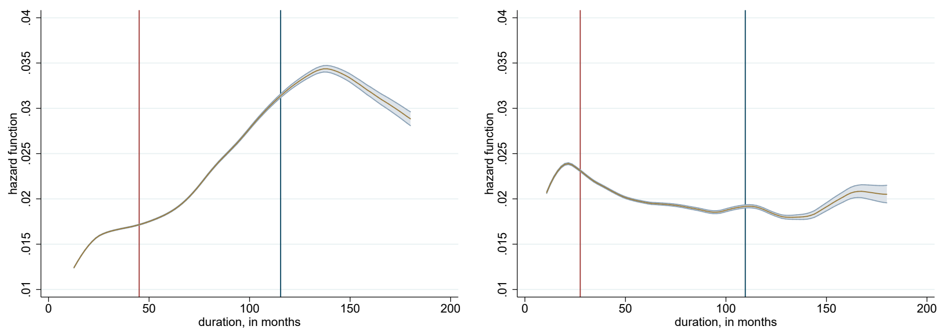


FIGURE 5: Instantaneous rate (probabilistic) of case resolution, full sample

Note: The figure shows the hazard function and confidence bands at 95%, for durations up to 15 years that cover the time span of the vast majority of cases. The red line is the median of case duration and the blue line the 90th percentile.
Sources: DGPJ and authors' calculations.



(A) Before the 2014 judicial year (B) 2014 judicial year and following

FIGURE 6: Instantaneous rate (probabilistic) of case resolution

Note: The figure shows the hazard function and confidence bands at 95%, for durations up to 15 years that cover the time span of the vast majority of cases. The red line is the median of case duration and the blue line the 90th percentile.
Sources: DGPJ and authors' calculations.

When the entire period under study is taken, the hazard function indicates positive duration dependence (Figure 5), i.e. the instantaneous rate of resolution of pending cases increases over time (except for very long durations, when a very small number of unresolved cases remains). However, this result is essentially determined by the profile in the first sample period (Figure 6A). In the sample beginning in the 2014 judicial year (Figure 6B), the resolution intensity does not have a marked trend, varying within a much more limited range. In addition, such intensity is higher in this second period, over a span that extends from the start of cases up to about 5 years of duration. This is the reason the median of resolved cases is reached much earlier.

In the pre-2003 reform regime, the intervention of judges in all enforcement cases would be a reason for an increasing profile of the hazard rate, on the assumption of a natural prioritization of older cases. Although the first sample starts already in 2007, the system still had a significant number of pending cases at the time, both pre-reform and entered in the changeover to the new regime. It is also to be expected that the profile estimated in figure 6A reflects the adjustment of the system to the new rules, notably with respect to the formation of a body of enforcement officers capable of responding to the volume of incoming cases.

The probabilistic profile of case resolution also depends on the several steps related to the seizure of debtor's assets. Thus, the greater intensity of resolution in the first years of the case and the absence of an upward trend in the second sample will also reflect the gains of speed in the procedures for identification and seizure of assets, including the impact of the effectiveness of these mechanisms in promoting voluntary compliance at an early stage.

Determinants of duration of enforcement cases

Methodology

The impact of the different explanatory variables on duration is studied on the basis of a semiparametric model, the Cox (1972) model, rather usual in this context (see, for example, Cameron and Trivedi 2005). The Cox model assumes that the hazard function of a case associated with the set of explanatory variables $\mathbf{x}_i$, $h(t|\mathbf{x}_i)$, is given by

$$h(t|\mathbf{x}_i) = h_0(t)exp(\beta_1 x_{1,i} + \beta_2 x_{2,i} + ... + \beta_k x_{k,i}), \tag{1}$$

where $h_0(t)$ is the baseline hazard function and $exp(\mathbf{x}\boldsymbol{\beta})$ the relative hazard. There is a proportional relationship between the hazard functions of any two cases $\mathbf{x}_j$ and $\mathbf{x}_i$, the proportionality ratio being given by $exp(\mathbf{x}_j\boldsymbol{\beta})/exp(\mathbf{x}_i\boldsymbol{\beta})$. In particular, if these cases differ only up to a characteristic expressed through a binary variable (which takes on, say, the value 1 if the case is filed by a company and the value 0 if the case is filed by an individual), the multiplicative constant relating the hazard functions is given by the exponential of the coefficient of that variable. The main advantage of the Cox model is that it does not require modelling the hazard function (although this can be estimated). Nevertheless, the Cox model assumes, in its simplest form, that all the regressors move this same function in a multiplicative way, that is, it assumes the proportionality of hazards.

The proportionality hypothesis can, however, be partially lifted through a stratified estimation procedure, in which it is assumed that the baseline hazard functions differ across the strata (1, 2, ...) corresponding to the values of one or more categorical variables for which one does not want to assume proportionality of hazards. Thus, there will be several hazard functions given by

$$\begin{aligned} h_1(t|\mathbf{x}_i) &= h_{01}(t)exp(\beta_1 x_{1,i} + \beta_2 x_{2,i} + ... + \beta_k x_{k,i}) \\ h_2(t|\mathbf{x}_i) &= h_{02}(t)exp(\beta_1 x_{1,i} + \beta_2 x_{2,i} + ... + \beta_k x_{k,i}) \\ &..., \end{aligned} \tag{2}$$

in which the coefficients of the explanatory variables not used for stratification are common to all strata.

In addition to the Cox model, the Weibull and Gompertz parametric models were estimated, both assuming the aforementioned proportional hazards hypothesis. These models are compatible with the increasing profile of the non-parametric risk function for the entire sample period (Figure 5). However, results (available upon request) are similar to those presented below for the Cox model (as in Vita 2012 and Bielen *et al.* 2016), so a separate analysis is not justified.

Variables

The explanatory variables cover the following characteristics of cases: case filed by a company versus individual, mass litigant status (companies), claim value in real terms, enforceable title underlying the case and procedural aspects (existence of *apensos*) - see section describing the characteristics of debt collection cases, for more details. The claim value has a strongly asymmetric distribution to the right, and was taken in logs. The nature of the plaintiff and the mass litigant status have a significant number of missing observations, particularly affecting pending cases (i.e. the censored observations), while the claim value has a residual number of missings. The missing observations in these variables were imputed through a multiple imputation procedure.[19]

In the database there is information about the *comarca* where the lawsuit ended in accordance with the configuration of the judicial map in place at the time. However, the territorial organization of the judicial system has

19. In the implementation of this procedure the standard routines available in Stata were used, the nature of the plaintiff and mass litigant status having been imputed on the basis of a logistic regression, and the claim value on the basis of a linear regression, on the various regressors, the failure indicator and the baseline cumulative hazard (see White and Royston 2009). The binary variables have, respectively, about 35% and 45% of missings, and the claim value about 2%. It was decided not to impute the activity sector of the plaintiff company, also available in the database, given the existence of a very high proportion of missings.

changed over the period covered by the study. In addition to the most recent revision, there was the establishment of *comarcas-piloto* in 2009. The current configuration was taken as a reference and extended to the whole horizon, through an exercise of territorial aggregation. The *comarca* of case resolution - thus determined - was used to stratify the sample, given that it did not seem appropriate to impose proportional hazards throughout all *comarcas* (by including fixed effects in the Cox regression). This procedure allows a more flexible modelling, while ensuring that the coefficients of the remaining variables are not affected by their correlation with the characteristics of *comarcas*. However, it has the disadvantage of not allowing to directly estimate the impacts of these same characteristics on the duration of cases.

Complementary specifications were estimated, including *comarca* variables, now retaining the actual geographical configuration of *comarcas* at the end of the lawsuit. In particular, one intended to study the impact of congestion on duration of cases. It is not possible to calculate a direct congestion measure for each *comarca* in a given year, considering completed and pending lawsuits, as there is only information about the *comarca* where the case ended and the judicial map underwent significant changes in the sample period. Alternatively, one considered the average duration of «peer» lawsuits, i.e. those that ended in the same *comarca* and year as the lawsuit concerned. A measure of the complexity of litigation was included as well, taking the log of the average claim value in the *comarca*. Other *comarca* indicators outside the judicial system were calculated, namely based on INE's Purchasing Power Index and the density of businesses (from data available at the municipality level), but these showed very high correlation with «peer» lawsuit duration and were not used.

The cases starting in a given year share common determinants, whether arising from the judicial system itself, such as changes in applicable legislation and procedures, whether outside the system, as changes to the type of transactions in the economy, or variations in non-performing credit rate due to the business cycle. In the benchmark econometric specification, the judicial year the case begins has also been taken as a stratification variable, as the graphical tests were unfavourable to the proportional hazards hypothesis. However, one also experimented with a fixed-effects type modelling, with almost no impact on the magnitudes and significance of other variables' coefficients. Regarding the remaining variables, graphical tests indicated that the proporcionality assumption approximately held, except for the existence of *apensos*. In fact, the inclusion of this last information into the regression should be seen as an approximation: an adequate modelling would require a time-varying model, in order to take into account the moment the *apensos* start.

As a modelling hypothesis one interacted the impact of some enforceable titles on duration with the claim value. In fact, the civil procedure law provides for a speedier enforcement procedure for titles such as court sentences and injunctions, and extrajudicial titles fulfilling certain requirements, in particular a claim value under a certain threshold (notwithstanding the various changes to those requirements over time). In the period covered by this study, are particularly relevant the regimes established by the 2003 and 2013 reforms which, as a general rule, do not require in such cases an order by the judge nor the delivery of notification to the defendant (see, for example, Passos 2012 and Pinto 2013).[20] There are a small number of older lawsuits in the database - about 10% - which began almost entirely after the entry into force of decree-law no. 274/97 of 8 October. In this period, for judicial titles and other titles with a claim value under a certain threshold (cumulatively with other conditions), the lawsuit began directly with the seizure of assets, without delivery of notification to the defendant, but following an order by the judge.[21] The need for an order by the judge in the period prior to 2003, either for delivery of notification to the defendant or seizure of the assets indicated by the plaintiff, was common to all titles and should in principle be captured by the indicator for the year the lawsuit starts.

Continuous variables, i.e. the claim value and the congestion and litigation complexity measures for *comarcas* - were centred on the median, so that the baseline hazard function was estimated by reference to that figure. In fact in the Cox model, this function is estimated at the origin of the covariates - the figure that makes the relative hazard equal to 1 in the expression (1).

Results for all comarcas

Table 1 presents the results for all enforcement cases in the period 2007-2016, stratifying by judicial year of case filing and *comarca*, as explained above, and also splitting up the sample into the period before and after the 2014 judicial year, in line with the section on case resolution profile. Two additional specifications have been estimated: including *comarca* variables (dropping the stratification by *comarca*), and interacting some titles with the claim value.

Lawsuits initiated by mass litigants within those filed by companies tend to proceed more rapidly (probabilistic finalization intensity about 40% higher), and the same holds for cases brought by individuals relative to companies (with a less marked effect: approximately 20% shift in resolution intensity). The first result may stem from mass litigants' allocating significant resources to this type of litigation, possibly resorting to experienced

20. The current threshold is 10 thousand euros, twice the minimum allowing appeal to second instance courts, as set out by Law no. 62/2013 of August 26; between 2003 and 2013, such threshold was 30 thousand euros, the minimum allowing appeal to the Supreme Court, as set out by decree-law no. 303/2007 of August 24, and previously, an amount equivalent to 14,963.94 euros, as set out by Law no. 3/99 of January 13.

21. The relevant threshold was then the minimum allowing appeal to second instance courts, set out at an amount equivalent to 3740.99 euros by Law no. 3/99 of January 13, and previously to 2493,99 euros by decree-law no. 38/87 of December 23.

