BANCO DE PORTUGAL
EUROSYSTEM

BANCO DE PORTUGAL ECONOMIC STUDIES

1

volume I

1
volume I

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, 2015 · www.bportugal.pt

BANCO DE PORTUGAL ECONOMIC STUDIES | Volume I - n.º 1 | Lisbon 2015 • Banco de Portugal Av. Almirante Reis, 71 | 1150-012 Lisboa • www.bportugal.pt • Edition Economics and Research Department • Design Administrative Services Department | Editing and Publishing Unit • ISSN 2183-5217 (online)

Content

Editorial

May 2015

We are currently in the midst of transforming the communication strategy of the economic research done at Banco de Portugal, whose main feature is the creation of products suited for each of the different audiences interested in that research. The creation of the *Banco de Portugal Economic Studies*, where signed articles comprising of finalized research projects will be published, is another step in this transformation. The journal is a companion, therefore, with the series of Working Papers – a collection of scientific works in progress aimed for future publication in international scientific journals – and the Economic Bulletin, which no longer will contain signed articles.

This inaugural issue of the *Economic Studies* contains three empirical articles, themselves revealing the scope of interests that catch the attention of Banco de Portugal's economists.

The long-term inflation expectations are critical for the conduction of monetary policy. The belief that these were apparently "anchored" substantially below 2% was an important aspect behind the recent policy shifts by the ECB. António Antunes's article, "Co-movement of revisions in short- and long-term inflation expectations," considers to what extent long-term expectations can really be considered well-anchored, in the sense that their revisions do not co-move with the fluctuations of short-run inflation expectations. The data for inflation expectations used in this article were obtained through the market prices of zero-coupon inflation swaps; the observations for short-term expectations were given by the expected inflation within a year for the following year, and the observations for long-term expectations were the expected inflation within five years for the following five years. Based on this data and modelling the co-movement between the two series using copulas – which are mathematical objects connecting two or more probability distributions – the article concludes that after 2012 long-term expectations could not be considered "anchored" because they tended to co-move with the extreme revisions of short-term expectations.

The second article, "Determinants of civil litigation in Portugal" by Manuel Pereira and Lara Wemans, seeks to analyze the determinants of the litigation rate, in particular for civil litigation. The authors construct a panel database for the civil law area, including information on incoming, resolved, and pending cases in 210 comarcas between 1993 and 2013. The inter- regional variation permitted identifying which socioeconomic characteristics – such as purchasing power, education level, or the number of enterprises – tend to increase the litigation rate. On the other hand, the length of the proceedings tends to reduce litigation.

Given the innate delays with which statistical information about production and its components is ascertained, the conjunctural assessment of the economy must be based on variables which are promptly observed and whose relation to GDP has certain desirable characteristics. The coincident indicators, which Banco de Portugal has published for almost three decades, take on an important role in the conjunctural analysis of the economy, seeking to synthesize in a single indicator a vast amount of information that at times can present contradictory developments. António Rua, in the article "Revising the monthly coincident indicators of Banco de Portugal," shows that the indicators used by Banco de Portugal to follow GDP and private consumption have been reliable even during turning points in the economic cycle.

Co-movement of revisions in short- and long-term inflation expectations

António Armando Antunes
Banco de Portugal and NOVA SBE

May 2015

Abstract
This article studies the co-movement between large daily revisions of short- and long-term inflation expectations using copulas. The main findings are: first, the co-movement between unusually large changes in short- and long-term inflation expectations increased markedly since mid-2012, which implies that long-term inflation expectations might not be, in a precise sense, well-anchored. Second, this co-movement measure is quite noisy. Finally, the result is shown not to be an artifact of the methodology or of the specific data used in the analysis. (JEL: C14, C46, G12)

Introduction

Market-based inflation expectations are widely used by market participants and policymakers for decision making and for inferring the likely monetary policy decisions of central banks. The alternative survey-based inflation expectations are also widely used but, for the purposes of this article, are not suitable given the lower frequency of available data. Market-based inflation expectations can be determined in several ways but perhaps the most popular method resorts to market prices of zero-coupon inflation swaps. These financial instruments are composed of a fixed leg and a variable leg and can be used to wedge against inflation fluctuations. For example, suppose that market participant A wants to insure herself against inflation fluctuations for holding a nominal asset for a period of five years starting from now. She can enter a zero-coupon inflation swap contract in the following terms: at the end of the five years, she receives the actual change in the relevant inflation index, which in the euro area can be for example the HIPC excluding tobacco, times the notional amount of the contract. This exactly compensates her for the changes of opposite sign in the real value of the nominal asset. At the same time, she pays the fixed leg of the contract to counter-party B, which is determined using the fixed rate

Acknowledgements: I thank Ildeberta Abreu, Rafael Barbosa, Nikolay Iskrev and Paulo Rodrigues, as well as participants in internal seminars, for useful discussions and help.
E-mail: aantunes@bportugal.pt

compounded for five years. Only one cash flow is exchanged at maturity, but the position can be closed at any moment by selling the contract in the market. The rate of the fixed leg of the contract is the expected inflation rate for the next five years. In fact, B enters the contract only if she believes that the fixed leg rate is going to be at least the actual inflation rate at maturity. On the other hand, A enters the contract only if she believes that the actual inflation is going to be at least the fixed leg rate. Of course there are additional effects involved. In particular, because A is effectively wedged against inflation risk, B has to be compensated through an inflation risk premium.

Using market-based inflation rate expectations, this article assesses the co-movement between daily revisions in short- and long-term inflation expectations using copulas, a special class of multivariate distribution functions. The main advantage of using copulas lies in their simple implications in terms of dependence of random variables, especially in the tails of the distribution. This allows for an assessment of the degree with which changes in long-term inflation expectations co-move with large swings in short-term inflation expectations. Moreover, certain copulas allow one to distinguish between upward and downward revisions in expectations.

Policymakers often mention that long-term inflation expectations are "well-anchored". However, this expression can mean different things. Sometimes it refers to the fact that the level of expectations is hovering close to a commonly accepted target level. On other occasions, the expression simply asserts that revisions of short-term inflation expectations should not per se imply revisions of long-term inflation expectations. One implication of this is that revisions in short- and long-term inflation expectations should not co-move significantly. The two meanings are not equivalent and have distinct implications in terms of the suitable methods for investigating whether inflation expectations are well-anchored. While the first focuses on levels and calls for a more traditional time series analysis, the second suggests using methods with an emphasis on correlation and co-movement, and not necessarily keeping track of the level of the inflation expectations. This article adopts the second type of approach. Moreover, special attention is paid to large innovations in inflation expectations, as these are more likely to represent fundamental changes in expectations than normal market fluctuations of smaller magnitude.

In a world where the central bank is deemed credible by market participants and with perfectly anchored long-term inflation expectations, one would expect that large revisions of short-term inflation expectations displayed little co-movement with large revisions of long-term inflation expectations. For example, a sudden oil price drop implying a large revision downwards of the short-term inflation expectations should not imply a revision of the same magnitude (in relative terms) in long-term inflation expectations.

Likewise, if one observes large revisions in long-term inflation expectations when there are large revisions in the short-term expectations, then the idea that long-term inflation expectations are solidly anchored becomes less obvious. In the limit, if one were to observe a one-to-one co-movement between these two measures, surely inflation expectations would not be anchored: they would be reacting immediately and significantly to the same information that produced swings in short-term expectations, with potentially highly disruptive effects in the effectiveness of monetary policy.

There is a relatively large literature on this topic which uses high frequency data and focuses mostly on the effects of news on long-term inflation expectations. This literature usually looks at the possibility of occurrence of structural breaks in a context of regression analysis (see, for example, Gürkaynak *et al.* 2010; Galati *et al.* 2011; Nautz and Strohsal 2015). In this article it is assumed that news are incorporated both in short- and long-term expectations but, if long-term inflation expectations are well-anchored, the effect on them would be small, whereas the effect on the short-term ones would be large. This should induce a low degree of co-movement between inflation expectations at long and short horizons. Using estimated copulas, it is shown that co-movement between changes in short- and long-term inflation expectations increased since 2012. This is in contrast with the absence of any significant co-movement in the previous low inflation period of end-2009. Moreover, these effects are shown not to be an artifact of the data, as simulations with random permutations of data eliminate them. Tail dependence between revisions in short- and lagged long-term inflation expectations persists but only for lags of one or two days, especially in the most recent portion of the sample. Finally, different choices for short- and long-term inflation expectations do not change the results in any meaningful way. While noisy, the observed co-movement in large swings suggests an increasing likelihood that long-term inflation expectations might have become de-anchored.

Inflation expectations and co-movement

In this article inflation expectations are taken from zero-coupon inflation swap rates. In terms of notation, average inflation prevailing from now until five years from now, for example, is denoted by π_{5y0y}, average inflation prevailing from next year for the following three years is π_{3y1y}, and average inflation prevailing five years from now for the following five years is π_{5y5y}. There are restrictions among these values, and these restrictions allow one to compute all relevant expectations based only on zero-coupon inflation swap rates. For instance, if market participants are risk neutral in perfectly competitive and frictionless markets, the equality $(1+\pi_{5y0y})^5 = (1+\pi_{2y0y})^2(1+\pi_{3y2y})^3$ must hold. Notice how the two zero-coupon rates can be used to estimate π_{3y2y}.

Another example: $(1+\pi_{5y5y})^5 = (1+\pi_{4y5y})^4(1+\pi_{1y9y})$ must hold. In this article, the value of the short-term inflation expectation will be the expected inflation one year ahead for one year (π_{1y1y}), and the long-term inflation expectation measure will be defined in the period five years ahead for five years (π_{5y5y}).

Data

Data are daily from Bloomberg and span the period from 22Jun04 until 17Feb15. Figure 1 presents the evolution over time of the two chosen variables, π_{1y1y} and π_{5y5y}, as well as observed inflation at monthly frequency. Table 1 presents summary statistics of the levels and first differences of π_{1y1y} and π_{5y5y}, along other variables (see below). The first differences correspond to the daily revisions of long- and short-term inflation expectations and constitute the focus of this article. Table 2 displays contemporaneous correlations among these variables.

Variable	Obs.	Mean	Std. Dev.	Min.	Max.	Autocorr.
π_{1y1y}	2781	1.787	0.504	0.293	3.751	0.978
π_{5y5y}	2781	2.304	0.205	1.483	2.803	0.987
$\triangle\pi_{1y1y}$	2780	-0.001	0.105	-1.163	1.132	-0.418
$\triangle\pi_{5y5y}$	2780	0.000	0.033	-0.196	0.220	-0.267
x	2780	-0.005	0.999	-5.968	11.334	-0.065
y	2780	0.000	1.000	-7.368	5.507	-0.019
u	2780	0.500	0.289	0.000	1.000	-0.005
v	2780	0.500	0.289	0.000	1.000	0.028

TABLE 1. Summary statistics. Daily data for period 22Jun04–17Feb14. π_{1y1y} and π_{5y5y} are market-based inflation rates one year from now during one year and five years from now during 5 years, respectively, and $\triangle\pi_{1y1y}$ and $\triangle\pi_{5y5y}$ are first differences; x and y are $\triangle\pi_{1y1y}$ and $\triangle\pi_{5y5y}$ filtered through an AR(1) process for the conditional mean and a GARCH(1,1) for the conditional variance; u and v correspond to the empirical quantiles of variables x and y, respectively. Values for π_{1y1y}, π_{5y5y}, $\triangle\pi_{1y1y}$ and $\triangle\pi_{5y5y}$ in percentages, except autocorrelations. Values for x, y, u, v and autocorrelations in natural units.

Sources: Bloomberg and author's calculations.

From the summary statistics it is readily seen that, historically, short-term inflation expectations have lower mean and higher volatility than long-term inflation expectations. In first differences, this behavior carries through for volatility but not for the mean, as expected. Level variables have strong persistence. In first differences there is negative autocorrelation, suggesting that increases in inflation expectations are often followed by corrections in the next trading day. Contemporaneous correlation between revisions of short- and long-term inflation expectations is only -0.007.

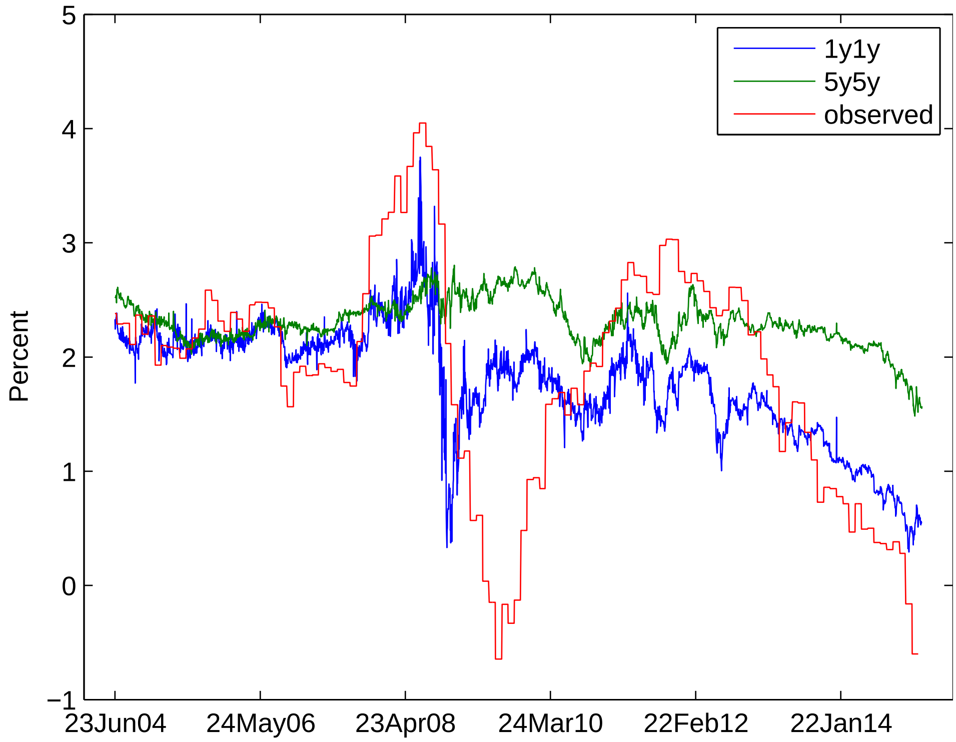


FIGURE 1: Market-based inflation rate expectations and observed inflation. Daily data for period 22Jun04–17Feb15. All values in percentage.

Source: Bloomberg.

	$\triangle\pi_{1y1y}$	$\triangle\pi_{5y5y}$	x	y	u	v
$\triangle\pi_{1y1y}$	1					
$\triangle\pi_{5y5y}$	-0.007	1				
x	0.761	0.047	1			
y	0.028	0.915	0.094	1		
u	0.681	0.049	0.893	0.088	1	
v	0.024	0.857	0.089	0.931	0.097	1

TABLE 2. Correlation matrix. Daily data for period 22Jun04–17Feb14. See legend of Table 1 for definitions of variables.

Sources: Bloomberg and author's calculations.

Conditional tail dependence

The study of co-movement between two random variables X and Y can be done in various ways. The first would be a simple correlation. This measure between -1 and 1 computes how X and Y co-move around their respective means. Sometimes this measure is enough for one's purposes. For example, the co-movement between two gaussian variables can be fully characterized by correlation. One problem with correlation as a measure of cross dependence is the fact that zero cross correlation does not in general imply independence. For example, the cross correlation between a random variable and its square is zero but they are clearly not independent. This in fact is a valid reason for not using correlation (or a linear regression coefficient) to study essentially unknown dependence among variables. Another problem with correlation is that it cannot be defined for certain distributions with heavy tails, as often is the case with financial returns (for examples of such distributions, see Resnick 2007).

An alternative way of studying co-movement between two variables is *conditional tail dependence*, and this is the focus of this article. To understand the notion of conditional tail dependence it is necessary first to define quantiles of a random variable. Quantile k of a random variable X is the value such that the probability of a random draw from X being less than or equal to that number is k. For example, the quantile 0.5 of a random variable is its median, and the interval defined by quantiles 0.025 and 0.975 is the 95% confidence interval for that random variable.

The idea of conditional tail dependence is simple: take values of variable X above a certain quantile k; compute the probability that the corresponding values of variable Y are above Y's quantile k; take the limit as k goes to 1. This is the so-called *upper tail dependence*. A similar measure can be computed for *lower tail dependence*, but in this case the limit is taken when k goes to 0. Intuitively, *this amounts to measuring the co-movement of two variables whenever either of them displays unusually large fluctuations*.[1]

This measure can be computed given the cumulative joint distribution function of the two variables, a function denoted by F. This function specifies the probability that a random realization of the two variables has both elements below the respective argument of F, so that for example $F(2, 1)$ is the probability that, in a random draw from the joint distribution of X and Y, the draw from X is lower than 2 and the draw from Y is lower than 1. The marginal cumulative distribution is the cumulative distribution of one of the variables unconditional on the other; for example $F_X(x) = F(x, +\infty)$ would be the marginal cumulative distribution of X.

One way to proceed would be to estimate some parametric form for distribution F and then compute tail dependence. In practice, however, such an estimation is difficult and suffers from frequent scale and domain problems in terms of variables X and Y. An easier route to compute conditional tail dependence is using copulas.

1. See Appendix A for formal definition of tail dependence.

Copulas: intuition

Copulas are a special class of cumulative distribution functions; see Patton (2006b) for the etymology of this designation and a rationale for the use of copulas in practical applications, and Nelsen (2006) and Patton (2012) for a detailed exposition of the theory and practical aspects of copulas. The distinguishing features of a copula are two: first, its underlying random variables are defined in the $[0, 1]$ interval; second, its marginal distributions are those of an uniform distribution. Using a copula involves specifying marginal cumulative distribution functions of each random variable along with a function (that is, the copula) that connects them. In this way, the researcher can separate the modeling of the marginal distributions from the dependence between the two variables. The copula specification implies a certain shape for the dependence between the marginal distributions. In the case where the copula is the product of the two marginal cumulative distribution functions, the two variables are independent and one can separately estimate each marginal. Otherwise, one can efficiently resort to the estimation of the joint distribution using a copula. Since the copula captures dependence structures for any shape of the marginal cumulative distribution functions,[2] the copula approach to modeling related variables can be very useful from an estimation perspective.[3]

Data transformation

As with many distribution functions, copulas can be fitted to the data using maximum likelihood methods. However, inflation rate expectations do not necessarily have to lie on the interval between 0 and 1, as required by copulas, nor do they exhibit temporally uncorrelated behavior. In order to clean up data, in this analysis the original data, π_{1y1y} and π_{5y5y}, will be transformed in three steps. First, the variables of interest (daily revisions) are obtained by computing the first differences of the levels, yielding $\triangle\pi_{1y1y}$ and $\triangle\pi_{5y5y}$.

