Banco de Portugal Economic Studies

Volume III

BANCO DE PORTUGAL
EUROSYSTEM

1

1

volume III

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, 2017 · www.bportugal.pt

BANCO DE PORTUGAL ECONOMIC STUDIES | Volume III - n.º 1 | Lisbon 2017 • Banco de Portugal Av. Almirante Reis, 71 | 1150-012 Lisboa • www.bportugal.pt • Edition Economics and Research Department • Design Communication Directorate | Image and Graphic Design Unit | Editing and Publishing Unit • ISSN 2183-5217 (online)

Content

Editorial

January 2017

The first issue of Banco de Portugal Economic Studies for 2017 contains four essays covering heterogeneous but very relevant topics. One is a theoretical topic of universal appeal: what is the best taxation policy regarding wealth? The other essays have an empirical nature and use data on the Portuguese economy and Public Administration to study the demographics of firms, the dating of business cycles and the workings of the civil justice system.

The first paper is "Productivity in civil justice in Portugal: A crucial issue in a congested system", by Manuel Coutinho Pereira and Lara Wemans. This essay deals with the "supply" side of the Justice system and provides a follow-up to an earlier study by the same authors on the "demand" side, published in 2015 in Banco de Portugal Economic Studies.

The area of justice has been identified as one of the Achilles' heels of the Portuguese economy, a situation aggravated in recent years by the increase in litigation generated by the Great Recession, and with surveys showing that the justice system delays are one of the top concerns of firms. The paper by Pereira and Wemans studies the determinants of productivity in civil justice in Portugal using data for the period between 1993 and 2013.

In a European context, Portugal has a number of judges and non-judge staff per capita below the European average and above, but close to, the average of the countries of legal French origin. In terms of financial resources, the budget for courts per inhabitant is slightly higher than the average, in particular when considering that budget as a percentage of GDP. The country is located in the set of countries with clearance rates (relation between cases resolved and incoming) below 100 percent and with slightly above average congestion indicators. This suggests that there is scope for improving the efficiency of the Portuguese justice system, bringing the country closer to European best practices.

In order to enable an analysis of these issues, Pereira and Wemans constructed a database on the courts of first instance with information on 210 *comarcas* and referring to the period between 1993 and 2013. This database combines three different sources of information covering the case flows and staffing between 1993 and 2013 and the budget execution of the courts between 2007 and 2013. Between 1993 and 2013 the number of civil cases, making up more than half of the cases resolved in the courts of first instance followed a growing trend. Throughout the period under review the justice system has generally resolved fewer civil cases than those incoming. The Portuguese judicial system thus has a very high level of congestion: given the figures for the cases settled and pending in 2013, it would take about

two years and three months to resolve all pending cases. Meanwhile, the average duration of the completed cases grew steadily between 1993 and 2007, showing a reduction afterwards. In the recent period, the average length of civil cases has fluctuated around 30 months.

The results of a careful analysis explaining econometrically the number of civil cases resolved per judge indicate that this variable responds positively to demand pressure, but in a way that differs between large and small *comarcas*, with a greater degree of resource utilization in the larger *comarcas*. On the other hand, there is a positive effect of specialization in the productivity of judges.

The results mentioned above and others in the study are pieces of information relevant to the management of resources in the justice system as they suggest areas with potential for significant productivity gains. This type of structural reforms, such as altering the human resources available to the courts, may stand to gain from cost-benefit studies and this article constitutes a relevant input for such analyzes.

The second essay in this issue is "Firm creation and survival in Portugal" by Sónia Félix. Every year thousands of new firms are born and thousands of existing firms die, just like living beings. The dynamics of the aggregate economy are heavily dependent on this cycle of firms' creation and demise. What do we know about their birth rates and their mortality rates, or to put it in a more fruitful way, their survival patterns? The study by Félix answers these questions using longitudinal data from the governmental Simplified Corporate Information (IES), a database that covers a vast majority of all Portuguese nonfinancial corporations and that became available in 2005. While defining firm creation poses no major empirical problem, one needs to operationalize the concept of a firm's death. Following the literature in this area, the author identifies a firm closure or exit as the period in which firms cease to report IES information for at least two consecutive years. The analysis uses data from 2005 up to 2014, which means firms are followed for a maximum of eight years beyond their creation year. The results show that, on average, slightly more than 11.5 thousand firms are created per year, a figure that corresponds to an average 3.4 percent entry rate. The longevity of these new firms is quantified by a Kaplan-Meier survival function where about 5 percent of the firms exit after one year, close to 50 percent exit after six years and only 41 percent survive after eight years. The study shows that the entry figures and entry rates are heterogeneous across industries, as are survival rates. Another result in the study concerns the size of the new firms: it is quite small with more than 95 percent of new firms employing less than ten workers. However, new firms with more than 10 workers have higher survival probabilities than the other new firms and this difference increases as they age. Finally, when considering the macroeconomic environment, firm entry seems to be pro-cyclical, in particular for small entrants.

The paper by António Rua, entitled "Dating the Portuguese business cycle", uses developments from the literature studying the definition and

dating of business cycles and applies those ideas and techniques to the study of the Portuguese business cycles from the beginning of 1977 up to the end of 2015. To begin with, there is a widespread idea that two consecutive quarters of negative real GDP growth define a recession. It turns out this is not exactly true. The dating of business cycles in the US by the National Bureau of Economic Research business cycle dating committee takes into account a richer set of information as well as the judgment of the committee members. The results sometimes differ from the "two-quarters rule". In order to capture the dating of business cycles in a quantitative rule, Bry and Boschan developed an algorithm in the early seventies that seems to be quite precise. Rua applies the Bry-Boschan algorithm to the Portuguese quarterly GDP data and produces estimates for the dates of peaks and troughs, identifying six recessions since 1977.

Rua then proceeds to analyze monthly data from a coincident indicator that is built from multiple data sources including, besides GDP quarterly data, monthly data from retail sales, heavy commercial vehicles and cement sales, data from manufacturing, from the labor market, etc. Using this more detailed dataset Rua identifies five recessions in the period under study, eliminating the first recession, in 1980, from the list identified by the earlier quarterly GDP approach. The last part of the essay examines the dating of the turning points by looking simultaneously at multiple time series with data from a diversified set of sources. In general the results reinforce the robustness of the previously obtained monthly business cycle chronology.

The fourth and last essay in this issue, by Pedro Teles with Joana Garcia, is entitled "Why Wealth Should Not Be Taxed". The authors motivate the paper by referring to the increase in wealth inequality that has accompanied the increase in income inequality in the last fifty years. One might think that correcting the problem would imply using all kinds of redistributive taxation including taxation of capital income and taxation of wealth. However, the Economics literature, reinforced by the latest contribution by Chari, Nicolini and Teles, has a surprising negative result: the best policy is not to tax both capital income and wealth even if we care only about the welfare of the workers. Not taxing wealth is a Pareto improvement compared with taxing wealth as all households, rich and poor, benefit from wealth accumulation not being taxed.

The results are obtained in an intertemporal model where wealth accumulation means increasing the capital stock. It turns out that in the model an initial unanticipated taxation of past-accumulated wealth may be a desirable way to fund present and future public expenditures but the tax policymakers must credibly commit to never tax wealth again. In others words a more precise statement of the result is that the optimal tax rate on future capital accumulation is zero. What explains this result? An intuitive answer may be glimpsed from two fundamental ideas. First, a classical result obtained in the 1970s by Diamond and Mirrlees is that in general it is not a good policy

to introduce inefficiencies in production by having taxes distorting producers' choices of inputs and technologies. Taxes are best located in the consumption of final goods. The second idea comes from the intertemporal nature of the model. If one takes a general view, consumption and labor (more precisely leisure) today, tomorrow and the day after tomorrow are final goods. But wealth is not a final good, it is part of the production technology used to generate final goods over time. Taxing wealth implies changes in the capital accumulation that ultimately have consequences for consumption in the future. Taxing the accumulation of wealth is akin to a distortion on production as it means taxing future consumption more than current consumption. It also means taxing current labor more than future labor. To the extent that stable taxation of consumption and labor over time is desirable, as indeed is the case for reasonable preference specifications, and that such a goal can be achieved by contemporaneous taxation of said consumption and labor, then it is efficient not to tax wealth.

This is a striking theoretical result. Does it have direct policy implications? It does as it reminds us that taxing wealth may have costs for workers that are not obvious. On the other hand, the results of the model are based on assumptions that may prove to be less than robust. For example, the model assumes that the allocation of capital to firms occurs in markets that are competitive and efficient. What if that is not the case and these markets have imperfections? The bottom line is that while drawing the implication that taxation of wealth should be zero here and now may be a little premature, the results obtained remind us that the costs to the economy of taxing wealth may be much larger than what is usually perceived.

Productivity in civil justice in Portugal: A crucial issue in a congested system

Manuel Coutinho Pereira
Banco de Portugal

Lara Wemans
Banco de Portugal

Abstract
This article focuses on the determinants of productivity in civil justice in Portugal using panel data covering the period from 1993 to 2013, from a strictly quantitative perspective. The performance indicators and the relationship between demand and resource distribution in the territory suggest that there is room for improving the allocation of resources. Such evidence is confirmed by a positive response of productivity to incoming cases per judge, taking into account the casemix. Moreover, productivity is positively impacted by both the number of judicial staff per judge and the proportion of cases resolved in judgeships that deal mainly with civil cases. (JEL: K40, H11, H40)

Introduction

The implementation of structural reforms that foster the growth potential of the Portuguese economy has been systematically suggested by several international institutions as a way to counteract low medium-term growth perspectives. At the same time, fiscal consolidation needs increased the pressure for efficiency of public policies and, in some areas, the reduction of available resources coincided with an increase in the demand for the services they provide. Justice was one of the sectors pressured during the crisis, namely as regards «economic» litigation (Correia and Videira 2015). In addition, the Portuguese justice system maintains a high congestion level which consistently places the country in the group of low performers in international comparisons (CEPEJ 2014). In this context, the justice sector has been at the centre of the discussions regarding structural reforms and the improvement in efficiency of the public sector.

Acknowledgements: The authors thank Direção-Geral de Política da Justiça and Direção-Geral da Administração da Justiça for providing the data on the judicial system and for valuable clarifications. The authors are also grateful for the comments made by João Amador, António Antunes, Manuela Espadaneira Lourenço, Miguel Gouveia and Nuno Garoupa and by the participants in a seminar from the Economic Research Department of Banco de 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

The connection between efficiency of the justice system and economic growth has been the focus of several articles which relate the reduction of delays in case resolution in economic litigation (mainly civil and tax) to foreign direct investment and firms entry rates, key factors for economic growth (Lorenzano and Lucidi 2014). In the Portuguese case, recent surveys as the Business Cost of Contexts Survey from Statistics Portugal, published in 2015, and Gouveia *et al.* (2012a) show that firms identify the justice system as one of the top constraints to their activity. These results suggest that improving the performance of the justice sector could have relevant economic effects. This has been the basis for the recent reforms in the area of economic litigation, in particular those implemented during the Economic and Financial Assistance Programme, as the changes to the procedure rules for enforcement cases.[1]

This article focuses on the determinants of productivity in civil justice between 1993 and 2013. Note that, although tax cases could also be particularly relevant from an economic perspective, the disclosure of statistics regarding administrative and tax courts, which in Portugal have a separate jurisdiction from judicial courts, is much scarcer.[2] The period under analysis comes just before the most recent change in the judicial map that took place in 2014. This does not hinder the relevance of the analysis which can indeed be useful in the evaluation of the results of this reform, when data for a significantly long period after its implementation become available.

Empirical evidence regarding the determinants of efficiency of the judicial sector is quite vast. In recent years, several cross-country studies focusing on this topic have been released, mainly using CEPEJ data as in Voigt and El-Bialy (2016). The main determinants analysed in the literature concern organizational aspects, as the size of courts and their specialization level, the allocation of human and financial resources, court management and incentives (Gouveia *et al.* 2016). Although CEPEJ information is quite detailed, justice systems have substantial differences in aspects that are hardly quantifiable, such as the different agents' culture or the procedure rules in place. In this context, it is important to complement this evidence with analyses focused on a single country, to better inform public policy decision makers. Additionally, cross-country studies are sometimes based on correlations between very aggregate indicators, at the country level, while a detailed analysis of judicial systems requires more disaggregated data. Regarding studies focused on the Portuguese system, the paper by Borowczyk-Martins (2010) stands out, focusing on the determinants of productivity using 2001 data from civil judgeships in first instance courts.

1. For further details concerning the measures implemented during the adjustment programme in this area and their effects, see Correia and Videira (2015, 2016) and Pompe and Bergthaler (2015).

2. There has been an improvement in the statistics released in this area more recently. Nonetheless they remain less detailed than judicial courts' statistics.

The main innovation of our paper is the reliance on a much broader sample, with the benefits of panel data. Moreover, it focuses on a more diverse set of explanatory variables.

There are several indicators that reflect the efficiency of justice systems, with the duration of resolved cases standing out as the one impacting economic agents' decisions more directly. In fact, the conclusions of the Business Cost of Contexts Survey, already mentioned, reinforce the importance of the justice system delays that are indicated as the top concern of firms. However, the duration of resolved cases is strongly influenced by internal procedures and practices of courts which may lead to a court resolving mainly very recent or particularly old cases in a certain year. As a result, this measure does not correctly gauges the efficiency of the system in each moment in time. In addition, procedure rules and potential procedural incidents can have an impact on this indicator that is not proportional to the effort made by the judge.

Considering the abovementioned limitations, we chose to focus on the ratio between the results achieved - number of resolved civil cases - and the number of judges. For simplicity, we refer to this ratio as a measure of productivity. A clear limitation of this measure is to ignore the complexity of cases, the so-called casemix (Gomes 2005). In order to overcome this limitation, the econometric approach followed accounts for both the heterogeneity across different *comarcas* and the caseload from other litigation areas, as outlined below. Another limitation of this indicator – common to all strictly quantitative indicators – is the fact that it disregards the quality of judicial decisions, an aspect which certainly also influences investment decisions but that was impossible to consider due to data limitations. The indicator commonly used to gauge the quality of judicial decisions is the reversal rate in higher courts.[3] Note, however, that according to the results from a survey to Portuguese firms presented in Gouveia *et al.* (2012a), negative evaluations are much more prevalent for the duration of cases than the quality of decisions, even for the companies that had a majority of unfavourable rulings.

This article is organized as follows. Firstly, there is a section dedicated to data, presenting the main characteristics of the database that covers 210 *comarcas* from 1993 to 2013. Secondly, the main indicators of resources and performance of the judicial system are presented, including a discussion on their territorial distribution and a brief international comparison. Thirdly, the main determinants of productivity are discussed, taking casemix into account. Finally, we make some concluding remarks.

3. See, for instance, Rosales-López (2008) for a discussion on the connection between productivity and quality of the decisions using this indicator.

Data

The database was constructed by merging three different datasets for first instance courts. Direção-Geral da Política de Justiça (DGPJ) provided one dataset on incoming[4], pending and resolved cases, and another one regarding judicial staff.[5] In addition, we gathered budgetary information from Direção-Geral da Administração da Justiça between 2007 and 2013. In order to merge this information into a single dataset, it was necessary to create a correspondence between the different classifications of courts.

As the territorial organization of the justice system changed several times in the period under analysis, we considered the broadest territorial definition of each *comarca*. As such, for the years with a more disaggregated territorial organization, the data were aggregated as if the *comarca* had kept the same territorial scope for the whole period. In addition to the courts identified as belonging to a specific *comarca*, the database includes information on courts that have a regional scope that is broader than the *comarca*, including courts specialized in more complex cases (*tribunais de círculo*) and courts specialized in a particular law area, namely in labour, family and minors or preliminary criminal enquiries.[6] Taking into account that this paper focuses on the relationship between available resources and case flows, we included these courts in the *comarca* where they are located.

This approach leads to the use of a definition of *comarca* that is different from the official one. However, it allows to overcome incomplete reporting (for instance, with the establishment of *tribunais de círculo*, the information regarding judicial staff in some *comarcas* includes the staff working in these courts) and maintains the correspondence between case flows and the human resources allocated to deal with them. Moreover, this approach is suited for the specific analysis made in this paper, which takes advantage of the heterogeneity across *comarcas*, independently of their actual territorial boundaries. The database excludes information on *tribunais de execução de*

4. The number of incoming cases was corrected whenever, as a result of the creation of new *comarcas* or new judgeships within a *comarca*, there was an unusually high number of incoming cases resulting from the transfer of cases which are reported as resolved in one organizational unit and incoming in another unit. For further details regarding this correction, see Pereira and Wemans (2015).

5. For around 3 per cent of the observations, although there were cases resolved, there was no judge assigned to the *comarca*. To fill this gap we obtained information on the judges sitting in *comarcas agregadas* from the nominal lists of judges available at the Conselho Superior da Magistratura site since 2005. These judges are allocated to a specific *comarca* but resolve cases in two different ones. Due to the lack of information regarding the time spent by the judge in each of them, a value of 0.5 was allocated to both. For the remaining observations – where information was only missing in some years – the number of judges and non-judge staff was interpolated.

6. For a description of the organization of the Portuguese Justice system, see Gouveia et al. (2012), volume I.

penas, for which DGPJ stopped reporting information in 2010, as well as courts with a national scope and the two commerce courts.

Regarding courts included in the database that have a scope broader than the *comarca*, it is important to address their relevance in terms of civil justice. In the case of *tribunais de círculo*, which were closed down in 2000 and dealt with more complex cases, their weight on resolved civil cases[7] was around 4 per cent. This percentage is similar when it comes to family and minors courts, which are mainly focused on issues related to minors. As regards labour courts, which deal primarily with labour law, their relevance in the civil area is residual (around 1 per cent of total civil cases). Even less significant, as expected, is the number of civil cases resolved in courts dedicated to preliminary criminal enquiries. It is important to note that, while almost all labour cases are resolved in labour courts, only half of the cases related to minors are resolved in family and minors courts. Therefore, the courts that have the *comarca* as the territorial scope resolve mainly civil and criminal cases, but also some minors' cases.

Main indicators of resources and performance of the judicial system

Evolution between 1993 and 2013

The number of first instance cases resolved in Portugal changed significantly during the period under analysis.[8] Excluding 1993 and 1994[9], there has been an overall increasing trend, with civil cases representing more than half of all resolved cases (Figure 1A). Within civil cases, enforcement cases, intended to demand the fulfilment of an obligation that was previously established, gained predominance over declarative ones, aiming at the definition of a particular right. This composition shift was namely related to the gradual generalisation of the injunction procedure (Figure 1B).[10] Note that the significant increase in the number of resolved cases in civil law in 2013 is related to the measures to end enforcement procedures set in decree-law no.