		Full sample Interaction value x title	**Comarca variables**	**Period before new map**	**Period after new map**
Plaintiff (Company)					
Individual	1.21***	1.21***	1.10***	1.26***	1.12***
	0.06	*0.06*	*0.04*	*0.09*	*0.01*
Company, mass litigant	1.39***	1.39***	1.24***	1.49***	1.12***
	0.05	*0.05*	*0.03*	*0.11*	*0.01*
Claim value	*0.86****	*0.86****	0.90***	*0.87****	*0.86****
	0.01	*0.01*	*0.01*	*0.01*	*0.00*
Enforceable title (Injunction)					
Court sentence	1.02	1.02	1.04	1.08**	0.85***
	0.05	*0.05*	*0.06*	*0.04*	*0.04*
Authentic document	1.01	1.03	1.03	1.03	1.00
	0.04	*0.05*	*0.04*	*0.07*	*0.03*
Contract	0.90***	0.86***	0.96**	0.89***	0.93***
	0.02	*0.02*	*0.02*	*0.02*	*0.02*
Private document	0.92**	0.93**	1.01	0.93**	0.90**
	0.03	*0.03*	*0.05*	*0.03*	*0.04*
Another title	0.99	0.97	1.02	1.02	0.93**
	0.02	*0.02*	*0.03*	*0.02*	*0.03*
Enf. title x value threshold					
Contract, value > threshold		1.22***			
		0.07			
Private doc., value > threshold		0.97			
		0.02			
Another title, value > threshold		1.13***			
		0.04			
Apensos					
Creditors' claims	0.69***	0.68***	0.64***	0.83***	0.36***
	0.02	*0.02*	*0.02*	*0.03*	*0.01*
Embargos/oppositions	0.90***	0.90***	0.82***	1.02	0.65***
	0.04	*0.04*	*0.03*	*0.05*	*0.05*
Other	0.80***	0.80***	0.84***	0.85***	0.59***
	0.03	*0.03*	*0.03*	*0.03*	*0.03*
Two or more *apensos*	1.03	1.03	1.01	1.04	0.89***
	0.02	*0.02*	*0.02*	*0.03*	*0.04*
Comarca variables					
Claim value *comarca*			0.08***		
			0.03		
Duration «peer» cases			0.96***		
			0.00		
Strat. by judicial year	yes	yes	yes	yes	yes
Strat. by *comarca*	yes	yes	no	yes	yes
No. observations	2 351 768	2 351 768	2 351 768	2 061 289	1 051 168

TABLE 1. Determinants of duration of enforcement cases

Notes: Hazard ratios estimated by the Cox regression. Regressions also include a binary variable for special enforcement cases. Continuous variables were centred at the median. Robust standard errors, clustering on *comarcas*, in italics. P-values: *<0.1; **<0.05; ***<0.01.

enforcement officers. The second result is surprising, as it would be expected that companies, even those that are not frequent litigants, would be more familiar with debt enforcement. The variable may be thus capturing specificities of cases filed by individuals, not captured by other variables in the model. Both coefficients show a reduction in magnitude from the first to the second subsample, but no conclusions can be drawn in this respect given that estimates in the second subsample are based on a very high proportion of missings for these two variables. Even for the whole sample, the estimated magnitudes should be read with some caution, taking into account the weight of imputed observations.

A higher claim value tends to lengthen enforcement cases, reflecting the specificities of the litigation associated with higher amounts not captured by other variables in the model, such as an enhanced complexity and a greater opposition on the debtor side. An increase of one standard error in this variable, at the median,[22] leads to a reduction in resolution intensity of about 20%. This magnitude also holds in each of the subperiods, before and after the implementation of the new judicial map.

Regarding the impact of the different enforceable titles on duration, measured vis-a-vis injunctions, cases based on private documents and contracts tend to be more time-consuming, a result common to both subsamples. Such a result could reflect the less swift enforcement procedure applicable to some of these titles relative to injunctions, in particular where the former assume a value above a certain threshold (as described above). However, when this is modelled by interacting the title with an indicator for a claim value above the threshold that determined different procedural rules at each moment, such an interaction is not significant or indicates a shorter duration, contrarily to what would be expected. Thus, other aspects associated with the title seem to explain the result. Court sentences - whose procedure has been broadly analogous to injunctions - tended to be comparatively faster in the period prior to the new judicial map and slower thereafter.

Procedural aspects, in the form of *apensos*, have a negative impact on the speed of proceedings, as it might be expected. This result is particularly pronounced for creditors' claims, where a case resolution rate around 30% lower is estimated. There are, however, interesting differences between the two subsamples considered. In the period prior to the 2014 judicial year, the impact of procedural aspects on the duration of enforcement lawsuits appears comparatively diluted, and the existence of *embargos*/oppositions and two or more *apensos* are not statistically significant. The greatest impact of procedural aspects in the second subsample could be due to the latter's already reflecting clearly the post-2003 regime of enforcement procedure as regards the lack of intervention of the judge in many cases. When there are *apensos* that require that intervention, making the case resolution depend on the end of a declarative action, this clearly results in an extended duration.

22. Corresponding to a change in the claim value from approximately 2500 to 12500 euros.

The coefficients of *comarca* variables, proxies of congestion and complexity of litigation, have the expected signs, as an increase in each of these covariates brings about a longer duration of enforcement cases. The estimated magnitudes translate into reductions in the probabilistic resolution intensity, respectively, of about 30% and 60%, for increases of one standard error in the variables, at the median.[23] These impacts are substantial, in general even higher than those of case characteristics, confirming that their duration depends heavily on the overall litigation directed to the *comarca*. Another interesting aspect is that the impacts of these variables on duration decrease greatly in the most recent subsample (coefficients not shown), with the congestion measure losing statistical significance, and an equivalent increase in the mean claim value bringing about a fall in resolution intensity of only about 10%.

Duration holding case characteristics constant

One may take up again the evolution of enforcement cases' duration, now considering the baseline survival function estimated by the Cox regression (which corresponds to the baseline hazard function, described in the section on methodology). This allows an analysis of such an evolution on a more comparable basis, in that the explanatory variables are held fixed, at the medians for continuous variables and omitted groups for binary variables (one compares cases based on injunctions without *apensos*, brought by companies that were not mass litigants). In line with the remainder of this section, separate regressions are run for the period before and after the 2014 judicial year, implying that the impacts of explanatory variables are also allowed to vary. These regressions were not stratified by *comarca*, nor by judicial year of start, because it was intended that the survival functions concerned the judicial system and the years in each subsample as a whole. Figure 7 confirms the evidence of a shorter duration in the most recent sample, with the medians resulting from the baseline survival functions being approximately equal to 55 months in the first period and 25 months in the second one. A related analysis consisting in modelling the judicial year of start as fixed-effects in the Cox regression (see the discussion in the section on explanatory variables) indicates a greater speed of the lawsuits entered in the most recent years. In general terms, these results should be explained by several factors, namely the reforms implemented over time, changes to the unobservable characteristics of cases and factors outside the judicial system, such as changes in the cyclical position of the economy. Identifying the contribution of each of these factors goes beyond the scope of this work.

23. Corresponding to the variation of duration of peer cases in the same *comarca* from 44 to 57 months and average claim value from approximately 20 thousand to 30 thousand euros.

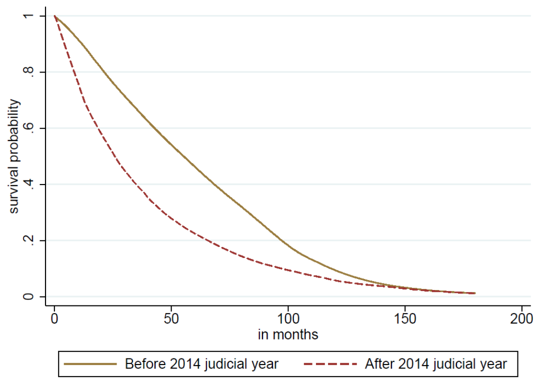


FIGURE 7: Survival functions

Note: Baseline survival functions estimated from the Cox regression, model with *comarca* variables and without stratification (for durations up to 15 years that cover the time span of the vast majority of cases).
Sources: DGPJ and authors' calculations.

Results considering territorial breakdowns

Table 2 presents the results for the Cox model splitting up the sample into the cases that ended, or were pending at the end of 2016, in enforcement judgeships and the remaining ones. Enforcement judgeships were initially created in highly congested *comarcas* and their number increased considerably with the implementation of the new judicial map in 2014 (Figure 1A). Results are also presented for the four largest *comarcas* in the sample in absolute terms, i.e. Lisboa and Porto as existing until the 2013 judicial year, and again Lisboa and Porto, as defined in the new judicial map, afterwards.

The estimated impacts for enforcement cases brought by mass litigants are smaller for both types of judgeships than the one for the sample as a whole (Table 1) that thus seems to reflect the interaction of the observations belonging to the two groups. Something similar occurs for the coefficient of lawsuits filed by individuals that loses statistical significance in the subsample of judgeships specialized in enforcement.

The effect of claim value on duration is, in each of the subsamples now considered, close to that for the judicial system as a whole. As far as the enforceable title is concerned, lawsuits based on contracts and private documents (vis-à-vis injunctions) remain slower, but for the latter titles this result is restricted to the enforcement judgeships. The coefficient of authentic documents now appears as statistically significant, indicating a shorter duration in judgeships not specialized in enforcement.

As regards procedural aspects, the negative impact on the speed of cases is relatively more pronounced in the enforcement judgeships, especially for

	Non-enforcement judgeships	Enforcement judgeships	Lisboa and Porto
Plaintiff (Company)			
Individual	1.08***	1.04	0.92*
	0.03	0.05	0.04
Company, mass litigant	1.22***	1.29***	1.20***
	0.02	0.06	0.03
Claim value	0.91***	0.86***	0.83***
	0.01	*0.02*	*0.03*
Enforceable title (Injunction)			
Court sentence	0.98	0.89**	1.10
	0.07	*0.05*	*0.10*
Authentic document	1.13***	0.91	0.80*
	0.03	*0.07*	*0.11*
Contract	0.96**	0.83***	0.86***
	0.02	*0.03*	*0.04*
Private document	0.99	0.79***	0.91
	0.02	*0.05*	*0.07*
Another title	1.04	0.88***	0.93
	0.03	*0.03*	*0.09*
Apensos			
Creditors' claims	0.71***	0.51***	0.54***
	0.02	*0.05*	*0.11*
Embargos/oppositions	0.85***	0.82**	0.99
	0.02	*0.08*	*0.12*
Other	0.89***	0.76***	0.76*
	0.02	*0.06*	*0.12*
Two or more *apensos*	0.96***	1.08*	1.15*
	0.01	*0.04*	*0.08*
Strat. by judicial year	yes	yes	yes
Strat. by *comarca*	yes	yes	yes
No. observations	994 421	1 357 347	853 421

TABLE 2. Determinants of duration of enforcement cases

Notes: Hazard ratios estimated by the Cox regression. Regressions also include a binary variable for special enforcement cases. Continuous variables were centred at the median. Robust standard errors, clustering on *comarcas*, in italics. P-values: *<0.1; **<0.05; ***<0.01.

credit claims. This suggests an interaction of the effects of such claims on duration with differentiating characteristics of the enforcement lawsuits that tend to be dealt with in specialized judgeships. Finally, the coefficients of *comarca* variables (not shown) have a magnitude similar to that presented in Table 1 for the congestion measure, in both types of judgeships. The measure of complexity of litigation has, however, a much stronger impact in the enforcement judgeships, indicating an enhanced role as a determinant of duration.

Conclusions

Changes to debt enforcement in Portugal have been very significant in recent decades. This paper presents alternative measures of duration of enforcement cases that indicate its considerable reduction in the recent period, contributing to the reassessment of this factor as a framework cost of economic activity in Portugal. The analysis of the probabilistic profile of case resolution in two different periods, before and after the 2014 judicial year, shows that there has been a structural change, with a much greater probability that the lawsuit is currently resolved in the first months after its start.

As mentioned, this improvement will certainly have been influenced by the very significant set of reforms in the area of debt enforcement implemented over time, although it cannot be ascribed to a specific policy measure. In addition, it will reflect the impact of other factors during the sample period, notably the evolution of the business cycle. The separation of the role of these factors is best done in a policy evaluation framework. The production of empirical evidence about the reforms introduced in this area could lead to improvements in other areas of justice in which performance indicators continue to disappoint. Likewise outside the scope of this paper are other considerations, such as an analysis of the impact of these reforms on the guarantees of defendants and other parties involved.

This article dealt only with one aspect of the effectiveness of debt enforcement. It would be interesting to address the evolution of the proportion of debt recovered, an indicator also influenced by factors outside the judicial system, such as the business cycle. This information is a relevant criterion for economic agents in the decision to bring an enforcement case to court. It would also be interesting to examine the impact of a faster resolution of cases on litigation costs. Continuing the research strand initiated by this study, one could explore the estimation of a duration model with time variation of some determinants, namely the moment of start of *apensos*, or a model of competing risks.[24] The identified risks would be, on the one hand, the effective resolution of the case by the judicial system and, on the other hand, the plaintiff dropping it (before the system has responded).

24. There has been some applications of this approach to court data that focus primarily on the distinction between the time elapsed until an out-of-court settlement is reached or the time elapsed until trial (Grajzl and Zajc 2017 and Bielen *et al.* 2017).

References

Araújo, Nuno (2009). "Primeiros passos das novas execuções." *Compilações Doutrinais Verbo Jurídico*.

Bielen, Samantha, Peter Grajzl, and Wim Marneffe (2017). "Understanding the Time to Court Case Resolution: A Competing Risks Analysis Using Belgian Data." *CESifo Working Paper Series*, 6450.

Bielen, Samantha, Peter Grajzl, Wim Marneffe, and Valentina Dimitrova-Grajzl (2016). "The duration of judicial deliberation: Evidence from Belgium." *CESifo Working Paper Series*, 5947.

Cameron, Colin and Pravin Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Christensen, Robert K. and John Szmer (2012). "Examining the Efficiency of the U.S. Courts of Appeals: Pathologies and Prescriptions." *International Review of Law and Economics*, 32(1), 30–37.

Cleves, Mario, Roberto Gutierrez, William Gould, and Yulia Marchenko (2010). *An Introduction to Survival Analisys using STATA, third edition*. Stata Press.

Correia, Pedro and Susana Videira (2015). "Troika's Portuguese ministry of justice experiment: an empirical study on the success story of the civil enforcement actions." *International Journal for Court Administration*, 7(1), 37–50.

Correia, Pedro and Susana Videira (2016). "Troika's Portuguese Ministry of Justice Experiment, Part II: Continued Positive Results for Civil Enforcement Actions in Troika's Aftermath." *International Journal for Court Administration*, 8(1), 20–31.

