Routinization and Covid-19: a comparison between United States and Portugal

Piero De Dominicis

November 2, 2020

Master Thesis

A Project carried out on the Master in Economics, under the supervision of

Prof. Pedro Brinca

Nova School of Business & Economics, Lisbon

Abstract

Earnings inequality has been increasing in Portugal and United States in the last 30 years and with the recent pandemic outbreak this trend is likely to reach even higher levels. The purpose of this paper is to identify what is the role of automatization in increasing wage inequality, making a comparison between the two countries. Using PSID and Quadros de Pessoal, we find that labor income dynamics are strongly determined by the variance of the individual fixed component. This effect is intensively reduced by adding information on workers' occupational tasks, confirming the upward effects that decreasing price of capital and the consequent replacement of routine manual workers have on wage inequality. During the current crisis, we find that the possibility of working from home is strongly related with the occupation type. As such, we simulate the impact of a permanent demand shock using an overlapping-generations model with incomplete markets and heterogeneous agents to quantitatively predict the impact of Covid-19 and lockdown measures on wage premium and earnings inequality. We find that wage premia and earnings dispersion increase, suggesting that earnings inequality will increase at the expenses of manual workers.

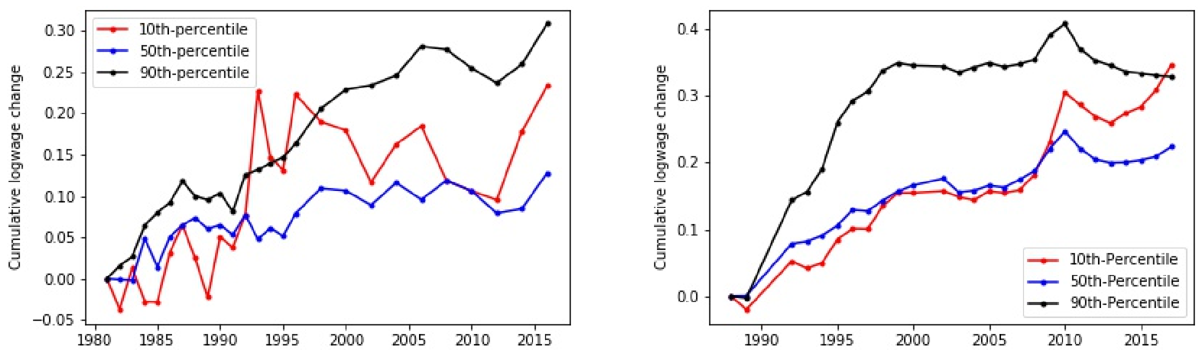
Key Words: Routinization, Wage Inequality, Labor Income Process, Covid-19, Teleworking.

A special thanks to professor Pedro Brinca for the constant help, support and, above all, for conveying me the passion of doing research at an excellent level. I am grateful to João B. Oliveira for the helpful tutoring and to Hans Holter for his technical advices. I dedicate this work to my parents, who made sacrifices to allow me to pursue my goals and always supported my ambitious projects, to my sister, for being her, to my friends in Teramo and my old teammates, who represent the other family I luckily had in my life.[1].

[1]Nova School of Business and Economics. This work used infrastructure and resources funded by Fundacao para a Ciencia e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

1 Introduction

Technological progress is considered one of the main drivers behind earnings inequality. Factor-biased technological change and skill-biased technological change represent two main sources of wage inequality. To this extent, we explore empirically the differences between workers in different categories, according to their occupation tasks, to assess how labor market has been impacted by task premia changes. This paper provides two main contributions to the existing literature. First, we use a 10-rolling window to estimate the evolution of determinants of dispersion in the labor income processes to investigate whether changes in task-premia represent a major source of labor income inequality. Second, we implement an overlapping generations model with incomplete markets to study the role of skill-based technological change in increasing wage inequality and to assess the potential impact of Covid-19 when people ability to continue working is mostly determined by the type of task they perform. We calibrate the model in order to match US and Portuguese economies using 2010 as benchmark year and we repeat the exercise targeting different working hours ratio per cognitive and manual workers in order to simulate the impact of demand side shocks.



(a) **Portugal**: 57,354,268 Observations (b) **US**: 245,316 Survey-Weighted Observations

Figure 1: Real wage increase per percentiles

Figure 1 shows the steady rise in wage inequality and wage growth at different percentiles suggesting that both Portugal and U.S. experienced wage polarization at two different time periods. In Portugal, low wages in routine task intensive occupations, combined with the same price of computer capital may have limited the gains of substituting workers by machines. We separate agents into non-routine and routine, according to their abilities substitutability with machines, and cognitive and manual, depending on the level of skills required to perform daily tasks. In this framework, we find that the wage premium of non-routine workers increases, following the drop in investment price and the decrease in tax progressivity[2], this mechanism is triggered by a drop in routine labor demand by firms and by cheaper capital accumulation. The trends in labor force composition, figure 2, confirms that Portugal experiences similar patterns of labor market polarization of the U.S., explained by technology advances such as computerization and automation which displace routine tasks, and complement cognitive tasks.

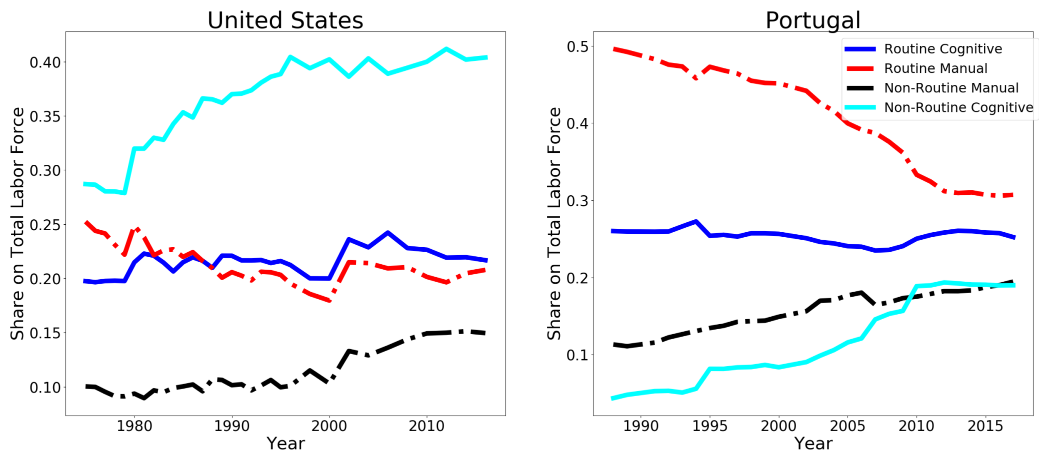


Figure 2: Labour force composition.

There is a clear increase in employment share of non-routine cognitive occupations, these workers are indeed complementary to capital and less likely to be substitute by machines. Both countries show a decrease in routine manual occupations, in Portugal the change is larger, as a matter of fact it decreases from 50% of the labor force in 1987 to 30% in 2017. Routine cognitive occupations remained approximately at the same level in both countries, driven by the increasing importance of the service sector. Non-routine workers, both cognitive and manual, show a positive trend which is steeper for cognitive occupations. For the U.S. there is a steady increase in non-routine cognitive employment share from 30% in 1976 to 40% in 2017, in Portugal the same occupation category increases from 3.5% in 1987 to 20% in 2017. The increase in demand for non-routine occupation confirms that also Portugal is experiencing labor market polarization but is lagging behind the

[2] Ferriere & Navarro (2014) and Nóbrega (2020).

United States in the adoption of computer capital[3].

Literature Review

Autor *et al.* (2003) first introduced the concept of routinization hypothesis as the decrease in labor input of routine manual tasks and the increase in labor input for non-routine cognitive tasks. Autor *et al.* (2006) pointed out that US wages structure widened due to an increase in demand for skills that was driven by skill-biased technical change and a slowdown in the growth of the relative supply of college workers. Acemoglu & Restrepo (2020) argues that difference in education are important source of inequality and Krusell *et al.* (2000) found that factor-biased technological change has the strongest impact in determining the increase in wage inequality. Acemoglu & Restrepo (2018) discuss the impact of increasing demand for skilled workers, who are able to perform more abstract tasks, outlining how automation can replace manual tasks in the long-run if the rental rate of capital remains less costly than wages. Also Guerreiro *et al.* (2017) found that substitutability is higher for routine occupations requiring low skills. Recent improvements in Artificial Intelligence brought astonishing changes in different fields and is expected to be even more disrupting in the future, Acemoglu & Restrepo (2018) investigate on the trade-off between the displacement effect, change in labour supply cause by automation of tasks which reduces demand for labor, and the overall increase in labor demand triggered by productivity-enhancing technologies. On the other side the creation of new tasks where human capital has a comparative advantage relative to machines , the reinstatement effect, may counterbalance the displacement effect. These mentioned effect do not grow equally faster, and different economies require different time to absorb efficiently and smoothing these processes, Goos & Manning (2007) argue that the "routinization" hypothesis is the driving factor of the increase in highest and lowest wage occupations in United kingdom since 1975 and Goos *et al.* (2009) extend the study to Western European group countries explaining job polarization using both routine biased technological change and offshoring. Krusell *et al.* (2000) and Karabarbounis & Neiman (2014) argues that the more recent decline in relative price of investment has been triggered by the investment-specific technological change. Eden & Gaggl (2015) shows that

[3]Workers in the two sample are unlikely to change occupation across the panel, meaning that changes in labor composition are driven by replacement with machines. This can be checked also in transition matrices 19-21 in the Appendix B.

the previously mentioned drop in demand for routine occupations was concurrent to the decrease in price of information and communication technology capital goods: this drop is responsible for 50% of the drop in labor share. Institutional differences between Portugal and US may be direct candidates for the wage inequality dynamics observed in the data. In the spirit of Fonseca *et al.* (2018) we replicated figure 7 finding that the increase in minimum wage had a positive impact for Portugal on the 10^{th} percentile as it may have impacted the wage convergence between different percentiles and the growth in wages for manual workers. The United States, on the other side, experienced a decrease in real minimum wage that may have widen the gap between wage growth rate at top 90^{th} and bottom 10^{th} of the distribution.