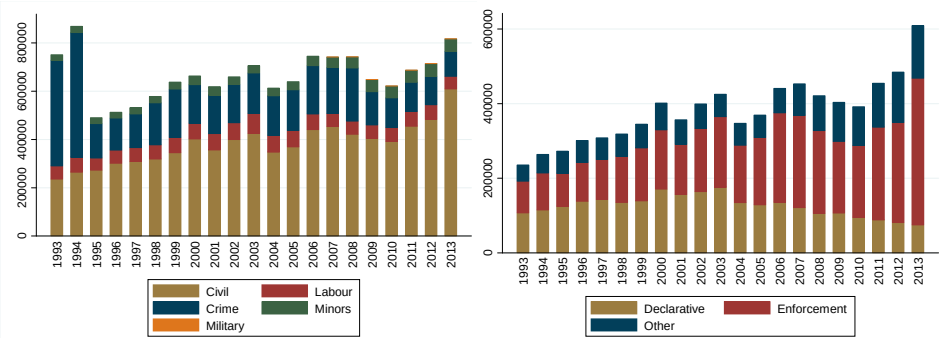
7. In this article, by resolved cases we mean the total resolved cases less the number of cases transferred (for further details, see Direção-Geral da Política de Justiça (2014b)).

8. Although the data for 2014 and 2015 were released in April 2016, this information is not considered here as it reflects the major changes related to the implementation of the New Judicial Map (see Introduction).

9. Note that, in 1995 the flows of criminal cases had a sharp reduction related to the decriminalization of some minor offenses, as discussed in Gomes (2006).

10. As mentioned in Pereira and Wemans (2015), the injunction procedure was created in 1993, but its use was rather limited. Legislative changes implemented in 1998, 2003 and 2005 led to a gradual increase in the number of injunction procedures filed.

4/2013.[11] During the period analysed, the change in the number of judges allocated to first instance courts was broadly in line with changes in the number of resolved cases, each judge ending around 550 cases per year (once again, excluding the first two years).



(A) By area of law (B) Civil litigation - by type of case

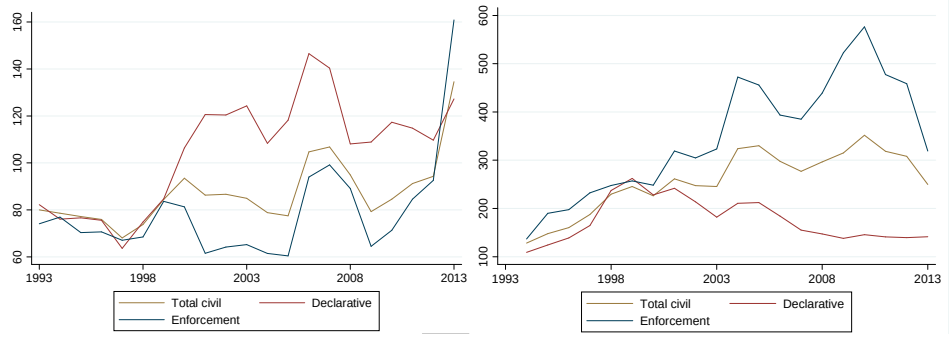
FIGURE 1: Resolved Cases

Sources: DGPJ and authors' calculations.

Focusing on civil justice in particular, it is important to highlight some indicators (constructed following the formulas presented in Appendix A) that capture the capacity of the system to deal with citizens' requests. The clearance rate – ratio between resolved and incoming cases – indicates that, with the exception 2006, 2007 and 2013 (years when measures to reduce pending cases were implemented) the number of civil cases resolved by the justice system has been always below the number of incoming cases, as the clearance rate has stood under 100 per cent (Figure 2A).[12] This explains the considerable growth in the number of pending cases that led to the increasing trend in the congestion rate (ratio of pending and resolved cases) (Figure 2B). The Portuguese judicial system is, thus, characterized by a very high congestion level: taking into account 2013 numbers of resolved and pending cases, the system would need two years and three months to solve all pending cases. The analysis of these indicators by type of case clearly shows that, particularly after 2000, the congestion problem is much more severe as regards enforcement cases than for declarative ones.

11. This decree-law established a number of measures to reduce pending enforcement cases, including the broadening of rules for the extinction of proceedings.

12. The values published in *estatísticas da justiça* for 2014 and 2015 point to the maintenance of clearance rates above 100 per cent (Direção-Geral da Política de Justiça 2016). If maintained for a considerable number of years, this would allow for a consistent reduction in the number of pending cases in civil justice.



(A) Clearance Rate (B) Congestion Rate

FIGURE 2: Performance Indicators - civil litigation

Note: The formulas for calculating these indicators are detailed in Appendix A.
Sources: DGPJ and authors' calculations.

Taking into account that justice system delays, namely as regards civil litigation, may contribute to an inefficient allocation of resources by economic agents, thereby restraining economic growth, it is important to look into the duration of cases. The average duration of resolved cases had an upward trend between 1993 and 2007, posting a decline thereafter, concentrated on declarative cases (Figure 3). In 2006 and 2007 the figures are affected by the measures to reduce the backlog of pending cases, encouraging the termination of old cases, as mentioned in Direção-Geral da Política de Justiça (2010). In parallel, the increase posted in 2013 for enforcement cases is most probably related with the implementation of the abovementioned decree-law no. 4/2013 (Direção-Geral da Política de Justiça 2014a) that led to the termination of a significant number of older cases. Overall the duration of civil cases stood at around 30 months, a figure which clearly signals the system lengthiness, especially for enforcement cases (around 40 months) as opposed to declarative ones (around 18 months).

International comparison

Although the high heterogeneity of justice systems hampers a direct comparability of summary indicators, the data regularly published by CEPEJ are an important benchmark. Indeed, these data serve as a reference to gauge how the Portuguese justice system compares with its European peers as regards resources allocated and efficiency. Taking into account the 2012 results, which are the most recent with information for Portugal, Table 1 presents some key indicators of judicial systems in the European Union

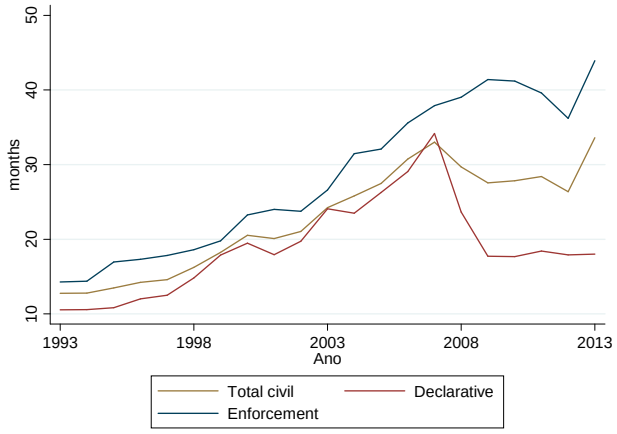


FIGURE 3: Average duration of resolved cases

Sources: DGPJ and authors' calculations.

countries followed by CEPEJ.[13] As justice systems with the same legal origin tend to be more homogeneous and, consequently, can be seen as more directly comparable, we include information on legal origin based on Djankov *et al.* (2007).

Regarding human resources allocated to the system, Portugal has a number of both judges and non-judge staff *per capita* below the overall average and above but close to the average of countries with a French legal origin. Concerning financial resources, the budget for courts per inhabitant is slightly above the European average[14], even for 2012, when temporary cuts to government employees salaries were in place.[15] Moreover, Portugal does not clearly stand out in terms of the level of litigation, as the number of incoming cases *per capita* is close to the average, as discussed in Pereira and Wemans (2015).

The litigation rate and performance indicators presented are for litigious cases other than criminal, which correspond, in the Portuguese system, to civil, labour and minors' cases. In addition, these indicators exclude enforcement cases that are precisely those which present worse performance

13. The most recent CEPEJ report was published in October 2016 (based on 2014 data) but it does not include results for Portugal regarding case flows.

14. Note that Portugal stands out more clearly if we consider this budget in percentage of GDP, showed in brackets in Table 1.

15. In Portugal, compensation of employees represents around 90 per cent of total court expenditure, the second highest percentage of the countries under analysis.

indicators among civil cases (see Figures 2A and 2B).[16] As regards the clearance rate, the results for Portugal are similar to the average for the countries with the same legal origin, but below the overall average. In addition, the country is among the group of 9 countries where pending cases increased in 2012. As an indicator of congestion, Table 1 includes the estimated clearance time. This indicator takes into account the ratio of pending to resolved cases and, considering the 2012 figures, the Portuguese justice system would take 369 days to end all pending cases, a figure that stands above the average.

In a nutshell, the international comparison of summary indicators of the justice system highlights that Portugal has a level of litigation and allocation of resources close to the average of other European countries with the same legal origin. In addition, the country shows up amongst the group of countries with clearance rates below 100 per cent, presenting a congestion level which is slightly above average, even excluding the enforcement cases that have deeper congestion problems, as shown in Figure 2B. This result suggests that there is a wide margin to improve efficiency of the Portuguese judicial system as far as allocation of resources is concerned, in order to close the gap to the top performers.

The explanatory factors of the judicial sector performance are highly complex and it is important to stress that, although this paper covers some relevant issues, there are several others that could only be analysed with case-level data, notably covering cases' procedural steps.[17] Indeed, differences in efficiency can arise from legislation, procedure rules or the behaviour of different players in the system, namely judges and lawyers. It should also be mentioned that, in order to deal with the backlog of pending cases, several measures were implemented over the last years, including changes to the legislation, which may not yet be visible in the data presented but may prove effective in the medium run.[18]

16. The exclusion of enforcement cases is justified by the different treatment of these cases in the countries under analysis, and leads to a better comparability of the results.

17. For an example of a study based on this type of data, see Gouveia *et al.* (2012b).

18. For instance, Portugal appeared as one of the countries with a higher degree of formality, according to Djankov *et al.* (2003), concerning the procedures needed to evict a tenant for non-payment. Such procedures were, however, considerably eased in the recent past.

Country (legal origin)	Litigation	Judges	Non-judge staff	Budget (% of GDP)	Clearance rate	Estimated clearance time
Austria (G)	1.2	15.7	54.8	- (-)	101	135
Belgium (F)	6.8	11.6	48.9	- (-)	-	-
Bulgaria (G)	-	16.3	82.6	17.2 (0.3)	-	-
Croatia (G)	4.3	32.3	162.6	36.7 (0.4)	95	457
Cyprus	-	10.4	49.0	35.4 (0.2)	84	-
Czech Republic (G)	3.5	17.7	86.9	35.3 (0.2)	103	174
Denmark (N)	0.8	4.6	32.5	43.4 (0.1)	109	165
Estonia	1.3	13.0	74.4	23.1 (0.2)	112	167
Finland (N)	0.2	13.7	40.8	46 (0.1)	103	325
France (F)	2.6	7.6	33.2	44.5 (0.1)	99	311
Germany (G)	2.0	18.5	66.9	103.5 (0.3)	100	183
Greece (F)	5.8	13.7	48.2	- (-)	58	469
Hungary (G)	4.4	16.9	82.2	32.9 (0.3)	105	97
Ireland (E)	3.9	3.0	20.6	23.3 (-)	-	-
Italy (F)	2.6	8.3	39.7	50 (0.2)	131	590
Latvia (G)	2.2	12.9	78.6	21.8 (0.2)	118	241
Lithuania (F)	3.6	22.8	87.2	17.7 (0.2)	101	88
Luxembourg	0.9	35.4	67.6	- (-)	173	73
Malta	1.0	8.1	85.4	27.4 (0.2)	114	685
Netherlands (F)	-	11.1	37.3	63.7 (0.2)	-	-
Poland (G)	2.8	24.5	106.0	35.8 (0.4)	89	195
Portugal (F)	**3.5**	**14.1**	**58.3**	**45.5 (0.3)**	**98**	**369**
Romania (F)	5.2	9.4	43.6	15.2 (0.2)	99	193
Slovakia (G)	3.0	16.1	82.8	28.2 (0.2)	82	437
Slovenia (G)	3.1	38.2	161.7	78 (0.5)	101	318
Spain (F)	3.8	7.9	97.3	80.9 (0.4)	100	264
Sweden (N)	0.7	8.0	54.1	66.7 (0.2)	99	179
Average	**2.9**	**15.2**	**69.7**	**42.3 (0.2)**	**103**	**278**
Average (F)	**4.2**	**11.8**	**54.8**	**45.4 (0.2)**	**98**	**326**

TABLE 1. Resources and performance indicators of judicial systems in 2012

Notes: Legal origin - German (G), French (F), English (E), Nordic (N). Litigation - non-litigious, other than crime incoming cases per 100 inhabitants. Judges - no. of professional judges in first instance courts per 100 inhabitants. Non-judge staff - no. of non-judge staff per 100 inhabitants. Budget - total budget of courts in euros per inhabitant. Clearance rate - non-litigious, other than crime cases (see appendix A). Estimated clearance time - Estimated clearance time in days for non-litigious, other than crime cases (see appendix A).
Sources: Djankov *et al.* (2007) and CEPEJ-STAT dynamic database (accessed on 10 October 2016).

Territorial distribution of demand, resources and congestion

The average number of incoming civil cases in a certain *comarca* can be seen as a measure of demand for civil justice directed to courts with jurisdiction there. Therefore, it is interesting to analyse the connection between this demand and

the human resources allocated to those courts.[19] As previously mentioned, these resources are not exclusively allocated to civil justice, but also to other litigation areas such as crime, labour and minors.

Taking into account the average number of incoming civil cases between 1993 and 2013, we separated out the *comarcas* into two groups of the same size.[20] Regarding the territorial distribution, small *comarcas* are almost exclusively located in inland *círculos* (which tend to have population density below average) or on the islands.

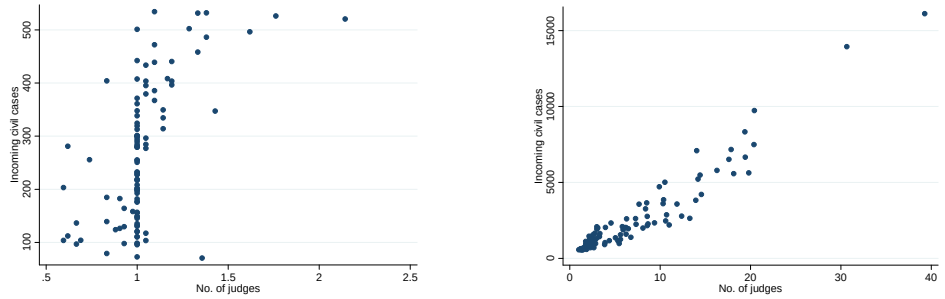
Furthermore, the restriction concerning the allocation of at least one judge to each *comarca*[21] is very biding in the group of small *comarcas* and makes the number of judges largely independent from demand for this group (Figure 4A). As a result, most of these *comarcas* have on average one judge, while the number of incoming civil cases ranges between less than 100 and more than 500. In contrast, for large *comarcas* there is a positive relationship between demand and the number of judges (Figure 4B). As a consequence, there is a lower dispersion in the distribution of incoming cases per judge for this last group, with a coefficient of variation of 0.43 as opposed to 0.57 in small *comarcas*. This distribution reflects, as expected, a higher pressure on judges located in large *comarcas* (the median is 234 incoming cases per judge in small and 381 in large *comarcas*), but there is a significant overlap of the two distributions (Figure 5). Regarding non-judge staff, their allocation does not face a similar restriction, and there is a positive relationship between incoming civil cases and the number of non-judge staff, regardless of the size of the *comarca*. The differences in the number of incoming cases per judge could be explained by a different pressure from other legal areas on judge's workload. However, the results remain valid if one considers all litigation, instead of civil litigation only. The only change is a smaller overlapping of the distributions on Figure 5.

Expenditure on judges' wages can be regarded as a proxy for the average experience of the judges allocated to the *comarca*, as there is evidence that the career of judges in Portugal is highly based on tenure (see Centeno and Pereira 2005). In particular, the connection between wages and experience is quite strong during the first half of the career, stage clearly overrepresented in a sample of first instance judges. In this context, there is some evidence that judges are, on average, more experienced in large *comarcas*, as spending

19. This relationship mainly mirrors decisions regarding the allocation of resources in the territory, but could also be influenced by the response of demand to changes in the availability of resources, as there is evidence of rationing by waiting line in the Portuguese judicial system (Pereira and Wemans 2015).

20. For the purpose of replicating the results, the list of the *comarcas* in each group is available upon request.

21. As mentioned in a previous footnote, there are exceptions to this rule as some very small *comarcas* are sometimes linked to neighbouring ones. However, these happens in very few cases.



(A) Small *Comarcas* (B) Large *Comarcas*

FIGURE 4: Incoming cases vs the number of judges

Note: Lisbon and Porto *comarcas* were excluded from figure B, as the average number of incoming cases in these areas is very high. The figures depict averages for each *comarca* between 1993 and 2013.
Sources: DGPJ and authors' calculations.

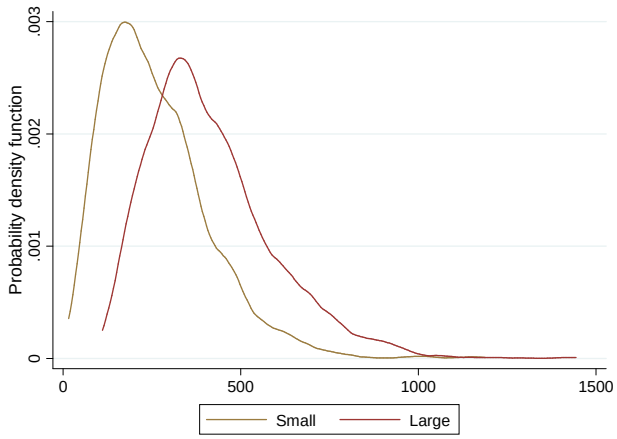


FIGURE 5: Distribution of the number of incoming cases per judge in small and large *comarcas*

Sources: DGPJ and authors' calculations.

on wages per judge in this group is higher than in small *comarcas*. However, the correlation between the size of the *comarca* and average wage per judge is quite low (0.12).

It is also relevant to gauge if there are significant differences between the two groups regarding performance indicators. Concerning the average duration of resolved cases, small *comarcas* usually present more favourable outcomes (Figure 6), although the difference is not very clear. In addition,

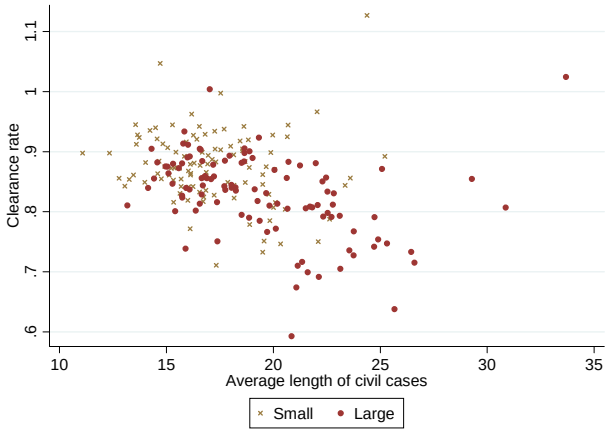


FIGURE 6: Duration and clearance rate for small and large *comarcas*

Note: The figure depicts averages for each *comarca* between 1993 and 2013.
Sources: DGPJ and authors' calculations.

regarding clearance rates the difference between the two groups is even less clear, as the distributions are very similar.

The analysis of the distribution of demand and resources in the territory shows a high heterogeneity in the ratio of incoming cases per judge in small *comarcas*. Additionally, although these *comarcas* have on average a lighter caseload per judge than large ones, this does not clearly translates into better performance as regards disposition time or congestion. Indeed, the prevalence of clearance rates below 100 per cent, which is on the basis of congestion problems, is common to both small and large *comarcas*.