Cox, David Roxbee (1972). "Regression models and life-tables." *Jounal of the Royal Statistical Society*, B(34), 187–220.

de Freitas, J. Lebre (2014). "A ação executiva à luz do Código do Processo Civil de 2013." *Coimbra Editora*.

European Commission (2014). "The Economic Adjustment Programme for Portugal 2011-2014." *Occasional Papers*, 202.

Fenn, Paul and Neil Rickman (2014). "Information and the Disposition of Medical Malpractice Claims: A Competing Risk Analysis." *Journal of Law, Economics, and Organization*, 30(2), 244–274.

Gomes, Conceição (ed.) (2007). *A ação executiva em avaliação: Uma Proposta de Reforma*. Observatório Permanente da Justiça Portuguesa – Centro de Estudos Sociais da Universidade de Coimbra.

Grajzl, Peter and Katarina Zajc (2017). "Litigation and the Timing of Settlement: Evidence from Commercial Disputes." *European Journal of Law and Economics*, 44(2), 287–319.

Grembi, Veronica and Nuno Garoupa (2013). "Delays in Medical Malpractice Litigation in Civil Law Jurisdictions: Some Evidence from the Italian Court of Cassation." *Health Economics, Policy and Law*, 8(4), 423–452.

Heise, Michael (2000). "Justice delayed? An Empirical Analysis of Civil Case Disposition Time." *Case Western Reserve Law Review*, 50(4), 813–849.

Intrum Justitia (2017). "European Payment Report 2017." *Intrum Justitia AB (Sweden)*.

Kiefer, Nicholas (1988). "Economic Duration Data and Hazard Functions." *Journal of Economic Literature*, 26(2), 646–679.

Leiras, Diana (2015). "Determinação dos bens a penhorar: reflexões." *Revista Solicitadoria e Ação Executiva Estudos*, 3, 49–110.

Lourenço, Paula Meira (2017). "O processo executivo." *40 anos de políticas de justiça em Portugal*, pp. 227–249.

Passos, Márcia (2012). *A fase introdutória da ação executiva*. Tese de Mestrado - Universidade Portucalense.

Pereira, Manuel Coutinho and Lara Wemans (2015). "Determinants of civil litigation in Portugal." *Banco de Portugal - Economic Studies*, 1(1), 21–50.

Pereira, Manuel Coutinho and Lara Wemans (2017). "Productivity in civil justice in Portugal: a crucial issue in a congested system." *Banco de Portugal - Economic Studies*, 3(1), 11–39.

Pimenta, Paulo (2012). "Tópicos para a reforma do Processo Civil Português." *Julgar*, 17, 109–134.

Pinto, Rui (2013). "Notas breves sobre a reforma do Código de Processo Civil em matéria executiva." *Revista da Ordem dos Advogados*, 73(1).

Statistics Portugal (2015). *Custos de contexto: a perspetiva das empresas 2015*.

Vita, Giuseppe Di (2012). "Factors Determining the Duration of Legal Disputes: An Empirical Analysis with Micro Data." *Journal of Institutional and Theoretical Economics*, 168(4), 563–587.

Voigt, Stefan (2016). "Determinants of Judicial Efficiency: a Survey." *European Journal of Law and Economics*, 42, 183–208.

White, Ian and Patrick Royston (2009). "Imputing missing covariate values for the Cox model." *Statistics in Medicine*, 25(5), 1982–1998.

Is the Phillips curve dead? - results for Portugal

Sara Serra
Banco de Portugal

April 2018

Abstract
This article assesses the forecasting performance over recent years of Phillips curves for Portugal. Phillips curves are estimated for a large variety of slack measures and evaluated in terms of their out-of-sample performance in real time. The issues of time variation of parameters and possible non-linearities are also explored. The findings suggest that there is no unique best specification for the Phillips curve over time, and therefore a set of these models, considering different slack measures, should be considered. Furthermore, there is some evidence of flattening of the traditional linear Phillips curve in recent years, which is possibly related to non-linearities in the model. Overall, the Phillips curve maintains some forecasting power for inflation when compared to a naïve benchmark. (JEL: E31, E37)

Introduction

The Phillips curve (PC), introduced in 1958 by A. W. Phillips, postulates the existence of a negative relationship between unemployment and inflation, or of a positive relationship between output and inflation. Given the importance of the link between inflation and economic activity for monetary policy, it quickly become popular as an instrument of economic analysis. Over time, the PC has been subject to some criticism, with its standard formulation in the literature being adjusted accordingly.[1] Initially taken as a long term economic relationship, in the late 70's, with the work of Phelps and Friedman, it became seen as a short-term trade-off, dependent on inflation expectations.

More recently, both in Europe and the US, the Great Recession brought along the so called missing disinflation: inflation appears to have reacted less to the amount of slack in the economy than suggested by PC models

Acknowledgements: I thank the members of the LIFT team, particularly those from the ECB, whose codes were the starting point for the ones used in this article. I also thank José Ramos Maria, Miguel Gouveia and the participants of a seminar at Banco de Portugal for useful comments and opinions. The opinions expressed in the article are those of the author and do not necessarily coincide with those of Banco de Portugal or the Eurosystem. Any errors and omissions are the sole responsibility of the author.

E-mail: srserra@bportugal.pt

1. See Macklem (1997) for a summary of the evolution of the theory surrounding the Philips curve.

(Albuquerque and Baumann (2017) and ECB (2014)). The ensuing recovery, on the other hand, has shown a weaker increase in inflation than expected given the reduction in unemployment ("missing inflation") (Ciccarelli and Osbat (2017)). PC models were one of the standard instruments used to determine this counterfactual path for inflation, which has led to questions about their reliability in explaining price developments.

The standard Phillips curve formulation is often limited, being unable to capture many aspects not related to the amount of slack in the economy (administered prices, taxes, import prices), and affected by uncertainty regarding the relevant measure of slack to be considered. It is also potentially flawed, given possible non-linearities (namely due to downward price rigidity), time varying parameters and state dependent behaviour. As stated by Dotsey *et al.* (2017), the Phillips curve is likely to be unstable, being a reduced form model which is a function of deeper structural parameters that change over time.

Notwithstanding the standard Phillips curve limitations, namely for forecasting (ECB (2014)), it has remained a central instrument of analysis for central bankers.[2] Teles and Garcia (2016) analyse the possible usefulness of Philips curves for monetary policy definition in the current context.

Several authors have shown that refinements of this instrumental are able to reduce the recent puzzles surrounding inflation. Some of these refinements deal with non-linearities in the Phillips curve.

This article aims at analysing PC models for Portugal, drawing from the work developed for the Low Inflation Team (LIFT) (Ciccarelli and Osbat (2017)). The issue of parameter instability over time and its potential relationship with nonlinearities is explored, and the forecasting ability of several PC specifications is assessed. Given that the measurement of slack has several caveats, specially when resorting to output or unemployment gap measures, the forecasting performance of the PC is assessed in real time.

The structure of the article is as follows: the following sections present the baseline specification for the PC and the details on the variables considered in estimation and data transformations. Then the selection process of PC specifications is described and the possibility of time-varying parameters and non-linear effects is explored. Finally, the forecasting performance of selected PC specifications is analysed through a real time exercise.

Phillips curve baseline specification

The baseline specification takes the form of the hybrid Phillips curve equation considered in Albuquerque and Baumann (2017), which is given by:

2. See for example Draghi (2017), Constâncio (2015) or Yellen (2013).

$$\pi_t = \theta_0 + \alpha E_t(\pi^*_{t+1}) + \sum_{i=1}^{n} \beta_i \pi_{t-i} + \sum_{j=0}^{m} \gamma_j pm_{t-j} + \delta \hat{y}_{t-1} + \varepsilon_t \quad (1)$$

where π is actual inflation, π^* is expected inflation, pm is a measure of import prices, $\hat{y}$ is a variable that measures available slack in the economy and E_t is the expectations operator. Explanatory variables are in general considered with a lag to make results more robust to potencial endogeneity (Albuquerque and Baumann (2017)).

The possibility of more restricted models (with the exclusion of one or several regressors), with the limit option of a purely autoregressive (AR) model is also explored.

A first exercise, along the lines of Albuquerque and Baumann (2017) and Ciccarelli and Osbat (2017), was to estimate PC models for a lage set of inflation, slack, import price and inflation expectations measures. This approach tries to address the fact that there is large uncertainty in the measurement of slack (Yellen (2013)) and inflation expectations and at the same time access which inflation concept is more suitable to be fitted by the PC.

Estimation details

This article will be focused on inflation measures stemming from consumer prices, namely the Harmonized Index of Consumer prices (HICP). A standard approach to Phillips curves is to consider wages as the relevant inflation measure, but administrative changes to wages during recent years hinder the quality of the data and may distort results. In addition, wage data has the additional problem of data revisions, which create an additional source of uncertainty in PC estimation.

Below there are some details on the variables considered. Data is in general seasonally adjusted, with the exception of some unemployment measures and of the inflation expectations targeted to the annual rate of change in prices.

- **Inflation measures**: the overal HICP and the HICP excluding energy and food (the most volatile components) are considered. Due to the importance of indirect tax increases in 2011 and 2012, which are administratively driven and may distort results, HICP and HICP excluding energy and food are also considered correcting for the impact of indirect taxes (see Ciccarelli and Osbat (2017) for motivation on the use of these variables, and the impact in the Portuguese case). The estimation is made on the basis of data expressed in annualized quarter-on-quarter rates of change.

- **Slack measures**: A wide range of slack measures was considered. These include several output gap estimates, both model based (Cobb Douglas, CES, UCM) and filter based (HP, BK, CF) (see Banco de Portugal (2017) for more details on these measures). The output gaps published by the European Commission (EC) and International Monetary Fund (IMF) are also considered.[3] Several measures related to unemployment were also considered. These include the unemployment rate, (both the headline and a broader measure[4]) and the unemployment gap. The short-term unemployment rate was also considered because some authors argue that it may be more representative of cyclical pressure to inflation than the headline unemployment rate (Dotsey *et al.* (2017)). The unemployment recession gap, defined as the difference between the current unemployment rate and the minimum unemployment rate over the current and previous eleven quarters, was also included (Stock and Watson (2010)). The combined unemployment and labour participation gap (UPRGAP), used by Albuquerque and Baumann (2017), aims at capturing existent slack in the labour market arising from workers that left it temporarily, like discouraged workers.[5] In addition, slack measures derived from the EC business surveys were also considered, namely capacity utilisation and demand and labour as factors limiting production in manufacturing. Finally, the real GDP and real unit labour costs were also included in this set of explanatory measures, expressed in annualized quarter-on-quarter rates of change. All other variables were considered in levels. In the case of the variables related to unemployment, the sign was flipped, to facilitate coefficient comparability.

- **Import price measures**: the options considered include overall import prices and import prices of goods. In addition, these two aggregates excluding energy are also considered. Data is expressed in annualized quarter-on-quarter rates of change.

- **Inflation expectation measures**: the information set includes past inflation (average of past four year-on-year rates of change), Consensus forecasts (both for current year and next year) and EC consumer survey expectations for price developments in the following 12 months. For the

3. This data is annual, and was converted to quarterly frequency using a cubic spline.

4. The broad measure of unemployment includes, along with unemployed, discouraged workers and a measure of involuntary part-time work. For more details see Statistics Portugal press release: `https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=281328836&DESTAQUEStema=5414314&DESTAQUESmodo=2`.

5. The trend labour participation rate implicit in this indicator was calculated as the HP filtered raw data for this variable.

latter variable the level difference vis-à-vis the same period of the previous year is considered, while for the remaining no transformation is applied.

Overall, this information set and subsets were one or several regressors are excluded amount to about 500 different specifications for each inflation aggregate.

Three estimation samples were considered. The start of all samples is 1996Q1 in the case of headline HICP measures or 1997Q1 in the case of "core" inflation measures, but in some cases could be more limited due to regressor availability. Sample 1 ends in 2007Q4, which allows an analysis of the PC behaviour over the Great Recession, a strongly disruptive period for the global economy. Sample 2 ends in 2011Q4, given that after that period Portugal and the euro area began a disinflation path not captured by traditional forecasting models (Ciccarelli and Osbat (2017)). Finally, the full sample, ending in 2017Q4, was also considered. For the smaller samples, a set of out-of-sample conditional (on the actual path of slack, import and expectation measures) forecasts were estimated. This allows an evaluation of models without the noise brought about by the need of forecasting regressors. Forecasts are dynamic, in the sense that the projected HICP for one period serves as autoregressive term in the following periods. For the HICP, the autoregressive lag order included in the model was set to three as a result of trial and error tests on the significant lag order while maintaining the expected sign of coefficients. For import prices the optimal lag to be included in the model was optimized on the basis of the Schwarz information criteria. For slack variables the first lag was considered given that is the most standard approach in the literature, as the use of contemporaneous values may lead to potential endogeneity problems.

Model selection

Table A.1 shows that the performance of PC forecasts is better for "core" aggregates, that do not include the more volatile components (food and energy), given that the average and median Root Mean Squared Errors (RMSE) are lower and in some cases the dispersion (min-max range) is also lower. For the overall inflation measures, the overwhelming majority of PC fail to forecast the decline in inflation that took place from 2008 and 2012 onwards.