2 Data

To divide the workers in different categories according to the level of automation of their job we followed Cortes *et al.* (2014). The main data sources for this work are Quadros de Pessoal (QP) for Portugal and Panel Study of Income Dynamics (PSID) for the US.

Quadros de Pessoal This database is a matched employer-employeee dataset created by the Portuguese Ministry of Labor in the 1980s, it includes Portuguese firms with at least one employee and does not take into account self-employed workers. The dataset cover the time period going from 1987 to 2017. The original occupations map was made by Cortes *et al.* (2014) on Census Occupational Codes, to map the Portuguese occupations we use different algorithms and crosswalks, details can be found in the Appendix. We propose a 4 digits mapping after 2007 and 3 digits between 1987 and 2006.

PSID The Panel Study of Income Dynamics (PSID) is one of the longest longitudinal study as it includes almost 5000 families followed from 1968 to 2017. Data are collected every year from 1968 to 1997 and biannually from 1997 to 2017. All the information collected are referred to the previous year. The survey contains information both at individual level and family level, in this work we focused on individuals. In particular, to define the sample used for the estimation of the labor income processes we followed Heathcote *et al.* (2010) approach. The only difference is that we split households to create a panel for singular individuals and we generate individual characteristics splitting variables based on household composition. Figure 3 shows that PSID sample, despite two minor divergences between 1995-1999 and after 2008, is representative for the US labor market[4]. The sample is made of only heads and spouses of the families where the greatest level of accuracy in the data is guaranteed.

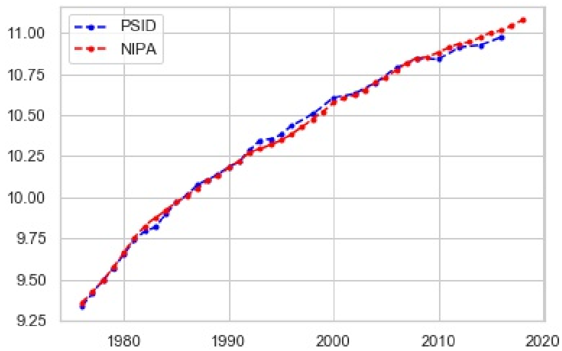


Figure 3

Observations with a wage lower than half of the minimum wage[5] have been dropped, also individual working less than 260 yearly hours have been dropped out of the samples. Table 1 and table 2 report the two samples that we use for our analysis. For Quadros de Pessoal we followed the approach of Fonseca *et al.* (2018) re-adapting their method to Heathcote *et al.* (2010) to have consistency between the two samples.

Table 1: PSID Sample Selection (Survey years 1969-2017)

	Dropped	Remaining
Initial Sample 1969-2017		453,969
Hourly Wage $\leq 0.5 \times$ min.wage	10,784	443,185
Age $25-64$	$126,072$	317,113
Annual Hours ≤ 260	8,388	308,725
Workers only/Wage $= 0$	$62,909$	245,816
≥ 10 years in the panel	83,165	162,651
year ≤ 1997	36,269	126,382
Only males	63,571	**62,667**

Table 2: QP Sample Selection (Database years 1987-2017)

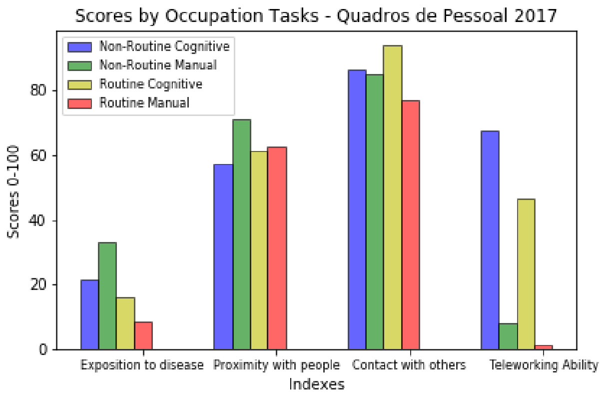
	Dropped	Remaining
Initial Sample 1987-2017		76,555,445
Missing Age	441,822	76,113,629
Age $25-64$	$11,550,875$	64,562,754
Missing Wage	1,193,496	63,369,258
Miscoded Infos/Wage $= 0$	6,156,393	57,212,865
Praticante/Ajudante/Estagiario	1,524,276	55,688,589
Monthly Wage $<0.5 \times$ min.wage	1,741,557	53,947,032
Monthyl Hours $\leq 260/12$	96,458	53,850,578
≥ 10 years in the panel	17,064,774	36,785,804
Only males	16,095,688	**20,690,116**

[4]Series for National Income and Product Account have been obtained from Bureau of Economic Analysis website. The series is obtain as the ratio between National Income from Wages and Salaries and Full-time equivalent employees, which includes employees on full-time schedules plus the number of employees on part-time schedules converted to a full-time basis.

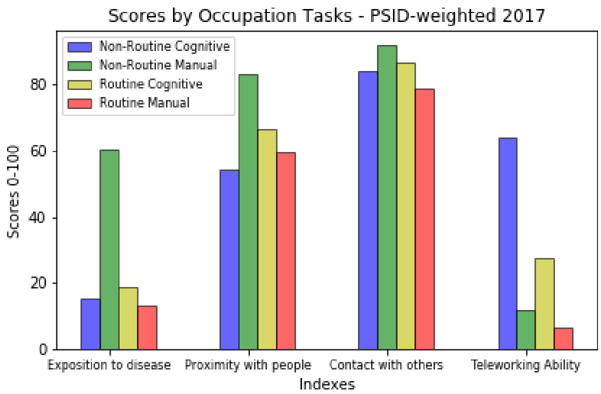
[5]Minimum wage is calculated hourly for US and monthly for Portugal, source: Federal Reserve Economic Data (US) and OECD Labour Data (Portugal).

Impact of Covid-19

The current pandemic situation and the lockdown measures adopted by governments in many countries obliged people to work from home but, simply, many occupations cannot be done from home. To understand and link our results to the recent developments in people working conditions we replicate and improve the mapping made by Dingel & Neiman (2020)[6] conforming it to the PSID and Quadros de Pessoal samples in order to define whether occupations can be performed at home or not. For U.S. we used the same crosswalk between SOCs and Census made for mapping occupation categories, for Portugal the method is described in details in the appendix. The teleworking index we use is based on two O-NET surveys questioning the "work context" and "generalized work activities" and in case that respondents' job need to be done outdoor, or require the use of specific machines for which the use of other facilities is needed, then that occupation cannot be performed at home and the occupation receives a teleworking index equal to 0. We also mapped every worker with three other indexes obtained from O-NET surveys: i) exposition to diseases or infections, ii) contact with others and iii) proximity with the others[7].



(a) **Portugal**: 57,354,268 Observations



(b) **US**: 245,316 Survey-Weighted Observations

For both Portugal and U.S. we observe a large difference between the possibility of working from home between workers that perform cognitive and manual jobs. This difference motivates our choice to delve into the sources of inequality generated by skill-biased technological change[8]. Within cog-

[6]They propose a mapping 6-digits code SOCs to 2-digit ISCOs and use a 2-digits occupational data for countries other than US using country-level data from ILOSTAT. We match with 4-digits precision.

[7]More details about these surveys and indexes can be found in Appendix B. For a comprehensive description of the teleworking index refer to Dingel & Neiman (2020) appendix.

[8]Coelho (2020) and Ferreira (2019).

nitive occupations the routine component of the occupation task has an important role in determining the possibility of teleworking; this effect is stronger for the US where the difference between non-routine cognitive and routine cognitive is approximately $40p.p.$ Among the other measures of infection riskiness, non-routine manual results the category most exposed to viruses and diseases due to many occupations involved in the health care industry, as for example dental hygienists, critical care nurses, hospitalists and respiratory therapists. Table 15 shows that for Portugal teleworking feasibility of tasks is increasing with wage, this is not the same for U.S., table 16, where there is no clear correlation between wage and teleworking ability[9]. The effects of restriction measures are not symmetric across sectors, figure 8 confirms that for Portugal many manual occupations cannot be performed at home. Moreover manual workers in manufacturing, wholesale, retail trade, construction and food service industries comprehend large part of the national labour force and produce a remarkable component of the national value added in GDP. This could have dramatic consequences for the economy if the restrictions continue to be strict.

For the U.S., figure 9, there is a clear separation between the non-routine cognitive share of each sector and the others categories; this difference in teleworking could further increase the demand for non-routine cognitive labor and decrease the demand for manual and routine workers. Furthermore, considering that a large part of the labor force is at the bottom of the teleworking scale, earning inequality is very likely to increase. Susceptibility index[10] is quite heterogeneous across sectors, both for the U.S. and Portugal.

Estimation of the labor income processes

One of the main contributions of this work is the estimation of the permanent component dispersion over time both using the previously described samples from PSID and Quadros de Pessoal. We estimate the evolution of the dispersion on the permanent and transitory components of labor income processes overtime following Brinca *et al.* (2016) and Chakraborty *et al.* (2015). Different characteristics determine the number of efficient units of labour the individual is endowed with, namely age j plus a set of year dummies $D_t'\xi_i$:

[9] Unfortunately PSID does not capture efficiently the heterogeneity between occupation as only a sample of families is chosen.