Some determinants of productivity in civil justice

Variables

The specification estimated in this article intendeds to explain the number of resolved cases per judge in civil justice (variable *ResCiv*). As mentioned in the introduction, this dependent variable appears, as a productivity indicator, preferable to the length of resolved cases that is also available in the database. The estimation is based on a panel dataset (210 *comarcas*, followed over the period 1993-2013). Given that this panel covers a relatively long time horizon, we chose a dynamic specification that includes the lagged dependent variable, that is, the number of civil cases completed per judge in the previous year.

A first group of explanatory variables refers to the caseload in each *comarca*, captured by the number incoming civil cases in the year and pending

ones at the end of the year before per judge (respectively, *IncCiv* and *PendCiv*). For the first of these variables, a quadratic term was also included, which aims to capture a possible non-linear response of resolved cases to incoming ones.

A second group of variables includes measures of specialization in civil litigation. The first one flags the *comarcas* where there were courts with specific jurisdiction in this area. Specific jurisdiction is defined as court specialization concerning the applicable form of procedure or the value of the claim, for example, civil judgeships (*varas cíveis*) that deal only with claims above a certain amount (*ClaimSpec*).[22] The second variable reflects the percentage of civil cases completed in judgeships that deal mostly with civil cases (*CivSpec*). This variable therefore reflects *de facto* specialization and not *de jure* specialization, flagging judgeships in which, in a given year, more than 80 percent of resolved cases were civil.[23] A third variable that reflects the proportion of enforcement cases in overall civil cases in the *comarca* (*WEnfor*) has also been included.

The number of cases resolved by a judge in a given *comarca* in other areas, such as criminal justice, will also have an impact on the number of resolved civil cases. In fact the judges considered deal in general with both civil and non-civil justice. Therefore there should be a negative «rivalry» effect between the resolution of civil and non-civil cases (in particular as regards the time spent by judges) that will be all the more intense the greater the degree of utilization of available resources. However, this effect may be mitigated by differences in the productivity of judges: a more productive judge will tend to resolve more cases whatever the litigation area.

This kind of effects is first captured by the number of resolved criminal cases per judge (*ResCrim*), in which we consider separately ordinary and special proceedings, misdemeanour and a residual category. In addition, in the *comarcas* whose data do not include labour courts or family and minors courts, cases pertaining to labour and minors law (*ResLab* and *ResMin*) were taken on board as well, as these are heard in the same courts as civil cases. In the *comarcas* whose data include both the cases and resources belonging to labour courts or family and minors courts (see section Data), binary variables have been added (*LabCou* and *FamCou*), in order to capture the impact on resolved civil cases. Note that *comarca* fixed-effects (see below) should capture to a large extent such an impact, as there have been only a few changes during the sample period in terms of the creation or extinction of these courts.

22. More specifically, the variable flags the *comarcas* in which there was at least one court with the following designation: *vara cível*, *vara mista*, *tribunal de pequena instância cível* or *juízo de execução*.

23. This percentage is intended to approximate a substantial degree of specialization in civil justice, but it is arbitrary. We experimented with 90 percent, and the results shown below did not change significantly.

With regard to resources allocated to the courts, we only have variables related to human resources: judicial staff per judge in the *comarca* (*JudSt*) and, for the period after 2006 only, a variable that intends to approximate the experience of judges (*Exper*) through their average salary. As said, a strong association between salary and experience is documented for this professional group.[24] It would have been useful to have variables capturing the availability of equipment, particularly as regards investment in IT, but such information was unavailable.

As for the judicial organization of *comarcas*, we experimented with including in the regression a binary variable for those where *tribunais de círculo* (that existed between 1993 and 1999) were located. However, this variable was not significant, perhaps due to the relatively small number of cases heard by these courts, despite the fact that such cases were more relevant, notably, as far as the value of the claim was concerned (for civil justice).

The regression has an indicator that measures purchasing power of *comarcas* (*PurcPower*) - see Pereira and Wemans (2015) - in logs, to approximate their economic development. The latter should impact the characteristics of civil litigation, e.g. its complexity. We also considered the number of incoming civil cases (in logs), intended to capture the «size» of *comarcas* (*Size*).

The regression includes *comarca* fixed-effects (α) that model a multiplicity of time-invariant factors impacting on the number of resolved cases. These include the differences among *comarcas* as to the characteristics of litigation (i.e. the casemix) and judicial organization, whenever there has been no substantial changes over the sample period. In particular, these fixed-effects will capture the bulk of the impact of including data on labour courts and family and minor courts in the *comarcas* where they are located.

Finally, the regression includes year fixed-effects (δ) that model the impact on resolved cases of factors affecting similarly the various *comarcas*, such as methodological changes to justice statistics[25] or the legislative measures to reduce the backlog of pending cases mentioned in the previous section.

The estimated specification for the complete sample period is thus as follows:

24. Naturally there are wage variations not due to changes in the experience of judges, such as changes in the wage scale and the wage cuts and reinstatements in the last few years of the sample period. Thus, the average salary of each *comarca* was taken against the average salary for all of them in a given year.

25. An example of these changes is the one in 2007 in the procedure for collecting information that started to be made directly from the courts' computer system.

$$
\begin{aligned}
ResCiv_{i,t} = {} & c + \beta_1 ResCiv_{i,t-1} + \beta_2 IncCiv_{i,t} + \beta_3 IncCiv_{i,t}^2 \\
& + \beta_4 PendCiv_{i,t-1} + \beta_5 CivSpec_{i,t} + \beta_6 ClaimSpec_{i,t} \\
& + \beta_7 WEnf_{i,t} + \beta_8 ResCrim_{i,t} + \beta_9 ResLab_{i,t} \\
& + \beta_{10} ResMin_{i,t} + \beta_{11} LabCou_{i,t} \\
& + \beta_{12} FamCou_{i,t} + \beta_{13} JudSt_{i,t} \\
& + \beta_{14} Size_{i,t} + \beta_{15} PurcPower_{i,t} + \alpha_i + \delta_t + \varepsilon_{i,t},
\end{aligned}
\tag{1}
$$

where i indexes the *comarca* and t the year. This specification is estimated for all *comarcas* and also for the subsets of large and small ones (the definition used in the previous section has been kept). Another estimated specification includes the indicator of experience of judges, for the subsample 2007-2013. Appendix B presents the descriptive statistics of the variables included in the regressions.

Taking into account that declarative and enforcement cases differ much in terms of substance and procedure, the weight of the latter in overall civil cases is held constant in the specification above. Nevertheless, we also estimate regressions taking resolved declarative and enforcement cases as the dependent variable (changing the definition of incoming and pending cases accordingly). The role of judges in the resolution of enforcement cases has been played down with the reform of the respective procedure rules (from 2003 onwards). Therefore, the explanatory power of the regression in which enforcement cases show up as the dependent variable will be smaller, for we do not include variables that capture the intervention of enforcement agents, which has been gaining prominence.

Econometric issues

The dynamic panel (1) can be estimated consistently by the Arellano and Bond (1991) estimator, under the usual conditions (see, for example, Wooldridge (2002, chapter 11)). The fixed-effects estimator for panel data is not consistent in this case, but its results are still presented as a benchmark. The Arellano-Bond estimator instruments the lagged dependent variable by a certain number of the respective lags. Given that pending cases may respond to the number of cases resolved in previous periods, such variable was instrumented in a similar way. In the implementation of the Arellano-Bond method, particularly in a long panel as ours, the problem of an excess of instruments arises as the number of lags used increases. To address this problem, the methods suggested in the literature (Roodman 2009) have been followed, namely, the use of a relatively small number of lags to construct the matrix of instruments as well as collapsing the latter. In addition, the robustness of coefficients to the change in the number of lags was checked for each one of the regressions. Coefficients are, in general, robust (the exceptions are indicated

in the text). In order to have a general indication about the validity of the instrumentation, the results of Hansen's overidentification test are presented.

Another relevant econometric point is that, as mentioned, the resolution of civil cases occurs simultaneously with the resolution of criminal, labour and minor' cases. Therefore, resolved non-civil cases cannot be considered exogenously determined in the regression above, and were instrumented by the number of incoming cases in the same litigation area. Such instrumentation strategy is justified, firstly, by the high degree of correlation between incoming and completed cases within each area. Moreover, it seems reasonable to assume that resolved civil cases do not respond directly to criminal, labour and minors' cases filed (although they may indirectly respond via the variables included in the regression above, particularly the number of incoming civil cases).

Results

Table 2 shows the estimation results for all of *comarcas* and taking overall civil cases as the dependent variable, in the full sample (including the results for the fixed-effects estimator) and in the most recent period. This last regression includes the indicator of average experience of judges in the *comarca*, allowing at the same time to study the changes vis-à-vis the entire sample period.[26] Table 3 presents separate estimates for declarative and enforcement cases, as well as for large and small *comarcas* (overall civil and full sample). The Hansen statistic indicates the non-rejection of the null hypothesis in the regressions for overall civil (shown in table 2). However, two of the regressions presented in table 3 have symptoms of endogeneity: the one that has enforcement cases as the dependent variable and the one referring to small *comarcas*. Even taking into account the conservative instrumentation strategy followed (precisely to avoid weakening Hansen's statistic), the results of these regressions must be viewed with caution. We present them nevertheless, in order to compare with the remaining results. In any case, the conclusions drawn in this section are based on evidence following from the full set of regressions run.

The coefficient of the lagged dependent variable is significant in the fixed-effects regression, but not in the Arellano-Bond regression for overall civil, both in the full sample and in the post-2007 sample. However, the results of the full-sample regression are sensitive to the instrumentation procedure, being significant and with a magnitude similar to the one in the fixed-effects regression when instruments are not collapsed. Moreover, in the regressions with declarative and enforcement cases as the dependent variable, the coefficient at issue is always significant. The smaller persistence of resolved

26. In fact the results changing the sample period only (i.e. without adding the new variable) are quite similar to those presented, both in terms of significance of regressors and size of the statistically significant coefficients.

declarative cases may have several explanations, such as the more important role of judges in their resolution that may lead to a greater fluctuation in the completion of cases, associated with their rotation across *comarcas*. Comparing the sets of estimates in the Arellano-Bond and fixed-effect regressions, these are generally close, with the exception of the coefficient of the lagged dependent variable, already mentioned, and the coefficient of the lagged pending cases (see below).

The number of resolved civil cases per judge varies positively with incoming ones (a result that holds across all regressions), meaning that productivity of judges responds to the pressure put by demand on the judicial system. Such an evidence helps explain the relative homogeneity in performance indicators between small and large *comarcas*, notwithstanding the differences in the volume of civil litigation they face, as documented above. In general terms, this type of effect is described both for Portugal (Borowczyk-Martins 2010) and other countries (e.g. Dimitrova-Grajzl *et al.* 2012, for Slovenia, and Beenstock and Haitovsky 2004, for Israel). Beenstock and Haitovsky interpret this increase in effort as a response to pressure as an attempt by judges to prevent an increase in congestion in the jurisdictions for which they are responsible.

In addition, in the regression for overall civil, the estimated coefficient of the quadratic term is negative and statistically significant, indicating that, as the number of incoming cases grows, resolved cases increase at a progressively lower rate. This is expectable given the more intense use of resources, as incoming and completed cases grow. Exemplifying with the estimates for the full-sample regression, combining the linear and non-linear terms in the average of incoming cases per judge, 100 additional cases filed lead to an increase by about 50 in resolved ones. Such nonlinear effect does not hold, however, for small *comarcas*, indicating less pressure on resources there. This evidence, coupled with the heterogeneity in the relationship between incoming cases and the number of judges documented for the smaller *comarcas*, suggests that there is scope for increasing efficiency with a more equitable sharing of caseload among judges through territorial aggregation, in the spirit of the New Judicial Map, implemented in 2014.

The impact of pending cases is negative and significant for overall civil in the full sample. However, this result is not robust to the variation in the number of lags in the implementation of the Arellano-Bond estimator, losing statistical significance when instruments are not collapsed. In the fixed-effects regression, the coefficient is positive and significant, but this change in sign could result precisely from the correction of endogeneity. For the more recent sample period and when declarative and enforcement cases are considered separately, the coefficients of pending cases are not statistically significant. This apparent lack of impact of pending cases on judges' productivity does not imply that they only deal with the new cases, but suggests that they establish

	Full sample	Sample 2007-13	Fixed-effects estimator
Resolved civil / judge(t-1)	0.05	0.16	0.16***
	0.16	*0.21*	*0.01*
Incoming civil/ judge (100 cases)	72.80***	76.99***	68.40***
	17.35	*21.15*	*4.09*
Incoming civil / judge^2	-3.03***	-2.83*	-1.74***
	0.94	*1.45*	*0.27*
Pending civil / judge(t-1) (100 cases)	-32.94**	-28.63	3.22***
	13.22	*21.23*	*0.47*
Specialization civil (perc.)	0.51***	0.76***	0.53***
	0.11	*0.20*	*0.06*
Claim type spec. (binary var.)	64.17**		6.45
	31.58		*10.03*
Weight enforc. cases (perc.)	2.60***	4.27***	1.86***
	0.37	*1.07*	*0.13*
Resolved criminal (ord.) / judge	-0.51**	-0.15	-0.51***
	0.21	*0.57*	*0.07*
Resolved labour / judge	-1.69	-5.05	-0.63**
	1.04	*5.44*	*0.31*
Resolved minors / judge	0.14	0.49	0.03
	0.53	*0.74*	*0.08*
Judicial staff / judge	8.00***	4.71	9.51***
	2.36	*4.29*	*0.96*
Ind. experience of judges		-27.83**	
		13.11	
Size of *comarca*	-88.22***	-57.22	-39.78***
	28.21	*56.02*	*7.10*
Purchasing power of *comarca*	-70.99*	-91.79	20.39**
	36.31	*116.79*	*10.26*
Hansen Test (p-value)	0.30	0.13	
N. of instruments	43	28	
N (*Comarcas*)	210	192	210
T (Years)	19	6	19

TABLE 2. Determinants of productivity in the resolution of civil cases

Notes: Regressions estimated by the Arellano-Bond method (except for the third column), instrumenting resolved and pending civil cases in the previous year by their lags (2nd to 5th) and collapsing the instruments as in Roodman (2009). In all regressions, resolved non-civil cases were instrumented by the incoming ones. In addition to the variables in the table, it is controlled for the existence of a labour court or family and minors court in the *comarca*, for the other categories of resolved criminal cases (special, misdemeanour and others) and *comarca* and year fixed-effects. Robust standard deviations in italics. P-values: * <0.1; ** <0.05; *** <0.01.

their objectives of resolution of cases with reference to the number of those entering in the year.

With regard to specialization, judges tend to be more productive in the resolution of civil cases in *comarcas* where judgeships dealing almost exclusively with this litigation area have more importance. This positive effect

	Declarative cases	**Enforcement cases**	**Small *comarcas***	**Large *comarcas***
Resolved civil / judge(t-1)	0.27***	0.42***	0.16	0.26
	0.06	*0.12*	*0.16*	*0.16*
Incoming civil / judge (100 cases)	70.64***	79.88***	68.79***	104.14***
	11.26	*10.83*	*14.19*	*18.68*
Incoming civil / judge^2	-3.83**	-3.04**	-1.91	-4.23***
	1.81	*1.35*	*1.55*	*0.94*
Pending civil / judge(t-1) (100 cases)	3.61	1.67	4.36	-27.55***
	2.7	*11.26*	*8.91*	*9.33*
Specialization civil (perc.)	0.02	0.54***	0.27***	0.61***
	0.04	*0.11*	*0.10*	*0.18*
Claim type spec. (binary var.)	1.61	16.25		50.21*
	12.08	*10.91*		*26.77*
Weight enforc. cases (perc.)			1.46***	4.02***
			0.30	*0.56*
Resolved criminal (ord.) / judge	-0.26***	-0.52***	-0.74*	-0.30
	0.08	*0.17*	*0.38*	*0.21*
Resolved labour / judge	-0.36	-1.34	-1.60	-0.60
	0.28	*1.14*	*1.93*	*0.46*
Resolved minors / judge	-0.25**	0.03	1.11***	-0.05
	0.11	*0.28*	*0.39*	*0.76*
Judicial staff / judge	2.89***	10.55***	8.44***	3.27
	0.92	*2.32*	*3.14*	*4.57*
Size of *comarca*	-19.37***	-49.62***	-63.31***	-182.64***
	5.15	*11.36*	*18.88*	*50.11*
Purchasing power of *comarca*	4.33	-59.14**	-49.18	8.10
	12.36	*28.95*	*40.08*	*47.97*
Hansen test (p-value)	0.14	0.00	0.03	0.50
N. of instruments	43	43	40	43
N (*Comarcas*)	210	210	105	105
T (Years)	19	19	19	19

TABLE 3. Determinants of productivity by case type and size of *comarca*

Notes: The size of *comarcas* is defined by reference to the number of incoming civil cases. Regressions estimated by the Arellano-Bond method, instrumenting resolved and pending civil cases in the previous year by their lags (2nd to 5th) and collapsing the instruments as in Roodman (2009). Resolved non-civil cases were instrumented by the respective incoming ones. In addition to the variables in the table, it is controlled for the existence of a labour court or family and minors court in the *comarca*, for the other categories of resolved criminal cases (special, misdemeanour and others) and *comarca* and year fixed-effects. In the regression for declarative (enforcement) cases, all case flow variables refer to them and it is controlled, in addition, for the number of enforcement (declarative) cases. Robust standard deviations in italics. P-values: * <0.1; ** <0.05; *** <0.01.

of specialization in civil litigation (relative to other areas) on the number of completed cases per judge shows up in all regressions, except the one for declarative cases. As regards judgeships of specific jurisdiction - specialization within the civil area, by the value of the claim - the evidence of an effect on productivity is less robust. In fact, this variable is only statistically significant

in the regression for overall civil in the full sample, and even there this depends on the instrumentation procedure. Note that, regardless of the impact on the number of resolved cases, there may be gains in terms of the quality of decisions which measures based on the number of completed cases do not capture. Quality is an important factor to consider in assessing the effects of specialization.

Productivity in the resolution of civil cases varies positively with the proportion of enforcement cases heard in the *comarca*, indicating that the time spent by judges to resolve them will be shorter than for declarative cases. In a *comarca* where this proportion is 1 percentage point higher, with everything else constant, about 2.5 additional cases per judge are completed (full sample). In addition, there is a stronger impact of this variable (around 4 cases per judge), when the sample is restricted to the more recent years, possibly reflecting the modification of procedure rules, playing down the role of judges. With regard to the number of cases of each type that a judge can resolve in a year, the resolution benchmark figures presented in annex 1 to Direção-Geral da Administração da Justiça (2012) point precisely in that direction: 6500 cases in enforcement judgeships vis-à-vis 550 cases in generic judgeships dealing with civil cases of other type (already taking into account the simplified enforcement procedure in force at the time of publication of that report). It is interesting to note that, despite requiring less time for the judge, figure 2 shows that the duration of completed enforcement cases is, on average, higher than that of declarative cases (increasing steadily over the period analysed). This results from backlog in pending cases that implies a higher proportion of older cases among resolved ones. In addition, note that the duration of enforcement cases can be extended by mere procedural issues that do not involve court intervention, such as instalment payment plans.