Second, because the sole interest of the analysis is dependence between variables, and to avoid spurious dependence stemming from persistence or heteroscedasticity, the resulting variables are filtered through an AR(1) model for the conditional mean and a GARCH(1,1) specification for the variance (for a similar approach, see for example Christoffersen *et al.* 2012). This yields standardized daily revisions in inflation expectations x and y, respectively for $\triangle\pi_{1y1y}$ and $\triangle\pi_{5y5y}$.

2. In fact, the dependence between the two distributions is, using a copula, invariant to monotonic transformations of the two random variables.

3. For a brief exposition of basic copula theory, as well as the notion of a dynamic copula, see Appendix B.

Third, standardized daily revisions in inflation expectations are mapped into numbers between 0 and 1 so that the resulting variables can be used to fit a copula. This is done through the computation of an empirical marginal cumulative distribution function. More specifically, take the time series of, say, the standardized revisions in inflation expectations one year ahead for one year, that is, the collection $\{x_t\}_{t=1,...,T}$. Then there is a certain empirical marginal cumulative distribution function $\tilde{F}_X$ so that $u_t = \tilde{F}_X(x_t)$. (This function is an empirical, non-parametric counterpart to F_X.) Do a similar procedure for the long-term inflation expectations, y. Figure 2 represents the two empirical distribution functions. Variables u and v thus obtained are by construction approximately uniformly distributed.

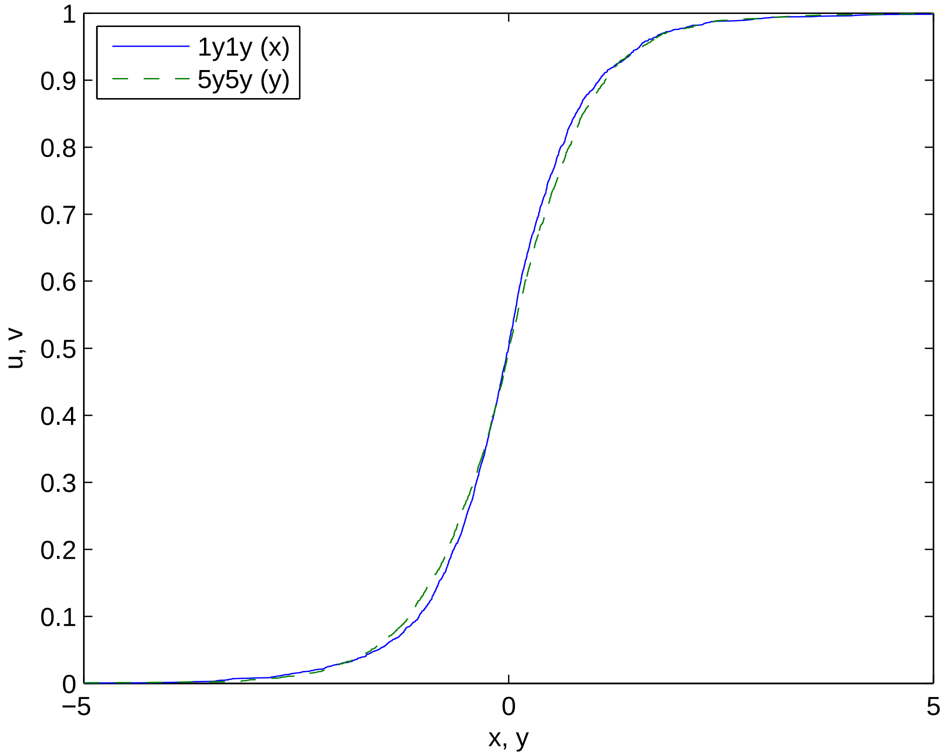


FIGURE 2: Empirical cumulative marginal distribution functions of x and y, the revisions of inflation expectations 1y1y and 5y5y standardized through applying an AR(1) conditional mean model and a GARCH(1,1) conditional variance model to the daily revisions of level variables.

Sources: Bloomberg and author's calculations.

Figures 3 and 4 present the daily innovations in inflation expectations, the standardized series and the uniform variables for the two variables of interest. Notice that there is substantial heteroscedasticity in both $\triangle\pi_{1y1y}$ and

$\triangle\pi_{5y5y}$, even though the latter exhibits less volatility, as previously seen. Heteroscedasticity is effectively removed by applying the filter mentioned above in both variables. Finally, the uniform transformations of x and y exhibit the expected behavior. Figure 5 shows a detail (observations during 2014) of the evolution of x and y.

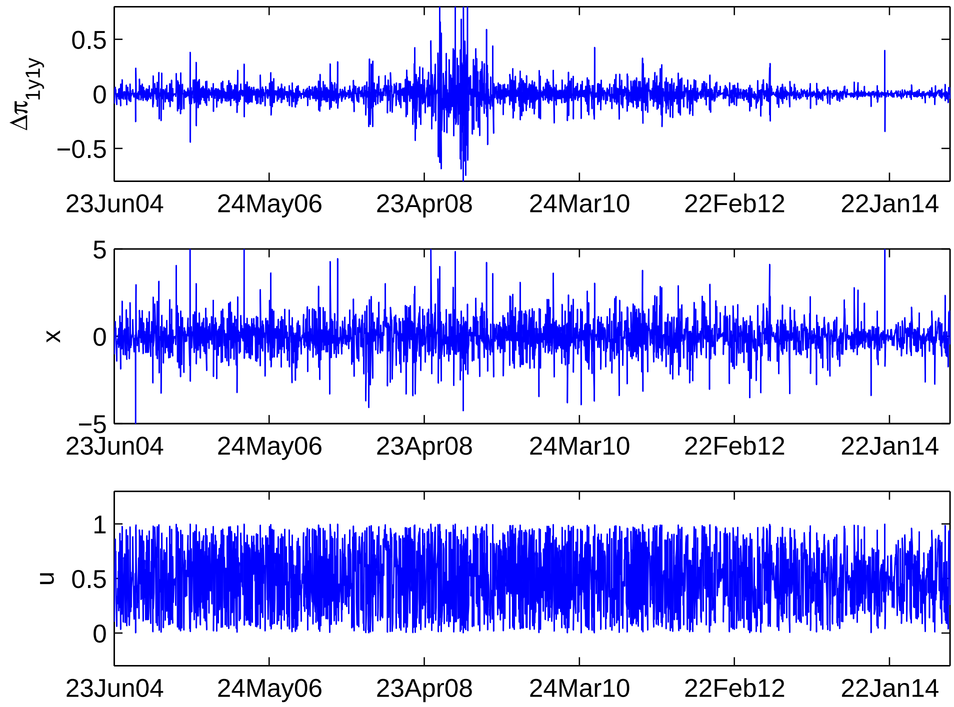


FIGURE 3: Evolution of $\triangle\pi_{1y1y}$, x and u. See legend of Table 1 for definitions of variables.

Sources: Bloomberg and author's calculations.

Going back to Tables 1 and 2, it can be seen that autocorrelation is mostly removed through the application of the AR(1) and GARCH(1,1) filters to the first differences of inflation expectations. Moreover, revisions of short- and long-term variables display relatively low contemporaneous correlation: the highest is u with v (0.097).

Results

The analysis consists of estimating several types of copulas in rolling windows of roughly one year, at the beginning of each quarter, and computing the

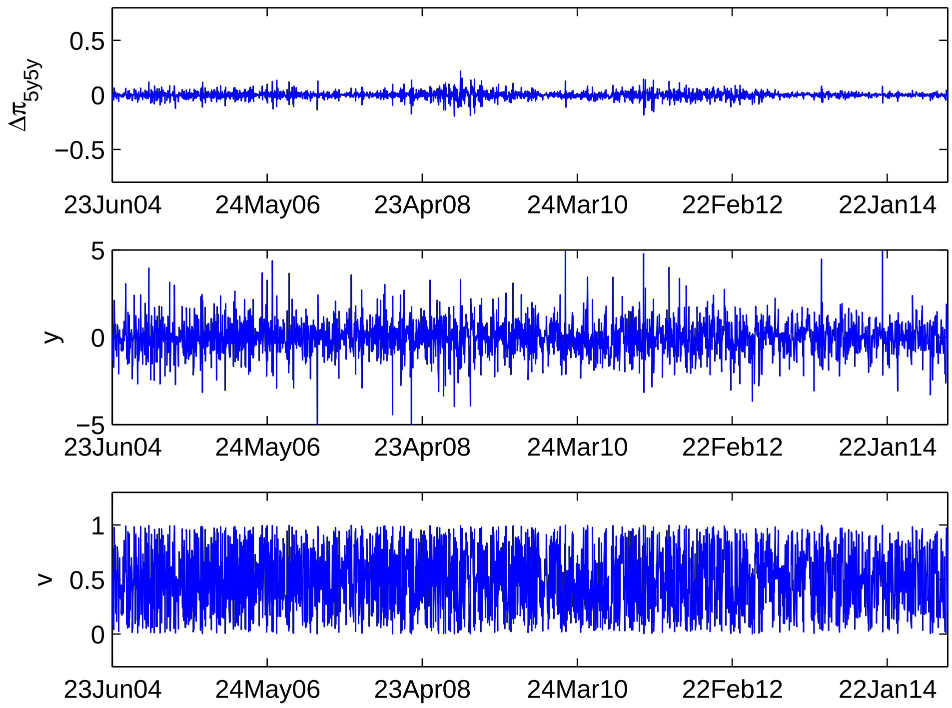


FIGURE 4: Evolution of $\triangle\pi_{5y5y}$, y and v. See legend of Table 1 for definitions of variables.

Sources: Bloomberg and author's calculations.

implied tail dependence. The estimated copulas differ in their parametric functional forms and, hence, in their characteristics in terms of symmetry and tail dependence.[4] A set of additional exercises and tests was also conducted but will only be briefly mentioned here.

Before looking at the evolution of tail dependence, a selection procedure was followed in which several different copulas were estimated. See Trivedi and Zimmer (2005) and Patton (2004, 2006a,b) and references therein for full descriptions of each copula. Table 3 summarizes the results. The ranking criterion was the number of times a copula is the best performer in each of the 39 quarters of the sample as measured by the value of its likelihood function. Under this criterion, the Student's t copula is the best performer, followed by

4. See Appendix B for a parametric example of a copula and references therein for full descriptions of copulas used in this section.

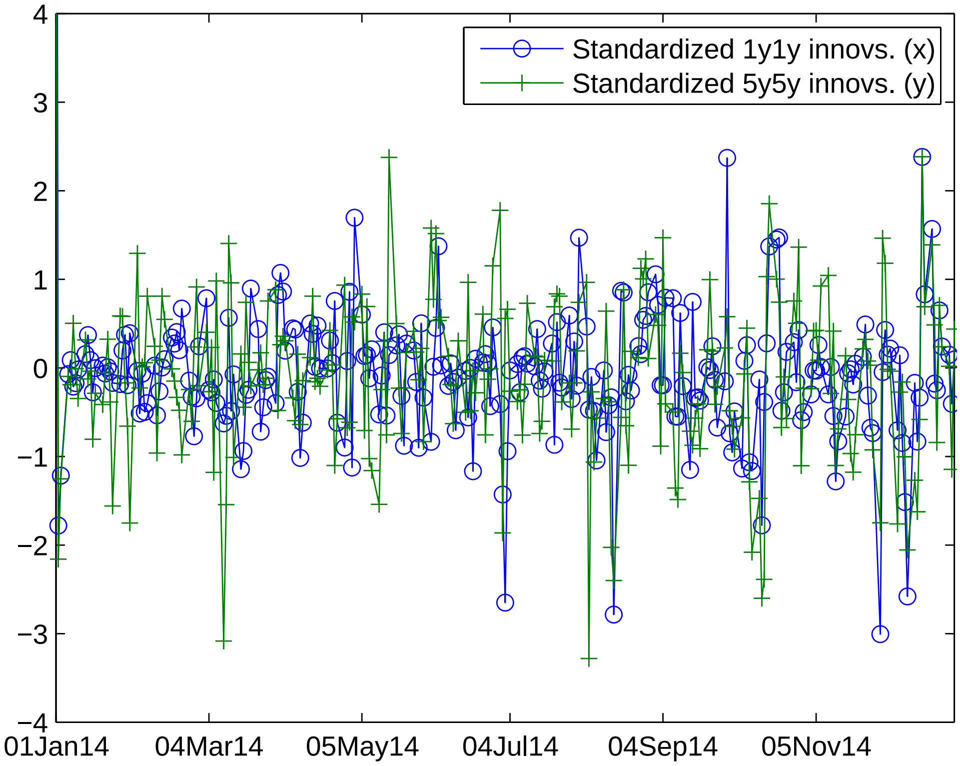


FIGURE 5: Evolution of x and y during 2014. See legend of Table 1 for definitions of variables.

Sources: Bloomberg and author's calculations.

the Normal, the Symmetrized Joe-Clayton (SJC), the Gumbel and the Rotated Gumbel.

At the beginning of each quarter, a copula was estimated using the available data of the previous 350 calendar days. The results are presented in Figures 6–8. The shaded areas are 90 percent confidence bands obtained through a bootstrap procedure (see Patton 2012). Looking at the results of Student's t copula (Figure 6), two features stand out. First, tail dependence is a noisy measure. The results are noisy and this volatility of the measure is still visible in the quarterly estimations reported in the figure.

The second salient aspect is that tail dependence increased markedly towards the end of the sample. The start of the increase in tail dependence can be traced back to 2012. The average tail dependence until 12Q3 was 0.011, and from 12Q4 on was 0.138. This is in stark contrast with the absence of

Copula	Tail dependence	# of quarters in which it was best
Student's	Yes, symmetric	20
Normal	No	9
Symmetrized Joe-Clayton (SJC)	Yes	7
Gumbel	Yes, upper tail	3
Rotated Gumbel	Yes, lower tail	0

TABLE 3. Ranking of estimated copulas according to the number of quarters that each copula performs the best.

Source: author's calculations.

any significant tail co-movement during the low inflation period of end-2009, when a fall in oil prices induced a marked decrease in inflation.

The figure also depicts the correlation parameter.[5] While at first tail dependence is fairly small, there is a period when, while there is correlation between the two series, the distribution becomes approximately Normal and no tail dependence occurs. After that, in 12Q4, tail dependence starts increasing consistently.

Among the copulas displaying tail dependence, the second best performer is the SJC copula and from the results depicted in Figure 7 one can see that upper tail dependence was higher than lower tail dependence during most of the sample. This means that large positive revisions in short- and long-term inflation expectations were more likely to be associated than large negative revisions. Towards the end of the sample (14Q2) lower tail dependence increases markedly. It should be noted that, since highly volatile data are being used, the distinction between upward revisions and downward revisions is not so clear-cut as with, say, quarterly data. Indeed, even when there seems to exist a secular trend to lower inflation, when one looks at longer spans of time (like, for example, during 2014 in Figure 1) daily filtered data still looks like white noise (see for example the filtered series in Figure 5), as expected, and there are as many upswings as there are downswings.

For the Gumbel copula, tail dependence decisively exceeds the 0.1 mark from 12Q3 on, and climbs to 0.4 towards the end of the sample. The Rotated Gumbel results are similar and hence not shown.

The Normal copula also performs well, although it has zero tail dependence. That is not surprising because the Student's t copula (which nests

5. Student's t copula estimation involves two parameters: correlation and degrees of freedom. When the estimated degrees of freedom of the copula become large, the copula converges to the Normal copula and there is no tail dependence.

the Normal copula as a particular case) has many degrees of freedom in many quarters, and this makes it very similar to the Normal copula in those quarters.

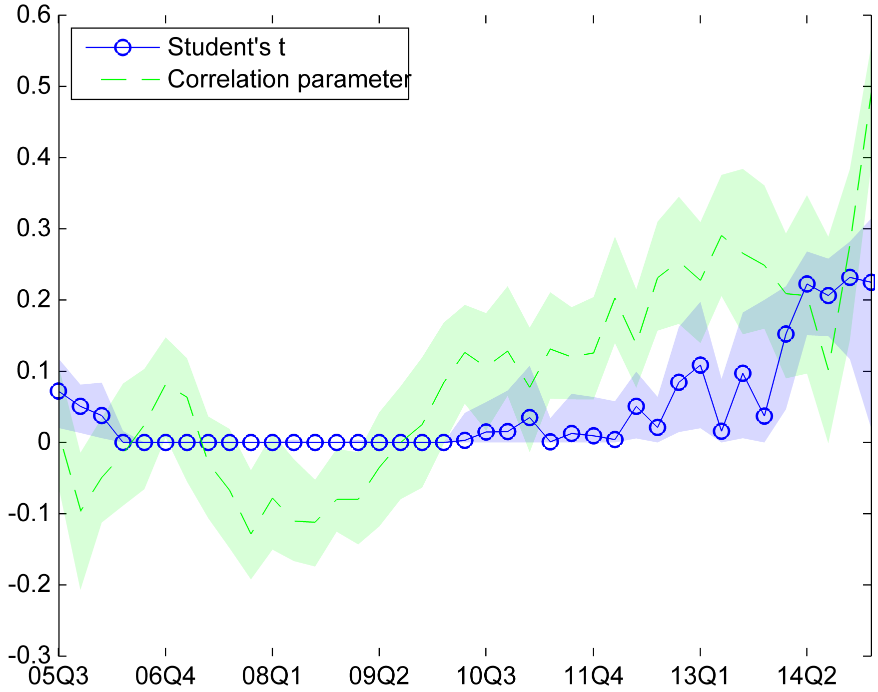


FIGURE 6: Tail dependence using estimated Student's t copulas at the beginning of each quarter using data from the previous 350 calendar days.

Source: author's calculations.

The general conclusion of the exercise is that the increase in tail dependence is very sharp since late 2012.