[10] Obtained as a combination of the previously stated 3 measures of infection riskiness.

$$w_{i,t} = e^{\gamma_1 j + \gamma_2 j^2 + \gamma_3 j^3 + D_t' \xi_i + u_{i,t}}, \tag{1}$$

The productivity shock u follows an AR(1) process given by:

$$u_{i,t} = \rho_u u_{i,t-1} + \alpha_i + \epsilon_{i,t} \tag{2}$$

where $\alpha_i \sim N(0, \sigma_\alpha^2)$ represents the individual permanent ability and $\epsilon_{i,t} \sim N(0, \sigma_\epsilon^2)$ the idiosyncratic shock to the productivity shock process. Thanks to this specification, we are able to separate the permanent component from the individual fixed effect and the random noise in the productivity process. This specification outlines the same sources of heterogeneity of Heathcote *et al.* (2017): (i) the individual fixed effect defines innate individual ability; (ii) the realization of idiosyncratic efficiency shocks determines individual fortune in labor market outcomes and (iii) experience of the individual in the labour market [11]. We inflation adjust the nominal wages using CPI inflation series from OECD with 2015 as base year. We found that the individual fixed component contribution to wage dispersion is increasing overt time, as the ratio between the variance of individual ability and the variance of idiosyncratic shock increases.

To understand their evolution over time, we estimated the above equation using a rolling window of 10 years, including year dummies in the wage equation:

$$ln(w_{it}) = D_t' \xi_i + \gamma_1 j + \gamma_2 j^2 + \gamma_3 j^3 + u_{i,t} \tag{3}$$

To assess the impact of skill-biased and factor-biases technological change, we included dummies for different occupation categories in the above equation and it becomes:

$$ln(w_{it}) = D_t' \xi_i + \gamma_1 j + \gamma_2 j^2 + \gamma_3 j^3 + NRM_{it} + NRC_{it} + RC_{it} + RM_{it} + u_{i,t} \tag{4}$$

This result is robust to different specification: for the US, having also non-workers in the initial sample, we use the Heckman estimation method used in Chakraborty *et al.* (2015) that use a two step approach to control for selection into the labor market, as described in Heckman (1976) and

[11] In Heathcote *et al.* (2017) they use individual working effort instead of labor market experience.

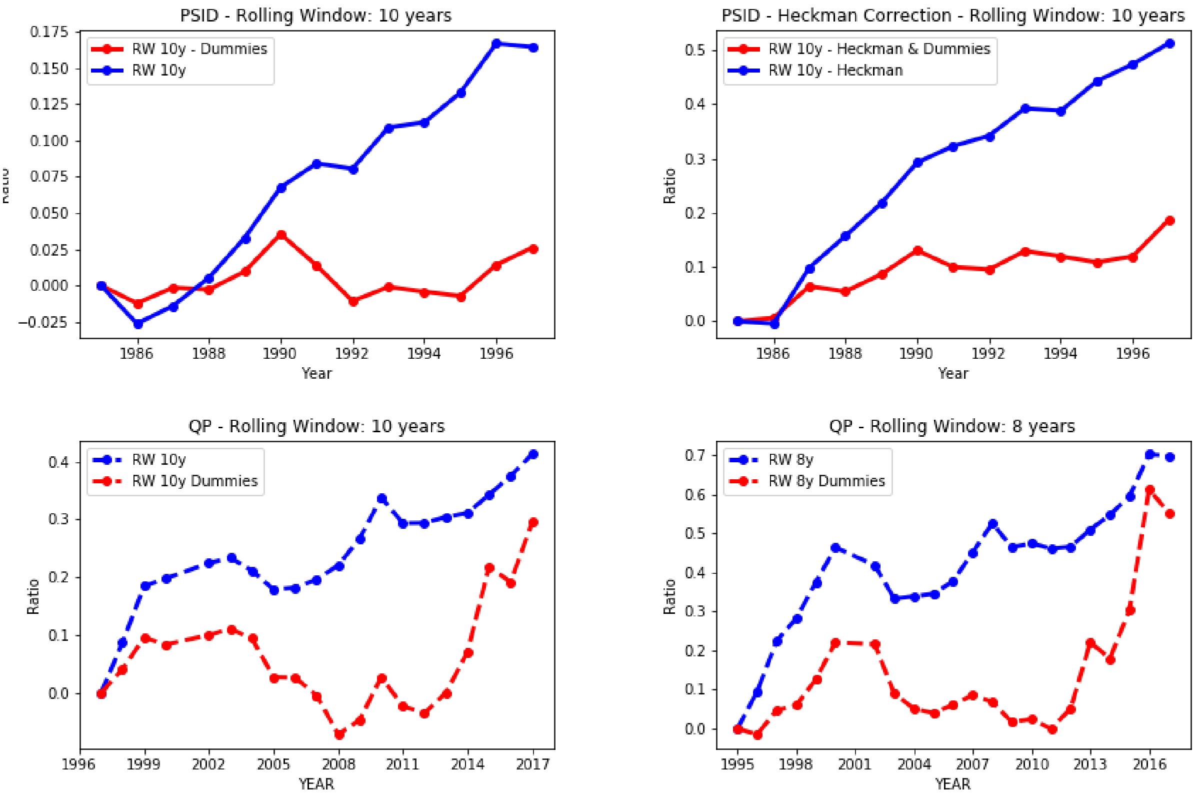


Figure 5: The blue lines represent are obtain using the base specification, the red lines are obtained from the wage equation that includes dummies. On the y-axis, we plotted the log-change in the ratio between the variance of the permanent component and the variance of the idiosyncratic shocks resulting from the residual of the wage equation.

Heckman (1977), whereas for Portugal, having only workers in the dataset, we use different size for the rolling window as robustness check. More information on the Heckman selection equation can be found in the Appendix.

This change in wage dispersion determinants is originated by different dynamics for U.S. and Portugal. For the U.S. , tables 6 and 7, the variance of individual ability is increasing over time more than the variance of the residual idiosyncratic shock. This increase, together with the decrease in permanent component persistency and the lower impact of individual experience on wage, is likely to have a large effect on long-run earnings, as suggested by Autor *et al.* (2006) and Acemoglu & Restrepo (2020). Including dummies for different tasks in the regression, the increase in individual ability dispersion is much lower compared to the baseline specification, meaning that different occupation categories explain two thirds of the total increase in the variance of labor income.

For Portugal, as shown in tables 9 and 10, the same increase in the ratio is driven by different dynamics[12] as now the noisy component dispersion is decreasing more than individual ability variance and the persistency of the residual is increasing across years, whereas the impact of individual experience increases particularly from 2006. When we include dummies in the wage regression these trends do not change, but the dispersion of individual ability decreases in size whereas the variance of transitory component remains approximately the same. This underlines the impact of investment-specific technological change[13] and the drop in the relative price of investment plays in explaining increases in wage premia and consequently income and earnings inequality.

3 The Model

The model is an incomplete markets economy with overlapping generations heterogeneous agents and partial uninsurable idiosyncratick risk generating both income and wealth distribution. Households are differentiated into Cognitive and Manual, according to the level of education required to perform daily tasks.

Demographics

In the economy there are J overlapping generations of households, who start life at age 20 and enter retirement at age 65. After retirement, households face an age-dependent probability of death $\pi(j)$ and when they reach 100 they die with certainty. Time is discrete and one period is 1 year, indeed there are 40 model periods of active work life. Population size is considered to be constant over time. We define the age-dependent probability of survive as $w(j) = 1 - \pi(j)$, so that the mass of retired agents of age $j >= 65$ still alive at any given period is $\Omega_j = \prod_{q=65}^{q=J-1} w(q)$. Given that there are no annuity markets, a fraction of households leave unintended bequests, denoted by Γ (per-household bequest), which are redistributed in a lump-sum manner between the household that are currently alive. Moreover, retired households receive a subsidy from the government Ψ. Households are subject to different persistent idiosyncratic productivity shocks, permanent ability, asset holdings and a discount factor $\beta \in (\beta_1, \beta_2, \beta_3, \beta_4)$ uniformly distributed across agents.

[12]We capture dynamics from 1987 for Portugal, period for which U.S. estimates are different.
[13]Brinca *et al.* (2019b)

Preferences

Agents utility is decreasing in work hours $n \in (0, 1]$ and increasing in consumption c and takes the following CRRA representation:

$$U(c, n) = \frac{c^{1-\sigma}}{1-\sigma} - \chi \frac{n^{1+\eta}}{1+\eta} \tag{5}$$

In the above equation, χ is the disutility from work and η the Frisch labor elasticity. For retired households, utility function is extended with the scrap value of the bequest they leave to living generations:

$$D(h'_t) = \phi log(h'_t) \tag{6}$$

Technology

By means of a linear production technology, intermediate inputs are transformed in consumption and investment goods. A quantity z_t^c of intermediate input is used to produce one unit of consumption good, that represents the numeraire and it is sold to households and government at price P_t^c. The transformation technology is:

$$C_t + G_t = z_t^c \tag{7}$$

with z_t^c being the quantity of input paid p_t^z from a representative intermediate goods firm. Assuming that we are in perfect competition the environment, the final consumption good will have price equals to its marginal cost of production, hence:

$$P_t^c = 1 = p_t^z \tag{8}$$

The investment good, X_t uses the transformation technology:

$$X_t = (\frac{1}{\xi_t}) z_t^x \tag{9}$$

where ξ_t is the level of technology used in the production of X_t relative to the final consumption good and $z_t^x(z)$ represents the quantity of input z used to produce the final investment good. Through the zero profit condition, the price of the investment good can be expressed as:

$$p_t^x = \xi_t p_t^z = \xi_t \tag{10}$$

and ξ_t can be interpreted as the relative price of the investment good. The production function used in the economy has constant return to scale and uses capital and labor as inputs, with form $y_t(z) = F(k_t(z), n_t^C(z), n_t^M(z))$, where r_t is the rental rate of capital and w_t^C and w_t^M are the costs for cognitive and manual labour. We measure aggregate demand $Y_t = C_t + G_t + \xi_t X_t$ in terms of the consumption good. Inputs for production maximize firms' profit function:

$$\Pi_t^z = p_t^z y_y - r_k K_t - w_t^C n_t^C - w_t^M n_t^M, \tag{11}$$

subject to:

