1

BANCO DE PORTUGAL
ECONOMIC STUDIES

1

Banco de Portugal Economic Studies

Volume VIII

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

Lisboa, 2022 • www.bportugal.pt

Banco de Portugal Economic Studies | Volume VIII – no. 1 | Lisbon 2022 • Banco de Portugal Av. Almirante Reis, 71 | 1150-012 Lisboa • www.bportugal.pt • Edition Banco de Portugal • Design Communication and Museum Department | Design Unit • ISSN (*online*) 2183-5217

Content

Editor's note[1]

Pedro Duarte Neves

January 2022

---

1. This issue of *Banco de Portugal Economic Studies* includes four studies. The first study examines the use of energy in Portugal, covering aspects such as energy dependence, energy intensity and the composition of the energy mix. The second study assesses the duration of insolvencies in Portugal, identifying the factors that are associated with the duration of these proceedings. The third study describes corporate income tax in Portugal and analyses the link between the various characteristics of firms and their effective tax rates. Finally, the fourth study assesses the link between firms' productivity and workers' skills. These studies share a feature: the use of individual corporate data to characterise the heterogeneity observed in the variables of interest.

2. Amador's study provides important findings on energy use in Portugal and in the euro area over the past thirty years. Using 2019 as a reference year, the following can be concluded: Portugal's degree of energy dependence was around 10 percentage points higher than that of the euro area (80% and 70% respectively); the degree of energy intensity, measured as energy consumption per unit of value added, was around 25% higher in Portugal; in terms of energy composition, the contribution of renewable energies in Portugal was around 10 percentage points higher (contributions of 27 and 16% respectively); finally, the share of electricity in total energy consumption was around 25%, both in Portugal and in the euro area.

The study also characterises – in different productive sectors and at firm level – the degree of energy intensity and the level of electricity use as a proportion of total energy in Portugal. The level of energy use shows an inverse relationship with firm size, measured by turnover or number of employees; in the case of electricity, the level of use also tends to decrease with firm size, although this relationship is not so clear.

3. Efficient insolvency procedures help non-viable debt to be resolved quickly and viable debt to be restructured in a sustainable way. International organisations – the European Commission, the European Central Bank, the OECD and the International Monetary Fund – have repeatedly expressed the need for more efficient insolvency mechanisms in Europe and also for convergence between national practices, which are

---

E-mail: pneves@bportugal.pt

1. The analyses, opinions and conclusions expressed in this editorial are entirely those of the editor and do not necessarily coincide with those of Banco de Portugal or the Eurosystem.

currently very diverse. Differences across national insolvency frameworks may result in consequences as varied as different financing costs, different rates of reallocation of productive resources in the course of an economic recovery and, ultimately, constitute an impediment to greater financial integration in the euro area.

As part of a set of recent proposals[2] to boost the Capital Markets Union, the European Commission is committed to submitting a legislative initiative in 2022 with a view to promoting greater harmonisation of the corporate insolvency framework and procedures, in order to ensure greater predictability for economic actors (firms, investors, employees, suppliers and the State itself). This is a matter of the utmost importance in the context of the ongoing recovery of the European economy. Against this background, it is highly important to have a good knowledge of insolvency procedures in Portugal, identifying aspects that could be improved.[3]

4. Duration is a key element in the ex-post assessment of the efficiency of insolvency proceedings. This is the focus of the second study in this issue of *Banco de Portugal Economic Studies*: Pereira and Wemans analyse the duration of insolvencies in Portugal from the period September 2014 to December 2024 – both for private individuals and firms. To this end, the study focuses on two concepts: duration until declaration of insolvency, where the debtor's assets are seized for the benefit of the insolvent estate and where pending enforcement proceedings and seizures are suspended; and the duration until the insolvency is concluded, which coincides with the distribution of the insolvent estate among creditors or with the start of the insolvency plan.

Overall, these timelines have been shortened: the median time-span until declaration of insolvency was six days for private individuals and 17 days for firms in 2020 (in the 75th percentile, 14 days and 60 days respectively); the median time-span until conclusion of the insolvency process was around four months for private individuals and around 15 months for firms also in 2020 (in the 75th percentile, six months and 45 months respectively).

The authors, by exploring a number of databases, pinpoint aspects that are likely to be associated with a longer duration of these proceedings, with the aim of identifying any room for efficiency gains in these legal instruments. As regards duration until completion of the insolvency proceedings, this study yields some interesting results. In the case of private individuals, the duration tends to increase if the applicant is a creditor, there is more than one debtor, the number of creditors increases and where there are private creditors. In contrast, timelines tend to be shorter where there are

2. "Capital Markets Union: Commission proposes new measures to boost Europe's capital markets", 25 November 2021, Brussels, press release.

3. The European Banking Authority Report on the benchmarking of national loan enforcement frameworks released in November 2020 provides a description of time to recovery indicators and recovery rates concerning bank loans to enterprises and private individuals in the euro area.

creditors from financial institutions, enterprises or general governments. In the case of enterprises, timelines tend to be longer, again, if the applicant is a creditor, the number of creditors is higher, there are private creditors, and also in accordance with the size of the enterprise, the amount owed, the assets to be disposed of and, finally, if there is real collateral. The timeline for insolvencies is particularly long in the construction sector.

5. The study by Braz, Cabral and Campos reviews corporate income tax (IRC) in Portugal. Portugal is currently one of the OECD countries with the highest statutory maximum corporate income tax rate; however, corporate income tax revenues as a percentage of Gross Domestic Product (GDP) are very close to (although slightly above) those of the OECD average.

Corporate tax systems are particularly complex as the actual tax burden depends on a range of factors such as benefits, incentives, deductions, surcharges and various legal provisions. In Portugal, corporate income tax is relatively concentrated, with large firms – accounting for 0.5% of the total number of firms – providing around 45% of total revenue, while micro firms – accounting for around 80% of the total number of firms – contribute just under 16%.

To better understand the impact of corporate taxation in Portugal, the authors build a statistical description of the effective tax rates at firm level. The main findings are that effective tax rates tend to drop with financial leverage and capital intensity – which is in line with the tax environment for interest expenditure and investment in fixed assets – and show a non-linear relationship with firm size and productivity level. Finally, the authors have not identified major differences by sector of activity.

6. The study by Cima, Pimenta, Portela and Silva looks at the relationship between firm productivity and worker skills, using data from *Quadros de Pessoal* and the *Sistema de Contas Integradas das Empresas* ("Integrated Enterprises Accounts System") for the period 2006-18. To measure worker skills, the authors develop an indicator that includes the number of years of education, age and a general skills estimate obtained from a Mincer wage equation.

In line with the economic literature, the authors use two statistical moments to describe the distribution of workers' skills: the mean value, which has a positive relationship with productivity, and a measure of dispersion, the standard deviation, which has a negative relationship with productivity. Both results are statistically significant and robust to alternative measures of productivity, skills and skill heterogeneity. This study, in line with relatively well-established results in the literature, suggests that firms with a more homogeneous workforce – in terms of skills and education – tend to show productivity gains over firms with a less homogeneous workforce.

Non-technical summary

January 2022

Energy mix and intensity in Portugal: Portraits from aggregate and firm-level data

João Amador

The consumption of energy per capita in Portugal is lower than in the euro area. Nevertheless, total energy consumption in Portugal increased by 34 per cent since 1990, which compares with a 2 per cent increase in the euro area in the same period. The energy dependence ratio concerns the proportion of energy that an economy must import. Portugal posts a higher energy dependence ratio than the euro area. In the period 2017-2019, the ratio stood at values slightly higher than 80 per cent, after having decreased from values close to 90 per cent in the mid 2000s. In the euro area the ratio has been increasing since 2013 to values close to 70 per cent in 2019.

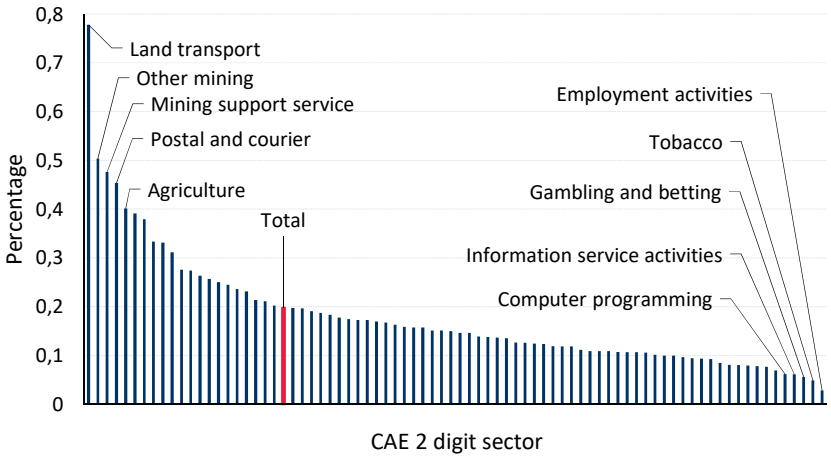
Energy intensities, expressed in kilotons of oil equivalent (Ktoe) per million of GVA, taken at constant prices, were close in Portugal and the euro area in 1995 and developments have been quite positive, even if improvements only started in Portugal in the mid-2000s. From 1996 up to 2017 the accumulated reduction in this indicator reached 9 and 25 per cent in Portugal and the euro area, respectively.

The energy supply mix is the structure of energy supply by primary source as a percentage of total energy supply in the country. A key dimension of interest, in connection with climate challenges, is the share of renewables in the energy supply mix and in the production of electricity, i.e, the role of hydro, geothermal, wind, biomass, waste and solar energy. In Portugal, there has been a steady increase in the share of renewables and biofuels since 2000, reaching 27 per cent in 2019. In contrast, in the euro area the share of renewables was about 16 per cent in 2019. This accrues to the share of nuclear energy, which accounted for about 15 per cent ot total energy supply in the euro area in 2019. As for the breakdown of electricity consumption along the different primary energy sources that generate it, there is still a limited share of electricity derived from renewable sources, especially in the euro area, which again links with the role of nuclear energy as a source of electricity (36 per cent in 2019).

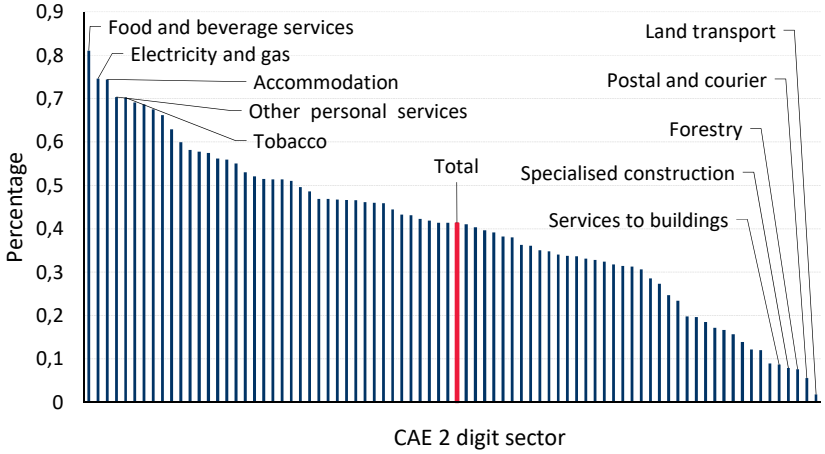
Firm-level data on electricity and liquid fuels expenditures, combined with gross value added, turnover and employment data provide additional important results. The top panel of Figure 1 ranks NACE 2-digit sectors from the highest to the lowest in terms of the average ratio of total energy expenditure on GVA in 2018. Results show important differences across sectors, with “land transport” posting a ratio of 78 per cent, in contrast with “employment activities”, “tobacco” with average ratios lower than 5 per cent. The bottom panel ranks sectors from the highest to the lowest in terms of the average ratio of electricity on total energy expenditure in 2018. Sectors “land transport” and “postal and courier” post ratios lower than 7 per cent, in strike contrast with “food and beverages

services", "electricity and gas" and "accommodation" with average ratios higher than 70 per cent.

Firms' size is a very important dimension of analysis as specific energy-related technologies may be dependent on scale, both in terms of technical feasibility and rate of return of investments. The relationship between firms size and energy consumption patterns is also relevant for the correct design of public policies. Using a simple regression approach, and controlling for firm and time fixed effects for different size classes of firms, we conclude that the correlation between firms energy intensity and their size, measured either in terms of total turnover or employment, is negative. However, the negative correlation between firms electricity share and size is not so strong. In addition, there is evidence that a higher share of electricity on total energy expenditure is negatively correlated with the energy intensity in the firms.



(A) Average share of energy expenditure on total GVA



(B) Average share of electricity in energy expenditure

FIGURE 1: Energy intensity and share of electricity across sectors in 2018

Energy mix and intensity in Portugal: Portraits from aggregate and firm-level data

João Amador
Banco de Portugal
Nova School of Business and Economics

January 2022

Abstract
This article presents the path of several aggregate energy indicators for Portugal and the Euro area in the last three decades. In addition, we use Portuguese firm-level data on electricity and liquid fuels expenditures to assess firms' electrification and efficiency, while also correlating those indicators with their size. The article ultimately aims at sheding light on the pace of transformation towards an electrically driven, renewable and energy efficient economy. Overall, we identify some progress in the renewable-based electrification of the Portuguese and Euro area economies, as well as sizeable progress in energy intensity in Portugal since the mid 2000s. Moreover, controlling for firms' heterogeneity, we find a robust negative correlation between energy intensity and firms' size and partial evidence of a negative correlation between the share of electricity on total energy expenditures and firms' size. Finally, we identify a negative correlation between the share of electricity on firms' energy expenditure and their energy intensity ratio. (JEL: Q40, L21, L25)

1. Introduction

Energy is indispensable for economic activity. Indeed, every single human activity requires some degree of energy consumption and it is hard to conceive the full impact in our lives of a continued collapse in energy supply. Nowadays, energy price hikes coupled with the need to restructure the energy sector and consumption patterns in order to phase out fossil fuels and meet targets for greenhouse gas reductions, have brought the topic to the forefront of economic debate.

The impact of energy prices and their pass-through on inflation dynamics is well known. Supply shortages due to geopolitical developments, natural disasters or the simple exertion of market power by producers, in a context where the demand curve is relatively rigid, lead to price spikes. These increases do not affect core inflation unless they are sustained and feed into expectations and wage updates. In this article we do not directly discuss the link between energy and inflation but we assess the dependence of

Acknowledgements: The author is thankful to Nuno Alves, António Antunes, Joana Garcia, Carlos Gouveia, Pedro Duarte Neves, Ana Catarina Pimenta and Cátia Silva for their very useful comments and suggestions. The opinions expressed are those of the author and do not necessarily coincide with those of Banco de Portugal or the Eurosystem.

E-mail: jamador@bportugal.pt

the Portuguese and euro area economies on external energy supplies, which is a measure of exposure to some of the referred risks. In addition, we analyse the primary energy mix, which depends on the natural conditions of countries and also has a bearing on their ability to diversify energy supply risks and manage prices.

The link between our analysis, the green transition and climate change challenges is tighter. By assessing the share of renewable-based electricity on the total amount of energy consumed in the economy we can infer how far it is from abandoning combustion engines and other fossil fuel burning technologies for the purposes of heating or transport. A strong renewable-based electrification of the economies is a condition for the fulfilment of targets regarding zero emissions. Another related feature is the total elimination of solid fossil fuels, oil and gas from the portfolio of primary energy sources. Nevertheless, such progress is quite dependent on technological solutions that are not yet totally available, for example in what concerns the storage of large amounts of energy.

The debate on energy and climate action is also inexorably linked to energy intensity. The ability to generate value-added with the smallest energy consumption possible is also key to meet the internationally agreed targets for greenhouse gas reductions. Energy intensity depends on the climatic conditions of countries and also on the type of technology used and the organization of production.

In this article we try to contribute to these debates by computing a set of energy-related indicators for Portugal and the Euro area in the period 1990-2019. This long time span makes it possible to gauge the pace of transition towards a scenario where fossil fuels are absent from the portfolio of primary energy sources and renewable energies prevail. In parallel, it is important to assess gains in terms of energy intensity, i.e., lower amounts of energy used for each unit of value-added produced in the economy, in industries and in individual firms. In the aggregate analysis, in order to set useful benchmarks, we compare domestic aggregate and industry-level energy intensities with those of the euro area.

A novel element in this article is the use of detailed firm-level data regarding expenditure in electricity and liquid fuels. The firm-level analysis discloses strong heterogeneity across firms in terms of nominal energy mix and intensity, a feature common to numerous other firm-level characteristics. In addition, the granular information provided by firm-level data provides insights for targeted economic policies. The "European Green Deal" targets a 55 per cent reduction in carbon emissions in the EU compared to 1990 levels by 2030, and aims the block to become carbon-neutral by 2050. Meeting these goals will forcefully imply policies targeted at energy adaptation, which may have a different impact depending on firms' characteristics. Firms' size is a very important dimension of analysis as energy-related technologies may be dependent on scale, both in terms of their technical feasibility and investments' rates of return. It is also worth noting that EU internal restrictive policies prompt the relocation of production to countries with lower carbon prices and softer legislation, which links with the proposal for a EU Carbon Border Adjustment Mechanism (CBAM), currently under discussion.

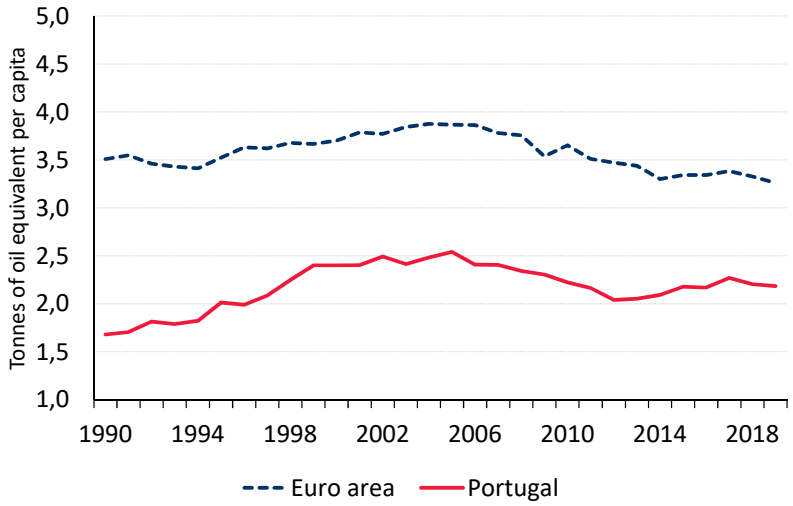


FIGURE 1: Energy consumption per capita

Notes: Author's calculations based on Eurostat energy balances and European Commission AMECO database.

The consumption of energy per capita in Portugal is lower than in the euro area (Figure 1). Nevertheless, total consumption of energy in Portugal increased by 34 per cent since 1990, which compares with a 2 per cent increase in the euro area in the same period. The rise in total consumption was extremely strong until the mid-2000s but it decreased until 2012 and broadly stabilized afterwards. A milder but qualitatively similar path was observed in the euro area. Overall energy consumption levels and developments depend on several important features that are identified in the article.

Portugal has made progress regarding the share of renewables and biodiesel, which represent close to a third of total energy consumption. The share of these primary energy sources in the euro area is smaller than in Portugal (about 16 per cent in 2019). Nevertheless, the Portuguese energy dependence remains substantially higher than that of the euro area, which poses challenges in terms of exposure to external shocks. As for energy intensity, Portugal has made substantial progress since the mid-2000s but gains in the euro area have started earlier and been steadier. As for the results derived from individual data on energy expenditure, there is strong heterogeneity across firms. The distribution of energy expenditure on total value added resembles a Pareto distribution and that of the share of electricity on total energy expenditure is bimodal in the tails. The correlation between firms' energy intensity and their size, measured either in terms of total turnover or employment, is negative. The negative correlation between firms' electricity share and size is not so strong. In addition, there is evidence that a higher share of electricity in total energy expenditure is negatively correlated with the energy intensity in the firms.

The article is organized as follows. In the next section we briefly review the literature on energy consumption in Portugal and in an international perspective. Section 3 provides information on the three databases used in the analysis. Section 4 presents

the path of energy dependence, energy intensity and energy mix at the aggregate and sectoral levels, always taking the euro area as a benchmark for the Portuguese situation. Section 5 uses firm-level data on yearly expenditures in energy items, GVA, turnover and employment to assess the energy intensity, electricity share and the relationship between these indicators and firms size. Finally, section 6 offers some concluding remarks.

2. Literature

The literature on energy mix and energy intensity is vast and such a survey is totally beyond the scope of this article. Nevertheless, we make brief reference to studies that relate with our work and that may be of interest to the reader.

Energy efficiency has been an important topic in EU energy policy and this has materialized in specific legislation. The first Directive on energy efficiency dates back to 2006 (The European Commission (2006)) and a revision is presently under discussion (The European Commission (2021)). From an academic perspective, the analysis of energy efficiency in the EU, with an emphasis on Italian and UK regulatory experiences, was studied in Malinauskaite *et al.* (2019).

Beyond the EU, other international organizations regularly trace cross-country developments in energy policies and in the main variables of interest. Examples of yearly analysis are OECD (2020), IEA (2021b) and IEA (2021a). In the same vein, Kaivo-oja *et al.* (2016) studies trends in electricity production and consumption in China, US, the Euro area and the EU in the period 1961–2011 using World Bank and IEA data. Energy intensity analysis, which requires information on value added, is mostly studied at a very aggregate level. Geller *et al.* (2006) reviews energy intensity trends for Japan, United States, and Western Europe since 1973, also considering the role of structural change.

As for Portugal, studies focusing on the path of the main energy variables are also scarce. Nunes (2018) takes a secular perspective on energy developments up to the mid-2000s, while Amador (2010) follows a similar approach analysing the period the period 1960-2008 and comparing Portugal with other European countries. Other contributions based on the Portuguese experience lay on the frontier between energy and environmental issues, often linking with the evaluation of the impacts of policies. Examples are Pereira and Pereira (2019) and Alves *et al.* (2010).

As for the analysis of energy issues at the firm level, the literature is quite scarce. One thoughtful contribution is Zhang *et al.* (2016), which uses a firm-level data envelopment analysis to study energy efficiency in the Swedish industry. In addition, the paper delves into causality to assess how the EU ETS, the carbon dioxide tax and the energy tax affect energy efficiency. Another recent contribution is Lee and Yu (2019), which uses a time-series panel vector model to analyse the interdependencies between energy usage, costs, the share of renewable, economic growth, and greenhouse gas emissions in the Korean industrial sector, with an emphasis on firms' size.

3. Data

This article combines analysis with aggregate and firm-level dimensions, which are not fully integrated due the different nature of the underlying data. The aggregate dimension uses energy data in real terms, i.e., measured in tonnes of oil equivalent (TOE), as well as gross value added (GVA) at constant prices for the overall economy, main sectors and selected manufacturing industries. The firm-level analysis is based in nominal values and corresponds to energy-related expenditures, GVA, turnover and employment for virtually the universe of Portuguese firms.

Energy commodities are often bought for their heat-rising properties and can be converted into derived fuels. Therefore, energy supply and consumption are expressed in specific units (terajoules or TOE) and the structure adopted to report the data is termed as "energy balance". This balance identifies the primary energy sources, breaks down its transformation into secondary energy sources and by the different sectors that consume them, all of it at a quite detailed level in terms of energy types. This allows us to assess the relative importance of the different fuels in terms of their contribution to energy production and consumption. In this article we use the detailed energy balances available online at the Eurostat webpage, covering all EU individual Member-states, the EU and euro area aggregates for the period 1990-2019.

The data on GVA at constant prices was collected from the 2019 release of the EU Klems database run by the Vienna Institute for International Economic Studies (wiiw). The database provides measures of economic growth, productivity, employment, capital formation, and technological change at the industry level for all EU member states, Japan and the US in the period 1995-2017. An overview of data construction issues and methodology is thoroughly presented in Stehrer *et al.* (2019).

The third database used in this article collects detailed balance sheet and income statement information for virtually the universe of Portuguese firms, including sole proprietorships, for the period 2011-2018. The "*Sistema de contas integradas das empresas*" is maintained by Statistics Portugal and derives from firms mandatory reporting to tax and social security authorities as well as the legal obligation to submit their balance sheets. Beyond detailed information regarding expenditure on electricity and liquid fuels, this dataset contains a large number of balance sheet and income statement variables, which allow us to identify firms' characteristics. The non-energy variables used in the article comprise turnover, employment and GVA.

4. Aggregate analysis

This section presents the path of basic energy indicators for Portugal and the euro area since the nineties. The indicators concern energy dependence, energy intensity, detailing along main sectors and industries, and energy mix.

4.1. Energy dependence

The energy dependence ratio concerns the proportion of energy that an economy must import. In this article we define it as net energy imports (imports minus exports) divided by total energy supply minus changes in stocks, expressed as a percentage. A negative dependence ratio indicates a net exporter of energy, while a dependence rate in excess of 100 per cent indicates that energy products have been stocked.

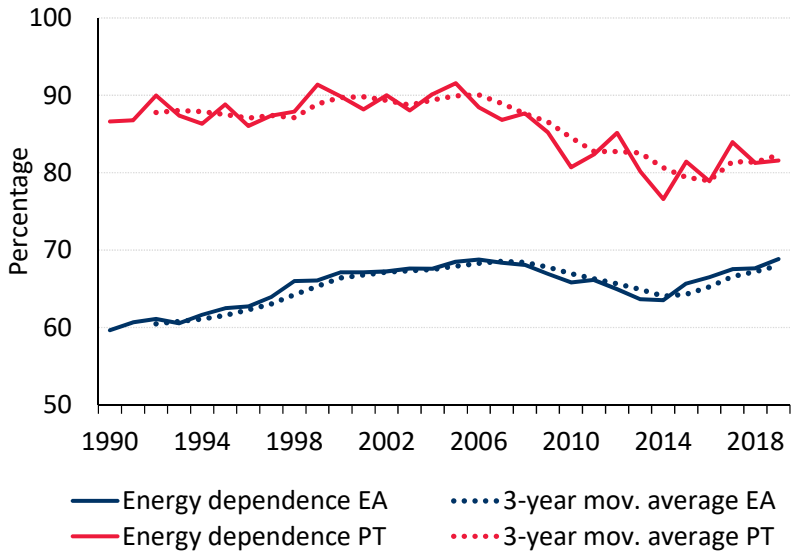


FIGURE 2: Energy dependence
Notes: Authors' calculations based on Eurostat energy balances

Figure 2 presents the energy dependence ratio from 1990 to 2019 in Portugal and the euro area. The figure also reports the 3-year moving average of the indicator in order to smooth out the impact of changes in stocks. Portugal posts a higher energy dependence ratio than the euro area. In the period 2017-2019, the ratio stood at values slightly higher than 80 per cent, after having decreased from values close to 90 per cent in the mid 2000s. In the euro area the ratio has been increasing since 2014 to values close to 70 per cent in 2019.

It is important to note that, although highly relevant, the dependence ratio does not provide full information about energy security, interpreted as the reliability of energy supply. The diversification of foreign energy suppliers, as well as their geographical positioning and political stability are also very important variables when it comes to reduce the exposure to the risk of supply shortages motivated by public health or political crises, conflict or natural disasters. In any case, since the endowment of primary energy sources roots on countries' natural conditions, international trade of energy goods is essential.

4.2. Energy intensity

Energy intensity is typically expressed in kilotons of oil equivalent (Ktoe) per unit of GVA taken at constant prices. It is important to note that energy intensity does not necessarily reflect energy efficiency, as the latter also depends on elements that are not taken into account by the simple measure of energy supply to GVA (e.g., climate and sectoral structure of the economy). Filippini and Hunt (2011) uses a parametric stochastic frontier analysis to obtain an energy demand function for 29 OECD countries over the period 1978 to 2006 and shows the differences between these two concepts. Nevertheless, the energy intensity indicator is very important to identify overall trends linking economic activity and energy use.

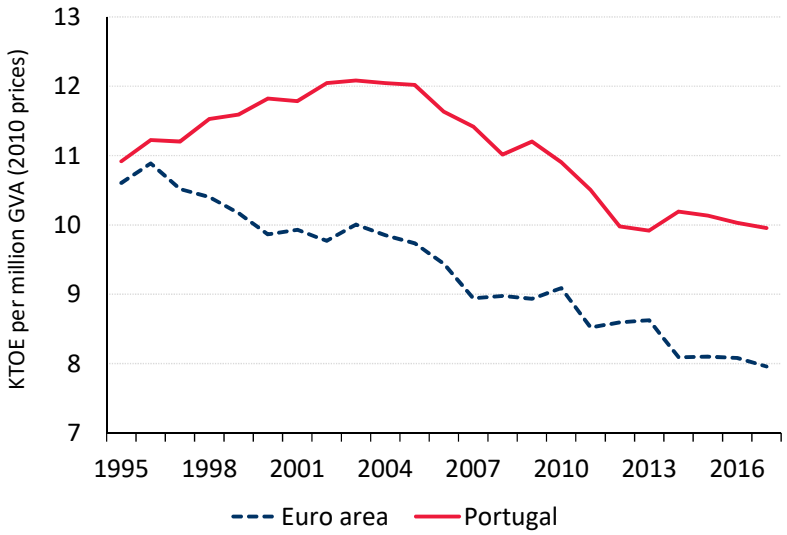
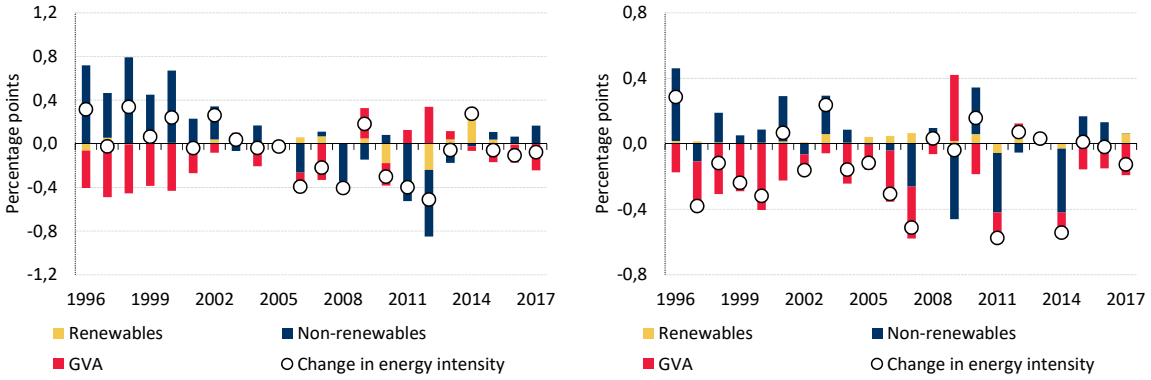


FIGURE 3: Energy intensity

Notes: Authors' calculations based on Eurostat energy balances and EU Klems database.

Figure 3 compares the level of energy intensity in Portugal and the euro area between 1995 and 2017. The level of the indicator was close in both regions in 1995 and developments have been quite positive, even if improvements only started in Portugal in the mid-2000s. From 1996 up to 2017 the accumulated reduction of Ktoe per million euro reached 9 and 25 per cent in Portugal and the euro area, respectively. Nevertheless, in Portugal this indicator increased by 10 per cent between 1996 and 2005. The yearly change in energy intensity can be broken down along the contributions of GVA growth and energy consumption, and also in terms of its renewable and non-renewable components. The panels of Figure 4 present this basic decomposition in Portugal and the euro area. The contribution of GVA developments (the denominator effect) is typically very important to explain yearly changes in the indicator, while the contribution of renewables is not. Since energy is an important input for production, its consumption is positively correlated with activity. Therefore, the contribution of energy consumption has typically the opposite sign of the GVA. This pattern was affected by the sovereign debt crisis in the euro area and by the Portuguese 2011-2014 economic



(A) Portugal (B) Euro area

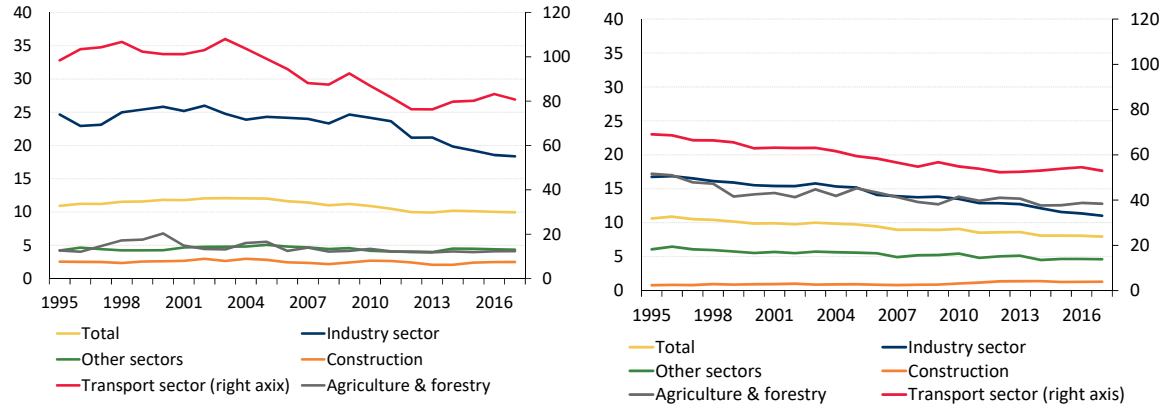
FIGURE 4: Decomposition of change in energy intensity

Notes: Authors' calculations based on Eurostat energy balances and EU Klems database.

and financial assistance program, as well as by progress in energy efficiency in the two economies.