Additional exercises

Three additional exercises were performed.[6] The first is a robustness check where the whole procedure is repeated with a random permutation of time series $\{y_t\}_{t=1,\ldots,T}$ instead of the original series. The idea is to check whether there are artifacts of the data not related to co-movement that induce tail dependence. Given that the permutation should destroy all the time and cross dependence, one should observe essentially no tail dependence between the

6. Detailed results available upon request.

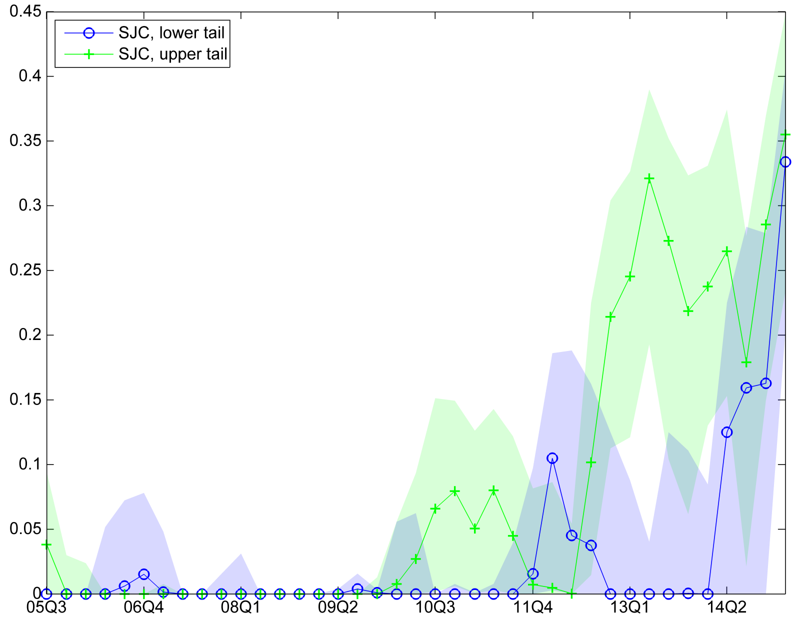


FIGURE 7: Tail dependence using estimated Symmetrized Joe-Clayton copulas at the beginning of each quarter using data from the previous 350 calendar days.

Source: author's calculations.

two variables. Indeed, the results show very low tail dependence throughout. The tail dependence parameters are found to be essentially zero. The second exercise was to perform the analysis with lags of one day and five days (which for this data set is one week) in variable y. The results for a 1-day lag display co-movement, although at a smaller level than the original estimates and concentrated in the final part of the sample. The co-movement dies out very fast and at a one-week lag it essentially has disappeared. All in all, this exercise suggests that there is time tail dependence at very short lags. The third exercise was to perform the analysis with different measures of short- and long-term inflation expectations, such as π_{2y1y} and π_{3y5y}. The results, however, remain essentially unaltered.

Concluding remarks

This article addresses the question of co-movement between revisions of short- and long-term inflation expectations. In particular, it focuses on a

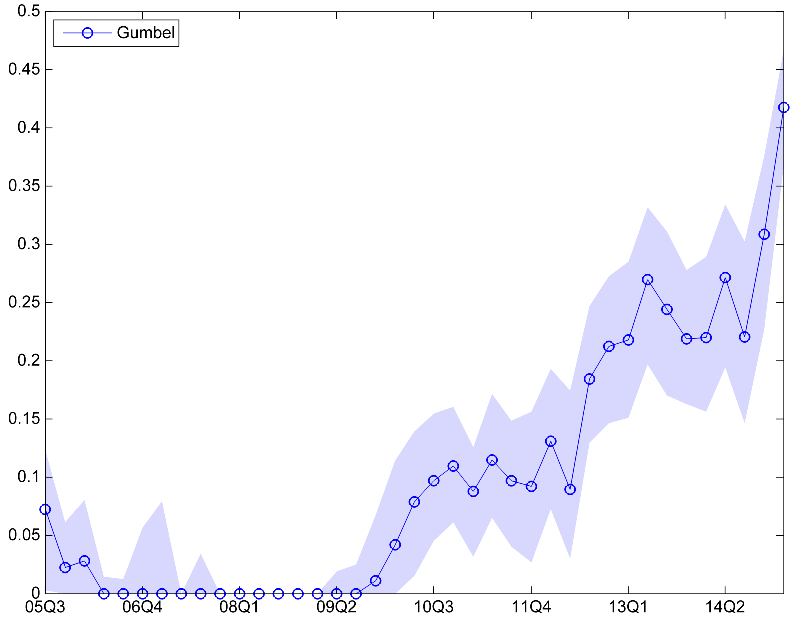


FIGURE 8: Upper tail dependence using estimated Gumbel copulas at the beginning of each quarter using data from the previous 350 calendar days.

Source: author's calculations.

measure called tail dependence, which looks at the probability that the two variables co-move when relatively large changes occur in one of them. Under the particular interpretation that inflation expectations are well-anchored when large innovations in short- and long-term inflation expectations do not co-move, this article shows that the case for well-anchored inflation expectations is not as strong since mid-2012 as it was before. This result is robust to different definitions of short- and long-term inflation expectations and does not seem to be an artifact of the data, produced for example by persistence or heteroscedasticity, and rapidly fades away when the data are not synchronous. Further work would include investigating the possibility of structural breaks in tail dependence in the context of copulas, and assessing the direction of causality, if any, in co-movement.

References

Braun, Valentin and Martin Grziska (2011). "Modeling Asymmetric Dependence of Financial Returns with Multivariate Dynamic Copulas." Mimeo, Goethe Universität, Frankfurt, and Ludwig Maximilian Universität, München.

Christoffersen, Peter, Vihang Errunza, Kris Jacobs, and Hugues Langlois (2012). "Is the Potential for International Diversification Disappearing? A Dynamic Copula Approach." *Review of Financial Studies*, 25(12), 3711–3751.

Galati, Gabriele, Steven Poelhekke, and Chen Zhou (2011). "Did the Crisis Affect Inflation Expectations?" *International Journal of Central Banking*, 7(1), 167–207.

Gürkaynak, Refet, Andrew Levin, and Eric Swanson (2010). "Does inflation targetting anchor long-run inflation expectations? Evidence from the US, UK, and Sweden." *Journal of the European Economic Association*, 8(6), 1208–1242.

Nautz, Dieter and Till Strohsal (2015). "Are US inflation expectations re-anchored?" *Economics Letters*, 127, 6–9.

Nelsen, R. (2006). *An Introduction to Copulas*. Second ed., Springer, New York.

Oh, Dong Hwan and Andrew J. Patton (2013). "Time-Varying Systemic Risk: Evidence from a Dynamic Copula Model of CDS Spreads." Mimeo, Duke University.

Patton, Andrew J. (2004). "On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation." *Journal of Financial Econometrics*, 2(1), 130–168.

Patton, Andrew J. (2006a). "Estimation of Multivariate Models for Time Series of Possibly Different Lengths." *Journal of Applied Econometrics*, 21, 147–173.

Patton, Andrew J. (2006b). "Modelling asymmetric exchange rate dependence." *International Economic Review*, 47(2), 527–556.

Patton, Andrew J. (2012). "A review of copula models for economic time series." *Journal of Multivariate Analysis*, 110, 4–18.

Resnick, Sidney I. (2007). *Heavy-tail phenomena: probabilistic and statistical modeling*. Springer.

Sklar, A. (1959). "Fonctions de répartition à N dimensions et leurs marges." *Publ. Inst. Statist. Univ. Paris*, 8, 229–231.

Sklar, A. (1973). "Random variables, joint distributions, and copulas." *Kybernetica*, 9, 449–460.

Trivedi, P. K. and D. M. Zimmer (2005). "Copula Modeling: An Introduction for Practitioners." *Foundations and Trends in Econometrics*, 1(1), 1–111.

Appendix A: Tail dependence

In this article, attention is restricted to the bivariate case; in most instances the theoretical generalization to the m-dimensional case is straightforward. It is useful to provide some theoretical background. Given two random variables X and Y, define the joint cumulative distribution function F as:

$$F(x, y) = \Pr\{X \leq x \text{ and } Y \leq y\} . \tag{A.1}$$

In order for F to qualify as a cumulative distribution function, it has to fulfill certain requirements. Intuitively, it is clear that F has to be 0 if any of its arguments is below the lowest value that the respective random variable can attain; it has to be 1 if all its arguments are higher than the highest value that each random variable can attain; and it must assign a non-negative value for the probability of any rectangle in its domain. Formally, these ideas would be expressed as $\lim_{x\to-\infty} F(x, y) = 0$ (and similarly for y), $\lim_{x,y\to+\infty} F(x, y) = 1$, and $F(x_1, y_1) + F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) \geq 0$ for any (x_1, y_1) and (x_2, y_2).

The one-dimensional margins are defined as $F_X(x) = \lim_{y\to+\infty} F(x, y)$ and $F_Y(y) = \lim_{x\to+\infty} F(x, y)$. Let x_k denote quantile k of variable X, that is, the value of x that solves equation $F_X(x) = k$, and similarly for y.[7] The *conditional upper tail dependence* is defined as

$$\lambda_U = \lim_{k\to1} \Pr\{y > y_k | x > x_k\} . \tag{A.2}$$

Similarly, it is possible to define the lower tail dependence λ_L taking the limit as k goes to zero and reversing the inequalities.

Appendix B: More about copulas

The first important characteristic of a copula is that its underlying random variables are defined in the $[0, 1]$ interval. The second important characteristic is that the copula's marginal distributions are uniform. Copulas are relevant because they connect multivariate distributions to their one-dimensional margins. Under pretty standard regularity conditions, a theorem due to Sklar (1959, 1973) states that there exists a copula C satisfying $F(x, y) = C(F_X(x), F_Y(y))$. In other words, any bi-dimensional cumulative distribution function can be decomposed into its marginal distributions and a copula. Moreover, the latter completely characterizes the dependence between the two variables. If the marginal cumulative distribution functions are continuous, this copula is unique.

7. The conditional cumulative distribution functions are, in case F is differentiable, $F_{X|Y}(x, y) = \frac{\partial F}{\partial y}(x, y)$ and $F_{Y|X}(x, y) = \frac{\partial F}{\partial x}(x, y)$.

One important consequence of this is that using the inverse of the marginal cumulative distribution function of X, F_X^{-1}, to transform a uniformly distributed variable in $[0,1]$, U, yields a variable that is distributed according to F_X. The same happens for Y and a uniformly distributed variable V in $[0,1]$. Therefore, $(F_X(x), F_Y(y))$ has copula C as its cumulative distribution function, and $(F_X^{-1}(u), F_Y^{-1}(v))$ has F as its cumulative distribution function. Because order relations in equation (A.2) are maintained between (x,y) and the corresponding uniformly distributed values (u,v), conditional tail dependence occurring for F will also occur for C.

Copulas turn out to be especially useful because tail dependence can be easily computed from their functional forms. Moreover, their domain fits nicely to the language of quantiles and percentiles necessary to study co-movement. There are a few notable copulas, some of which will be used in the body of this article. See Trivedi and Zimmer (2005) and Nelsen (2006) for a thorough exploration of different copulas and their properties. It is enough here to give just one example, which will be the Gumbel copula. Its expression is

$$C(u,v) = \exp\left(-\left((-\log(u))^{\theta} + (-\log(v))^{\theta}\right)^{\frac{1}{\theta}}\right), \tag{B.1}$$

where $\theta \in [1, +\infty]$. If θ is 1, the copula collapses to $C(u,v) = uv$, which is the case where variables are independent. If θ goes to $+\infty$, then $C(u,v) = \min\{u,v\}$, which corresponds to maximum dependence; this would imply correlation 1 between the two variables. This copula does not exhibit lower tail dependence, which may or may not be an obstacle to its utilization, but in turn can display arbitrarily large upper tail dependence. If one is interested in focusing on the co-movement between large upward revisions of short-term inflation expectations and long-term inflation expectations, then a Gumbel copula would be appropriate.[8] The formula above also allows for the computation of the upper tail dependence as expressed by equation (A.2); the result is $\lambda_U = 2 - 2^{\frac{1}{\theta}}$. As θ approaches 1 upper tail dependence approaches 0, which means no dependence; as θ approaches $+\infty$ upper tail dependence approaches 1, which means full correlation between the upper tails of the two variables. Figure B.1 provides a visual representation of the Gumbel copula for several levels of tail dependence: θ equal to 1, 1.3, 2.5 and $+\infty$, which entail upper tail dependence of 0, 0.3, 0.68 and 1, respectively. Several of the typical characteristics of copulas are evident. First, the marginal distributions are uniform, as can be seen from the straight line segments connecting (1,0,0) to (1,0,1), and (0,1,0) to (0,1,1). Second, as tail dependence increases from the independence case ($\theta = 1$) to the full correlation case ($\theta \to \infty$) the iso-probability curves (the "level curves" in the copula graph) go from hyperbolas

8. In fact, it is also possible to study lower tail dependence using the so-called Rotated Gumbel copula, whose expression is that in (B.1) with the arguments replaced by $1-u$ and $1-v$.

(with equation $k = uv$) to two segments connected at right angles at points such that $u = v$.

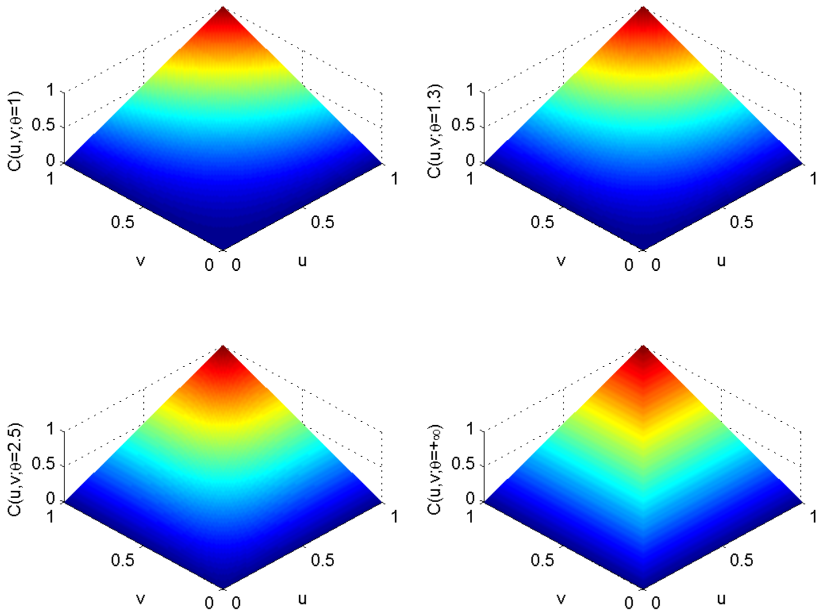


FIGURE B.1: Gumbel copula for several values of θ.

Source: author's calculations.

A last topic in terms of copulas concerns dynamic copulas. Dynamic copulas were first introduced by Patton (2006b) and are essentially the same as static copulas except that a subset of, or all, the parameters governing dependence is allowed to change over time. Patton (2006a), Braun and Grziska (2011) and Oh and Patton (2013) provide examples of dynamic copulas. The way in which parameters evolve over time is somehow arbitrary. Several dynamic copulas were also estimated for the data used here. The results do not differ significantly from those reported in this article and are available from the author upon request.

Determinants of civil litigation in Portugal

Manuel Coutinho Pereira
Banco de Portugal

Lara Wemans
Banco de Portugal

May 2015

Abstract
This article studies the evolution of the resort to civil justice in Portugal in the last two decades, particularly seeking to identify the main determinants of the litigation rate observed in the different regions, benefiting from a dataset with information by *comarca*. We conclude that the length of proceedings tends to reduce litigation and, therefore, there is evidence of rationing by waiting list in the access to justice. At the same time there is some evidence of demand inducement by lawyers. Socioeconomic characteristics as the illiteracy rate, purchasing power and the location of enterprises influence the level of litigation in the different regions of the country. Moreover there are significant spatial spillovers in the generation of litigation - not only the characteristics of the *comarca* itself, but also those of the neighbouring ones, play a relevant role. (JEL: K41, R10)

Introduction

The discussion about euro area growth potential, and in particular that of the countries most affected by the sovereign debt crisis, has assumed a central role in the last years. In this context, several international institutions have advocated the implementation of structural reforms of the judicial system as a way to promote competitiveness of the countries in the euro area periphery. The relation between economic growth and an adequate functioning of the judicial system has been recurrently addressed in the literature. More recent analyses include Lorenzano and Lucidi (2014), by the European Commission, and Palumbo *et al.* (2013), by the OECD. The proliferation of papers addressing this theme was stimulated by the substantial progress in the production and dissemination of international statistics in this field, specifically through the reports of the *Commission*

Acknowledgements: The authors thank Direção-Geral de Política da Justiça for providing the data on the judicial system and for valuable clarifications. The authors are also grateful for the comments made by Jorge Correia da Cunha, José Tavares, Manuela Espadaneira, Nuno Alves and Nuno Garoupa and by the participants in a seminar from the Economic Research Department of Banco de Portugal, especially Álvaro Novo and Pedro Portugal. 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

Européenne Pour l'Efficacité de la Justice (CEPEJ), an organization of the Council of Europe. In this vein, it is also important to mention the creation by the European Commission, in 2013, of the EU Justice Scoreboard. However, it is worth stressing that, despite the effort (namely by the CEPEJ) to improve data comparability, the large differences between legal systems hamper a direct comparison of synthetic indicators for different countries, requiring a critical analysis of the results.

The justice system plays a central role in modern societies, characterized by a multiplicity of social relations with a high degree of formality and by a widespread use of deferred payment methods. Regarding the impact of the functioning of justice systems on economic growth, several hypotheses have been explored by economic literature, primarily those related to the internalization in investment decisions of the benefits associated with the degree of predictability of court decisions and the ability to enforce contracts. The transmission mechanisms between the efficiency of the judicial system and economic growth have been discussed by studies which evaluate the impact of this efficiency on a wide range of economic indicators, including the size of companies (Posada and Mora-Sanguinetti 2013) and the functioning of credit markets (Jappelli *et al.* 2005). Note that the analyses of the justice system as a relevant factor for economic growth are typically centered on civil justice, also known as economic justice (Gouveia *et al.* 2012), as this area deals mainly with the resolution of economic disputes between private agents.

In the Portuguese case, the reform of the justice system has been a regular topic in the public debate. In fact, this was one of the areas covered by the Memorandum of Understanding, signed in 2011 in the framework of the Portuguese Economic and Financial Assistance Programme. In particular, the justice system was subject to deep organizational changes, notably through the implementation of the reform of the organization of judicial courts, in 2014. The pressure to implement reforms in this area in Portugal comes mainly from the unfavourable position of the country in most international comparisons related to the efficiency of the system. In addition, some studies indicate that the justice reform has a high potential to foster economic growth in the Portuguese case (Tavares 2004; European Commission 2014).

As regards the effectiveness of the Portuguese justice system, the CEPEJ report elaborated with 2012 data (CEPEJ 2014) highlights as main drawbacks the congestion and excessive length of civil proceedings in first instance courts. Conversely, the most recent analyses concluded for a comparatively better performance of higher instance courts (Garoupa and Pinheiro 2014a). These authors state that, in international comparisons about the performance of the judicial system, Portugal is often associated with countries with similar legal systems (also based on continental law), in line with the literature on the theory of legal origins (Porta *et al.* 2008). In this vein, the contribution by

Djankov *et al.* (2003) is noteworthy, stating that legal origins have an impact on the efficiency of the systems, notably through different degrees of formalism.[1]

The level of provision of civil justice can be seen as the result of an equilibrium which translates into the number of resolved cases, in a market where demand materialises as an inflow of cases, and supply corresponds to the services provided by the judicial system that is responsible for its resolution. According to this approach, reforms that try to tackle the congestion of civil justice can be divided essentially into two groups. Firstly, reforms which focus on supply through the expansion of resources allocated to the system or the reorganization of the functioning of courts. Secondly, policies that influence demand, particularly by changing the incentives faced by economic agents when filling cases.