$$y_t = z_t^c + z_t^X = C_t + G_t + \xi_t X_t = Y_t \tag{12}$$

which implies $p_t^z = P_t^C = 1$, $N_t^s(z) = N_t^s$,$z_t^c = C_t + G_t$, $z_t^X = \xi_t X_t$, $y_t = Y_t = C_t + G_t + \xi_t X_t$, and $Y_t = F(K_t, N_t^C, N_t^M)$, where K_t, N_t^C, N_t^M are the aggregate values of capital and the labor varieties. The production function of the representative intermediate goods firm takes the following functional form:

$$Y_t = F(K_t, N_t^C, N_t^M) = A_t \left(\phi_1 Z_t^{\frac{\sigma-1}{\sigma}} + (1-\phi_1) N_t^{M,\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{13}$$

$$Z_t = \left(\phi_2 K_t^{\frac{\rho-1}{\rho}} + (1-\phi_2) N_t^{C,\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}}, \tag{14}$$

where A_t is the total factor productivity, ρ the elasticity of substitution between capital and non-routine labor, ϕ_1 and ϕ_2 are factor shares, σ is the elasticity of substitution between the composite of those factors and routine labor. Capital depreciates at rate δ and X_t represents the aggregate gross investment; the transition equation is:

$$K_{t+1} = (1-\delta)K_t + X_t. \tag{15}$$

Government

The government manages the social security system balancing tax rates for employees and employers, defined respectively by τ_{ss} and $\overline{\tau}_{ss}$, and benefits paid to retirees Ψ. Expenditures on pure public consumption goods G_t, interest payment on the national debt rB_t and the lump sum redistribution g_t, are assumed to be separable in the utility function and are financed by the government through taxes on consumption (τ_c), labor (τ_l) and capital (τ_k) income. The government uses flat rates on τ_c and τ_k, whereas the labour income tax follows a non-linear functional form as in Bénabou & Tirole (2002) and Heathcote *et al.* (2020):

$$y_a = 1 - \theta y^{-\theta_1} \tag{16}$$

where y is the pre-tax labour income, y_a the after-tax labour income and θ_1 and θ_2 represent respectively the level and progressivity of the tax schedule. The government budget constraint is defined as follows:

$$g_t(\sum_{j\geq 45} \Omega_j) = T_t - G_t - r_t B_t \tag{17}$$

$$\Psi_t(\sum_{j\geq 45} \Omega_j) = R_t^{ss} \tag{18}$$

with R_t^{ss} being the social security revenues and T_t the other tax revenues.

Asset Structure

The economy has two types of assets, capital (k) and government bonds (b). The relative price of the equipment good is constant as there is no investment-specific technological change in the steady-state , i.e. $\xi = \xi'$. Moreover, the return rate on the bond must satisfy:

$$\frac{1}{\xi}[(\xi + (r - \xi\delta)(1 - \tau_k)] = 1 + R(1 - \tau_k), \tag{19}$$

that follows the non-arbitrage condition ensuring that investing in capital has the same return as investing in bonds. The state variable observed by the consumer when taking decision is:

$$h \equiv [\xi + (r - \delta\xi)(1 - \tau_k)]k + (1 + R(1 - \tau_k))b. \tag{20}$$

With the non-arbitrage condition the previous equation can be rewritten as:

$$h = \frac{1}{\xi}[\xi + (r - \delta\xi)(1 - \tau_k)](\xi k + b), \tag{21}$$

Household Problem

In every period the agent is endowed with certain characteristics, as age j, asset position h, time discount factor $\beta \in \{\beta_1, \beta_2\}$, permanent ability α, a persistent idiosyncratic productivity shock

u and, according to his skills level, a labor variety supply constant over time $s \in \{C, M\}$. Consumption c, hours worked n_C and nM and future asset holdings h' are the control variables of the optimization process. Each household is subject to the budget constraint :

$$c(1+\tau_c) + \xi k' + b' = (\xi + (r - \delta\xi)(1-\tau_k))k + (1 + R(1-\tau_k))b + \Gamma + g + Y^N \tag{22}$$

which with equations 20 and 21 becomes:

$$c(1+\tau_c) + qh' = h + \Gamma + g + Y^N \tag{23}$$

where Y^N is the labour income of the household after deductions. Hence, the household problem assumes the following recursive form:

$$V(j, h, \beta, \alpha, u) = \max_{c,n,h'} \left[U(c, n) + \beta \mathbb{E}_{u'} \left[V(j+1, h', \beta, \alpha, u') \right] \right]$$

$$s.t. : \qquad c(1+\tau_c) + qh' = h + \Gamma + g + Y^N$$

$$Y^N = \frac{nw(j, a, u)}{1 + \tilde{\tau_{ss}}} \left(1 - \tau_{ss} - \tau_l \left(\frac{nw(j, a, u)}{1 + \tilde{\tau_{ss}}} \right) \right)$$

$$n \in [0, 1], \quad h' \geq -\underline{h}, \quad h_0 = 0, \quad c > 0$$

When household retires the optimization problem is characterized by the age dependent probability of dying $\pi(J)$, retirement benefits and the bequest motive[14] $D(h')$ and it can be defined as:

$$V(j, h, \beta) = \max_{c,h'} \left[U(c, n) + \beta(1 - \pi(j)) V(j + \ h', \beta) + \pi(j) D(h') \right]$$

$$s.t. : \qquad c(1+\tau_c) + qh' = h + \Gamma + g + \Psi$$

$$h' \geq -\underline{h}, \quad c > 0$$

[14]Scrap value of the dynamic problem introduced by Brinca *et al.* (2016)

4 Calibration

The benchmark calibration of the model matches the US and Portuguese economies in 2010. The exogenous parameters are set to match the data, the endogenous parameters are estimated through simulated method of moments (SSM).

Preferences

The Frisch elasticity parameter follows Brinca *et al.* (2016) and is set to 1.0, at the same level of the risk aversion parameter.

Taxes and Social Security

We use the previously described labor income tax function proposed by Bénabou & Tirole (2002) for both US and Portugal, we estimate tax income level and progressivity parameters, respectively θ_0 and θ_1, using labor income tax data provided by the OECD. We then compute the weighted average over the population of θ_0 and θ_1 for different individuals, depending on whether they are single or married and on the number of children. Social Security parameters, $\tilde{\tau}_{ss}$ and τ_{ss}, are estimated from OECD Tax Data and τ_c and τ_k are taken from Trabandt & Uhlig (2011).

Parameters calibrated using SMM

We use simulated methods of moments to calibrate parameters that do not have an empirical counterpart. This method is used to estimate $\psi, \beta_1, \beta_2, \beta_3, \beta_4, h, \chi, T_C, T_M, \sigma_C$ and σ_M minimizing the loss function between moments from the model and moments observed in the data:

$$L(\psi, \beta_1, \beta_2, \beta_3, \beta_4, h, \chi, T_C, T_M, \sigma_C, \sigma_M) = ||M_m - M_d|| \tag{24}$$

used to match 75-100/all, $\bar{n}_C$, $\bar{n}_M$, K/Y,w_C/w_M, $\sigma_{ln(w);C}$, $\sigma_{ln(w);M}$, Q_{20}, Q_{40}, Q_{60} and Q_{80}. Table 3 and table 4 contains the estimated parameters and table 5 the endogenously calibrated parameters.

Table 3: Calibration Fit - United States

Data Moment	**Description**	**Source**	**Target**	**Model Value**
75-100/all	Average wealth of households 75 and over	US Census	1.31	1.33
$\bar{n}_C$	Fractions of hours worked - Cognitive	PSID	0.489	0.489
$\bar{n}_M$	Fractions of hours worked - Manual	PSID	0.501	0.51
K/Y	Ratio between capital and output	BEA	3.0	3.0
w_C/w_M	Wage Premium	PSID	0.519	0.518
var $ln(w)$ Cogn./Man.	Variance of the log wages	PSID	0.707 ;0.651	0.7067; 0.651

Table 4: Calibration Fit - Portugal

Data Moment	**Description**	**Source**	**Target**	**Model Value**
75-100/all	Average wealth of households 75 and over	Assumption	1.31	1.295
$\bar{n}_C$	Fractions of hours worked - Cognitive	QP	0.472	0.479
$\bar{n}_M$	Fractions of hours worked - Manual	QP	0.527	0.532
K/Y	Ratio between capital and output	PWT	3.229	3.20
w_C/w_M	Wage Premium	QP	0.623	0.624
var $ln(w)$ Cogn./Man.	Variance of the log wages	QP	0.388;0.154	0.374;0.155

Table 5: Parameters Calibrated Endogenously - US & Portugal

Parameters	**Description**	**Value - US**	**Value - PT**
ψ	Bequest utility	4.15	4.8
$\beta_1, \beta_2, \beta_3, \beta_4$	Discount factors	0.979;0,9355 0.9235;0.9235	0.981;0,942 0.940;0.925
h	Borrowing limit	0.115	0.075
χ	Disutility from work	2.55	2.0
T_C	lab.augmenting tech.Cognitive	1.1	1.0
T_M	lab.augmenting tech.Manual	0.9	1.1
σ_C, σ_M	Standard Deviations of ability	0.4725; 0.773	0.520; 0.291

5 Quantitative Results

Our main experiment consists in estimating how wage and earnings inequality change following the demand shocks caused by the pandemic outbreak. We argue that demand for many jobs that cannot be performed from home, as occupations in the hospitality and leisure services sector, will fall in the long run. Brinca *et al.* (2020) separate between demand and supply shocks, finding evidences of a predominant negative supply shock in the short run and correlation between both demand and supply shocks and teleworking ability for occupations. In this context, we estimate the impact of COVID-19 outbreak by applying the drop in working hours aggregating the drop in demand for each sector and weigthing occupations by teleworking ability, as we expect firms to adapt to the new social distancing norms. We found a large decrease in monthly hours worked for manual workers in almost every sector and a modest drop in hours worked by cognitive workers. Quadros de Pessoal, for structural reasons, gives a better representation of the effects on the whole labor market, as it includes employees from every industry, PSID includes only a panel of selected families so it does not capture entirely the heterogeneity of demand shocks.