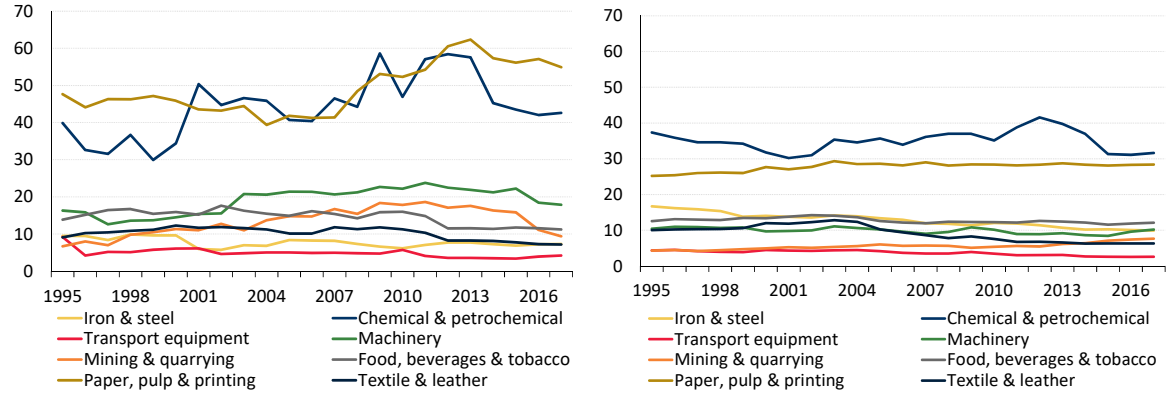
It is relevant to compare energy intensity levels and developments across the main economic sectors. The two panels of Figure 5 present values for Portugal and the euro area and make it clear that "transportation" is, by far, the most energy intensive sector. The level of the indicator in 2017 in this sector is 53 per cent higher in Portugal than in the euro area. The other energy intensive sectors are "industry" and, in the euro area, "agriculture and forestry". If we detail the manufacturing sector the "chemicals and petrochemicals" and "paper, pulp and printing" are the most energy intensive industries both in Portugal and in the euro area (Figure 6). These unsurprising results accrue to the specific nature of these activities, where energy saving technologies would bring important gains.

The two panels of Figure 7 compare energy intensity in 1995 and 2017 in Portugal and in the euro area, while signalling the relative importance of each main sector in energy consumption in 2019. The figure highlights the importance of the transport sector as a user of energy in both economies (36.6 and 31.3 per cent in Portugal and in the euro area, respectively) and also the existence of important progress in Portugal in period considered. The share of industry in energy consumption is smaller (27.3 and 24.3 per cent in Portugal and the euro area, respectively) and also smaller than in "other sectors", which includes "commercial and public services" and "households". In the same vein, the two panels of Figure 8 compare energy intensity in 1995 and 2017 in Portugal and the euro area in selected manufacturing sectors. In the case of the euro area the share of manufacturing sectors on energy consumption is more evenly distributed than in Portugal, where "paper, pulp and printing" stands out, and represents about 9 per cent of total energy consumption.



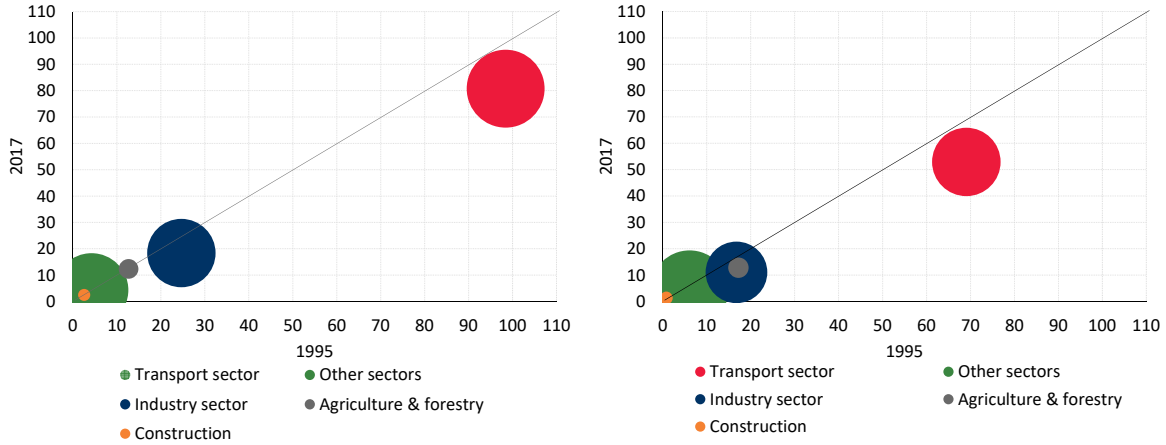
(A) Portugal (B) Euro area

FIGURE 5: Energy intensity in main sectors (Ktoe per million GVA at 2010 prices)



(A) Portugal (B) Euro area

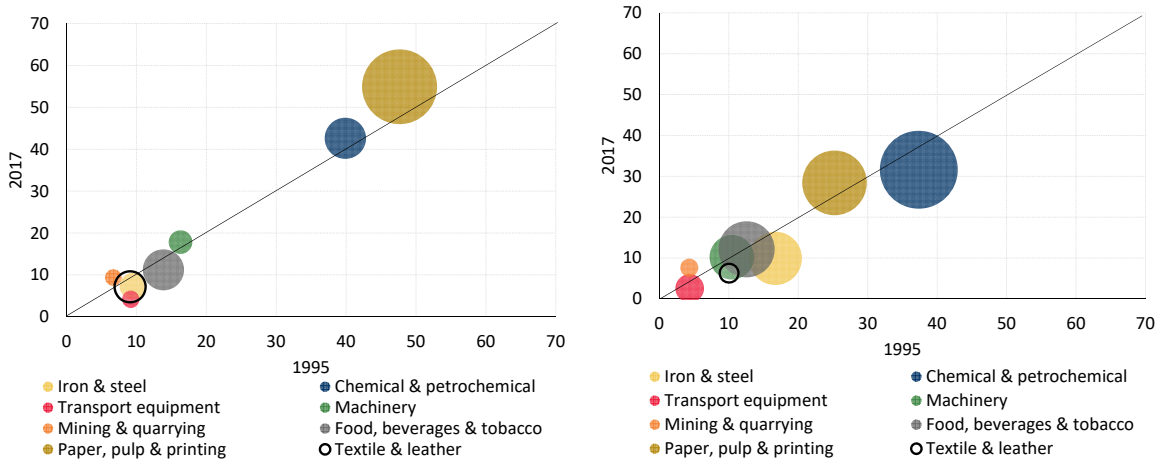
FIGURE 6: Energy intensity in main manufacturing industries (Ktoe per million GVA at 2010 prices)



(A) Portugal (B) Euro area

FIGURE 7: Energy intensity: 1995 vs 2017 - Main sectors (Ktoe per million GVA at 2010 prices)

Note: The diameter of the circles is proportional to the importance of sector in total energy consumption in 2019.



(A) Portugal (B) Euro area

FIGURE 8: Energy intensity: 1995 vs 2017 - Manufacturing (Ktoe per million GVA at 2010 prices)

Note: The diameter of the circles is proportional to the importance of each manufacturing sector in total energy consumption in 2019.

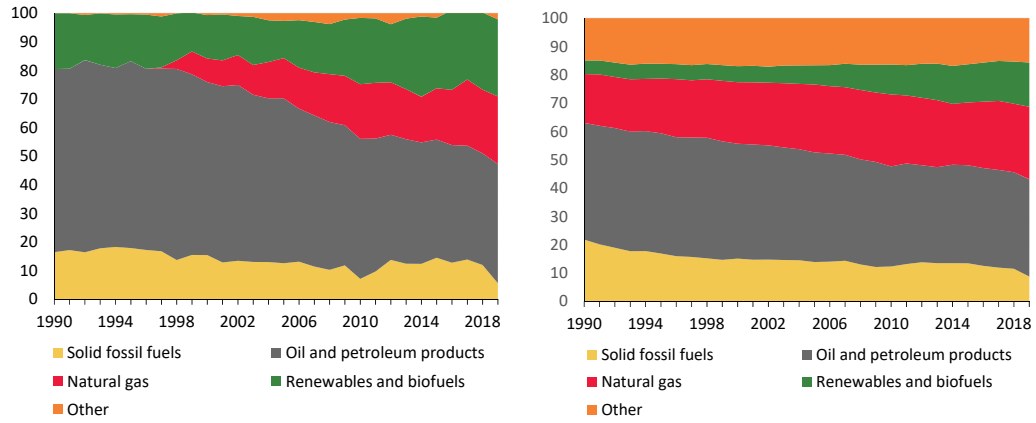
4.3. Energy mix

The energy supply mix is the structure of energy supply in terms of primary energy source as a percentage of total energy supply in the country. A key dimension of interest, in connection with climate challenges, is the share of renewables in the energy supply mix and in the production of electricity, i.e, the role of hydro, geothermal, wind, biomass, waste and solar energy.

The two panels of Figure 9 present the energy supply mix in Portugal and the euro area in the period 1990-2019. The most relevant features in Portugal are the rise of natural gas as a primary energy source in the late nineties, reaching nearly one fourth of total energy supply in 2019, and the steady increase in the share of renewables and biofuels since 2000, reaching 27 per cent in 2019. In contrast, in the euro area there is a much lower share of renewables (about 16 per cent in 2019) and a larger share of "Other" category. The latter difference accrues to the share of nuclear energy, which accounted for about 15 per cent ot total energy supply in the euro area in 2019.

The two panels of Figure 10 detail the share of different primary renewable energy sources in Portugal and the euro area. Except for the larger share of solar energy in the euro area, differences in this structure are not striking in the latest years. One difference is the higher volatility in the share of hydro energy in Portugal, which can be easily understood by the uneven yearly rain patterns, which are averaged out in the larger euro area aggregate.

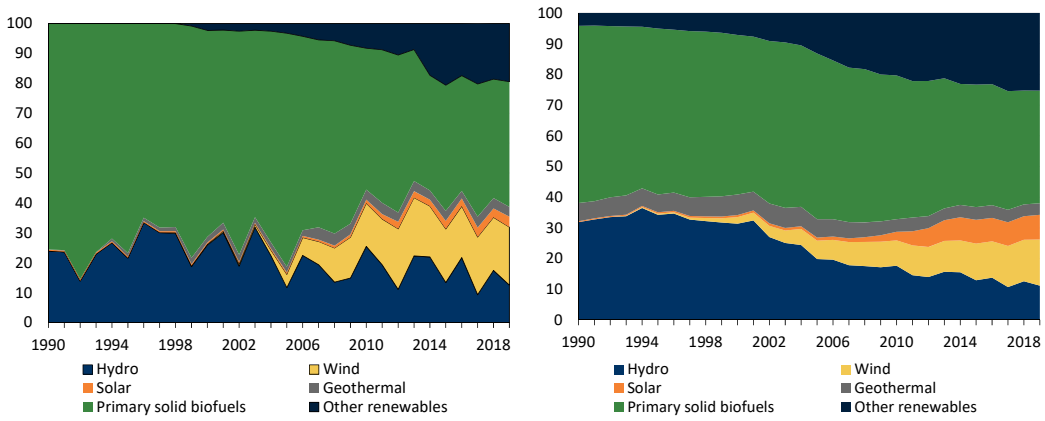
The primary energy mix documented above is transformed into secondary energy sources that are consumed by households and firms. As previously mentioned, it is particularly important to quantify the share of electricity on total energy consumption and its primary energy sources. Indeed, carbon neutrality is expected to be achieved through a renewable-based electrified economy. Figure 11 presents the share of



(A) Portugal (B) Euro area

FIGURE 9: Energy mix in Portugal and in the Euro area

Sources: Eurostat energy balances and author's calculations.
Note: The category "Other" in the graph is mostly composed by nuclear energy.



(A) Portugal (B) Euro area

FIGURE 10: Renewables mix in Portugal and in the Euro area

Sources: Eurostat energy balances and author's calculations.

electricity in total energy consumption in Portugal and in the euro area in the period 1990-2019. This figure shows an upward trend in this share in both regions, but electricity represents only about one-quarter of total energy consumption.

The two panels of Figure 12 present the breakdown of electricity consumption along the different primary energy sources that generate it, both in Portugal and in the euro area. An important result is the still limited share of electricity derived from renewable sources, especially in the euro area. In Portugal this share was 40 per cent in 2019, while in the euro area it was only 22 per cent. Nevertheless, the strong role of nuclear energy as a source of electricity in the euro area must be highlighted (36 per cent in 2019). Although beyond the period under analysis in this article, it is worth remarking the elimination of coal to produce electricity in Portugal in 2021.

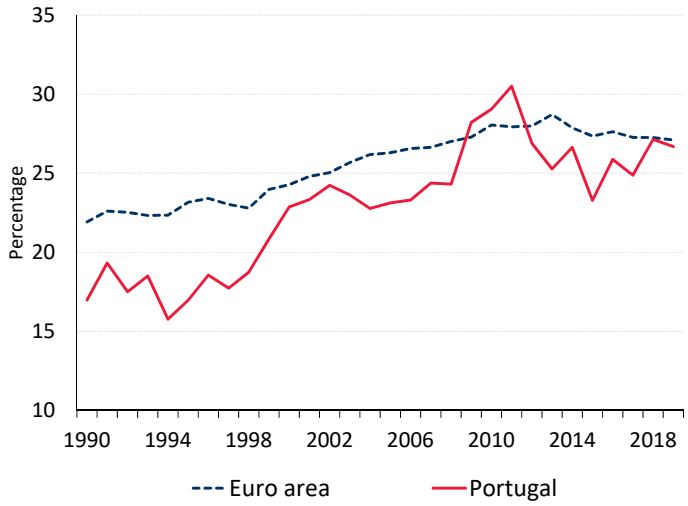
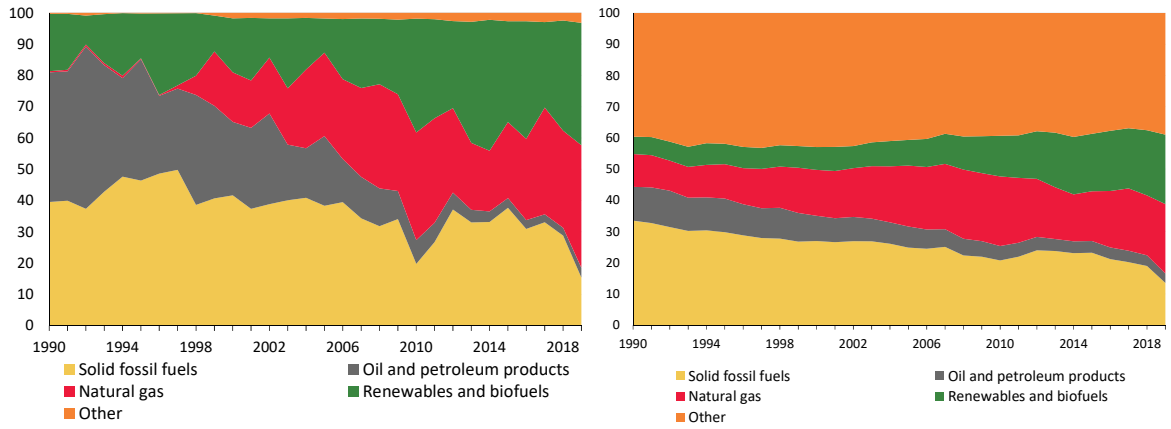


FIGURE 11: Share of electricity in energy consumption
Notes: Authors' calculations based on Eurostat energy balances



(A) Portugal (B) Euro area

FIGURE 12: Breakdown of electricity consumption by primary sources that generate it in Portugal and in the euro area

Sources: Eurostat energy balances and author's calculations.
Note: The category "Other" in the graph is mostly composed by nuclear energy.

5. Firm-level analysis

In the second part of the article we turn to firm-level data on yearly electricity and liquid fuels expenditures to complement the portrait obtained from aggregate energy data and GVA. A limitation of this type of data is the fact that they are expressed in nominal terms, thus combining energy prices and quantities consumed. Firm-level data on energy consumed in real terms is only available for a sample of larger manufacturing firms, thus not sufficiently describing the landscape of firms in the economy.

Although focusing strictly on quantities of energy consumed would be preferable, the analysis of expenditure data is useful and interpretable. Moreover, if the analysis

is conducted in terms of ratios, where prices affect both numerator and denominator, comparisons between firms and density distributions convey relevant information. The fluctuation of prices along time can be overcome by focusing on a cross section of a specific year or by adding time fixed effects in the context of a regression. Another element that may confound the results is the possibility of having different energy prices for firms of different sizes. In theory, large consumers may bargain lower prices from their energy suppliers or public policies may distort energy prices faced by firms of different sizes due to subsidization or taxation rules. In order to overcome this potential difficulty, we detail results for the subset of micro, small, medium and large firms in our sample, classified along the definition used by the European Commission.[1] As a standard cleaning procedure, we eliminate all observations in the database with negative GVA, turnover or expenditures with goods and services.

5.1. Energy intensity of firms

Figure 13 plots the non-weighted kernel distribution of the ratio of energy expenditures on GVA for all firms in the sample in 2018, truncated at the 1st and 99th percentiles. This ratio informs on the energy intensity of firms. The distribution is strongly right skewed, with a large density of firms with low energy expenditures on total GVA and a small number of them with very large values. This Pareto-like distribution is associated with the sectoral specificities of firms' activity. In this vein, Figure 14 ranks NACE 2-digit sectors from the highest to the lowest in terms of the average of the ratio of total energy expenditure on GVA in 2018. Results show again important differences across sectors, with "land transport" posting a ratio of 78 per cent, in contrast with "employment activities" and "tobacco" with average ratios lower than 5 per cent. If the ranking is constructed basing on the median of the ratio in the sector, results are broadly unaltered. The 2-digit level values for the mean, median and interquartile range are presented in Table A.1 in the Appendix.

Firms' size is a very important dimension of analysis as specific energy-related technologies may be dependent on scale, both in terms of technical feasibility and return. The relationship between firms size and energy consumption patterns is also relevant for the correct design of public policies. Therefore, we statistically assess the correlation between the energy intensity and the size of firms, measured either by the logarithm of total turnover or employment, while bearing in mind the sectoral heterogeneity and the problems that emerge from dealing with effects associated to different prices by larger consumers versus smaller ones. These concerns are addressed by considering firm and time fixed effects and separate regressions for firms in different size categories.

1. According to the Recommendation of the European Commission 2003/361/EC, the category of micro-sized firms includes those that employ fewer than 10 workers and have an annual turnover or total annual balance sheet not in excess of 2 million euros. In turn, small firms employ fewer than 50 workers and have an annual turnover or total annual balance sheet not in excess of 10 million euros. The medium-sized firms employ fewer than 250 workers and have an annual turnover not in excess of 50 million euros or a total annual balance sheet not in excess of 43 million euros. Large firms are those that do not belong to any of the aforementioned categories.

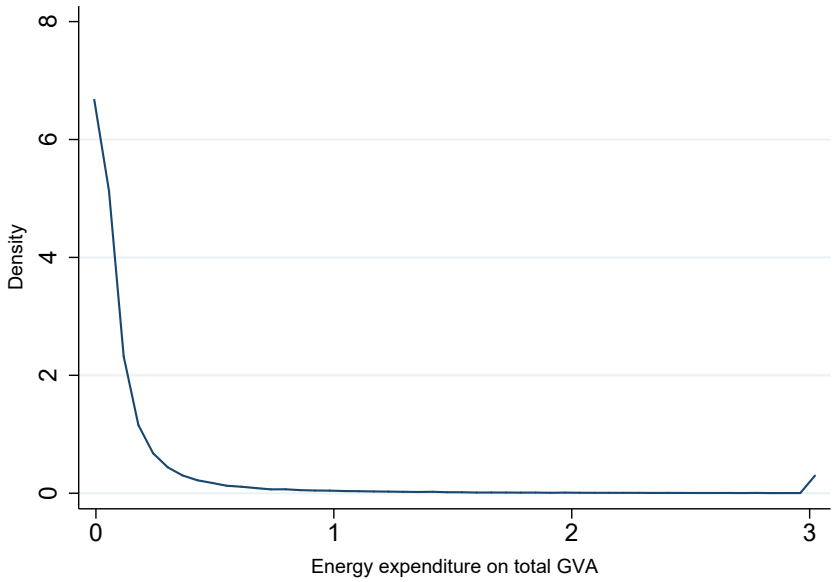


FIGURE 13: Energy expenditure on total GVA in 2018

Notes: Energy expenditure comprises electricity and liquid fuels. Non-weighted observations.

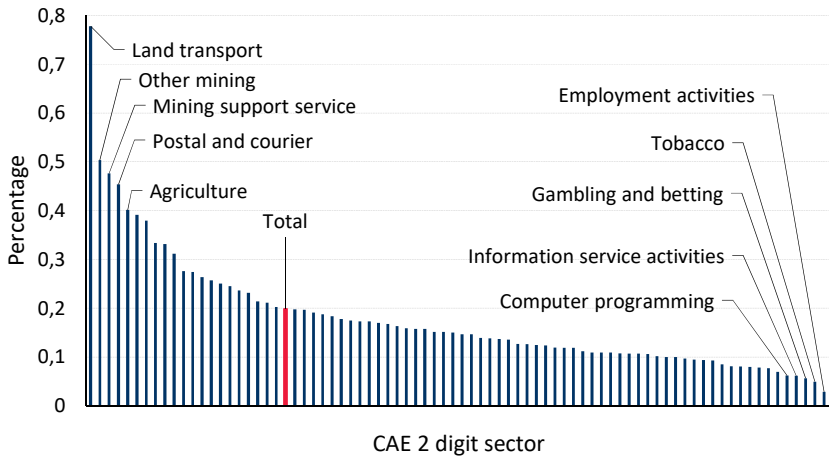


FIGURE 14: Average share of energy expenditure on total GVA across sectors in 2018

Moreover, the inclusion of firm fixed effects in regressions will capture other time invariant firms' characteristics that may affect their energy consumption patterns.

It could be argued that exploring variability across firms in each sector would be the easiest approach to establish a link between the ratio of energy expenditures on GVA and the size of firms. This would imply running regressions with sectoral fixed effects instead of firm fixed effects. However, the heterogeneity of activities within sectors is still quite large (e.g., in agriculture, greenhouse flowers differ from extensive olive oil

production) and other firms' characteristics (e.g. power and reliability of the electric grid where it is located) advise for considering firm fixed effects.

Table 1 presents estimated coefficients for the semi-elasticity between the ratio of energy expenditures on GVA and the logarithm of turnover in the period 2011-2018, which are consistently negative and statistically significant, meaning that larger firms post a lower energy expenditure for each euro of GVA generated, i.e., they are less energy intensive. Table 2 repeats the exercise above, taking the logarithm labour as the proxy for size and results are quite consistent, except for the class of large firms where the coefficient turns out non significant.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Total	Micro	Small	Medium	Large
ln turnover	-0.051***	-0.055***	-0.047***	-0.031***	-0.015***
	(0.001)	(0.001)	(0.002)	(0.004)	(0.005)
Constant	0.796***	0.837***	0.790***	0.601***	0.352***
	(0.007)	(0.007)	(0.028)	(0.063)	(0.091)
Observations	2,184,001	1,819,364	282,101	48,396	9,208
Adjusted R^2	0.507	0.493	0.691	0.672	0.689
Time FE	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 1. Energy intensity and size of firms measured by turnover

Notes: The dependent variable corresponds to the ratio of total energy expenditures (electricity plus liquid fuels) on firms' GVA. Significances are computed using robust clustered errors.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Total	Micro	Small	Medium	Large
ln nb. employees	-0.040***	-0.045***	-0.031***	-0.018***	-0.006
	(0.001)	(0.001)	(0.002)	(0.006)	(0.005)
Constant	0.242***	0.243***	0.221***	0.191***	0.128***
	(0.001)	(0.001)	(0.006)	(0.025)	(0.031)
Observations	2,184,001	1,819,364	282,101	48,396	9,208
Adjusted R^2	0.504	0.489	0.689	0.671	0.688
Time FE	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 2. Energy intensity and size of firms measured by employment

Notes: The dependent variable corresponds to the ratio of total energy expenditures (electricity plus liquid fuels) on firms' GVA. Significances are computed using robust clustered errors.

5.2. Energy mix of firms

Figure 15 plots the non-weighted kernel distribution of the share of electricity on total energy expenditure for all firms in the sample in 2018. This ratio proxies the energy

mix of firms. The distribution is bimodal, with a larger density of firms that use little electricity on their activity and on firms that use almost exclusively this energy item. This pattern results from firms technological decisions and the specific nature of their business. For example, even if electric vehicles are available, transportation firms spend almost exclusively on liquid fuels, while services firms spend almost exclusively on electricity. As expected, many firms combine both types of energy expenditures. The possibility of considering other types of energy expenditure or energy self-production in the denominator was abandoned due to the lack of data.

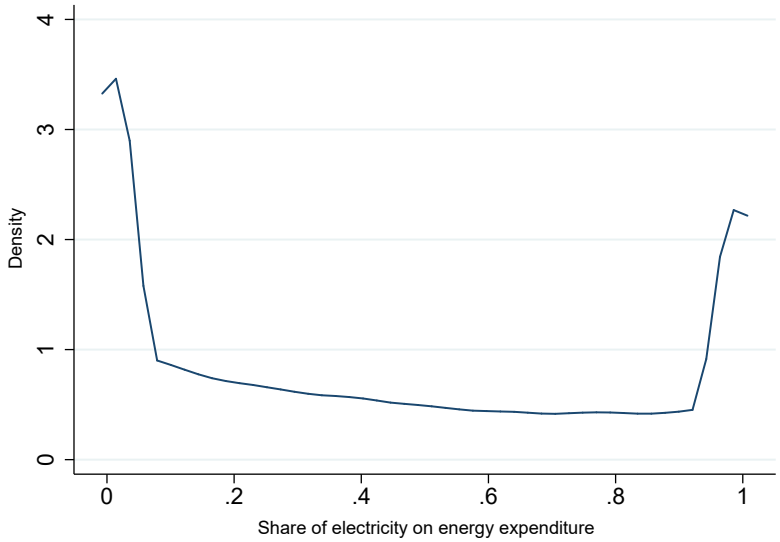


FIGURE 15: Share of electricity on total energy expenditure in 2018

Notes: Total energy expenditure comprises electricity and liquid fuels expenditure. Non-weighted observations.

In Figure 16, using a NACE 2-digit classification, sectors are ranked from the highest to the lowest in terms of the average of the ratio of electricity on total energy expenditure in 2018. Results show important differences across sectors, with "land transport" and "postal and courier" posting ratios lower than 7 per cent, in strike contrast with "food and beverages services", "electricity and gas" and "accommodation" with average ratios higher than 70 per cent. The detailed 2-digit level values for the mean, median and interquartile range are presented in Table A.1 in the Appendix.

At this point we turn to testing the association between the share of electricity expenditure on total energy expenditure (electricity and liquid fuels) and firm size, proxied both by the logarithm of turnover and employment, while controlling for time and firm specific effects. Table 3 reports the results of the regression, considering the overall sample of firms and subsamples for the four categories of firms in the period 2011-2018. Results are not as strong as those obtained for the association between energy intensity and size. Estimated coefficients for the semi-elasticities are negative and significant only for the overall sample and for the sets of micro and small firms.

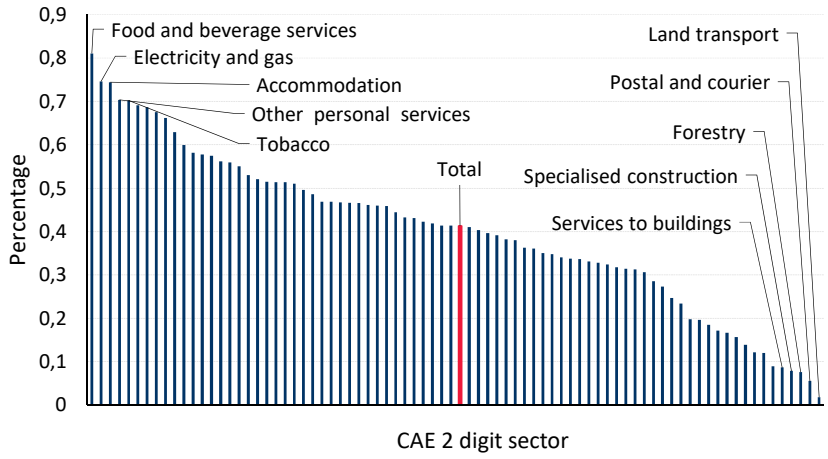


FIGURE 16: Average share of electricity on total energy expenditure across sectors

VARIABLES	(1) Total	(2) Micro	(3) Small	(4) Medium	(5) Large
ln turnover	-0.016***	-0.017***	-0.012***	-0.002	-0.008
	(0.000)	(0.000)	(0.001)	(0.003)	(0.007)
Constant	0.607***	0.607***	0.593***	0.539***	0.682***
	(0.004)	(0.005)	(0.017)	(0.052)	(0.121)
Observations	1,902,385	1,548,971	274,230	47,206	8,995
Adjusted R^2	0.860	0.856	0.907	0.924	0.925
Time FE	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 3. Share of electricity on total energy expenditure and size of firms measured by turnover

Notes: The dependent variable corresponds to the ratio of total energy expenditures (electricity plus liquid fuels) on firms' GVA. Significances are computed using robust clustered errors.

Table 4 repeats the previous exercise considering the logarithm of the number of employees as the indicator of size, instead of total turnover. Results are consistent with those obtained above. The coefficient of size, proxied by the logarithm of employment, in the overall sample is negative and significant and the same result exists for the subsample of micro firms. Nevertheless, coefficients for small, medium and large firms are not statistically different from zero.

VARIABLES	(1) Total	(2) Micro	(3) Small	(4) Medium	(5) Large
ln nb. employees	-0.010***	-0.012***	-0.002	0.007	0.014
	(0.000)	(0.001)	(0.002)	(0.004)	(0.009)
Constant	0.425***	0.416***	0.430***	0.475***	0.465***
	(0.001)	(0.001)	(0.004)	(0.017)	(0.051)
Observations	1,902,385	1,548,971	274,230	47,206	8,995
Adjusted R^2	0.860	0.855	0.907	0.924	0.925
Time FE	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 4. Share of electricity on total energy expenditure and size of firms measured by employment

Notes: The dependent variable corresponds to the ratio of total energy expenditures (electricity plus liquid fuels) on firms' GVA. Significances are computed using robust clustered errors.

5.3. Correlation between firms energy intensity and the share of electricity

As previously stated, lower energy intensity and a higher share of electricity in total energy consumption are desirable. High electrification would open the door to satisfy firms energy needs through renewable energy sources, while a lower energy intensity would prompt the economy to lower overall energy consumption. In this vein, we explicitly test the correlation between these two variables, while controlling for time and sector specific effects.

VARIABLES	(1) Total	(2) Micro	(3) Small	(4) Medium	(5) Large
Electricity share	-0.045***	-0.044***	-0.020***	-0.020***	-0.077***
	(0.001)	(0.001)	(0.002)	(0.004)	(0.009)
Constant	0.246***	0.265***	0.151***	0.127***	0.136***
	(0.000)	(0.001)	(0.001)	(0.002)	(0.006)
Observations	1,974,701	1,624,935	290,766	49,707	9,290
Adjusted R^2	0.134	0.126	0.289	0.235	0.240
Time FE	YES	YES	YES	YES	YES
Sector FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 5. Correlation between firms' energy intensity and the share of electricity in total energy expenditure

The dependent variable corresponds to the ratio of total energy expenditures on firms' GVA and the explanatory variable "electricity share" stands for the share of electricity on total energy expenditure. Significances are computed using robust clustered errors at the firm level.

Table 5 presents the coefficients of a regression where the independent variable is the share of electricity expenditures on total energy expenditure and the dependent variable is the ratio of energy expenditure on firm's GVA. The coefficients are negative and strongly significant for the overall sample and for the four different firm size classes

taken separately. This result signals that underperformance in these two dimensions goes hand in hand in most firms.

6. Final remarks

The challenges posed by climate change and the need to meet the calendars established for carbon neutrality are extremely ambitious. Therefore, the adjustment effort required by economic agents is very large. From an aggregate perspective, this involves increasing the share of renewables in total primary energy, as well as reducing the amount of energy consumed for each unit of value added created in the economy. We identify progress in Portugal and in the euro area in both dimensions. Nevertheless, the pace of transformation should be increased if targets are to be met.

Policy action at firm level is particularly challenging and needs to informed by empirical micro-based evidence. Although we do not have real energy consumption data for firms, energy expenditure data provides some initial insights. We document substantial heterogeneity at the sectoral and firm size dimensions. The goal of lower energy intensity seems to be facilitated by having larger firms, but this is not so clear for the purpose of increasing electrification. Nevertheless, underperformance in these two dimensions seems to go hand in hand in most firms.

There is large room to proceed with this research agenda, notably in the firm-level dimension and especially in the link between energy consumption patterns and firms' size. One natural extension is to test causality between firms' size and both their energy mix and energy intensity ratio, i.e., going beyond simple correlations, as tested in this article. In turn, establishing a link between energy intensity and labour productivity seems problematic as GVA is present in the denominator of both dependent and independent variables. However, a focus on total factor productivity may bear interesting results. Moreover, assessing the role of international trade and digitalization in connection to firms' energy consumption patterns are also interesting research questions. Finally, mapping the impact of energy consumption patterns on emissions is highly relevant in what concerns the climate agenda.

References

Alves, M., M. Rodríguez, and C. Roseta-Palma (2010). "Sectoral and regional impacts of the European Carbon Market in Portugal." Working Paper 21, GEE.

Amador, João (2010). "Energy Production and Consumption in Portugal: Stylized Facts." *Economic Bulletin Articles of Banco de Portugal*.

Filippini, Massimo and Lester Hunt (2011). "Energy Demand and Energy Efficiency in the OECD Countries: A Stochastic Demand Frontier Approach." *The Energy Journal*, 32, 59–80.

Geller, Howard, Philip Harrington, Arthur H. Rosenfeld, Satoshi Tanishima, and Fridtjof Unander (2006). "Polices for increasing energy efficiency: Thirty years of experience in OECD countries." *Energy Policy*, 34(5), 556–573.