The number of ordinary criminal cases[27] resolved per judge has a negative impact on the resolution of civil cases, perhaps reflecting the aforementioned rivalry effect, accentuated by the priority that criminal cases generally enjoy. With regard to completed labour and minors' cases (not dealt with in the specialized courts), there is a lack of effect on overall civil, although there is still a negative effect of resolved minors' cases, when declarative cases figure as the dependent variable. This lack of an effect of completed labour cases may reflect an uninformative sample, given the small proportion of cases of this type heard outside labour courts (see Appendix B). When the sample is divided into small and large *comarcas*, the coefficient of resolved family cases is positive for the first group, perhaps reflecting productivity differences among judges, as the overwhelming majority of such *comarcas* has only one judge (Figure 4A).

27. This category is the most important one within criminal litigation and covers the majority of crimes, with the exception chiefly of certain petty crimes that are included in the special criminal category.

Judicial staff per judge has a positive and statistically significant impact on resolved civil cases - of a greater magnitude for enforcement cases - highlighting the importance of considering jointly the allocation of judges and remaining resources making up the judicial system. The importance of staff in judicial proceedings in Portugal is evidenced in Gomes (2005) who, analysing the procedural acts practiced in a sample of declarative cases, concludes that 61 percent of such acts are conducted by judicial staff.

We estimate a negative effect of the indicator of experience of judges on the number of completed civil cases, a result that may reflect several factors, such as incentives to the resolution of cases and the balance between quantity and quality of judicial decisions. As regards the second interpretation, this result is in line with Backes-Gellner *et al.* (2011) who, for second instance labour courts in Germany, find a negative impact of experience on the number of completed cases, but a positive one on the quality of judicial decisions - measured by the proportion of appeals upheld by a higher court. As already mentioned, one of the important limitations of our data is that we only have strictly quantitative indicators of productivity. It is possible to cite other literature that finds evidence of an improvement in the quality of judicial decisions as judges become more experienced, such as Kosma (1998), although there are also studies that do not find this kind of relationship, such as Posner (1995), both looking at higher courts for the United States.

The size of *comarcas* has a negative coefficient in the various regressions presented, a result that can be read in several ways. It is conceivable that an increase in size is negatively correlated with the availability of physical resources, omitted in the regression, or has implications in terms of court organization, with negative repercussions on productivity. However, given that there is an association between the size of *comarcas* and litigation characteristics, a negative coefficient may also arise from the variable being capturing features that make case resolution more difficult. The regression includes a purchasing power indicator that intends to model the complexity of litigation and also has a negative coefficient for overall civil (and for enforcement cases), pointing to a greater complexity in more developed *comarcas*. However, although this indicator (along with *comarca* fixed-effects) captures certain characteristics of litigation, others may be captured by the size indicator. In fact it is difficult to distinguish the impact on productivity of demand and supply factors based on the size of *comarcas*, given that this variable stems from the litigation itself, but at the same time has implications from the viewpoint of judicial organization. This is mirrored by the high correlation, around 70 percent, between the indicators of purchasing power and size of *comarcas*.

Conclusions

This work focused on the determinants of productivity in civil justice in Portugal, presenting at the same time some descriptive evidence, based on data by *comarca* for the period from 1993 to 2013. The summary indicators of performance of the Portuguese judicial system point to a congestion problem in this litigation area, much more marked for enforcement than for declarative cases. It will be necessary to keep clearance rates well above 100 per cent over a considerable period of time, to substantially bring down congestion and allow the country to move closer to the set of countries with fastest justice systems.

Given the ineffectiveness of backlog reduction plans for solving the structural problems of the system and the medium-term budgetary constraints on resource increases, it is essential to act on the determinants of productivity in the resolution of civil cases. The results of this study indicate that resolved cases per judge respond positively to demand pressure, but in a different way in large and small *comarcas*. In fact, there is evidence of a greater degree of resource use in large *comarcas*. In this framework, more flexible human resource management, in the spirit of the New Judicial Map, will tend to increase productivity while allowing a more balanced distribution of the caseload within the system.

With regard to specialization and in a purely quantitative dimension of productivity, there is a positive effect of specialization in civil cases vis-à-vis other litigation areas. Another result to be highlighted is the importance of judicial staff in case resolution, reinforcing the idea that resources allocated to the system should be considered as a whole in decision making. In the analytical framework of this article, some aspects could not be addressed due to the lack of data. For instance, it would be important to introduce in the analysis the quality of judicial decisions, notably through a measure of the rate of reversal. In addition, it would be interesting to assess the impact of changes in the size of *comarcas* on efficiency, measured taking into account total financial resources.

As regards future research on the impact of a wide range of other factors on productivity, already mentioned, it seems crucial to use disaggregated data at the case level (as in (Gomes 2005) and (Gouveia *et al.* 2012b)), naturally anonymised. The use of this type of data would allow, in particular, to identify the main bottlenecks in court procedures. Finally, the recent improvement in the statistics released for administrative and tax courts should make it possible to carry out quantitative studies focusing on this area.

References

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.

Backes-Gellner, Uschi, Martin R. Schneider, and Stephan Veen (2011). "Effect of Workforce Age on Quantitative and Qualitative Organizational Performance: Conceptual Framework and Case Study Evidence." *Organization Studies*, (32(8)), 1103–1121.

Beenstock, Michael and Yoel Haitovsky (2004). "Does the appointment of judges increase the output of the judiciary?" *International Review of Law and Economics*, (24), 351–369.

Borowczyk-Martins, Daniel (2010). "The Supply of Judicial Services in Portugal: An Empirical Investigation." *Proceedings of the 5th edition of Banco de Portugal's conference Portuguese Economic Development in the European Context*, pp. 47–88.

Centeno, Mário and Manuel Coutinho Pereira (2005). "Wage determination in General Government in Portugal." *Banco de Portugal - Autumn Economic Bulletin*.

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

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

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

Dimitrova-Grajzl, Valentina, Peter Grajzl, Janez Sustersic, and Katarina Zajc (2012). "Court output, judicial staffing, and the demand for court services: Evidence from Slovenian courts of first instance." *International Review of Law and Economics*, (32), 19–29.

Direção-Geral da Administração da Justiça (2012). *Ensaio para reorganização da estrutura judiciária*. Ministério da Justiça, Lisboa.

Direção-Geral da Política de Justiça (2010). *Plano de Acção para o Descongestionamento dos Tribunais - Relatório de Monitorização Janeiro a Dezembro de 2009*. Ministério da Justiça, Lisboa.

Direção-Geral da Política de Justiça (2014a). "Alguns Indicadores Estatísticos sobre os processos nos tribunais judiciais de 1ª instância, 2007-2013." *Destaque Estatístico*, (28).

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

Direção-Geral da Política de Justiça (2016). "Estatísticas da Justiça – Primeiros resultados: Movimento processual nos tribunais judiciais de 1ª instância, 1996-2015." *Destaque Estatístico*, (40).

Djankov, Simeon, Caralee McLiesh, and Andrei Shleifer (2007). "Private Credit in 129 Countries." *Journal of Financial Economics*, 84(2), 299–329.

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

Gomes, Conceição (ed.) (2005). *Os actos e os tempos dos juízes: contributos para a construção de indicadores da distribuição processual nos juizos cíveis*. Observatório Permanente da Justiça Portuguesa – Centro de Estudos Sociais da Universidade de Coimbra.

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, Ana F., Sílvia Santos, and Corinna Herber (2016). "The impact of structural reforms of the judicial system: a survey." *Gabinete de planeamento, estratégia, avaliação e relações internacionais do Ministério das finanças e da administração pública*, 2016(5).

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

Gouveia, Mariana, Nuno Garoupa, and Pedro Magalhães (eds.) (2012b). *Justiça económica em Portugal: Consulta de processos judiciais - uma análise económica*. Fundação Francisco Manuel dos Santos.

Instituto Nacional de Estatística (ed.) (2015). *Custos de contexto: a perspetiva das empresas 2015*.

Kosma, M. (1998). "Measuring the influence of Supreme Court Judges." *Journal of Legal Studies*, 27, 333–372.

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

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

Pompe, Sebastiaan and Wolfgang Bergthaler (2015). "Reforming the legal and institutional framework for the enforcement of civil and commercial claims in Portugal." *IMF Working Paper*, 279.

Posner, R. (ed.) (1995). *Ageing and old age*. University of Chicago Press.

Roodman, David (2009). "A Note on the Theme of Too Many Instruments." *Oxford Bulletin of economic and statistics*, 71, 136–158.

Rosales-López, Virginia (2008). "Economics of court performance: an empirical analysis." *European Journal of Law and Economics*, 25(3), 231–251.

Voigt, Stefan and Nora El-Bialy (2016). "Identifying the determinants of aggregate judicial performance: taxpayers' money well spent?" *European Journal of Law Economics*, 41(2), 283–319.

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

Appendix A: Performance Indicators

$$CongestionRate_t = Pending_{t-1}/Resolved_t, \tag{A.1}$$

$$ClearanceRate_t = Resolved_t/Incoming_t, \tag{A.2}$$

$$EstimatedClearanceTime_t = Pending_t/Resolved_t * 365, \tag{A.3}$$

Note: When we present results for these indicators, whenever feasible, both incoming and resolved cases are corrected for transferred cases (moved between courts).

Sources: DGPJ and CEPEJ.

Appendix B: Descriptive Statistics

Variable	Unit	Observations	Mean	Standart deviation	Min.	Max.
Resolved civil cases per judge	No. / judge	4410	278.7	148.1	3.0	1325.0
Resolved civil cases per judge - declarative	No. / judge	4410	77.7	48.0	1.0	443.8
Resolved civil cases per judge - enforcement	No. / judge	4410	165.6	104.9	1.0	1080.0
Resolved labour cases per judge - Comarcas without LC	No. / judge	4410	1.1	6.6	0.0	150.0
Resolved criminal cases per judge	No. / judge	4410	167.1	233.3	4.0	5793.0
Resolved criminal cases per judge - common	No. / judge	4410	76.8	48.6	0.0	453.5
Resolved criminal cases per judge - special	No. / judge	4410	31.8	25.6	0.0	214.0
Resolved criminal cases per judge - misdemeanour	No. / judge	4410	53.9	214.9	0.0	5657.0
Resolved criminal cases per judge - other	No. / judge	4410	4.6	12.2	0.0	190.0
Resolved minors' cases per judge - Comarcas without FMC	No. / judge	4410	25.8	23.3	0.0	442.0
Incoming cases per judge	100 cases / judge	4410	3.4	1.8	0.2	14.4
Pending cases per judge	100 cases / judge	4410	6.1	4.4	0.3	33.2
Civil specialization	Percentage	4410	13.7	30.9	0.0	100.0
Type of claim specialization	Binary variable	4410	0.0	0.2	0.0	1.0
Percentage of enforcement cases	Percentage	4410	57.4	12.0	1.6	93.6
Judicial staff per judge	No. / judge	4410	7.3	2.6	1.5	30.0
Proxy for the seniority of judges	Salary per judge / average	1327	1.0	0.3	0.2	3.7
Size of the Comarca	100 incoming civil cases	4410	21.4	89.2	0.2	1805.4
Purchasing power index	Index base 100	4410	71.0	27.9	18.9	314.2
Labour Court (LC)	Binary variable	4410	0.2	0.4	0.0	1.0
Family and Minors Court (FMC)	Binary variable	4410	0.1	0.3	0.0	1.0

TABLE B.1. Descriptive statistics - all *comarcas*

Variable	Unit	Observations	Mean	Standard deviation	Min.	Max.
Resolved civil cases per judge	No. / judge	2205	222.9	129.4	3.0	1142.0
Incoming cases per judge	100 cases / judge	2205	2.6	1.5	0.2	11.9
Pending cases per judge	100 cases / judge	2205	4.5	3.6	0.3	33.1
Civil specialization	Percentage	2205	2.1	14.4	0.0	100.0
Type of claim specialization	Binary variable	2205	0.0	0.0	0.0	0.0
Percentage of enforcement	Percentage	2205	56.0	12.7	1.6	93.6
Non judge staff per judge	No. / judge	2205	6.7	2.4	1.5	18.0
Size of the Comarca	100 incoming civil cases	2205	2.6	1.6	0.2	9.5
Purchasing power index	Index base 100	2205	55.6	13.8	18.9	139.9

TABLE B.2. Descriptive statistics - small *comarcas*

Variable	Unit	Observations	Mean	Standard deviation	Min.	Max.
Resolved civil cases per judge	No. / judge	2205	334.5	144.5	56.5	1325.0
Incoming cases per judge	100 cases / judge	2205	4.1	1.8	1.1	14.4
Pending cases per judge	100 cases / judge	2205	7.6	4.5	1.0	33.2
Civil specialization	Percentage	2205	25.3	37.8	0.0	100.0
Type of claim specialization	Binary variable	2205	0.1	0.3	0.0	1.0
Percentage of enforcement	Percentage	2205	58.7	11.1	15.7	89.7
Non judge staff per judge	No. / judge	2205	7.9	2.7	1.8	30.0
Size of the Comarca	100 incoming civil cases	2205	40.2	123.4	2.0	1805.4
Purchasing power index	Index base 100	2205	86.4	29.9	28.2	314.2

TABLE B.3. Descriptive statistics - large *comarcas*

Firm creation and survival in Portugal

Sónia Félix
Banco de Portugal

January 2017

Abstract
In this study we use a very rich panel dataset that allows us to identify new firms at founding and follow them over time. We provide a comprehensive characterization of the dynamics of firm entry and firm exit in Portugal, in the period between 2005 and 2012. In particular, we analyze the distribution of new firm creation and survival by sector of economic activity, size class, and over the business cycle. The results suggest that entry rates are fairly high while survival rates are small. Moreover, the share of new entrants' sales on total sales and the employment share suggest that new firms are in general small. Entry rates and employment shares show as pro-cyclical for smaller firms. (JEL: L11)

Introduction

Newly created firms are an important driver of innovation and job creation. Haltiwanger *et al.* (2013) document that new firms are responsible for most of new jobs in the U.S. and Adelino *et al.* (2016) show that firm entry account for most of net employment creation in response to local demand shocks in the U.S.. Nevertheless, despite the number of firms that starts activity ever year, new firms fail at a significant rate in their first years of life.

The dynamics of firm entry and firm exit have been widely studied in the literature of industrial organization. Geroski (1995) provides a survey of empirical work on the determinants of firm entry and the likelihood of firms to survive over time. The author documents that firm entry is common, with a high number of firms entering most markets in most years, mainly for firms operating at the small-scale. A second stylized fact is that entry rates are rarely high or persistently low over time in particular industries, and that firm entry is generally not synchronized across industries. Moreover, entry and exit rates are highly positively correlated, which is consistent with the organizational ecology population theory developed by Hannan and Freeman

Acknowledgements: I am grateful to Luísa Farinha, Miguel Gouveia, and Pedro Portugal for helpful comments and suggestions. The views expressed are those of the author and do not involve the responsibility of the Banco de Portugal or the Eurosystem. Any errors and omissions are the sole responsibility of the author.

E-mail: scfelix@bportugal.pt

(1989), in which new firms are more likely to survive in populations with a small number of other competing new entrants. More recently, Geroski *et al.* (2010) document that firms that enter in industries with lower entry rates are more likely to survive. Additionally, Geroski (1995) pinpoints that the survival rate of most entrants is low and successful entrants may take a long time to achieve a size comparable to the average incumbent.

Empirical evidence for Portugal suggests that the aforementioned stylized facts hold for Portuguese newly created firms. Mata (1993) presents an overview of the determinants of entry for Portuguese manufacturing firms according to the type of entrant. Geroski *et al.* (2010), Mata *et al.* (1995), and Mata and Portugal (1994) show that market-specific conditions are important determinants of firm survival. It is therefore important to understand market dynamics across economic sectors and over time for Portuguese firms. In fact, little is known about the size and economic activity sector distribution of new firms in Portugal, and how firm creation and survival respond to aggregate economic conditions.

In this study we provide a comprehensive characterization of the dynamics of firm creation and survival in Portugal in the period between 2005 and 2012, using a rich dataset that allows us to identify firms at birth and follow them over time. In particular, we analyze the distribution of newly created firms by firm's sector of economic activity, size class, and over the business cycle.

The results suggest that entry is common across sectors of economic activity and that new firms are in general much smaller than their incumbent counterparts. This finding is corroborated by the low employment share of new firms by firm size[1]. Moreover, while it seems that barriers to entry are modest, barriers to survival seem to be very relevant. In fact, about only 48 percent of new firms survive throughout the sampling period. These low survival rates are independent of the firm's economic activity sector. Moreover, entry rates and employment shares for smaller firms show as pro-cyclical, suggesting that the likelihood of entry is higher during upturns.

This paper is organized as follows: Section 2 describes the data. Section 3 presents the main descriptive facts of firm entry and survival by sector of economic activity and size class. Section 4 studies the dynamics of new firms over the business cycle. Section 5 concludes.

Data

The dataset we use in this study is the Portuguese dataset Simplified Corporate Information - IES (*Informação Empresarial Simplificada*) that covers

1. It is important to highlight that, while the share of new firms' employment in total employment is fairly low, the share in job creation may be important.

the population of virtually all Portuguese nonfinancial corporations[2]. Data on firm's employment are obtained in *Quadros de Pessoal* (QP), which is a dataset compiled by the Portuguese Ministry of Employment and is an annual mandatory employment survey addressed to establishments employing at least one wage earner. QP is available only until 2013.

IES data consists of a new system to collect firm mandatory annual economic, financial, and accounting information for a single entity. Firms report detailed balance sheet information as well as information on several important variables, namely the legal form of the firm. Even though data on IES started being collected in 2006, there was a report collecting data in 2005 that was also taken into consideration in the analysis. Data are available for the period between 2005 and 2014. Our sample consists of firms with limited liability and with at least one employee during the sampling period. Moreover, firms belonging to an economic group are not considered as new firms in the analysis.

We follow the empirical literature on firm survival and identify a firm exit as a firm closure. Then, the time of exit is found by identifying the moment in which firms cease to report IES information. We require that a firm is absent from the survey for at least two years in order to identify an exit because temporary non-reporting may occur for a number of reason other than cessation of activity. This means that a firm exits at time t if it is absent from IES at time $t+1$ and $t+2$. If a firm does not report information temporarily, meaning that the firm is in the survey at time $t-1$ and $t+1$ but not at time t, we considered the firm as active and inputed data as the average of variables between $t-1$ and $t+1$. This means that we use data only until 2012 in the analysis of firm survival because data for 2013 and 2014 are considered to determine a firm exit. In turn, the founding year of the firm is available in the dataset and used to identify new firms.

Dynamics of firm entry and firm survival

In this section we describe the main facts related to firm natality and firm survival for Portuguese new firms over the period between 2005 to 2012. We begin the analysis by considering the aggregate evolution of new firms and proceed by distinguishing the distribution of new firms by sector of economic activity and size class.

2. The sampling method consists of non-financial corporations covering all sectors of activity defined in the Portuguese Classification of Economic Activities with the following exceptions: financial intermediation, general government, private households with employed persons, international organizations, and other non-resident institutions.