Aggregating results we found that for Portugal the share of cognitive workers increases from 47.2% to 93.1% of the labor force, whereas manual workers decreasese to 6.8% from the pre-covid 52.7%. For the U.S. the impact has the same magnitude, going from 48.9% to 88.1% for cognitive workers and from 51.07% to 11.9% for manual workers. The effects in the short run[15] are quite strong although we expect that once the restrictions measures will be relieved the shock will be smoother and, in the long-run, many occupations will be readapted such that they can be performed from home. This will reduce the overall impact on hours worked but many manual occupation may be permanently replaced. The objective of this experiment is to study the heterogeneous impact of Covid-19 on cognitive and manual workers, and to do that we assume that only 20% of the observed demand shock will be permanent[16], so the demand shock will be -15.6% for the U.S. and -17.4% for Portugal and the share of hours worked by manual workers will respectively drop to 43.1% and 43.5%. Recalibrating the model to match the decrease in working hours for manual workers, we find that wage premium between cognitive and manual workers increase from the initially observed 0.52 to 1.83 for the U.S. and from 0.62 to 2.19 for Portugal, and the variance of log-earnings from 0.63 to 1.81 for the U.S. and from 0.44 to 1.49 for Portugal. The U.S. are characterized by higher inequality within same occupation-task group but are more advanced in the adoption of technological capital and have a higher share of skilled human capital. Portugal delay in using new technologies will foster a higher demand for cognitive-task occupations, which, in turn, will raise wage premium for cognitive workers.

6 Conclusions

In this paper we study the role of task complementarity in explaining an important component of earnings inequality, namely the task wage premia. As the relative price of capital drops, workers

[15]Figure 6

[16]Calculated on the shock estimated from data.

whose tasks are complementary[17] with capital tend to observe an increase in demand, whereas workers whose main tasks are substitutable[18], observe a drop. Empirical findings show that Portugal is experiencing the same labor market trends but is still lagging behind behing the U.S. due to the lower supply of skilled human capital which slows down the adoption of computer capital. We estimate income processes for US and Portugal, based on PSID and Quadros de Pessoal respectively, and find that in both instances, the variance of wages that is explained by an increase in the variance of permanent differences across individuals relative to the variance of transitory shocks. Under the assumption that workers tend to say in the same task-type occupations over their life course, the impact of changes in the relative demand of routine vs non-routine type of work on wage premia is going to be captured mainly through individual fixed effects. When we include dummies for the type of occupation the worker has, we can explain about two thirds of the total increase in the relative variance of earnings for the US and about 30% of the same increase for Portugal in the overall sample. This stresses the role that investment-specific technological change and the drop in the relative price of investment plays in explaining increases in wage premia and consequently income and earnings inequality. The recent Covid-19 pandemic is also likely to have an impact on earnings inequality, as low wage manual and routine workers are being disproportionally affected, since these tasks typically involve physical contact and cannot be performed from home. In order to study the impacts that social distancing may have on inequality in the future, we simulate a permanent change in the demand for workers in those occupations. We study these counterfactuals in a structural model and find that wage premium and variance of log-earnings increase significantly for both the US and Portugal, even if only a fifth of the observed drop in the relative demand for manual workers is observed in the long run. This relative drop in demand is justified by the fact that manual workers tend to be over-represented in jobs that are most affected by social distancing policies and less doable from home. In future works, we want to study the effects of the pandemic on wage and earnings inequality from the supply side and divide workers according to the four categories initially used in the empirical analysis. This would allow us to capture entirely the heterogeneous effects of demand and supply shocks on different workers categories.

[17] In our taxonomy, workers who perform mostly non-routine tasks involving cognitive work.

[18] Workers who perform mostly routine tasks involving manual work.

APPENDIX A

Figure 6: Decomposition of demand shocks between sectors in April 2020.

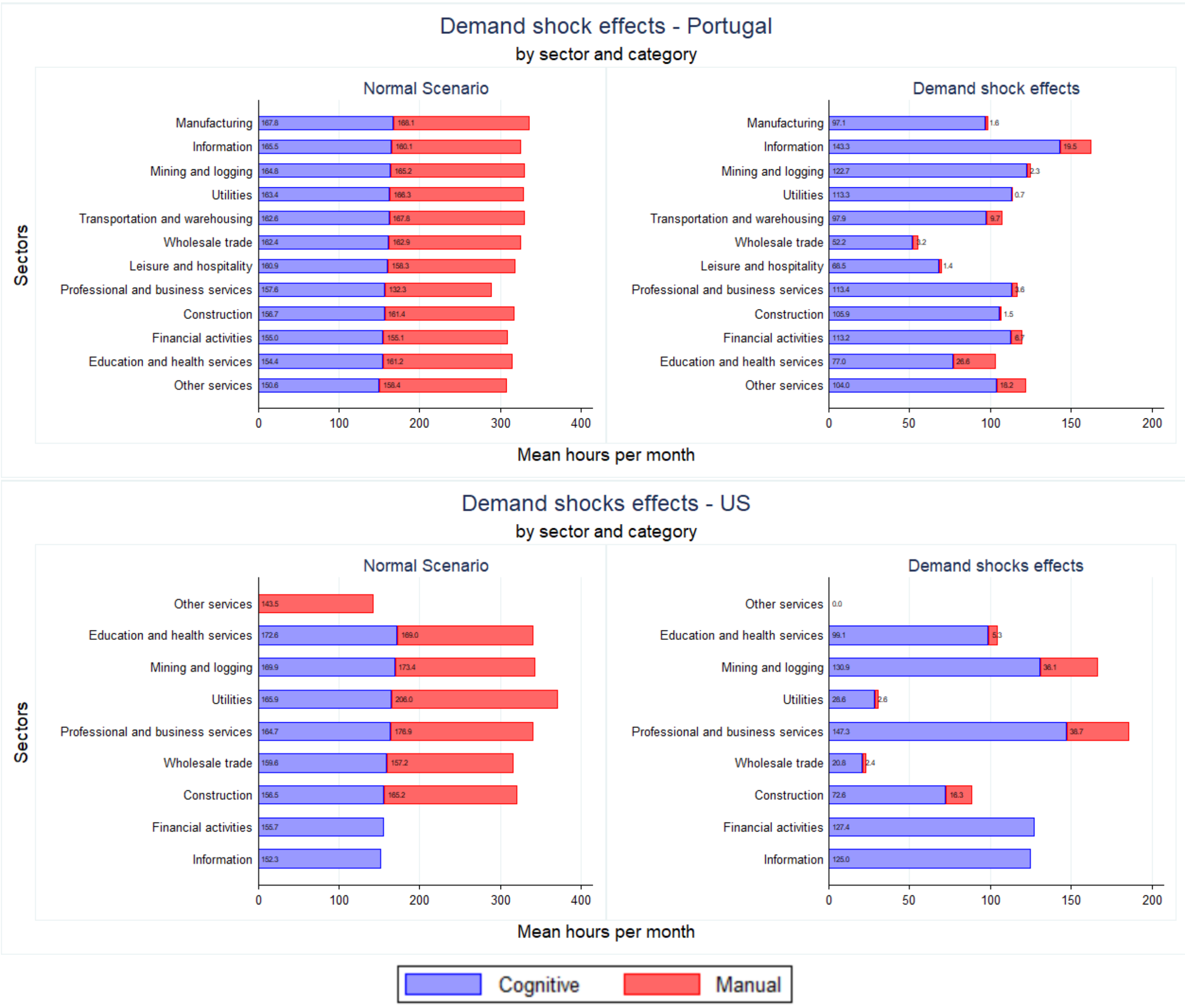


Table 6: U.S. - Heckman

Year	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
σ_α^2	0.401	0.401	0.424	0.437	0.454	0.473	0.475	0.485	0.504	0.505	0.519	0.525	0.540
σ_ϵ^2	0.316	0.317	0.318	0.319	0.321	0.322	0.319	0.322	0.327	0.328	0.328	0.327	0.330
ρ	0.278	0.282	0.276	0.267	0.258	0.242	0.246	0.238	0.220	0.215	0.202	0.186	0.165
γ_1	0.237	0.213	0.201	0.181	0.155	0.141	0.130	0.112	0.0864	0.0668	0.0562	0.0488	0.0389

Table 7: U.S. - Heckman with dummies

Year	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
σ_α^2	0.386	0.389	0.400	0.397	0.404	0.418	0.417	0.424	0.438	0.440	0.443	0.446	0.466
σ_ϵ^2	0.278	0.279	0.279	0.278	0.279	0.282	0.286	0.291	0.296	0.299	0.303	0.303	0.306
ρ	0.225	0.220	0.227	0.225	0.219	0.205	0.211	0.207	0.201	0.198	0.191	0.170	0.147
γ_1	0.188	0.162	0.153	0.140	0.133	0.125	0.117	0.105	0.0877	0.0767	0.0746	0.0771	0.0762

Figure 7: Task wage percentiles and minimum wage.