This paper addresses the demand for civil justice, seeking to understand which factors influenced the number and territorial distribution of cases filed in first instance courts in Portugal between 1993 and 2013. The decision to file a case usually translates into an inflow at the level of this instance[2], while cases brought before higher courts mostly originate in this flow. Regarding the measurement of demand for justice, the information used identifies the cases that were filed in a particular territorial jurisdiction, which may not correspond to cases brought to court as a result of disputes occurring specifically in that territory (Gomes 2006). Given that the characteristics of a given geographical area may have an impact on the number of incoming cases in other areas, the demand for justice is modelled in this paper considering spatial interaction effects.

Several determinants of civil litigation have been addressed in the literature. For example, the costs of justice, including the rules to allocate them to different parties, and a number of institutional characteristics of the judicial system, such as its effectiveness as perceived by economic agents, the clarity of the law, the development of alternative dispute resolution mechanisms, among others (Palumbo *et al.* 2013). Other studies focus on determinants related to incentives of particular agents, as Carmignani and Giacomelli (2010), who discuss the possibility of demand inducement by lawyers. In addition, socioeconomic factors with an impact on the volume or complexity of economic transactions, as the sectoral composition of the economy, the level of schooling or the purchasing power, are equally mentioned as determinants of litigation, as they influence the type and volume of the litigation that is brought to court (Palumbo *et al.* 2013). Finally, it is important to highlight the effect of the economic cycle on the number of incoming civil cases. This paper discusses the importance, in the Portuguese context, of some

1. The index calculated by Djankov *et al.* (2003) places Portugal as a country with high formalism, even within civil law countries (as opposed to those based in common law).

2. For a description of the organization of the Portuguese judicial system, see Gouveia *et al.* (2012), volume I.

of these determinants, particularly those for which quantitative information was available, including several socioeconomic characteristics of *comarcas* and indicators linked to the judicial system, such as the length of proceedings and the concentration of lawyers.

A deeper understanding of the major factors behind the decision to file a civil case allows anticipating the effects of public policies in the area of justice and improving resource allocation to the different regions of the country. However, the literature in this vein focusing on the Portuguese reality is relatively scarce. Some of the few exceptions are the eminently descriptive work by Gomes (2006), along with Garoupa *et al.* (2006) who analyze the evolution of demand for justice using a time series, and Garcia *et al.* (2008) who investigate the territorial distribution of demand based on a panel for 2003 and 2004.

The paper is organized as follows. Firstly, the variables and the method used in the construction of the database by *comarca* are discussed. Secondly, a descriptive analysis of civil litigation in Portugal is presented, including a brief regional perspective. Thirdly, the econometric modelling of the determinants of litigation is addressed, with an emphasis both on factors that impact on the evolution of litigation over time and structural factors at the *comarca* level, considering contagion effects. Finally, some concluding remarks are made.

Data

As previously mentioned, this study is confined to civil justice which is the most relevant law area with respect to interactions with economic activity. As a result, cases related to criminal, labour and family law were excluded. Cases filed in administrative and tax courts were also disregarded.

A panel database for the civil law area was put together, including information on incoming, resolved[3] and pending cases in each *comarca*, between 1993 and 2013. The territorial organization of the Portuguese justice system includes, in descending hierarchical order, *distritos judiciais*, *círculos* and *comarcas*, the latter being its basic territorial unit (Gomes 2006). However, the geographical boundaries of *comarcas* were subject to several changes between 1993 and 2013. The most relevant ones include, firstly, the division of *comarcas* located in densely populated areas in the nineties and, more recently, several mergings as a result of the creation of pilot-*comarcas*, in the

3. In the statistics of justice a resolved case is defined as «a case in which a final decision has been reached, in the form of a judgement or order, in the respective instance and regardless of *res judicata*. Cases resolved at a given organizational unit also include the cases transferred or sent to another unit, where these are registered as incoming.» (Direção-Geral da Política de Justiça 2014, pp. 45, authors' translation)

context of the staged implementation of the 2008 law on the organization and functioning of judicial courts (Law No. 52/2008). In order to ensure time consistency, one considered 210 *comarcas* corresponding to the broadest territorial definition of each of them during the period under analysis.

The focus on the *comarca* level led to the exclusion of the civil cases filed in courts with a broader territorial scope. These courts can be divided into three types. Firstly, courts with jurisdiction over the whole territory, such as those related to competition, regulation and supervision (*Tribunal da Concorrência, Regulação e Supervisão*), to intellectual property (*Tribunal da Propriedade Intelectual*) and the maritime court (*Tribunal Marítimo*). Secondly, courts ruling in a particular law area and with a regional scope, as for instance, labour and family courts. Thirdly, the *tribunais de círculo*. It should be noted that, between 1993 and 2013, the configuration of the courts excluded from the sample changed often. For instance, the *tribunais de círculo* that covered a significant number of *comarcas* early in the sample were closed down in the late nineties. Furthermore, a number of labour and family courts were created and extinguished. In general, all these courts competed with those in the sample in terms of inflow of cases.[4] Nevertheless, the weight of the cases filed in these courts corresponded, on average, to only 5 per cent of total civil cases.

In the study of the determinants of litigation, only the cases that first enter a court matter, not those that move between courts (known as transferred cases). The information available in the statistics of justice allows the correction of incoming cases from the transferred ones at the national level. However, in the analysis by *comarca* this correction is unfeasible as there is information regarding the cases resolved through transfer at each court, but the court they were sent to is unknown. In order to overcome this constraint, one identified the situations in which, by the creation of new *comarcas* or new judgeships within a *comarca*, there was an unusually high number of cases resolved through transfer and it was possible to infer which *comarca* they had entered.[5] In such situations, the number of incoming cases was corrected from the transferred ones.

The database includes, in addition to the judicial system variables, several socioeconomic indicators published by INE at the municipal level. These include averages for the available time period of the purchasing power, population density, illiteracy rates, and the number of small and medium

4. The pilot-*comarcas* established in 2008 include specialised judgeships. The cases filed in these judgeships were also disregarded assuming that, before the creation of such *comarcas*, those cases would have been judged in courts outside the sample.

5. In these cases, monthly data supplied by the Direção-Geral de Política da Justiça was used to crosscheck if the movements of cases entered and resolved through transfer were consistent with the assumptions made.

enterprises and large enterprises *per* inhabitant.[6] Appendix A presents the details about this time period and the definition of the variables, including some descriptive statistics. In allocating the data by municipality to different *comarcas*, the major source of information was Gomes (2006, Appendix E), supplemented with Gouveia *et al.* (2012) with respect to the composition of pilot-*comarcas*. It was not possible to match the municipalities that spread throughout several *comarcas*, but this happened only in 10 out of 308 municipalities. The variables by *comarca* result from the average weighted by the population of the values for the corresponding municipalities.

Building on research applied to the Italian (Carmignani and Giacomelli 2010; Bounanno and Galizzi 2010) and Japanese (Ginsburg and Hoetker 2006) cases, suggesting that the concentration of lawyers increases litigation, one collected data regarding lawyers registered in the *círculo* corresponding to each *comarca*. The choice of a higher territorial level comes from a limitation of the data, since there is no information for *comarcas*. However, this territorial unit may even prove to be more appropriate, especially for regions with less litigation, where it is likely that a lawyer is not working exclusively in one *comarca*.

Information which could be useful to understand the evolution of total litigation in the last decades was also gathered. In this context, litigation growth is commonly related to the phenomenon of mass debt collection claims (Gomes 2006), associated with the filing of a significant number of similar claims by a reduced number of litigants. Indeed, technological development multiplied the number of services offered to the majority of the population through deferred payment methods. A clear example of this is the massification of contracts related to mobile telecommunications. There is no information on the number of cases filed by mass litigants. Nevertheless, some indirect information can be obtained from the list of commercial societies which generated a large number of cases in the recent period (published in accordance with *Portaria* No. 200/2011). An analysis of these enterprises shows that the financial sector along with water, electricity, gas and telecommunications suppliers are the most represented sectors. Therefore, data related to these sectors were collected, in particular the amount of non-performing loans to enterprises and households and the number of contracts for mobile telecommunications (published, respectively by Banco de Portugal and ANACOM). For these variables, only the country total and not their distribution by *comarca* is available.

Finally, a significant limitation is the scarcity of information regarding the cost of filing a case. This cost is one of the factors taken into account by economic agents in the decision making, even if it may have a smaller

6. In this paper the number of employees (above or below 250) is used as a criterion to define the two types of enterprises.

relevance in Portugal than in other countries. Indeed, according to the 2015 Doing Business report (World Bank 2014), the cost of enforcing a contract as a percentage of the value of the claim is particularly low in Portugal (13.8 per cent, which is the 12^{th} lowest in a list of more than 180 countries). There are virtually no indicators for the costs of access to justice in Portugal, except for the value of unit of account that is the benchmark for setting out court fees.[7] However, changes in the value of the unit of account explain only a fraction of the evolution of such fees, which are also influenced by revisions in the number of units of account that is due for the different proceedings. The most important change in this area during the sample period was enacted by the Decree-Law No. 34/2008, which revoked the *Código das Custas Judiciais* and introduced the *Regulamento das Custas Processuais*.[8] Other relevant indicator in this context would be the evolution of lawyer fees. This information would be particularly relevant as the aforementioned World Bank report states that these fees represent the majority of the costs to file a case in Portugal (10.7 per cent of the value of the claim). Data on lawyer fees are, however, virtually unavailable.[9]

Descriptive analysis of civil litigation

Evolution between 1993 and 2013

The number of incoming civil cases in first instance courts increased significantly between 1993 and 1997, rising from less than 300 thousand a year to above 450 thousand. In the following years, the demand for civil justice stabilized, presenting in 2013 a level similar to the one in 1997. In this period, the number of resolved cases stood generally below the number of incoming ones, explaining the continuous rise in pending cases (Figure 1).[10] This occurred every year except in 2006, 2007 and 2013, when administrative measures to reduce court congestion were implemented, as stated in Garoupa and Pinheiro (2014b). In this context, the litigation rate in Portugal (calculated as the ratio between incoming first instance civil cases and the population)

7. The unit of account (*unidade de conta*) is the index used in the calculation of the court fees to be paid by the different parties to a given dispute.

8. For an analysis of the impact of this change in revenues, see Correia and Joaquim (2013).

9. In the database *Quadros de Pessoal* it is possible to obtain the wages of lawyers working as employees. However, taking into account a survey conducted by the Bar Association in 2003 (Caetano 2003) and the number of observations available in that database, employees represent a small fraction (around 5 percent) of the total of lawyers, which strongly limits the use of this information.

10. For a detailed analysis of flows and main performance indicators of the Portuguese justice system, see the publications available on the website of Direção-Geral da Política de Justiça and also Gouveia *et al.* (2012).

presented a trend similar to that of civil litigation, increasing from 3.0 to 4.5 *per* 100 inhabitants from 1993 to 1997, and hovering around this value in subsequent years.

Figure 2 shows the breakdown of civil cases by main types, namely declarative, aiming at the definition of a particular right, and enforcement, intended to demand the fulfilment of a previously set obligation.

The relative stabilisation of the number of incoming cases after the end of the nineties was related to the generalisation of the injunction procedure, which allows the creditor to obtain an enforceable order, so as to require the recovery of a debt. Indeed, the legislative changes enacted over the period widened the scope of this procedure[11], increasing its use, inasmuch as it avoids the resource to declarative cases. This shift in the nature of injunctions is reflected on the strong growth of these procedures considered jointly with declarative actions, in particular from the late nineties on, in contrast with the sharp decline in the latter. In addition, the creation of the National Desk for Injunctions (*Balcão Nacional de Injunções*) in 2008 (*Portaria* No. 220-A/2008), implementing the dematerialization of this procedure, temporarily led to a considerable increase in the number of incoming injunctions. In turn, enforcement claims sustained a clear upward trend until 2003 and stabilized at around 200 thousand a year thereafter. Finally, the inflow of other types of cases including, for example, corporate reorganization and bankruptcy and credit claiming proceedings, went up, especially from 2006 onwards.

It is also important to consider how the resource to alternative dispute resolution and the Courts of Peace (*Julgados de Paz*) has evolved since these could be viewed as substitutes to civil litigation. However, the available data regarding arbitration and Courts of Peace signals that the evolution of these mechanisms is still at an early stage, with around 10 thousand incoming cases in 2013 in each of these mechanisms. Furthermore, the results from a survey to a group of Portuguese corporations presented in Gouveia *et al.* (2012) reinforce the idea of little use of these mechanisms, showing that in the previous three years only 5 per cent of the respondents were part to a dispute resorting to them. Strictly speaking, it is not clear that an increased use of such mechanisms leads to a reduction in incoming cases, arguing Garoupa and Pinheiro (2014a) that, in practice, they generate more litigation.

11. The injunction was introduced by the Decree-law (DL) No. 404/93 having as a limit a value equal to half of the lower bound for claims entering appeal courts, but its use was rather limited (see preamble to DL No. 269/98). The DL No. 269/98 doubled the aforementioned limit and removed procedural obstacles. Subsequently, the DL No. 32/2003 extended the scope to all late payments in commercial transactions, regardless of the amount owed, and the DL No. 107/2005 increased the limit to the lower bound for claims entering the Supreme Court of Justice (for more detail, see Gomes 2006).

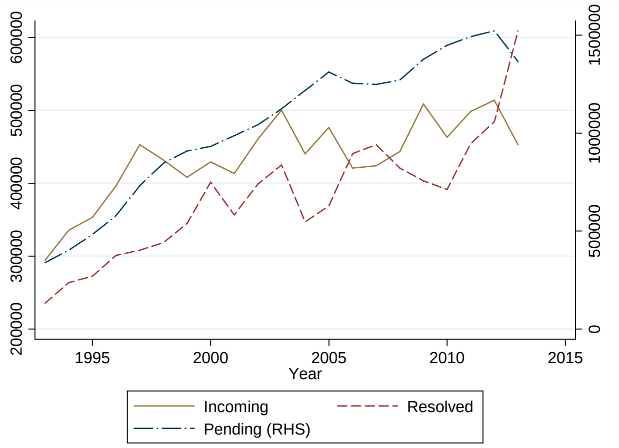


FIGURE 1: Incoming, resolved and pending civil justice cases

Note: Incoming and resolved cases do not include those transferred (moved between courts).

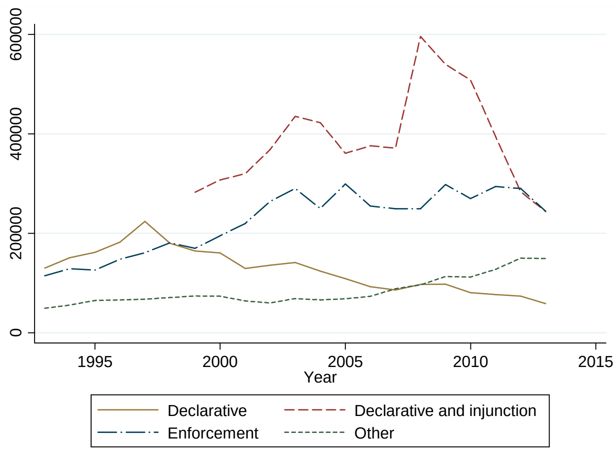


FIGURE 2: Incoming, pending and resolved civil justice cases including injunctions

International comparison

Litigation rates calculated with CEPEJ data for 2010 (CEPEJ 2012) are presented in Palumbo *et al.* (2013) for several advanced economies. These data are shown in table 1, including an update of this indicator for 2012, based on authors' calculations. It should be noted that, in order to ensure comparability of data internationally, the CEPEJ uses a classification of cases by judicial areas that differs from the classification adopted in Portugal. Thus, the litigation rate in table 1 considers not only incoming civil cases but also those related to labour and family, excluding, however, enforcement claims.

For 2010, the civil litigation rate in Portugal is close to that observed in France and Germany, placing Portugal among the countries with highest rates, but below the levels presented by Italy and Spain, for example. However, the replication of calculations for 2012 generates different results for some countries, showing that the comparison of data among countries with very dissimilar justice systems should be performed with caution.

	2010 a)	**2012 b)**
Finland	0.3	0.2
Norway	0.4	0.4
Luxembourg	0.4	0.9
Sweden	0.7	0.7
Denmark	1.3	0.8
Austria	1.3	1.2
Estonia	1.6	1.2
Poland	2.1	2.8
Hungary	2.2	4.4
Switzerland	2.2	2.9
Slovenia	2.2	1.8
Slovak Republic	2.4	3.0
France	2.8	2.6
Portugal	**3.0**	**3.5**
Germany	3.5	2.0
Italy	4.0	2.6
Greece	4.0	5.8
Spain	4.2	3.8
Czech Republic	4.5	3.5
Russia	9.6	4.5

TABLE 1. Litigation rates in different European countries (incoming cases for 100 inhabitants)

Sources: Palumbo *et al.* (2013) for 2010 and authors' calculations with data from CEPEJ and Eurostat for 2012.

Regional distribution

Litigation in the different regions of the country has presented a very heterogeneous evolution. During the period under analysis, there was a significant reduction in the average of incoming cases *per capita* in the two *comarcas* with the highest levels of litigation, Lisbon and Oporto.[12] By contrast, average litigation rates in other *comarcas*, both those located in coastal *círculos* and in inland *círculos*, showed an increasing trend between 1993 and 2013

12. Note that, taking into account the broadest territorial definition in force between 1993 and 2013, the *comarca* of Oporto comprises the *comarcas* of Oporto, Valongo, Gondomar and Maia, by reference to the judicial map of 2013.

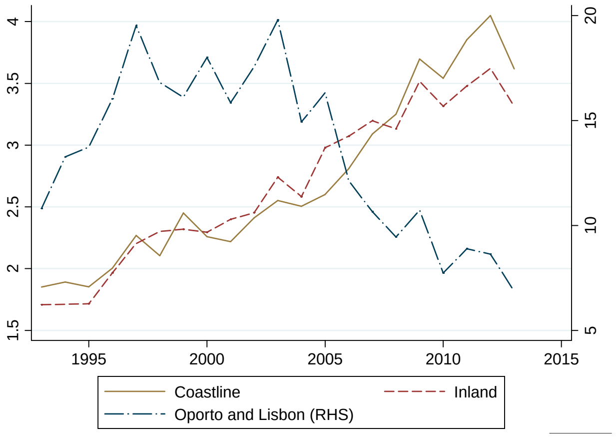


FIGURE 3: Litigation rates evolution in the coastline, inland and Lisbon and Oporto

Note: Inland (coastline) includes *comarcas* in inland (coastline) *círculos*, except Lisbon and Oporto.