Aggregate market dynamics

Table 1 reports the number of new firms by sampling year and survival rates by age cohort of new firms. The estimates suggest that survival rates calculated without accounting for firm heterogeneity at the sector of economic activity or size class level, seem to be independent of the age cohort. At a first glance, entry rates seem lower after 2009, which may suggest the presence of an economic crisis effect in firm creation. Nevertheless, the relationship between entry rates and the business cycle is analyzed in more detail in Section 4.

Cohort	Start-ups	Entry	Survival rates by life duration of the firm (in percentage)							
		rate	1	2	3	4	5	6	7	8
2005	12,514	3.42	99	92	82	73	65	59	53	48
2006	14,227	3.81	94	85	74	65	58	52	46	
2007	15,100	3.92	93	82	71	63	55	48		
2008	14,642	3.77	94	83	72	62	55			
2009	9,721	3.00	93	83	72	63				
2010	8,883	3.24	95	86	76					
2011	10,143	3.72	95	85						
2012	8,205	3.16	95							
2013	8,476	3.25								

TABLE 1. New firms and survival rates by cohort.

The sampling period goes from 2005 to 2013. Firm exits are identified only until 2012. For more data details see Section 2.

Figure 1 depicts the survival rates of new firms obtained through the estimation of a Kaplan-Meier survival function. It follows from the nature of the dataset that the amount of information available to estimate survival rates is different in each sampling year because firms are observed over a different number of years. The maximum age attained by a firm born in 2005 is equal to eight years and for a firm born in 2011 is equal to two years. Nevertheless, these results suggest that new firms fail at a significant rate, with approximately 50 percent of new firms exiting operation before their seventh year of life. The results also suggest that a considerable fraction of new firms fails in their first year of life and that about only 48 percent of new firms survive for eight years. These results tally with the high mortality of new firms documented in the industrial organization empirical literature.

Market dynamics by sector of economic activity

In this section we analyze new firm entry and survival by sector of economic activity, with economic activity sectors defined at the 2-digit NACE. Table 2 reports the number of new entrants, entry rates, and the share of sales of

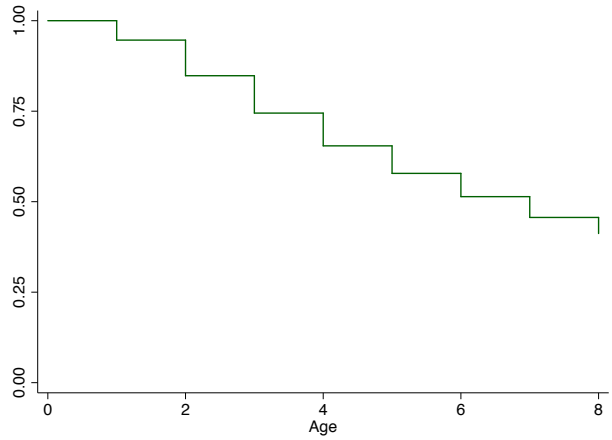


FIGURE 1: Kaplan-Meier survival function.

new firms on total activity sector sales, calculated for 2005[3]. The entry rate is defined at the sector of economic activity level and is calculated as the number of new firms divided by the total number of firms (entrants plus incumbents) in a given year. The share on sales equals sales by entrants divided by total sales in a given sector of economic activity. The survival rate is defined as the percentage of new firms surviving in a given sector of economic activity up until 2012 as of the total number of new firms created in 2005.

The results in Table 2 indicate that 12,514 new firms were created in 2005. The sector of economic activity with a higher entry activity is trade with 3,969 new firms and the one with less entry activity is agriculture with 210 new firms. The entry rate ranges from 3.15 percent in the transportation and storage sectors to 8.19 percent in the information and communication sectors, which are high entry rates[4]. These figures suggest that entry is common in most sectors of economic activity. Nevertheless, the share of sales by entrants on total sales in a given sector of economic activity is fairly small, ranging from 0.20 percent in the information and communication sectors to 1.61 percent in the accommodation and food service activities sector. Geroski (1995) suggests that this difference between entry rates and entry penetration is due to the much smaller scale of entrants than their incumbent counterparts. An estimate of the average size of entrants relative to that of all firms can be obtained by

3. It follows from the nature of the data that survival rates are calculated using a different amount of information for different cohorts, being the 2005 cohort the one that conveys more information. Furthermore, the results in the previous section suggest that survival rates are independent of the age cohort.

4. Audretsch (1995) documents that entry rates for Netherlands range from 2.53 percent to 4.72 percent across manufacturing sectors.

dividing the entrants' share on sales by entry rates. According to the results reported in Table 2, entrants are estimated to be on average approximately 15 percent of the average size of Portuguese firms in 2005.

	2005			2012	
	New firms	Entry rate (%)	Share on sales (%)	Surviving firms	Survival rate (%)
Agriculture	210	4.13	1.29	147	70.50
Manufacturing	1,217	3.60	0.31	640	48.05
Construction	1,750	5.63	0.57	831	41.12
Trade	3,969	5.51	0.66	2,044	46.88
Transporting	476	3.15	0.66	269	50.86
Accommodation	1,184	5.50	1.61	559	39.55
Information	321	8.19	0.20	170	49.41
Real estate	290	4.07	0.44	148	45.18
Others	3,097	6.98	0.53	1,778	52.87
Total	12,514	5.63	0.66	6,586	47.51

TABLE 2. New firm entry and survival rates by sector of activity.

Agriculture stands for agriculture, forestry, and fishing, *Trade* for wholesale and retail trade, *Transporting* for transporting and storage, *Accommodation* for accommodation and food service activities, *Information* for information and communication, and *Others* includes all other sectors.

The survival rates of newly created firms in 2005 and that are still operating in 2012 range from approximately 39.55 percent in the accommodation and food service activities sector to 70.5 percent in the agriculture sector. Moreover, the survival rate calculated for all firms born in 2005 is approximately 48 percent.

These findings are consistent with the stylized facts identified by Geroski (1995) regarding the start-up size and survival rates of new firms. Furthermore, the fact that new firms are in general small and that their lives are typically short suggest that new firms play a modest role in shaping industry structure and industry performance[5].

One interesting result is that the coefficient of variation estimated for the entry rate equals 0.32 and for the survival rate is approximately 0.18, which suggests that survival rates show considerably lower variability than entry rates. This result is apparently inconsistent with the industrial organization literature on market dynamics, which posits that entry rates show considerably lower variance than survival rates. Furthermore, we find that entry rates show greater cross-sector variation than within sector variation, which is also not in line with previous research (see Geroski (1995) and Audretsch (1995), for example). Nevertheless, this inconsistency may arise from the fact that in this study we consider the full economy and

5. Mata and Portugal (1995) document that the competitive disciplining role played by new firms on incumbent firms is rather modest.

distinguish across sectors of economic activity, while most of the industrial organization literature on market dynamics considers only industries within the manufacturing sector.

Market dynamics by size class

In this section we consider the entry rate and the entrants' employment share by firm size to exploit the extent of new firm creation in the Portuguese economy. The entry rate is defined in the previous section and the entrants' employment share is obtained by dividing the employment in new firms by the total employment. Mata (1996) calls this measure the entry share. Both measures are computed by size class.

Tables 3 and 4 report entry rates and entry shares by size class and time-averaged over the sampling period, with size classes defined using the number of employees. According to the estimates reported in Table 3, the most striking result is that new firms are in general quite small, with approximately 95 percent of new firms employing less than ten workers. The fraction of new firms employing more than 50 workers at birth is very small. The results regarding the employment share of new firms show that firms with less than ten workers are responsible for the creation of 65 percent of the total jobs created by new firms, and only 4.3 percent of job creation is attributed to new firms employing more than 100 workers. The results reported in Table 4 corroborate the previous findings and suggest that entry rates and entry shares are higher in the size classes with few employees.

	Total	<5	5-9	10-49	50-99	>100
Entry rate and relative contributions (%)	5.013	85.07	10.34	4.30	0.19	0.07
Entry share and relative contributions (%)	1.51	41.16	23.66	26.30	4.57	4.31

TABLE 3. New firm entry by size class.

Entry rates and entry shares are time-averaged over the period between 2005 and 2013, and calculated as a proportion of the total number of firms. Figures reported in size classes correspond to the relative contribution of each size class to the *total* entry rate and *total* entry share, respectively.

	<5	5-9	10-49	50-99	>100
Entry rate (%)	6.56	2.82	1.56	0.71	0.37
Entry share (%)	5.09	2.74	1.39	0.69	0.18

TABLE 4. New firm entry by size class.

Entry rates and entry shares and time-averaged over the period between 2005 and 2013, and calculated as a proportion of the number of firms in a given size class.

Figure 2 shows survival rates of new firms by distinguishing between firms with less than 10 employees at birth and firms with at least 10 employees at birth. Survival rates of firms with at least 10 employees are considerably higher than of their counterparts after the third year of life. Moreover, the difference in survival rates of the two groups seems to widen with the age of firms. This result is in line with the industrial organizational literature on firm survival that states that large firms experience higher survival probabilities than their smaller counterparts.

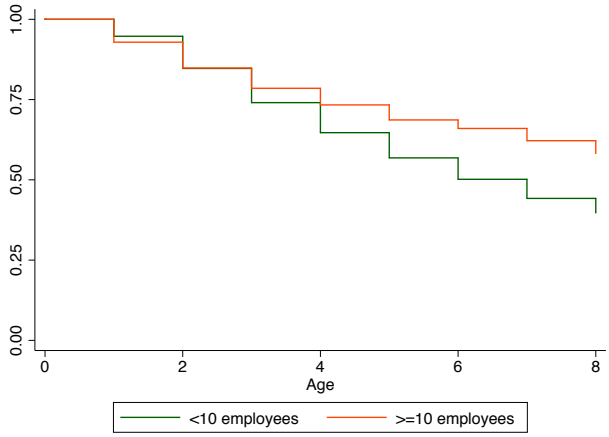


FIGURE 2: Kaplan-Meier survival function by number of employees at birth.

New firm dynamics over the business cycle

A strand of the industrial organization literature on market dynamics pinpoints that periods of high firm creation follow periods of relatively depressed conditions because unemployed individuals are more likely to create new firms than employed ones (see Highfield and Smiley (1987) and Evans and Leighton (1989)). An alternative strand of this literature posits that firm entry is pro-cyclical, meaning that firm creation is particularly important during good times because profit opportunities are greater and, therefore, new firms are more likely to survive. In this section we study the behavior of firm creation according to aggregate macroeconomic conditions.

Figures 3 and 4 depict entry rates and entry shares by firm size and real GDP growth rates, respectively. We follow Mata (1996) and consider real GDP growth lagged by one period because firm creation at time t is expected to respond to GDP growth registered in the previous year $t-1$. At first glance, entry rates and employment shares seem to respond to macroeconomic conditions in a pro-cyclical fashion in the case of the smaller firms. In turn,

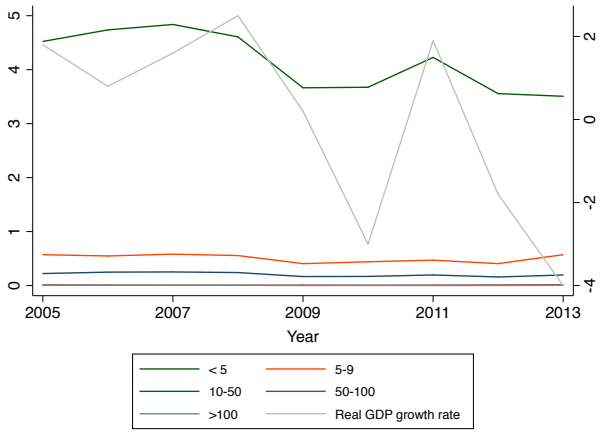


FIGURE 3: Entry rates by size class and real GDP growth rate.

Entry rates by size class and real GDP growth rate (rhs scale), in percentage. Real GDP growth rates are lagged by one year. Source for real GDP growth rate: Eurostat.

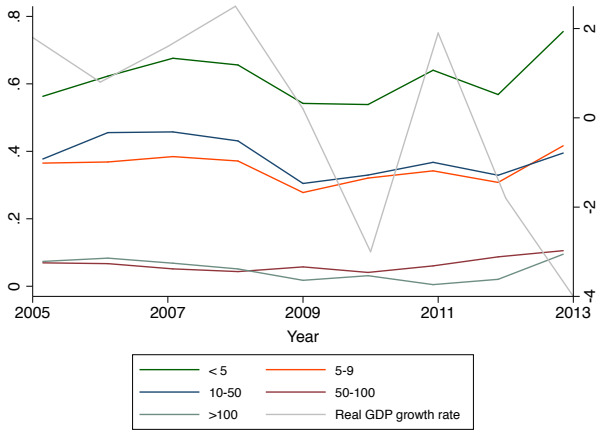


FIGURE 4: Entry shares by size class and real GDP growth rate.

Entry shares by size class and real GDP growth rate (rhs scale), in percentage. Real GDP growth rates are lagged by one year. Source for real GDP growth rate: Eurostat.

no pattern can be found in the case of larger firms. These results are in line with the findings of Mata (1996), who show that small firms are created mostly when aggregate conditions are more favorable. Empirical evidence for the U.S. also shows that firm entry is less common in recessions and that in general new firms are smaller in bad times (see Moreira (2015)).

The empirical literature on new firm survival documents that macroeconomic conditions do matter for firm survival (see Geroski *et al.* (2010), Boeri and Bellmann (1995), and Ilmakunnas and Topi (1999)). Figure 5 shows the behavior of exit rates over the business cycle. In general, it seems that exit rates have increased steadily over the sampling period, and that periods of economic recovery were not followed by lower exit rates[6].

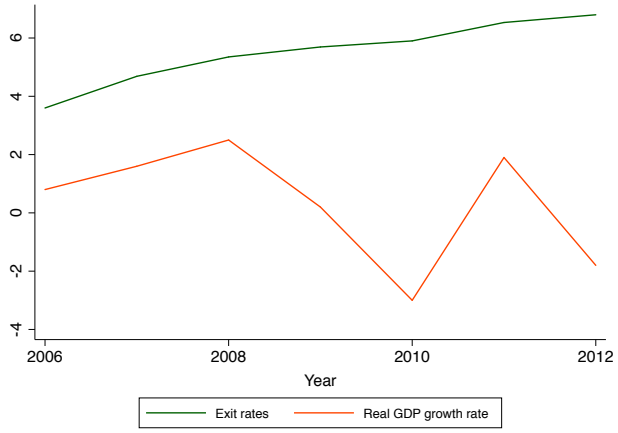


FIGURE 5: Exit rates and real GDP growth rate.

Exit rates and real GDP growth rate, in percentage. Real GDP growth rates are lagged by one year. Source for real GDP growth rate: Eurostat.

Conclusions

Newly created firms are an important driver of innovation and job creation. However, new firms fail at a significant rate. In this study we use a comprehensive dataset that allows us to identify firms at birth and follow them over their lives. We analyze firm creation and firm survival, exploring heterogeneity at the sector of economic activity level and size class. Furthermore, we study the behavior of firm creation over the business cycle.

The results suggest that entry rates are fairly high but represent a small share of the total sales in a given sector of economic activity, meaning that new firms are in general much smaller than their incumbent counterparts. This result is corroborated by the low employment shares of new firms.

6. The analysis of exit rates over the business cycle starts only in 2006 because IES information started to be collected in 2006. Even though the report with information regarding 2005 was also taken into consideration in the remaining analysis, exit rates in 2005 are most likely biased because data for 2005 was collected in 2006.

These findings suggest that new firms play a limited role in shaping industry structure and industry performance.

The results also indicate that a considerable fraction of new firms fail in their initial years of life and about only 48 percent of new firms survive throughout the sampling period. These high firm mortality rates are independent of the age cohort. Furthermore, we document that entry rates and employment shares for smaller firms seem to be pro-cyclical.

It is important to highlight that the aim of this study is a descriptive one and no causal effects should be attempted based on this analysis.

References

Adelino, Manuel, Song Ma, and David T Robinson (2016). "Firm age, investment opportunities, and job creation." *Journal of Finance*, forthcoming.

Audretsch, David B (1995). "Innovation, growth and survival." *International Journal of Industrial Organization*, 13(4), 441–457.

Boeri, Tito and Lutz Bellmann (1995). "Post-entry behaviour and the cycle: Evidence from Germany." *International Journal of Industrial Organization*, 13(4), 483–500.

Evans, David S and Linda S Leighton (1989). "Some empirical aspects of entrepreneurship." *American Economic Review*, 79(3), 519–535.

Geroski, Paul A (1995). "What do we know about entry?" *International Journal of Industrial Organization*, 13(4), 421–440.

Geroski, Paul A, José Mata, and Pedro Portugal (2010). "Founding conditions and the survival of new firms." *Strategic Management Journal*, 31(5), 510–529.

Haltiwanger, John, Ron S Jarmin, and Javier Miranda (2013). "Who creates jobs? Small versus large versus young." *Review of Economics and Statistics*, 95(2), 347–361.

Hannan, Michael T and John H Freeman (1989). "Organizational Ecology." *Cambridge, MA: Harvard University*.

Highfield, Richard and Robert Smiley (1987). "New business starts and economic activity: An empirical investigation." *International Journal of Industrial Organization*, 5(1), 51–66.

Ilmakunnas, Pekka and Jukka Topi (1999). "Microeconomic and macroeconomic influences on entry and exit of firms." *Review of Industrial Organization*, 15(3), 283–301.

Mata, José (1993). "Entry and type of entrant: Evidence from Portugal." *International Journal of Industrial Organization*, 11(1), 101–122.

Mata, José (1996). "Small firm births and macroeconomic fluctuations." *Review of Industrial Organization*, 11(2), 173–182.

Mata, José and Pedro Portugal (1994). "Life duration of new firms." *Journal of Industrial Economics*, 42(3), 227–245.

Mata, José and Pedro Portugal (1995). "The survival of new plants: Start-up conditions and post-entry evolution." *International Journal of Industrial Organization*, 13, 459–481.

Mata, José, Pedro Portugal, and Paulo Guimarães (1995). "The survival of new plants: Start-up conditions and post-entry evolution." *International Journal of Industrial Organization*, 13(4), 459–481.

Moreira, Sara (2015). "Firm Dynamics, Persistent Effects of Entry Conditions, and Business Cycles." Mimeo, University of Chicago.

Dating the Portuguese business cycle

António Rua
Banco de Portugal

Abstract
The aim of this article is to establish a reference business cycle chronology for Portugal over the last four decades. Drawing on a non-parametric approach embedding the NBER's business cycle dating procedure, a monthly business cycle chronology is provided and its robustness is assessed resorting to a large data set. (JEL: C23, C55, E32)

Introduction

The study of business cycles has been a key research area for a long time in economics. In this respect, one should mention that there are two types of business cycles in the economic literature: the classical business cycle and the growth cycle. Classical cycles refer to alternating periods of contraction and expansion whereas growth cycles refer to alternating periods of acceleration and slowdown of economic activity. In general, before the economy gets into a recession, there is a deceleration of activity and it usually accelerates before attaining an expansionary phase. Moreover, there might be decelerations that do not translate into recessions or accelerations that do not correspond to expansionary phases. Hence, the timing of the turning points does not necessarily coincide between the two kinds of cycles.