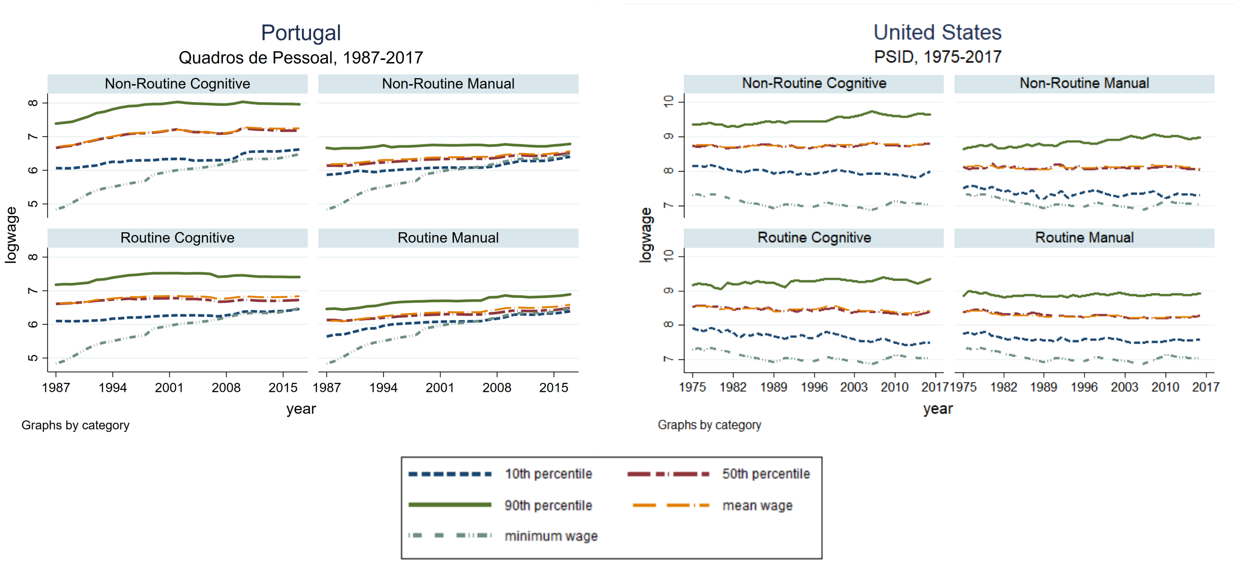


Table 8: 2010 Benchmark calibration for US

Description	Parameter	Value	Source
Preferences			
Inverse Frisch elasticity	η	1.000	Brinca et al.(2016)
Risk aversion parameter	λ	1.000	Brinca et al.(2016)
Labour Productivity			
Depreciation rate equipment	δ_e	0.105	BEA
Depreciation rate structures	δ_s	0.033	BEA
Parameter 1 age profile of wages	γ_1	0.236	Authors' Calculations
Parameter 2 age profile of wages	γ_2	-0.0012	Authors' Calculations
Parameter 3 age profile of wages	γ_3	$1.58e^{-06}$	Authors' Calculations
Variance of idiosyncratic shock	σ_u	0.330	Authors' Calculations
Persistence idiosyncratic risk	ρ_u	0.165	Authors' Calculations
Technology			
Share of income which goes to structures	α	0.151	Authors' Calculations
Share of the ICT cap/Cognitive composite	ϕ	0.469	Eden and Gaggl (2018)
Share of the ICT cap in the ICT Cognitive composite	ϕ	0.300	Eden and Gaggl (2018)
Elasticity of substitution of the ICT cap/Cognitive composite	ρ	1.558	Eden and Gaggl (2018)
TFP	A	1.000	Normalization
Relative price of investment	I_p	1.000	Normalization
Employment share (headcount) Cognitive group	emp_C	0.650	Authors' Calculations
Government and Social Security			
Consumption tax rate	τ_c	0.054	Trabandt and Uhlig (2011)
SS tax employer	τ_{ss}	0.078	OECD Tax Data
SS tax employee	$\tilde{\tau_{ss}}$	0.077	OECD Tax Data
Capital income tax rate	τ_k	0.469	Trabandt and Uhlig (2011)
Tax scale parameter	θ_0	0.850	Implied value from θ_1
Tax progressivity parameter	θ_1	0.160	Ferriere and Navarro (2016)
Government debt to GDP	B/Y	0.880	(FRED) Average 2008-2012
Government spending to GDP	G/Y	0.213	FRED

Table 9: QP: 10 years RW.

Year	σ^2_α	σ^2_ϵ	ρ	γ_1
1997	0.450	0.238	0.110	0.0284
1998	0.459	0.232	0.109	0.0147
1999	0.468	0.226	0.105	0.00280
2000	0.463	0.222	0.122	-0.00827
2002	0.467	0.221	0.125	-0.0131
2003	0.457	0.215	0.151	-0.0106
2004	0.441	0.210	0.188	-0.00412
2005	0.428	0.207	0.216	0.00650
2006	0.422	0.204	0.232	0.0192
2007	0.420	0.201	0.239	0.0336
2008	0.423	0.200	0.236	0.0493
2009	0.428	0.198	0.229	0.0654
2010	0.439	0.196	0.210	0.0810
2011	0.426	0.194	0.237	0.0974
2012	0.422	0.193	0.246	0.112
2013	0.420	0.191	0.251	0.123
2014	0.415	0.188	0.257	0.131
2015	0.416	0.185	0.257	0.138
2016	0.415	0.182	0.260	0.142
2017	0.414	0.178	0.261	0.146

Table 10: QP: 10 years RW with dummies.

Year	σ^2_α	σ^2_ϵ	ρ	γ_1
1997	0.366	0.255	0.129	0.0312
1998	0.370	0.252	0.133	0.0218
1999	0.371	0.246	0.136	0.0144
2000	0.364	0.243	0.155	0.00729
2002	0.366	0.242	0.150	0.00358
2003	0.363	0.239	0.162	0.00510
2004	0.353	0.234	0.194	0.00889
2005	0.337	0.232	0.239	0.0159
2006	0.332	0.228	0.260	0.0245
2007	0.338	0.236	0.257	0.0335
2008	0.325	0.235	0.295	0.0433
2009	0.325	0.232	0.305	0.0536
2010	0.340	0.234	0.265	0.0638
2011	0.327	0.230	0.297	0.0746
2012	0.320	0.227	0.316	0.0839
2013	0.320	0.223	0.320	0.0912
2014	0.325	0.218	0.314	0.0971
2015	0.339	0.212	0.291	0.101
2016	0.324	0.205	0.320	0.104
2017	0.332	0.199	0.300	0.104

Table 11: 2010 Benchmark calibration for Portugal

Description	**Parameter**	**Value**	**Source**
Labour Productivity			
Parameter 1 age profile of wages	γ_1	0.0638	Authors' Calculations
Parameter 2 age profile of wages	γ_2	-0.0020	Authors' Calculations
Parameter 3 age profile of wages	γ_3	$1.25e^{-4}$	Authors' Calculations
Variance of idiosyncratic shock	σ_u	0.196	Authors' Calculations
Persistence idiosyncratic risk	ρ_u	0.210	Authors' Calculations
Technology			
Employment share (headcount) Cognitive group	emp_C	0.472	Authors' Calculations
Government and Social Security			
Consumption tax rate	τ_c	0.215	Trabandt and Uhlig (2011)
SS tax employer	τ_{ss}	0.238	OECD Data
SS tax employee	$\tilde{\tau_{ss}}$	0.110	OECD Data
Capital income tax rate	τ_k	0.276	Trabandt and Uhlig (2011)
Tax scale parameter	θ_0	0.937	Implied value from θ_1
Tax progressivity parameter	θ_1	0.136	OECD Tax Data
Government debt to GDP	B/Y	0.447	IMF Data [19]
Government spending to GDP		0.37	OECD

[19] B/Y is the average of net public debt from 2008-12, IMF Data.

Heckman correction on returns to experiences and shocks processes

We use Heckman's selection model to control for selection bias only for PSID, as it contains information on non-workers, through a two-step statistical approach that will correct for the non-randomly selected sample. The first step consists in estimating the probability of entering the labor force through the selection equation:

$$\Phi(participation) = \Phi\left(Z_{it}'\epsilon + v_{it}\right) \tag{25}$$

where Z includes education, age, marital status and number of children. As we are we are using rolling window to capture the dynamics in the income process, time dummies for the specific window are used together with an interaction term between education and age. From these estimates the inverse of the Mills ratio, λ_i, is stored for each observation ($\lambda_i = \frac{\phi(z_i\epsilon_{it})}{\Phi(z_i\epsilon_{it})}$, with ϕ being the normal density and Φ the normal CDF), and we use it to obtain consistent estimate of the conditional expectation of logwage:

$$E\left[ln(w_{it})|X_{it}, workers = 1\right] = D_t'\xi + \gamma_1 j + \gamma_2 j^2 + \gamma_3 j^3 + \rho\sigma_u\lambda\left(Z_{it}'\epsilon\right) + u_{it} \tag{26}$$

u_{it} is then modelled as an AR(1) with panel data to separate the individual fixed effect from the permanent and the idiosyncratic components,

$$u_{i,t} = \rho_u u_{i,t-1} + a_i + \epsilon_{i,t}. \tag{27}$$

Stationary Recursive Competitive Equilibrium

An agent with characteristics (j, h, β, a, u) has measure $\Phi(j, h, \beta, a, u)$. We define the recursive competitive equilibrium in the following way:

1. The household's optimization problem is solved dynamically through the value function $V(j, h, \beta, a, u)$ and the policy functions $c(j, h, \beta, a, u)$, $h'(j, h, \beta, a, u)$ and $n(j, h, \beta, a, u)$, given factor prices and initial conditions.

2. Markets clear:

$$[\xi + (r - \xi\delta)(1 - \tau_k)]\left(K + \frac{1}{\xi}B\right) = \int h + \Gamma d\Phi$$

$$N^C = \int_{a>a^*} nd\Phi,$$

$$N^M = \int_{a\leq a^*} nd\Phi,$$

$$C + \xi X + G = Y.$$

3. Assuming perfect competition, firms' factor prices equalize marginal products:

$$r = \left[A^{\sigma-1}Y\right]^{\frac{1}{\sigma}} \phi_1 Z^{\frac{\sigma-\rho}{\rho\sigma}} \phi_2 \left(\frac{1}{K}\right)^{\frac{1}{\rho}},$$

$$w^C = \left[A^{\sigma-1}Y\right]^{\frac{1}{\rho}} \phi_1 Z^{\frac{\sigma-\rho}{\rho\sigma}} (1-\phi_2) \left(\frac{1}{N^C}\right)^{\frac{1}{\rho}},$$

$$w^M = (1-\phi_1)\left(\frac{A^{\sigma-1}Y}{N^M}\right)^{\frac{1}{\sigma}}.$$

4. The government budget balances:

$$g\int d\Phi + G + RB = \int \left(\tau_k (r/\xi - \delta)\left(\frac{h+\gamma}{\xi + (r-\xi\delta)(1-\tau_k)}\right) + \tau_c c + n\tau_l \left(\frac{nw(a,u,j)}{1+\tilde{\tau}_{ss}}\right)\right) d\Phi.$$

5. The social security system balances:

$$\int_{j\geq 45} \Psi d\Phi = \frac{\tau_{ss} \tilde{+} \tau_{ss}}{1+\tilde{\tau_{ss}}}\left(\int_{j<45} nwd\Phi\right).$$

6. The assets of the deceased at the beginning of the period are uniformly distributed among the living:

$$\Gamma \int w(j) d\Phi = \int (1 - w(j)) h d\Phi.$$

References

Acemoglu, Daron, & Restrepo, Pascual. 2018. The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, **108**(6), 1488–1542.