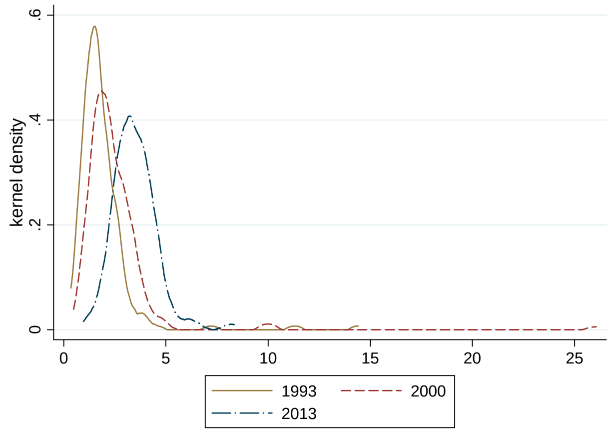


FIGURE 4: Distribution of civil litigation rates by *comarca* in 1993, 2000 and 2013

(Figure 3). Observing the distribution of litigation over time, there is clearly a trend of increased dispersion (Figure 4) which occurred simultaneously with the reduction of the concentration of cases in Lisbon and Oporto. While the number of cases filed in these two *comarcas* accounted for over 40 per cent of nationwide incoming cases until 2006, such percentage had already halved in 2010.

This trend is related to the entry into force of Law No. 14/2006, which amended the definition of the territorial jurisdiction of courts, imposing the residence of the defendant as a rule for actions relating to the fulfilment of obligations (except where both parties reside in the metropolitan areas of

Lisbon or Oporto or the defendant is a corporation, in which case it is possible to opt for the place where the obligation should have been fulfilled). As the head offices of large litigants are highly concentrated in Lisbon and Oporto (respectively, 67 and 14 per cent)[13], these *comarcas* are the main attractors of litigation related to the fulfilment of obligations. Therefore, the legislative amendment in question would have contributed to a lower concentration of cases in the two *comarcas* and, in general, to a greater geographical spread of litigation.

Determinants of litigation: panel regressions

Explanatory variables

The analysis of the determinants of litigation, measured by the log of the number of civil cases brought before courts *per capita* (*LitigRate*), is primarily carried out through econometric regressions including explanatory factors that vary simultaneously over time and among *comarcas* (equation (1)). The explanatory variables comprise, apart from the lagged litigation rate, the average length in months of cases resolved (*Length*) also lagged, for the length of cases resolved in the current year is not part of the information set of economic agents at the time the decision to file a case is taken. In addition, they include a variable meant to capture the change in the territorial jurisdiction of courts in 2006 (*Ref2006*). This variable takes on the value 1 for the period after the reform in the *comarcas* of Lisbon and Oporto, on the assumption that these would have been the most affected (see previous section). The log of incoming non-civil cases *per capita* (*NCLitigRate*) was also included, as one wants to model civil litigation stemming from other law areas. The validity of this regressor rests on the assumption that the converse effect, i.e. the existence of litigation in other areas originating in civil litigation, has a minor expression. This assumption is supported in particular by the significant proportion of civil cases in criminal judgeships, but not the opposite.

Lastly, to control for the impact of *tribunais de círculo* that existed for a limited period of time and covered about two-thirds of *comarcas*, the number of cases filed in those courts *per* 100 inhabitants (*CircCourts*) was included in the regression. One cannot control, correspondingly, for the cases filed in the other courts with jurisdiction beyond a single *comarca* but still of sub-national scope, because the available information does not allow to allocate them to a set of *comarcas*. In any case, these courts received, on average, only 4 percent of the total annual civil cases. Finally, with regard to national courts, their impact

13. Considering large litigants companies which generated more than 500 cases both in 2013 and in 2014.

on the inflow of cases in the courts in the database will tend to be captured by year fixed effects, as described below.

$$\begin{aligned} LitigRate_{i,t} = c + \beta_1 LitigRate_{i,t-1} + \beta_2 Length_{i,t-1} + \beta_3 Ref2006_{i,t} \\ + \beta_4 NCLitigRate_{i,t} + \beta_5 CircCourts_{i,t} + \alpha_i + \delta_t + \varepsilon_{i,t}, \end{aligned} \quad (1)$$

where i indexes the *comarca*, t the year and, in addition to the abovementioned variables, the regression includes fixed effects for *comarcas* (α_i), in order to capture their specific characteristics, and year fixed effects (δ_t), in order to control for the specificities of a given year with a cross-*comarca* impact. Year fixed effects may capture a multiplicity of factors, in particular, changes in the configuration of courts ruling in a particular law area and in the code of civil procedure (such as the spreading of injunctions, mentioned in the previous section), including changes in court fees and the effect of several factors such as the economic cycle and mass debt collection. In addition, the compilation of information for statistics of justice changed in 2007, when it started to be carried out directly from the courts' IT system. Year fixed effects will also capture any impact this change may have on the data, provided that it is felt across *comarcas*. Regressions were run not only for civil litigation as a whole, but also for declarative and enforcement cases separately, in order to understand if there could be different determinants.

Results

Table 2 presents the estimates of equation (1) by the Arellano and Bond (1991) estimator (see Appendix B for more details). The results show a negative relationship between the length of proceedings and the respective litigation rate, which may indicate the existence of a congestion effect. A negative impact of case length on litigation is plausible on the assumption that such variable is, at the time of decision making about starting a case, used as an indicator of its expected length, a factor that can be of great importance. Garoupa *et al.* (2006) find evidence of a positive relationship between the litigation rate and the expansion of the judicial system, measured by the number of judges. This is consistent with the evidence of rationing-by-queuing due to a negative relationship between the length of proceedings and the litigation rate. It is not excluded that the true impact of length is underestimated, for the coefficient estimate may be picking up another effect of this variable on litigation, of a positive sign, to the extent that the slowness of the judicial system can be detrimental to contract compliance. However, taking into account that this issue will depend on the overall perception about the efficiency of the system, it may also be captured by the year fixed effects.

By type of civil cases, the estimated length effect is stronger and more precise for enforcement ones. Such an evidence may result from the latter being, relative to declarative cases, more often filed by companies[14] that may have a better perception regarding the length of proceedings. The impact of litigation in other law areas on civil litigation is positive, reflecting the fact that criminal, labour and family cases give rise to civil cases. This variable may be also capturing the effect of common determinants (omitted in the equation) of civil and non-civil cases.

Explanatory variable	**Arellano-Bond estimator**			*Fixed effects estimator*
	Civil cases	**Declarative**	**Enforcement**	*Civil cases*
Litigation rate(t-1)	0.584***	0.569***	0.578***	0.503***
	(0.041)	(0.044)	(0.039)	(0.025)
Length of proceedings(t-1)	-0.012***	-0.004*	-0.008***	0.002
	(0.003)	(0.003)	(0.003)	(0.002)
2006 Reform	-0.753**	-0.846	-1.118***	-0.521***
	(0.312)	(0.533)	(0.399)	(0.057)
Tribunais de círculo	-0.06	-0.342***	-0.53	-0.06
	(0.054)	(0.124)	(0.397)	(0.05)
Non-civil litigation	*0.188****	*0.204****	0.231***	0.141***
	(0.027)	*(0.034)*	(0.039)	(0.015)
Hansen test (p-value)	0.312	*0.548*	0.299	-
N (Comarcas)	210	210	210	210
T (Years)	19	19	19	19

TABLE 2. Determinants of litigation: panel regressions

Notes: Estimates of equation (1), with the log of the respective litigation rate as the dependent variable, by the Arellano-Bond estimator, instrumenting the variables not strictly exogenous (*LitigRate(t-1)*, *Length(t-1)* and *CircCourts*) by their lags (2 to 6), in levels. The last column presents the results for the fixed effects estimator which assumes that all variables are strictly exogenous. Standard deviations are in parentheses. P-values: * <0.1; ** <0.05; *** <0.01.

The existence of *tribunais de círculo* has an impact for declarative cases only, indicating a smaller inflow in the *comarcas* under the jurisdiction of such courts. A greater relevance for declarative actions may stem from a larger weight of *tribunais de círculo* for them, which stood on average at 3.4 percent in the period 1993-1999, against 1.3 percent for enforcement actions. The results point to a lower concentration of cases in Lisbon and Oporto in the period the territorial jurisdiction of courts associated with the residence of the defendant was already in force. This redirected to other *comarcas* litigation relating to fulfilment of obligations where the defendant did not reside in the metropolitan areas of Lisbon or Oporto. Such an effect is only significant for enforcement actions that have been the most affected by this legislative

14. Gomes (2006) concludes that, on average, between 2000 and 2004 companies filed 63 percent of declarative cases and almost 90 percent of enforcement ones.

change. Note that the aim here is not to assess the impact of this reform, but essentially to control for its effects on the regional distribution of litigation.

The estimates of year fixed effects in equation (1) would ideally be regressed on a set of possible determinants of the evolution of litigation over time (an approach similar to that followed for the *comarca* fixed effects, in the next section). Nevertheless, having only 19 years of data makes this approach unfeasible. In an attempt to obtain some evidence on this issue, one analysed the correlation between the changes in litigation, measured by the estimates of year fixed effects, and some factors of potential relevance in this context. One considered the mass debt collection indicators mentioned in the section devoted to the data, namely the stock of non-performing loans to companies and to individuals and the number of contracts for the mobile telephone service, and the value of the unit of account used in the calculation of court fees. The correlations have the expected signal, i.e. positive for the first three variables and negative for the fourth, but lack statistical significance. One also considered the correlation with the real GDP growth rate, this being negative but statistically not significant as well. A counter-cyclical variation of incoming cases may reflect a greater tendency for contract-breaching with the deterioration in macroeconomic conditions. It should be noted, however, that economic growth and litigation may be positively associated in the long run, to the extent that an increase in the volume and complexity of economic relationships may contribute to greater litigation.

Determinants of litigation: structural characteristics of *comarcas*

Explanatory variables

Following the estimation of equation (1), the dependent variable is now the estimate of *comarca* fixed effects.[15] The explanatory variables are taken as the average for the available time frame and include: the number of lawyers *per capita* enrolled in the *círculo* the *comarca* belongs to (*Lawyers*), the number of small and medium (*SME*) and large enterprises (*LargEnterp*) in the *comarca* relative to the population, the population density (*PopDens*), the purchasing power index (*PPI*) and the illiteracy rate (*IllitRate*).

The literature modelling the determinants of demand for justice has not considered the existence of spatial interaction effects, not allowing the litigation in a given region to depend also on the characteristics of the surrounding regions. Indeed, to the extent that a party to a case does not

15. This estimate is obtained by averaging the composite residuals by *comarca*. More specifically, one takes the fixed effects not conditioning to litigation in the previous year, whose expression is given by $\alpha_i/(1-\beta_1)$, where β_1 is the coefficient of the lagged litigation rate. The composite residuals are an estimate of $(c+\alpha_i+\varepsilon_{i,t})$ in equation (1).

reside in the *comarca* where it is brought, there will be interdependence among *comarcas*, particularly neighbouring ones, in the determination of litigation. This phenomenon is modeled by means of a spatial econometric model (see, for example, Paelinck and Klaassen 1981; Anselin *et al.* 1995, and 2004) that uses information about the location of the geographical units in the sample (*comarcas*), summarized in a matrix of spatial weights. The construction of this matrix may follow several criteria, including a contiguity criterion, which would confine the interdependence to contiguous geographic units, or a distance criterion. In this paper the second option is chosen, inasmuch as it will primarily be the distance (irrespective of the existence of a common spatial boundary) to determine the intensity of spatial effects.

Spatial models may include spillover effects of various types, notably the so-called exogenous effects that in this context consist of making the litigation in a given *comarca* depend on the characteristics of the neighbouring ones (taking the distance as a weight). The estimation may also incorporate interaction effects in the error term, under the assumption that the omitted explanatory factors have the same kind of spatial dependence.[16] Equation (2) shows the specification to be estimated, including the abovementioned explanatory variables and also the spatially weighted averages of the values taken by each of these variables in other *comarcas*, except for the concentration of lawyers.[17] The same spatial correlation mechanism is assumed for the error term.

$$\begin{aligned} \hat{\alpha}_i &= c + \beta_1 Lawyers_k + \beta_2 SME_i + \beta_3 LargEnterp_i + \beta_4 PopDens_i \\ &\quad + \beta_5 PPI_i + \beta_6 IllitRate_i + \mathbf{w}_i \boldsymbol{X} \boldsymbol{\gamma} + e_i, \\ e_i &= \lambda \mathbf{w}_i \mathbf{e} + \varepsilon_i, \end{aligned} \tag{2}$$

where i indexes the *comarca* and k the *círculo*, $\hat{\alpha}_i$ is the estimate of the *comarca* fixed effect, $\mathbf{w}_i$ is the row of the matrix $\mathbf{W}$ of spatial weights that corresponds to *comarca* i, $\boldsymbol{X}$ is a matrix whose columns contain the covariates which were spatially interacted, and $\boldsymbol{\gamma}$ contains the coefficients of such covariates. The w_{ij} element of matrix $\mathbf{W}$ results from the inverse of the distance between the *comarcas* i and j of mainland[18] if $i \neq j$, and equals 0 if

16. It does not appear adequate to take on board endogenous interaction effects (that concern the dependent variable), since the decision by an agent to bring a case in a given *comarca* does not usually depend on similar decisions by other agents in other *comarcas* (once controlling for spatial interactions in observed and omitted explanatory variables). Note that the law prescribes the *comarca* in which a case must be brought and, even where there is a choice, this is much restricted.

17. As explained in the section devoted to the data, the variable *Lawyers* is defined for *círculos*, which are geographically much broader than *comarcas*.

18. The weights are determined so that the closest *comarcas* have a greater weight and their sum for a given *comarca* is equal to 1. The distance between *comarcas* that underlies this matrix is

$i = j$. The weight w_{ij} was truncated to 0 in the case of *comarcas* more than 100 km apart. One experimented with shorter distances for this truncation and also not imposing it, but the results did not change much. In the equation for the error term, λ is the spatial autocorrelation coefficient.

Results

Table 3 presents the results for the model with spatial spillovers, taking civil cases as a whole, and also declarative and enforcement ones separately. The last column presents for comparison the estimates for the model without spatial effects, i.e. imposing the restriction $\boldsymbol{\gamma} = 0$ and $\lambda = 0$ in equation (2) (see Appendix B for the estimation methods). Given that the explanatory variables are measured in quite different units (see table A.2 in Appendix A), the figure 5 shows the percentage impact on the civil litigation rate of changes in socioeconomic regressors on a more comparable basis (with the magnitude of 1 standard deviation). The results of this section concern the *comarcas* located in mainland Portugal only. Note that the inclusion of the *comarcas* on Azores and Madeira (in the model without spatial interactions) does not change the findings as to the variables that are statistically significant and respective coefficients.

The evidence presented in table 3 is clear about the joint significance of the covariates interacted with the spatial weights matrix. As to the presence of spatial autocorrelation in the errors, this is significant for enforcement claims only. The specification with spatial interactions is more informative, as it allows disentangling the geographical origin of the impacts. Indeed, the impact of a given factor on litigation may differ depending on whether it stems from the *comarca* concerned or the neighbouring ones (as shown below). As the neighbouring *comarcas* tend to have similar characteristics, there is a relatively strong and positive correlation in this regard.[19] Therefore, in the regression without spatial effects, the coefficients capture a mixture of impacts originating in the *comarca* itself and surrounding ones.

calculated as the Euclidean distance between the respective centroids, and the latter have been obtained from data by municipality available in the *Carta Administrativa Oficial de Portugal* of 2014 at the site of the *Direção-Geral do Território* (www.dgterritorio.pt).

19. The correlation between a given covariate (say, $\boldsymbol{x}$) and its version interacted with the spatial weights matrix ($\mathbf{W}\boldsymbol{x}$) ranges from 0.43 for the concentration of large companies to 0.70 for the illiteracy rate.

Explanatory variable	Civil cases	Declarative	Enforcement	*memo item: civil cases no spatial effects*
Constant	2.004***	1.144***	0.420	0.642***
	(0.234)	(0.24)	(0.403)	(0.105)
Small and medium enterprises	3.647***	2.850**	4.500***	-0.196
	(1.273)	(1.388)	(1.651)	(1.34)
Large enterprises	0.590*	0.706**	0.587	0.993**
	(0.326)	(0.354)	(0.427)	(0.437)
Purchasing power	0.273**	0.279**	0.326**	0.003
	(0.128)	(0.14)	(0.166)	(0.128)
Illiteracy rate	-1.355***	-0.190	-2.454***	-2.194***
	(0.47)	(0.51)	(0.616)	(0.397)
Population density	0.520	0.164	0.492	0.238
	(0.44)	(0.469)	(0.565)	(0.48)
Lawyers	0.801***	1.073***	0.987***	1.191***
	(0.166)	(0.174)	(0.206)	(0.219)
W*Small and medium enterprises	-22.522***	-12.503***	-27.634***	
	(4.788)	(4.666)	(7.982)	
W*Large enterprises	-2.392	-1.375	-2.498	
	(2.468)	(2.565)	(3.557)	
W*Purchasing power	1.615***	-0.051	2.175**	
	(0.551)	(0.546)	(0.892)	
W*Illiteracy rate	-5.052***	-10.026***	-1.204	
	(1.532)	(1.65)	(2.308)	
W*Population density	-6.870***	-5.080***	-6.421**	
	(1.789)	(1.742)	(2.993)	
Lambda	0.103	-0.065	0.500**	
	(0.286)	(0.329)	(0.214)	
Signif. of interactions - W's (p-value)	[0.00]	[0.00]	[0.00]	
N (comarcas)	192	192	192	192

TABLE 3. Impact of characteristics of *comarcas* on litigation

Notes: Results of the estimation of equation (2) for the mainland *comarcas*, with the *comarca* fixed effects as the dependent variable, by a maximum likelihood estimator (LeSage, 2004). The last column presents the results of the estimation by least squares of a model without spatial effects. Standard deviations are in parentheses. P-values: * <0.1; ** <0.05; *** <0.01.