While the former concept relies on the level of economic activity, the latter draws on deviations from a long-run trend. From a practical point of view, the first is more tractable as the second one entails a decomposition in trend and cycle which are unobservable components. Thus, the analysis of growth cycles is conditional on the method chosen to de-trend macroeconomic time series.

Typically, the business cycle chronologies refer to the dating of classical cycles. The most notable and well-known case of dating peaks and troughs in economic activity is the National Bureau of Economic Research (NBER) for the United States.[1] At the time NBER was established in 1920 by Wesley Mitchell and colleagues, the study of business cycles was settled as one of the primary objectives. In this respect, one should mention the pioneer work

E-mail: antonio.rua@bportugal.pt

1. For further details visit the NBER site at http://www.nber.org/cycles/main.html

"Measuring Business Cycles" written jointly with Arthur Burns in 1946 where a now widely accepted definition of business cycles was provided (p. 3):

> Business cycles are a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises: a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle; in duration, business cycles vary from more than one year to ten or twelve years; they are not divisible into shorter cycles of similar characteristics with amplitudes approximating their own.

The NBER started publishing its first business cycle dates in 1929 and since 1978 a business cycle dating committee chaired by Robert Hall determines the business cycle turning points. More recently, following NBER's approach, the Economic Cycle Research Institute (ECRI) and the Centre for Economic Policy Research (CEPR) also determine business cycle turning points for twenty other countries[2] and the euro area, respectively.[3]

Naturally, the establishment of a business cycle chronology is certainly a complex task. The NBER business cycle dating committee mentions that:

> The chronology comprises alternating dates of peaks and troughs in economic activity. A recession is a period between a peak and a trough, and an expansion is a period between a trough and a peak. During a recession, a significant decline in economic activity spreads across the economy and can last from a few months to more than a year. Similarly, during an expansion, economic activity rises substantially, spreads across the economy, and usually lasts for several years.

As it is clear from above, the definitions of recession and expansion used by the NBER are vague and involve judgment. In 1971, Gerhard Bry and Charlotte Boschan introduced a non-parametric algorithm at the NBER that comes closest to translating the NBER's definition into practice. Basically, the algorithm relies on a set of filters and rules to spot local minima and maxima in the level (or log-level) of the series. A local minimum is a trough and the following local maximum a peak so that the period between trough and peak is an expansion, and from peak to trough a recession. The algorithm assumes that a full business cycle (peak to peak or trough to trough) should last at least fifteen months, each business cycle phase (peak to trough or trough to peak) should last at least five months and peaks and troughs should alternate.

As stressed, for example, by Watson (1994), the algorithm developed by Bry and Boschan (1971) is able to replicate quite well the turning points selected by experts. The Bry-Boschan algorithm has been applied by King

2. Canada, Mexico, Brazil, Germany, France, United Kingdom, Italy, Spain, Switzerland, Sweden, Austria, Russia, Japan, China, India, Korea, Australia, Taiwan, New Zealand and South Africa.

3. For further details visit the ECRI site at https://www.businesscycle.com/ecri-business-cycles/international-business-cycle-dates-chronologies and CEPR site at http://cepr.org/content/euro-area-business-cycle-dating-committee

and Plosser (1994), Watson (1994), Artis and Osborn (1997), Mönch and Uhlig (2005), Stock and Watson (2010, 2014), among many others.

The aim of this article is to establish a reference business cycle chronology for Portugal over the last forty years. We start by applying the Bry-Boschan algorithm to quarterly real GDP. For comparison purposes, we also report the resulting business cycle chronology using the popular rule-of-thumb of at least two consecutive quarters of decline in economic activity to define a recession. Bearing in mind the caveats of relying solely on a single series while aiming at establishing a monthly chronology, we then resort to the coincident indicator for the Portuguese economy (see Rua (2004)). The monthly coincident indicator is a composite indicator representative of a wide spectra of economic activity which is regularly released by Banco de Portugal. Drawing on the coincident indicator, a monthly reference business cycle chronology is established. Finally, we resort to a large monthly dataset for the Portuguese economy to assess the robustness of the previously obtained monthly chronology following Stock and Watson (2010, 2014).

Dating with quarterly GDP

As it is widely recognized, GDP is a natural proxy for measuring aggregate economic activity. Actually, Burns and Mitchell (1946, p. 72) state that:

> Aggregate [economic] activity can be given a definite meaning and made conceptually measurable by identifying it with gross national product

Hence, we firstly rely on real GDP to obtain a quarterly business cycle chronology. In particular, we apply the Bry-Boschan algorithm (BB hereafter) to the log-level of real GDP. However, since GDP is available only on a quarterly basis, the original BB algorithm cannot be applied as it was developed for monthly series. Hence, we resort to the modified BB algorithm proposed by Harding and Pagan (2002) which shares the same features of the original BB algorithm but adapted to the quarterly frequency (the so-called BBQ).

In the case of Portugal, due to data availability constraints, the time period under analysis ranges from the beginning of 1977 up to the end of 2015. In particular, as quarterly real GDP is currently released by INE only for the period since the first quarter of 1995, we use the historical series regularly updated and disclosed by Banco de Portugal which start in the first quarter of 1977 on a seasonally adjusted basis. One should mention that such series coincide with the quarterly GDP series released by INE since 1995.

The resulting quarterly business cycle chronology is presented in Table 1 and the log-level of Portuguese quarterly real GDP is displayed in Figure 1 along the recessionary periods denoted by the shaded areas.

Business cycle dates		Duration (in quarters)			
Peak	**Trough**	**Contraction**	**Expansion**	**Cycle**	
		Peak to trough	Previous trough to this peak	Trough from Previous Trough	Peak from Previous Peak
1980 Q2	1980 Q4	2	-	-	-
1983 Q1	1984 Q1	4	9	13	11
1992 Q2	1993 Q2	4	33	37	37
2002 Q1	2003 Q2	5	35	40	39
2008 Q1	2009 Q1	4	19	23	24
2010 Q3	2012 Q4	9	6	15	10
Average		*5*	*20*	*26*	*24*

TABLE 1. Business cycle chronology based on the quarterly real GDP

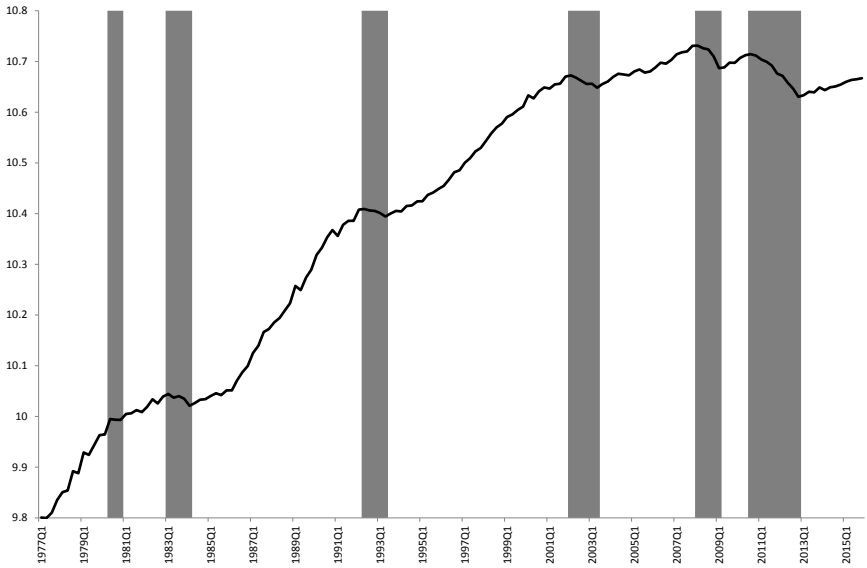


FIGURE 1: Log-level of Portuguese quarterly real GDP

The BBQ algorithm determines six peaks and six troughs since 1977. In particular, six recessionary periods are identified with the shortest recession lasting only two quarters in 1980 whereas the latest recession has been the longest one spanning through nine quarters. A stylized business cycle feature, also present in the Portuguese case, is the asymmetry between expansions and contractions. The average duration of recessions is five quarters whereas the average length of expansions is 20 quarters. This corresponds to an average duration of the Portuguese business cycle of 25 quarters.

An informal and commonly used rule-of-thumb for dating the business cycle, identifies recessions as periods recording at least two consecutive quarters of negative real GDP growth, as suggested by Julius Shiskin in an 1979 New York Times article. In Table 2, we report the peaks and troughs one would obtain by considering the above-mentioned rule-of-thumb. One can see that in most cases the peaks and troughs coincide. However this is not always the case. In particular, the peak in 1983 is dated differently, the trough in the early 2000's recession is not the same and the rule-of-thumb identifies one extra recessionary period than the BBQ algorithm in 2004.

Turning point	BBQ	Rule-of-thumb
P	1980 Q2	1980 Q2
T	1980 Q4	1980 Q4
P	1983 Q1	1983 Q3
T	1984 Q1	1984 Q1
P	1992 Q2	1992 Q2
T	1993 Q2	1993 Q2
P	2002 Q1	2002 Q1
T	2003 Q2	2002 Q4
P	-	2004 Q2
T	-	2004 Q4
P	2008 Q1	2008 Q1
T	2009 Q1	2009 Q1
P	2010 Q3	2010 Q3
T	2012 Q4	2012 Q4

TABLE 2. Turning points with BBQ algorithm and rule-of-thumb

Note: P and T denote Peak and Trough, respectively.

In this respect, the NBER business cycle dating committee states that most of the recessions in the United States do consist of two or more quarters of declining real GDP, but not all of them, and that the committee's procedure for identifying turning points differs from the two-quarter rule in a number of ways. For instance, the depth of the decline in economic activity is considered. In the Portuguese case, the recessionary period identified by the rule-of-thumb in 2004 corresponds to an accumulated decrease over that period of only 0.3 per cent which was not classified by the BBQ algorithm as a recession. Hence, notwithstanding the similitude of dates reported in Table 2, one should

be cautious when relying solely on the rule-of-thumb of two consecutive quarters of negative real GDP growth.

Dating with the monthly coincident indicator

Although real GDP can be a useful proxy for measuring aggregate economic activity, it suffers from several shortcomings. For instance, GDP is available only at a quarterly frequency and may suffer from measurement errors. In this respect, Burns and Mitchell (1946, p. 73) qualified the previously quoted sentence by referring that:

> Unfortunately, no satisfactory series of any of these types is available by months or quarters for periods approximating those we seek to cover.

Moreover, the U.S. Department of Commerce (1984, p. 65) acknowledges that:

> Aggregate economic activity cannot be defined precisely, and no single time series measures it adequately; however, a variety of statistical series measure some of its major aspects.

In fact, one the features of the NBER committee's procedure is not to rely solely on real GDP but using a range of other indicators as well. Furthermore, a considerable emphasis is put on monthly indicators in order to attain a monthly chronology.

A possible approach is to consider a monthly composite coincident index for the whole economy. In this respect, one should mention the pioneering work of Stock and Watson (1989, 1991, 1993) who considered a factor model in order to extract a common factor summarizing the co-movements from a small number of indicators. Such a composite indicator is aimed at capturing the overall state of the economy and can be used to date the business cycle. This is the so-called average-then-date approach for dating the business cycle.

In the Portuguese case, a monthly coincident indicator for economic activity is disclosed by Banco de Portugal every month since June 2004. This composite indicator has been proposed by Rua (2004) drawing on the methodology of Azevedo *et al.* (2006).[4] In particular, from a starting set of hundreds of series a subset of variables have been chosen according to several criteria namely availability on a monthly frequency, timeliness, a reasonable time span, a noteworthy co-movement with the economic cycle and with the aim of obtaining a broadly based activity measure. Hence, besides real quarterly GDP, the set of information of the coincident indicator includes retail sales volume which intends to capture private consumption developments. Regarding investment, it considers the sales of heavy commercial vehicles

4. See Rua (2015) for a historical assessment of the performance of the coincident indicator in tracking Portuguese economic developments.

reflecting Gross Fixed Capital Formation (GFCF) in transportation equipment as well as cement sales which portray GFCF in the construction sector. From the supply side, the manufacturing production index captures the industrial sector behavior which is typically a strongly cyclical sector. In order to take on board the evolution of income and wealth, it includes the households' assessment of their current financial situation. Concerning the labor market, new job vacancies are considered. Finally, to reflect external environment, it includes a weighted average of the current economic situation assessment of the Portuguese main trade partners, where the weights are each country's share in Portuguese exports. The trend-cycle underlying the coincident indicator is available on a monthly frequency since January 1977.[5]

As the monthly coincident indicator for the Portuguese economic activity is a composite indicator that merges information both from real GDP as well as from other relevant economic variables and being available at a monthly frequency, it seems particularly suitable for dating the business cycle.

Hence, we now determine the peaks and troughs of the coincident indicator, in its trend-cycle format, through the use of the monthly BB algorithm. The monthly business cycle chronology in presented in Table 3 and the log-level of the trend-cycle underlying the coincident indicator along with the identified recessions is displayed in Figure 2.

Business cycle dates		Duration (in months)			
Peak	**Trough**	**Contraction**	**Expansion**	**Cycle**	
		Peak to trough	Previous trough to this peak	Trough from Previous Trough	Peak from Previous Peak
March 1983 (Q1)	February 1984 (Q1)	11	-	-	-
July 1992 (Q3)	June 1993 (Q2)	11	101	112	112
April 2002 (Q2)	February 2003 (Q1)	10	106	116	117
November 2007 (Q4)	April 2009 (Q2)	18	56	74	66
September 2010 (Q3)	April 2013 (Q2)	31	17	48	35
Average		*16*	*70*	*88*	*83*

TABLE 3. Business cycle chronology based on the monthly coincident indicator

Note: The corresponding quarterly dates are presented in parentheses.

5. Note that, although the coincident indicator is released as the year-on-year change of the estimated trend-cycle, herein we naturally consider the log-level of the trend-cycle for the purpose of dating classical business cycles.

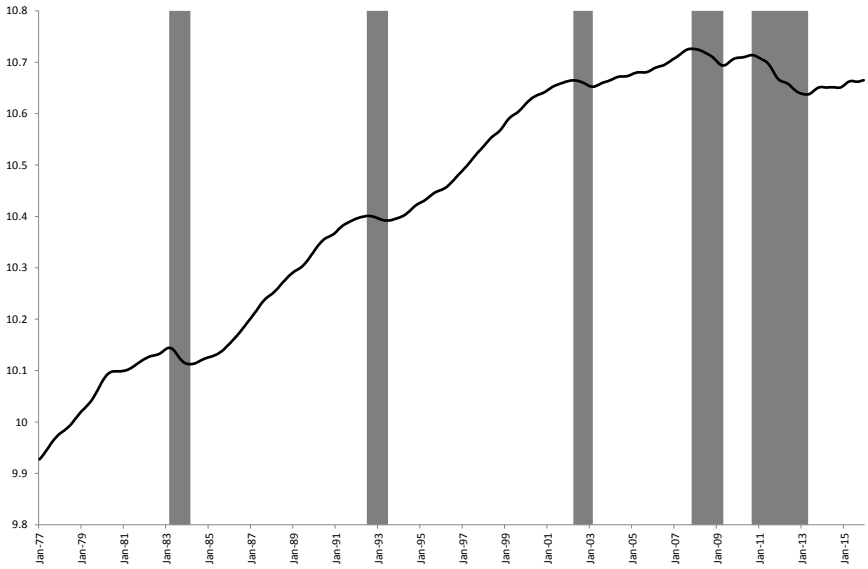


FIGURE 2: The trend-cycle underlying the monthly coincident indicator

According to this resulting monthly business cycle chronology, five recessions occurred over the last four decades. The first three lasted around 11 months, the one recorded at the time of the so-called US Great Recession lasted 18 months whereas the most recent one was by far the longest one (31 months). The business cycle asymmetry, in terms of duration, is once again present with the average recession lasting 16 months in contrast with the average length of expansions of 70 months. This corresponds roughly to an average duration of the business cycle of 86 months. However, one should bear in mind the noteworthy heterogeneity across business cycles.

When comparing with the previously discussed quarterly dating, one should note that in the monthly business cycle chronology only five recessions show up in contrast with the six recessionary periods identified with quarterly real GDP. The difference lies in the period from the second quarter of 1980 up to the fourth quarter of 1980 which is identified as a recession when drawing solely on real GDP but not when resorting to the monthly coincident indicator. In fact, the decrease in real GDP over such a short-lived period is quite marginal (-0.2 per cent). Regarding the other recessions, the monthly dating is relatively close to the quarterly one, in the sense that the month of the peak or trough typically falls within or in the adjacent quarter previously

identified with real GDP. A notable exception concerns the end of the last recession, which is determined to be in April 2013 instead of the last quarter of 2012.

Dating with a large dataset

We now turn to the use of a wide range of time series to date the business cycle. The underlying idea is to first determine the turning points individually in a large number of series and then obtain a common set of turning points. This corresponds to the date-then-average approach.

Following Stock and Watson (2010, 2014), let us consider a panel of N time series. For each time series one can determine a specific chronology, resorting to the BB algorithm, and denote τ_{is} as the turning point for series i in episode s, $i = 1, \ldots, N$, $s = 1, \ldots, S$.[6] Once the individual chronology has been established for all the N series, one can consider the following model

$$\tau_{is} = D_s + k_i + \eta_{is}$$

where D_s is the reference cycle turning point date in episode s, k_i is the mean lag of series i relative to the reference cycle, and η_{is} is a discrepancy. This model can be estimated by fixed effects panel data regression with an unbalanced panel and missing observations. In particular, one obtains estimates for the reference cycle turning point dates, D_s, as well as the corresponding standard errors.

For the empirical application of this method, we resort to the large monthly data set considered by Dias *et al.* (2015, 2016) for the Portuguese economy which comprises 126 series. It includes both hard and soft data covering business and consumers surveys (43 series), retail sales (4 series), industrial production (7 series), turnover in industries and services (20 series), employment, hours worked and wage indices in industries and services (24 series), tourism nights spent in Portugal (3 series), car sales (3 series), cement sales, vacancies and registered unemployment (5 series), energy consumption (3 series), nominal exports and imports of goods (10 series), real effective exchange rate, Portuguese stock market index and ATM/POS series.

However, Dias *et al.* (2015, 2016) considered data only since 1995. Since the aim is to cover the last forty years, longer series have been collected for most of the variables. In particular, for business and consumer surveys, which account for a large fraction of the data set, the series go back to the second half of the 80's (only the services survey starts in the late 90's). For several variables it was

6. An episode denotes a non-overlapping period which contains a single turning point of unknown date. Note that if series i has no turning point or is unavailable in episode s then τ_{is} is treated as missing data.

possible to collect data since 1977 namely industrial production, vacancies, tourism nights spent in Portugal and gasoline consumption. Other series such as car sales and cement sales begin in the early 80's whereas turnover in industries and external trade data start around 1990. Most of the remaining series are available from mid-90's onwards.[7] All series are seasonally adjusted and considered in levels (or log-levels).

Turning point		Deviation in months	
P	March 1983	2	(1.2)
T	February 1984	3	(1.4)
P	July 1992	0	(1.0)
T	June 1993	2	(1.0)
P	April 2002	1	(1.0)
T	February 2003	4	(0.9)
P	November 2007	3	(0.9)
T	April 2009	2	(0.9)
P	September 2010	3	(1.0)
T	April 2013	0	(0.9)

TABLE 4. Turning points based on a large data set (in deviations from the monthly chronology)

Note: The figures are rounded and the corresponding standard errors are presented in parentheses.