IEA (2021a). "Renewables 2021: Analysis and forecasts to 2026." Report, International Energy Agency.

IEA (2021b). "World Energy Outlook." Report, International Energy Agency.

Kaivo-oja, Jari, Jarmo Vehmas, and Jyrki Luukkanen (2016). "Trend analysis of energy and climate policy environment: Comparative electricity production and consumption benchmark analyses of China, Euro area, European Union, and United States." *Renewable and Sustainable Energy Reviews*, 60, 464–474.

Lee, Jaeseok and Jongmin Yu (2019). "Heterogenous Energy Consumption Behavior by Firm Size: Evidence from Korean Environmental Regulations." *Sustainability*, 11-3226.

Malinauskaite, J., H. Jouhara, L. Ahmad, M. Milani, L. Montorsi, and M. Venturelli (2019). "Energy efficiency in industry: EU and national policies in Italy and the UK." *Energy*, 172, 255–269.

Nunes, Adélia (2018). "Energy changes in Portugal: An Overview of the Last Century." *Méditeranée*, 130-2018.

OECD (2020). "Environment at a Glance." Report, Organization for Economic Cooperation and Development.

Pereira, R. and A. Pereira (2019). "Financing a Renewable Energy Feed-in Tariff with a Tax on Carbon Dioxide Emissions: A Dynamic Multi-Sector General Equilibrium Analysis for Portugal." Working Paper 123, GEE.

Stehrer, R., A. Bykova, K. Jäger, O. Reiter, and M. Schwarzhappel (2019). "Industry level growth and productivity data with special focus on intangible assets." Statistical Report 8, wiiw.

The European Commission (2006). "Directive of the European Parliament and of the Council on energy end-use efficiency and energy services and repealing Council Directive 93/76/EEC." Tech. rep., Office of the United Nations High Commissioner for Human Rights.

The European Commission (2021). "Proposal for a Directive of the European Parliament and of the Council." Tech. rep., Office of the United Nations High Commissioner for Human Rights.

Zhang, Shanshan, Tommy Lundgren, and Wenchao Zhou (2016). "Energy efficiency in Swedish industry: A firm-level data envelopment analysis." *Energy Economics*, 55, 42–51.

Appendix

		Electricity share			Energy intensity		
Code	NACE 2-digit sector	Average	P50	IQR	Average	P50	IQR
1	Agriculture	0,34	0,28	0,54	0,40	0,18	0,36
2	Forestry	0,08	0,00	0,02	0,39	0,25	0,38
3	Fishing	0,09	0,00	0,03	0,38	0,19	0,30
7	Mining of metal ores	0,69	0,96	0,78	0,12	0,11	0,10
8	Other mining	0,31	0,22	0,48	0,50	0,34	0,47
9	Mining support service	0,17	0,00	0,38	0,48	0,09	0,62
10	Food	0,57	0,58	0,59	0,27	0,17	0,17
11	Beverages	0,47	0,47	0,49	0,18	0,08	0,11
12	Tobacco	0,70	0,70	0,12	0,05	0,04	0,05
13	Textiles	0,55	0,57	0,55	0,16	0,08	0,12
14	Wearing apparel	0,51	0,51	0,45	0,08	0,04	0,05
15	Leather and its products	0,52	0,57	0,41	0,08	0,04	0,05
16	Wood products, except furniture	0,43	0,43	0,45	0,20	0,10	0,14
17	Paper and its products	0,47	0,45	0,45	0,16	0,08	0,09
18	Printing and reproduction	0,51	0,51	0,46	0,11	0,07	0,07
19	Coke and refined petroleum	0,35	0,25	0,42	0,15	0,07	0,10
20	Manufacture of chemicals and its products	0,42	0,33	0,58	0,18	0,09	0,13
21	Pharmaceutical	0,46	0,51	0,71	0,08	0,04	0,06
22	Rubber and plastic	0,66	0,78	0,44	0,20	0,13	0,16
23	Other non-metallic mineral products	0,49	0,48	0,49	0,26	0,14	0,19
24	Basic metals	0,51	0,50	0,59	0,15	0,08	0,13
25	Metal products, except machinery and equipment	0,36	0,29	0,46	0,13	0,07	0,08
26	Computer, electronic and optical	0,47	0,42	0,76	0,11	0,04	0,07
27	Electrical equipment	0,40	0,34	0,46	0,12	0,06	0,08
28	Machinery and equipment n.e.c.	0,38	0,34	0,43	0,11	0,06	0,07
29	Motor vehicles, trailers and semi-trailers	0,58	0,60	0,51	0,11	0,05	0,06
30	Other transport equipment	0,46	0,42	0,60	0,10	0,05	0,08
31	Manufacture of furniture	0,50	0,50	0,49	0,15	0,09	0,09
32	Other manufacturing	0,47	0,42	0,54	0,10	0,05	0,07
33	Repair and installation of equipment	0,20	0,10	0,28	0,14	0,07	0,10
35	Electricity and gas	0,75	1,00	0,52	0,09	0,01	0,02
36	Water collection, treatment and supply	0,46	0,50	0,80	0,26	0,09	0,17
37	Sewerage	0,39	0,11	0,90	0,28	0,16	0,25
38	Waste collection, treatment and disposal	0,27	0,14	0,40	0,25	0,14	0,23
39	Remediation and other waste management	0,19	0,07	0,43	0,14	0,08	0,09
41	Construction of buildings	0,16	0,00	0,12	0,12	0,06	0,10
42	Civil engineering	0,12	0,02	0,09	0,19	0,09	0,17
43	Specialised construction	0,08	0,00	0,07	0,17	0,09	0,12
45	Trade and repair of motor vehicles and motorcycles	0,44	0,39	0,57	0,15	0,06	0,09
46	Wholesale trade, except vehicles and motorcycles	0,23	0,12	0,31	0,17	0,08	0,14
47	Retail trade, except vehicles and motorcycles	0,56	0,57	0,79	0,17	0,08	0,13
49	Land transport	0,02	0,00	0,00	0,78	0,49	0,74
50	Water transport	0,12	0,01	0,09	0,33	0,12	0,30

TABLE A.1. Distribution of the ratios: Electricity on total energy expenditure, and Energy expenditure on total GVA, for NACE 2-digit sectors, in 2018

Notes: p50 stands for the median and IQR stands for interquartile range.

		Electricity share			Energy intensity		
Code	NACE 2-digit sector	Average	P50	IQR	Average	P50	IQR
51	Air transport	0,34	0,08	0,65	0,24	0,01	0,18
52	Warehousing and support for transportation	0,33	0,17	0,60	0,19	0,04	0,13
53	Postal and courier	0,06	0,00	0,02	0,45	0,28	0,43
55	Accommodation	0,74	0,83	0,40	0,21	0,10	0,15
56	Food and beverage services	0,81	0,90	0,28	0,33	0,15	0,22
58	Publishing activities	0,38	0,26	0,69	0,08	0,02	0,06
59	Motion picture, video, TV and music	0,32	0,15	0,61	0,10	0,03	0,09
60	Programming and broadcasting	0,68	0,79	0,56	0,17	0,08	0,14
61	Telecommunications	0,20	0,00	0,17	0,15	0,07	0,15
62	Computer programming	0,32	0,17	0,53	0,06	0,01	0,05
63	Information service activities	0,41	0,26	0,92	0,06	0,01	0,05
68	Real estate activities	0,60	0,71	0,82	0,11	0,02	0,08
69	Legal and accounting activities	0,53	0,44	0,80	0,07	0,03	0,05
70	Head offices; management consultancy	0,35	0,17	0,73	0,08	0,01	0,06
71	Architectural and engineering activities	0,33	0,18	0,56	0,11	0,04	0,09
72	Scientific research and development	0,34	0,25	0,56	0,08	0,02	0,06
73	Advertising and market research	0,31	0,19	0,50	0,11	0,03	0,10
74	Other professional, scientific and technical act.	0,36	0,19	0,78	0,10	0,03	0,09
75	Veterinary activities	0,58	0,58	0,74	0,12	0,05	0,07
77	Rental and leasing activities	0,17	0,02	0,18	0,25	0,07	0,25
78	Employment activities	0,29	0,13	0,44	0,03	0,00	0,02
79	Travel agency and related activities	0,43	0,30	1,00	0,20	0,03	0,14
80	Security and investigation activities	0,14	0,07	0,14	0,12	0,06	0,11
81	Services to buildings	0,09	0,00	0,07	0,16	0,08	0,14
82	Office and other business support	0,31	0,12	0,59	0,14	0,04	0,12
85	Education	0,47	0,39	0,85	0,16	0,05	0,13
86	Human health activities	0,41	0,31	0,84	0,09	0,04	0,07
87	Residential care activities	0,63	0,65	0,46	0,09	0,06	0,06
88	Social work activities without accommodation	0,56	0,52	0,78	0,11	0,04	0,07
90	Creative, arts and entertainment activities	0,25	0,04	0,36	0,14	0,04	0,12
91	Cultural activities	0,42	0,30	0,82	0,13	0,06	0,11
92	Gambling and betting	0,69	0,84	0,63	0,06	0,03	0,05
93	Sports, amusement and recreation	0,41	0,23	0,92	0,31	0,11	0,29
94	Activities of membership organisations	0,51	0,55	0,21	0,23	0,08	0,18
95	Repair computers, personal and household goods	0,40	0,27	0,68	0,17	0,07	0,13
96	Other personal services	0,70	1,00	0,67	0,21	0,08	0,14
	All sectors	0,41	0,30	0,82	0,20	0,07	0,15

TABLE A.1. Distribution of the ratios: Electricity on total energy expenditure, and Energy expenditure on total GVA, for NACE 2-digit sectors, in 2018

Notes: p50 stands for the median and IQR stands for interquartile range.

Non-technical summary

January 2022

Characteristics of parties and duration of insolvency cases in Portugal

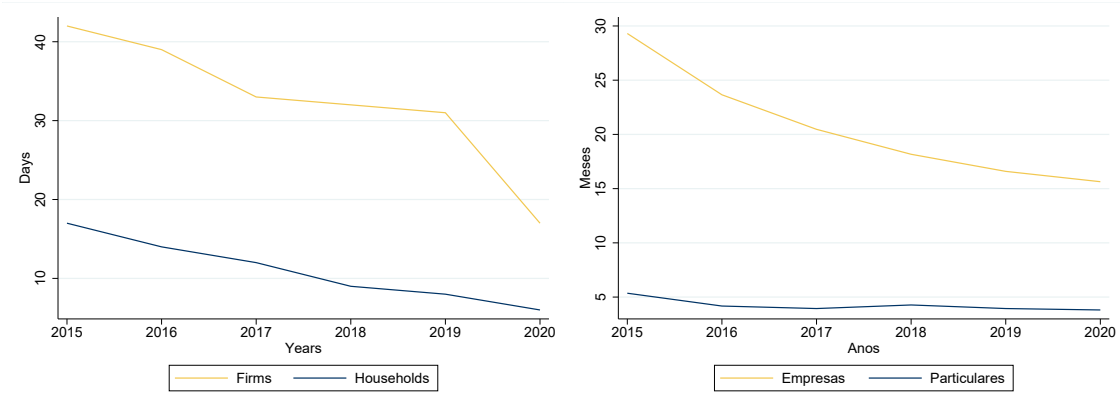
Manuel Coutinho Pereira and Lara Wemans

Duration is a central element in the efficiency of the insolvency procedure. Benefiting from detailed information on corporate and private insolvencies that were dealt with in the Portuguese judicial system between September 2014 and 2020, an analysis focused on the duration of these cases is presented.

Regarding insolvency, there are two main milestones. The first is insolvency declaration, which seizes the assets from the debtor for the benefit of the insolvency estate and suspends all the ongoing enforcement and seizures on the debtor. The second is case closure, which takes place, for corporations with the distribution of the insolvency estate to the creditors or with the approval of the insolvency plan. For households, case closure may coincide with the beginning of the period when the income above a minimum subsistence level is transferred to creditors.

A descriptive analysis shows that in the last years there was a clear reduction of duration to insolvency declaration by the judge (Figure 1A). Around half the cases from households presented in 2020 a duration below 6 days (17 days in 2015). As regards corporations, median duration is higher, also posting a clear reduction, to 17 days (40 days in 2015). The median duration to case closure also posted a reduction to 4 months for households and 16 months to corporations which compare with, respectively, 5 and 29 months in 2015 (Figure 1B).

The analysis of the case features which explain duration of insolvency may give important elements for fostering the discussion on the possible actions to increase the



(A) until declaration of insolvency (B) until case closure

FIGURE 1: Time estimated up to the resolution of half of the insolvency cases

Nota: Median from survival functions, estimated with complete information on the insolvencies that went through the system each year.

efficiency of this legal instrument. Both for private and corporate debtors, when it is the creditor requesting the insolvency, that increases duration to insolvency declaration and to case closure. This may be related to the necessary proceedings to identify the debtor and also to the level of cooperation. Moreover, the number of parties, namely creditors and, for private insolvencies, also debtors increases both duration measures. This effect would be related to the relationship between this variables and case complexity and coordination challenges. Other factor which increases duration, but only up to case closure, is the existence of private creditors as they may have a less «efficient» participation in the case and may need more support in decision making.

For companies, in the analysis of the duration to case closure, a wider set of characteristics related to their activity and financing was considered. The size of the company, as well as the amount of debt and assets to liquidate increase insolvency duration. The same happens with the existence of real collateral, which grants priority guaranteed credits (up to the amount the profit from the liquidation of guaranteed assets). Corporations in the construction sector have specially long-lasting insolvencies. Finally, insolvencies where the state holds the majority of the debts proceed more rapidly, as it is easier to align positions between public creditors in the creditors meeting.

Characteristics of parties and duration of insolvency cases in Portugal

Manuel Coutinho Pereira
Banco de Portugal

Lara Wemans
Banco de Portugal

January 2022

Abstract
Insolvency proceedings play an important role in shaping credit conditions and the allocation of productive resources. The paper focuses on duration analysis of insolvencies judged by Portuguese courts between September 2014 and 2020. The features which contribute to higher duration are the case not being filed by the debtor, a higher number of parties, both creditors and debtors, and the existence of households as creditors. For corporate insolvencies, firm's size, the amount of liabilities and fixed assets, activity in the construction sector and the existence of real collateral guaranteeing debts to the financial sector, also increase duration. (JEL: K40, H11, C41)

1. Introduction

Insolvency is the legal proceeding which aims at satisfying creditors, namely through the recovery of the company or, when that is not possible, through the liquidation of the assets of the insolvent (Insolvency and Corporate Recovery Code - CIRE, article 1). Less than 0.5% of companies and households in Portugal are subject to an insolvency request each year.[1] However, the effectiveness of these proceedings has an impact that goes much beyond those that face insolvency.

Firstly, this proceeding has a direct impact on a wide set of economic agents which includes the creditors of the insolvent and, in the case of companies, also their suppliers, clients and workers (Titman 1984). For the workers in firms that become insolvent, wage losses can be relevant and long-lasting (Graham *et al.* 2019).

Secondly, differences between legal jurisdictions regarding creditor protection, amounts recovered and costs have an effect on credit contractual conditions (Roberts and Sufi 2009). For companies, Davyddenko and Franks (2008) state that insolvency practices in the United Kingdom, France and Germany have very different durations

Acknowledgements: The authors thank the data collection and support for its use from BPLIM - Banco de Portugal Microdata Research Laboratory, and the valuable clarifications obtained from Direção-Geral da Política de Justiça. The authors are also grateful for the discussion with participants in a seminar at the Economics and Research Department and the comments made by António Antunes, Diana Bonfim, João Amador, Nuno Garoupa, Manuela Espadaneira Lourenço, Pedro Duarte Neves e Rui Conde Morais. The analyses, opinions and conclusions expressed herein are the sole responsibility of the authors and do not necessarily reflect the opinions of Banco de Portugal or the Eurosystem.

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

1. CEPEJ data indicates that the number of insolvencies per inhabitant in Portugal in close to the euro area average.

and recovery rates and that banks demand more collateral when legal provisions are less favourable to the creditor. Recent studies found evidence of an impact of insolvency effectiveness on credit access conditions for riskier companies (Rodano 2021) and on investment (Ponticelli and Alencar 2020), even among legal jurisdictions within the same country. Regarding households' insolvency, Albanesi and Nosal (2015) conclude that these costs have an impact on the decision to file a case and Antunes *et al.* (2019) argue that a change in these costs can have relevant macroeconomic effects on consumption.

Finally, there is evidence that corporate performance deteriorates much before liquidation (Almus 2004). Insolvency rules, as they have an effect on the timing companies exit the market, will also play an important role in resource reallocation. Osterhold and Gouveia (2020) study the survival of low productive companies in Portugal and conclude that a more efficient exit mechanism promotes the restructuring of viable firms.

McGowan and Andrews (2018) construct an indicator based on responses to a questionnaire on insolvency policies. The results for Portugal, referring to 2016, point to an intermediate position regarding the features of insolvency legislation promoting economic growth. Nevertheless, as studies from Rodano (2021) and Ponticelli and Alencar (2020) indicate, legislation is only one of the relevant features for the efficiency of this procedure as implementation can lead to significant differences, namely regarding duration.

Households' insolvency has grown in Portugal in the last decade, with the increase in cases filed during the financial crisis being only partially reverted in recent years. Nowadays, private debtors represent three in every four insolvency cases. A more efficient insolvency proceeding for households may contribute to a reduction in the costs of credit, especially for those with higher default risk, and to a swifter recovery of over-indebted households. Results point to a reduction of duration of these cases in the last years, though a quarter of them still takes at least 14 days up to the declaration of insolvency and six months up to closure.

Corporate insolvencies also increased significantly with the financial crisis, but returned to pre-crisis levels in 2016. Duration also registered a downward trend, but these cases are usually lengthy, with the 75th percentile for duration currently at two months up to insolvency declaration, and three and a half years until case closure.

This article analyses the duration of insolvency cases pending in Portuguese courts in September 2014 and those filed from there onwards up to the end of 2020, excluding cases concerning public and financial sector companies, not covered by CIRE. In fact, duration is, together with administrative costs, one of the most important characteristics of these procedures. Studies such as Morrison (2007) point to the existence of a strong relation between duration and administrative costs for the parties involved. Moreover, indirect costs from corporate insolvency should also increase with duration, through the retention of assets which could have been more swiftly reallocated (Bricongne *et al.* 2016).

Amongst the characteristics that delay both the declaration of insolvency and case closure, the following are noteworthy: case requested by the creditors (as opposed to being presented by the insolvent); number of parties and the participation of private

creditors. As regards corporate insolvency, several other characteristics delay case closure, such as the size of liabilities and fixed assets, the existence of real collateral guaranteeing debts to the financial sector and activity in the construction sector.

There are several studies focusing on duration analysis of corporate insolvency and reorganization.[2] Bris *et al.* (2006) find evidence of a relevant «judge effect» on insolvencies and reorganizations in the US and conclude that these procedures have a similar duration, but reorganizations yield higher recovery rates, especially to unsecured creditors. A higher number of creditors increases duration. Morrison (2007) focuses on restructuring in the US, estimating a competing risk model which studies duration to closure (frequently through insolvency) or restructuring. It concludes that uncertainty regarding the value of the company, proportion of cash holdings and leverage reduce case duration. Lastly, Kwon and Hahn (2010) apply a similar model to reorganizations of listed companies in South Korea and conclude that firm size increases duration.

This literature uses samples of, at most, some hundred cases, which compromises the precision of econometric estimates, but has the benefit of considering very exhaustive information about case characteristics. The innovation of this article lies on the focus on Portugal and on a very comprehensive fraction of all private and corporate insolvencies in the last years (around 85%), including around 100 thousand observations, which allows for a higher reliability of the estimates of insolvency duration. However, the database used does not include information on the debtor, when it is a household, and has few details on procedural aspects and on the way the insolvency ended.

This article is organized as follows. Section 2 briefly describes insolvency proceeding in Portugal. Section 3 presents the main characteristics of insolvencies, including the evolution cases filed and duration in the recent past. Section 4 presents the methodology and the main variables. Sections 5 and 6 discuss the determinants of duration to case closure and to insolvency declaration, respectively. Section 7 concludes.

2. Insolvency procedure in Portugal

Under Portuguese law, an entity is insolvent if it cannot fulfil overdue obligations or, in the case of a corporation, if it has liabilities clearly above its assets (CIRE, article 3). The case is filed in the jurisdictional unit (*comarca*) of residence or where the firm develops its main economic activity. Cases are judged in specialized courts (*tribunais de comércio*) whenever they exist.[3]

Figure 1 presents the main milestones of these proceedings, which start with a request made by the insolvent or by its creditors. When made by the creditors, debtors need to be notified, and can oppose. The second relevant milestone is the insolvency declaration. It implies the attachment of seizable assets and the suspension of enforcement and seizures

2. Under the US legal system, Chapter 7 is the procedure comparable to insolvencies in Portugal, while Chapter 11 reorganizations are dealt under *Processo Especial de Revitalização* which will not be analysed in this article.

3. Currently only 6 from 23 jurisdictional units do not have specialized courts.

on the debtor. Afterwards, an initial meeting of creditors may occur (this became optional in 2012). In this meeting, creditors appreciate the report from the insolvency practitioner. They also take a stand, in the case of companies, for the liquidation or the continuation of the firm and, in the case of households, for the discharge of debts or the approval of a payments plan (upon debtor's proposal). Votes at creditor meetings are set according to the proportion of credits.

Finally, case closure takes place, for companies, at the beginning of the application of the insolvency plan or the proportional distribution to the creditors of the proceeds from the sale. The insolvency plan must be approved by a two-thirds majority and certified by the judge. It may foresee firm's restructuring or liquidation, even under different conditions than set out by CIRE, as long as creditors are treated equally. The sale of the company can occur as a going concern, ensuring business continuation, or in parts, leading to liquidation.

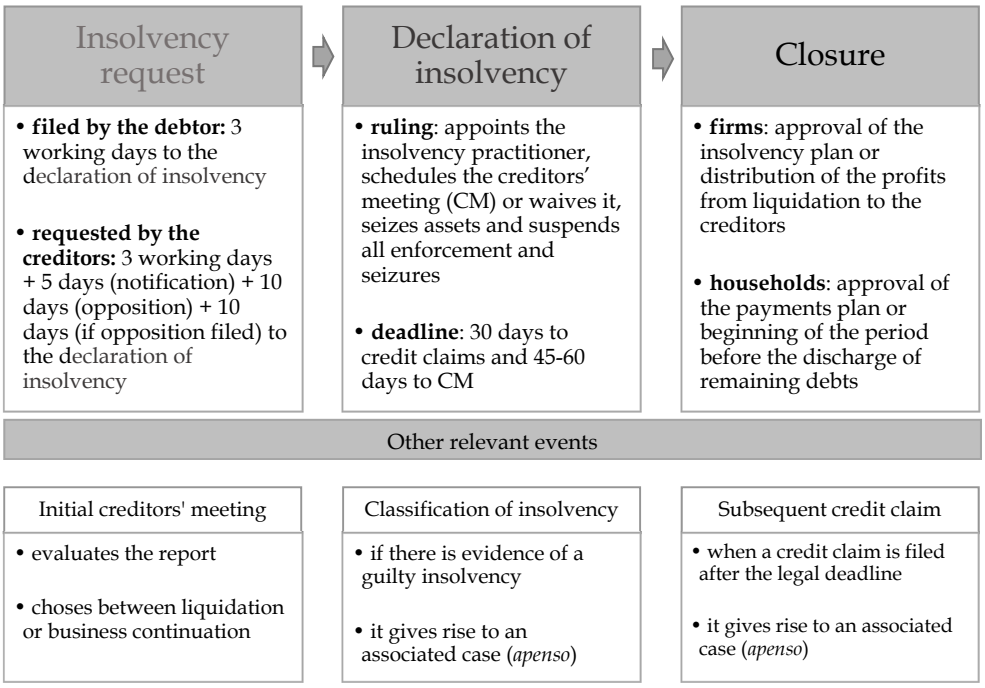


FIGURE 1: Milestones of the insolvency procedure - simplified presentation

For households, case closure can also take place for two reasons: i) the beginning of the payments plan, supported by a majority of two thirds at the creditors meeting and approved by the judge, involving debt restructuring and enabling the debtor to keep seizable assets or ii) the beginning of a period until the discharge of the remaining debts. In this latter option, creditors receive, in addition to the proceeds from the sale of seizable assets, the income of the insolvent above a certain minimum subsistence threshold during 5 years. After this period, the remainder debts are cancelled, with the exception of tax and contributory debts. Some assets are not seizable, such as goods vital to the household economy or of a reduced economic value, goods in co-ownership and a bank account balance up to the minimum wage.

Several procedural aspects (*apensos*) can be filed under insolvency cases, the most relevant being: i) classification of insolvency, which became non-mandatory in 2012 and evaluates whether there were acts which led to a deterioration of debtors' assets and also acts entailing simultaneously personal benefits to the insolvent or a third party; ii) subsequent credit claim, when a creditor seeks the recognition of debts after the initial deadline to do so (usually 30 days after the declaration of insolvency). These procedural aspects do not prevent case closure.

Except for such procedural aspects, judge's intervention is scarce, being focused on the insolvency declaration, the endorsement of the insolvency or payments plan and credit recognition and ranking. The insolvency practitioner has a central role in managing the case (Bernstein 2017). The compensation of insolvency practitioners includes a fixed component and a variable component related to the amount and proportion of recovered credits in the insolvency plan or asset sales. If the insolvency practitioners have to manage a company day-to-day business, they are also paid for that.

Since the approval of CIRE, in 2004, the most significant changes occurred in 2012, in the context of the Economic and Financial Assistance Programme (European Commission - DG-ECFIN 2014), and in 2017, with the creation of special procedures, respectively, for the recovery of firms (*Processo Especial de Revitalização*) and households (*Processo Especial para Acordo de Pagamento*). These procedures will not be analysed in this article. In 2012, insolvency was simplified through the resource to electronic communication and publication of procedural acts and the shortening of some legal deadlines. An example was the reduction of the deadline for holding the creditors' meeting from 75 to 60 days. In 2017, the use of electronic means was broadened, while the possibility of creditors to choose the insolvency practitioner was restricted to complex cases (Rodrigues *et al.* 2017 and Abreu Advogados 2017). In 2021, a ministerial order was passed regulating the direct electronic access of insolvency practitioners to the databases of tax administration, social security and commercial, vehicles and property register, with full implementation expected in February 2022. This access may reduce insolvency duration, similarly to what happened to enforcement cases (Pereira and Wemans 2018).

3. Insolvency description

3.1. General characteristics

The main dataset used was extracted from information published online[4] that allows the identification of the date, type and jurisdictional unit of the judicial acts in insolvency cases taken between September 2014 and December 2020[5]. This database also includes

4. www.citius.mj.pt/portal/consultas/ConsultasCire.aspx.

5. Including only cases with acts up to 31 December 2020 may lead to some underestimation of the number os cases entering in 2020, as cases may be registered with some delay.

several characteristics of insolvencies, such as the date of the request, debtors, creditors and those filing the request. The system for exchange of reference information (*Sistema de Partilha de Informação de Referência* - SPAI) was used to identify the sector of activity of creditors and both the Central Balance Sheet Database (*Central de Balanços* - CB) and the Central Credit Register (*Central de Responsabilidades de Crédito* - CRC) to obtain some characteristics of the debtors if they were corporations.[6] Cases with errors in the identification of relevant dates were discarded as well as 105 cases (0.1% of the sample) which started before September 2004, when CIRE came into force.

Private insolvency cases are almost always requested by the sole insolvent and are dealt with in specialized courts (both 90%). On average, there are 5.2 creditors, 2.1 financial institutions, 0.9 non-financial corporations, 0.3 public - mainly tax authority and social security - and 0.2 households.[7] There are still, on average, 1.7 unidentified creditors. Around a quarter of cases has more than one debtor (Table A1 in appendix).

The vast majority of corporate insolvency cases are also dealt with in specialized courts (92%), but the percentage of cases filed by the debtor is much lower (50%) and such cases include, on average, a higher number of creditors (18, of which 5.8 unidentified). Among the identified creditors, on average 7 are non-financial corporations, for instance suppliers and 3.5 are households, which may be workers. Besides those, there is an average of 1.3 financial corporations and 0.7 public sector creditors. Only a very small fraction of cases has more than one debtor (0.5%).

Regarding the cases successfully matched with CB and CRC, 76% are micro firms, in 10% the public sector holds the majority of the debts and in 34% the firm is not active in the year of the insolvency request.[8] Taking into account CRC information, 22% have loans with real collateral, 10% with financial collateral and 50% with other collateral types, such as personal guarantees. Compared to all companies covered by CB, those facing insolvency are on average larger, both in number of workers and in total liabilities. Regarding economic sectors, companies which requested insolvency are concentrated in the trade, food and accommodation (37%), manufacturing (22%) and construction (16%) sectors.[9]

Figure 2A depicts the evolution of cases filed and resolved, which have a very similar behaviour, as the duration up to insolvency declaration is, in general, very short. Official statistics only provide the split between private and corporate insolvencies for resolved cases, shown in Figure 2B but for the abovementioned reasons, this would be similar to that of filed cases. The number of insolvency requests in Portugal grew significantly

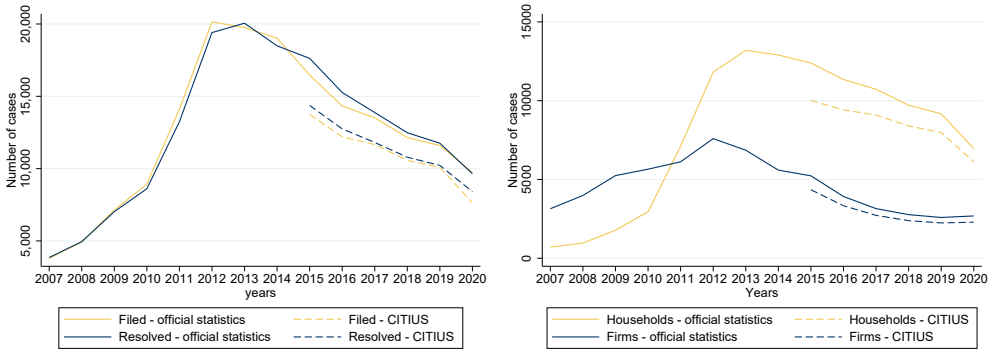
6. Data is matched in the closer year to the insolvency request going back at most 4 years. For simplification, January information from CRC was considered. There was no available information from CB or CRC for 2020 yet and information from CSI starts in 2006. For confidentiality reasons, it is not possible to match the data for households in CRC.

7. Financial sector corresponds on the Portuguese Classification of Economic Activities (CAE) to sector K and public sector to CAE O.

8. Besides firms classified under suspension or ceasure of activity in CB, we also considered as inactive firms those that did not file the CB in the year of the insolvency request nor in the two previous years.

9. Trade, food and accommodation corresponds to CAE G and I, manufacturing to CAE C, D and E and construction to CAE F.

during the financial crisis. A downward trend started in 2014 that led to the return to pre-crisis levels in the case of firms, while for households the recovery was incomplete. The year of the outbreak of the pandemic crisis saw a sharp reduction in private insolvencies that may be related to supply bottlenecks, taking into account the restrictions to the operation of services, and also to demand issues, in a context of reduced mobility, proliferation of credit under moratorium and the suspension of tax and contributory enforcement actions. Figures 2A and 2B also show that the database constructed from information published on CITIUS covers an important fraction of all insolvency cases (around 85%).



(A) resolved and filed cases (B) resolved cases, by type of debtor

FIGURE 2: Evolution of the number of insolvencies in Portugal

Sources: DGPJ and CITIUS (authors' calculations).
Note: Official data exclude transferred cases. Data collected from CITIUS may underestimate filed cases in 2020 (see section 3).

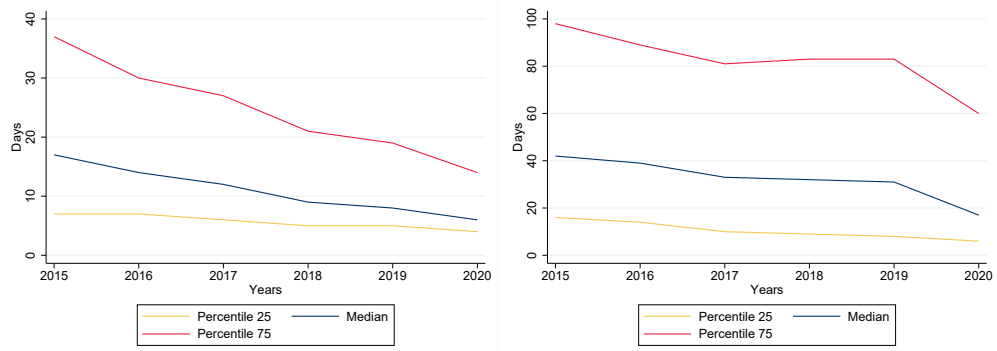
3.2. Duration

The duration of insolvency cases can be analysed according to two approaches. The first, used in official statistics, classifies a case as resolved when insolvency is declared or when the insolvency request is denied. This article considers as duration up to insolvency declaration the time span between the case being filed and one of these two events. For simplicity, taking into account that only around 0.5% of requests in our database are denied[10], the term duration until declaration also covers such cases. The second approach, perhaps of higher economic interest, is centred on the duration between the insolvency request and case closure, identified by the closing date, the publication of the initial decision of the discharge of the remaining liabilities or the approval of a payments or insolvency plan.