The small and medium-sized enterprises located in the *comarca* appear as a litigation generator and, as far as declarative actions are concerned, this holds as well for large companies. Such an evidence is consistent with Gomes (2006) who concludes that between 2000 and 2004 about 3/4 of civil cases were filed by companies. By contrast litigation *per capita* varies negatively with the concentration of small and medium enterprises in neighbouring *comarcas* (which, in the regression without spatial effects, disturbs the estimate for the *comarca* itself). As for large companies, spatial interactions lack statistical significance. The negative sign of the relationship between litigation and the concentration of small and medium enterprises in neighbouring areas may be associated with the fact that the proportion of economic relationships within-*comarca* goes down with the increase of such concentration. Therefore part of the litigation that stems from these

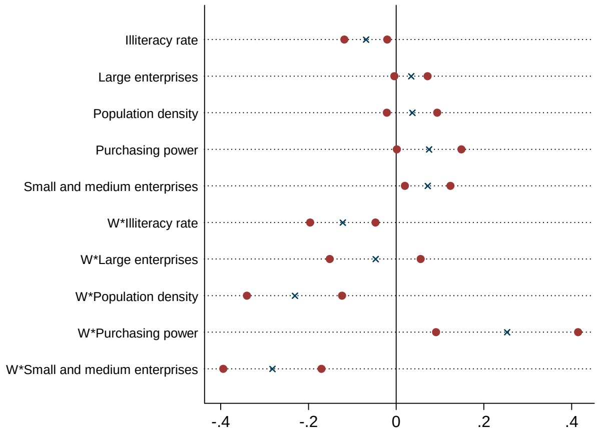


FIGURE 5: Percentage change in the civil litigation rate as a result of changes in the socioeconomic covariates of comparable magnitude

Note: One-standard-deviation increases in each regressor; the figure shows the point estimate (x in blue) and the confidence interval at 95 percent (circles in red).

relationships will be diverted to the surrounding *comarcas* (in particular, where they constitute the defendant's domicile). Moreover, concerning litigation between individuals and companies (assuming that the companies generally have the initiative, in this case), such result may also reflect a preference by companies for bringing cases in their own *comarcas*.[20]

In interpreting the impact of the concentration of large companies on litigation, it should be noted that this explanatory variable is not appropriate to capture mass debt collection disputes. Indeed, a significant part of mass litigants are large companies - mainly located in the *comarcas* of Lisbon and Oporto and in the service sector - but they make up a small part of this universe. In fact, it is not easy to study this phenomenon with the information available. In particular, a binary variable identifying the *comarcas* where the headquarters of at least one large litigant is located (see definition in footnote 13) is not suitable for this purpose, because it would work as a fixed effect for the most developed *comarcas* in the country.

In the specification with spatial spillovers, the *comarcas* with greater purchasing power have higher levels of civil litigation, which is not surprising given the likely association between disposable income and credit granting or deferred-payment contracts. Litigation in this type of legal relationships is mainly related to debt default and gives rise to enforcement actions, but also

20. This possibility was wider before the change in the territorial jurisdiction of courts in 2006. Even after this reform and for cases that concern the fulfilment of obligations, there is still a choice if the parties reside in the metropolitan areas of Lisbon and Oporto.

to declarative ones for debt recognition. An effect of neighbouring *comarcas* is, for the purchasing power, confined to enforcement claims (also showing up for the overall civil). In this case, the coefficient is positive, supporting a complementarity between *comarcas* in the generation of litigation, unlike the concentration of companies that brings about a substitution effect between them. In this regression, the purchasing power is capturing the general level of economic development of *comarcas*, as other indicators, such as the concentration of corporations, are held fixed.

A higher illiteracy rate has a negative impact on the civil litigation rate, as expected, given that *comarcas* with lower educational levels will feature less formal social relationships, and this may reduce the propensity to litigate. Such a result holds for enforcement actions but not for declarative ones. In this case there is also a complementarity in the estimated effects between *comarcas*.

It is interesting that some socioeconomic factors, such as the density of small and medium enterprises and purchasing power, appear to play a more important role for litigation when stemming from the surrounding *comarcas* than from the *comarca* itself, despite the uncertainty of the estimates. This importance of spatial spillovers probably reflects the small size of the basic territorial units underlying the organization of justice vis-à-vis the geographical extent of the transactions between economic agents. Such an evidence highlights the need to consider the environment surrounding each *comarca* in the definition of justice policies.

Impact of the concentration of lawyers on litigation

The results in table 3 support the hypothesis of a positive impact of the concentration of lawyers on litigation, both for declarative and enforcement actions. Demand inducement by lawyers is plausible in this market, because it concerns a service based on trust[21] which, owing to its high degree of technical complexity, is characterized by strongly asymmetric information between the provider and the client. This asymmetry extends to essential aspects in the decision to bring a case, as the expected length and probability of success. Such an evidence may also result from lower fees being charged in more competitive markets which, in turn, would make litigation financially more attractive. Indeed in Portugal, unlike in some other European countries (Palumbo *et al.* 2013), fees are not fixed by law, being only regulated by the Bar Association. Unfortunately, as mentioned in the section devoted to data, it was not possible to collect sound data on lawyer fees, and so this issue could not be pursued.

21. Such goods and services are called credence goods (Darby and Karni 1973). Dulleck and Kerschbamer (2006) feature a literature review about this topic, addressing the expected effects given the characteristics of each market.

Nevertheless, the evidence presented in table 3 has, as a major limitation, the likely endogeneity of the concentration of lawyers that should respond positively to the volume of litigation in the *comarca*. Accordingly, Appendix C presents the results when that variable is instrumented by the distance of a given *comarca* to the *comarca* where the nearest law college is located, a procedure also followed in Carmignani and Giacomelli (2010) (see Appendix B for details on the estimation). The use of this instrumental variable assumes that it is (negatively) correlated with the number of lawyers *per capita*, but it does not directly affect litigation. As regards the first assumption, there is a negative and statistically significant effect of the instrumental variable on the number of lawyers, controlling for the other explanatory variables in regression (2).

Firstly, it is interesting to compare the new estimates for the remaining variables with those in table 3, as the endogeneity of the concentration of lawyers could also bias them. The estimates do not significantly change, as regards the signs and magnitudes of the coefficients. However, some of the variables lose statistical significance, as the density of large companies and the purchasing power at the *comarca* level, which may reflect the increase in variance that is characteristic of instrumental variables estimators. The evidence of demand inducement by lawyers remains unchanged. This is consistent with the abovementioned studies (Carmignani and Giacomelli 2010; Bounanno and Galizzi 2010; Ginsburg and Hoetker 2006). Nevertheless, contrary to what one would expect, the estimated coefficient does not come down from table 3, even if the respective confidence intervals at 95 percent intersect. The results for the concentration of lawyers must thus be interpreted with caution, as there is uncertainty about an effective correction of its endogeneity in the second estimation.

Conclusions

This paper presents an econometric approach to the determinants of civil litigation in Portugal, using a panel database that covers 210 *comarcas* over a period of about 20 years. There is evidence that the decision to bring a case is negatively influenced by the length of proceedings, indicating that this variable plays a role in rationing access to justice. In addition, the inflow of non-civil cases (associated with labor, criminal and family law) has a positive impact on the inflow of civil cases, suggesting that the implementation of policy measures in a given law area will benefit from an integrated perspective of the system.

The results support the existence of important spillover effects between *comarcas*. Therefore, an assessment of resources allocated to a *comarca* should take into account the characteristics of the surrounding areas. The socioeconomic indicators considered indicate a positive relationship between the level of development and litigation, particularly visible for the illiteracy rate and purchasing power. There is also evidence that the location of companies is a strong attractor of litigation, and differences between neighbouring *comarcas* with regard to the concentration of small and medium enterprises divert litigation from each other. Finally, trying to correct for potential endogeneity issues, there is evidence of demand inducement by lawyers.

The findings of this study should be seen as complementary to those achieved by other scientific approaches. In this regard, a multidisciplinary analysis of justice policy issues appears beneficial. In terms of future research, it would be useful to compare the factors underlying the geographical distribution of litigation, studied in this article, with the changes in the territorial distribution of resources allocated to justice. Furthermore, it would be important to deepen the understanding of the impact of costs borne by litigants on the demand for justice, an aspect that could not be addressed due to data limitations.

References

Anderson, Theodore W. and Cheng Hsiao (1982). "Formulation and estimation of dynamic models using panel data." *Journal of Econometrics*, 18, 47–82.

Anselin, Luc, Raymond Florax, and Sergio Rey (eds.) (1995). *New directions in spatial econometrics: advances in spatial science*. Springer.

Anselin, Luc, Raymond Florax, and Sergio Rey (eds.) (2004). *Advances in spatial econometrics: methodology, tools and applications*. Springer.

Arellano, Manuel and Stephen Bond (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *Review of Economic Studies*, 58, 277–297.

Bounanno, Paolo and Matteo M. Galizzi (2010). "Advocatus, et non latro? Testing the supplier induced demand hypothesis for Italian courts of justice." *Working Papers Fondazione Eni Enrico Mattei*, (52).

Caetano, António (ed.) (2003). *Inquérito aos advogados portugueses: uma profissão em mudança. Ordem dos Advogados Portugueses*.

Carmignani, Amanda and Silvia Giacomelli (2010). "Too many laywers? Litigation in Italian civil courts." *Banca D'Italia Working papers*, (745).

CEPEJ (2012). "European judicial systems – Edition 2012 (data 2010)." *Council of Europe Publishing*.

CEPEJ (2014). "European judicial systems – Edition 2014 (data 2012)." *Council of Europe Publishing*.

Correia, Pedro and Júlio Joaquim (2013). "O regulamento das custas processuais implicou uma diminuição das receitas para o Estado? O problema da ausência de avaliação prévia de impacto." *Scientia Iuridica LXII*, (331), 107–126.

Darby, Michael and Edi Karni (1973). "Free Competition and the Optimal Amount of Fraud." *Journal of Law and Economics*, 16(1), 67–88.

Direção-Geral da Política de Justiça (2014). *Os Números da Justiça 2013*. Ministério da Justiça, Lisboa.

Djankov, Simeon, Rafael La Porta, Florentio Lopez de Silanes, and Andrei Shleifer (2003). "Courts." *Quarterly Journal of Economics*, pp. 453–517.

Drukker, David, Peter Egger, and Ingmar Prucha (2013a). "On two-step estimation of a spatial autoregressive model with autoregressive disturbances and endogenous regressors." *Econometric Reviews*, 32, 686–733.

Drukker, David, Ingmar Prucha, and Rafal Raciborski (2013b). "A command for estimating spatial-autoregressive models with spatial-autoregressive disturbances and additional endogenous variables." *Stata Journal*, 13, 287–301.

Drukker, David, Ingmar Prucha, and Rafal Raciborski (2013c). "Maximum likelihood and generalized spatial two-stage least-squares estimators for a spatial-autoregressive model with spatial-autoregressive disturbances." *Stata Journal*, 13, 221–241.

Dulleck, Uwe and Rudolf Kerschbamer (2006). "On doctors, mechanics, and computer specialists: the economics of credence goods." *Journal of Economic Literature*, 44(1), 5–42.

European Commission (2014). "Reforms at work: in Italy, Spain, Portugal and Greece." *European Economy*, (5).

European Commission (2015). "EU Justice Score Board." *Communication from the Commission to the European Parliament, the Council, the European Central Bank, the European Economic and Social Committee and the Committee of the Regions*, (COM(2015) 116, final).

Garcia, Sofia, Nuno Garoupa, and Guilherme Vilaça (2008). "A justiça cível em Portugal: uma perspectiva quantitativa." *Fundação Luso-Americana*.

Garoupa, Nuno and Zélia Gil Pinheiro (2014a). *A reforma da justiça e implicações para o orçamento e a economia*, chap. in Para uma Reforma Abrangente da Organização e Gestão do Sector Público – Comunicações e Comentários, pp. 167–204. Ciclo de seminários Sextas da Reforma.

Garoupa, Nuno and Zélia Gil Pinheiro (2014b). "Repensar a justiça em Portugal."

Garoupa, Nuno, Ana Simões, and Vitor Silveira (2006). "Ineficiência do sistema judicial em Portugal: uma exploração quantitativa in Análise Económica do Direito – Parte II." *Sub Judice – Justiça e Sociedade*, (34).

Ginsburg, Tom and Glenn Hoetker (2006). "The unreluctant litigant? An empirical analysis of Japan's turn to litigation." *Journal of Legal Studies*, (35).

Gomes, Conceição (ed.) (2006). *A geografia da justiça – para um novo mapa judiciário*. Observatório Permanente da Justiça Portuguesa – Centro de Estudos Sociais da Universidade de Coimbra.

Gouveia, Mariana, Nuno Garoupa, and Pedro Magalhães (eds.) (2012). *Justiça económica em Portugal*, vol. I-III. Fundação Francisco Manuel dos Santos.

Jappelli, Tulio, Marco Pagano, and Magda Bianco (2005). "Courts and banks: effects of judicial enforcement on credit markets." *Journal of Money, Credit, and Banking*, 37(2), 223–244.

LeSage, James (2004). "Maximum likelihood estimation of spatial regression models." *Spatial Econometrics Course Lectures, Faculty of Economics, University of Coimbra, Portugal*.

Lewis, Jeffrey B. and Drew A. Linzer (2005). "Estimating Regression Models in Which the Dependent Variable Is Based on Estimates." *Political Analysis*, 13, 345–364.

Lorenzano, Dimitri and Federico Lucidi (2014). "The economic impact of civil justice reforms." *European Commission - European Economy Economic Papers*, (530).

Paelinck, Jean and Leo Klaassen (1981). "Spatial econometrics." *Gower*.

Palumbo, Giuliana, Giulia Giupponi, Luca Nunziata, and Juan Mora-Sanguinetti (2013). "Judicial performance and its determinants: a cross-country perspective." *OECD Economic Policy Papers*, (05).

Porta, Rafael La, Florencio Lopez de Silanes, and Andrei Shleifer (2008). "The economic consequences of legal origins." *Journal of economic literature*, 46(2), 285–332.

Posada, Miguel Garcia and Juan Mora-Sanguinetti (2013). "Firm Size and Judicial Efficacy: Evidence for the New Civil Procedures in Spain." *Bank of Spain Working Paper*, (1303).

Tavares, José (2004). "Institutions and economic growth in Portugal: a quantitative exploration." *Portuguese Economic Journal*, 3, 49–79.

Wooldridge, Jeffrey M. (2002). *Econometric analysis of cross section and panel data*. The MIT Press, Cambridge, MA.

World Bank (2014). "Doing Business 2015: going beyond efficiency."

Appendix A: Descriptive Statistics

Variable	Unit	Observations	Mean	Standard deviation	Min.	Max.
Litigation rate	No. / 100 pop	4410	2.79	1.77	0.26	29.95
Litigation rate - declarative	No. / 100 pop	4410	0.69	0.69	0.04	19.54
Litigation rate - enforcement	No. / 100 pop	4410	1.39	1.06	0.01	18.34
Litigation rate - non civil	No. / 100 pop	4410	1.50	1.13	0.23	19.93
Average length of resolved cases	months	4410	18.30	6.21	1.65	60.39
Average length of resolved cases- declarative	months	4410	16.48	6.18	1.17	72.71
Average length of resolved cases - enforcement	months	4410	21.70	9.34	2.13	66.78
Tribunais de círculo	No. / 100 pop	4410	0.06	0.13	0.00	1.07
Tribunais de círculo - declarative	No. / 100 pop	4410	0.02	0.06	0.00	0.58
Tribunais de círculo- enforcement	No. / 100 pop	4410	0.01	0.02	0.00	0.23

TABLE A.1. Descriptive statistics – variables in the first set of regressions

Variable	Unit	Period	Observations	Mean	Standard deviation	Min.	Max.
Purchasing power	index base 1	1993-2011 (biennial)	210	0.70	0.27	0.38	2.63
Illiteracy rate	No. / 100 pop.	1991, 2001, 2011	210	0.12	0.05	0.03	0.30
Lawyers	No. / 100 pop.	2006-2013	210	0.15	0.12	0.07	1.59
Population density	Pop. / dam^2	1993-2013	210	0.03	0.06	0.00	0.57
Small and medium enterprises	No. / pop.	2004-2012	210	0.10	0.02	0.06	0.18
Large enterprises	No. / 1.000 pop.	2004-2012	210	0.04	0.06	0.00	0.42

TABLE A.2. Descriptive statistics – variables in the second set of regressions

Appendix B: Estimation

Equation (1)

Equation (1) cannot be estimated by the usual fixed effects estimator for panel data. Indeed, the dynamic nature of the panel brought about by the inclusion of the lagged dependent variable - that is not strictly exogenous - leads to the inconsistency of this estimator. At the same time, it is plausible that there is an impact of litigation on court congestion and, indirectly, on the length of proceedings. In this case, such length would depend on the litigation in previous periods, leading as well to the violation of the strict exogeneity assumption. Given that it is still reasonable to assume the length of proceedings, similarly to lagged litigation, to be predetermined relative to future litigation, these variables can be instrumented by their own lags (Wooldridge 2002, Chapter 11). Equation (1) was thus estimated by the generalized method of moments of Arellano and Bond (1991) which, in addition to being consistent in the presence of predetermined variables, is more efficient than alternative estimators, such as Anderson and Hsiao (1982). As it was felt that the creation of *tribunais de círculo* could somehow also respond to past congestion, this variable was instrumented in the same way. Taking into account the number of years in the panel, one followed the usual procedure of restricting the number of instruments, which were limited to the sixth lag. However, augmenting the number of lags does not significantly alter the estimates.

Equation (2)

Equation (2) is estimated by the maximum likelihood method that in the presence of spatial autocorrelation is more efficient than the least squares estimator (LeSage 2004). The fact that the dependent variable in this equation is estimated may bring about heteroskedasticity in the variable ε_i (Lewis and Linzer 2005). A generalized least squares estimator, robust to heteroskedasticity, was also tried but with a negligible change in the estimates. Drukker *et al.* (2013b) presents an implementation of this estimator and the maximum likelihood estimator in STATA.

In the regression instrumenting the concentration of lawyers, one uses the generalized method of moments estimator with instrumental variables developed by Drukker *et al.* (2013a). See also Drukker *et al.* (2013c) for an implementation of this estimator in STATA.