Results for the core aggregates with constant taxes increase in accuracy from 2012 onwards given that the direct impact of indirect taxes increases that took place in 2011 and 2012 is excluded. This improvement is not however enough to generate a better performance than for the overall HICP excluding energy and food, which was chosen as the relevant aggregate of analysis.

For each subsample, only specifications for which the slack variable was significant with the expected sign were selected. For the other variables, this selection process allowed for the possibility that they were not significant, but if so, only included when they had the expected sign. The group of model specifications that satisfied these restrictions simultaneously over the three samples was selected, thus focusing the analysis on PC specifications that are relatively stable over time. This requirement implied the exclusion of some slack variables, despite the fact that most of them would be included if only the full sample was analysed: real unit labour costs, capacity utilization and demand as a restriction to production, the IMF output gap, the unemployment recession gap and real GDP.

Import prices are rarely significant in sample 1, appearing more frequently as a significant regressor with expected sign as the sample size increased. For the top 20 out-sample performing models, these variables are never present in the sample up to 2008 and rarely present in the sample up to 2012. This suggests that this variable only gains importance in the most recent period, which is a sign of parameter instability of the PC and may be a result of globalization (Constâncio (2015)). The same results apply to inflation expectations measures. Given their apparent growing importance over time, one import price variable (goods excluding energy) and one expectation variable (consumers' survey price expectations) were selected, given that they are present in the top out-of-sample performing specifications for the sample ending in 2011Q4.

There is not a close relationship between the best performing models in terms of the RMSE of out-of-sample errors and in terms of in sample fit. The R^2 (coefficient of determination) is nevertheless relatively low for all specifications in all samples, never reaching a value much above 50%.

After this process of selection, a group of about 50 models is left. Figures (1) and (2) show the conditional forecasts generated by this set of specifications for the top 20 group with lower RMSE.

Conditional forecasts for inflation - year-on-year rate of change

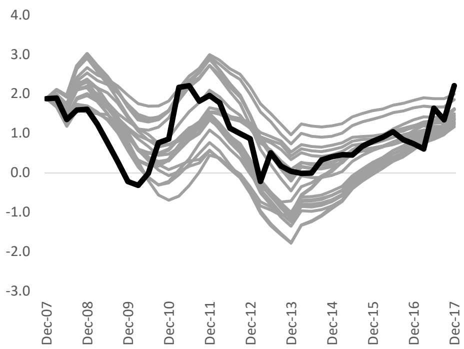


FIGURE 1: Sample 1 (2008Q1-2017Q4)

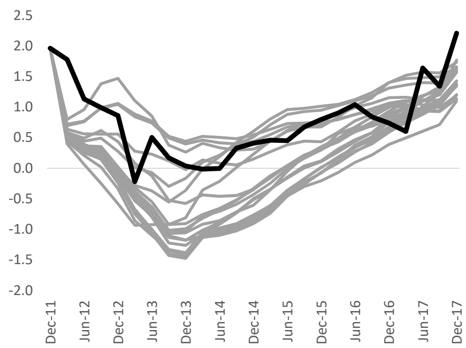


FIGURE 2: Sample 2 (2012Q1-2017Q4)

Note: Inflation is measured by the year-on-year rate of change of the HICP excluding energy and food.

The specifications chosen on the basis of sample 1 appear to capture inflation developments quite well over the out-of-sample period, despite some lag in reflecting the 2008 desinflation and not being able to fully follow the most recent increase. On the other hand, the majority of models estimated over sample 2 tend to overstate the 2012 desinflation. This may partly reflect the increases in indirect taxes that affected this period, that the PC is not able to capture.[6]

Given that the selected models are relatively similar in terms of out-of-sample performance, the remaining part of this article focuses on an even more restricted sample of models. These were selected on the basis of the criteria that they are the top best performing models in terms of out-of-sample forecasts for sample 1, while simultaneously being in the top RMSE also for sample 2 and in the top R^2 for the full sample. Tables A.2 to A.4 report the main estimation results for the top 20 performing models in terms of RMSE in case of samples 1 and 2 and in terms of R^2 for sample 3. The selected measures of slack are the short-term unemployment rate, the survey question related to labour as a limiting factor to production in the manufacturing industry and the Cobb Douglas and CES production function output gaps.[7] Import prices and inflation expectations are not included in any of these "best" models, or when included they are not significant, possibly because the series considered capture the impact of supply shocks and inflation expectations in an imperfect way.

Table A.5 shows the main estimation results for the models that include the selected slack variables and exclude both import prices and expectation measures.

The coefficients on slack variables are all strongly significant. For the output gap variables, which are expected to be nil in the long run, we can compute the long-term expected inflation as the value of the constant divided by 1 minus the sum of the autoregressive coefficients. This yields values close to 2 per cent for the three sample periods. The long run coefficients on slack, computed in the same vein, yield about 0.6, a value broadly in line with those found for other euro area countries (Ciccarelli and Osbat (2017)). The output gap measures clearly outperform the other in terms of RMSE in sample 2, while results are more similar across specifications for sample 1.

The results also show some time variation in the coefficients pertaining to slack variables, namely for all measures considered except the survey question there is a decline in the coefficient when moving from sample 1 to sample

6. The PC for the HICP excluding energy and food with constant taxes, which excludes the impact of these factors, is however even worse for the same sample period. This is because inflation is also underestimated from 2013 onwards, but it is grossly overestimated in 2012.

7. The European Commission output gap would also be a selected indicator according to these criteria, but was not included because import prices are not significant with expected sign, but when excluded from the equation the slack variable becomes non-significant.

3. This is related to the possibility of PC flattening that arose with the 2012 missing inflation puzzle (Constâncio (2015)). To test this possibility, along with a more general one of parameter instability in the PC due to non linearities, a rolling window exercise was performed. This is presented in the following section.

Parameter instability and non-linear Phillips curves

The initial window considered for the rolling sample was sample 1, and from then onwards the model was reestimated moving the window forward by one period.

Rolling Regression estimates of coefficients on slack measures

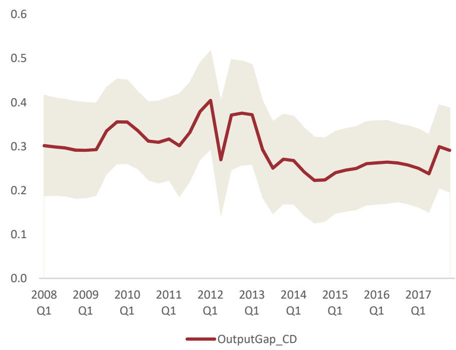


Figure 3: Slack measure: Cobb-Douglas output gap

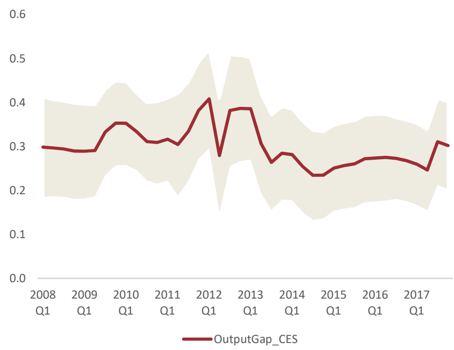


Figure 4: Slack measure: CES output gap

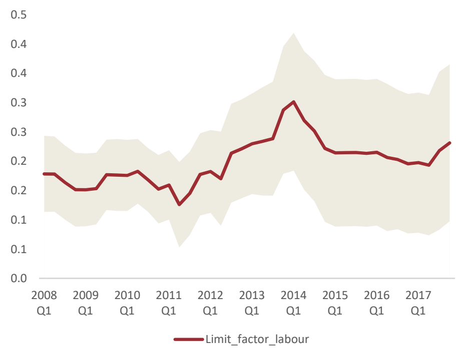


Figure 5: Slack measure: labour as a factor limiting production

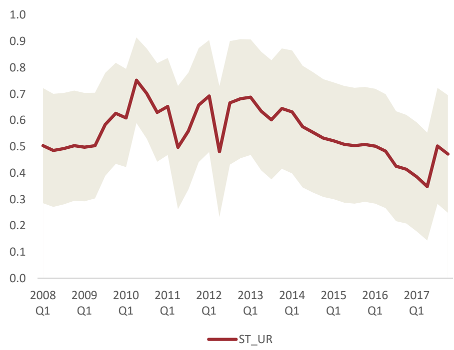


Figure 6: Slack measure: short-term unemployment rate

Note: The shaded area is defined by parameter estimate +/- one standard deviation. The dates on the x-axis define the last quarter included in the rolling regression.

Results, shown in Figures 3 to 6, display in most cases an increase in the coefficient of the PC on slack over the periods of the last two recessions (considering the 2009 decline in GDP as a separate recession) and a posterior decline to levels below those observed prior to the 2008 financial crisis, supporting the thesis of flattening of the Phillips curve. However, this conclusion is contingent on the slack measured considered (the survey-based indicator yields the opposite conclusion) and on the relevant concept of inflation considered (Ciccarelli and Osbat (2017)). This evidence however supports the idea vastly found in the literature (Dotsey *et al.* (2017)) that PC coefficients are time varying, what may be the result of non-linearities in the model. Several hypothesis for the flattening of the PC have been put forward in the literature (see Constâncio (2015) for a summary). One possibility is that inflation only reacts to slack when changes are large enough, given that there are menu costs to changing prices. Another possibility is that during a recovery firms have unused spare capacity, and do not feel the pressure to raise prices until the installed capacity is reached (Macklem (1997)). This inertia effect should be stronger in a low inflation environment.

To test for this hypothesis a threshold model was estimated, defined in equation 2:

$$\pi_t = \theta_0 + \theta_1 I_{outT} + \sum_{i=1}^{n} \beta_i \pi_{t-i} + \delta_1 I_{outT} \hat{y}_{t-1} + \delta_2 (1 - I_{outT}) \hat{y}_{t-1} + \varepsilon_t \tag{2}$$

where I_{outT} is a dummy variable that takes the variable of 1 when $\hat{y}$ falls outside the thresholds and zero otherwise.

The thresholds considered are given by the 30th and 70th percentile of the distribution of $\hat{y}$.[8]

Table A.6 shows the result of this estimation. For the specifications in which the slack variable is the output gap, only the values outside the threshold are significant to explain changes in inflation, for all samples considered. There is no evidence of a change in the constant of the equation when slack variables lie outside the threshold. For the other two slack variables considered results suggest that this type of disaggregation does not help explaining inflation. The out-of-sample forecasting performance of the models where thresholds are significant worsens vis-à-vis the previous exercise, but this possibly reflects the fact that the non significant regressors are being used to produce these forecasts.

Real time forecasting exercise

As a final exercise, the forecasting performance of selected models is evaluated in real time. The real time analysis is particularly important for output gap measures, where data revisions can be substantial (Banco de Portugal (2017)).

8. Results are qualitatively similar with thresholds of 20th-80th and 25th-75th percentiles.

Therefore, adding to the potential forecasting errors of the PC arising from uncertainty regarding the model used and the projected path of the regressors, there is additional uncertainty due to possible revisions in potential output estimates. The other two slack variables, the short-term unemployment rate and the survey question related to labour as a limiting factor to production, are not subject to revisions, and have no projections available. In this case the actual values were used to produce the conditional forecasts[9], which favours PC results in this case, given that the uncertainty associated to regressors' projections is ruled out. In this exercise, the relative performance of the PC is confronted with a naïve random walk model and with the half-yearly Eurosystem projections. The choice of the random walk benchmark allows an evaluation of the PC performance against a very simple and standard reference in the literature, but which provided very good results in terms of inflation forecasting on a short to medium run horizon (Teles and Garcia (2016)). On the other extreme, the Eurosystem projections provide a very demanding benchmark, given that they are computed at a very detailed level and benefit from experts' judgement. In this case the relevant question is whether a simple PC model can provide inflation forecasts as accurate as this benchmark. Threshold models that include only the values of the slack variables outside the treshold are also considered for the cases where they were found significant, namely for output gap specifications. In this case the thresholds were defined also in real time, i.e., taking into account the distribution of data available at each vintage.

Forecasts are produced from 1 to 8 steps ahead and evaluated in terms of their relative RMSE vis-à-vis the benchmarks for each horizon. Moreover, the significance of these relative differences is tested with the Diebold-Mariano (Diebold and Mariano (1995)) test with Harvey *et al.* (1997) correction, considering a 10% significance level. The vintages available for this evaluation range from the June 2007 projection exercise (with observed data for the HICP up to 2007Q1) to the December 2017 projection exercise (with observed data for the HICP up to 2017Q3), thus 22 vintages in total.

The relative RMSE of the forecasts vis-à-vis the benchmark of the random walk and the Eurosystem projections are presented respectively in Tables A.7 and A.8. The relative performance of PC models is in general better than the random walk, with the exception of the short-term unemployment rate model. This outperformance is statistically significant for some medium to long term horizons. The threshold models do not perform better than their standard counterpart. On the other hand, the Eurosystem projections are only better than the random walk in a statically significant way for one step ahead forecasts. Considering the Eurosystem projections as the benchmark,

9. When necessary, the dataset for the short-term unemployment variable was extended beyond 2017Q4 with the quarterly changes in the latest unemployment projections of the Eurosystem.

the PC relative RMSE are also in general lower than 1 for medium to long-run horizons, but this difference is not statistically significant. The AR and random walk models have a relative RMSE higher than 1, which is only significant for short-term horizons.