As discussed in Pereira and Wemans (2018), duration of resolved cases is not the most informative measure of case duration in a given year, as many cases resolved in that year have entered the system in previous years. Therefore, this indicator may

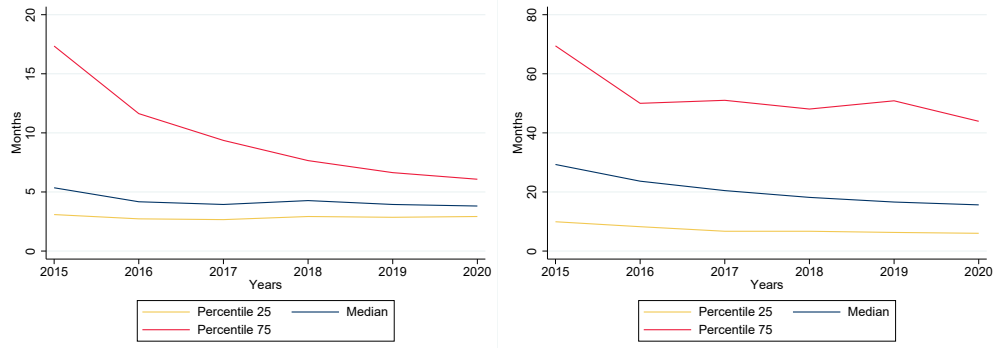
10. If there would be a tendency by judges to do not register procedural acts in CITIUS when insolvencies get denied, this percentage would not be representative of the universe of insolvencies.

increase (decrease) if there is a focus on the resolution of older (more recent) cases. As an alternative, duration analysis considering all cases that went through the system year-by-year allows for the estimation of the time until 25%, 50% and 75% of cases being resolved (Figures 3A to 3D).



(A) until declaration - households

(B) until declaration - corporations



(C) until closure - households

(D) until closure - corporations

FIGURE 3: Duration of insolvencies in Portugal between 2015 and 2020

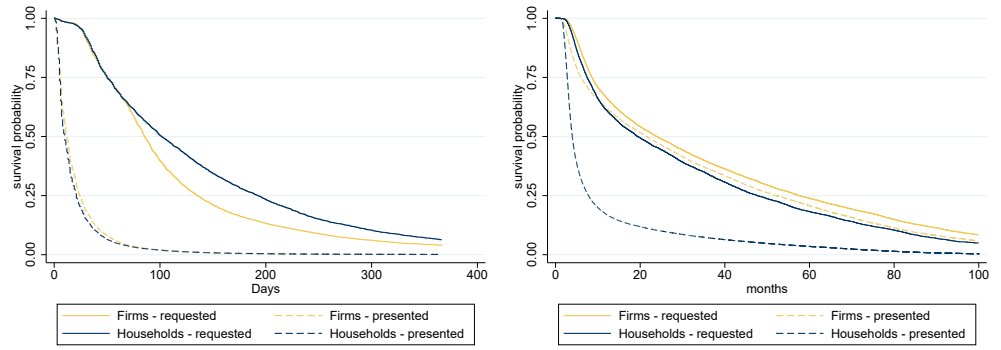
Note: Percentiles from survival functions, estimated with complete information on the insolvencies that went through the system each year.

Between 2015 and 2020, the estimates of the time span until insolvency is declared for half of the cases declined, in the case of households, from 17 to 6 days and, for corporations, from 40 to 17 days. Regarding duration until case closure, the median for households has stabilized around 4 months (in contrast with the 75th percentile that posted a significant reduction). In the case of corporations, case length was reduced from 29 to 16 months. This evolution may be associated to improving economic activity and to the increase in the use of swifter instruments to identify and sell assets by insolvent practitioners.

The survival function indicates the estimated probability of cases to remain open (on the y-axis) as a function of time since the request was filed (on the x-axis). As regards duration until insolvency declaration, whether the requests is made by the debtor or by creditors should be an important determinant of duration. This because, as posted in Figure 1, in the first case the law establishes a deadline of 3 working days for the issuance of the declaration of insolvency, while in the second case this period is increased

by up to 25 days. Figures 4A and 4B confirm a much higher duration for cases requested by creditors. In fact, the median duration until the issuance of the declaration in cases presented by the insolvent is 10 days, higher than established by law but still much lower than the 90 days for the rest of the cases.

Regarding duration until case closure, only cases presented by the debtor when this is a household are clearly swifter than the rest. The reasons for household insolvency to be quicker should be linked to its lower complexity, while the difference between cases filed by the debtor and by others may be associated with the cooperation of the insolvent in the case, which may be more important as regards private insolvencies.



(A) until declaration (B) until closure

FIGURE 4: Survival functions by insolvent and according to presentation by the debtor or requested by the creditors

Note: Figures depict Kaplan-Meier survival functions until the 99th percentile of duration for the complete sample.

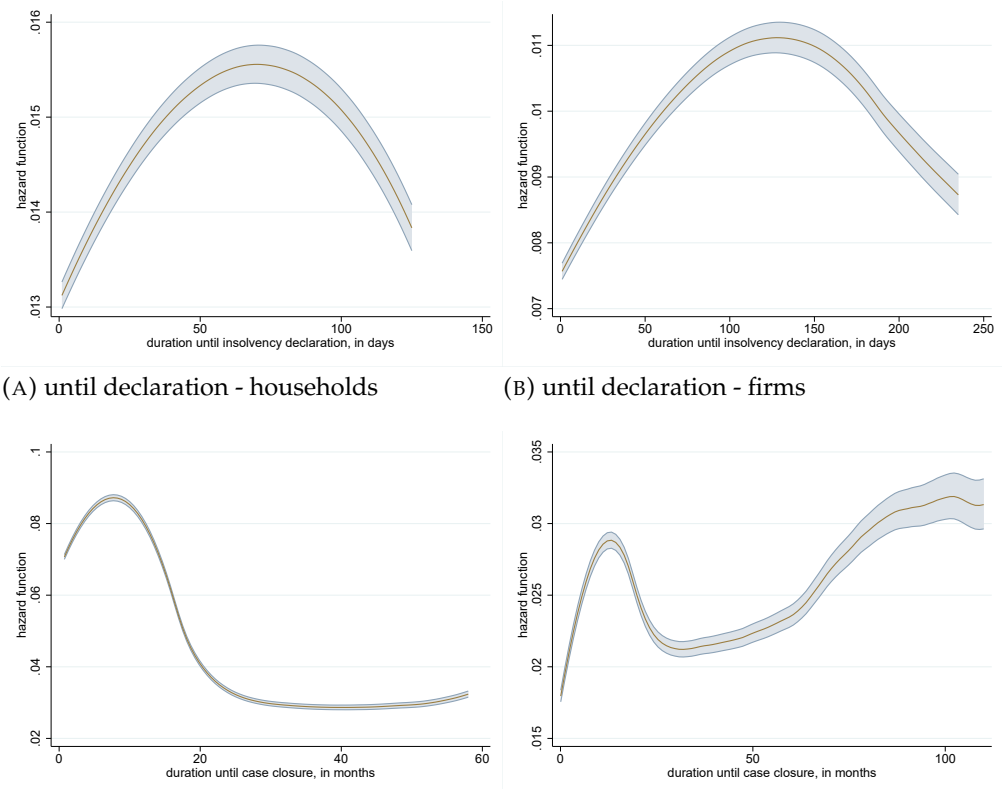
Furthermore, there is a significant variability in duration among different jurisdictional units, without a clear connection with size. For instance, median duration until closure in the quickest jurisdictional unit is of 3 months for households and 8 months for firms, which compares to 14 and 57 months, respectively, in the slowest one. This may be associated to differences in the efficiency of the proceeding or in average case complexity among jurisdictional units. It is important to consider that the majority of the insolvency practitioners works in a significant number of jurisdictional units, which makes significant efficiency variability less likely.

3.3. Time profile of insolvency declaration and case closure

Along with the survival functions, the distribution of duration can be studied taking into account hazard functions, which measure the instantaneous rate of case resolution, reported to a given moment after case start, conditioning to the case being still pending.

Figures 5A and 5B depict these functions, respectively, for households and firms until insolvency declaration. They post an increasing resolution rate up to a maximum that is reached more rapidly for households than for firms. Figure 5C posts the same function but for the duration to case closure of private insolvencies indicating that the rate of resolution of pending cases is higher in an initial phase, declining up to a duration

of 25 months and remaining more less constant after that. In the case of corporate insolvencies, the hazard function for the duration to case closure (Figure 5D) has, after an initial increasing phase, a U shape - decreasing and subsequently slightly increasing.



(A) until declaration - households (B) until declaration - firms

(C) until closure - households (D) until closure - firms

FIGURE 5: Hazard functions by type of insolvent

Note: Figures depict hazard functions for durations up to the 95th percentile.

4. Econometric methodology and explanatory variables

The impact of explanatory variables on duration is studied by using the Cox (1972) model, which assumes that each regressor shifts the baseline hazard function in a multiplicative way, i.e. there is proportionality of hazards. Thus, the hazard function of an insolvency case associated with the explanatory variables $\mathbf{x}_i$, $h(t|\mathbf{x}_i)$, is given by $h(t|\mathbf{x}_i) = h_0(t)\exp(\mathbf{x}_i\boldsymbol{\beta})$, where $h_0(t)$ is the baseline hazard function and $\exp(\mathbf{x}_i\boldsymbol{\beta})$ the relative hazard. There is a proportionality relationship between the hazard functions of any two cases (associated with $\mathbf{x}_j$ and $\mathbf{x}_i$), given by $\exp(\mathbf{x}_j\boldsymbol{\beta})/\exp(\mathbf{x}_i\boldsymbol{\beta})$. The Cox model has a semiparametric nature, as it does not require the formalization of the baseline hazard function.

The proportionality assumption can be tested for the regression as a whole and for specific regressors. It can also be dropped for one or more categorical variables through a stratified estimation procedure, assuming that the baseline hazard functions

are differentiated within strata defined by the values of these variables, while the coefficients of the remaining explanatory variables are common to all strata. It is also possible to estimate differentiated impacts of regressors by segments of overall duration, thus restricting the proportionality assumption to such segments. Both procedures have been used in this article, as explained below. Furthermore, parametric models have been estimated in a robustness section, which require the specification of the distribution associated with the duration model, without imposing proportional hazards.

The explanatory variables in the econometric analysis of duration to closure capture some general features of cases, common to private and corporate insolvencies. For the latter insolvencies, a set of variables concerning firms' activity and financing has been added (see Section 3.1 for more details on the data).

Variables relating to the case comprise (i) whether the case has been filed by the insolvent or by creditors, (ii) number of debtors, in the case of private insolvencies, (iii) number of creditors and (iv) type of creditors involved (financial, public administration, business and private). The variables specific to corporate insolvencies include (i) sector of activity, (ii) existence of real, financial or other collateral, guaranteeing debts to the financial sector (iii) a size indicator, (iv) value of fixed assets, (v) value of liabilities, (vi) if the majority of liabilities are with the public administration and (vii) if the company is no longer active in the year of the insolvency request. The value of fixed assets approximates the size of assets that generally will have to be liquidated in the course of the proceedings. The value of liabilities captures the amount of claims to be satisfied, complementing information on the number and type of creditors. Fixed assets and liabilities were taken in logarithms. Cessation of activity measures the extent to which firm's situation has deteriorated at the start of the insolvency case.

The economic cycle may also have an influence on case duration, namely through the volume of incoming cases and the pace of liquidation of the insolvent's assets. Thus, an economic activity indicator over the lifetime of the case has been derived as the average variation of the coincident indicator of Banco de Portugal (with monthly frequency). Ideally, the regression should also include variables capturing procedural events giving rise to *apensos* (see Section 2), in a specification allowing time-variation of such variables. However, it has not been possible to identify such events for all cases in the database. Nevertheless, these *apensos* do not prevent the closure of insolvencies and should have a less significant impact on duration than in other types of cases, such as enforcement ones. The variables referring to the type of creditor have an important number of missing observations, which were imputed through a multiple imputation procedure.[11]

In the Cox models estimated in this article, the jurisdictional unit where the insolvency case has been dealt with has been used to stratify the sample, in order to control for factors specific to that unit not captured by the regressors considered, for example, the degree of congestion or the average complexity of insolvency proceedings.

11. This procedure is based on «chained» logistic regressions in which the regressors include the other variables to be imputed, the other regressors in the main regression and, as additional information, the proportion of credits belonging to the financial sector and the public sector, as well as the failure indicator and the baseline cumulative hazard (White and Royston 2009).

Note that this is preferable to the inclusion of fixed effects, which assumes the proportionality of the hazard functions across all jurisdictional units, an assumption that has been tested and is violated in the data. Even after stratification by jurisdictional unit, the proportionality assumption does not hold for some regressors for which it is important to assess impacts on duration. However, such an assessment would not be possible if these regressors were modelled as additional stratification variables. Thus, a complementary specification was estimated, with variable coefficients along three segments of case duration.[12] This specification is more flexible, as proportionality of risks is imposed within each segment only.

Duration to case closure is the one of greatest economic interest. However, duration to the insolvency declaration by the judge, the first stage of the case, has been studied as well. In this second analysis, only the case-related explanatory variables have been considered.

5. Duration to closure of insolvency proceedings

Tables 1 and 2 present the results of the estimation of a Cox model for private and corporate insolvencies, respectively, both assuming proportional hazards for the whole duration and restricting it to segments of duration. Estimation results (shown as the exponentials of coefficients) indicate the shift in the baseline hazard function stemming from the change in regressors, i.e. the impact on the (probabilistic) rate of resolution of pending cases. Therefore, when the exponential of a coefficient is equal to 1, the regressor has no impact on the resolution rate. If the exponential is lower than 1, say 0.9, such rate is reduced by 10%, extending duration. If the exponential is higher than 1, say 1.1, the rate is increased by 10%, shortening duration.

5.1. Determinants of duration relating to the case

The filing of the insolvency case by the insolvent (rather than by a creditor) shortens duration strongly, especially in the first duration tertile, and in private insolvencies. This is due, from the outset, to the absence of the need to notify the debtor at the beginning of the proceedings, as well as of a possible opposition to insolvency. Furthermore, the cooperation of the insolvents in the identification of the assets and, more generally, the fact that they consider the case to be in their own interest tend to shorten duration.

Private insolvencies involving more than one debtor tend to proceed more slowly; coefficients by duration segments are always significant for this regressor, but they are close to each other. The involvement of a larger number of creditors tends to prolong the duration of insolvencies, both when the debtors are individuals and firms, which may be attributable to a greater difficulty in reconciling the interests of the parties, for example, at creditors' meetings. For private insolvencies, the estimated coefficients by

12. The segments were calculated according to the 33rd and 66th percentiles of the survival functions, separately for private insolvencies (3rd and 7th months) and corporate insolvencies (11th and 41st months).

	Proportionality full duration	Proportionality by duration segments up to 3 months	3 to 7 months	over 7 months
Requested by debtor (creditor)	2.53***	15.54***	4.38***	1.64***
	0.04	*1.40*	*0.14*	*0.03*
Several debtors	0.89***	0.88***	0.88***	0.91***
	0.01	*0.02*	*0.02*	*0.01*
Number of creditors (less than 4)				
4 or 5 creditors	0.89***	0.91***	0.90***	0.91***
	0.01	*0.02*	*0.02*	*0.02*
over 6 credores	0.81***	0.85***	0.80***	0.84***
	0.01	*0.02*	*0.02*	*0.02*
Type of creditor				
Financial	1.07***	1.09**	1.03	1.08**
	0.02	*0.05*	*0.03*	*0.04*
Public administration	1.08***	1.04*	1.06***	1.11***
	0.01	*0.02*	*0.02*	*0.02*
Business	1.07***	1.02	1.06***	1.09***
	0.01	*0.02*	*0.02*	*0.02*
Private	0.89***	0.88***	0.86***	0.90***
	0.01	*0.02*	*0.02*	*0.02*
Economic activity (non-interacted)	1.03***		1.03***	
	0.00		*0.00*	
Nº observations	71,800		71,800	

TABLE 1. Determinants of duration to closure, private insolvencies

Notes: Table shows exponentials of coefficients, i.e. hazard ratios, estimated by the Cox regression, stratifying by jurisdictional unit; omitted groups in parenthesis; standard-errors (in italics) adjusted to account for the variability stemming from the multiple imputation of the creditor type variables; p-values: * <0.1; ** <0.05; *** <0.01.

duration tertiles are again significant and quite close to each other, while for firms there is only statistical significance for the short to the intermediate durations.

The presence of public administration creditors accelerates the course of insolvencies, particularly when the debtors are firms. This effect is attributable to two factors. On the one hand, debts to the State (mostly to tax authorities and social security) have a very standardized nature, similar across insolvencies, which facilitates the tasks of the creditor and the insolvency practitioner. Furthermore, given the privileged status of the State vis-à-vis some of the other creditors, it is possible that the debts to the State may, in some cases, absorb the entire insolvency assets, simplifying the proceedings. The binary variable that captures the cases in which the public sector is the majority creditor (in the case of corporate insolvencies) indicates an additional acceleration effect. This can be explained by the ease of position alignment among public creditors at creditors' meetings, where some important decisions are made by majority.

In private insolvencies, there is also a speeding-up effect when financial and business creditors are present. In fact, financial institutions and larger firms will be in a position to carry out a professional follow-up of insolvencies, and their presence may, in this way, speed up the proceedings. However, such an effect is almost absent in corporate insolvencies. This may reflect the fact that the variable is also capturing other characteristics of cases (not included in the model) associated with financial and business creditors, which give rise to an increased duration. Insolvency proceedings that

	Proportionality full duration	Proportionality by duration segments up to 11 months	11 to 41 months	over 41 months
Requested by debtor (creditor)	1.16***	1.50***	1.03	1.11***
	0.02	*0.04*	*0.03*	*0.03*
Number of creditors (less than 6)				
6 to 13 creditors	0.94*	0.85***	0.94	1.08
	0.03	*0.04*	*0.05*	*0.06*
over 13 creditors	0.76***	0.57***	0.72***	0.93
	0.03	*0.04*	*0.04*	*0.06*
Type of creditor				
Financial	0.96	0.91**	1.02	0.99
	0.03	*0.04*	*0.05*	*0.05*
Public administration	1.40***	1.29***	1.36***	1.43***
	0.05	*0.07*	*0.06*	*0.07*
Business	1.06	0.95	1.11*	1.24***
	0.04	*0.05*	*0.07*	*0.10*
Private	0.94*	0.83***	0.95	1.03
	0.03	*0.04*	*0.04*	*0.05*
Public creditors in majority	1.11***	1.10**	1.22***	1.06
	0.03	*0.05*	*0.05*	*0.05*
Cessation of activity	1.10***	1.12***	1.13***	1.05*
	0.02	*0.03*	*0.03*	*0.03*
Liabilities	0.95***	0.95***	0.94***	0.94***
	0.00	*0.01*	*0.01*	*0.01*
Fixed assets	0.97***	0.96***	0.98***	0.99***
	0.00	*0.00*	*0.00*	*0.00*
Microenterprise	1.11***	1.16***	1.11***	1.16***
	0.02	*0.06*	*0.04*	*0.04*
Type of collateral				
Real collateral	0.79***	0.76***	0.71***	0.83***
	0.02	*0.04*	*0.03*	*0.03*
Financial collateral	0.95*	1.08	0.88***	0.91**
	0.03	*0.07*	*0.04*	*0.03*
Other collateral	1.08***	0.97	1.00	1.26***
	0.02	*0.03*	*0.03*	*0.04*
Debtor's activ. sector (manuf. ind.)				
Agriculture and mining	0.98	1.12	0.81*	1.04
	0.07	*0.13*	*0.10*	*0.12*
Construction	0.78***	0.75***	0.71***	0.85***
	0.02	*0.04*	*0.03*	*0.03*
Trade, food and accommodation	1.05**	1.04	1.05	1.02
	0.02	*0.04*	*0.04*	*0.03*
Other services	0.98	1.00	0.90***	1.01
	0.02	*0.05*	*0.04*	*0.04*
Economic activity	*1.23***		1.22***	
(non-interacted)	*0.01*		*0.01*	
Nº observations	24,542		24,542	

TABLE 2. Determinants of duration to closure, corporate insolvencies

Notes: Table shows exponentials of coefficients, i.e. hazard ratios, estimated by the Cox regression, stratifying by jurisdictional unit; omitted groups in parenthesis; standard-errors (in italics) adjusted to account for the variability stemming from the multiple imputation of the creditor type variables; p-values: * <0.1; ** <0.05; *** <0.01.

include private creditors tend to proceed more slowly, perhaps because they will be less familiar with the procedural details.

The duration of insolvencies tends to decrease in response to the expansion of economic activity, especially when firms appear as debtors. This may reflect a particular sensitivity to the business cycle of sales of insolvent companies' assets, either as a going concern or in parts.

5.2. *Determinants of duration specific to firms*

The insolvent's sector of activity - vis-à-vis insolvent companies in manufacturing industry - has a clear impact on duration for the construction sector only, where proceedings tend to last longer, perhaps reflecting a composition of insolvency assets which makes their sale more difficult. A larger volume of fixed assets lengthens the proceedings, probably to the extent that this entails a more cumbersome process of liquidation. In addition, this variable is an approximation to the claim value and may capture procedural aspects that depend on this indicator.[13] The overall liabilities measure the size of the claims involved in the case, having a positive impact on duration.

Microenterprise insolvencies tend to be quicker, even controlling for variables such as the value of assets and liabilities, reflecting additional features associated with the size of the firms that simplify the proceedings. Furthermore, the insolvencies of firms that have ceased activity tend to be shorter. In this case, insolvency practitioners will have their tasks simplified, as they do not have to deal with management issues and the liquidation of the firm itself may be easier.

The impacts of the explanatory variables relating to corporate insolvencies analysed so far are statistically significant along the duration tertiles, but, at the same time, there is no marked variation across them.

The existence of real collateral guaranteeing debts to financial institutions has a clear impact of slowing down the proceedings. Credits that benefit from real guarantees have priority over the other credits in the insolvency, up to the value of the assets given as a guarantee. Thus, in such cases, one will have to wait for the sale of these assets, in order to calculate the remaining debt that will compete with common credits (i.e. that are neither guaranteed nor have a privileged nature[14]). The existence of financial collateral tends to increase duration as well, but the impact is smaller and only visible for intermediate to long durations. In turn, the existence of personal guarantees, captured by the other collateral, tends to accelerate insolvencies, but the impact is confined to the last tertile. It should be noted, however, that these last two types of collateral do not have a statistically significant impact on duration, in the parametric model estimated in the next section.

5.3. *Robustness analysis*

One now carries out a robustness analysis, by estimating parametric models that assume a probabilistic distribution for duration time. The generalized gamma distribution was chosen because it approximates well the shapes, respectively, decreasing and approximately bathtub of the hazard functions presented in Graphs 5C and 5D.[15]

13. In insolvency proceedings, the claim value is measured by the amount of the assets.

14. Examples of credits with a privileged nature include debts to workers and to the State.

15. This family of distributions has other frequently used parametric models as particular cases, such as the exponential, Weibull and lognormal. The constraints associated with these models were tested and rejected against the more general model.

Table 3 presents the results of the estimation of the specifications above, for private and corporate insolvencies, based on the generalized gamma distribution. This model is estimated in the accelerated failure-time metric, and does not have a proportional hazards representation, such as the one underlying the Cox model. The estimation results (presented as the exponentials of coefficients) capture the multiplicative effect of regressors on the time to case closure, in terms of shortening or extending it. Thus, when the exponential of the coefficient is equal to 1, the regressor has no impact on duration to closure. If the coefficient exponential is less than 1, say 0.9, duration is reduced by 10%. If the exponential is greater than 1, say 1.1, duration is increased by 10%. Note that the interpretation of coefficients in this type of models is the opposite of the interpretation in the proportional hazards model, in which coefficients smaller (larger) than 1 mean an extension (shortening) of time to resolution.

Although the different metrics underlying the Cox model and the generalized gamma model do not allow comparing coefficient values, it is possible to compare the respective sign and statistical significance, as well as the relative sizes among regressors. The results in the two methodologies are, in general, very aligned.

As in the Cox model, the fact that the case is filed by the insolvent tends to shorten duration, while the intervention of more than one debtor (for private insolvencies) and of a larger number of creditors tends to prolong it. For private insolvencies, the impact on duration of the variables related to the type of creditor is more mitigated in the generalized gamma model. Indeed, in this case only the presence of private and financial creditors has a clear statistical significance, respectively delaying and accelerating the course of proceedings. In corporate insolvencies, the evidence for these regressors is entirely consistent across models, with an impact on duration being confined to the presence of public administration creditors, which speeds up proceedings.

The effects of both the business cycle throughout case lifetime and the firm-specific regressors are very much aligned between the semi-parametric and the parametric models. The extension of duration that comes from the insolvent's belonging to the construction sector and from the existence of real collateral guaranteeing debts to the financial sector stand out, as above, for its magnitude. In the parametric model, however, the existence neither of financial collateral nor of other collateral appear to have an impact on duration.

6. Duration to declaration of insolvency

In this section, one carries out an analysis of duration to declaration of insolvency. This period until the declaration of insolvency is the initial stage of the case, when it is still completely under the jurisdiction of a judge, prior to the appointment of the insolvency practitioner (see Section 2). Table 4 presents the impact of the variables relating to the case on duration according to both the Cox model and a parametric model based on the logistic distribution (estimated in the accelerated failure-time metric). The choice of this distribution is justified by the arc-shaped hazard function - see Graphs 5A and 5B. As before, we estimate a specification of the Cox model allowing time-varying coefficients,

	Private insolvencies	Corporate insolvencies
Requested by debtor (creditor)	0.50***	0.87***
	0.00	*0.01*
Several debtors	1.02***	
	0.01	
Number of creditors		
4 - 5 (priv.) / 6 - 13 (corp.)	1.02***	1.05
	0.01	*0.03*
Mais de 6 (priv.) / 13 (corp.)	1.03***	1.27***
	0.01	*0.05*
Type of creditor		
Financial	0.97**	1.04
	0.01	*0.03*
Public administration	0.99*	0.76***
	0.01	*0.02*
Business	1.00	0.94
	0.01	*0.04*
Private	1.04***	1.06*
	0.01	*0.03*
Public creditors in majority		0.92***
		0.02
Cessation of activity		0.92***
		0.01
Liabilities		1.05***
		0.00
Fixed assets		1.02***
		0.00
Microenterprise		0.91***
		0.02
Type of collateral		
Real collateral		1.25***
		0.02
Financial collateral		1.03
		0.03
Other collateral		0.98
		0.02
Debtor's activ. sector (manuf. ind.)		
Agriculture and mining		1.04
		0.07
Construction		1.26***
		0.03
Trade, food and accommodation		0.97
		0.02
Other services		1.04*
		0.02
Economic activity	0.98***	0.87***
(non-interacted)	*0.00*	*0.01*
Constant	4.95***	50.75***
	0.23	*5.57*
N° observations	71,800	24,542

TABLE 3. Determinants of duration to closure, generalized gamma model

Notes: Table shows exponentials of coefficients, i.e. time-to-failure ratios, estimated assuming the generalized gamma distribution, stratifying by jurisdictional unit; omitted groups in parenthesis; standard-errors (in italics) adjusted to account for the variability stemming from the multiple imputation of the variables related to creditor type; p-values: * <0.1; ** <0.05; *** <0.01.

but only within 2 segments of duration, respectively below and above the median of

the survival function (15 days). Recall that the Cox model and the accelerated failure-time models have opposite coefficient readings, as coefficients greater than 1 mean a shortening of duration, in the first case, and an extension of it, in the second.

	Cox model: proportionality full duration	**Cox model: proport. by duration segments**		**Loglogistic model**
		up to 15 days	**over 15 dias**	
Requested by debtor (creditor)	5.47***	53.81***	4.08***	0.12***
	0.06	*3.40*	*0.06*	0.00
Several debtors	0.82***	0.85***	0.82***	1.16***
	0.01	*0.01*	*0.01*	0.01
Number of creditors				
4 - 5 (priv.) / 6 - 13 (corp.)	0.93***	0.93***	0.95***	1.05***
	0.01	*0.01*	*0.02*	*0.01*
Mais de 6 (priv.) / 13 (corp.)	0.90***	0.91***	0.92***	1.05***
	0.02	*0.02*	*0.02*	*0.01*
Type of creditor				
Financial	1.05***	1.05**	1.05**	0.96***
	0.02	*0.02*	*0.02*	*0.01*
Public administration	1.01	1.00	1.02	1.00
	0.01	*0.02*	*0.02*	*0.01*
Business	1.06***	1.03*	1.07***	0.98
	0.01	*0.02*	*0.02*	*0.01*
Private	0.98	0.99	0.99	1.00
	0.02	*0.02*	*0.02*	*0.01*
Private insolv. (corporate)	1.07***	1.22***	0.96**	0.91***
	0.01	*0.02*	*0.02*	*0.01*
Constant				0.56***
				0.02
Nº observations	72,885	72,885		72,885

TABLE 4. Determinants of duration to private and corporate insolvency declaration

Notes: Table shows exponentials of coefficients, i.e. hazard ratios (Cox regression) and time-to-failure ratios (generalized gamma regression), stratifying by jurisdictional unit; omitted groups in parenthesis; standard-errors (in italics) adjusted to account for the variability stemming from the multiple imputation of the variables related to creditor type; p-values: * <0.1; ** <0.05; *** <0.01.

The shortening of duration that comes from the case being initiated by the debtor is now even more visible than in Tables 1 and 2, especially in the first segment considered (up to 15 days). In fact, in such instances, there is no notification of the debtor, nor a possible opposition from his/her side which otherwise occupy a relevant length of time until the declaration of insolvency. The number of debtors and creditors also has a clear impact in terms of prolonging duration up to insolvency declaration, which may have to do with procedures for identifying the relevant creditors and debts. All these results hold both for the Cox model and in the parametric model.

The effects of the type of creditor variables differ from those previously presented for duration to closure, for example, as far as the lack of an impact of public administration creditors is concerned. Considering the evidence for parametric and semiparametric models jointly, only the presence of financial creditors - shortening duration to insolvency declaration - is statistically significant. Indeed credit institutions may be particularly routined in the initial steps of insolvency proceedings. Finally, the fact that the court is dealing with private insolvencies (vis-a-vis corporate ones) speeds up insolvency declaration, possibly reflecting a lesser complexity.

7. Conclusions

Efficiency of insolvency proceedings can have relevant macroeconomic impacts and duration will be one of the key factors determining such an efficiency. Indeed, the speediness of proceedings is crucial not only from the point of view of safeguarding the interests of those involved, but also of the reallocation of productive resources. Most insolvency requests of households in Portugal are made by themselves. In these cases, duration of proceedings will be particularly relevant, in order for them to recover, as quickly as possible, from a vulnerable financial situation. Therefore identifying factors that shorten duration can provide important insights from the viewpoint of public policies. Other factors that determine the quality of the insolvency procedure, outside the scope of this article, include the recovery rate of claims and safeguarding the priorities of different creditors and the rights of insolvents.

A transversal result to private and corporate insolvencies is the increase in duration when the case is filed by a creditor and when there is a greater number of parties, both debtors and creditors. This occurs at the beginning of the case, until the insolvency declaration, as well as at later stages, until its closure. Therefore, it could be useful to analyse whether there would be room to speed up that declaration, particularly in such instances, without calling into question the rights of the parties. As regards duration to closure, there is evidence that the intervention of private creditors leads to greater delays, suggesting that more support in decision-making by this type of creditors could be beneficial. With regard to other features that extend the duration of corporate insolvencies, it should be mentioned the size of firms, the volume of debt and assets to be sold, the existence of real collateral and activity in the construction sector. A detailed analysis of the cases with such features could clarify the reasons for that.

The database used in this article allows us to identify the firms that went into insolvency in recent years. It would be interesting to compare the situation of companies before, during and after this procedure, with that of companies in similar conditions, but which have not resorted to insolvency. As for private insolvencies, which have increased significantly in recent years, the possibility of combining this information with other individual databases, namely regarding the employment situation and indebtedness, will open, when feasible, interesting research opportunities.

References

Abreu Advogados (2017). "Alterações ao Código de Insolvência e Recuperação de Empresas." *Newsletter - grupo de Insolvência e Recuperação de Empresas*.

Albanesi, Stefania and Jaromir B. Nosal (2015). "Insolvency after the 2005 Bankruptcy Reform (April 2015)." *CEPR Discussion Paper*, 10533.

Almus, Matthias (2004). "The Shadow of Death – An Empirical Analysis of the Pre-Exit Performance of New German Firms." *Small Business Economics*, 23, 1253–1303.

Antunes, António R., Tiago Cavalcanti, Caterina Mendicino, Marcel Peruffo, and Anne Villamil (2019). "Tighter Credit and Consumer Bankruptcy Insurance." *Banco de Portugal Working Papers*, 21.

Bernstein, Donald S (ed.) (2017). *The Insolvency Review*. fifth ed., The Law Reviews.

Bricongne, Jean-Charles, Maria Demertzis, Peter Pontuch, and Alessandro Turrini (2016). "Macroeconomic Relevance of Insolvency Frameworks in a High-debt Context: An EU Perspective." *Economic Economy Discussion Paper - European Commission*, 32.

Bris, Arturo, Ivo Welch, and Ning Zhu (2006). "The costs of bankruptcy: Chapter 7 liquidation versus Chapter 11 reorganization." *The Journal of Finance*, 61, 189–201.

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

Davyddenko, Sergei A. and Julian R. Franks (2008). "Do bankruptcy codes matter? A study of defaults in France, Germany, and the UK." *Journal of Finance*, 63(2), 565–608.

Direção-Geral da Política de Justiça (2020). "Estatísticas trimestrais sobre processos de insolvências, processos especiais de revitalização e processos especiais para acordo de pagamento, 2020." *Destaque Estatístico Trimestral*, (78, 81, 84 e 87).

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

Graham, John R., Hyunseob Kim, Si Li, and Jiaping Qiu (2019). "Employee Costs of Corporate Bankruptcy." *NBER Working Papers*, 25922.

Kwon, Sehoon and Sang Buhm Hahn (2010). "Duration analysis of corporate bankruptcy in the presence of competing risks." *Applied Economics Letters*, 17(15), 1513–1516.

McGowan, Müge Adalet and Dan Andrews (2018). "Design of insolvency regimes across countries." *OECD Economics Department Working Papers*, 1504.