Appendix C: Estimates taking into account the endogeneity of the concentration of lawyers

Explanatory variable	Civil cases	Declarative	Enforcement
Constant	1.744***	0.929***	-0.074
	(0.345)	(0.322)	(0.964)
Small and medium enterprises	3.307**	2.573*	3.908**
	(1.377)	(1.474)	(1.879)
Large enterprises	0.436	0.581	0.329
	(0.362)	(0.386)	(0.51)
Purchasing power	0.131	0.141	0.070
	(0.176)	(0.185)	(0.249)
Illiteracy rate	-1.934***	-0.773	-3.491***
	(0.689)	(0.72)	(0.99)
Population density	-0.290	-0.533	-0.851
	(0.714)	(0.748)	(1.079)
Lawyers	1.606***	1.797***	2.356**
	(0.609)	(0.63)	(0.933)
W*Small and medium enterprises	-20.967***	-12.203**	-25.710**
	(5.843)	(5.535)	(10.757)
W*Large enterprises	-2.031	-0.615	-1.888
	(2.836)	(2.9)	(4.224)
W*Purchasing power	1.531**	-0.033	2.195*
	(0.67)	(64)	(126.4)
W*Illiteracy rate	-3.325	-8.229***	1.978
	(2.291)	(2.273)	(4.075)
W*Population density	-5.242**	-4.015*	-3.845
	(2.224)	(2.146)	(3.796)
N (comarcas)	192	192	192

TABLE C.1. Impact of characteristics of *comarcas* on litigation

Notes: Results of the estimation of equation (2) for the mainland *comarcas*, with *comarca* fixed effects as the dependent variable, by a generalized moments estimator with instrumental variables (Drukker *et al.*, 2013a). Standard deviations are in parentheses. P-values: * <0.1; ** <0.05; *** <0.01.

Revisiting the monthly coincident indicators of Banco de Portugal

António Rua
Banco de Portugal

May 2015

Abstract
After a decade releasing the monthly coincident indicators of Banco de Portugal, this article revisits the main features of these indicators which play an important role in the conjunctural assessment of the Portuguese economy. In particular, it is analyzed its behavior as underlying measures of the evolution of the corresponding macroeconomic aggregates as well as their real-time behavior in monitoring economic developments. (JEL: C10, E32)

Introduction

For macroeconomic policymaking it is crucial to have tools that allow assessing the current economic evolution. In particular, the coincident indicators play an important role for monitoring the ongoing economic developments by synthesizing in a single indicator a larger information set which may present heterogeneous behavior.

Banco de Portugal has a long tradition in compiling and releasing coincident indicators for the Portuguese economy. The first wave of coincident indicators goes back to the work of Dias (1993) with the construction of a quarterly coincident indicator for economic activity based on the approach developed by Stock and Watson (1989). Following the same methodology, Gomes (1995) suggested a quarterly coincident indicator for private consumption. The compilation and release of these two indicators ended up being replaced by the monthly coincident indicators proposed by Rua (2004, 2005) for economic activity and private consumption, respectively. Drawing on the methodology proposed by Azevedo *et al.* (2006), these indicators encompass a larger information set and are available on a monthly frequency in contrast with its predecessors.

A decade after the beginning of its release, it is useful to revisit the coincident indicators in several dimensions.[1] On the one hand, it is formally

E-mail: antonio.rua@bportugal.pt

1. One should mention that, in a similar fashion, Dias (2003) also conducted an analysis of the quarterly coincident indicator for the economic activity after a decade.

analyzed the role of the monthly coincident indicators as underlying measures of the evolution of the corresponding macroeconomic aggregate. Given that, by construction, the coincident indicators do not aim to pinpoint at each moment in time the year-on-year change of the corresponding reference variable, it is important to assess if the coincident indicators display a set of features desirable for underlying measures. This evaluation allows to reinforce the role of the coincident indicators in the economic conjunctural analysis and to contribute to a more educated reading by the users.

On the other hand, this article intends to assess the real-time behavior of the coincident indicators namely through the analysis of the information content of the estimates available at each point in time. This exercise can be very useful as the assessment of the past experience may contribute to a better use in the future. In particular, the characterization of the behavior around turning points, which constitute challenging episodes by all means for real-time analysis, intends to frame what one could expect in future similar periods.

Some general features of the coincident indicators

In the spirit of the well-known work developed by Burns and Mitchell (1946) in the United States, one of the main building blocks underlying the construction of the coincident indicators is that the business cycle consists of expansions and recessions occurring in several economic activities. Hence, under this assumption, it seems natural that the cyclical component can be better identified resorting to a larger information set instead of relying solely on a single variable.

Based on the above idea, besides the natural inclusion of the corresponding macroeconomic aggregates, namely real quarterly GDP and private consumption, it were included other series in the construction of the coincident indicators. Among the several selection criteria one could mention the availability on a monthly frequency, timeliness, a reasonable time span, a noteworthy co-movement with the economic cycle and the aim of obtaining a broadly based measure.

In the case of the coincident indicator for economic activity the set of information includes the retail sales volume (retail trade survey) which intends to capture private consumption developments. Regarding investment, the sales of heavy commercial vehicles reflect GFCF in transportation equipment while cement sales are linked to the GFCF in the construction sector. From the supply side, the manufacturing production index captures the industrial sector behavior which is typically the most cyclical sector. In order to take on board the evolution of income and wealth, it is included the households' assessment of their current financial situation. Concerning the labor market, new job vacancies were included. Finally, to reflect external

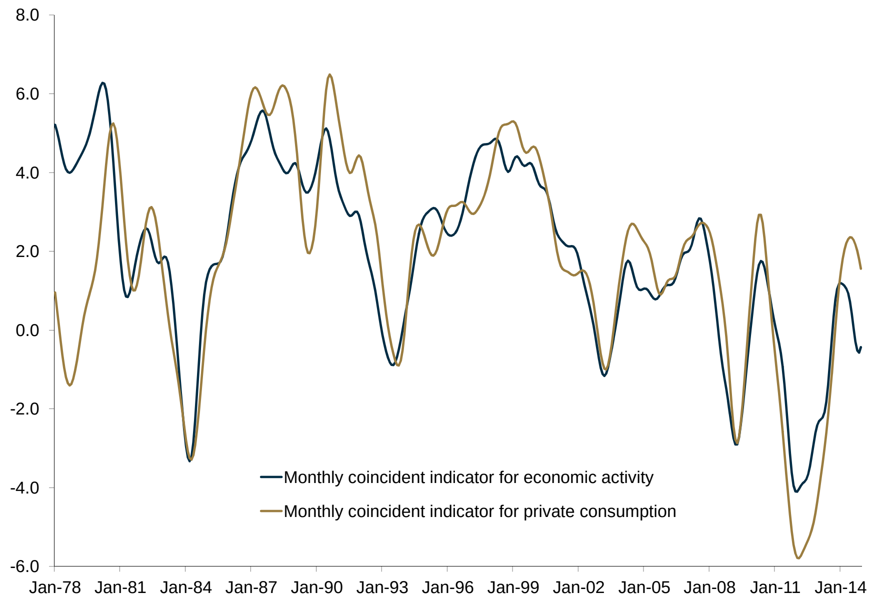


FIGURE 1: Monthly coincident indicators

environment, it was included a weighted average of the current economic situation assessment (consumers' survey) of the Portuguese main trade partners, where the weights are each country's share in Portuguese exports.

In what concerns the coincident indicator for private consumption it was considered the real retail trade turnover index and light passenger vehicles sales which provide quantitative data on the evolution of goods consumption, both durables and non-durables. These data is complemented with qualitative one, namely retail sales volume assessment (retail trade survey). The inclusion of the number of nights spent in Portugal by residents intends to capture, to some extent, developments in services consumption. From the supply side, it is considered the real industrial turnover index of consumer goods in the domestic market. Given that income and wealth are key determinants of consumption behavior, it was included the households' assessment of their current financial situation. Finally, in order to take into account the macroeconomic situation, consumers' opinion on the general economic situation is also considered.

From a methodological stance, the coincident indicators are estimated following the approach suggested by Azevedo *et al.* (2006) which corresponds to the generalization to the multivariate case of the trend-cycle modeling developed by Harvey and Trimbur (2003) which allows extracting a smooth cycle. The model can be casted in state-space form and estimated by maximum likelihood. The coincident indicator is obtained by computing the year-on-year rate of change change of the trend-cycle resulting from the estimated model, with the cyclical component being common to all series while the trend is the one that results implicitly for the reference variable.

The monthly coincident indicators for economic activity and private consumption are presented in Figure 1.

Indicators of the underlying evolution of the corresponding macroeconomic aggregates

Given that, by definition, the coincident indicators focus on the trend-cycle component, they do not aim to match exactly the year-on-year evolution of the reference variable at each point in time. In particular, the coincident indicators are composite indicators intended to reflect the underlying movement of the year-on-year change of the corresponding macroeconomic aggregates. Therefore, one should not expect a total concordance between the evolution of both variables. Although the average difference is almost null, from time to time, there may be significant differences in absolute terms.[2] From Figure 2 it is visible that such differences are larger in the case of GDP than in private consumption given the higher irregularity in the first case.

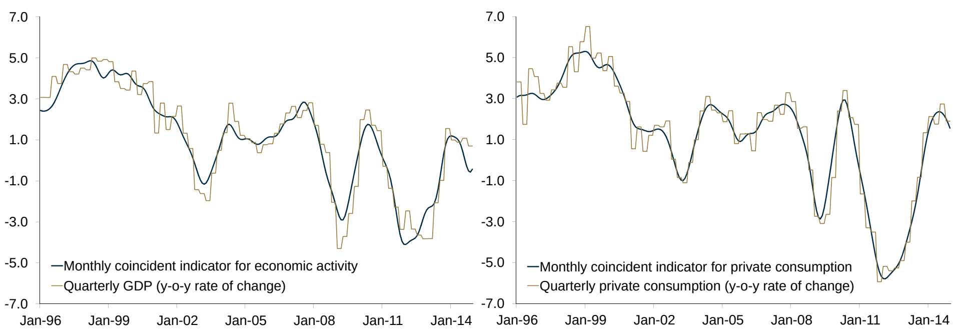


FIGURE 2: The coincident indicators and corresponding macroeconomic aggregates

One possible way of assessing the information content of the coincident indicators in tracking the underlying evolution of macroeconomic aggregates is through its comparison with the year-on-year change of the trend-cycle of GDP and private consumption. To isolate the trend-cycle component it is used the filter proposed by Christiano and Fitzgerald (2003) which presents several

2. In quarterly terms, the average absolute difference is 0.8 p.p. in the case of the coincident indicator for economic activity and 0.6 p.p. in the case of the coincident indicator for private consumption.

advantages over other popular filters in the literature (such as the well-known Hodrick and Prescott 1997 or the filter suggested by Baxter and King 1999).

By comparing the quarterly evolution of the coincident indicators with the year-on-year rate of change of the trend-cycle of the corresponding reference variables (see Figure 3), one can conclude that the coincident indicators seem to capture quite well what they intend to track.[3] In fact, both coincident indicators have a correlation around 0.97 with the year-on-year change of the trend-cycle obtained with the above mentioned statistical filter.

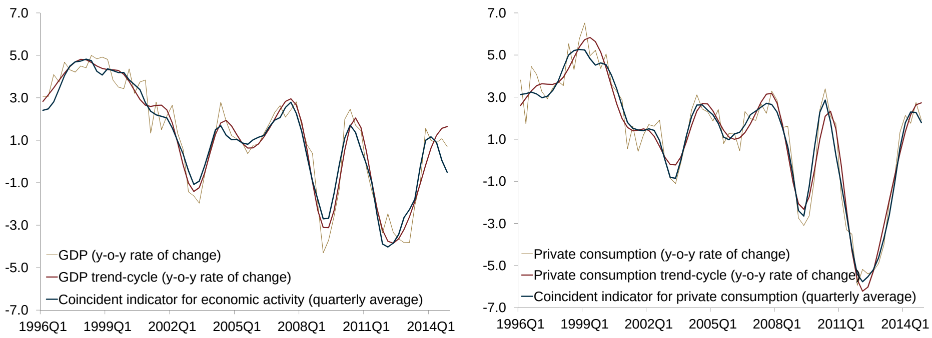


FIGURE 3: The coincident indicators and the year-on-year change of the trend-cycle

A complementary exercise consists in testing formally if the coincident indicators present a set of desirable features in indicators aimed to track the underlying evolution of the reference variable. In particular, Marques *et al.* (2003) proposed a set of conditions that properly adjusted to the current context can be described as follows. The coincident indicator, IC, constitutes an indicator of the underlying behavior of the reference variable, y, if:

$i)$ There is no systematic difference between the coincident indicator and the reference variable.

$ii)$ The coincident indicator should be an attractor of the reference variable. This means that the reference variable should converge to the coincident indicator in the sense that if y is above (below) IC than one should expect that y decreases (increases) and converges to IC.

$iii)$ The reference variable does not Granger cause the coincident indicator. That is, past values of y do not contribute to anticipate the evolution of IC otherwise it would be very difficult to infer about the behavior of y based on IC. One should mention that this condition also implies that the reference variable does not act as an attractor of the coincident indicator.

3. One should mention that the evolution in the most recent period should be read with additional caution as it is potentially subject to revisions as discussed in the next section.

Through the estimation of appropriate econometric models, it is possible to statistically test the above conditions (see Appendix). From the empirical results obtained one can conclude that both the coincident indicator for economic activity and the coincident indicator for private consumption fulfill all the conditions required to be underlying evolution indicators.

Real-time behavior

In practice, the assessment of the underlying evolution of the economy in real-time is extremely hard. On the one hand, the split between signal and noise in the most recent period is difficult due to the uncertainty regarding the future behavior. In fact, the distinction between the observed and the underlying evolution for a given point in time reflects both the preceding and subsequent behavior to that point in time. Hence, the assessment of the underlying evolution can be potentially revised with the arrival of new information. On the other hand, statistical data can be subject to revisions that result from the natural incorporation of additional information in its compilation or due to changes, from time to time, of a more methodological nature.

In light of this, it is relevant to document the real-time behavior of the coincident indicator for economic activity and the coincident indicator for private consumption. Hence, all the real-time estimates released by Banco de Portugal were collected since the beginning of its publication, namely June 2004 in the case of the coincident indicator for economic activity and October 2005 in the case of the coincident indicator for private consumption.

The evaluation of the real-time reliability focus on assessing how close a given estimate is from the subsequent ones. In this context, it is reported in Tables 1 and 2 a set of descriptive statistics regarding the different estimates of the coincident indicators. In particular, it is analyzed the monthly revisions between the first and second estimates, between the second and the third, between the third and the fourth and between the fourth and fifth estimates. In addition, it is assessed the revisions between the first and fifth estimates, so as to capture the cumulative sequence of revisions, as well as the revisions between the first and the final estimate, considered as the latest estimate. The report of the revisions up to the fifth estimate is to assure that the monthly estimate that takes on board the quarterly information of the reference variable for the quarter to which that month belongs is considered. By its turn, the revision of the final estimate vis-à-vis the first estimate reflects the whole cumulative process of revision over time.

	1^{st} vs 2^{nd} estimate	2^{nd} vs 3^{rd} estimate	3^{rd} vs 4^{th} estimate	4^{th} vs 5^{th} estimate	1^{st} vs 5^{th} estimate	1^{st} vs Final estimate
Mean revision	0.0	0.0	0.0	0.0	−0.1	0.0
1^{st} quartile	−0.2	−0.1	−0.1	−0.1	−0.4	−0.5
Median	0.0	0.0	0.0	0.0	0.0	0.2
3^{rd} quartile	0.1	0.1	0.1	0.0	0.2	0.6
Mean absolute revision	0.2	0.2	0.1	0.1	0.4	0.2
Standard deviation	0.3	0.2	0.2	0.2	0.5	0.8
Noise-to-signal ratio	0.15	0.12	0.10	0.09	0.27	0.44
Sign concordance	0.94	0.97	0.94	0.95	0.89	0.83
Direction concordance	0.86	0.94	0.99	0.97	0.87	0.80

TABLE 1. Monthly revisions of the coincident indicator for economic activity

	1^{st} vs 2^{nd} estimate	2^{nd} vs 3^{rd} estimate	3^{rd} vs 4^{th} estimate	4^{th} vs 5^{th} estimate	1^{st} vs 5^{th} estimate	1^{st} vs Final estimate
Mean revision	−0.1	0.0	0.0	0.0	−0.2	0.0
1^{st} quartile	−0.2	−0.1	−0.1	−0.1	−0.7	−0.4
Median	0.0	0.0	0.0	0.0	−0.1	0.1
3^{rd} quartile	0.1	0.1	0.1	0.1	0.3	0.6
Mean absolute revision	0.3	0.2	0.2	0.1	0.6	0.9
Standard deviation	0.4	0.4	0.3	0.2	0.8	1.3
Noise-to-signal ratio	0.16	0.13	0.11	0.09	0.28	0.48
Sign concordance	0.94	0.95	0.96	0.98	0.88	0.83
Direction concordance	0.92	0.94	0.95	0.95	0.86	0.82

TABLE 2. Monthly revisions of the coincident indicator for private consumption

Firstly, one can conclude that the mean revision is close to zero which means that there is no evidence in terms of bias of the different estimates. The median is also approximately nil with a large proportion of the revisions presenting a relatively small size.[4]

In terms of the size of absolute revisions, both the mean absolute revision and the standard deviation of the revisions point to relatively low figures in the initial estimates while increasing with the revision horizon. The signal-to-noise ratio, computed as the ratio between the standard deviation of the revisions and the standard deviation of the final estimate and which allows to assess the relative importance of the revisions, point to the same findings. One should note that these results also reflect the revisions of the Quarterly National Accounts (see Cardoso and Rua 2011). The latter revisions can be noteworthy when Annual Accounts are released, with a typical lag of two years, or due to methodological changes like base changes which can influence significantly the comparison between the first and final estimates.

Besides the analysis of the size of the revisions, it was also computed measures like the concordance both in terms of sign and direction between the different estimates of the coincident indicators. In the first case, it measures the percentage of times that both estimates share the same sign whereas the latter measures the percentage of times that the estimates give the same qualitative indication in terms of increase/decrease. In both cases, it was recorded very high concordance rates.

Additionally, to have an idea of the uncertainty associated with the coincident indicators over time, it was calculated the standard deviation of the different estimates for each month since the beginning of its publication (see Figure 4). Despite the relatively small standard deviation, it is visible that the uncertainty is higher around turning points as expected.

Given that turning points are usually the points in time which are more challenging for real-time performance, one now focus in more detail on the reliability of the coincident indicators around turning points. Hence, it was considered four turning points namely two local maxima and two local minima. For a comparison term, it was also considered the performance that one would have in real-time with the year-on-year change of the trend-cycle extracted with the Christiano-Fitzgerald filter. One should note that the latter only takes on board data regarding the reference variable, that is, quarterly GDP or private consumption. Given the typical end-of-sample filtering problem, the latest values can be subject to substantial revisions. One possible way to mitigate such problem consists in using multivariate information, which is basically one of the main building blocks of the methodology used to compile the coincident indicators. In fact, the use of

4. One should note that since both coincident indicators were re-estimated around mid-2009 the revisions might be a bit larger in that period of time.