Conclusion

Despite some difficulties in coping with inflation fluctuations since the Great Recession, Phillips curves remain a staple of economic analysis for central bankers. This article resorted to a large diversity of slack measures to estimate Phillips curves for Portugal. These models have some forecasting power for inflation, but results have shown that the best slack measure is not constant over time, and therefore it is preferable to rely on a diversified set of Phillips curves. There is some evidence that some nonlinearities are present in Phillips curves estimation, but further work is needed on this issue and how best to tackle it for forecasting purposes.

Appendix: Tables

	HICP	HICPxENFOOD	HICP_CT	HICP_CT_xENFOOD	HICP	HICPxENFOOD	HICP_CT	HICP_CT_xENFOOD
Sample up to:	2007Q4				2011Q4			
Max	5.4	2.7	4.8	6.7	3.1	3.3	3.4	2.6
Min	1.1	0.6	1.1	0.6	0.5	0.4	0.5	0.5
Average	1.7	1.3	1.8	1.4	1.6	1.0	1.5	1.0
Median	1.6	1.3	1.6	1.2	1.5	0.9	1.5	0.9

TABLE A.1. RMSE of conditional forecasts

Slack variable	slack coefficient	constant	Import prices Included?	Import price coefficient	Expectations Included?	Expectations coefficient	RMSE	R^2
EC_OG	0.22	0.99	yes	not significant	no		0.62	0.36
Limit_factor_labour	0.18	0.17	no		no		0.63	0.39
Limit_factor_labour	0.18	0.30	no		yes	not significant	0.64	0.45
OutputGap_CES	0.29	1.10	yes	not significant	no		0.68	0.41
Limit_factor_labour	0.19	0.37	yes	not significant	no		0.70	0.45
OutputGap_CD	0.29	1.08	yes	not significant	no		0.72	0.41
ST_UR	0.50	2.95	no		no		0.72	0.36
OutputGap_CES	0.28	0.94	no		yes	not significant	0.74	0.42
OutputGap_BK	0.37	0.88	no		no		0.75	0.37
OutputGap_CES	0.30	0.90	no		no		0.77	0.38
OutputGap_CD	0.28	0.93	no		yes	not significant	0.79	0.42
OutputGap_BK	0.32	1.05	yes	not significant	no		0.82	0.37
OutputGap_CD	0.30	0.88	no		no		0.83	0.38
OutputGap_HP	0.34	1.05	no		no		0.84	0.36
ST_UR	0.68	3.99	yes	not significant	no		0.85	0.48
OutputGap_CF	0.20	1.19	no		no		0.97	0.37
ST_UR	0.70	3.98	no		yes	not significant	0.99	0.48
OutputGap_UCM	0.44	1.90	yes	not significant	no		1.08	0.44
OutputGap_UCM	0.43	1.78	no		yes	not significant	1.10	0.44
OutputGap_CF	0.20	1.38	yes	not significant	no		1.14	0.39

TABLE A.2. Main estimation results for sample 1997Q1-2007Q4

Notes: the acronyms for slack variables stand for (in the order they are presented): the EC output gap, the survey question related to labour as a restrictive factor to production, the CES output gap, the Cobb-Douglas output gap, the short-term unemployment rate, the Baxter-King output gap, the Hodrick-Prescott output gap, the Christiano-Fitzgerald output gap and the unobserved components model output gap. Banco de Portugal (2017) provides details on the output gap measures.
The shaded variables denote the selected slack variables.

Slack variable	slack coefficient	constant	Import prices Included?	Import price coefficient	Expectations Included?	Expectations coefficient	RMSE	R^2
Limit_factor_labour	0.16	0.09	no		yes	not significant	0.45	0.48
ST_UR	0.46	3.14	no		no		0.46	0.39
Limit_factor_labour	0.16	0.16	no		no		0.46	0.39
OutputGap_CF	0.16	0.51	no		no		0.47	0.36
Limit_factor_labour	0.17	0.17	yes	not significant	no		0.48	0.43
ST_UR	0.58	3.87	yes	not significant	no		0.56	0.47
EC_OG	0.31	0.73	yes	not significant	no		0.56	0.43
Unemployment_gap	0.28	1.60	no		no		0.68	0.39
Unemployment_gap	0.29	1.68	yes	not significant	no		0.71	0.43
OutputGap_UCM	0.33	1.67	no		no		0.73	0.41
ST_UR	0.65	4.17	no		yes	not significant	0.74	0.54
OutputGap_UCM	0.34	1.75	yes	0.04	no		0.76	0.47
UPRGAP	0.31	1.77	no		no		0.81	0.40
UPRGAP	0.32	1.83	yes	not significant	no		0.84	0.44
OutputGap_CES	0.35	1.12	yes	not significant	no		0.95	0.49
OutputGap_CES	0.36	1.11	no		no		0.95	0.45
Labour_slack	0.21	4.21	no		no		0.97	0.41
OutputGap_CES	0.32	1.00	no		yes	not significant	1.00	0.51
OutputGap_CD	0.35	1.11	yes	not significant	no		1.01	0.48
OutputGap_CD	0.36	1.10	no		no		1.01	0.45

TABLE A.3. Main estimation results for sample 1997Q1-2011Q4

Notes: the acronyms for slack variables stand for (in the order they are presented): the survey question related to labour as a restrictive factor to production, the short-term unemployment rate, the Christiano-Fitzgerald output gap, the EC output gap, the unemployment gap, the unobserved components model output gap, the combined unemployment and labour participation gap, the CES output gap, the measure of unemployment in broad sense and the Cobb-Douglas output gap. Banco de Portugal (2017) provides details on the output gap measures.
The shaded variables denote the selected slack variables.

Slack variable	slack coefficient	constant	Import prices Included?	Import price coefficient	Expectations Included?	Expectations coefficient	R^2
OutputGap_CF	0.21	0.67	no		no		0.34
OutputGap_HP	0.25	0.83	no		no		0.34
OutputGap_BK	0.30	0.89	no		no		0.36
UR	0.22	3.39	no		no		0.38
Labour_slack	0.17	4.16	no		no		0.38
Limit_factor_labour	0.22	0.23	no		no		0.38
UR	0.21	3.35	yes	0.05	no		0.45
Unemployment_gap	0.26	2.05	no		no		0.38
UPRGAP	0.26	2.14	no		no		0.39
OutputGap_UCM	0.29	2.07	no		no		0.39
OutputGap_CF	0.18	0.70	yes	0.05	no		0.40
OutputGap_CD	0.27	1.64	no		no		0.41
OutputGap_CES	0.28	1.64	no		no		0.41
Labour_slack	0.16	4.02	no		yes	0.02	0.42
OutputGap_BK	0.27	0.89	yes	0.05	no		0.42
OutputGap_CD	0.24	1.56	no		yes	not significant	0.44
EC_OG	0.28	1.18	yes	0.05	no		0.44
OutputGap_CES	0.25	1.57	no		yes	not significant	0.44
Limit_factor_labour	0.20	0.36	no		yes	0.02	0.44
UPRGAP	0.24	2.15	no		yes	0.02	0.44

TABLE A.4. Main estimation results for sample 1997Q1-2017Q4

Notes: the acronyms for slack variables stand for (in the order they are presented): the Christiano-Fitzgerald output gap, the Hodrick-Prescott output gap, the Baxter-King output gap, the unemployment rate, the measure of unemployment in broad sense, the combined unemployment and labour participation gap, the unobserved components model output gap, the Cobb-Douglas output gap, the CES output gap and the EC output gap. Banco de Portugal (2017) provides details on the output gap measures.
The shaded variables denote the selected slack variables.

Slack variables	sample 1: 1997Q1-2007Q4 sum of AR coefficients	slack	constant	RMSE	R^2	sample 3: 1997Q1-2011Q4 sum of AR coefficients	slack	constant	RMSE	R^2	sample 2: 1997Q1-2017Q4 sum of AR coefficients	slack	constant	R^2
ST_UR	0.60	0.50	2.95	0.72	0.36	0.47	0.46	3.14	1.12	0.39	0.35	0.49	3.53	0.38
		(0.22)	(1.08)				(0.18)	(1.06)				(0.14)	(0.9)	
Limit_factor_labour	0.44	0.18	0.17	0.63	0.39	0.46	0.16	0.16	1.13	0.39	0.27	0.22	0.23	0.38
		(0.07)	(0.56)				(0.06)	(0.4)				(0.06)	(0.28)	
OutputGap_CD	0.53	0.30	0.88	0.83	0.38	0.41	0.36	1.10	0.50	0.45	0.23	0.27	1.64	0.41
		(0.11)	(0.52)				(0.1)	(0.38)				(0.07)	(0.38)	
OutputGap_CES	0.52	0.30	0.90	0.77	0.38	0.41	0.36	1.11	0.51	0.45	0.23	0.28	1.64	0.41
		(0.11)	(0.52)				(0.1)	(0.38)				(0.07)	(0.37)	

TABLE A.5. Main estimation results for selected models

Notes: figures between brackets refer to the standard deviation of the corresponding coefficients. The acronyms for slack variables stand for (in the order they are presented) the short-term unemployment rate, the survey question related to labour as a restrictive factor to production, the Cobb-Douglas output gap and the CES output gap.

slack variables	sample 1: 1997Q1-2007Q4 Sum of AR coefficients	OutT	InT	OuT Dummy	Constant	RMSE	R^2	sample 2: 1997Q1-2011Q4 Sum of AR coefficients	OutT	InT	OuT Dummy	Constant	RMSE	R^2	sample 3: 1997Q1-2017Q4 Sum of AR coefficients	OutT	InT	OuT Dummy	Constant	R^2
ST_UR	0.66	-0.09	-0.25	1.89	-0.69	1.53	0.47	0.53	0.33	0.37	0.45	2.32	0.59	0.42	0.38	0.48	0.49	0.30	3.28	0.39
		(0.63)	(0.47)	(0.73)	(2.12)				(0.2)	(0.19)	(0.39)	(1.21)				(0.14)	(0.16)	(0.37)	(0.37)	
LF_labour	0.64	-0.01	0.11	1.62	-0.04	1.35	0.46	0.52	0.13	0.18	0.50	-0.14	0.58	0.40	0.22	0.23	0.31	0.08	0.06	0.39
		(0.15)	(0.23)	(0.79)	(0.99)				(0.08)	(0.14)	(0.46)	(0.55)				(0.07)	(0.12)	(0.42)	(0.42)	
OutputGap_CD	0.48	0.61	-0.39	-0.51	0.74	2.15	0.45	0.29	0.62	0.02	-0.89	1.56	2.21	0.55	0.22	0.27	0.41	-0.31	1.89	0.41
		(0.23)	(0.57)	(0.38)	(0.5)				(0.12)	(0.31)	(0.38)	(0.5)				(0.07)	(0.35)	(0.38)	(0.5)	
OutputGap_CES	0.48	0.60	-0.42	-0.49	0.72	2.02	0.45	0.28	0.62	0.02	-0.91	1.58	2.12	0.56	0.21	0.28	0.42	-0.33	1.91	0.42
		(0.22)	(0.57)	(0.68)	(0.75)				(0.12)	(0.32)	(0.37)	(0.46)				(0.07)	(0.36)	(0.37)	(0.37)	

TABLE A.6. Main estimation results for threshold models

Notes: figures between brackets refer to the standard deviation of the corresponding coefficients. The acronyms for slack variables stand for (in the order they are presented) the short-term unemployment rate, the survey question related to labour as a restrictive factor to production, the Cobb-Douglas output gap and the CES output gap. OutT and IntT stand for the slack variable values outside and inside the thresholds, respectively. OutT Dummy stands for the dummy variable that has an unit value outside the thresholds.

	Slack variable	OutputGap CD	OutputGap CES	OutputGap CD (outT)	OutputGap CES (outT)	LF_labour	ST_UR	AR Model	Eurosystem projections	Random Walk
Steps ahead	1	0.8	0.8	0.8	0.8	0.9	0.9	1.0	0.5	1.0
	2	0.8	0.8	0.8	0.8	0.9	1.0	1.0	0.7	1.0
	3	0.7	0.7	0.7	0.7	0.9	1.0	1.2	1.0	1.0
	4	0.7	0.7	0.7	0.8	0.9	1.1	1.1	0.9	1.0
	5	0.7	0.7	0.8	0.8	0.9	1.0	1.1	1.0	1.0
	6	0.8	0.8	0.9	0.9	0.8	1.1	1.1	0.7	1.0
	7	0.8	0.8	0.9	0.9	0.8	1.1	1.1	0.6	1.0
	8	0.9	1.0	1.1	1.1	0.8	1.1	1.1	0.7	1.0

TABLE A.7. Relative RMSE - Benchmark Random Walk

Notes: shaded values stand for statistically significant differences between the forecasts according to the Diebold Mariano test. The acronyms ST_UR and LF_labour stand for, respectively, the short-term unemployment rate the survey question related to labour as a restrictive factor to production. The "outT" models refer to threshold models.