Morrison, Edward R. (2007). "Bankruptcy decision making: an empirical study of continuation bias in small-business bankruptcies." *Journal of Law and Economics*, 50(2), 381–419.

Osterhold, Christian and Ana Gouveia (2020). "Productivity, Zombie Firms and Exit Barriers in Portugal." *International Productivity Monitor*, 38, 29–49.

Pereira, Manuel Coutinho and Lara Wemans (2018). "How long does it take for a debt to be enforced by the Portuguese judicial sector?" *Banco de Portugal - Economic Studies*, 4(2), 1–28.

Ponticelli, Jacopo and Leonardo S. Alencar (2020). "Court Enforcement, Bank Loans, and Firm Investment: Evidence from a Bankruptcy Reform in Brazil." *The Quarterly Journal of Economics*, 131(3), 1365–1413.

Roberts, Michael R. and Amir Sufi (2009). "Financial contracting: A survey of empirical research and future directions." *Annual Review of Financial Economics*, 1, 207–226.

Rodano, Giacomo (2021). "Judicial efficiency and bank credit to firms." *Banca D'Italia Working Paper*, 1322.

Rodrigues, Gabriela Cunha, Laurinda Gemas, and Margarida Paz (eds.) (2017). *Insolvência e processo especial de revitalização*. Centro de Estudos Judiciários - Coleção formação contínua.

Titman, Sheridan (1984). "The effect of capital structure on a firm's liquidation decision." *Journal of Financial Economics*, 13, 137–151.

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

Appendix: Descriptive Statistics

Variable	Unit	Observations	Mean	Standard deviation	Min.	Max.
Households						
requested by the debtor (creditor)	binary variable	72,244	0.90	0.30	0	1
specialised court	binary variable	72,244	0.90	0.30	0	1
several debtors	binary variable	72,244	0.24	0.43	0	1
debtors	number	72,244	1.24	0.43	1	6
several creditors	binary variable	72,244	0.91	0.29	0	1
creditors	number	72,244	5.17	3.71	1	228
private creditors	number	72,244	0.16	0.68	0	48
public creditors	number	72,244	0.32	0.56	0	6
financial creditors	number	72,244	2.11	1.82	0	20
corporations credits	number	72,244	0.91	1.68	0	144
unidentified creditors	number	72,244	1.23	1.80	0	107
Corporations						
requested by the debtor (creditor)	binary variable	29,706	0.50	0.50	0	1
specialised court	binary variable	29,706	0.92	0.26	0	1
several debtors	binary variable	29,706	0.01	0.07	0	1
debtors	number	29,706	1.0	0.18	1	29
several creditors	binary variable	29,706	0.90	0.30	0	1
creditors	number	29,706	18.14	41.12	1	1,532
private creditors	number	29,706	3.48	14.94	0	1,423
public creditors	number	29,706	0.66	0.78	0	16
financial creditors	number	29,706	1.25	1.85	0	20
corporations credits	number	29,706	6.98	20.73	0	1,207
unidentified creditors	number	29,706	5.34	16.45	0	972
Central Balance Sheet Database variables						
micro firm	binary variable	25,572	0.76	0.43	0	1
liabilities	millions of euros	25,572	1.97	52.3	0	7,890
fixed asset	millions of euros	25,572	0.68	43.8	0	6,850
public creditor with majority	binary variable	25,565	0.10	0.31	0	1
company with no activity	binary variable	25,572	0.34	0.47	0	1
agriculture and extractive industry	binary variable	25,571	0.01	0.12	0	1
industry	binary variable	25,571	0.22	0.41	0	1
construction	binary variable	25,571	0.16	0.37	0	1
retail, accommodation and restaurants	binary variable	25,571	0.37	0.48	0	1
other services	binary variable	25,571	0.23	0.42	0	1
Central Credit Register variables						
real collateral	binary variable	26,747	0.22	0.41	0	1
financial collateral	binary variable	26,747	0.10	0.30	0	1
other collateral	binary variable	26,747	0.50	0.50	0	1

TABLE A.1. Descriptive Statistics

Note: Liabilities and fixed assets are on 2016 prices.

Non-technical summary

January 2022

A micro-level analysis of corporate income taxation in Portugal

Cláudia Braz, Sónia Cabral and Maria Manuel Campos

Understanding the design and functioning of corporate income tax (CIT) systems is crucial in a world marked by globalisation, where tax competition and firms' tax planning practices have been gaining prominence. This article aims to provide a thorough analysis of corporate income taxation in Portugal in the last decade.

The article presents an encompassing description of the Portuguese CIT framework, alongside a characterisation based on aggregated indicators and an international comparison. Portugal stands out as one of the countries with higher CIT top statutory rates amongst OECD countries (Figure 1 - Panel A). In Portugal, the top statutory rate currently stands at 31.5%. Although the CIT general rate was cut over time, following the international trend, progressivity increased in the last decade reflecting the creation and subsequent reinforcement of a State surcharge.

Corporate tax systems are typically complex. Besides the structure of statutory rates, firms' actual tax burdens reflect other elements of regulations such as tax benefits, incentives and deductions. Hence, the literature commonly takes observed effective tax rates (ETR) as indicators of a firm's tax burden. ETRs can be broadly defined as the ratio of tax expenses to a metric of pre-tax income. In this article, we use a large and detailed micro-level database to compute two measures of firms' ETR: a metric using earnings before taxes (EBT) in the denominator, which, in essence, is more comparable to statutory rates; and a measure using earnings before interest, taxes, depreciation and amortisation (EBITDA), more useful to compare sectors and firms. We analyse effective taxation and investigate the relationship with several characteristics of non-financial corporations in a sample covering a significant part of Portuguese firms in the period from 2010 to 2019.

Some well-known features of CIT collection in Portugal are evident in the results. In particular, the large proportion of micro firms (81%) and their much smaller contribution to total tax expenses and GVA (around 16% in both cases), with the opposite occurring with large firms. Results also show that the average ETR (based on EBT) derived from micro data stood broadly stable at around 25% in the last years. Although caution is warranted when comparing the magnitudes of ETRs and statutory rates, Panel B of Figure 1 shows that they evolved broadly in tandem over the last decade. The fact that the ETR stands below both the top statutory rate reflects the progressivity stemming from the rate structure and the existing tax benefits, incentives and deductions. Moreover, the micro-based ETR is close to a "weighted" average statutory rate computed taking into account the distribution of firms' EBT in the sample.

Regression analysis suggests the existence of non-linear relations between firms' ETRs, using EBITDA in the denominator, and their size and productivity. The estimates

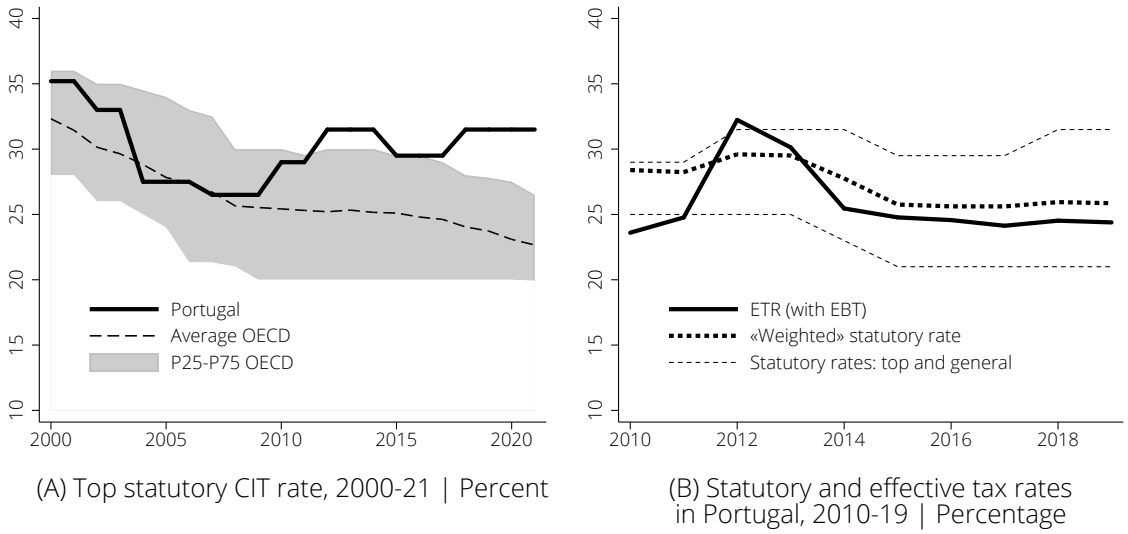


FIGURE 1: CIT rates in Portugal: statutory and effective tax rates

Sources: OECD - Tax database, IES and authors' calculations.
Notes: The top statutory CIT rate shown in Panel A is labelled as the "combined CIT rate" by the OECD. It takes into account the basic combined central and sub-central government statutory CIT rates. Sub-central government CIT rate shows the representative municipal rate. The State surtax is included. The ETR (in Panel B) uses current income tax expenses (including autonomous taxation) in the numerator. The "weighted" statutory rate corresponds to an average of the different statutory rates (reduced, general, and general plus surcharges) weighted by the share of firms subject to each bracket in each year. EBT was used as a proxy of taxable income in the computations. Both the top and the "weighted" statutory rates include the representative local surcharge of 1.5%.

support the expected negative association between ETRs and financial leverage, given that interest expenses are deductible for tax purposes. Firm's capital intensity is also negatively associated with ETR, confirming the hypothesis that the Portuguese tax framework is favourable for firms that invest more in fixed assets. Even if ETR levels differ by sector of activity, the signs and magnitudes of the estimated coefficients of firms' characteristics do not vary considerably by sector.

A micro-level analysis of corporate income taxation in Portugal

Cláudia Braz
Banco de Portugal

Sónia Cabral
Banco de Portugal

Maria Manuel Campos
Banco de Portugal

January 2022

Abstract
This article analyses corporate income taxation in Portugal. It provides an encompassing description of the Portuguese corporate income tax system. In addition, firm-level effective tax rates (ETR) are computed using a micro database, and their relation with several firms' characteristics is examined in the period 2010-2019. In terms of results, Portugal stands out as one of the countries with higher top statutory tax rates amongst OECD countries. Although the general rate was cut over time, progressivity increased substantially and collection is very concentrated on a small number of large firms. Regression estimates suggest the existence of non-linear relations between firms' effective taxation and their size and productivity, and negative associations between ETR and both financial leverage and capital intensity. (JEL: H25, H26, L25)

1. Introduction

In the last decades the increasing digitalisation of the economy and the wide-ranging effects of globalisation have been posing challenges to corporate income tax systems worldwide. Typically, corporate tax receipts are not the main source of government's revenue and their underlying tax base is extremely sensitive to legal conditions in jurisdictions all over the world.[1] Indeed, firms frequently engage in tax planning strategies by exploiting gaps and mismatches in legislation to reduce their tax burden, often implying shifting profits to low tax locations. The resulting increase in tax competition between countries has led to a general declining trend of corporate

Acknowledgements: The authors are especially grateful to Luís Morais Sarmento for his useful comments and suggestions. We also thank the editor, Pedro Duarte Neves, and António Antunes, Jorge Correia da Cunha, Joana Garcia, Lara Wemans and participants in an internal seminar at Banco de Portugal for their comments. The analyses, opinions, and findings expressed in this article are those of the authors and do not necessarily coincide with those of Banco de Portugal or the Eurosystem.

E-mail: crbraz@bportugal.pt; scabral@bportugal.pt; mmcampos@bportugal.pt

1. According to the classical theory of optimal taxation, capital income taxation is undesirable because it becomes very distortive over long time horizons. Equity and efficiency arguments may, however, justify taxing capital income. In the case of the corporate income tax, the major constraint is the possibility for firms to relocate their activities abroad but practical arguments may favour its existence (like complementing personal income taxation, as difficulties subsist in taxing all types of household's income, and making less attractive to shift income between the two taxes). For a recent survey on how capital should be taxed, see Bastani and Waldenström (2020).

tax rates and to the establishment of preferential regimes at the national level. With regard to international taxation, efforts to protect corporate income tax bases against profit shifting and to promote coordination on tax collection have been implemented. In this context, a deep understanding of the design and functioning of corporate tax systems is of utmost importance to assess the different policy options on the table.

Tax research has a multidisciplinary nature, often analysing relevant questions from the viewpoint of accounting and corporate finances, law and/or economics. As a result, the empirical literature on corporate taxation is very diverse. In terms of data, this field of research either relies on macroeconomic aggregates or is based on micro databases, which allow for a more detailed analysis.

Corporate tax systems are typically complex and subject to numerous modifications over time. Besides the structure of statutory rates, corporate tax rules usually encompass various other elements relevant for the calculation of a firm's tax burden, like benefits, incentives and deductions, with both domestic and international dimensions. Given this diversity, statutory rates do not perfectly reflect the tax burden of firms. An indicator of a firm's tax burden commonly used in the literature is the observed effective tax rate (ETR), defined as the ratio of income tax expenses to a metric of pre-tax income.

The aim of this article is to analyse corporate income taxation in Portugal. With recourse to a large and detailed micro database, it derives backward-looking ETRs for the period from 2010 to 2019. In addition to describing developments over the last decade, the article investigates the relation between effective income taxation and several firms' characteristics, like sector of activity, size, leverage, capital intensity and productivity. The main patterns and relations highlighted in an exploratory analysis are further examined with simple multivariate regressions. Results are put into context with legislation and the use of macro-based indicators, also from the perspective of an international comparison.

Presently, Portugal stands out as one of the countries with higher corporate income tax top statutory rate amongst OECD countries. Indeed, the international trend of reduction of statutory tax rates was not followed in the last decade. The general rate was cut substantially but its overall impact appears to have been offset by the introduction of a State surcharge in 2010, stepped up subsequently both in terms of reference rates and underlying progressivity. In the same period, total receipts as a ratio to GDP have remained broadly constant, as the collection is very concentrated on a low number of large non-financial firms. The average ETR in the sample, measured using earnings before taxes (EBT) as a proxy for taxable income, stood relatively stable at 25% since 2014. Even if some caution is warranted when comparing to statutory rates, this value stands well bellow the top statutory rate (31.5% in 2019) and slightly below an average statutory rate computed using as weights the EBT from the micro data.

Although evidence in the literature is not consensual, we find several statistically significant associations between firms' features and their ETRs, using earnings before interest, taxes, depreciation and amortisation (EBITDA) in the denominator. Estimates point to the existence of non-linear relations between firms' effective taxation and their size and productivity. The results also support the expected negative association between ETRs and financial leverage, given that interest expenses are deductible for tax

purposes while dividends are not. Moreover, capital intensity is negatively associated with ETR, confirming the hypothesis that the tax framework is favourable for firms that invest more in fixed assets. Finally, even if ETR levels differ by sector of activity, the signs and magnitudes of the estimated coefficients of firms' characteristics do not vary considerably. It is important to note, however, that these multivariate regressions have no intention to establish causality.

The article is organised as follows. Section 2 discusses some of the related literature that frames this study. Section 3 provides a general overview of corporate income taxation in Portugal, including an international comparison. Section 4 describes the database and the main variables used. A detailed descriptive analysis of the evolution of the effective income taxation of Portuguese firms is reported in Section 5. Section 6 complements the previous section by estimating some multivariate regressions. Section 7 concludes.

2. Related literature

The review of the vast literature on the several aspects of taxation is beyond the scope of this article. Instead, this section offers a non-exhaustive list of references that are related to this article and provide a framework for the analysis, with a particular focus on studies using firm-level data.

Corporate income taxation has been an important area of research over the last decades and a number of new methods and techniques have emerged in distinct areas of the literature. A first thoughtful review of the empirical tax research in the accounting literature until 2000 is presented by Shackelford and Shevlin (2001). Hanlon and Heitzman (2010) provide a valuable summary of more recent developments in the literature and an extended discussion of the various measures of tax avoidance. More recently, a survey of the literature on corporate tax planning over the previous decades is provided by Wilde and Wilson (2018), while Wang *et al.* (2020) synthesise the major findings of the research on tax avoidance from the accounting and finance literature. Beer *et al.* (2020) review the rapidly growing empirical research on international tax avoidance by multinational corporations. As defined in this literature, tax avoidance ranges from the reduction of the corporate tax burden by legitimate use of tax rules to, on the other extreme, violation of tax laws (tax evasion). Thus, aggressive corporate tax strategies do not necessarily indicate that the firm uses illegal procedures in tax reporting. The most common metric for tax avoidance is the observed ETR: firms that are more tax aggressive have lower ETRs.

There are several definitions of ETRs in the literature. In general, we can identify two broad types of firm-specific ETRs: forward-looking and backward-looking (see Nicodème (2001) for a detailed discussion of the pros and cons of these two concepts).

Forward-looking studies do not compute observed ETRs, but rely on theoretical features of the tax system to obtain implicit tax rates. These studies calculate the net present value of an hypothetical potential investment using specific sources of financing both in the presence and in the absence of taxes. The implicit taxation is derived from

this difference, under several calibrated assumptions. The method was formalised at the country-level by Devereux and Griffith (1998, 2003) and extended to firm-level data by Egger *et al.* (2009).

The approach of micro backward-looking ETRs, followed in this article, uses firms' financial statements to derive effective taxation, usually as ratios of tax expenses on other accounting items such as pre-tax profit or gross operating profit. These ETRs are central to the research on corporate taxation and, as mentioned above, regularly appear as a proxy for the tax burden of a firm. An advantage of this methodology is that it uses observed data, thus allowing all elements of taxation to be taken into account. A second advantage is that it facilitates studying effective taxation at sectoral level and for groups of firms, being useful to examine the relation between tax liabilities and firms' characteristics. It is, however, a backward looking measure of taxation as it is determined by past decisions of the firm, including tax planning. Hence, it cannot be used to assess firms' behavioural reactions.

There are several methodological contributions in the accounting literature on the details of the computation of observed ETRs from firms' financial statements (see Omer *et al.* (1991) or Plesko (2003) for reviews). Regardless of the specificities of the metrics, observed ETRs are incapable of isolating the effects of specific features of national tax systems. The measure is encompassing, capturing all forms of tax reduction relative to pre-tax income, whether through tax sheltering, location decisions, income shifting, tax preferences within the tax code, or changes in legislation (e.g., Dyreng *et al.* 2017).

Irrespective of the specific metric of a firm's tax burden, there is ample evidence in the literature of a relation between corporate income taxation and different firms' characteristics. Gupta and Newberry (1997) was one of the first studies on effective income taxation with longitudinal firm-level data. Most of the subsequent research follows the covariates identified therein. Firm attributes like size, financial debt and capital intensity are used in most studies but the results are not consensual, in particular concerning the relation between firms' size and effective taxation (a review of this literature can be found in Delgado *et al.* 2014).

Several studies investigate the relation between effective taxation and firm characteristics in European countries. For Romania, Lazăr (2014) finds that capital intensity and leverage negatively affect firms' ETRs, while firm size has no effect. Janssen (2005) concludes that ETRs do not differ much from statutory tax rates in the Netherlands, even if capital intensity is negatively associated with ETRs. Stamatopoulos *et al.* (2019) show that larger firms in Greece face higher ETRs than smaller ones and that firm's capital intensity is negatively associated with ETRs. For Germany, using a quantile regression approach, Delgado *et al.* (2018) find positive estimates for the relation between ETRs and firm size in the first quantiles and negative ones in the upper end of the distribution. The opposite result is found for leverage: the sign of the relationship goes from negative to positive. Nicodème (2002) computes firm-level ETRs for eleven European countries, the US, and Japan, and shows that tax burdens are more favourable for large firms and for specific sectors.

For a sample of Chinese listed firms, Hsieh (2012) detects a negative relation between firm size and effective taxation, but Liu and Cao (2007) find no significant effects

of firm size and capital intensity on ETRs, while the impact of leverage is negative and significant. For Australia, the results of Richardson and Lanis (2007) indicate that corporate ETRs are negatively associated with firm size, leverage and capital intensity. Using data for Ethiopia, Mascagni and Mengistu (2019) show that small firms face a higher effective tax burden than larger firms, while middle-sized firms face the lowest burden. They also find evidence of a negative relation between leverage and capital intensity and ETRs. Fernández-Rodríguez *et al.* (2021) study the determinants of ETRs in nine emerging countries and find that firm size and inventory intensity positively affect the ETR, while leverage and capital intensity have a negative effect. In a different vein, Bartolini (2018) uses firm-level data for six OECD countries from 1998 to 2014 and documents the existence of a tax burden gap alongside the productivity gap: firms at the productivity frontier enjoy lower effective taxation.

There are some studies on effective taxation using micro-level data for Portugal, mostly made as master's dissertations, but using relatively small databases (e.g., Costa *et al.* 2012, Bessa 2016, Praça 2018, Topa 2018). Their conclusions are broadly similar even if the periods and samples examined differ. There is evidence of a reduction of the ETR after the corporate income tax reform in 2014. Leverage was found to have a negative relation with effective taxation, while the sign of the other covariates varied with the specific ETR measure used. Our article contributes to this literature by studying the effective income taxation of Portuguese firms, using a large and detailed database in the period 2010-2019. It also presents an encompassing description of the Portuguese CIT framework, alongside a characterisation based on macro indicators and an international comparison.

3. The Portuguese Corporate Income Tax System

3.1. Information for Portugal

The reform of direct taxation implemented in 1989 in Portugal laid down the foundation of a modern tax system. In a nutshell, several schedular taxes on different types of income and an encompassing income tax were replaced by two taxes structured according to the nature of different groups of taxpayers: the personal income tax (PIT, Portuguese acronym: *IRS*) and the corporate income tax (CIT, Portuguese acronym: *IRC*). Since then, corporate taxation has been subject to changes but the initial underlying structure still prevails.

The CIT is generally levied on all corporate entities that are resident or have a permanent establishment in Portugal. These companies are taxable on their worldwide income. Taxes paid abroad on foreign-source income may be credited against CIT liability. The taxable income is based on the profit and loss accounts made under the applicable accounting framework, whose result is adjusted according to the rules set forth in the CIT code. Afterwards, eligible tax losses from previous years and tax benefits may be deducted from the taxable income. This is the so-called direct method for the

determination of taxable income.[2] Figure 1 schematically represents the determination of the amount to be paid under CIT in Portugal.

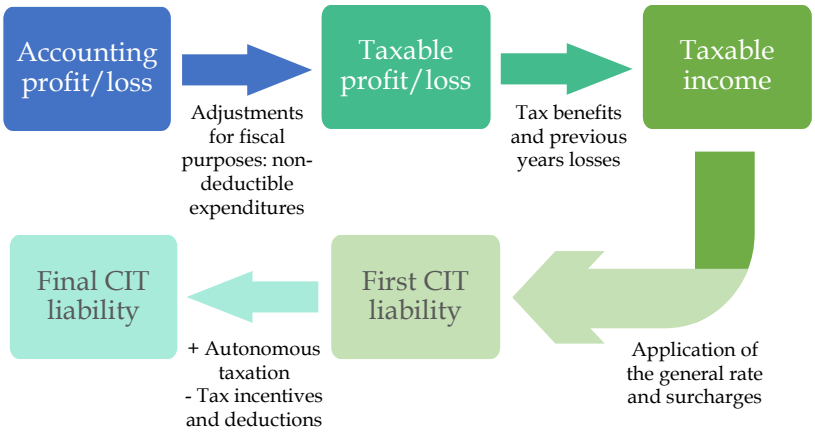


FIGURE 1: Determination of the corporate income tax in Portugal

Expenses are deductible for CIT purposes if they are documented and incurred by a company in order to generate or guarantee taxable income, but there are also non-deductible expenses. All fixed assets, except land, can be amortised for tax purposes. As a general rule, fixed assets are depreciated under the straight line method and the maximum and minimum rates are set in legislation. Since 2014, interest expenses are deductible up to the highest of (i) €1 million or (ii) 30% of EBITDA. Worldwide capital gains are regarded as regular income and subject to CIT. Losses considered for tax purposes generated as of tax year 2017 may only be carried forward for 5 years, but those incurred between 1 January 2014 and 31 December 2016, or after that year in case of small and medium-sized enterprises (SMEs[3]), may be carried forward for the next 12 years.[4] Such losses may only be offset up to a maximum of 70% of the taxable income. Tax losses carry-back is not allowed.

Tax incentives and deductions are considered in the determination of the final CIT liability, but the tax due cannot be less than 90% of the CIT a company would pay in

2. If the application of the direct method is not possible, the tax base is determined on the basis of circumstantial evidence - indirect method.

3. According to the Decree-Law 372/2007, the definition of micro, small and medium-sized enterprises (SMEs) is made according to the official EU classification by size categories, as described in the Recommendation 2003/361/EC of the European Commission of 6 May 2003. Following this definition, SMEs are firms which employ less than 250 persons and that have an annual turnover not exceeding €50 million or an annual balance sheet total not exceeding €43 million. Within the SME category, a small firm is defined as a firm which employs less than 50 persons and whose annual turnover or annual balance sheet total does not exceed €10 million. A micro-firm is defined as a firm which employs less than 10 persons and whose annual turnover or annual balance-sheet total does not exceed €2 million. All other firms not classified as SMEs are considered as large firms.

4. Due to the pandemic crisis, the supplementary budget of 2020 extended by two additional years the deduction of losses generated between 2014 and 2019 and to 12 years for losses generated in 2020 and 2021, regardless of the type of firm.

their absence. Examples of tax incentives under national legislation are the Madeira free-trade zone regime, contractual incentives granted for big industrial investment projects, tax credits and regimes for investment in general and, in particular, for R&D-related investment, corporate reorganisations, the urban property rehabilitation regime, among others. Firms may also qualify for a notional deduction in case of an equity injection by the shareholders (*Remuneração Convencional do Capital Social*), equivalent to 7% per year during 6 years (up to an injection of €2 million). This is meant to incentivise the financing of firms through equity and reduce the preferential treatment of debt in the CIT context.

To reach the final CIT liability, an autonomous taxation is also imposed separately on certain (listed) expenses. The autonomous taxation was introduced in 2001 with the aim of mitigating tax fraud and evasion and it is levied on expenses considered as not directly related to the companies' activity (undocumented expenses, related to ownership or use of passenger vehicles, bonuses to managers, among others).

In 2021, the general CIT rate in mainland Portugal is 21% (Table 1). A State surtax (*derrama estadual*) is levied on companies with higher taxable profits. A reduced rate of 17% applies to the first €25,000 of taxable income earned by SMEs. Also, if companies have a turnover smaller than €200,000 and a total balance sheet not exceeding €500,000, and fulfil other requirements, they can opt for a simplified taxation regime. Lastly, it is worth mentioning that the Autonomous Regions of Madeira and Azores have reduced general rates: currently, they stand at 20% and 16.8%, respectively.

		Rate (%)
General rate		21
SMEs (taxable income up to €25,000)[a)]		17
State surcharge ('*derrama estadual*') for companies with taxable income:[b)]	€1.5 million to €7.5 million	3
	€7.5 million to €35 million	5
	above €35 million	9
Local surcharge[c)]		1.5

TABLE 1. Corporate income tax rates in mainland Portugal, 2021 | Percentage

Notes: a) The reduced rate is applicable to the first €25,000 of taxable income, while for the amounts in excess the general rate applies.
b) For taxable income exceeding €1.5 million: i) when taxable income is higher than €7.5 million and up to €35 million, a rate of 3% is applied to the following €6 million taxable income, while a 5% rate applies to taxable income in excess of €7.5 million; ii) when taxable income is higher than €35 million, a rate of 3% is applied to the following €6 million taxable income, 5% to the next €27.5 million taxable income and 9% to taxable income above €35 million.
c) The maximum rate is 1.5% but the municipalities are allowed to levy a lower surtax on companies with a turnover not exceeding €150,000.

Although the architecture of CIT has remained stable since its inception, several changes were introduced over the years (Figure 2). The main changes stemmed from arguments of tax competition, the incorporation of rules approved at the EC/EU level, the general goal of promoting investment and the fight against tax evasion and fraud. Concerns over a possible loss of tax competitiveness as other economies decreased their rates justified successive reductions in the general tax rate from 36.5% in 1990 to 21% in 2021. However, a national surcharge was introduced in 2010, at that time as part

of a fiscal consolidation package aiming at accelerating the reduction of the excessive deficit and curbing public debt growth. Since then, progressivity increased substantially both through brackets and rates.[5] Progressivity in CIT may also stem from tax benefits, incentives and deductions. Given their high number and frequent changes, their overall impact is very difficult to assess. According to the report by the group mandated to analyse tax benefits in Portugal in 2019 (*Grupo de Trabalho para o Estudo dos Benefícios Fiscais* 2019), the number of tax benefits in force in the legal system exceeds 500[6], of which 121 refer to CIT. Further, the group concludes that, for more than 50% of these benefits, there is no quantified expenditure or it is not possible to quantify it on the basis of available information.

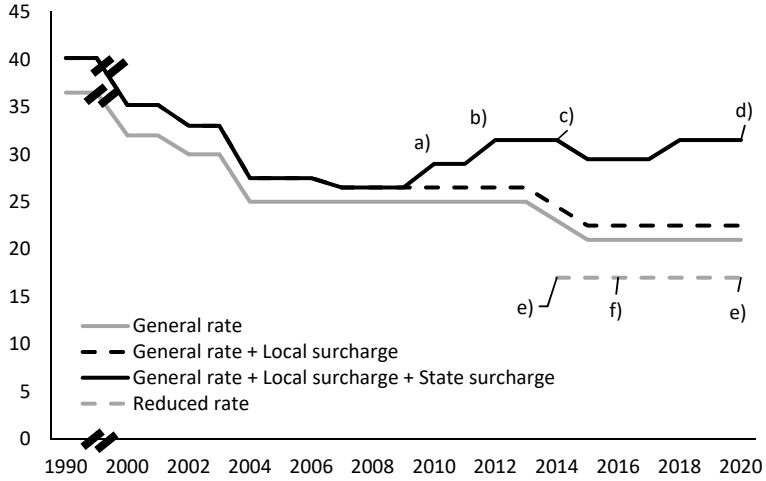


FIGURE 2: Evolution of statutory corporate income tax rates in mainland Portugal | Percentage

Notes: Regarding the State surcharge, the details in each legislative change were the following:
a) 2010 - 2.5% for taxable income above €2 million;
b) 2012 - 3% from €1.5 to €10 million taxable income and 5% above €10 million;
c) 2014 - 3% from €1.5 to €7.5 million taxable income, 5% from €7.5 to €35 million taxable income and 7% above €35 million;
d) 2018 - 3% from €1.5 to €7.5 million taxable income, 5% from €7.5 to €35 million taxable income and 9% above €35 million. No additional changes until 2021.
The reduced rate was introduced in 2014 and applies only to SMEs:
e) to the first €25,000 of taxable income in 2014-2015 and 2020-2021.
f) to the first €15,000 of taxable income in 2016-2019.

5. A tax system is progressive when the marginal tax rate is greater than the average tax rate. In this case, it is assessed at firm-level, having as reference taxable income. The progressivity of the Portuguese CIT system can be illustrated with some simple examples using the rates of Table 1. Consider that firm A is a SME and has a taxable income of €100,000 euros at the end of 2021. Up to €25,000 the applicable rate is 17% and the remainder of that limit will be taxed at 21% (general rate). Now, consider that firm B is a large firm with a taxable income of €40 million at the end of 2021. The general rate of 21% applies up to the limit of €1.5 million euros. The excess of that limit is divided in three parts: one equal to €6 million to which a State surcharge of 3% applies; another equal to €27.5 million to which a surcharge of 5% applies, and another equal to the taxable income exceeding €35 million to which a surcharge of 9% applies. For the sake of simplicity it was excluded in these examples, but a local surcharge is added at the aforementioned rates (0% to 1.5% of taxable income depending on the municipality where the firm is located).

6. This high number does not include benefits related to municipal taxes or decided by local authorities, nor preferential non-standard VAT rates (reduced and intermediate rates).

Regarding CIT administration, the accounting and tax periods coincide with the calendar year. Corporate taxpayers must electronically file a tax return for a given year in the first half of the following year. The tax liability is computed by the taxpayer (self-assessment). Companies must make a prepayment of CIT during the tax year. The prepayment is 80% of the preceding tax year's CIT liability for companies whose turnover does not exceed €500,000. For companies with a turnover in excess of this amount, the prepayment is 95% instead. The prepayments must be done in three instalments: July, September and December. In addition, companies must make a special prepayment which is meant to function as a minimum tax and is gradually being eliminated. An additional prepayment applies to firms subject to the State surcharge, made in three instalments and simultaneously with the ordinary prepayments. All the prepayments are creditable against the taxpayer's final CIT liability.

In terms of revenue, CIT is the third biggest tax in Portugal. On average, over the last two decades, it represents 8.7% of the tax burden, which compares to 23% and 16.9% in the cases of VAT and the personal income tax, respectively. Its weight as a ratio to GDP is relatively small when compared with the two other main taxes, 3% on average over 2000-2020, and it has been fairly stable, in spite of the several changes to the tax code, in particular regarding rates. The relative stability of the CIT revenue to GDP ratio over time is also observed for the EU average (Nicodème *et al.* 2018).