In table 4, we report the estimated turning points using the above panel data model as deviations from the monthly business cycle reference chronology discussed earlier along with the corresponding standard errors.[8] It is reassuring that the dates largely line up with the previously presented monthly business cycle reference chronology. In fact, the two chronologies are within a few months of each other and, in most cases, are not statistically different at a standard significance level. Even regarding the trough of the last recession identified as April 2013, which as mentioned earlier differs more markedly from the dating drawing solely on real GDP, the estimated turning point date coincides when a large data set is considered for the Portuguese economy. Hence, the above results reinforce the robustness of the business cycle chronology provided in Table 3.

7. In the Appendix, we report the full list of series along with the corresponding starting date for each series.

8. In the estimation of the panel data model, we considered as an episode the period centered at the month of the turning point identified in the previous section with a window size of 12 months.

Conclusions

In this article, it is proposed a business cycle reference chronology for the Portuguese economy. In particular, one draws on a non-parametric approach that mimics the expert system developed at the NBER for the identification of business cycle turning points. Firstly, one assesses the business cycle chronology based solely on quarterly real GDP. Then, embedding the NBER dating committee spirit, one acknowledges the caveats that may derive from relying on a single series and aim at providing a monthly business cycle reference chronology. In particular, it is considered the monthly coincident indicator for the Portuguese economy which has been regularly released by Banco de Portugal. A monthly business cycle reference chronology is established and its soundness is assessed resorting to a large monthly data set for Portugal.

References

Artis, Kontolemis Z.G., M.J. and D.R. Osborn (1997). "Business cycles for G7 and European countries." *Journal of Business*, 70(2), 249–279.

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.

Bry, G and C. Boschan (1971). *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. National Bureau of Economic Research, New York.

Burns, A. and W. Mitchell (1946). *Measuring business cycles*. National Bureau of Economic Research, New York.

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

Dias, F., M. Pinheiro, and A. Rua (2016). "A bottom-up approach for forecasting GDP in a data rich environment." Banco de Portugal Economic Studies, 2(3), 1-19, Banco de Portugal.

Harding, D. and A. Pagan (2002). "Dissecting the cycle: A methodological investigation." *Journal of Monetary Economics*, 49, 365–381.

King, R. and C. Plosser (1994). "Real business cycles and the test of the Adelmans." *Journal of Monetary Economics*, 33(2), 405–438.

Mönch, E. and H. Uhlig (2005). "Towards a Monthly Business Cycle Chronology for the Euro Area." *Journal of Business Cycle Measurement and Analysis*, 2005(1), 43–69.

Rua, A. (2004). "Um novo indicador coincidente para a economia portuguesa." Boletim Económico, Junho, 21-29, Banco de Portugal.

Rua, A. (2015). "Revisiting the monthly coincident indicators of Banco de Portugal." Banco de Portugal Economic Studies, 1(1), 49-64, Banco de Portugal.

Stock, J. and M. Watson (1989). "New Indexes of Coincident and Leading Economic Indicators." in O. Blanchard and S. Fischer (eds.), NBER Macroeconomics Annual, Cambridge, Mass.: MIT Press.

Stock, J. and M. Watson (1991). "A Probability Model of the Coincident Economic Indicators." in K. Lahiri and G.H. Moore (eds.), Leading Economic Indicators: New Approaches and Forecasting Records, Cambridge: Cambridge University Press.

Stock, J. and M. Watson (1993). "A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues and Recent Experience." in J.H. Stock and M.W. Watson (eds.), Business Cycles, Indicators and Forecasting, Chicago: University of Chicago Press for NBER.

Stock, J. and M. Watson (2010). "Indicators for dating business cycles: cross-history selection and comparisons." *American Economic Review: Papers and Proceedings*, pp. 16–19.

Stock, J. and M. Watson (2014). "Estimating turning points using large data sets." *Journal of Econometrics*, 178, 368–381.

U.S. Department of Commerce (1984). *Handbook of Cyclical Indicators: A Supplement to the Business Conditions Digest*. Bureau of Economic Analysis, Washington D. C.

Watson, M. (1994). "Business-cycle durations and postwar stabilization of the U.S. economy." *American Economic Review*, 84(1), 24–46.

Appendix

Series	Starting date
Economic Sentiment Indicator	January 1987
Consumer Confidence Indicator	June 1986
Financial situation over last 12 months	June 1986
Financial situation over next 12 months	June 1986
General economic situation over last 12 months	June 1986
General economic situation over next 12 months	June 1986
Major purchases at present	June 1986
Major purchases over next 12 months	June 1986
Unemployment expectations over next 12 months	June 1986
Savings at present	June 1986
Savings over next 12 months	June 1986
Price trends over last 12 months	June 1986
Price trends over next 12 months	June 1986
Statement on financial situation of household	June 1986
Construction Confidence Indicator	January 1989
Building activity development over the past 3 months	January 1989
Assessment of order books	January 1989
Employment expectations over the next 3 months	January 1989
Prices expectations over the next 3 months	January 1989
Industrial Confidence Indicator	January 1987
Production trend observed in recent months	January 1987
Assessment of order-book levels	January 1987
Assessment of export order-book levels	January 1987
Assessment of stocks of finished products	January 1987
Production expectations for the months ahead	January 1987
Selling price expectations for the months ahead	January 1987
Employment expectations for the months ahead	January 1987
Retail trade Confidence Indicator	January 1989
Business activity over recent months	January 1989
Assessment of stocks	January 1989
Expected business activity	January 1989
Orders placed with suppliers	January 1989
Employment expectations	January 1989
Services confidence indicator	June 1997
Business situation development over the past 3 months	June 1997
Evolution of the demand over the past 3 months	June 1997
Expectation of the demand over the next 3 months	June 1997
Evolution of the employment over the past 3 months	June 1997
Expectations of the employment over the next 3 months	June 1997
Economic Sentiment Indicator - Germany	January 1985
Economic Sentiment Indicator - Spain	April 1987
Economic Sentiment Indicator - France	February 1985
Economic Sentiment Indicator - UK	January 1985

Series (continued)	Starting date
Industrial Production Index - Total	January 1977
Industrial Production Index - Manufacturing	January 1977
Industrial Production Index - Consumer goods	January 1980
Industrial Production Index - Consumer non-durable goods	January 1995
Industrial Production Index - Consumer durable goods	January 1995
Industrial Production Index - Investment goods	January 1980
Industrial Production Index - Intermediate goods	January 1980
Industrial turnover index - Total	January 1990
Industrial turnover index - Manufacturing	January 1990
Industrial turnover index - Consumer goods	January 1990
Industrial turnover index - Consumer durable goods	January 1990
Industrial turnover index - Consumer non-durable goods	January 1990
Industrial turnover index - Intermediate goods	January 1990
Industrial turnover index - Investment goods	January 1990
Industrial turnover index - Total - Domestic Market (DM)	January 1995
Industrial turnover index - Consumer goods - DM	January 1995
Industrial turnover index - Consumer durable goods - DM	January 1995
Industrial turnover index - Consumer non-durable goods - DM	January 1995
Industrial turnover index - Intermediate goods - DM	January 1995
Industrial turnover index - Investment goods - DM	January 1995
Industrial turnover index - Total - External Market (EM)	January 1995
Industrial turnover index - Consumer goods - EM	January 1995
Industrial turnover index - Consumer durable goods - EM	January 1995
Industrial turnover index - Consumer non-durable goods - EM	January 1995
Industrial turnover index - Intermediate goods - EM	January 1995
Industrial turnover index - Investment goods - EM	January 1995
Services turnover index - Total	January 2000
Vacancies	January 1977
Unemployment	December 1977
New applications for employment by the unemployed	January 1979
New job vacancies	January 1979
New occupied jobs	December 1977
Industrial employment index - Total	January 1990
Industrial employment index - Manufacturing	January 1990
Industrial employment index - Consumer goods	January 1990
Industrial employment index - Consumer durable goods	January 1990
Industrial employment index - Consumer non-durable goods	January 1990
Industrial employment index - Intermediate goods	January 1990
Industrial employment index - Investment goods	January 1990
Industrial wages index - Total	January 1995
Industrial wages index - Manufacturing	January 1995
Industrial wages index - Consumer goods	January 1995
Industrial wages index - Consumer durable goods	January 1995

Series (continued)	Starting date
Industrial wages index - Consumer non-durable goods	January 1995
Industrial wages index - Intermediate goods	January 1995
Industrial wages index - Investment goods	January 1995
Hours worked index - Total industry	January 1995
Hours worked index - Manufacturing	January 1995
Hours worked index - Consumer goods	January 1995
Hours worked index - Consumer durable goods	January 1995
Hours worked index - Consumer non-durable goods	January 1995
Hours worked index - Intermediate goods	January 1995
Hours worked index - Investment goods	January 1995
Services employment index - Total	January 2000
Services wages index - Total	January 2000
Hours worked index - Total services	January 2000
Merchandise imports - Total	January 1988
Merchandise imports - Total excluding fuels	January 1990
Merchandise imports - Consumer goods	January 1990
Merchandise imports - Intermediate goods	January 1990
Merchandise imports - Investment goods	January 1990
Merchandise exports	January 1988
Merchandise exports - Total excluding fuels	January 1990
Merchandise exports - Consumer goods	January 1990
Merchandise exports - Intermediate goods	January 1990
Merchandise exports - Investment goods	January 1990
Retail trade turnover index - Total	January 1995
Retail trade turnover index - Food	January 1995
Retail trade turnover index- Non-Durable Non-Food	January 1995
Retail trade turnover index - Durable goods	January 1995
Tourism - Number of nights spent in Portugal	January 1977
Tourism - Number of nights spent in Portugal by residents	January 1977
Tourism - Number of nights spent in Portugal by non-residents	January 1977
Light passenger vehicle sales	January 1982
Light commercial vehicle sales	January 1982
Heavy commercial vehicle sales	January 1982
Cement sales	January 1982
Consumption of electricity	January 1987
Consumption of gasoline	January 1977
Consumption of diesel	January 1987
Real effective exchange rate index	January 1993
PSI-20	January 1988
ATM/POS	September 2000

Why wealth should not be taxed

Pedro Teles with Joana Garcia
Banco de Portugal, Católica Lisbon
SBE

January 2017

Abstract
Even if all the wealth in the economy was owned by one agent alone, with zero weight in the social welfare function, the accumulation of wealth should not be taxed. The workers lose by having the capital-rich pay taxes on the accumulation of capital. (JEL: E60, E61, E62)

Introduction

The wealth distribution in the U.S., and also in most of the developed world, has become increasingly more concentrated in the last fifty years, or so. After a long period, from the late twenties to the late seventies, in which wealth became more evenly distributed, that process was reversed in the last five decades. In particular, for the U.S., the top 0.01% of households own today roughly 10% of total net wealth, a figure that is as high as the high levels of the late twenties/early thirties (see Figure 1).[1]

Given the increasing concentration of wealth in the developed world, should wealth taxes be used to redistribute wealth back to the majority of households that have been made relatively poorer? Should the accumulation of capital be taxed, so that taxes on labor may be lowered? How should capital income be taxed relative to labor income?

Based on the work of Chari, Nicolini and Teles (2016) that builds on a large literature, it is shown again here in a simpler framework and using a wealth tax, that the accumulation of wealth should not be taxed. This is the case independently of how wealth and capital are distributed across the households in the model. Even if wealth was all concentrated in the hands of one agent, and this agent had zero weight in the welfare function, even in that case, the accumulation of wealth by that household should not be taxed. All households, rich and poor benefit from wealth accumulation not being taxed. The capital-rich benefit because they are not taxed directly; the workers are

E-mail: pteles@ucp.pt; jomgarcia@bportugal.pt

1. In Portugal, according to the Household Finance and Consumption Survey, in 2013, the top 1% of the population had 15% of total net wealth.

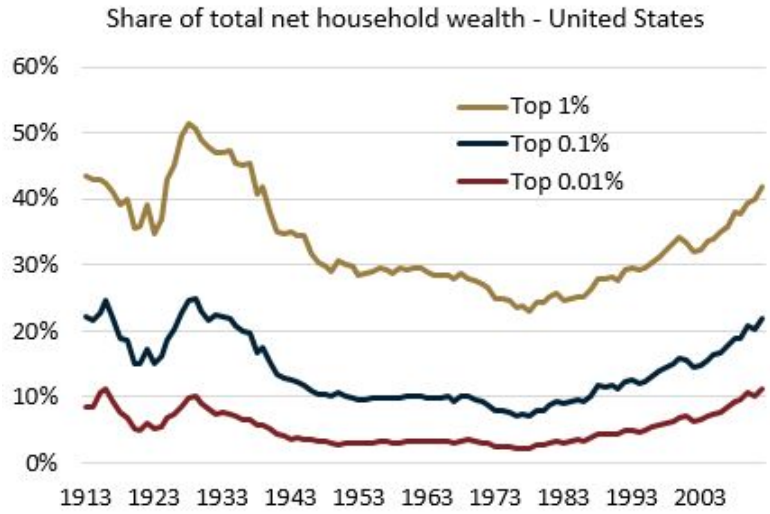


FIGURE 1: Share of total net household wealth - United States
Source: Saez and Zucman (2016)

taxed but their wages net of taxes are higher. The increase in labor income more than compensates the higher taxes.

We solve a simple optimal taxation problem with capital-rich and poor agents and taxes on labor income, wealth and capital income. The model is the standard neoclassical growth model with constant elasticity preferences. The main result is that the optimal taxes on the accumulation of wealth are zero. While the accumulation of wealth should not be taxed, a completely different matter is whether the initial wealth should be taxed. The initial wealth in the model is the one the households start with, today, when the policy change is being considered. Taxing today's wealth does not distort marginal decisions because the accumulation of today's wealth was decided in the past. It may be optimal to tax the initial wealth in order to transfer resources to the government and across households, depending on the distribution of wealth and on the social welfare function.

The main conclusion of this article is that the only tax on wealth that may be desirable in a standard macro model with capital-rich and poor agents is a confiscatory tax, never to be repeated. Now a tax that is levied, not to be repeated is a difficult tax to levy. Because if the tax is not to be repeated, how can it be levied in the first place? The initial confiscatory tax would be defrauding previous implicit promises or expectations. Otherwise there would be much less, possibly none, wealth to confiscate. There is a sense in which, if the government must confirm previous expectations of the returns on assets net of taxes, then, even in the initial period, wealth would not be

taxed, because taxing it would imply defrauding those expectations (see Chari *et al.* (2016) for a discussion of this).

What is the reason for this striking result, that the poor should pay taxes, and not the rich? Further discussion of the intuition follows next.

Understanding why wealth should not be taxed

Why is it that it is never optimal to tax capital accumulation, even if capital is concentrated in the hands of only a few? In the model economy the preferences are such that the price elasticity of consumption is the same for all periods, and, similarly, the wage elasticity of labor is also the same for all periods. The simple principle of optimal taxation, that goods that are equally elastic should be taxed at the same rate, should apply here. For this reason, consumption in every period would be taxed at the same rate and labor in every period would also be taxed at the same rate.[2] This means that capital accumulation would not be taxed, because what the taxation of capital does is tax consumption and labor at different rates over time.[3] This argument should apply independently of issues of distribution. All agents have the same preferences, with the same constant elasticities. All goods, in all periods, for all agents, should be taxed at the same rate.

There is another basic principle of optimal taxation, that should also apply here: Pure rents ought to be taxed, both to transfer resources at zero cost from the private agents to the government and to distribute resources across agents. The two principles, of uniform taxation and taxation of rents, could conflict.

Not so, in the model in this article. In the model here, there are instruments to tax pure rents that are independent from the instruments used to tax consumption and labor at uniform rates. The pure rents in this model are the rents from the initial wealth. And, in this model, it is possible to deal directly with the initial confiscation of wealth, independently of how future wealth is taxed. There is no need to deviate from uniform taxation in order to confiscate the initial wealth.

The results in this article, after Chari *et al.* (2016), differ from the results in the influential papers of Chamley (1986) and Judd (1985) and more recently Straub and Werning (2014) because the conflict between those two principles, of uniform taxation and taxation of pure rents, is present in that literature. The reason is that there are additional restrictions on the tax instruments. In particular they only consider a capital income tax that cannot exceed 100%.

2. This is a well known result in public finance attributed to Atkinson and Stiglitz (1972) which is also an application of the result in Diamond and Mirrlees (1971) on the optimality of productive efficiency.

3. A positive tax on capital accumulation taxes consumption tomorrow more than consumption today, and labor today more than labor tomorrow.

Here, instead, we consider a wealth tax, also restricted not to exceed 100%. In Chamley (1986) and Judd (1985), it is shown that it is optimal to fully tax capital income for a while in order to partially confiscate the initial installed capital. In Straub and Werning (2014), it is shown that the full taxation of capital income could actually last forever, also as a way to confiscate the initial wealth.

The optimal tax on the accumulation of wealth is zero starting today. This is the case independently of the concentration of wealth and the weights of the different agents in the welfare function. The fact that capital accumulation should not be taxed even if capital is all owned by one agent with zero weight, means that the workers without capital benefit from the owners of capital being exempt from taxes on capital accumulation. But it does not mean that capital should not be taxed at all. Today's installed capital will in principle be taxed, depending again on the distribution of wealth and on the welfare function. Taxing the installed capital today does not distort marginal decisions, so the only reason not to tax it is distributional. Now, does this make sense, that future capital should never be taxed, independently of the distribution of capital, while today's installed capital should be taxed fully, except for reasons of distribution?

If future capital should never be taxed, then in the future when future capital becomes today's installed capital, how can it then be taxed fully? How can the government be committed not to tax in the future, when it is free to tax in the present? Indeed, the only reason why it is optimal to tax today's installed capital is because the tax payers are surprised by a tax that was not taken into account in the past. If a government is committed not to defraud expectations on net returns, that will rule out the confiscation of the initial installed capital, whether that is done directly, or indirectly through the taxation of future capital.

As it turns out, the result that the taxes on capital accumulation are exactly zero hinges on the preferences, standard in macro models, that have constant elasticity for consumption and labor. In general, with elasticities that may be time varying with the allocation, it is going to be optimal to tax consumption in different periods at different rates, and labor in different periods also at different rates. In that case, it may be optimal to either tax or subsidize future capital, depending on whether future elasticities are larger than current elasticities. In the steady state, the allocation is constant and therefore, the elasticities are also constant. Therefore, in the steady state capital should not be taxed, as argued by Chamley (1986) and Judd (1985) among others.[4] Even if not fully general, the result that capital accumulation should never be taxed is a very useful benchmark. It suggests that there may be no major

4. See also Atkeson *et al.* (1999) and Chari *et al.* (1994).

economic justification for a recurrent tax on wealth, regardless of how wealth is distributed in the economy.

The remaining of this article contains the technical proofs of the results for the optimal taxes, first, in an economy with a representative agent and, second, in economies with capital-rich and poor agents.