	Slack variable	OutputGap CD	OutputGap CES	OutputGap CD (outT)	OutputGap CES (outT)	LF_labour	ST_UR	AR Model	Eurosystem projections	Random Walk
Steps ahead	1	1.6	1.6	1.6	1.6	1.8	1.7	2.0	1.0	2.0
	2	1.1	1.1	1.1	1.1	1.3	1.4	1.5	1.0	1.4
	3	0.7	0.7	0.7	0.7	0.9	1.0	1.2	1.0	1.0
	4	0.8	0.8	0.8	0.8	1.0	1.2	1.2	1.0	1.1
	5	0.7	0.7	0.8	0.8	0.8	1.0	1.1	1.0	1.0
	6	0.7	0.7	0.8	0.8	0.8	1.3	1.5	1.0	1.2
	7	0.8	0.7	0.8	0.8	0.7	1.2	1.5	1.0	1.2
	8	0.8	0.8	0.9	0.9	0.7	1.3	1.5	1.0	1.2

TABLE A.8. Relative RMSE - Benchmark Eurosystem projections

Notes: shaded values stand for statistically significant differences between the forecasts according to the Diebold Mariano test. The acronyms ST_UR and LF_labour stand for, respectively, the short-term unemployment rate the survey question related to labour as a restrictive factor to production. The "outT" models refer to threshold models.

References

Albuquerque, Bruno and Ursel Baumann (2017). "Will US inflation awake from the dead? The role of slack and non-linearities in the Phillips curve." *Journal of Policy Modeling*, 39(2), 247–271.

Banco de Portugal (2017). "Potential output: challenges and uncertainties." *Economic Bulletin*.

Ciccarelli, Matteo and Chiara Osbat (2017). "Low inflation in the euro area: Causes and consequences." Occasional Paper Series 181, European Central Bank.

Constâncio, Vítor (2015). "Understanding Inflation Dynamics and Monetary Policy: Panel remarks at the Federal Reserve Bank of Kansas City Economic Symposium, Jackson Hole, Wyoming." https://www.ecb.europa.eu/press/key/date/2015/html/sp150829.en.html.

Diebold, Francis X and Roberto S Mariano (1995). "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics*, 13(3), 253–263.

Dotsey, Michael, Shigeru Fujita, and Tom Stark (2017). "Do Phillips Curves Conditionally Help to Forecast Inflation?" Working Papers 17-26, Federal Reserve Bank of Philadelphia.

Draghi, Mario (2017). "Accompanying the economic recovery: speech at the ECB Forum on Central Banking." https://www.ecb.europa.eu/press/key/date/2017/html/ecb.sp170627.en.html.

ECB (2014). "The Phillips Curve Relationship in the Euro Area." *ECB Monthly Bulletin*, 45(July), 99–114.

Harvey, David, Stephen Leybourne, and Paul Newbold (1997). "Testing the equality of prediction mean squared errors." *International Journal of Forecasting*, 13(2), 281–291.

Macklem, Tiff (1997). "Capacity constraints, price adjustment, and monetary policy." *Bank of Canada Review*, 1997(Spring), 39–56.

Stock, James H. and Mark W. Watson (2010). "Modeling Inflation After the Crisis." NBER Working Papers 16488, National Bureau of Economic Research, Inc.

Teles, Pedro and Joana Garcia (2016). "How can the Phillips curve be used for today's policy?" *Economic Studies*, 2(2).

Yellen, Janet (2013). "Labor Market Dynamics and Monetary Policy: a speech at the Federal Reserve Bank of Kansas City Economic Symposium, Jackson Hole, Wyoming ."

Forecasting exports with targeted predictors

Francisco Dias
Banco de Portugal

Nuno Lourenço
Banco de Portugal and Nova SBE

António Rua
Banco de Portugal and Nova SBE

Abstract
This article applies factor models to forecast monthly Portuguese exports by resorting to an international dataset covering the country's main trading partners. We find noteworthy forecasting gains up to 12-month ahead when soft indicators for these countries are pooled and predictors are pre-selected prior to factor estimation. Resorting solely on national data and with no pre-selection of predictors yields greater forecasting accuracy when nowcasting. Hence, data from Portugal's main trading partners is more informative to produce h-step ahead forecasts. In turn, when hard and soft data are pooled, forecast accuracy is, in general, not enhanced. (JEL: C38, C55, F47)

Introduction

Forecasting macroeconomic time series is of utmost importance for fiscal and monetary policymakers to monitor or assess developments in any economy. Recent advances on short-term forecasting have drawn on the use of large datasets, where progress in information technology allows nowadays to access and handle hundreds of economic time series in real-time. Hard and soft data are at the core of this data-rich environment. While the former are based on quantitative information, the latter builds on surveys of economic activity that are characterized by the qualitative nature of their questions (e.g. regular harmonised surveys conducted by the European Commission for different sectors in the European Union). The interest in relying on soft data to forecast macroeconomic variables has been emphasized in the literature (see, for instance, Bańbura and Rünstler (2011) and Hansson *et al.* (2005) for an application to forecast GDP growth). A key advantage of qualitative indicators lies on their timeliness, as most surveys are released a few days after the reference period. Their high signal-to-noise ratio provides substantial informational content on the state of the economy and their encompassing nature allows for a wide sectoral coverage. Furthermore, since some questions concern future developments, they provide early information on the possible evolution of the economy. As these soft data series are not subject to revisions, real-time reliability is assured.

E-mail: fadias@bportugal.pt; nalourenco@bportugal.pt; arua@bportugal.pt

The use of soft data in addition to hard data for short-term forecasting has been notably attractive in Europe, where surveys are widely available for a long time (see, for example, Schumacher (2007) for an application to forecast GDP for Germany, Rünstler *et al.* (2009) to forecast GDP for several European countries and Angelini *et al.* (2011) for the euro area). While exploiting information from a data-rich environment has been widely documented in the literature, few authors have focused on the explicit role of adding foreign data to forecast national macroeconomic variables. In this direction, one should highlight the contributions by Brisson *et al.* (2003), where they take on board the predictive content of the United States variables as well as from other countries' to forecast Canadian real economic activity and inflation. Likewise, Schumacher (2009) considers the role of euro area and remaining G7 countries to forecast German GDP.

In a data-rich environment, forecasting macroeconomic variables amounts to extracting useful information from a large number of predictors. Factor models have been quite popular for such exercises, where the informational content from a large panel of time series is summarized in a few number of factors which are then used for forecasting purposes. Amongst the applications on the use of factor models stand out the seminal contributions by Stock and Watson (1999, 2002a,b) to forecast major macroeconomic variables for the United States, Marcellino *et al.* (2003) for euro-wide inflation and real activity and Rünstler *et al.* (2009) for a cross-country study comprising several European countries.

However, the use of additional data for forecasting with factor models might not improve forecast accuracy. In fact, extending a dataset for factor estimation can lead to worse forecasting results if the additional series are noisy or if forecasting power is provided by a factor that is dominant in a smaller dataset but turns out to be a dominated factor in a larger dataset (see Boivin and Ng (2006)). Reducing the influence of uninformative predictors to forecast a macroeconomic variable has given rise to a new stream of literature. In this regard, Bai and Ng (2008) proposed penalized regression techniques to target predictors, in particular, Least-Angle Regression with Elastic Net (henceforth LARS-EN), where selection of a subset of predictors prior to factor estimation is conducted to forecast US inflation. To further illustrate the importance of screening predictors prior to factor estimation, Schumacher (2007) provides an application to forecast German GDP and Li and Chen (2014) focus on several important variables in tracking the economy and monetary policies in the United States.

We investigate the role of information contained in foreign data to forecast international trade flows, with emphasis on exports of goods for Portugal, by extending the dataset to cover the country's main trading partners.[1] Thus, we contribute to the strand of the literature that relies on international data to forecast national variables. The high degree of interrelatedness of the Portuguese economy with the rest of the world lies at the core of this heterogeneous dataset. Given its large size, we then use LARS-EN pre-selection of predictors and infer on the usefulness of selection of predictors prior to factor estimation to enhance forecast performance. We exploit timely monthly data to nowcast and forecast Portugal's exports of goods on a monthly basis up to 12-month ahead.

The article proceeds as follows. Section 2 provides a quick glance at Portuguese exports. Section 3 introduces the analytical framework used in the forecasting exercise. Section 4 describes the underlying dataset used in the empirical application. The empirical results are discussed in Section 5 and Section 6 concludes.

A quick glance at Portuguese exports

Portugal has made notable progress in increasing its integration in world trade in the last few decades, with the accession to the European Economic Community in 1986 showing a particular leap forward. The relative importance of exports in the economy has grown gradually. However, it declined sharply in 2009 with the collapse of world trade, having gained pace thereafter (Figure 1).

The country's share of exports of goods in GDP in nominal terms has increased around 11 percentage points since 1993, standing roughly at 27 per cent in 2016. In the aftermath of the recent economic and financial crisis, Portugal experienced a gradual reallocation of inputs towards the production of goods for foreign markets. As a result, the relative importance of exports of goods increased markedly following the Great Recession.

A cross-country comparison with the euro area initial member states plus Greece (EA-12) portrays that Portugal stands as one of the countries that experienced the sharpest increase in the share of exports of goods in GDP during the period under study (Figure 2). It ranks ahead of the majority of the EA-12 member states, being surpassed only by The Netherlands and Germany.

1. Previous work on forecasting Portuguese exports of goods include Cardoso and Duarte (2006), who focus on forecasting nominal exports of goods using a small number of soft indicators through a bridge model.

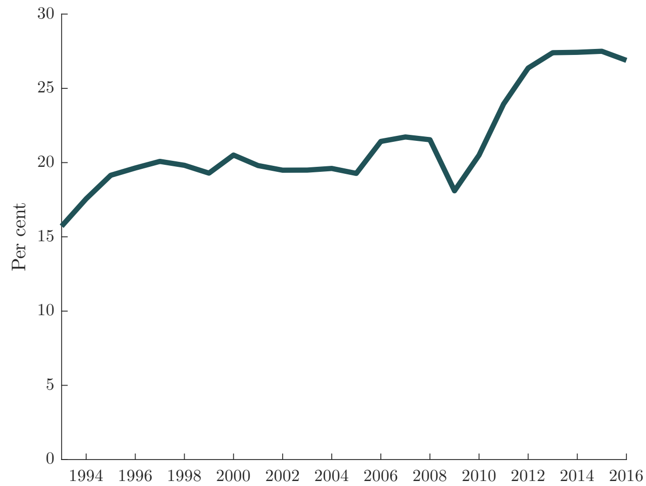


FIGURE 1: Share of exports of goods in GDP in nominal terms.

Source: Statistics Portugal.

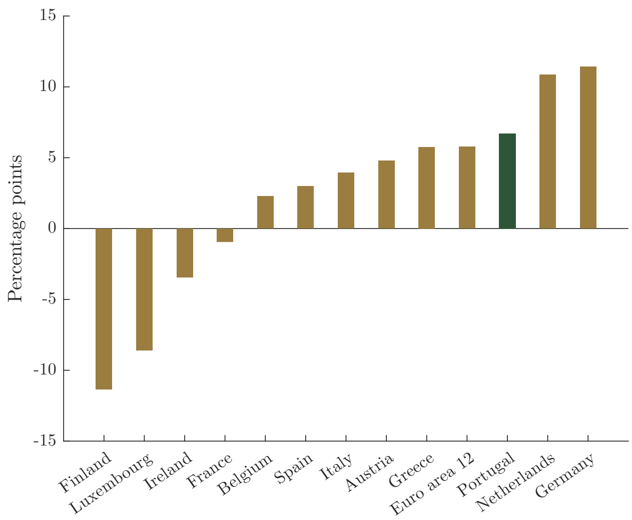


FIGURE 2: Change in the share of exports of goods in GDP in nominal terms between 2000 and 2016.

Source: Eurostat.

Concerning the main destinations of Portugal's exports of goods, euro area countries account for a large fraction. In Table 1, the main trading partners in 2016 are listed. In particular, exports to Spain comprise more than one fourth of total exports of goods, whereas France and Germany account for more than 10 per cent.

Main trading partners	**Shares (in per cent)**
Spain	26.2
France	12.6
Germany	11.6
United Kingdom	7.0
United States	4.9
Netherlands	3.7
Italy	3.4
Angola	3.0
Belgium	2.4

TABLE 1. Main destinations of Portuguese exports of goods in 2016.

Econometric framework

Factor models

We begin with a discussion of the factor model representation that motivates forecasting in a data-rich environment. Let X_t be an N-dimensional column vector of time series of predictor variables, observed for $t = 1, ..., T$. The aim of the exercise lies in representing these variables with a factor model representation and using the estimated factors to derive h-step ahead forecasts of the variable of interest, y, that is, y_{t+h}, where h denotes the forecast horizon. The variables in X_t are represented as the sum of two orthogonal components: the common component, driven by a small number of unobserved common factors that accounts for most of the co-movement among the variables; and the idiosyncratic component, driven by variable-specific shocks.