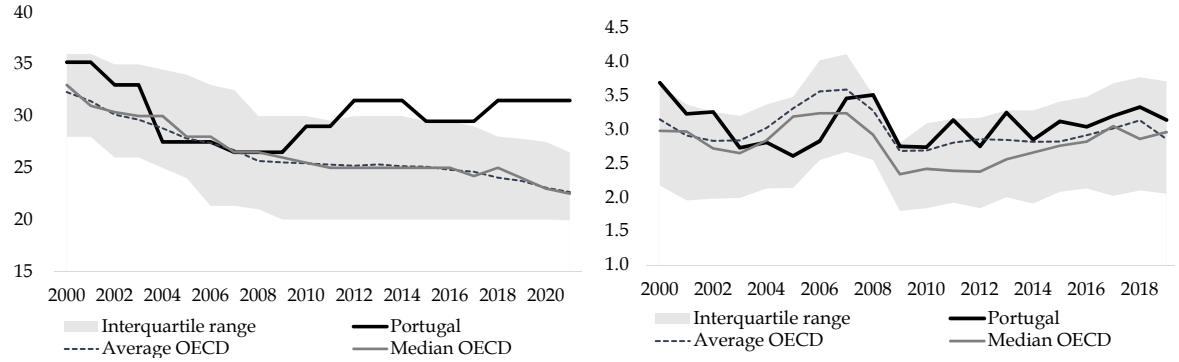
3.2. International comparison

In OECD countries, the decline in CIT rates has been a steady and widespread trend in the last two decades (Figure 3 - Panel A). For countries with lower tax rates (percentile 25), the bulk of the adjustment seems to have occurred up to the onset of the 2008/2009 financial crisis. The same happened for countries with higher rates (percentile 75), although in this case there is also a noticeable acceleration in the reduction after 2015, leading to further convergence in CIT rates across countries. Portugal emerges as a different case. The reduction in the general rate gained momentum after 2000 but it was more than offset by the introduction and further increase in the State surcharge for higher taxable income brackets. As a result, Portugal stands out as one of the countries with higher CIT top statutory rate: above or equal to percentile 75 since 2012 and in the top 3 since 2018.

In terms of CIT receipts as a percentage of GDP, Portugal stood in the last decade systematically above the OECD average, but below percentile 75 (Figure 3 - Panel B). This difference was even widening in the more recent period, which contrasts with tax rates' developments. The fact that the Portuguese CIT system became more progressive with the introduction and increased importance of the State surcharge may partly explain this result. Another possible explanation may be related to the relative generosity of tax benefits, incentives and deductions. Other reasons can be connected to differences in taxable bases and in the size of the corporate sector among countries (Nicodème *et al.* 2018).

For the euro area countries, it is possible to have an estimate on the magnitude of the joint effect of the rates' progressivity and tax benefits, incentives and deductions.

The European Commission regularly publishes an implicit tax rate on corporate income based on National Accounts data. It essentially uses as a proxy for taxable income the net operating surplus of financial and non-financial corporations, adding net interest, rents, dividends and insurance property income. The results for 2019 are depicted in Figure 4. For all countries except Slovenia, the implicit tax rate is smaller than the top statutory corporate income tax rate. Portugal has one of the highest values in both measures, but the difference between them is relatively small: around 25%, the seventh smaller in this group of countries.



(A) Top statutory CIT rate, 2000-2021 | Percentage (B) CIT receipts, 2000-2019 | Percentage of GDP

FIGURE 3: Top statutory CIT rate and CIT receipts - Portugal and OECD countries

Source: OECD - Tax database.
Notes: The shaded area corresponds to the interquartile range, i.e., to the difference between 75th and 25th percentiles of the respective distributions of OECD countries. The reported top statutory rate is labelled as "combined CIT rate" by the OECD. It takes into account the basic combined central and sub-central government statutory CIT rates. Sub-central government CIT rate shows the representative municipal rate. The State surtax is included.

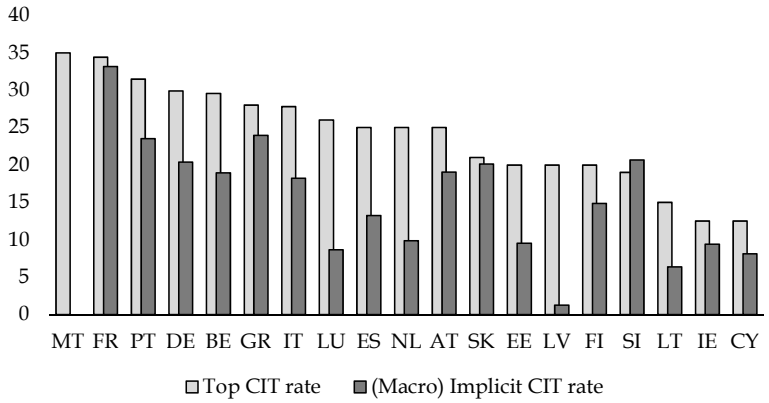


FIGURE 4: Top statutory corporate income tax rate (including surcharges) and (macro) implicit tax rate on corporate income - Euro area countries (2019) | Percentage

Source: European Commission - Taxation data.
Note: Data for the implicit tax rate in Malta are not available and for Luxembourg the last available year is 2018.

4. Database and variables definition

4.1. The measure of effective tax rate (ETR)

As stated in Section 2, the micro backward-looking methodology of ETRs is especially useful to examine differences in effective taxation for distinct types of firms. These ETRs are broadly defined as the ratio of income tax expenses to pre-tax income. However, the choice of the most appropriate indicator is not obvious and there are several alternatives in the literature, regarding both the numerator and the denominator.

Firms disclose income tax expenses split into two components: current and deferred. The latter relates to past events, like losses or the revaluation of assets, which have an impact on future tax liabilities. As such, we chose not to consider the deferred component and focus only on the current portion of tax expenses. Regarding autonomous taxation, in our database it is separately reported by some firms, while others include it in the current component of tax expenses. Thus, we chose to consider it in the numerator of our metrics of ETRs.

Regarding the denominator, the literature presents numerous alternatives that range from turnover to income before taxes (see, for instance, Lazăr 2014 for a discussion). The most commonly used denominator is EBT, resulting in an ETR that allows some comparison with statutory tax rates and with the implicit tax rate on corporate income based on National Accounts. At the margin and for a firm without preferential tax treatments, using EBT in the denominator should result in an ETR equal to the statutory tax rate. However, to investigate the relation of firm-specific characteristics with their tax burden, the informational content of a ratio with EBT may not be the most appropriate. If both the numerator (income tax expense) and the denominator (income) reflect behavioural responses related to tax preferences, then any systematic variation in ETRs because of firms' tax planning activities will not be properly detected. For instance, *ceteris paribus*, a higher amount of debt leads to higher interest expenses, resulting in lower tax expenses but also in a lower EBT, which, in turn, may lead to an increase in the ETR of that firm. Therefore, the economic literature using micro data tends to include also other indicators in the analysis.

An alternative denominator that excludes several tax planning–induced distortions is gross operating profits, i.e., EBITDA. A ratio computed with EBITDA has the advantage of isolating in the numerator the tax-minimising effect of deductible items, as interest expenses and depreciations. Consider, for example, the situation in which a firm relies more heavily on debt financing rather than equity financing. Given that interest expenditure is tax deductible while dividends are not, firms with higher leverage are expected to pay relatively lower taxes. An ETR computed with EBITDA will adequately capture this effect. Thus, focusing on EBITDA enhances comparability across firms and sectors with different financing structures and capital intensity.[7] A caveat of using the

7. Using EBITDA as the denominator does not completely account for all tax planning–induced distortions. For instance, it does not take into account more sophisticated tax planning-strategies developed by multinational firms in order to lower their effective tax burden, such as changes in transfer pricing or

EBITDA is that the corresponding ETR levels will be lower that those computed using EBT and should not be compared with statutory rates. If we define a firms' income as the residual value available to remunerate shareholders, after remunerating all other productive factors, then EBITDA does not correspond to this definition of income and tends to introduce a downward bias in the levels of the resulting ETRs.

Driven by these considerations, we follow Nicodème (2002) and Lazăr (2014) and favour using EBITDA in the denominator (ETR2) when comparing sectors and firms and in the regression analysis. We also present selected evidence based on EBT (ETR1), particularly as it is more comparable to statutory rates. However, some caution is still warranted in the comparisons of ETR1 with statutory tax rates. First, EBT is merely a proxy for taxable income, which is not reported in firms' financial statements. Second, two additional factors that work in opposite directions are worth mentioning: tax incentives and deductions, which reduce the ETR *vis-à-vis* statutory rates, and autonomous taxation, which increases it.

4.2. Description of the database

Our firm-level balance sheet data is based on annual information for Portuguese firms reported under Simplified Corporate Information (*Informação Empresarial Simplificada*, Portuguese acronym: IES). The IES follows the new accounting standards system from 2010 to 2019, and it covers virtually the universe of Portuguese non-financial corporations.[8] The universal coverage of IES emerges from its nature, as it is the system through which corporations report mandatory information to the tax administration and statistical authorities. Tax-related information is, however, less encompassing than that provided through other reporting mechanisms (e.g., *Modelo 22*). Under IES, firms provide detailed annual balance sheet, profit and loss accounts. It further contains information on firms' characteristics such as number of employees, age and main sector of economic activity according to the Portuguese industrial classification Revision 3 – *Classificação Portuguesa das Actividades Económicas* (CAE).

Some filters were imposed on the data to eliminate erroneous, inconsistent or missing observations. Firstly, the analysis was restricted to firms for whom there was information available for a set of key variables, such as age, regional location and sector of activity. Secondly, we further restricted the sample to firms with strictly positive values for production, intermediate inputs, gross value-added, employment, labour costs and total assets. Moreover, the analysis focuses only on firms located in mainland Portugal. The Autonomous Regions of Madeira and Azores have distinct statutory tax rates and the existence of the Madeira free trade zone could influence the results.

Further sample restrictions emerge from the definition of our main variable of interest: the ETR. As detailed above, we define the ETR as a ratio between tax expenses and a pre-tax income metric. Negative figures in either the numerator or

---

the strategic choice of location of intangible assets (e.g., Beer *et al.* 2020). The information available in our database does not allow to control for these practices.

8. IES registry takes firms individually, regardless of whether they are part of an economic group or not.

the denominator result in measures of ETR that lack a proper economic interpretation. Hence, as often done in this strand of literature, these observations were excluded from the analysis, even if restricting the sample to firms with positive income and taxes may induce some selection issues. In addition, very small values of the denominator can result in ratios of ETR of unreasonable magnitudes, so only firms that display profits larger than their tax liability are included. In practice, we limited our sample to firms whose ETR lies between 0% and 100% and have strictly positive pre-tax profits and tax expenses. This restriction implies dropping around 37.5% of observations in the whole period, of which 87.5% refer to micro firms. However, we still retain information representing approximately 70% of total assets, turnover, gross value-added and employment. Moreover, in each year, firms in the final sample account for an average of 90% of total CIT paid by non-financial corporations in Portugal and 70% of overall CIT collected by the government (Figure 5).

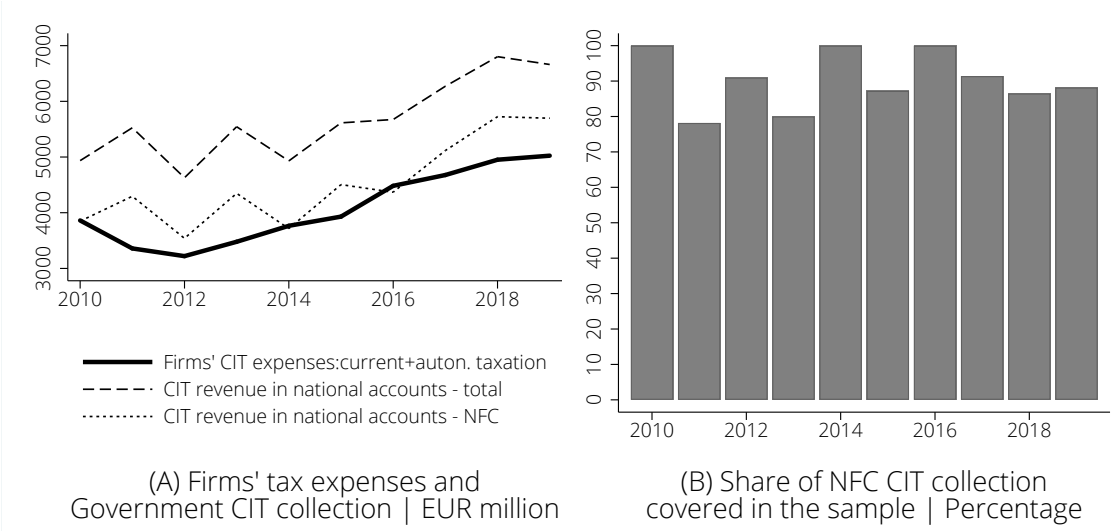


FIGURE 5: Final sample firms' tax expenses and government revenue from the CIT

Sources: Own calculations based on IES and Statistics Portugal.
Note: NFC stands for non-financial corporations.

The final sample is an unbalanced panel of 369,526 distinct firms with 1,564,579 observations. On average, firms show up in the sample 4.2 times over 2010-2019. Approximately 23.7% of firms show up only once, whereas 6.7% are followed throughout the ten years. The number of firms in the final sample ranges from 123,217 in 2012 to 193,465 in 2019.

A preliminary analysis of our IES sample highlights a number of well-known features of the Portuguese economy. The sample is clearly dominated by micro firms[9], which represent more than 80% of observations in 2010-2019, but account for less than 16% of total income tax expenses (Figure 6). In contrast, large firms account for only 0.5% of the sample but are the most relevant taxpayers when it comes to CIT, making up for

9. For more details on the definition of the size categories, see footnote 3.

almost 45% of total income taxes paid in the period. Regarding sectoral classification[10], the majority of observations refers to firms in the services' sector (more than 70%). Manufacturing and construction represent around 13% and 11% of total observations, respectively. In terms of age, the average stands at 14.4 years. Firms with up to 5 years of activity represent 27% and firms with more than 20 years account for 23% of the total sample. These features remained virtually unchanged throughout the 2010-2019 period.

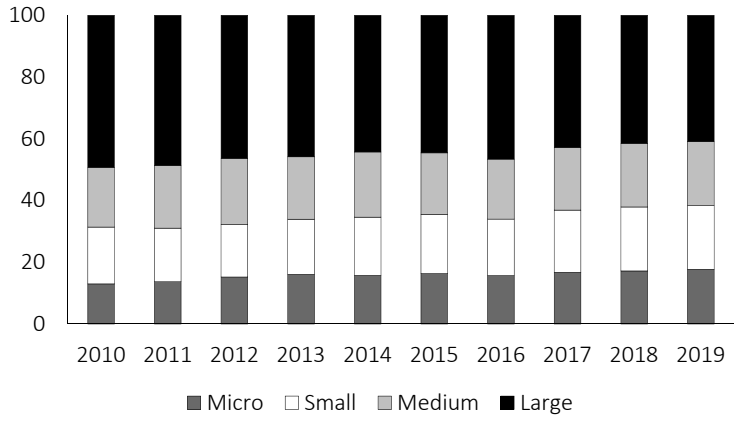


FIGURE 6: Percentage shares in total income tax expenses, by firm size category
Note: For more details on the definition of the size categories, see footnote 3.

There is wide variability across firms and sectors as regards important dimensions from a tax perspective (Figure 7). Firms operating in the utilities' sector feature the highest levels of profits (measured as EBITDA to total assets) and of labour productivity (proxied by gross value-added per worker). Along with firms in the primary sector, utilities' firms rank the highest as regards the leverage ratio (financial debt as a percentage of total assets) and capital intensity (share of tangible assets in total assets). In contrast, firms in the construction sector feature the lowest leverage ratio (following a sharp deleveraging process in the early 2010s) and are considerably less capital-intensive than their counterparts in other sectors. Further differences can also be noted across size categories and age groups. Micro firms feature the lowest leverage ratio and their share of tangibles in total assets is also the smallest across size categories. Regarding age, more mature firms appear to have, on average, lower profitability (scaled by total assets) and comparatively higher debt and capital intensity ratios.

10. The broad sectors are defined according to the sections of CAE rev.3. The primary sector comprises sections A (agriculture, forestry and fishing) and B (mining and quarrying) of CAE. Manufacturing refers to section C. The sector of utilities includes sections D and E and construction refers to section F. Services comprise all sections of CAE from G onwards.

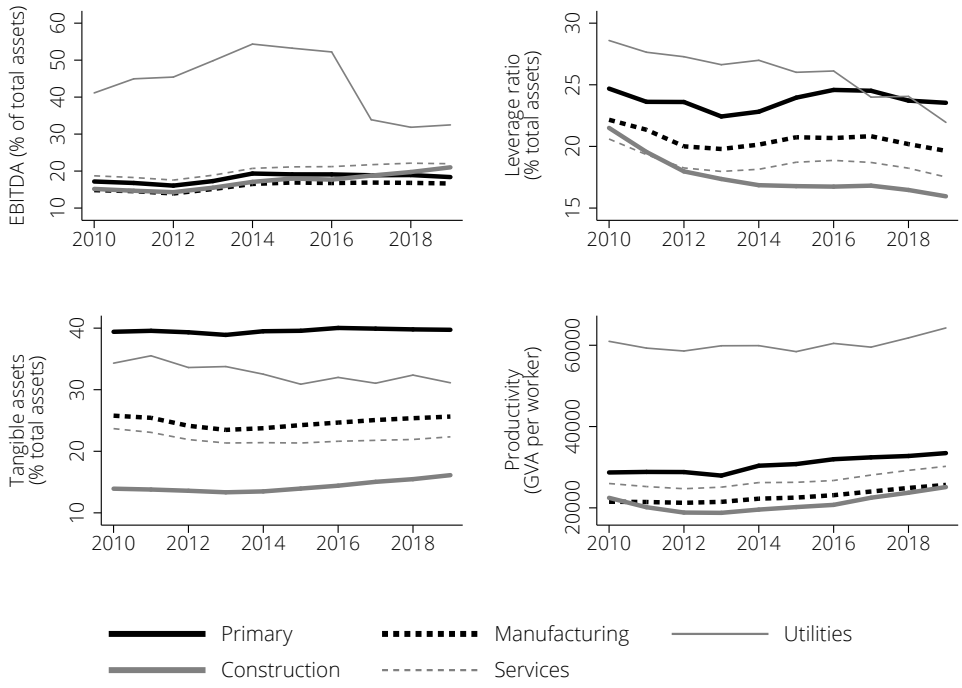


FIGURE 7: Characterisation of firms by sector of activity, 2010-2019
Note: For a definition of the sectors of activity, see footnote 10.

5. Exploratory analysis

This section provides an initial descriptive analysis of corporate effective income taxation using the database and metrics described in the previous section. In 2010, the average ETR1 (as measured on the basis of EBT) was 23.6%, and its distribution was fairly concentrated at relatively low levels (Figure 8). It stood below an average "weighted" statutory rate obtained by weighting the existing rates by the share of firms to which they apply. In 2012, the average ETR increased, coinciding with the lowering of the threshold for income subject to the State surcharge, together with an increase in the rates applicable to higher profits. In 2014, the general statutory rate was cut from 25 to 23%, while a reduced rate applicable up to a certain taxable income threshold was introduced for SMEs. Since then, the general rate was further brought down, to 21%, but the progressivitiy via the State surcharge increased. Hence, there was a gradual reversal of the previous increase in the effective tax burden and the average ETR broadly converged to the "weighted" statutory rate, standing at around 25% in recent years. Overall, the changes implemented in the last decade led to an increase in firms and taxable income subject to special rates and the distribution of ETR1 shifted to the right.

Panel C of Figure 8 provides a comparison of the "weighted" statutory tax rate and ETRs as computed on the basis of EBT and EBITDA. The ETR1 and ETR2 measures exhibit essentially the same evolution over time, though the former was more volatile in 2012-2014 reflecting the fact that firms' EBT dropped to a greater extent than EBITDA in 2012.

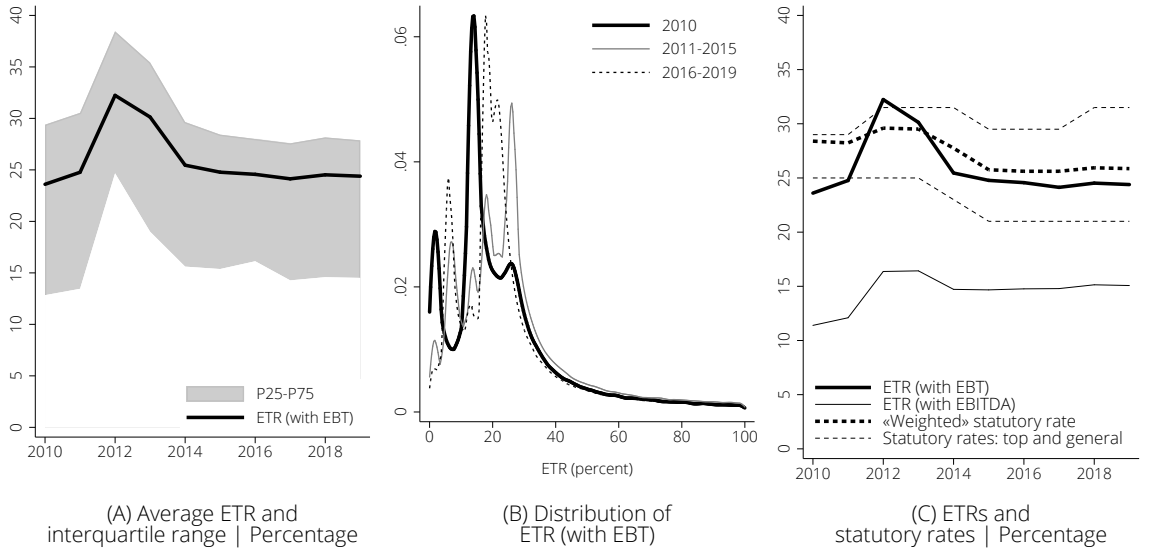


FIGURE 8: Effective tax rates: levels and distribution

Notes: The "weighted" statutory rate corresponds to an average of the different statutory rates (reduced, general, and general plus surcharges) weighted by the share of firms subject to each bracket in each year. EBT was used as a proxy of taxable income in the computations. Both the top and the "weighted" statutory rates include the representative local surcharge of 1.5%.

Taking EBITDA as the denominator when computing ETRs (ETR2) improves comparability across firms and sectors that feature distinct financing structures and capital intensity. ETR levels computed as such differ considerably across sectors and firms mirroring differences in various dimensions (Figure 9). Firms operating in services and in construction (which together make up for over 80% of the firms in our sample) face the highest tax burden along 2010-2019. In contrast, the lower levels of effective income taxation are observed in the primary sector. However, over the last decade, the evolution of the average ETR2 was essentially similar across sectors. Differences in the tax burden across size and age categories are less prominent[11] (Panels B and C of Figure 9). Still, it is worth highlighting that micro firms and those in the lowest age cohort seem to have been more affected by the 2012-2014 developments. Younger firms also feature, on average, a higher ETR than their older counterparts.

All in all, combined evidence in this and the previous sections suggests that the effective tax burden tends to be higher in sectors, size categories and age groups in which the shares of financial debt and tangible assets in total assets are lower. Similarly, higher debt and capital-intensity ratios appears to be associated with lower ETR as measured taking EBITDA in the denominator (ETR2). Indeed, our data suggests a decline in the effective tax burden as one moves up along the distributions of the leverage and

11. The results of a simple regression of firms' ETR2 on a set of broad sector dummies reveals that the average differences in effective taxation between sectors over this period are statistically significant. The same result applies to firms' size categories and age groups but the magnitude of the estimates is much smaller.

capital intensity ratios (Figure 10). The relation between the ETR and firms' assets or productivity is less clear-cut and the inspection of Figure 10 suggests the possibility of non-linear effects. These associations are further explored in the next section.

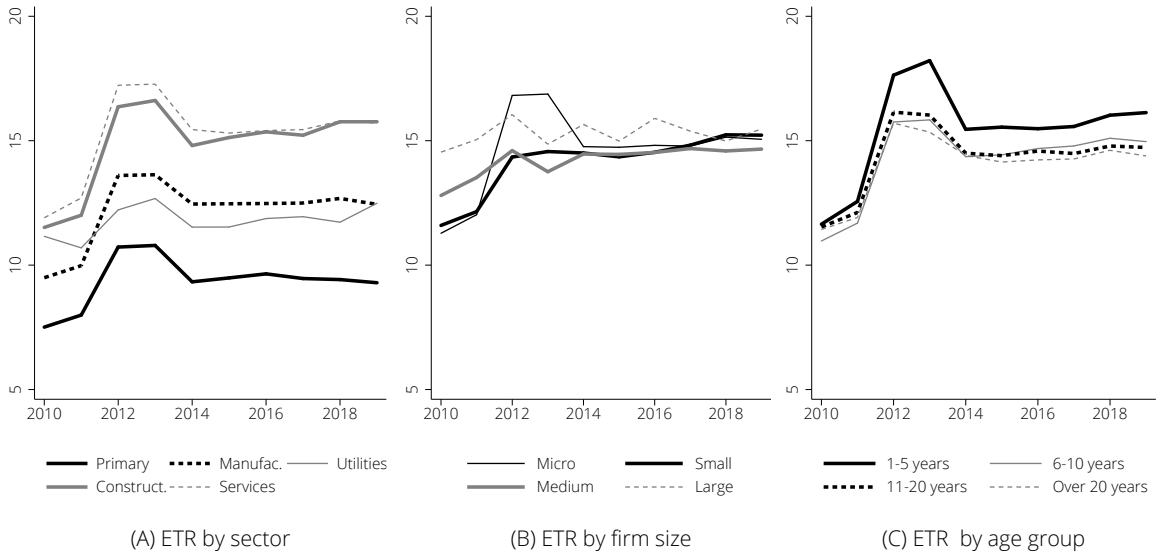


FIGURE 9: Effective tax rates (ETR2) across sectors, size categories and age groups | Percentage

Notes: Effective income taxation measured as the ratio of current income tax expenses (including autonomous taxation) to EBITDA (ETR2). For more details on the definition of the size categories, see footnote 3. For a definition of the sectors of activity, see footnote 10.

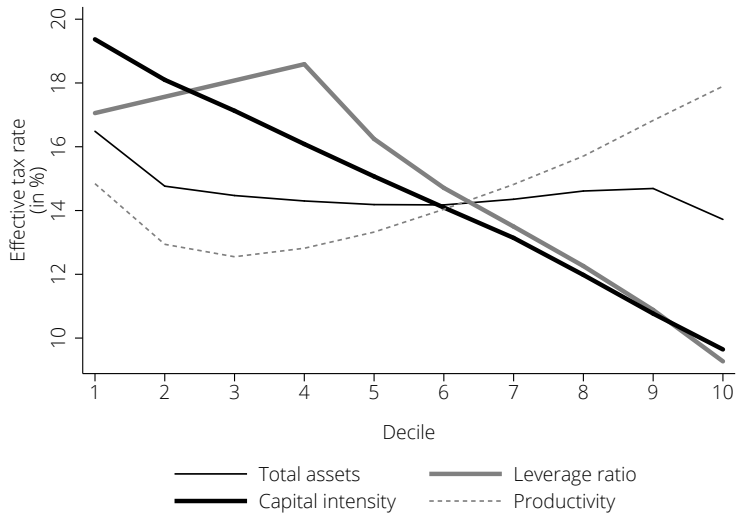


FIGURE 10: Effective tax rates (ETR2) across the distributions of some firms' characteristics

Notes: ETR2 is the ratio of current income tax expenses (including autonomous taxation) to EBITDA. The leverage ratio is measured as financial debt over total assets, capital intensity is computed as the book value of tangible assets in percentage of total assets, and labour productivity is defined as gross value-added per worker.

6. Regression analysis

The exploratory analysis of the previous sections is strongly suggestive of the key results emerging from the data. In this section, we further test differences in ETRs among Portuguese firms in a multivariate framework to isolate several aspects considered before. As mentioned, regression results are obtained only for firms with both strictly positive pre-tax incomes and tax expenses. Moreover, these regression estimates result from a simple empirical model, thus no causal inference can be drawn from them and extrapolation for the whole economy and for other ETR metrics should be avoided.

Motivated by the related literature on corporate effective taxation and by the descriptive analysis of the previous sections, we estimate the following equation at the firm-year level for the period 2010-2019:

$$Y_{it} = \beta_0 + \beta_1 size_{it} + \beta_2 size_{it}^2 + \beta_3 lever_{it} + \beta_4 capint_{it} + \beta_5 prod_{it} + \beta_6 prod_{it}^2 + \gamma_i + \gamma_t + \varepsilon_{it},$$

where Y_{it} is the dependent variable of interest, the effective tax rate of firm i in year t. $size$ is firm size proxied by total assets, $lever$ is the leverage ratio measured as financial debt over total assets, $capint$ is capital intensity defined as the book value of tangible assets scaled by total assets, and $prod$ is labour productivity measured as gross value-added per worker. $size^2$ and $prod^2$ are the quadratic terms of size and productivity, respectively, which were included to capture possible non-linear relations. All covariates were logarithmised. γ_i are firm fixed effects and γ_t are year fixed effects. ε_{it} is the error term. Robust standard errors are clustered at the firm-level. As an outlier treatment, we winsorised the top and bottom 1 percentiles of the covariates.

Table 2 presents our baseline results using ETR2 as the dependent variable, i.e., the ratio of current tax expenses (including autonomous taxation) over EBITDA. Including deferred income taxes expenses in the numerator makes almost no difference in the estimated relations between effective income taxation and firms' characteristics in Portugal.[12] The measure ETR1, which uses EBT in the denominator, does not take fully into account to what extent firms can minimise their tax liabilities with actions that operate through the tax base, given that interest payments and depreciations are considered in the denominator. Hence, we opted to focus the regression analysis of this section on ETR2. However, the point estimates of these regressions should not be interpreted as the causal effect of the covariates in a firm´s tax burden, as they represent only correlations with this specific ETR metric.

Our baseline estimates for the total economy are reported in column (1) of Table 2. The results suggest the existence of a non-linear relationship between firms' size and effective taxation: as firms grow, they experience an increase in the ETR but at a decreasing rate and, after a certain point, a further increase in size can even be associated with smaller ETRs. The turning point of the marginal effects occurs at the 60th percentile of the size distribution. The positive linear effect probably reflects the progressivity

12. All regressions were also run using vectors of firm, sector (at a 2-digit level), district and year fixed effects and the estimates remain unchanged. The full set of detailed results is available from the authors upon request.

of the Portuguese CIT system as described in Section 3. A rational for the negative quadratic coefficient is that the largest firms are likely to have greater scope for tax planning and for the adoption of accounting practices that minimise their effective taxation.

The opposite results are found for labour productivity: a negative linear coefficient and a positive quadratic coefficient, with a turning point at the 40th percentile. For the negative association at the lower part of the productivity distribution, it can be argued that as firms become more productive they have greater resources to engage in tax planning and organise their activities to achieve optimal tax savings. The positive quadratic term may reflect the fact that there is a limit for what firms can do to reduce their tax burdens, no matter how productive they are.

Firms' financing and investment decisions are likely to be correlated with their ETRs because the tax code stipulates a differential treatment to equity versus debt financing and to current versus non-current assets. Given that interest expenditure is tax deductible up to certain threshold, while dividends are not, and capital injections by shareholders benefit only from a limited notional deduction, firms with higher leverage are expected to have lower effective taxation. The negative and significant relation of $lever$ with ETR2 confirms this hypothesis. For the capital intensity measure ($capint$), the results also indicate that it has a negative association with ETR2. This finding is consistent with the preferential tax treatment for firms that invest in their fixed assets, given the deductibility of assets' depreciation and the tax incentives for investment provided by law.

There is evidence in the literature that firms' effective taxation can also depend on their sector of activity (e.g., Nicodème 2002). First, specific sectors may benefit disproportionately from some preferential tax treatments. Second, some attributes of firms, like size or capital intensity, can vary systematically by sector. In fact, the analysis of the previous two sections revealed important differences in terms of both firms' characteristics and ETR levels by sector. Hence, individual regressions were estimated by broad sectors in order to better understand if the associations obtained for the total sample are still observable. The results are reported in columns (2) to (6).

In general, the signs and magnitudes of the estimates do not vary considerably across sectors. The coefficients associated with leverage and capital intensity are always negative and statistically significant for all sectors. The estimates of the other covariates change more by sector of activity. In the primary sector, the relation of ETR2 with firms' dimension is not statistically significant, but both coefficients of productivity are sizeable. The effective taxation of firms in the utilities' sector has a strong association with size and productivity. However, the sample of firms in this sector is small, which can make the estimates more sensitive to extreme observations. On the contrary, in the construction sector, the linear coefficient of size is statistically non-significant. Comparing the estimates of the regressions for the manufacturing industry and for services, there is a higher association of ETR2 with firms' size and productivity in the manufacturing industry.

	Total (1)	Primary (2)	Manufacturing (3)	Utilities (4)	Construction (5)	Services (6)
$size$	1.445***	-0.958	3.939***	8.317***	-0.735	0.994***
	(0.197)	(0.853)	(0.500)	(3.185)	(0.588)	(0.244)
$size^2$	-0.058***	0.009	-0.172***	-0.354***	0.062**	-0.039***
	(0.008)	(0.035)	(0.020)	(0.125)	(0.025)	(0.010)
$lever$	-0.578***	-0.548***	-0.680***	-0.569***	-0.659***	-0.546***
	(0.010)	(0.038)	(0.026)	(0.132)	(0.031)	(0.012)
$capint$	-1.202***	-1.493***	-1.472***	-1.024***	-1.112***	-1.172***
	(0.014)	(0.071)	(0.042)	(0.260)	(0.039)	(0.016)
$prod$	-5.672***	-11.875***	-9.975***	-10.548**	-4.346***	-4.878***
	(0.310)	(1.018)	(1.024)	(4.500)	(0.908)	(0.366)
$prod^2$	0.294***	0.661***	0.568***	0.596***	0.223***	0.242***
	(0.016)	(0.051)	(0.052)	(0.219)	(0.046)	(0.018)
N	1,477,125	55,234	196,268	6,181	164,058	1,052,860
Adj. R^2	0.519	0.515	0.523	0.573	0.452	0.519

TABLE 2. Effective income taxation (ETR2) and firms' characteristics, total and by broad sector of activity, 2010-2019

Notes: The dependent variable is ETR2 defined as the ratio of current income tax expenses (including autonomous taxation) over EBITDA. For a definition of the sectors of activity, see footnote 10. All regressions include a constant and the vectors of firm and year fixed-effects. All covariates are in log form. See the main text for more details. Standard errors in parenthesis are clustered at the firm level and are robust to heteroscedasticity. Stars indicate significance levels of 10% (*), 5% (**), and 1%(***).