The neoclassical growth model with taxes

The model is the deterministic neoclassical growth model with taxes. This is the standard model used in the literature on capital income taxation, in particular in Chari *et al.* (2016). The preferences of a representative household, over consumption c_t and labor n_t, are described by a standard utility function with constant elasticity in consumption and labor,

$$U = \sum_{t=0}^{\infty} \beta^t \left[\frac{c_t^{1-\sigma} - 1}{1-\sigma} - \eta n_t^{1+\psi} \right], \tag{1}$$

with $\sigma > 0$ and $\psi > 0$. σ is the elasticity of the marginal utility of consumption with respect to consumption, which is the inverse of a price elasticity of consumption and ψ is the elasticity of the marginal utility of labor with respect to labor, which is the inverse of a wage elasticity of labor.

The production technology is described by

$$c_t + g_t + k_{t+1} - (1-\delta) k_t \leq F(n_t, k_t), \tag{2}$$

where k_t is capital, g_t is exogenous government consumption, and δ is the depreciation rate. F is constant returns to scale.

The household owns the capital stock and rents it to a representative firm every period at rate u_t.[5] The household accumulates the capital stock k_{t+1}, as well as public debt b_{t+1}. There is a capital income tax τ_t^k, paid by the household on the rental rate of capital with an allowance for depreciation. There is also a wealth tax l_t, paid by the household, levied on the value of capital and outstanding debt. The household also pays a labor income tax τ_t^n. We abstract from other taxes, such as consumption and dividend taxes, because they do not change the problem in fundamental ways.

The flow of funds for the household can be described by

$$\begin{aligned} \frac{1}{1+r_{t+1}} b_{t+1} + k_{t+1} &= (1-l_t)\left[b_t + (1-\delta) k_t + u_t k_t - \tau_t^k (u_t - \delta) k_t\right] \\ &\quad + (1-\tau_t^n) w_t n_t - c_t, \end{aligned}$$

5. If instead capital was accumulated by firms, the results would not change.

for $t \geq 0$. The household maximizes utility (1), subject to the budget constraint obtained from this flow of funds together with a no-Ponzi games condition. The single budget constraint of the household can be written as

$$\sum_{t=0}^{\infty} q_t \left[c_t - \left(1-\tau_t^n\right) w_t n_t\right] \leq \left(1-l_0\right)\left[b_0 + k_0 + \left(1-\tau_0^k\right)\left(u_0 - \delta\right) k_0\right],$$

where $q_t = \frac{1}{(1+r_1)(1-l_1)...(1+r_t)(1-l_t)}$ for $t \geq 1$, with $q_0 = 1$. At the optimum for the household, the constraint holds with equality.

The marginal conditions of the household problem are

$$-\frac{u_{c,t}}{u_{n,t}} = \frac{1}{\left(1-\tau_t^n\right) w_t}, \tag{3}$$

$$u_{c,t} = \left(1-l_{t+1}\right)\left(1+r_{t+1}\right)\beta u_{c,t+1}, \tag{4}$$

and

$$1 + r_{t+1} = 1 + \left(1-\tau_{t+1}^k\right)\left(u_{t+1} - \delta\right), \tag{5}$$

for all t, where $u_{c,t}$ and $u_{n,t}$ are the marginal utilities of consumption and labor in period t.

The representative firm maximizes profits

$$\Pi_t = F\left(k_t, n_t\right) - w_t n_t - u_t k_t. \tag{6}$$

The price of the good must equal marginal cost,

$$1 = \frac{w_t}{F_{n,t}} = \frac{u_t}{F_{k,t}}, \tag{7}$$

where $F_{n,t}$ and $F_{k,t}$ are the marginal productivity of labor and capital, respectively.

Using both the conditions for the household and the firm, the marginal conditions can be written as

$$-\frac{u_{c,t}}{u_{n,t}} = \frac{1}{\left(1-\tau_t^n\right) F_{n,t}}, \tag{8}$$

$$\frac{u_{c,t}}{\beta u_{c,t+1}} = \left(1-l_{t+1}\right)\left[1 + \left(1-\tau_{t+1}^k\right)\left[F_{k,t+1} - \delta\right]\right]. \tag{9}$$

It follows that

$$\frac{u_{n,t}}{\beta u_{n,t+1}} = \frac{\left(1-l_{t+1}\right)\left(1-\tau_t^n\right)}{1-\tau_{t+1}^n}\frac{F_{n,t}}{F_{n,t+1}}\left[1 + \left(1-\tau_{t+1}^k\right)\left[F_{k,t+1} - \delta\right]\right]. \tag{10}$$

These conditions show how the different taxes distort the marginal choices. Indeed, the first best allocation would have the marginal conditions

above with all the taxes set to zero, written as follows,

$$-\frac{u_{c,t}}{u_{n,t}} = \frac{1}{F_{n,t}}, \tag{11}$$

$$\frac{u_{c,t}}{\beta u_{c,t+1}} = 1 + F_{k,t+1} - \delta, \tag{12}$$

$$\frac{u_{n,t}}{\beta u_{n,t+1}} = \frac{F_{n,t}}{F_{n,t+1}} \left[1 + F_{k,t+1} - \delta\right]. \tag{13}$$

The budget constraint of the household can be used, together with the resource constraints, to write the budget constraint of the government.

The marginal conditions of the household and firm can be used to write the budget constraint of the household, as the following condition, which is commonly called an implementability condition,[6]

$$\sum_{t=0}^{\infty} \beta^t \left[u_{c,t} c_t + u_{n,t} n_t\right] \geq u_c\left(0\right)\left(1 - l_0\right)\left[b_0 + k_0 + \left(1 - \tau_0^k\right)\left(F_{k,0} - \delta\right) k_0\right]. \tag{14}$$

We impose restrictions on the tax rates, that they cannot be higher than 100%, so that the tax revenue cannot exceed the base. These are constraints that are usually imposed in this literature, even if somewhat arbitrary.

The implementability condition (14) together with the resource constraints (2) are the only equilibrium restrictions on the sequences of consumption, labor and capital. It is possible that even the implementability condition would not restrict the allocations. If the initial confiscatory taxes could be used to build up enough assets for the government, such that the path of future public consumption could be fully financed, then the only restrictions on the allocations would be the resource constraints. In that case the first best could be achieved. We assume this is not the case. We assume that even if the initial wealth of the household is fully confiscated, the tax revenue is not enough to pay for government consumption. Given the high levels of government consumption and transfers in most developed countries, this assumption is reasonable.

To show that the set of implementable allocations is fully characterized by the implementability condition (14) together with the resource constraints (2), it is necessary to show that all the remaining equilibrium conditions are satisfied. This is indeed the case, since all the other equilibrium condition, other than (14) and (2), are satisfied by other variables as follows:

$$1 = \frac{w_t}{F_{n,t}} \tag{15}$$

6. Allowing for nonnegative public transfers to the household, the condition can be written with greater than or equal.

determines w_t;

$$\frac{w_t}{F_{n,t}} = \frac{u_t}{F_{k,t}} \tag{16}$$

determines u_t;

$$-\frac{u_{c,t}}{u_{n,t}} = \frac{1}{(1-\tau_t^n)\, w_t} \tag{17}$$

determines τ_t^n;

$$u_{c,t} = (1-l_{t+1})\,(1+r_{t+1})\,\beta u_{c,t+1} \tag{18}$$

and

$$1 + r_{t+1} = 1 + \left(1-\tau_{t+1}^k\right)(u_{t+1} - \delta) \tag{19}$$

determine l_{t+1} and r_{t+1}, given τ_{t+1}^k, for all $t \geq 0$.

Notice that the restrictions that the taxes cannot be larger than 100% would not bind here. Notice also that the capital income tax rate was not used for the implementation. This means that it is a redundant tax that can be set equal to zero. Indeed the wealth tax, here, plays the same role of the capital income tax with a gain. While the capital income tax can only tax the net income on capital, $(u_{t+1} - \delta)$, the wealth tax can tax the gross return, $1 + \left(1-\tau_{t+1}^k\right)(u_{t+1} - \delta)$.

Future wealth should not be taxed

The optimal Ramsey policy can be obtained by solving the problem of maximizing utility subject to the implementability condition (14) and the resource constraints (2). The first straightforward result is on the optimal initial confiscation. In this economy it is optimal to fully confiscate initial wealth. The household benefits, because the marginal choices are not affected by the initial confiscation, and the higher that revenue is, the lower must the future distortionary taxes be.

The Ramsey problem then becomes the maximization of utility subject to the implementability condition with $l_0 = 1$,

$$\sum_{t=0}^{\infty} \beta^t \left[u_{c,t} c_t + u_{n,t} n_t\right] \geq 0, \tag{20}$$

together with the resource constraints (2). The first order conditions of this problem are the following:

$$-\frac{u_{c,t}}{u_{n,t}} = \frac{1+\varphi\,(1+\psi)}{1+\varphi\,(1-\sigma)}\frac{1}{F_{nt}},\, t \geq 0, \tag{21}$$

$$\frac{u_{c,t}}{\beta u_{c,t+1}} = 1 + F_{k,t+1} - \delta,\, t \geq 0, \tag{22}$$

where φ is the multiplier of the implementability condition. It follows that

$$\frac{u_{n,t}}{\beta u_{n,t+1}} = \frac{F_{n,t}\left[1 + F_{k,t+1} - \delta\right]}{F_{n,t+1}}, \text{ for all } t \geq 0. \tag{23}$$

Notice that if the multiplier of the implementability condition was zero, $\varphi = 0$, then the conditions above would be the conditions of the first best, $(11) - (13)$.

With a strictly positive multiplier, then the marginal Ramsey conditions, $(21) - (23)$, mean that the optimal intratemporal wedge is constant over time, while the optimal intertemporal wedges are zero. In this economy it is optimal not to distort intertemporally starting in period zero.

By comparing the Ramsey marginal conditions to the equilibrium ones, $(8) - (10)$, distorted by the marginal tax rates, it becomes apparent how taxes should be optimally chosen. The optimal labor tax needs to be constant over time, and, in general, positive, so that tax revenue may be raised to finance government spending. On the other hand, capital accumulation should not be distorted and therefore both taxes on capital income and wealth ought to be set to zero, $l_{t+1} = 0$ and $\tau^k_{t+1} = 0$, for all $t \geq 0$.

Future wealth should not be taxed independently of the distribution of wealth

In the economy studied above it was shown that the optimal way to tax wealth is to do it once and for all, ex-post, without distorting future accumulation of wealth. Now, the economy we studied was one with a single representative agent, which is a useful construct to analyze macroeconomic aggregate behavior, but it is not necessarily the best model to answer the question of whether labor or capital should be taxed. In order to do this, it is important to allow for households that are capital-rich or poor and inquire whether in that case, it may turn out that capital accumulation should be distorted. As it turns out, it is still the case that even with capital unevenly distributed in the economy, it is not optimal to tax capital accumulation.

To see this, we now consider an economy with two agents, 1 and 2. The social welfare function is

$$\theta U^1 + (1 - \theta) U^2,$$

with weight $\theta \in [0, 1]$. The individual preferences are assumed to be the standard constant elasticity preferences, and are the same for the two agent types,

$$U = \sum_{t=0}^{\infty} \beta^t \left[\frac{\left(c^i_t\right)^{1-\sigma} - 1}{1 - \sigma} - \eta \left(n^i_t\right)^{1+\psi} \right]. \tag{24}$$

The resource constraints are

$$c_t^1 + c_t^2 + g_t + k_{t+1}^1 + k_{t+1}^2 - (1-\delta)\left(k_t^1 + k_t^2\right) \leq F\left(n_t^1 + n_t^2, k_t^1 + k_t^2\right).$$

Tax rates are assumed not to discriminate across agents.

With heterogeneous agents it is no longer the case, that the initial wealth should always be fully taxed. That will depend on the distribution of wealth and on the weights.

The implementability conditions can be written as

$$\sum_{t=0}^{\infty} \beta^t \left[u_{c,t}^1 c_t^1 + u_{n,t}^1 n_t^1\right] = u_{c,0}^1 \left(1 - l_0\right) V_0^1 \tag{25}$$

and

$$\sum_{t=0}^{\infty} \beta^t \left[u_{c,t}^2 c_t^2 + u_{n,t}^2 n_t^2\right] = u_{c,0}^2 \left(1 - l_0\right) V_0^2, \tag{26}$$

with $V_0^i = \left[b_0^i + k_0^i + \left(1 - \tau_0^k\right)\left(F_{k,0} - \delta\right) k_0^i\right]$. Since the taxes must be the same for the two agents an implementable allocation must also satisfy the following marginal conditions

$$\frac{u_{c,t}^1}{u_{c,t}^2} = \frac{u_{n,t}^1}{u_{n,t}^2}$$

and

$$\frac{u_{c,t}^1}{u_{c,t}^2} = \frac{u_{c,t+1}^1}{u_{c,t+1}^2}$$

that equate the marginal rates of substitution across agents. These conditions can be written as

$$u_{c,t}^1 = \gamma u_{c,t}^2 \tag{27}$$

$$u_{n,t}^1 = \gamma u_{n,t}^2, \tag{28}$$

where γ is a choice variable for the planner.

Let φ^1 and φ^2 be the multipliers of the two implementability conditions, (25) and (26). The first order conditions for $t \geq 1$ imply

$$u_{c,t}^2 \frac{\gamma\left[\theta + \varphi^1\left(1-\sigma\right)\right]\frac{\sigma}{c_t^2} + \left[\left(1-\theta\right) + \varphi^2\left(1-\sigma\right)\right]\frac{\sigma}{c_t^1}}{\frac{\sigma}{c_t^2} + \frac{\sigma}{c_t^1}} = \lambda_t,\ t \geq 1, \tag{29}$$

together with

$$-\lambda_t + \beta\lambda_{t+1}\left[F_{k,t+1} + 1 - \delta\right] = 0.$$

Since, from (27), c_t^1 must be proportionate to c_t^2, $c_t^1 = \gamma^{-\frac{1}{\sigma}} c_t^2$, then it follows that (29) can be written as

$$u_{c,t}^2 \frac{\gamma\left[\theta + \varphi^1\left(1-\sigma\right)\right]\sigma\gamma^{-\frac{1}{\sigma}} + \left[\left(1-\theta\right) + \varphi^2\left(1-\sigma\right)\right]\sigma}{\sigma\gamma^{-\frac{1}{\sigma}} + \sigma} = \lambda_t, \tag{30}$$

and therefore

$$\frac{u^i_{c,t}}{\beta u^i_{c,t+1}} = F_{k,t+1} + 1 - \delta, t \geq 1. \tag{31}$$

Similarly, for labor, the first order conditions for $t \geq 1$ can be written as

$$u^2_{n,t} \frac{\gamma \left[\theta + \varphi^1 (1+\psi)\right] \psi (\gamma)^{\frac{1}{\psi}} + \left[(1-\theta) + \varphi^2 (1+\psi)\right] \psi}{\psi (\gamma)^{\frac{1}{\psi}} + \psi} = -\lambda_t F_{n,t}, t \geq 1, \tag{32}$$

so that,

$$\frac{u^i_{n,t}}{\beta u^i_{n,t+1}} = \frac{F_{n,t}}{F_{n,t+1}} \left[F_{k,t+1} + 1 - \delta\right], t \geq 1, \tag{33}$$

and therefore the intertemporal wedge for labor is also zero. The intratemporal wedge is constant.

From the derivations above it follows that intertemporal margins should not be distorted from period one on. This means that it is not optimal to tax the accumulation of capital after period one. But period one is not the initial period. What about capital accumulation from period zero to period one? Should it be taxed?[7]

The first order conditions for period zero will have additional terms associated with the value of the initial wealth for the different households. It could in principle be desirable to distort capital accumulation in that initial period, between periods zero and one, in order to change the value of the initial wealth, and distribute from the households to the government and across households. As it turns out, that intertemporal distortion is not part of the optimal policy. In this economy with heterogeneous agents there are no restrictions on the initial tax rates, other than the upper bound of 100%. It is always preferable to use those initial tax rates, l_0 and τ^k_0, to optimally confiscate the initial wealth of the two agents, rather than using the distortion on the accumulation of capital in the initial period.

The effects of distorting capital accumulation in the initial period, on the value of the initial wealth, are through the prices, and the prices affect the two agents in the same proportion, as do the two tax rates. To see this, notice that the first order condition for consumption of type one in period zero has an additional term associated with the valuation of the initial wealth,

$$\theta u^1_{c,0} + \varphi^1 u^1_{c,0} (1-\sigma) + \mu^c_0 u^1_{cc,0} - u^1_{cc,0} (1-l_0) \left(\varphi^1 V^1_0 + \varphi^2 \frac{V^2_0}{\gamma}\right) = \lambda_0. \tag{34}$$

The derivative of the lagrangian with respect to l_0 can be written as $u^1_{c,0} \left[\varphi^1 V^1_0 + \varphi^2 \frac{V^2_0}{\gamma}\right]$. At the optimum, either this derivative is zero, if the

7. This could appear to be a technical detail. But it is not. This issue is at the origin of the presumption that capital ought to be taxed at high rates for a while (as in Chamley (1986) and Judd (1985)) or even forever (as in Straub and Werning (2014)).

solution is interior, or else $l_0 = 1$, if the solution is at the upper bound of a 100% tax rate. Either way, the last term in the first order condition, (34), is zero. It follows that the first order condition for period zero, (34), has the same form as the ones for $t \geq 1$.[8] There are also additional terms for labor in the first order conditions at time zero. These are also zero at the optimum.

This means that regardless of how the initial confiscation takes place, whether it is full confiscation or not, it is never optimal to distort the accumulation of future capital. This is the case independently of the weights of the two agents.

8. See Werning (2007) for a related argument.

References

Atkeson, Andrew, Varadarajan V Chari, and Patrick J Kehoe (1999). "Taxing capital income: a bad idea." *Federal Reserve Bank of Minneapolis. Quarterly Review-Federal Reserve Bank of Minneapolis*, 23(3), 3–18.

Atkinson, Anthony B and Joseph E Stiglitz (1972). "The structure of indirect taxation and economic efficiency." *Journal of Public Economics*, 1(1), 97–119.

Chamley, Christophe (1986). "Optimal taxation of capital income in general equilibrium with infinite lives." *Econometrica: Journal of the Econometric Society*, 54(3), 607–622.

Chari, Varadarajan V, Lawrence J Christiano, and Patrick J Kehoe (1994). "Optimal fiscal policy in a business cycle model." *Journal of Political Economy*, 102(4), 617–652.

Chari, Varadarajan V, Juan Pablo Nicolini, and Pedro Teles (2016). "Optimal taxation revisited." *mimeo*.

Diamond, Peter A and James A Mirrlees (1971). "Optimal taxation and public production I: Production efficiency." *The American Economic Review*, 61(1), 8–27.

Judd, Kenneth L (1985). "Redistributive taxation in a simple perfect foresight model." *Journal of Public Economics*, 28(1), 59–83.

Saez, Emmanuel and Gabriel Zucman (2016). "Wealth inequality in the United States since 1913: Evidence from capitalized income tax data." *The Quarterly Journal of Economics*, 131(2), 519–578.

Straub, Ludwig and Ivan Werning (2014). "Positive long run capital taxation: Chamley-Judd revisited." *National Bureau of Economic Research*, No. w20441.

Werning, Ivan (2007). "Optimal fiscal policy with redistribution." *The Quarterly Journal of Economics*, 122(3), 925–967.