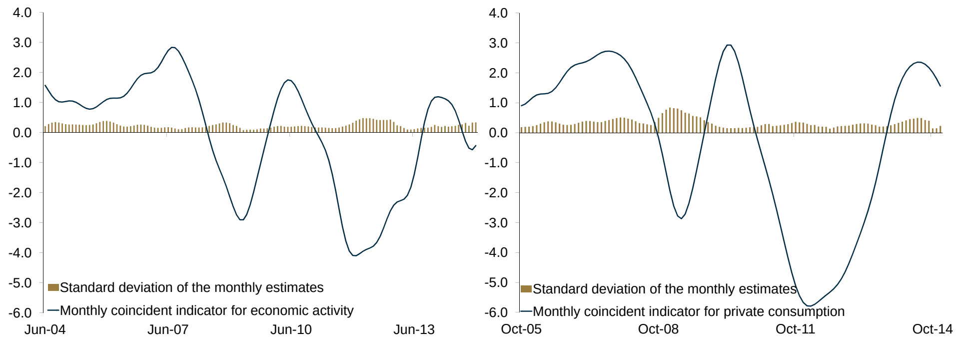
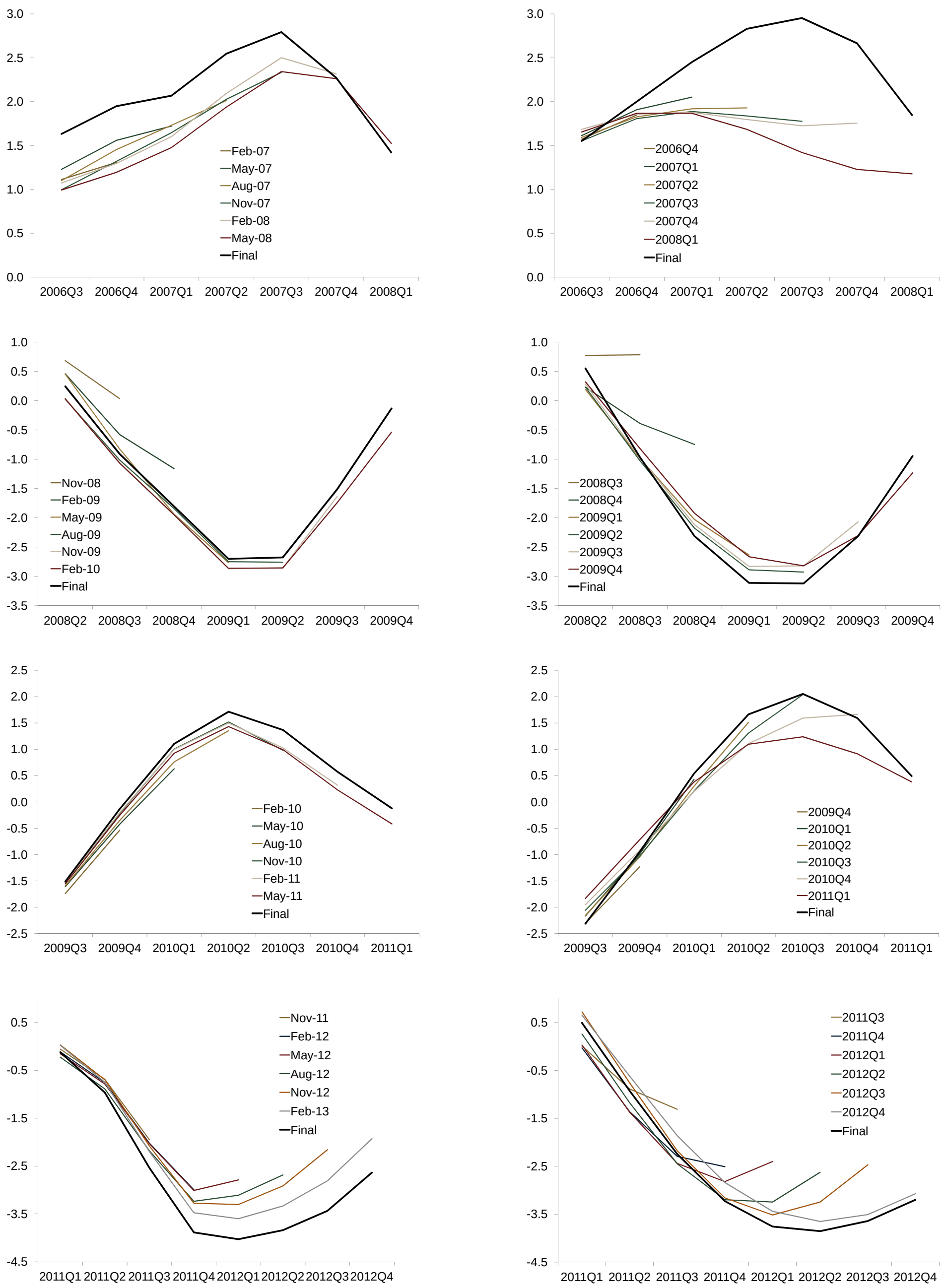


FIGURE 4: Coincident indicators and corresponding standard deviation of the monthly estimates

information beyond that conveyed by reference variable may allow mitigating the size of the revisions and reinforcing the real-time information content.

In Figures 5 and 6, the real-time estimates of the coincident indicators as well as the year-on-year change of the trend-cycle extracted with the above mentioned filter are presented. Note that the subsequent analysis is conducted on quarterly frequency so as to allow for the above mentioned comparison and all the estimates have been obtained using the available vintages at each point in time. In the case of the coincident indicators, it corresponds to the public releases by Banco de Portugal whereas in the case of the statistical filter it was required to collect all the vintages for GDP and private consumption. One should mention that the estimates presented in the figures correspond to the estimates obtained at the time of release of the Quarterly National Accounts by INE being therefore comparable in terms of closing date of information.

From Figures 5 and 6 it is possible to conclude that the coincident indicators have presented a higher information content in real-time in the identification of the turning points since the qualitative indication provided in real-time does not seem to have changed significantly with the arrival of new information. By its turn, the estimates obtained with the above mentioned filter are much more sensible to the inclusion of additional quarterly observations for the macroeconomic aggregate.



Coincident indicator | Christiano-Fitzgerald filter

FIGURE 5: Coincident indicator for economic activity in quarterly terms around turning points

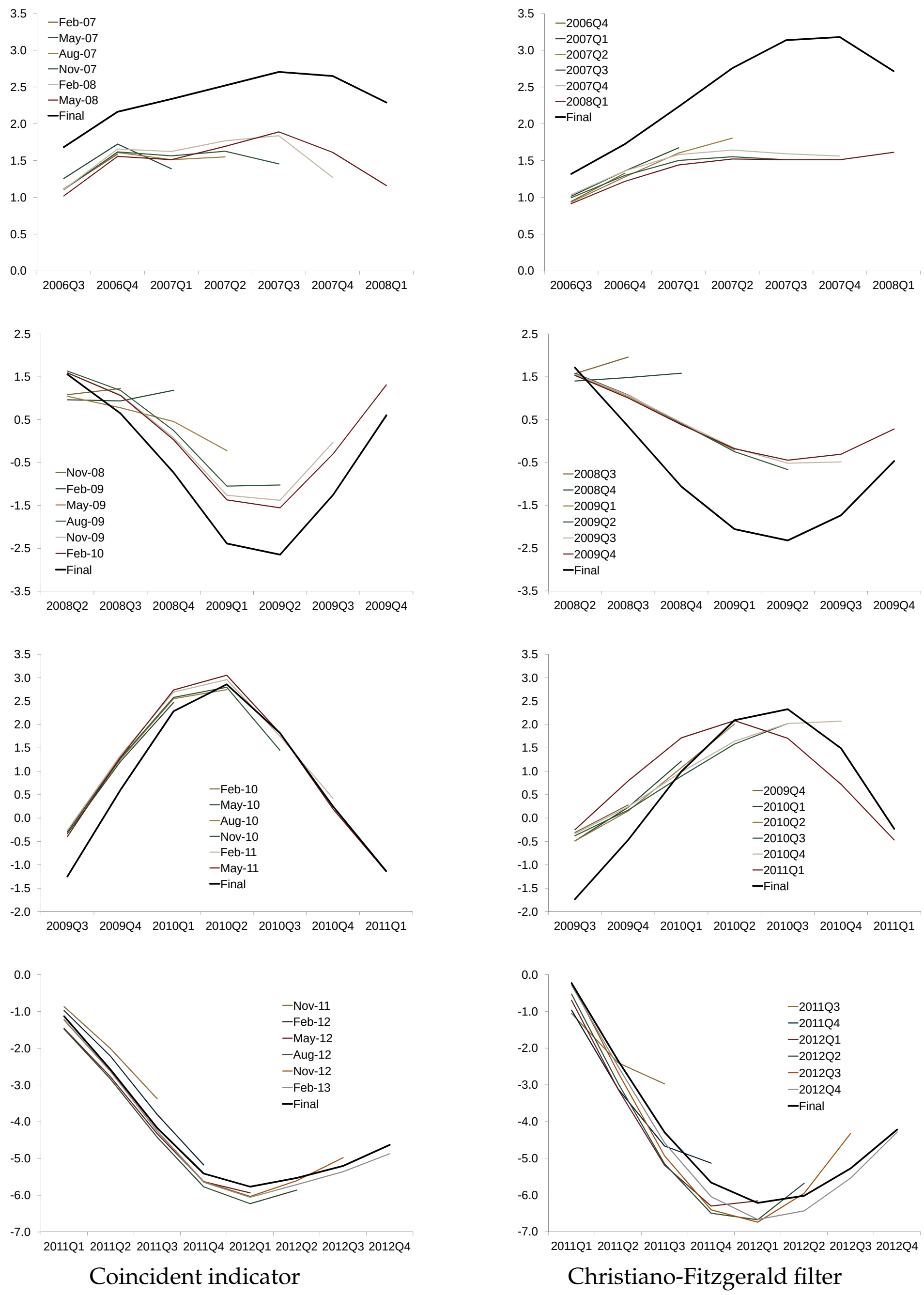


FIGURE 6: Coincident indicator for private consumption in quarterly terms around turning points

	Revisions vis-à-vis the final estimate			Concordance vis-à-vis the final estimate	
	Mean	Median	Standard deviation	Sign	Direction
Quarterly estimate of the coincident indicator at the time of release of the Quarterly National Accounts	−0.1	0.1	0.6	0.96	0.91
Y-o-y rate of change of the trend-cycle extracted with Christiano-Fitzgerald filter	−0.2	−0.1	0.8	0.96	0.74
Memo					
First quarterly estimate of the coincident indicator for a given quarter	−0.3	0.0	0.8	0.91	0.87

TABLE 3. Quarterly revisions of the coincident indicator for economic activity around turning points

	Revisions vis-à-vis the final estimate			Concordance vis-à-vis the final estimate	
	Mean	Median	Standard deviation	Sign	Direction
Quarterly estimate of the coincident indicator at the time of release of the Quarterly National Accounts	−0.1	0.0	1.0	0.96	0.83
Y-o-y rate of change of the trend-cycle extracted with Christiano-Fitzgerald filter	−0.3	−0.2	1.1	0.91	0.74
Memo					
First quarterly estimate of the coincident indicator for a given quarter	−0.4	−0.3	1.5	0.78	0.74

TABLE 4. Quarterly revisions of the coincident indicator for private consumption around turning points

From a quantitative point of view, the different estimates of the coincident indicators are relatively close to both the preceding and immediately subsequent estimates while the indication given by the year-on-year change of the trend-cycle extracted in real-time with the statistical filter is subject to larger revisions. One should recall that due to the above mentioned reasons the revisions can be large. In particular, in 2007 there was a base change in the national accounts involving methodological changes in the Annual National Accounts which contributed to a larger discrepancy between the real-time estimates and the final estimate. In particular, the release of the Annual National Accounts led to a revision, for 2007 as a whole, of 0.5 p.p. in the case of GDP and around 1.0 p.p. in the case of private consumption. Moreover, regarding private consumption it was also recorded a revision of -1.5 p.p. in 2009 with the release of the corresponding Annual National Accounts.

The information conveyed by Figures 5 and 6 is complemented with a set of descriptive statistics presented in Tables 3 and 4 respectively.

Tables 3 and 4 support quantitatively the above findings. In light of all the statistical measures presented, both the coincident indicator for economic activity and the coincident indicator for private consumption are more reliable in real-time than what would be possible to achieve by resorting to the statistical filter. In particular, the mean revision and the median are almost null for both coincident indicators. By its turn, the standard deviation of the revisions vis-à-vis the final estimate is higher for private consumption than for economic activity which may reflect the larger revisions recorded in the corresponding macroeconomic aggregate in the national accounts. In terms of concordance, it is observed high percentages both in terms of signal and direction. Besides being available at a higher frequency which allows a monthly monitoring of the economic evolution, the coincident indicators also present a more reliable real-time behavior in quarterly terms when compared with the statistical filter.

Additionally, it was also assessed the first quarterly estimate available for the coincident indicators for a given quarter. This estimate corresponds to the one which can be typically obtained two months before the other estimates analyzed. As expected, since the information available is more scarce and preliminary, such estimate is subject to larger revisions and presents lower information content when compared with the estimate for the coincident indicator at the time of the release of the national accounts. However, one should highlight that in the case of economic activity, this first estimate delivers a similar performance to that of the statistical filter with a higher concordance in terms of direction. In the case of the coincident indicator for private consumption, the first quarterly available estimate presents a behavior close to the one of the statistical filter.

Conclusions

After a decade of releases by Banco de Portugal, this article revisited the monthly coincident indicators in several dimensions. On the one hand, the role of the coincident indicators as measure of the underlying evolution of the corresponding macroeconomic aggregates was evaluated. Based on a formal approach, it was possible to conclude that both the coincident indicator for economic activity and the coincident indicator for private consumption present a set of features desirable for underlying indicators.

On the other hand, the revisions of the coincident indicators since the beginning of its publication were also analyzed. Although by resorting to a multivariate set of information one can eventually mitigate the revisions and improve the real-time information content, one should bear in mind that the classical problem of evaluating the underlying evolution in real-time will be hardly ever fully overcome whatever the tool used.

The analysis of the real-time behavior of the coincident indicators has shown that these composite indicators have proved to be quite useful in monitoring and tracking the evolution of the Portuguese economy, even during turning points which are the most challenging periods of time marked by a higher uncertainty. In fact, the monthly coincident indicators have presented a reliable real-time behavior in terms of both the size of the revisions and concordance.

References

Azevedo, J., S. Koopman, and A. Rua (2006). "Tracking the business cycle of the Euro area: a multivariate model-based band-pass filter." *Journal of Business & Economic Statistics*, 24(3), 278–290.

Baxter, M. and R. G. King (1999). "Measuring business cycles: approximate band-pass filters for economic time series." *The Review of Economics and Statistics*, 81, 575–593.

Burns, A. and W. Mitchell (1946). *Measuring business cycles*. NBER.

Cardoso, F. and A. Rua (2011). "The quarterly national accounts in real-time: an analysis of the revisions over the last decade." Economic Bulletin, Autumn, 137-154, Banco de Portugal.

Christiano, L. J. and T. J. Fitzgerald (2003). "The band-pass filter." *International Economic Review*, 44, 435–465.

Dias, F. (1993). "A composite coincident indicator for the Portuguese economy." Working Paper 18/93, Banco de Portugal.

Dias, F. (2003). "The coincident indicator of the Portuguese economy: An historical assessment after ten years." Economic Bulletin, September, 93-100, Banco de Portugal.

Gomes, F. (1995). "A coincident indicator and a leading indicator for private consumption." Economic Bulletin, September, 73-80, Banco de Portugal.

Harvey, A. and T. Trimbur (2003). "General model-based filters for extracting cycles and trends in economic time series." *The Review of Economics and Statistics*, 85, 244–255.

Hodrick, R. and E. Prescott (1997). "Postwar U.S. business cycles: an empirical investigation." *Journal of Money, Credit, and Banking*, 29, 1–16.

Marques, C. R., P. D. Neves, and L. M. Sarmento (2003). "Evaluating core inflation indicators." *Economic Modelling*, 20, 765–775.

Rua, A. (2004). "A new coincident indicator for the Portuguese economy." Economic Bulletin, June, 21-28, Banco de Portugal.

Rua, A. (2005). "A new coincident indicator for the Portuguese private consumption." Economic Bulletin, Autumn, 65-72, Banco de Portugal.

Stock, J. and M. Watson (1989). "New indexes of coincident and leading economic indicators." NBER Macroeconomics Annual, NBER.

Appendix

In this appendix, it is presented the econometric models used to assess the set of desirable features in indicators aimed to track the underlying evolution of the reference variable.

The condition of no systematic difference between the coincident indicator, IC, and the reference variable, y, can be assessed by estimating the following model

$$y_t = \alpha + \beta IC_t + \upsilon_t$$

and testing jointly if $\alpha = 0$ and $\beta = 1$.

Regarding the condition that the coincident indicator should be an attractor of the reference variable, the model to be estimated is given by

$$\Delta y_t = \sum_{j=1}^{m} \alpha_j \Delta y_{t-j} + \sum_{j=1}^{n} \beta_j \Delta IC_{t-j} - \gamma\left(y_{t-1} - IC_{t-1}\right) + \varepsilon_t$$

and test if $\gamma \neq 0$.

The condition that the reference variable does not Granger cause the coincident indicator can be tested through the model

$$IC_t = \mu + \sum_{j=1}^{r} \delta_j IC_{t-j} + \sum_{j=1}^{s} \theta_j y_{t-j} + \eta_t$$

and assess if $\theta_1 = \theta_2 = \ldots = \theta_s = 0$.

The main estimation results of the above mentioned models are reported in Table 5. One can conclude that both coincident indicators fulfill all the conditions.

	Coincident indicator for economic activity	Coincident indicator for private consumption
Condition $i)$ $\alpha = 0$ and $\beta = 1$	$\underset{(0.109)}{\widehat{\alpha} = -0.034}$ $\underset{(0.040)}{\widehat{\beta} = 1.036}$ $F(2, 146) = 0.507$ [0.603]	$\underset{(0.062)}{\widehat{\alpha} = 0.027}$ $\underset{(0.021)}{\widehat{\beta} = 1.013}$ $F(2, 146) = 0.568$ [0.567]
Condition $ii)$ $\gamma \neq 0$	$\underset{(0.182)}{\widehat{\gamma} = 1.767}$	$\underset{(0.170)}{\widehat{\gamma} = 1.124}$
Condition $iii)$ $\theta_1 = \theta_2 = \ldots = \theta_s = 0$	$F(4, 135) = 0.333$ [0.855]	$F(4, 135) = 1.256$ [0.290]

TABLE 5. Conditions for underlying evolution indicators

Note: In round brackets are reported the heteroscedasticity and autocorrelation consistent standard errors while in square brackets appear the p-value of the test statistics.