7. Concluding remarks

This article analyses corporate income taxation in Portugal. The top corporate income tax (CIT) statutory rate in Portugal is very high when compared to international peers, particularly in the most recent period: Portugal stood in the top 3 amongst OECD and euro area countries in 2019. Presently, the top statutory rate is 31.5% and the degree of progressivity of CIT increased substantially in the last decade with the introduction and further reinforcement of a State surcharge. This offset the reduction of the general rate that was aligned with the international trend. Based on macro data, the implicit tax rate obtained using a National Accounts proxy for taxable income reached 23.5% in 2019. This value is around 25% below the top statutory rate, which provides a broad estimate for progressivity stemming from both the rate structure and tax benefits, incentives and deductions. Their overall impact in collection is very difficult to assess, given their high number and frequent changes. In terms of revenue, although CIT is the third biggest tax in Portugal, its weight in total tax burden or GDP is relatively small and rather stable, as, on average, in other EU countries (8.7% and 3%, respectively, over the last two decades).

The recourse to a large micro database allows a more detailed characterisation of CIT in Portugal in the 2010-2019 period. The analysis is based on the observed firm-level ETR, broadly defined as the ratio of income tax expenses to a metric of pre-tax profits. In this article we compute two measures of firms' ETR: using either EBITDA or EBT in the denominator. The first measure is more appropriate when investigating the relation with firms' characteristics as it isolates the effects of preferential tax treatments related to debt and investment in the numerator. The second measure is more suitable for benchmarking against statutory tax rates and macro implicit tax rates. Still, as EBT is merely a proxy for taxable income, caution is warranted in such comparisons. In the

more recent period, the average micro-based ETR (using EBT) supported by Portuguese non-financial firms hovers around 25%, slightly below a weighted average of statutory rates. Some well-known features of CIT collection in Portugal are also evident in our results. Micro firms, which represent more than 80% of observations, account for less than 16% of total tax expenses in the sample. In contrast, large firms, which account for only 0.5% of the sample, are the most relevant taxpayers, making up for almost 45% of total income taxes paid.

In the literature, there is ample evidence of a relation between effective corporate income taxation and different firms' characteristics but results are not consensual. Our regression estimates, which have no intention to establish causality, suggest the existence of a negative association between ETRs and both financial leverage and capital intensity. This confirms expectations, given that interest outlays and depreciation of fixed assets have a favourable tax treatment. Regarding firms' size and labour productivity, the results point to the existence of non-linear relations with ETR. The fact that the largest firms may have more capacity to tax planning that minimises their tax burden can work as a rationale for this result. In spite of the different levels of ETRs across sectors, the estimates do not vary considerably by sector of activity.

Understanding the functioning of CIT systems is crucial in a world marked by globalisation and digitalisation, where tax competition and firms' tax planning practices have been gaining prominence. A major step towards limiting international tax avoidance was taken in 2013 by the OECD with the Base Erosion and Profit Shifting (BEPS) initiative. The BEPS process was an enormous effort and culminated with a multilateral convention signed by 89 countries between July 2017 and August 2019 (with the notable exception of the United States). Many of the OECD BEPS proposals have been implemented at the EU level through the Anti Tax Avoidance Directive (ATAD), with some concrete application to Member States from 2019 onwards. Very recently, an historical agreement on a minimum tax deal was adopted at the G7 and G20 meetings. The proposal is anchored in two pillars: i) The largest and most profitable multinationals will be required to pay CIT in the countries where they operate and not just where they have their headquarters (based on formulary apportionment); ii) A global minimum rate will be set that ensures multinationals pay corporate taxes of at least 15% in each country they operate. The actual implementation of this deal will certainly limit tax avoidance and reduce tax competition and the associated race to the bottom, but ensuing negotiations and extension to all firms will most likely be a lengthy process. Still, a minimum CIT rate of 15% is rather low and illustrative of the sort of challenges that Portugal's corporate taxation will face in the near future.

References

Bartolini, David (2018). "Firms at the productivity frontier enjoy lower effective taxation." OECD Economics Department Working Paper 1475, OECD.

Bastani, Spencer and Daniel Waldenström (2020). "How Should Capital Be Taxed?" *Journal of Economic Surveys*, 34(4), 812–846.

Beer, Sebastian, Ruud de Mooij, and Li Liu (2020). "International Corporate Tax Avoidance: A Review of the Channels, Magnitudes, and Blind Spots." *Journal of Economic Surveys*, 34(3), 660–688.

Bessa, Ana Teixeira (2016). "A taxa efetiva de imposto e a reforma do IRC de 2013." Dissertação de mestrado 201540290, ISCTE.

Costa, Ana, Francisco Vitorino Martins, and Elísio Brandão (2012). "Effective Tax Rate in Portugal: Determinants and Financial Reporting Impact." SSRN Working Paper 2180032.

Delgado, Francisco J., Elena Fernández-Rodríguez, and Antonio Martínez-Arias (2014). "Effective Tax Rates in Corporate Taxation: a Quantile Regression for the EU." *Engineering Economics*, 25(5), 487–496.

Delgado, Francisco J., Elena Fernández-Rodríguez, and Antonio Martínez-Arias (2018). "Corporation effective tax rates and company size: evidence from Germany." *Economic Research-Ekonomska Istraživanja*, 31(1), 2081–2099.

Devereux, Michael P. and Rachel Griffith (1998). "The taxation of discrete investment choices." IFS Working Papers W98/16, Institute for Fiscal Studies.

Devereux, Michael P. and Rachel Griffith (2003). "Evaluating Tax Policy for Location Decisions." *International Tax and Public Finance*, 10(2), 107–126.

Dyreng, Scott D., Michelle Hanlon, Edward L. Maydew, and Jacob R. Thornock (2017). "Changes in corporate effective tax rates over the past 25 years." *Journal of Financial Economics*, 124(3), 441–463.

Egger, Peter, Simon Loretz, Michael Pfaffermayr, and Hannes Winner (2009). "Firm-specific forward-looking effective tax rates." *International Tax and Public Finance*, 16(6), 850–870.

Fernández-Rodríguez, Elena, Roberto García-Fernández, and Antonio Martínez-Arias (2021). "Business and institutional determinants of Effective Tax Rate in emerging economies." *Economic Modelling*, 94(C), 692–702.

Grupo de Trabalho para o Estudo dos Benefícios Fiscais (2019). *Os benefícios fiscais em Portugal*. Cadernos de Ciência e Técnica Fiscal, Centro de Estudos Fiscais e Aduaneiros.

Gupta, Sanjay and Kaye Newberry (1997). "Determinants of the variability in corporate effective tax rates: Evidence from longitudinal data." *Journal of Accounting and Public Policy*, 16(1), 1–34.

Hanlon, Michelle and Shane Heitzman (2010). "A review of tax research." *Journal of Accounting and Economics*, 50(2-3), 127–178.

Hsieh, Yao-Chih (2012). "New evidence on determinants of corporate effective tax rates." *African Journal of Business Management*, 6(3), 1177–1180.

Janssen, Boudewijn (2005). "Corporate Effective Tax Rates in the Netherlands." *De Economist*, 153(1), 47–66.

Lazăr, Sebastian (2014). "Determinants of the Variability of Corporate Effective Tax Rates: Evidence from Romanian Listed Companies." *Emerging Markets Finance and Trade*, 50(sup4), 113–131.

Liu, Xing and Shujun Cao (2007). "Determinants of Corporate Effective Tax Rates: Evidence from Listed Companies in China." *Chinese Economy*, 40(6), 49–67.

Mascagni, Giulia and Andualem Mengistu (2019). "Effective tax rates and firm size in Ethiopia." *Development Policy Review*, 37(S2), O248–O273.

Nicodème, Gaëtan (2001). "Computing effective corporate tax rates: comparisons and results." European Economy - Economic Papers 153, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Nicodème, Gaëtan (2002). "Sector and size effects on effective corporate taxation." European Economy - Economic Papers 175, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Nicodème, Gaëtan, Antonella Caiumi, and Ina Majewski (2018). "What Happened to CIT collection? Solving the Rates-Revenues Puzzle." CEPR Discussion Paper 13385, Centre for Economic Policy Research (CEPR).

Omer, Thomas C., Karen H. Molloy, and David A. Ziebart (1991). "Measurement of effective corporate tax rates using financial statement information." *Journal of the American Taxation Association*, 13(1), 57–72.

Plesko, George A. (2003). "An evaluation of alternative measures of corporate tax rates." *Journal of Accounting and Economics*, 35(2), 201–226.

Praça, Ana Oliveira (2018). "Determinantes de Corporate Tax Planning: O Caso Português." Dissertação de mestrado 202215288, FEP.

Richardson, Grant and Roman Lanis (2007). "Determinants of the variability in corporate effective tax rates and tax reform: Evidence from Australia." *Journal of Accounting and Public Policy*, 26(6), 689–704.

Shackelford, Douglas A. and Terry Shevlin (2001). "Empirical tax research in accounting." *Journal of Accounting and Economics*, 31(1-3), 321–387.

Stamatopoulos, Ioannis, Stamatina Hadjidema, and Konstantinos Eleftheriou (2019). "Explaining corporate effective tax rates: Evidence from Greece." *Economic Analysis and Policy*, 62(C), 236–254.

Topa, Pedro Sousa (2018). "Capital Origin and Financial Determinants of Effective Tax Rate." Dissertação de mestrado 202217906, FEP.

Wang, Fangjun, Shuolei Xu, Junqin Sun, and Charles P. Cullinan (2020). "Corporate Tax Avoidance: A Literature Review And Research Agenda." *Journal of Economic Surveys*, 34(4), 793–811.

Wilde, Jaron H. and Ryan J. Wilson (2018). "Perspectives on Corporate Tax Planning: Observations from the Past Decade." *Journal of the American Taxation Association*, 40(2), 63–81.

Non-technical summary

January 2022

Workforce skills and firm productivity

Joana Cima, Ana Catarina Pimenta, Miguel Portela and Marta Silva

Firms' productivity is one of the key features to assess the countries' economic performance and resilience. In Portugal, at the aggregate level, there was a stabilization in firms' value-added per worker in the last two decades. The workforce skills have been pointed out in the literature as one of the key determinants of the firms' outcomes. An analysis at the firm-level contributes to identifying the potential drivers of the productivity dynamics.

This paper analyses the relationship between firm's productivity and two moments of the workers' skill distribution, the average and the dispersion. Workers' skills should capture several individual characteristics, ranging from formal education to general aptitudes obtained in the labour market, combined with innate or developed capacities, often unobservable. To capture these multiple dimensions, we use a skill index that comprises years of schooling, age and worker's unobserved ability, estimated from a wage equation.

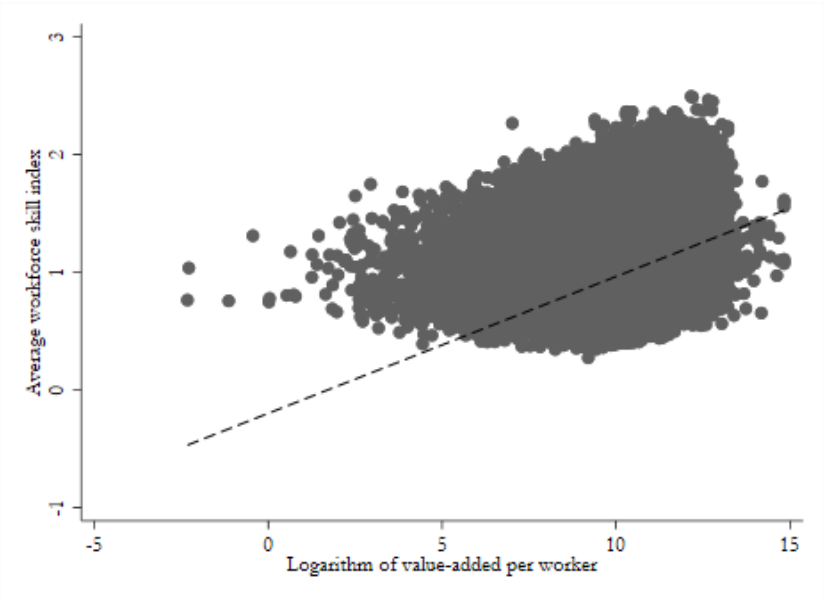


FIGURE 1: Relationship between firm's productivity and average workforce skill index

Sources: Own computations using data from *Quadros de Pessoal* and *Sistema de Contas Integradas das Empresas* (2006-2018).
Note: The dashed line represents the fitted values.

Figure 1 shows a positive correlation between the average workers' skill index and firm's productivity, measured by the value-added per worker, for the period 2006-2018. Our estimates confirm this finding and show that the standard deviation of workers'

skill index, conditional on its average, is negatively associated with firm's productivity. Our results are in line with the literature showing that firms with a higher productivity level also hire workers of similar ability, education and age.

We assess the sensitivity of our results to an alternative measure of productivity (i.e., the value-added per hour worked), different proxies for skill (i.e., education, age, unobserved ability, and a composite index with the two observed characteristics) and different measures of within-firm skill heterogeneity (i.e. percentile ratios, coefficient of variation and variance). The estimates remain qualitatively similar in all specifications.

Finally, we evaluate whether the estimates change across the conditional productivity distribution. Relying on quantile regression, we conclude that there is an increasing positive association between average skill and firm's productivity across the conditional productivity distribution. In contrast, the negative relation between the workforce skill heterogeneity and the firm's productivity is stable.

Workforce skills and firm productivity

Joana Cima
NIPE/Universidade do Minho

Ana Catarina Pimenta
Banco de Portugal

Miguel Portela
NIPE/Universidade do Minho
IZA, Bonn

Marta Silva
Banco de Portugal

January 2022

Abstract
We study the relationship between workforce skills and firm's productivity using Portuguese data for the period 2006-2018. We use a multi-dimensional index that incorporates worker's education, age, and unobserved ability to measure workers' skill. The analysis shows that the average skill of the workforce is positively associated with productivity. However, we find a negative relationship between the dispersion of the workforce skills and the value-added per worker. We also estimate quantile regressions and observe that the positive association between average skill and productivity is increasing across the conditional productivity distribution, while the negative association with skill dispersion is stable. (JEL: C23, J24)

Keywords: labour productivity, skill index, quantile, workforce heterogeneity.

1. Introduction

The differences in productivity levels across firms have been a central theme in economic research (Syverson 2011). The literature has pointed out several internal sources to the firms for these differences, including product innovation, investments in information technology and R&D, firm structure decisions, or human resource management practices, such as pay incentives, teamwork and investment in training (e.g., Acemoglu and Pischke 1998, Ichniowski *et al.* 1997). This article contributes to the literature that assesses how the workforce skill composition impacts productivity (e.g., Ilmakunnas and Ilmakunnas 2011).

If, on the one hand, a more heterogeneous workforce composition can positively affect productivity through the knowledge transfer effect, on the other hand, it can lead

Acknowledgements: We would like to thank Fernando Alexandre, João Amador, António Antunes, Hélder Costa, Paulo Guimarães, Attila Gyetvai, Pedro Duarte Neves, Pedro Portugal, Hugo Reis and participants at NIPE's seminar and an internal seminar of the Economics and Research Department of Banco de Portugal for helpful comments and suggestions. The analyses, opinions and conclusions expressed herein are the sole responsibility of the authors and do not necessarily reflect the opinions of Banco de Portugal or the Eurosystem. This paper is financed by National Funds of the FCT - Portuguese Foundation for Science and Technology, projects UIDB/03182/2020 and PTDC/EGE-ECO/29822/2017 ("It's All About Productivity: contributions to the understanding of the sluggish performance of the Portuguese economy").

E-mail: joana.cima@eeg.uminho.pt; apimenta@bportugal.pt; miguel.portela@eeg.uminho.pt; msilva@bportugal.pt

to opportunistic or free-riding behaviour from the part of the labour force, impacting negatively on average productivity (Hamilton *et al.* 2003). Although the literature has presented several approaches to assess which is the dominant effect on productivity, there are some limitations that this article intends to address.

Hamilton *et al.* (2003) analyse how teams' heterogeneity, computed considering the ratio between the maximum and the minimum individual productivity levels among all team members, impacts productivity. The authors conclude that the introduction of teams by the firm increases worker productivity and that more heterogeneous teams are, on average, more productive. Additionally, Mas and Moretti (2009) argue that not only do the most productive workers directly contribute to an increase in firms' output, as they also contribute indirectly through the increase of their co-workers productivity. However, since those studies are focused on specific firms, their findings may not be valid across a broad set of firms or sectors.

The literature that analyses the relationship between the workforce composition of a firm and its performance typically uses workers' observed characteristics, such as age, gender, educational attainment and race (e.g., Haltiwanger *et al.* 1999, Hellerstein *et al.* 1999, Mendes *et al.* 2010, Pfeifer and Wagner 2014). Although the observed characteristics may explain differences in productivity across workers and firms, it is restrictive to assume they entirely account for worker and firm idiosyncrasies. It is plausible that there is unobserved heterogeneity for both workers and firms, which conditions individual and firm-level productivity. Bender *et al.* (2018) and Iranzo *et al.* (2008) are examples of the literature that addresses this limitation and uses a measure that is not observed by standard variables. Both articles estimate the skill as the worker's fixed effect obtained from the AKM model – which decomposes wages into worker and firm fixed effects – proposed by Abowd *et al.* (1999). This worker specific component measures wages due to the worker's pure ability, regardless of the firm and net of the personal time-variant characteristics included as controls. Bender *et al.* (2018) use the average of the estimated worker fixed effects as a proxy for the average human capital at the firm and find that firms with a more skilled workforce have higher productivity. Torres *et al.* (2018) also rely on worker fixed effects to proxy for the workforce quality and highlight the importance to consider job title (i.e. occupational tasks) fixed effects as another source of labour heterogeneity in the production function.

Although the dispersion measure most commonly used in the literature is the standard deviation of the workers' skill level, several studies have proposed alternative measures. Kremer and Maskin (1996) propose a segregation index by skill proxied by wages, education or occupational categories. Ilmakunnas and Ilmakunnas (2011) and Parrotta *et al.* (2014) add other measures of dissimilarity and Herfindahl diversity indexes to infer how the dispersion in specific workforce characteristics affects firm's productivity. Ilmakunnas and Ilmakunnas (2011) find that age diversity impacts positively on total factor productivity (TFP), while educational diversity has a negative impact. Conversely, Parrotta *et al.* (2014) find that educational diversity significantly enhances firm productivity, while ethnic and demographic heterogeneity have the opposite effect. Finally, Iranzo *et al.* (2008) decompose the total skills dispersion into within-firm and between-firm components, showing that skill dispersion within

occupational status groups (production and nonproduction workers) is positively related to firm's productivity. In contrast, the dispersion between these groups is negatively related to firm's productivity.

In this paper, and given that the one-dimensional skill measures may have limitations in capturing the overall impact of workforce composition on productivity, we use the multi-dimensional skill index developed by Portela (2001). This index measures worker's skill combining several observed components, such as schooling, age and worker's unobserved ability in line with Bender *et al.* (2018) and Iranzo *et al.* (2008). In this regard, our strategy compares with that used by Rocha *et al.* (2019) which also uses this skill index to evaluate the effect of the initial workforce average quality on firm's performance.

Our analysis further explores the relationship between firm's productivity and two moments of the workers' skill distribution, the average and the dispersion. We use the standard deviation of the workforce skills computed within the firm to assess the heterogeneity in each year. Our measure of firms' productivity is the value-added per worker. [1]

This article presents novel evidence for the Portuguese economy about the relationship between the workforce composition and firm's productivity. Using a very rich linked employer-employee dataset, we compute a composite index to study the relationship between workforce skills and productivity not only at the mean but also across the productivity distribution. We find a positive and significant relationship between the average workforce skills and firm's productivity. Moreover, this relation appears more relevant at the top than at the bottom half of the conditional productivity distribution. We also report a negative association between a more heterogeneous workforce and value-added per worker, conditional on the average workers' skill, which is relatively stable across the conditional productivity distribution. Our results align with the literature and provide additional evidence on the importance of considering the complementarity between several dimensions of worker's skill when assessing its effects on firm outcomes.

We assess the sensitivity of our results to an alternative measure of productivity (i.e., the value-added per hour worked), different proxies for skill (i.e., composite index with education and age, and each variable included in the skill index individually) and different measures of within-firm skill heterogeneity (i.e. percentile ratios, coefficient of variation and variance). The estimates remain qualitatively similar in all specifications.

The article is organised as follows. Section 2 describes the worker skill index and the heterogeneity measure used in the analysis, Section 3 introduces our econometric methodology. Section 4 describes the main data sources and presents some descriptive statistics. Then, Section 5 discusses the main results and Subsection 5.2 assess the sensitivity of our findings. Section 6 concludes.

1. The option of not using Total Factor Productivity as a proxy for productivity is because the data do not contain a precise measure of capital stock for the entire period under analysis.

2. Skill index and workforce heterogeneity

The search for the most accurate measure of worker's skill has been at the core of the most recent debates in the empirical labour economics. It should capture several individual characteristics, ranging from formal education, to general aptitudes obtained in the labour market, combined with innate or developed capacities, which are often unobservable.

In the vein of Portela (2001), we construct an aggregate skill index which will be at the core of our empirical analysis. The main advantages of this index, over the ones typically used in the literature, are that it allows us to integrate in a composite measure several skill dimensions, as well as variables measured in different units.

We compute the worker skill index, $Skill_{it}$, using the dimensions education, age and (unobserved) ability, according to the following specification,

$$Skill_{it} = a_{it,school} \times a_{it,age} \times a_{it,unobserved} \tag{1}$$

where the subscripts i and t denote the worker and the year, respectively. Each skill component $a_{it,school}$, $a_{it,age}$ and $a_{it,unobserved}$ represents the worker's position in the education, age and (unobserved) ability distribution in each year, respectively.

To compute each component we consider the cumulative logistic distribution, corrected by the factor 0.5. This functional form ensures that the main changes occur around the mean, while changes far from the mean have smaller impacts. The correction factor 0.5 ensures that each component is bounded between 0.5 and 1.5. The specification for each component is given by equations (1a), (1b) and (1c).

The contribution of education to the skill index is defined by,

$$a_{it,school} = 0.5 + \frac{e^{(school_{it} - mschool_t)/sschool_t}}{1 + e^{(school_{it} - mschool_t)/sschool_t}} \tag{1a}$$

where $school_{it}$ corresponds to the years of schooling of worker i in year t. The $mschool_t$ and $sschool_t$ correspond to the average and the standard deviation of schooling in year t, respectively. By definition, $a_{it,school}$ is higher than 1 when the number of years of schooling is above the average in the economy, while years of schooling below the average are associated with a value of less than 1.

Similarly, age's component is computed as,

$$a_{it,age} = 0.5 + \frac{e^{(age_{it} - mage_t)/sage_t}}{1 + e^{(age_{it} - mage_t)/sage_t}} \tag{1b}$$

where age_{it} corresponds to the age of worker i in year t. $mage_t$ and $sage_t$ correspond to the average and the standard deviation of age in year t, respectively. As before, a worker older than the average in the economy has a value for $a_{it,age}$ greater than 1.

Finally, worker's (unobserved) ability contribution to this estimated overall skill is formulated as,

$$a_{it,unobserved} = 0.5 + \frac{e^{(FE_i - mFE_t)/sFE_t}}{1 + e^{(FE_i - mFE_t)/sFE_t}} \tag{1c}$$

where FE_i corresponds to the unobserved skill of worker i. mFE_t and sFE_t correspond to the average and standard deviation of the unobserved ability in year t, respectively.

To obtain the worker's unobserved component we estimate a wage equation with high-dimensional fixed effects:

$$wage_{ift} = \psi + X'_{ift}\varphi + \tau_i + \mu_f + \lambda_t + \omega_{ift} \tag{2}$$

where $wage_{ift}$ corresponds to the logarithm of real hourly wage for worker i in firm f and year t. X is a vector with the time-varying worker's observed characteristics (schooling years, a second order polynomial on both age and tenure) and firm's observed characteristics (logarithm of firm size and its square); τ_i is the worker fixed effect, μ_f is the firm fixed effect; λ_t corresponds to year-dummies and ω_{ift} is the usual white noise error-term.[2] We use the estimated worker's fixed effects as a proxy for FE_i in equation (1c). This variable represents the worker's unobserved ability.

Having computed the worker's skill index, $Skill_{it}$, we are able to measure the workforce skills and heterogeneity for each firm/year. Table 1 summarizes some of the alternative measures of workforce heterogeneity proposed in the literature. In this article we use the within-firm standard deviation of the skill index to capture the firm's skill diversity.

3. Econometric methodology

We estimate the following regression model to assess the impact of the average and dispersion of the workforce skills on firms' productivity,

$$y_{ft} = \alpha + \bar{s}'_{ft}\gamma + \theta'_{ft}\delta + X'_{ft}\beta + \eta_f + \vartheta_t + \varepsilon_{ft} \tag{3}$$

where y_{ft} is the logarithm of the gross value-added per worker of firm f in year t. $\bar{s}_{ft}$ and θ_{ft} represent the average and the standard deviation of the skill index for firm f in year t, respectively. The parameters of interest are γ and δ which capture the effect of the average and dispersion of the workforce skills on firms' productivity, respectively.

The control variables in X_{ft} include a second order polynomial of firm size, measured by the logarithm of the number of workers, the share of part-time workers, the share of female workers, and a second order polynomial of the average firm tenure. The model

2. The model is estimated using the algorithm of Guimarães and Portugal (2010) through the Stata command *reghdfe* (Correia 2016). To identify the worker fixed effect we restrict the data to the largest connected set of workers and firms dropping approximately 0.4% of the observations. We report the estimates of this model in Table A1 and present the density of the worker fixed effects in the Figure A1 of the Appendix. Note that the age coefficient is not identified due to the inclusion of worker fixed effects and year dummies.

Papers	Measure of heterogeneity
Hamilton *et al.* (2003)	Ratio between the maximum and the minimum individual productivity levels among all team members
Pfeifer and Wagner (2014); Haltiwanger *et al.* (1999)	Share of workers by category (e.g., age, gender, education, qualification)
Kremer and Maskin (1996)	Segregation index equal to 0 if all firms have the same workforce skill composition and 1 in the case of complete segregation. Skill is measured with observed variables, such as wages, education, or occupational categories
Ilmakunnas and Ilmakunnas (2011)	Standard deviation and dissimilarity, variety and diversity indexes for age and education
Parrotta *et al.* (2014)	Herfindahl indexes to measure the cultural, educational and demographic (age and gender) diversity
Iranzo *et al.* (2008)	Total within-firm skill dispersion decomposed into within and between-occupations. The skill is measured by the worker's fixed effect obtained from a wage equation

TABLE 1. Measures of heterogeneity discussed in the literature

also includes year-dummies (ϑ_t) to account for the macroeconomic conditions and firm fixed effects (η_f) to control for time-invariant unobserved factors that are specific to the firm and can impact productivity. This term also helps to mitigate the potential bias arising from the fact that the firm may endogenously select the optimal workforce mix to maximize productivity (e.g., Parrotta *et al.* 2014). ε_{ft} is an *i.i.d.* error term.

This specification allows us to conclude about the effect of the workforce skills composition on the productivity of the average firm. However, this effect may differ across the productivity distribution. In order to assess whether the effect is heterogeneous, we expand our analysis by estimating the specification above at selected quantiles of the conditional firm productivity distribution using the Method of Moments Quantile Regression estimator proposed by Machado and Santos Silva (2019). As argued by the authors, this approach has the advantage of allowing the fixed effects to have different effects over the conditional productivity distribution instead of being just a location shift as most of the other methods available.

4. Data

4.1. Data sources

The main data source of this article is the linked employer-employee data *Quadros de Pessoal* (QP) collected by the Portuguese Ministry of Labour, Solidarity and Social Security since the 1980s. The report of these data is mandatory for all Portuguese firms with at least one employee. Besides the high coverage, this dataset provides detailed information at the firm and establishment-level (location and main sector of activity, for example) and at the worker-level (such as age, gender, schooling, wage, occupation, tenure and hours of work) with reference to the month of October.

We match this dataset with *Sistema de Contas Integradas das Empresas* (SCIE), which provides economic and financial information for non-financial firms operating in Portugal. This dataset is collected through the Simplified Corporate Information since 2006 and compiled by Statistics Portugal. These data report to the whole fiscal period and allows us to compute the value-added per worker as a proxy for firm's productivity. Since this information is only available for corporations we restrict the analysis to this type of firm. Both QP and SCIE provide unique identifiers that allow us to match them and follow the same firm over time.

Our sample covers the firms located in Mainland Portugal for the period between 2006 and 2018. The least representative sectors are excluded.[3] To calculate the skill measures of the workforce we consider employees with non-missing information on the main variables of interest, aged between 16 and 64 years old, and with contracted weekly hours of work between 10 and 40. Since our study focuses on skill heterogeneity at the firm-level, we only consider the observations of the firms with at least five employees.

The final panel dataset includes information for 136,709 unique firms for the period 2006-2018. Table 2 describes the variables and the corresponding data sources.

4.2. Descriptive statistics

Table A2 in the Appendix presents summary statistics for the variables included in the analysis for the period 2006–2018. These statistics are obtained in the sample of our main econometric specification, i.e., without missing values in the variables included in the regression (first column of Table 3). We also split the sample into sector categories and show the statistics for the two most representative, i.e., manufacturing and services.

3. The excluded sectors are the primary sector (sectors 1-9, according to NACE Rev. 3); the manufacture of tobacco products (sector 12); remediation activities and other waste management services (sector 39); the activities of households as employers of domestic personnel (sector 97); undifferentiated goods and services-producing activities of private households for own use (sector 98) and activities of extraterritorial organizations and bodies (sector 99).

Variable	Description	Source
Workforce characteristics		
Wage	Real hourly wage (base wage and regular benefits divided by the normal monthly hours of work) in 2019 euros	QP
Schooling	Number of schooling years [a]	QP
Tenure	Number of years at the firm	QP
Age	Worker's age	QP
Firm characteristics		
Log of value-added per worker	Logarithm of gross value-added [b] in 2019 euros divided by the number of workers	SCIE
Log of value-added per hour	Logarithm of gross value-added in 2019 euros divided by the number of hours worked (normal and overtime monthly hours multiplied by the 11 months of work per year)	SCIE/QP
Percentage of female	Share of female workers at the firm	QP
Percentage of part-time	Share of part-time workers at the firm	QP
Log of firm size	Logarithm of the number of workers at the firm	QP
Average tenure	Average of workers' tenure at the firm	QP

TABLE 2. Variables' description and corresponding data source

a. The data reports the highest level of education completed by the worker which we convert in years of schooling. After correcting inconsistent values on this variable we attribute years of education to each worker according to the following rule: 0 years of education (workers who do not know how to read or write), 2 years (workers with less than 4 years of schooling), 4 years (first cycle of basic education), 6 (second cycle of basic education), 9 years (third cycle of basic education), 12 years (upper secondary education), 13 years (post-secondary education), 15 years of schooling (workers with polytechnic or bachelor degree), 17 years (master degree) and 21 years (PhD).

b. We apply the winsorize technique at the 1% and 99% for value-added in order to reduce the effect of outliers.

The two measures of the apparent labour productivity, i.e. value-added per worker and per hour, show that the average firm in the services sector is in general more productive than in manufacturing, which is in line with the official statistics for Portugal.

The average firm in the services sector has more skilled workers, as measured by the multi-dimensional skill index presented in Section 2. These results remain unchanged when we use different skill measures, as the skill index using the two observed characteristics: education and age.[4] Nevertheless, the average firm in the services sector is slightly more heterogeneous in terms of skills than in the manufacturing, as measured by the standard deviation of both skill indices. This is also an expected result, as the services include highly differentiated activities.

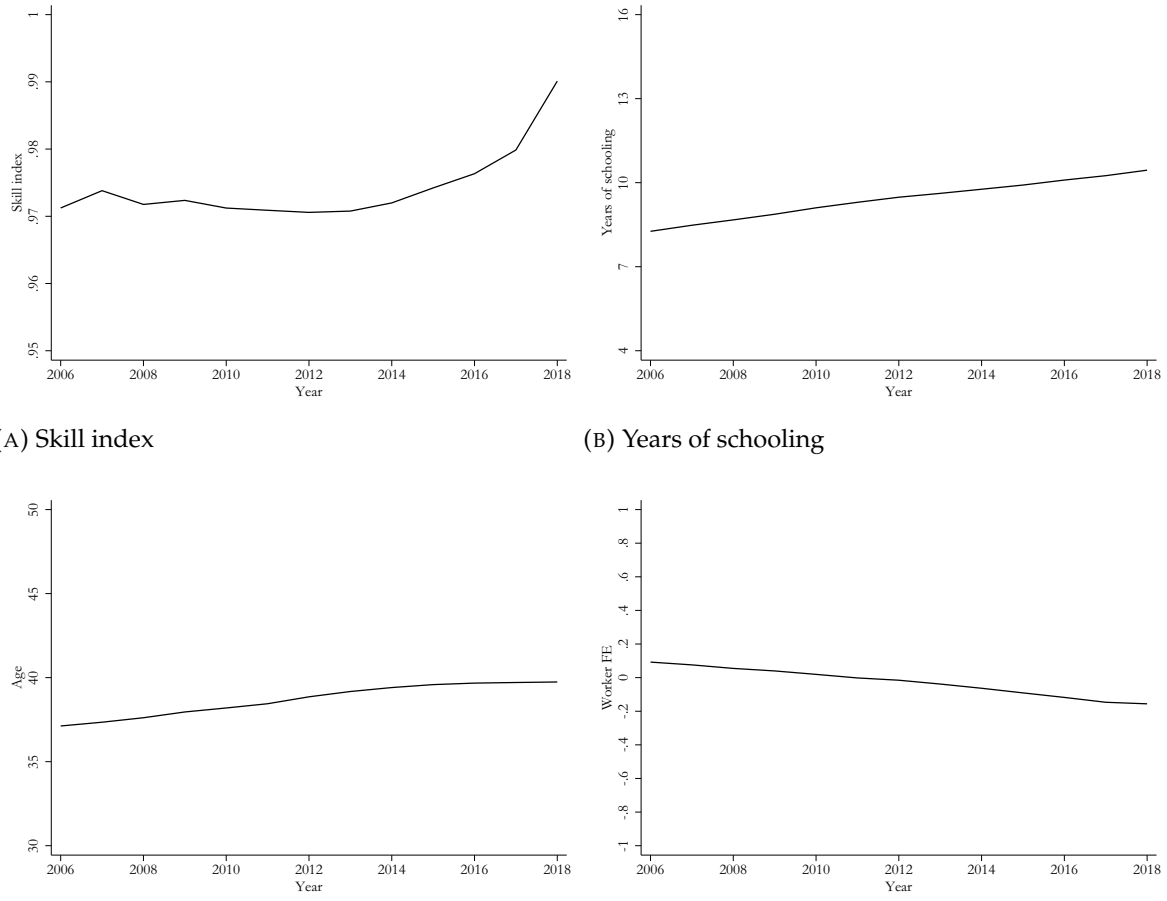
Considering the variables included in the skill index individually, the average number of years of schooling is also higher in services compared to manufacturing. The average workforce in services is also younger and stay at the firm for a shorter period of time.