The data generating process for X_t admits a static factor representation written as:

$$X_t = \Lambda F_t + \xi_t \tag{1}$$

where $F_t = (f_{1t}, ..., f_{rt})'$ is an $(r \times 1)$ vector of non-observable factors, Λ is an $(N \times r)$ matrix of unknown factor loadings and ξ_t denotes an N-dimensional vector of idiosyncratic terms. As pointed out by Stock and Watson (2002b), unobserved factors can be estimated consistently through principal components under fairly general assumptions.

Factor estimation by principal components aims at maximizing the explained variance in the whole dataset. Typically, the first few top-ranked principal components capture a sizeable fraction of the co-movement among the series in the dataset. Once the number of factors has been selected, the variable to be forecast is projected eventually on its lags and on the set of r estimated factors. This yields the following forecasting equation for the variable of interest:

$$y_{t+h} = \alpha_0 + \sum_{i=1}^{r} \alpha_i \hat{F}_{t,i} + \sum_{j=0}^{p} \delta_j y_{t-j} + \varepsilon_{t+h} \tag{2}$$

where α_0 is a constant term, α_i denotes the coefficients pertaining to $\hat{F}_{t,i}$, i.e., the principal component estimates of the factors in Equation (1), y_{t-j} accounts for the autoregressive component of the regression, where δ_j are the corresponding coefficients and p the number of lags.

The LARS-EN algorithm

Typically, when factors are estimated, the informational content other than the one conveyed by the small set of r factors is ignored, thus, it can disregard useful information for the variable to be forecast or the forecast horizon at stake. Following Bai and Ng (2008), forecasting using targeted predictors is considered. In other words, the relationship between y_{t+h} and X_t is analyzed in order to select a subset of predictors $X_{t,\mathcal{A}} \subseteq X_t$ prior to factor estimation.

We now describe a method based on penalized regressions that performs subset selection and shrinkage by dropping uninformative regressors. Put differently, the regression coefficients of those variables that are less informative for predicting the targeted variable are penalized. Following Zou and Hastie (2005), Bai and Ng (2008) suggest the use of the EN optimization problem which is given by:

$$\underset{\beta}{min} \left\{ RSS + \lambda_1 \sum_{j=1}^{N} |\beta_j| + \lambda_2 \sum_{j=1}^{N} \beta_j^2 \right\} \tag{3}$$

where RSS is the residual sum of squares from a regression of y_{t+h} on all available regressors, and λ_1 and λ_2 penalize with the L_1- and L_2-norm of β, respectively.

The L_1 penalty solves

$$\underset{\beta}{min} \left\{ RSS + \lambda_1 \sum_{j=1}^{N} |\beta_j| \right\} \tag{4}$$

where the tuning parameter λ_1 controls for the amount of shrinkage, and thus for the number of parameters that are set to zero. The method adds λ_1 regularization to ordinary least squares regression, yielding solutions that are sparse in terms of the regression coefficients. This is also know as the Least Absolute Shrinkage and Selection Operator (LASSO) solution-type of Tibshirani (1996).

In turn, the L_2 penalty solves

$$\underset{\beta}{min}\left\{RSS+\lambda_2\sum_{j=1}^{N}\beta_j^2\right\} \tag{5}$$

which for $0 \leq \lambda_2 < \infty$ shrinks toward zero the coefficients of the uninformative predictors. This is also know as the L_2 penalty of ridge regression.

By combining both penalties, i.e., the virtues of LASSO and ridge regression, the EN in Equation (3) allows for shrinkage of coefficients, elimination of regressors and efficient selection of variables within the dataset.

LARS provides an efficient algorithm to solve the EN minimization problem (see Zou and Hastie (2005)). The LARS algorithm estimates β and selects the subset of predictors $X_{t,\mathcal{A}} \subseteq X_t$ by solving the optimization criterion in Equation (3), given the parameters λ_1 and λ_2. In practice, the calibration of λ_1 is recast as a rule for the maximum number of variables with non-zero β_j included in the analysis, i.e., the number of regressors $N_{\mathcal{A}} \leq N$ to be included in $X_{t,\mathcal{A}}$. The procedure works as follows. It starts with all coefficients equal to zero and finds the most correlated predictor with the variable to be forecast. It takes the largest step possible in the direction of this predictor until a second predictor has as much correlation with the current residual. Instead of continuing along the first predictor, LARS proceeds in a equiangular direction between the two predictors until a third variable earns its way into the most correlated set. LARS then proceeds equiangularly between the three predictors, that is, along the least angle direction, until a fourth variable enters and so on. The algorithm builds up estimates $\hat{\mu} = X\hat{\beta}$ in successive steps, each step adding one covariate to the model, so that after k steps just k of the $\hat{\beta}_j$'s are non-zero (see Efron *et al.* (2004) for details).

Data

The forecasting exercise comprises forecasting the growth rate of a key macroeconomic variable, nominal exports of goods for Portugal. This variable is released on a monthly basis 40 days after the reference period by Statistics Portugal without any seasonal or calendar adjustment.

We focus on forecasting the year-on-year growth rate of the series. Besides allowing to tackle deterministic seasonality, this choice can be motivated by several other reasons, such as the high volatility underlying month-on-month growth rates of nominal trade data or the larger resemblance between variables measured in year-on-year terms and the profile of several qualitative indicators. By considering year-on-year rates of change, noise in the data is reduced and data irregularities are smoothed out. For further discussion, see Esteves and Rua (2012). However, even when modelling the dependent variable as a year-on-year growth rate, calendar effects or moving holidays can be sizeable and are expected to impact the outcome variable. These effects are addressed resorting to deterministic variables to be described in the forecasting exercise.

Data for Portugal builds on a monthly dataset described in detail in Dias *et al.* (2015, 2018) comprising business and consumer surveys, retail sales, industrial production, turnover in industry and services, employment, hours worked and wage indices in industry and services, overnight stays in Portugal, car sales, cement sales, vacancies and registered unemployment, energy consumption, imports of goods, real effective exchange rate, Portuguese stock market index and ATM/POS series. Furthermore, we extend this dataset to include disaggregated data on consumer and producer prices.

Besides national data, we augment the dataset with international monthly data to cover information for Portugal's main trading partners. These include Spain, France, Germany, United Kingdom, United States, Netherlands, Italy and Belgium. Data for Angola are scarce, hence this country was disregarded from the dataset. The monthly dataset spans january-2000 to december-2016.

For each trading partner, the panel of variables includes the main quantitative measures of economic activity, as well as qualitative assessments amounting, on average, to 80 series per country and to 766 series overall. The series were selected to represent broadly business and consumer surveys of economic activity, prices, retail trade, manufacturing and services and labour market. The splitting of the number of variables into hard and soft data is provided in Table 2.[2]

	Number of series	**Soft data**	**Hard data**
Portugal	145	39	106
Spain	82	41	41
France	81	40	41
Germany	80	39	41
United Kingdom	80	39	41
United States	75	20	55
The Netherlands	77	39	38
Italy	80	39	41
Belgium	66	41	25
Total	766	337	429

TABLE 2. Composition of the dataset.

2. A list of all series and data sources is available from the authors upon request.

In the case of Portugal and Spain, for a limited number of series we resort to the Expectation-Maximization algorithm suggested by Stock and Watson (2002a) to balance the dataset at the beginning of the sample period, since some series were available over a shorter time span. In general, with the exception of survey data, logarithms were taken for all non-negative series that were not already in rates or percentage units. Most series were differenced to achieve stationarity. Additionally, the series were further screened for outliers, where the adjustment corresponded to replace observations of the transformed series with absolute deviations exceeding six times the interquartile range by the median value of the preceding five observations, following Stock and Watson (2005).

Forecasting exercise

We begin this section with a detailed description of the design of the forecasting exercise. This entails fully recursive parameter estimation and factor estimation after the selection of the targeted predictors using the LARS-EN algorithm. Thus, we do not restrict the set of targeted predictors to be the same for each time period. Instead, predictors are selected at each point in time for each horizon and the forecasting equation is re-estimated after the new factors are estimated. We also consider the case where no pre-selection of predictors is applied, i.e., using the standard factor model approach.

As the benchmark model, we consider the usual AR(p) with the number p of autoregressive terms determined by the standard BIC criterion. We augment this model to account for calendar effects with three deterministic variables: the number of working days in each month and two dummy variables for the two moving holidays, Easter and Carnival.

An out-of-sample exercise is conducted to assess the relative performance of the factor model with targeted predictors against the benchmark. The number of estimated factors to be included in the forecasting equation is determined by minimizing a modified version of the BIC criterion suggested by Stock and Watson (1998).

The out-of-sample period spans january-2009 to december-2016, corresponding to half of the available sample period and the forecasting exercise is based on rolling window estimation with a window size equal to 96 monthly observations (8 years), which coincides with the typical average length of the business cycle. Rolling window estimation enhances model flexibility and time-varying parameters to cope with potential varying predictive content of the dataset. All the potential predictors are available for time t when exports of goods are also known. However, in the case of soft data, when exports for $t-1$ are released, data for period t are known. Hence, when considering only soft data, one can consider nowcasting besides forecasting from 1 to 12-month ahead.

Model performance is evaluated using the Mean-Squared Forecast Error (MSFE) and we compute the relative MSFE using the augmented autoregressive model as the benchmark. Hence, a ratio lower than unity means that the competing model outperforms the benchmark. We evaluate to what extent the forecasting accuracy gains are statistically significant through the Clark and West (2007) test procedure.

In the empirical analysis to follow, we examine two alternative panels, where different datasets are considered. First, we analyze soft data driven forecasts. Thus, we exploit survey-based indicators for Portugal and for its main trading partners. Secondly, we extend the dataset, so that hard and soft data for the countries are pooled for the forecasting exercise.

Soft data driven forecasts

Table 3 reports the forecasting results with soft-based datasets with targeted predictors, i.e., with LARS-EN pre-selection setting $\lambda_2 = 0.25$ as in Bai and Ng (2008) and Schumacher (2009), and with no pre-selection of predictors. In case pre-selection of predictors has been applied, the number of chosen predictors is discretized in each row, $N_{\mathcal{A}} = \{30, 40, ..., 150\}$. Entries in the bottom of the table correspond to the case where no pre-selection is applied. The underlying datasets comprise only soft indicators amounting to 337 series when considering all countries and 39 series when resorting only on national data. Entries in the table refer to the relative MSFEs of the factor model *vis-à-vis* the augmented univariate autoregressive forecast for different forecast horizons. Shaded entries denote the minimum relative MSFE for each forecast horizon.

	Forecast horizon												
	0	1	2	3	4	5	6	7	8	9	10	11	12
Targeted predictors ($N_{\mathcal{A}}$)													
30	1.12	1.06	0.89	0.81	0.83	0.90	0.89	1.04	0.81	0.90	1.23	1.04	1.17
40	1.15	1.03	0.71	0.73	0.71	1.06	0.66	0.91	0.85	0.93	0.89	0.71	1.01
50	0.95	0.96	0.64	0.64	0.73	1.11	0.53	0.81	0.90	0.83	0.76	0.69	1.17
60	0.92	0.83	0.63	0.70	0.73	1.08	0.49	0.70	0.83	0.77	0.81	0.65	1.04
70	0.91	0.84	0.63	0.66	0.73	0.97	0.50	0.69	0.92	0.67	0.85	0.53	0.90
80	0.94	0.87	0.66	0.69	0.76	1.17	0.51	0.77	0.97	0.76	0.96	0.59	0.82
90	0.87	0.88	0.68	0.72	0.78	0.91	0.56	0.83	1.10	0.82	0.97	0.63	0.97
100	0.93	0.94	0.65	0.77	0.65	0.81	0.69	0.81	1.09	0.96	0.88	0.66	0.99
110	0.96	0.84	0.69	0.75	0.51	0.91	0.63	0.82	1.03	0.98	0.79	0.69	0.99
120	0.89	0.94	0.70	0.73	0.50	0.94	0.72	0.96	0.95	1.04	0.83	1.03	0.92
130	0.95	0.92	0.74	0.70	0.52	1.07	0.84	1.09	0.97	0.88	0.86	1.15	0.81
140	0.96	0.89	0.73	0.71	0.52	1.10	0.93	0.96	1.06	0.97	1.03	1.21	0.87
150	0.91	0.85	0.72	0.71	0.58	1.14	1.02	0.90	1.05	0.98	1.13	1.35	0.93
No pre-selection													
All series	0.89	0.82	0.71	0.68	0.83	1.39	1.60	1.85	2.19	1.39	1.43	1.54	1.42
PT series only	0.78	0.86	0.87	0.78	0.84	0.97	0.96	1.07	0.94	1.07	1.05	2.14	2.45

TABLE 3. Relative MSFE of soft data driven forecasts *vis-à-vis* the benchmark.

A quick overview of the results reveals most of the entries are below one, showing that there are, in general, forecasting gains using factor models *vis-à-vis* the benchmark.

When nowcasting, greater forecasting gains are achieved by using national soft data only and with no pre-selection of predictors, and these exceed 20 per cent. This may reflect that data from trading partners convey more informational content for forecasting purposes. In this regard, forecasting accuracy gains are noteworthy when soft data for Portugal's trading partners are exploited and these gains are further enhanced if LARS-EN pre-selection of predictors is applied. For forecasting 1-month ahead, although the maximum gain is near 20 per cent attained with all series, i.e., without pre-selection, a similar figure can be delivered considering only 60 targeted predictors. For the forecast horizons from 2- up to 12-month ahead, the use of targeted predictors is consistently a dominant strategy delivering gains ranging from almost 20 per cent up to around 50 per cent *vis-à-vis* the benchmark. In general, the best forecast performance is achieved with no more than 70 variables chosen out of 337 and the forecasting accuracy gains are statistically significant over the forecast horizons.