Acemoglu, Daron, & Restrepo, Pascual. 2020. Robots and jobs: Evidence from US labor markets. *Journal of Political Economy*, **128**(6), 2188–2244.

Autor, David H, Levy, Frank, & Murnane, Richard J. 2003. The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, **118**(4), 1279–1333.

Autor, David H, Katz, Lawrence F, & Kearney, Melissa S. 2006. The polarization of the US labor market. *The American economic review*, **96**(2), 189–194.

Baumeister, Christiane, & Hamilton, James D. 2015. Sign restrictions, structural vector autoregressions, and useful prior information. *Econometrica*, **83**(5), 1963–1999.

Bénabou, Roland, & Tirole, Jean. 2002. Self-confidence and personal motivation. *The quarterly journal of economics*, **117**(3), 871–915.

Brinca, Pedro, Holter, Hans A, Krusell, Per, & Malafry, Laurence. 2016. Fiscal multipliers in the 21st century. *Journal of Monetary Economics*, **77**, 53–69.

Brinca, Pedro, Homem Ferreira, Miguel, Franco, Francesco A, Holter, Hans Aasnes, & Malafry, Laurence. 2019a. Fiscal consolidation programs and income inequality. *Available at SSRN 3071357*.

Brinca, Pedro, Duarte, João B, Holter, Hans A, Oliveira, João G, *et al.* 2019b. *Investment-Specific Technological Change, Taxation and Inequality in the US*. Tech. rept. University Library of Munich, Germany.

Brinca, Pedro, Faria-e Castro, Miguel, Homem Ferreira, Miguel, & Holter, Hans. 2019c. The nonlinear effects of fiscal policy. *FRB St. Louis Working Paper*.

Brinca, Pedro, Duarte, Joao B., & e Castro, Miguel Faria. 2020. *Measuring Sectoral Supply and Demand Shocks during COVID-19*. Tech. rept. Federal Reserve of St.Louis.

Chakraborty, Indraneel, Holter, Hans A, & Stepanchuk, Serhiy. 2015. Marriage stability, taxation and aggregate labor supply in the US vs. Europe. *Journal of Monetary Economics*, **72**, 1–20.

Coelho, José. 2020. Universal basic income and skill-biased technological change.

Cortes, Guido Matias, Jaimovich, Nir, Nekarda, Christopher J, & Siu, Henry E. 2014. *The micro and macro of disappearing routine jobs: A flows approach.* Tech. rept. National Bureau of Economic Research.

Dingel, Jonathan I, & Neiman, Brent. 2020. *How many jobs can be done at home?* Tech. rept. National Bureau of Economic Research.

Eden, Maya, & Gaggl, Paul. 2015. *On the welfare implications of automation.* The World Bank.

Ferreira, Ana. 2019. Skill-biased technological change and inequality in the US.

Ferriere, Axelle, & Navarro, Gaston. 2014. The heterogeneous effects of government spending: It's all about taxes.

Fonseca, Tiago, Lima, Francisco, & Pereira, Sonia C. 2018. Job polarization, technological change and routinization: Evidence for Portugal. *Labour Economics*, **51**, 317–339.

Goos, Maarten, & Manning, Alan. 2007. Lousy and lovely jobs: The rising polarization of work in Britain. *The review of economics and statistics*, **89**(1), 118–133.

Goos, Maarten, Manning, Alan, & Salomons, Anna. 2009. Job polarization in Europe. *American economic review*, **99**(2), 58–63.

Guerreiro, Joao, Rebelo, Sergio, & Teles, Pedro. 2017. *Should robots be taxed?* Tech. rept. National Bureau of Economic Research.

Heathcote, Jonathan, Perri, Fabrizio, & Violante, Giovanni L. 2010. Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006. *Review of Economic dynamics*, **13**(1), 15–51.

Heathcote, Jonathan, Storesletten, Kjetil, & Violante, Giovanni L. 2017. Optimal tax progressivity: An analytical framework. *The Quarterly Journal of Economics*, **132**(4), 1693–1754.

Heathcote, Jonathan, Storesletten, Kjetil, & Violante, Giovanni L. 2020. Optimal progressivity with age-dependent taxation. *Journal of Public Economics*, 104074.

Heckman, James J. 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Pages 475–492 of: Annals of economic and social measurement, volume 5, number 4*. NBER.

Heckman, James J. 1977. *Sample selection bias as a specification error (with an application to the estimation of labor supply functions)*. Tech. rept. National Bureau of Economic Research.

Karabarbounis, Loukas, & Neiman, Brent. 2014. The global decline of the labor share. *The Quarterly journal of economics*, **129**(1), 61–103.

Krusell, Per, Ohanian, Lee E, Ríos-Rull, José-Víctor, & Violante, Giovanni L. 2000. Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica*, **68**(5), 1029–1053.

Mendoza, Enrique G, Razin, Assaf, & Tesar, Linda L. 1994. Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption. *Journal of Monetary Economics*, **34**(3), 297–323.

Mirrlees, James A. 1971. An exploration in the theory of optimum income taxation. *The review of economic studies*, **38**(2), 175–208.

Nóbrega, Valter. 2020. Optimal Taxation and Investment-Specific Technological Change.

Trabandt, Mathias, & Uhlig, Harald. 2011. The Laffer curve revisited. *Journal of Monetary Economics*, **58**(4), 305–327.

APPENDIX B

Quadros de Pessoal matching

Algorithm for matching occupations in Quadros de Pessoal Following Fonseca *et al.* (2018), we use the same algorithm that they implemented which re-codes occupations based on the most frequent changes. The procedure is as follows: let $occupation_t^i$ be the occupation of worker i in year t, so we generate the matrix of $occupation_t^i$ and $occupation_{t+1}^i$, where the worker i is observed in both t and $t+1$ and finally we aggregate the results by the mode of $occupation_{t+1}^i$. This algorithm was used for consolidating the matching already generated by the official crosswalks between *CPP 2010* → *CNP 1994* between 2010 and 2009, *CNP 1994 4d* → *CNP 19943d* between 2007 and 2006 and *CNP 1994 3d* → *CNP 1985 3d* between 1995 and 1994. Our algorithm is matching with 4 digits precision when used between 2007-17 and 3 digits-precision between 1987-2007[20].

Matching Occupation from Census to Isco To apply the Cortes *et al.* (2014) task-based occupations split, we started from Census 2010 Occupational Code and mapped them to ONET-SOC Code 2010[21]. The method is describe in details in Appendix A. that has an almost unique one-to-one match with Census [22]; the latter is better matched to the ISCO-08 (International Standard Classification of Occupations). ISCO-08 is already embedded into the Portuguese Classification of Occupations 2010 (CPP 2010), the latest occupational code used in Portugal. In this way it is possible to create a consistent correspondence between Census Code 2010 and CPP 2010. This method covers the period 2010-2017. In some cases there is not a unique matching between Census-ISCO occupations and some codes have multiple values and each ISCO-08 is mapped to multiple Census Code 2010 values. After having created a full correspondence between the three codes, we defined a multiple dictionary that maps every ISCO-08 code to multiple Census values. The approach we followed here is based on Dingel & Neiman (2020) and occupations categories are defined by counting how many times ISCO-08 values fall in each category range, according to

[20]Fonseca *et al.* (2018) matching is at 2 digits level.

[21]We use the official crosswalks documents from the Bureau of Labor Statics. Some Official Crosswalks have been used in combination with files available on David Author's website.

[22]For multiple matching, we used the first occurrence in the list manually checking their consistency.

Cortes *et al.* (2014), in case of tie the occupation code is defined as 'Ambiguous' [23].

Matching Occupation across years To recover previous years mapping in Portugal we then use the crosswalk CPP 2010 to CNP 1994 [24]. To create a unique correspondence between occupations we implemented a specific algorithm that work as follows: starting from CPP 2010 values, if it has a unique correspondence, then the dictionary is updated with a one-to-one key to value object, otherwise when there are multiple values, the correct matching is recovered empirically, so the algorithm searches for the most common value in the panel containing common workers between 2009 and 2010, and assign the CNP 1994 code that is more recurrent, at the condition that it is above a certain recurrence threshold [25]. Crosswalks used for the analysis can be found in Appendix B. In doing that, we took into account also the changes that were made in Cortes *et al.* (2014) when passing from Census 2010 to Census 2002, in order to have a consistent mapping between US and Portugal. With method we covered the period 2010-1995. In 2007 the Occupational Code reduces to 3 digits only and for the majority of them a one-to-one matching is feasible, when there are multiple matching the same algorithm described before is used.

Table 12: **NACE Sectors**

Key	NACE Sector	Key	NACE Sector
A	Agriculture, forestry and fishing	K	Financial and insurance activities
B	Mining and quarrying	L	Real estate activities
C	Manufacturing	M	Professional, scientific and technical activities
D	Electricity, gas, steam and air conditioning supply	N	Administrative and support service activities
E	Water supply; sewerage, waste management and remediation activities	O	Public administration and defence; compulsory social security
F	Construction	P	Education
G	Wholesale and retail trade; repair of motor vehicles	Q	Human health and social work activities
H	Transportation and storage	R	Arts, entertainment and recreation
I	Accommodation and food service activities	S	Other service activities
J	Information and communication	U	Activities of extraterritorial organisations and bodies

[23] These cases represent only a small portion of the workers in the data, on the file sample, this group is made of 1,062,687, representing the 1,97% of the whole sample.