4. We consider the first two components of equation (1): $a_{it,school} \times a_{it,age}$.

Regarding other control variables included in our econometric specification, the percentage of part-time workers and the percentage of female workers are higher in services than in manufacturing for the period under analysis. Also, manufacturing firms are, on average, larger than those in services sector.

Figure 1 shows the evolution of the skill index defined by equation (1) as well as its variables. Regarding the skill index, we observe a period of relative stability followed by an increasing trend. This occurs in parallel with an increase in the workforce average education and age over the period. The postponement of the entry into the labour market during the crisis period, as well as the progressive increase in the retirement age, may contribute to these patterns. In turn, the unobserved ability presents a subtle decreasing trend.

Figure 2 shows a positive correlation between the average workers' skill index and firm's productivity, which we analyse in detail in the following sections.



(A) Skill index

(B) Years of schooling

(C) Age

(D) Unobserved ability (worker FE)

FIGURE 1: Evolution of the skill index and its components (2006-2018)

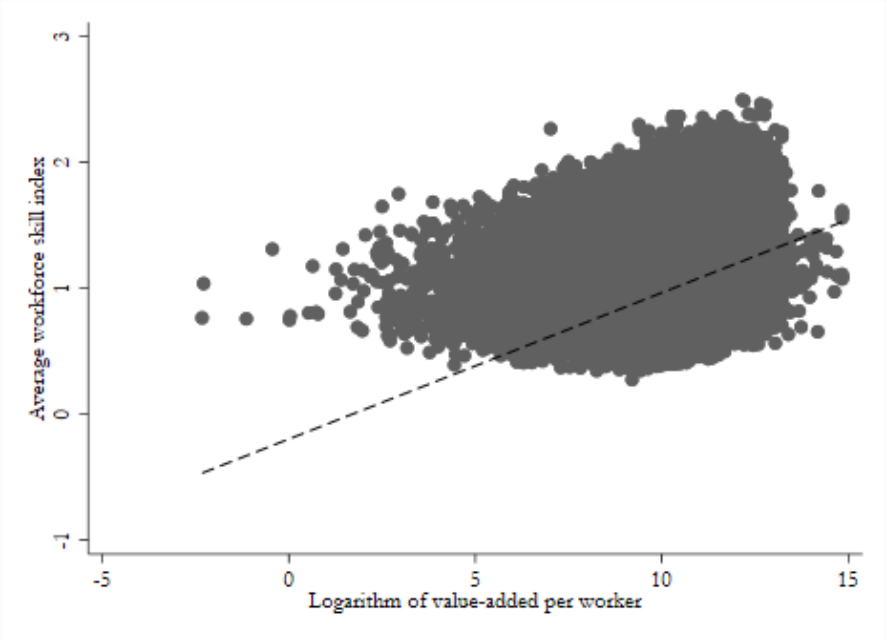


FIGURE 2: Relationship between firm's productivity and average workforce skill
Note: The dashed line represents the fitted values.

5. Results

5.1. Impact of workforce skills on firms' productivity

Table 3 presents the results of our main specification for the relationship between the two moments of workers' skill distribution, i.e., average and standard deviation, and firm's productivity, measured by value-added per worker. The first estimation column concerns the entire sample, while the second focuses on manufacturing and the last column refers to the services sector.

We find that the average workforce skills within the firm is positively related to its productivity. More specifically, a one standard deviation increase in the average worker skill is associated, on average, to an increase in firm's productivity by approximately 3.5% (product of the standard deviation of average skill index in Table A2, 0.23, by the estimated coefficient in Table 3, 0.1514, by 100%).[5] This is a consistent result in the literature that suggests that firms with a high-skilled workforce are also more productive, regardless of how skills are measured (e.g., Bender *et al.* 2018; Haltiwanger *et al.* 1999). There is also another strand of literature that corroborates this result, but considering on-the-job training. Barron *et al.* (1987), Dearden *et al.* (2006) and Konings and Vanormelingen (2015) are part of the research that found that workers' training increases firm's productivity.

5. Multiplying the standard deviation of the explanatory variable by the estimated coefficient gives an interpretation of the coefficient independent of the scale.

Regarding the standard deviation of the workers' skill index within the firm, our estimates indicate that more heterogeneous firms are also less productive. Specifically, a standard deviation increase in the dispersion of the skill index within the firm is associated to a decrease in firm's productivity by approximately 0.6%.[6]

	All	Manufacturing	Services
Average worker skill	0.1514***	0.1397***	0.1306***
	(0.0121)	(0.0256)	(0.0153)
Worker skill dispersion (SD)	-0.0523***	-0.0551**	-0.0389**
	(0.0130)	(0.0280)	(0.0166)
Share of part-time workers	-0.0023***	-0.0021***	-0.0022***
	(0.0001)	(0.0004)	(0.0002)
Share of females	-0.0009***	-0.0011***	-0.0009***
	(0.0001)	(0.0002)	(0.0001)
Average worker tenure	0.0064***	0.0105***	0.0076***
	(0.0010)	(0.0025)	(0.0012)
Average worker tenure squared	-0.0002***	-0.0004***	-0.0002***
	(0.00004)	(0.0001)	(0.00005)
Firm size (log)	0.4988***	0.5422***	0.4686***
	(0.0109)	(0.0212)	(0.0138)
Firm size (log) squared	-0.0855***	-0.0836***	-0.0873***
	(0.0020)	(0.0037)	(0.0025)
Adjusted R^2	0.714	0.705	0.730
Number of observations	722,494	192,578	415,134

TABLE 3. Workforce skills and firm's productivity (2006–2018)

Notes: Standard errors clustered at the firm level in parentheses. Significance levels: ***, 1%; **, 5%. The dependent variable is the logarithm of gross value-added per worker. The regressions include year dummies and firm fixed effects. Manufacturing corresponds to 2-digit NACE Rev. 3 codes 10 to 33; Services corresponds to NACE Rev. 3 codes 45 to 96. "All" stands for all firms in the sample. "SD" represents the standard deviation.

6. Since we cannot exclude the possible bias resulting from the simultaneity between the firm's workforce selection and productivity maximization decision, we also estimate the model with all the independent variables lagged by one period. The results for the main variables of interest are qualitatively similar in this specification.

Comparing the estimates of the main parameters between manufacturing and services, the positive relation between the average worker skill and the value-added per worker is slightly higher for manufacturing than for services. Also, the coefficient associated with skill dispersion is statistically significant for both sectors, despite being slightly larger, in absolute terms, for manufacturing.

Table A3 in the Appendix presents the results considering only the firms with at least 10 workers in all periods. Although this condition is very restrictive considering the small average size of the firms in Portugal (e.g., Braguinsky *et al.* 2011; Banco de Portugal 2021), the impact of the average worker skill on firm's productivity remains significant and is even higher. Regarding the impact of the worker skill standard deviation, it is negative in all samples and loses significance in manufacturing.[7] The lowest worker turnover levels in the manufacturing sector or the smaller number of firm observations in the sample comparing to the services sector may contribute to explain this result. It is also important to highlight the large heterogeneity, within and between manufacturing and services, in terms of the activities and the occupational composition. For example, in the manufacturing sector engineers and skilled technicians coexist with workers performing repetitive tasks. The effect of skill heterogeneity on productivity may differ for white and blue-collar workers depending on the level of substitutability between them (e.g., Iranzo *et al.* 2008; Parrotta *et al.* 2014).

Regarding the control variables, the share of part-time workers is negatively related to the firm's value-added per worker. Furthermore, the lower level of productivity in firms with a higher share of female workers is also a common result in the literature (e.g., Ilmakunnas and Ilmakunnas 2011; Parrotta *et al.* 2014; Pfeifer and Wagner 2014). Finally, the average tenure at the firm is positively related to its productivity, which is in line with previous studies (e.g., Parrotta *et al.* 2014), and we also observe a concave tenure-productivity profile. The firm size has also an inverted U-shaped relation with productivity as found by Pfeifer and Wagner (2014).

5.2. Sensitivity analysis

In this subsection, we assess the sensitivity of our estimates to an alternative measure of productivity, different proxies for skill and different approaches to quantify within-firm skill heterogeneity.

5.2.1. Productivity measure

We re-estimate equation (3) with the gross value-added per hour worked as the dependent variable.[8] The results shown in Table 4 are qualitatively similar to those obtained for the value-added per worker. On average, a standard deviation increase

7. This result also holds if we consider small and medium firms with at least 10 and up to 249 workers in all time periods.

8. SCIE data do not provide information on hours worked. Therefore, we use the total number of normal and overtime hours reported with reference to the month of October in QP data multiplied by the 11 working months assuming that each worker is absent from the firm, on average, for one month.

in the average worker skill is associated with an increase of approximately 2.9% in value-added per hour worked, while skill dispersion is associated with a decrease of approximately 0.9%. We also confirm the previous conclusion that these effects tend to be larger in manufacturing than in the services sector.

	All	Manufacturing	Services
Average worker skills	0.1274***	0.1893***	0.0875***
	(0.0123)	(0.0277)	(0.0154)
Worker skills dispersion (SD)	-0.0761***	-0.0925***	-0.0705***
	(0.0135)	(0.0300)	(0.0170)
Adjusted R^2	0.705	0.696	0.725
Number of observations	722,494	192,578	415,134

TABLE 4. Sensitivity analysis – Productivity measured by value-added per hour worked

Notes: Standard errors in parentheses are clustered at the firm level. Significance levels: ***, 1%. The dependent variable is the logarithm of gross value-added per hour worked. The regressions include the following controls: percentage of female and part-time workers, tenure and tenure squared and the logarithm of size and its square, year dummies and firm fixed effects. "All" stands for all firms in the sample. "SD" represents the standard deviation.

5.2.2. Skill measure

In this subsection we analyse to what extent the findings discussed in the previous section are sensitive to some alternative skill measures.

The estimation of the worker's unobserved ability using the procedure described in Section 2 hinges upon having enough variability in the observed characteristics to disentangle the observed and unobserved effects. In order to alleviate this restriction we compute the skill index defined in equation (1) with the two observed components of skill: education and age. The results are shown in Table 5. The coefficients remain qualitatively unchanged but lose statistical significance in the services sector using this alternative skill index. The statistical significance is kept unchanged, however, if we consider only firms with at least 10 workers in all periods.[9] Ilmakunnas and Ilmakunnas (2011) find that a two-dimensional age-education diversity measure is not significantly correlated with productivity using Finnish data.

The choice over the skill variable matters for the empirical evidence, as shown by Ilmakunnas and Ilmakunnas (2011). They find that productivity is negatively associated with educational diversity but positively correlated with age diversity. Therefore, it is relevant to understand the association between firm's productivity and each one of the variables that are used in the skill index proposed in Section 2.

9. These results are available upon request.

	All	Manufacturing	Services
Average worker skills with education and age	0.0781***	0.0822**	0.0348
	(0.0168)	(0.0342)	(0.0219)
Worker skills dispersion with education and age (SD)	-0.0767***	-0.1133***	-0.0337
	(0.0195)	(0.0404)	(0.0255)
Adjusted R^2	0.713	0.705	0.730
Number of observations	722,725	192,630	415,276

TABLE 5. Sensitivity analysis – Skill index with observed characteristics

Notes: Standard errors in parentheses are clustered at the firm level. Significance levels: ***, 1%; **, 5%. The dependent variable is the logarithm of gross value-added per worker. The regressions includes the following controls: percentage of female and part-time workers, tenure and tenure squared and the logarithm of size and its square, year dummies and firm fixed effects. "All" stands for all firms in the sample. "SD" represents the standard deviation.

Table 6 shows the relationship between the average and standard deviation of the years of education at the firm and value-added per worker. As expected, the average years of education of the workforce are positive and significantly associated with firm's productivity – one standard deviation increase in the average worker education is associated with an increase of 2.8% in productivity.

The larger dispersion in terms of years of education is associated, on average, with a decrease in firm's productivity. However, this effect is relatively low – one standard deviation increase in the dispersion of the years of education is associated with a decrease of 0.3% in productivity – only statistically significant for the services sector.

	All	Manufacturing	Services
Average worker education	0.0109***	0.0084***	0.0107***
	(0.0012)	(0.0024)	(0.0016)
Worker education dispersion (SD)	-0.0028**	0.0039	-0.0045***
	(0.0013)	(0.0026)	(0.0017)
Adjusted R^2	0.713	0.705	0.730
Number of observations	722,725	192,630	415,276

TABLE 6. Sensitivity analysis – Education

Notes: Standard errors in parentheses are clustered at the firm level. Significance levels: ***, 1%; **, 5%. The dependent variable is the logarithm of gross value-added per worker. The regressions includes the following controls: percentage of female and part-time workers, tenure and tenure squared and the logarithm of size and its square, year dummies and firm fixed effects. "All" stands for all firms in the sample. "SD" represents the standard deviation.

Table 7 shows the impact of the firm's workforce age composition on value-added per worker. In line with the reported evidence of an inverse U-shaped relationship between age and productivity (e.g., Pfeifer and Wagner 2014, Cardoso *et al.* 2011), we consider a slightly different specification by including a second order polynomial for the average workers' age. The results confirm a concave relationship between the workforce average age and productivity. The workers' age dispersion is negatively associated with firm's productivity – one standard deviation increase in the age dispersion is associated with a decrease of 0.9% in productivity – although not statistically significant at the usual significance levels for manufacturing.

	All	Manufacturing	Services
Average worker age	0.0342***	0.0250***	0.0312***
	(0.0023)	(0.0048)	(0.0030)
Average worker age squared	-0.0004***	-0.0003***	-0.0004***
	(0.00003)	(0.0001)	(0.00004)
Worker age dispersion (SD)	-0.0036***	-0.0015*	-0.0043***
	(0.0004)	(0.0009)	(0.0006)
Adjusted R^2	0.714	0.705	0.730
Number of observations	722,725	192,630	415,276

TABLE 7. Sensitivity analysis –Workers' age

Notes: Standard errors in parentheses are clustered at the firm level. Significance levels: ***, 1%; *, 10%. The dependent variable is the logarithm of gross value-added per worker. The regressions include the following controls: percentage of female and part-time workers, tenure and tenure squared and the logarithm of size and its square, year dummies and firm fixed effects. "All" stands for all firms in the sample. "SD" represents the standard deviation.

Most recent papers also use worker fixed effects estimated in a first stage Mincerian wage equation as a proxy for worker (unobserved) ability (Iranzo *et al.* 2008). The results obtained with this measure are shown in Table 8 and are qualitatively similar to those obtained using the skill index. A one standard deviation increase in the average worker ability is associated with an increase of 6% in productivity. In comparison, a standard deviation increase in the dispersion of the workers estimated fixed effects is associated with a decrease in value-added per worker by 0.3% but not statistically significant at the usual significance levels. Although the coefficient of the dispersion of workforce ability within the firm is positive for the manufacturing sector, it is not statistically significant.

These results are consistent with the idea that workforce diversity can affect firms' productivity through different dimensions (Parrotta *et al.* 2014). Our results show that firms, and especially those in the services sector, may have productivity gains by hiring workers of similar ability, education and age. The skill index used in this article is a comprehensive measure that considers this evidence.

	All	Manufacturing	Services
Average worker FE	0.2227***	0.1899***	0.2159***
	(0.0103)	(0.0207)	(0.0133)
Worker FE dispersion (SD)	-0.0206*	0.0179	-0.0359**
	(0.0113)	(0.0220)	(0.0145)
Adjusted R^2	0.714	0.706	0.730
Number of observations	722,494	192,578	415,134

TABLE 8. Sensitivity analysis – Unobserved ability

Notes: Standard errors in parentheses are clustered at the firm level. Significance levels: ***, 1%; **, 5% *, 10%. The dependent variable is the logarithm of gross value-added per worker. The regressions include the following controls: percentage of female and part-time workers, tenure and tenure squared and the logarithm of size and its square, year dummies and firm fixed effects. "All" stands for all firms in the sample. "SD" represents the standard deviation.

5.2.3. Dispersion measure

Finally, we consider it is relevant to assess the robustness of our results to other dispersion measures. To this end, we re-estimate equation (3) replacing the standard deviation of the skill index by the variance, coefficient of variation and the ratio between different percentiles of the skills distribution in order to assess the consistency of the correlation of skills dispersion and firm's productivity (Table 9). The estimates are broadly consistent with those discussed above, irrespective of the dispersion measure used.

The ratio between the skill level of the worker at the 90^{th} percentile and that of the worker at the 10^{th} or the 50^{th} percentile of the skill index distribution is negatively associated with firm's productivity. However, the coefficient of the ratio between the skill level at the median and that at the 10^{th} percentile of the skill index distribution is not statistically significant. This provides evidence that the dispersion at the bottom is not as relevant as the dispersion at the top half of the skill distribution.

5.3. Workforce skills and productivity distribution

In this subsection, we intend to verify whether the estimated coefficients of our main econometric specification change across the productivity distribution. We, therefore, estimate regression quantiles with firm fixed effects using the Method of Moments Quantile Regression proposed by Machado and Santos Silva (2019). According to Machado and Santos Silva (2019) when the number of observations is large compared to the number of time periods we may face asymptotic bias issues. As such, the results in this subsection should be read with caution.

	All	All	All	All	All
Average worker skills	0.1363***	0.1344***	0.1337***	0.1353***	0.1489***
	(0.0111)	(0.0111)	(0.0111)	(0.0111)	(0.0120)
Worker skills dispersion (P90/P10)	-0.0067***				
	(0.0016)				
Worker skills dispersion (P90/P50)		-0.0156***			
		(0.0032)			
Worker skills dispersion (P50/P10)			-0.0019		
			(0.0030)		
Worker skills dispersion (Coeff. Var.)				-0.0388***	
				(0.0120)	
Worker skills dispersion (Variance)					-0.0645***
					(0.0185)

TABLE 9. Sensitivity analysis – Alternative dispersion measures

Notes: Standard errors in parentheses are clustered at the firm level. Significance levels: ***, 1%. The dependent variable is the logarithm of gross value-added per worker. The regressions include the following controls: percentage of female and part-time workers, tenure and tenure squared and the logarithm of size and its square, year dummies and firm fixed effects. The number of observations is 722,494. "All" stands for all firms in the sample. "P90" represents percentile 90; likewise for the other percentiles. "Coeff. Var." is "Coefficient of variation".

Table 10 presents the estimates for five percentiles (10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th}) of the conditional productivity distribution.[10] We can observe that the coefficients associated with the average and dispersion of worker skill preserve the statistical significance in the five percentiles. Regarding the magnitude of the coefficients, the results are similar at the mean and median of the conditional distribution. However, the hypothesis that the coefficients are the same across all quantiles is rejected, i.e., the impact of the average worker skill on firm's productivity varies depending on the position of each firm in the productivity distribution.

We find an increasing positive association of the average worker skill with firm's productivity over the conditional quantiles of the distribution. Therefore, in more productive firms a marginal increase in the average workers' skill index is associated with a larger increase in productivity than in less productive firms, controlling for the share of females and part-time workers and average tenure at the firm, firm's size and firm and time fixed effects.

10. These estimates were obtained in the same sample of our main econometric specification.

	P10	P25	P50	P75	P90
Average worker skill	0.1235***	0.1358***	0.1534***	0.1672***	0.1766***
	(0.0141)	(0.0119)	(0.0097)	(0.0097)	(0.0107)
Worker skill dispersion (SD)	-0.0508***	-0.0514***	-0.0524***	-0.0531***	-0.0536***
	(0.0163)	(0.0137)	(0.0110)	(0.0108)	(0.0118)

TABLE 10. Workforce skills and firm's productivity distribution

Notes: These estimates are obtained in the sample of our main econometric specification. We use 1000 bootstrap replications to obtain estimates for standard errors in parentheses. Significance levels: ***, 1%. The dependent variable is the logarithm of gross value-added per worker. The regression includes the following controls: percentage of female and part-time workers, tenure and tenure squared and the logarithm of size and its square, year dummies and firm fixed effects. The number of observations is 722,494. "P90" stands for percentile 90 and the same applies to the other percentiles. "SD" represents the standard deviation.

The skill dispersion is negatively associated with firm's productivity in line with the estimates at the mean. The hypothesis of coefficient equality over the different quantiles cannot be rejected which provides evidence that the relationship between worker skill dispersion and firm's productivity is relatively homogeneous over the conditional productivity distribution.

6. Conclusion

We use Portuguese linked employer-employee data to investigate the relationship between firm's productivity (value-added per worker) and the two first moments of the workers' skill distribution (average and standard deviation) for 2006-2018.

Unlike most previous empirical studies, which focus on a single component of worker's skill, we use a multi-dimensional skill index to comprehensively measure three of the most debated dimensions of workforce skills: worker's formal education, age and unobserved ability. This last dimension corresponds to the worker fixed effect obtained from a Mincerian wage equation.

We find a positive and significant relationship between the average workforce skills and firm's productivity, both in the manufacturing and the services sector. This result is robust to different skill measures and increases across the conditional productivity distribution.

On the other hand, the standard deviation of workers' skill index, conditional on its average, is negatively associated with firm's productivity. This effect is roughly the same across firms with different productivity levels.

Our reduced-form analysis deserves further exploration to identify causal relations between skill composition and firm's productivity. Additionally, the skill index can be extended to include firm-specific human capital and managerial skills that the literature singles out to be relevant for firm's outcomes. Finally, it would also be pertinent to

analyse the productivity dynamics in the post-COVID-19 period, since the pandemic represents a shock to the organisation of work, namely in terms of the technologies used and how the workers interact, which may have heterogeneous effects across sectors of activity.

References

Abowd, John M, Francis Kramarz, and David N Margolis (1999). "High wage workers and high wage firms." *Econometrica*, 67(2), 251–333.

Acemoglu, Daron and Jorn-Steffen Pischke (1998). "Why Do Firms Train? Theory and Evidence." *The Quarterly Journal of Economics*, 113(1), 79–119.

Banco de Portugal (2021). "The evolution of the average firm size in Portugal." *March 2021 Economic Bulletin*, pp. 27–39.

Barron, John M., Dan A. Black, and Mark A. Loewenstein (1987). "Employer Size: The Implications for Search, Training, Capital Investment, Starting Wages, and Wage Growth." *Journal of Labor Economics*, 5(1), 76–89.

Bender, Stefan, Nicholas Bloom, David Card, John Van Reenen, and Stefanie Wolter (2018). "Management Practices, Workforce Selection, and Productivity." *Journal of Labor Economics*, 36(S1), S371–S409.

Braguinsky, Serguey, Lee G Branstetter, and André Regateiro (2011). "The incredible shrinking Portuguese firm." *National Bureau of Economic Research*.

Cardoso, Ana Rute, Paulo Guimarães, and José Varejão (2011). "Are Older Workers Worthy of Their Pay? An Empirical Investigation of Age-Productivity and Age-Wage Nexuses." *De Economist*, 159(2), 95–111.

Correia, Sergio (2016). "Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator." Duke University Working Paper.

Dearden, Lorraine, Howard Reed, and John Van Reenen (2006). "The Impact of Training on Productivity and Wages: Evidence from British Panel Data." *Oxford Bulletin of Economics and Statistics*, 68(4), 397–421.

Guimarães, Paulo and Pedro Portugal (2010). "A simple feasible procedure to fit models with high-dimensional fixed effects." *The Stata Journal*, 10(4), 628–649.

Haltiwanger, John C., Julia I. Lane, and James R. Spletzer (1999). "Productivity Differences across Employers: The Roles of Employer Size, Age, and Human Capital." *The American Economic Review*, 89(2), 94–98.

Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan (2003). "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation." *Journal of Political Economy*, 111(3), 465–497.

Hellerstein, Judith K, David Neumark, and Kenneth R Troske (1999). "Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations." *Journal of Labor Economics*, 17(3), 409–446.

Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi (1997). "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines." *The American Economic Review*, 87(3), 291–313.

Ilmakunnas, Pekka and Seija Ilmakunnas (2011). "Diversity at the workplace: Whom does it benefit?" *De Economist*, 159(2), 223–255.

Iranzo, Susana, Fabiano Schivardi, and Elisa Tosetti (2008). "Skill dispersion and firm productivity: An analysis with employer-employee matched data." *Journal of Labor Economics*, 26(2), 247–285.

Konings, Jozef and Stijn Vanormelingen (2015). "The Impact of Training on Productivity and Wages: Firm-Level Evidence." *The Review of Economics and Statistics*, 97(2), 485–497.

Kremer, Michael and Eric Maskin (1996). "Wage inequality and segregation by skill." Working Paper w5718, National bureau of economic research.

Machado, José A.F. and J.M.C. Santos Silva (2019). "Quantiles via moments." *Journal of Econometrics*, 213(1), 145–173.

Mas, Alexandre and Enrico Moretti (2009). "Peers at Work." *American Economic Review*, 99(1), 112–145.

Mendes, Rute, Gerard J. van den Berg, and Maarten Lindeboom (2010). "An empirical assessment of assortative matching in the labor market." *Labour Economics*, 17(6), 919–929.

Parrotta, Pierpaolo, Dario Pozzoli, and Mariola Pytlikova (2014). "Labor diversity and firm productivity." *European Economic Review*, 66, 144–179.

Pfeifer, Christian and Joachim Wagner (2014). "Age and gender effects of workforce composition on productivity and profits: Evidence from a new type of data for German enterprises." *Contemporary Economics*, 8(1), 25–46.

Portela, Miguel (2001). "Measuring skill: a multi-dimensional index." *Economics Letters*, 72(1), 27–32.

Rocha, Vera, Mirjam Van Praag, Timothy B Folta, and Anabela Carneiro (2019). "Endogeneity in strategy-performance analysis: An application to initial human capital strategy and new venture performance." *Organizational Research Methods*, 22(3), 740–764.

Syverson, Chad (2011). "What Determines Productivity?" *Journal of Economic Literature*, 49(2), 326–365.

Torres, Sónia, Pedro Portugal, John T. Addison, and Paulo Guimarães (2018). "The sources of wage variation and the direction of assortative matching: Evidence from a three-way high-dimensional fixed effects regression model." *Labour Economics*, 54, 47–60.

Appendix

	All
Schooling (years)	0.0058***
	(0.0002)
Age squared	-0.0003***
	(6.1e-06)
Tenure	0.0080***
	(0.0003)
Tenure squared	-0.0001***
	(0.00001)
Firm size (log)	0.0561***
	(0.0120)
Firm size (log) squared	-0.0020
	(0.0016)

TABLE A1. Estimates of the wage equation (2) (2006–2018)

Notes: Standard errors in parentheses are clustered at the firm level. Significance levels: ***, 1%. The dependent variable if the logarithm of real hourly wages. The regression includes firm, worker and year fixed effects. The number of observations is 24,643,358. "All" stands all firms in the sample.

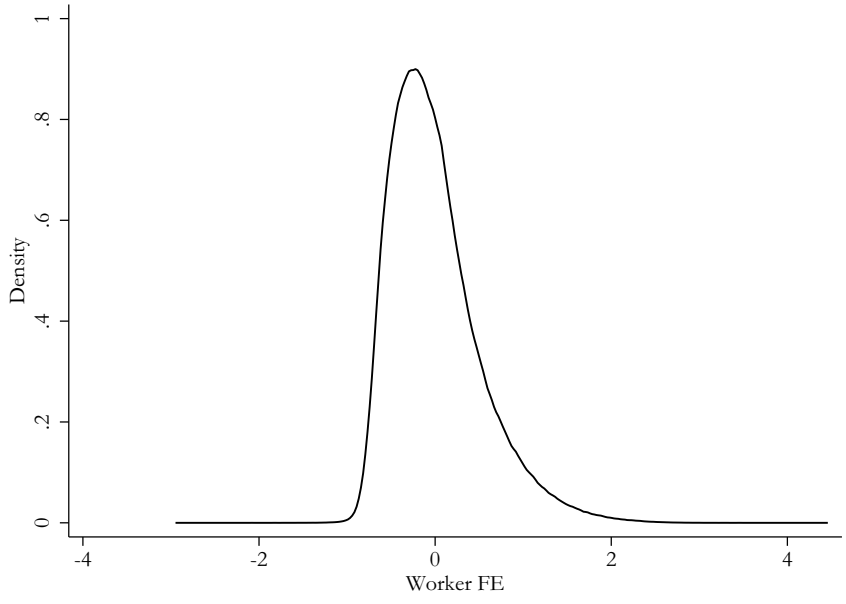


FIGURE A1: Density of the worker fixed effects

	All					Manufacturing					Services				
	Mean	SD	P25	P50	P75	Mean	SD	P25	P50	P75	Mean	SD	P25	P50	P75
Workforce characteristics															
Average skill index	0.95	0.23	0.79	0.91	1.08	0.87	0.18	0.74	0.85	0.97	1.01	0.25	0.83	0.97	1.15
Standard deviation of the skill index	0.31	0.12	0.22	0.29	0.38	0.29	0.11	0.21	0.28	0.36	0.32	0.12	0.23	0.31	0.39
Average skill index with education and age	0.95	0.14	0.85	0.93	1.04	0.89	0.11	0.81	0.88	0.96	0.99	0.15	0.89	0.98	1.09
Standard deviation of the skill index with education and age	0.19	0.06	0.15	0.19	0.23	0.18	0.06	0.14	0.18	0.22	0.20	0.06	0.15	0.19	0.23
Years of schooling	9.04	2.57	7.06	8.75	10.80	7.80	1.87	6.43	7.61	9.00	10.00	2.56	8.12	9.89	12.00
Standard deviation of the years of schooling	2.59	1.06	1.87	2.62	3.29	2.62	0.99	1.99	2.68	3.30	2.51	1.04	1.76	2.51	3.21
Age	39.19	5.68	35.21	39.20	43.13	39.78	5.35	36.10	39.77	43.43	38.65	5.79	34.54	38.60	42.67
Standard deviation of the age	9.22	2.62	7.51	9.30	10.93	9.50	2.38	7.99	9.55	11.03	9.04	2.75	7.18	9.10	10.88
Worker FE	-0.05	0.27	-0.23	-0.08	0.10	-0.11	0.23	-0.27	-0.13	0.04	-0.02	0.30	-0.23	-0.06	0.14
Standard deviation of the worker FE	0.35	0.14	0.25	0.33	0.43	0.34	0.13	0.25	0.33	0.42	0.36	0.14	0.26	0.34	0.44
Percentage of part-time workers	2.74	8.73	0.00	0.00	0.00	1.04	4.64	0.00	0.00	0.00	3.65	10.39	0.00	0.00	0.00
Percentage of female workers	39.78	31.86	12.50	33.33	66.67	41.19	32.17	14.29	33.33	66.67	47.26	30.69	20.00	43.75	72.81
Average tenure	6.62	5.33	2.50	5.36	9.55	8.20	5.89	3.57	7.11	11.75	6.28	5.10	2.35	5.00	9.00
Firm characteristics															
Log of value-added per worker	9.88	0.72	9.49	9.87	10.29	9.76	0.63	9.40	9.75	10.15	9.96	0.76	9.58	9.97	10.40
Log of value-added per hour	2.55	0.71	2.17	2.55	2.95	2.41	0.63	2.04	2.40	2.79	2.63	0.75	2.26	2.64	3.05
Log of firm size	2.55	0.92	1.79	2.30	3.00	2.80	0.98	2.08	2.56	3.33	2.47	0.90	1.79	2.20	2.83

TABLE A2. Summary statistics (2006-2018)

Notes: Manufacturing corresponds to 2-digit NACE Rev. 3 codes 10 to 33; Services corresponds to NACE Rev. 3 codes 45 to 96. "SD" stands for standard-deviation. "P25", "P50" and "P75" represents percentile 25, median and percentile 75, respectively.

	All (>=10 workers)	Manufacturing (>=10 workers)	Services (>=10 workers)
Average worker skills	0.3322***	0.1677***	0.3343***
	(0.0285)	(0.0441)	(0.0390)
Worker skills dispersion (SD)	-0.1322***	-0.0555	-0.1254***
	(0.0310)	(0.0502)	(0.0433)
Adjusted R^2	0.785	0.768	0.802
Number of observations	232,122	87,373	115,821

TABLE A3. Workforce skills and firm's productivity – firms with 10 or more workers

Notes: Standard errors in parentheses are clustered at the firm level. Significance levels: ***, 1%. The dependent variable is the logarithm of gross value-added per worker. The estimation includes firms with at least 10 workers in all time periods. The regressions includes the following controls: percentage of female and part-time workers, tenure and tenure squared and the logarithm of size and its square, year dummies and firm fixed effects. "All" stands for all firms in the sample. "SD" represents the standard deviation.

www.bportugal.pt