To shed some light on the composition of the set of targeted predictors used for factor estimation, Figures 3 and 4 provide plots for the average share of targeted predictors from each country and sectoral survey, respectively, for different number of predictors (N_A) and forecast horizons (h). By looking at Figure 3, one can see that the average share of selected series from Portugal increases with both the forecast horizon and the number of selected predictors, going from less than 5 per cent to more than 15 per cent. Focusing on the most important trading partners, the same holds broadly for Spain. In the case of France, the average share of series decreases with the forecast horizon and with the number of predictors, going from around 30 per cent to less than 10 per cent. In turn, for Germany, the average share is particularly important for shorter horizons (around 15 per cent) and less relevant for horizons close to one year, being relatively stable across the number of selected predictors.

In Figure 4, one can see that the manufacturing survey variables are very relevant to forecast at shorter horizons, attaining around 50 per cent for a small number of predictors, and its importance decreases with the forecast horizon and number of predictors. In contrast, when considering consumers' survey, the share increases with the forecast horizon and to a less extent with the number of predictors, reaching around 40 per cent. Shares in the remaining surveys are relatively stable, with services representing around 10 per cent and retail trade and building around 15 per cent.

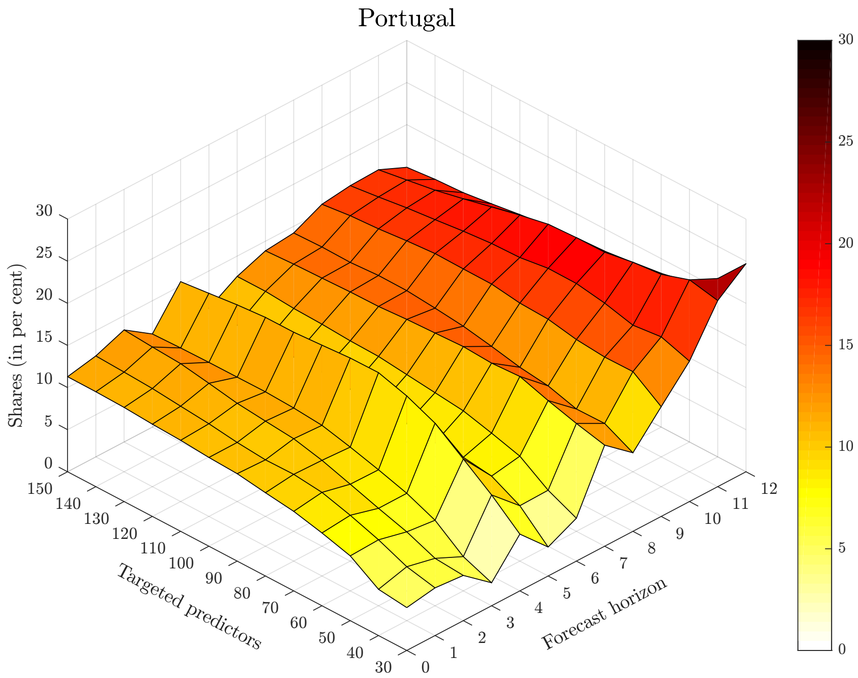


FIGURE 3: Average share of targeted predictors from each country for different number of predictors and forecast horizons.

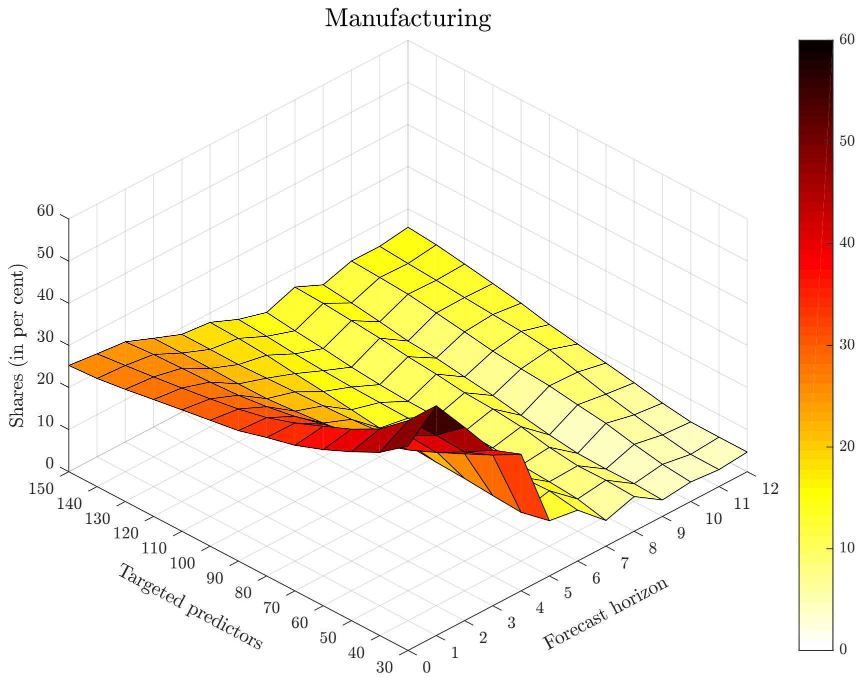


FIGURE 4: Average share of targeted predictors from each survey for different number of predictors and forecast horizons.

The EN algorithm involves the choice of λ_2, which penalizes with the L_2-norm of β. As in Bai and Ng (2008), we considered $\lambda_2 = \{0.5, 1.5\}$ as a robustness check. We find that the results are not very sensitive to this choice, which is in line with the findings by Bai and Ng (2008) and Schumacher (2009).

Pooling hard and soft data

We now extend the analysis where hard and soft data are pooled. As such, we exploit hard indicators for Portugal and its main trading partners, besides the previouly used soft-based dataset. The results are reported in Table 4. A quick glance at the results shows that the inclusion of hard data does not seem to bring additional predictive power for the longer horizons *vis-à-vis* the benchmark. In turn, the forecasting gains at shorter horizons are, in general, similar to those obtained when one resorts solely on soft data. One should note that LARS-EN pre-selection of predictors enhances forecast accuracy in comparison to no pre-selection and seems to play a role when forecasting at shorter horizons.

	Forecast horizon											
	1	2	3	4	5	6	7	8	9	10	11	12
Targeted predictors ($N_{\mathcal{A}}$)												
30	1.12	0.90	1.17	0.88	1.25	1.31	1.33	1.50	0.94	1.18	1.82	1.16
40	1.10	0.92	1.12	0.76	1.18	0.84	1.18	1.50	0.95	1.22	1.12	1.15
50	0.98	0.87	1.10	0.77	1.17	0.82	1.27	1.05	1.36	1.34	1.36	1.13
60	0.98	0.77	0.99	0.76	1.16	0.95	0.78	1.24	1.20	1.33	1.44	1.25
70	0.84	0.67	0.98	0.72	1.52	0.87	0.92	1.30	1.24	1.56	1.37	1.27
80	0.91	0.56	1.05	0.64	1.47	0.89	0.86	1.40	1.32	1.67	1.37	1.42
90	0.83	0.57	1.06	0.58	1.60	0.96	0.89	1.44	1.45	1.88	1.57	1.48
100	0.98	0.53	1.15	0.53	1.68	1.08	1.04	1.32	1.28	1.64	1.69	1.61
110	0.86	0.61	1.14	0.49	1.49	1.26	1.01	1.42	1.48	1.67	1.62	1.50
120	0.98	0.69	1.00	0.57	1.40	1.30	1.19	1.56	1.44	1.63	1.68	1.61
130	0.97	0.68	1.01	0.62	1.37	1.46	1.30	1.67	1.39	1.75	1.76	1.72
140	0.90	0.74	1.12	0.64	1.36	1.58	1.44	1.62	1.43	1.87	1.97	2.01
150	0.92	0.72	1.07	0.66	1.47	1.69	1.46	1.80	1.72	2.08	2.05	1.97
No pre-selection												
All series	0.86	0.82	1.18	1.46	2.26	2.43	2.73	3.78	4.06	3.57	4.28	4.64
PT series only	0.92	0.87	1.15	1.52	1.96	1.40	2.57	2.53	2.10	3.22	2.00	1.46

TABLE 4. Relative MSFE of soft and hard data driven forecasts *vis-à-vis* the benchmark.

Concluding remarks

This article exploits the role of international datasets for forecasting in a data-rich environment the Portuguese exports of goods on a monthly basis. Drawing on the informational content of the country's main trading partners, we document noteworthy forecasting gains up to 12-month ahead when soft indicators for these countries are pooled and predictors are pre-selected prior to factor estimation through the LARS-EN algorithm. In general, the best forecast performance is achieved with no more than 70 variables chosen. We find that forecasting accuracy gains delivered by factor forecasts using targeted predictors are statistically significant. Moreover, pooling hard data with soft data does not seem to bring additional predictive power for forecasting exports of goods.

References

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin, and G. Rünstler (2011). "Short-term forecasts of euro area GDP growth." *Econometrics Journal*, 14, C25–C44.

Bańbura, M. and G. Rünstler (2011). "A look into the factor model black box: Publication lags and the role of hard and soft data in forecasting GDP." *International Journal of Forecasting*, 27, 333–346.

Bai, J. and S. Ng (2008). "Forecasting economic time series using targeted predictors." *Journal of Econometrics*, 146, 304–317.

Boivin, J. and S. Ng (2006). "Are more data always better for factor analysis?" *Journal of Econometrics*, 132, 169–194.

Brisson, M., B. Campbell, and J. W. Galbraith (2003). "Forecasting Some Low-predictability Time Series Using Diffusion Indices." *Journal of Forecasting*, 22, 515–531.

Cardoso, F. and C. Duarte (2006). "The Use of Qualitative Information for Forecasting Exports." *Banco de Portugal, Economic Bulletin, Winter 2006*, pp. 67–74.

Clark, T. E. and K. D. West (2007). "Approximately normal tests for equal predictive accuracy in nested models." *Journal of Econometrics*, 138, 291–311.

Dias, F., M. Pinheiro, and A. Rua (2015). "Forecasting Portuguese GDP with factor models: Pre- and post-crisis evidence." *Economic Modelling*, 44, 266–272.

Dias, F., M. Pinheiro, and A. Rua (2018). "A bottom-up approach for forecasting GDP in a data-rich environment." *Applied Economics Letters*, 25, 718–723.

Efron, B., T. Hastie, I. Johnstone, and R. Tibshirani (2004). "Least angle regression." *Annals of Statistics*, 32, 407–499.

Esteves, P. S. and A. Rua (2012). "Short-Term Forecasting For The Portuguese Economy: A Methodological Overview." *Banco de Portugal, Economic Bulletin, Autumn 2012*, pp. 137–149.

Hansson, J., P. Jansson, and M. Löf (2005). "Business survey data: Do they help in forecasting GDP growth?" *International Journal of Forecasting*, 21, 377–389.

Li, J. and W. Chen (2014). "Forecasting macroeconomic time series: LASSO-based approaches and their forecast combinations with dynamic factor models." *International Journal of Forecasting*, 30, 996–1015.

Marcellino, M., J. Stock, and M. Watson (2003). "Macroeconomic forecasting in the euro area: country specific versus euro wide information." *European Economic Review*, 47, 1–18.

Rünstler, G., K. Barhoumi, S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitiene, P. Jelonek, A. Rua, K. Ruth, and C. Van Nieuwenhuyze (2009). "Short-Term Forecasting of GDP Using Large Datasets: A Pseudo Real-Time Forecast Evaluation Exercise." *Journal of Forecasting*, 28, 595–611.

Schumacher, C. (2007). "Forecasting German GDP Using Alternative Factor Models Based on Large Datasets." *Journal of Forecasting*, 26, 271–302.

Schumacher, C. (2009). "Factor forecasting using international targeted predictors: The case of German GDP." *Economic Letters*, 107, 95–98.

Stock, J. and M. Watson (1998). "Diffusion indexes." *NBER Working Paper no. 6702*.

Stock, J. and M. Watson (1999). "Forecasting inflation." *Journal of Monetary Economics*, 44, 293–335.

Stock, J. and M. Watson (2002a). "Macroeconomic forecasting using diffusion indices." *Journal of Business and Economics Statistics*, 20, 147–162.

Stock, J. and M. Watson (2002b). "Forecasting using principal components from a large number of predictors." *Journal of the American Statistical Association*, 97, 1167–1179.

Stock, J. and M. Watson (2005). "Implications of dynamic factor models for VAR analysis." *NBER Working Paper Series 11467*.

Tibshirani, R. (1996). "Regression Shrinkage and Selection via the LASSO." *Journal of Royal Statistical Society Series B*, 58, 267–288.

Zou, H. and T. Hastie (2005). "Regularization and variable selection via the elastic net." *Journal of Royal Statistical Society Series B*, 67, 301–320.

www.bportugal.pt