[24] Source: Official Crosswalk CPP 2010 → CNP 1994 Istituto Nacional de Estatistica.

[25] If the match is lower than 50% the occupation is defined as "Ambiguous"

Figure 8: **Portugal**

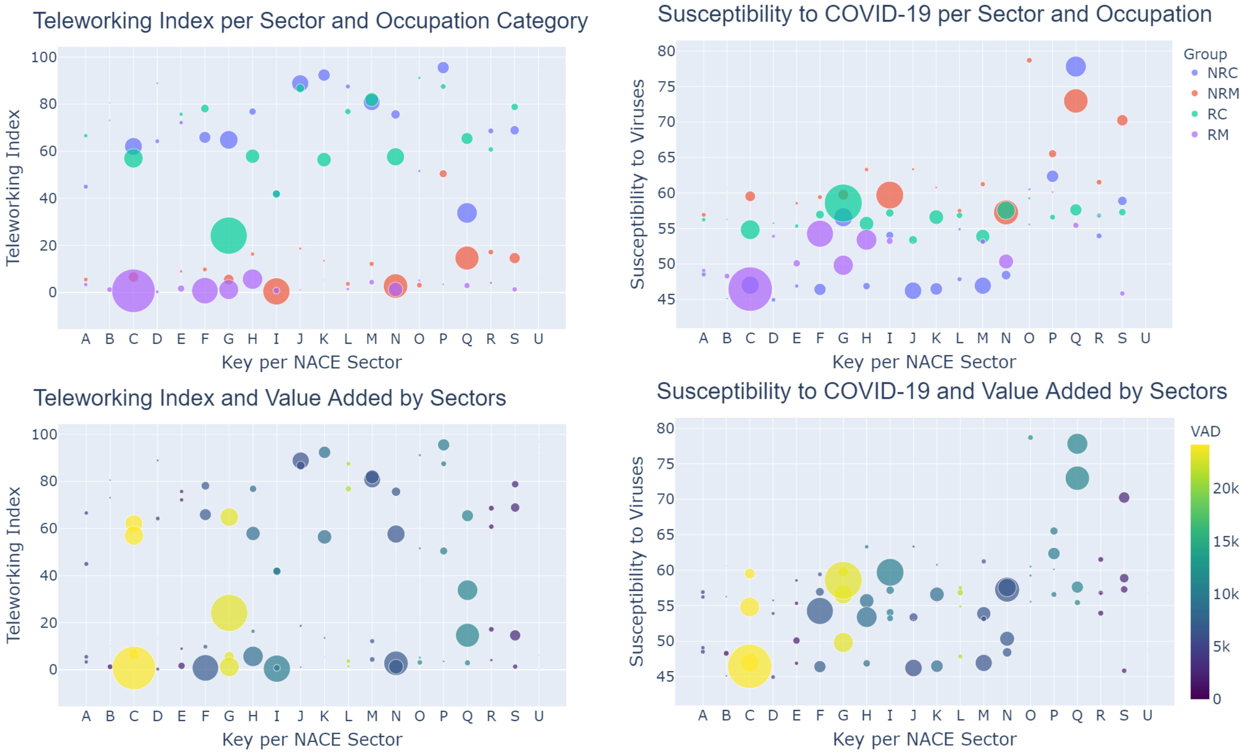
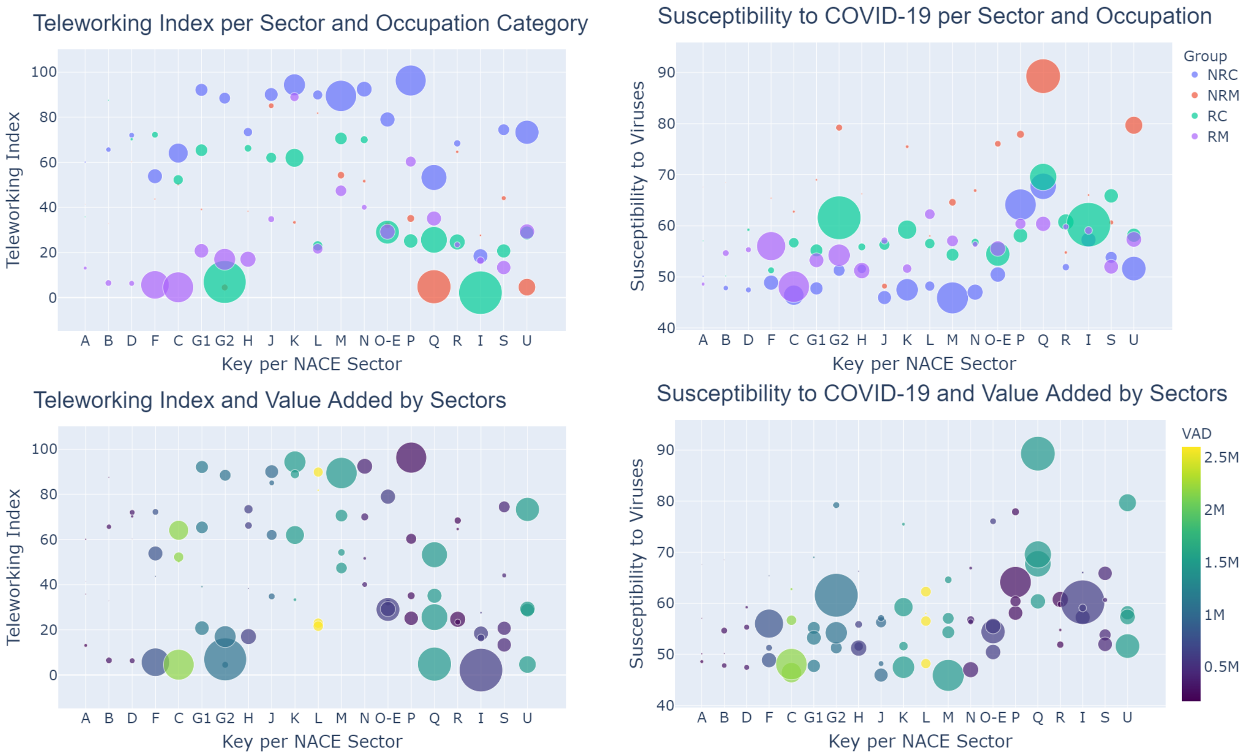


Figure 9: **United States**



Bubble size represents the number of workers in each sector, as benchmark routine manual Manufacturing has respectively 371,041 workers for Portugal and 5,186,890 for the US. Both value added are scaled in millions. U.S. sector have been mapped from NAICS to NACE, the former splits wholesale and retail trade in G1-G2.

Teleworking and Susceptibility to Covid-19 by earnings percentiles

Table 13: **Employment share** per percentile group - Portugal

Occupation Categories	Bottom 10%	10-25%	25-50%	50-75%	75-90%	Top 10%
Non-Routine Cognitive	0.91%	1.09%	1.58%	3.9%	6.05%	6.38%
Non-Routine Manual	4.08%	5.6%	6.46%	3.67%	0.62%	0.14%
Routine Cognitive	2.06%	2.85%	6.45%	8.19%	4.65%	2.47%
Routine Manual	2.74%	5.3%	10.09%	9.2%	4%	1.33%

Table 14: **Employment share** per percentile group - United States

Occupation Categories	Bottom 10%	10-25%	25-50%	50-75%	75-90%	Top 10%
Non-Routine Cognitive	1.48%	2.5%	5.1%	11.2%	10.63%	10.13%
Non-Routine Manual	2.4%	4.06%	4.17%	2.52%	1.2%	0.68%
Routine Cognitive	1.92%	3.38%	6.5%	5.9%	2.37%	1.82%
Routine Manual	1.5%	2.87%	4.77%	6.67%	3.8%	1.4%

Table 15: **Teleworking Index** per percentile group - Portugal

Occupation Categories	Bottom 10%	10-25%	25-50%	50-75%	75-90%	Top 10%
Non-Routine Cognitive	68.27	68.08	63.54	61.61	62.08	77.42
Non-Routine Manual	4.202	7.691	10.31	8.316	10.43	12.76
Routine Cognitive	33.61	34.18	36.07	48.01	59.93	71.15
Routine Manual	1.177	1.079	1.111	1.475	1.756	2.923

Table 16: **Teleworking Index** per percentile group - United States

Occupation Categories	Bottom 10%	10-25%	25-50%	50-75%	75-90%	Top 10%
Non-Routine Cognitive	66.15	76.78	70.67	66.58	62.72	54.46
Non-Routine Manual	6.482	12.29	13.54	16.18	10.20	7.075
Routine Cognitive	35.71	29.69	36.11	25.92	13.33	7.461
Routine Manual	7.873	6.448	8.110	6.718	6.422	3.138

Table 17: **Susceptibility Index** per percentile group - Portugal

Occupation Categories	Bottom 10%	10-25%	25-50%	50-75%	75-90%	Top 10%
Non-Routine Cognitive	56.49	55.18	54.86	56.38	58.41	51.11
Non-Routine Manual	59.70	61.95	65.56	64.73	63.99	64.54
Routine Cognitive	58.13	58.66	58.50	57.31	55.72	53.94
Routine Manual	50.21	48.25	49.19	49.72	50.27	52.59

Table 18: **Susceptibility Index** per percentile group - United States

Occupation Categories	Bottom 10%	10-25%	25-50%	50-75%	75-90%	Top 10%
Non-Routine Cognitive	48.11	47.40	51.71	51.14	51.13	52.28
Non-Routine Manual	80.82	77.60	77.46	76.18	81.60	82.45
Routine Cognitive	55.58	57.71	56.69	55.73	58.51	63.22
Routine Manual	48.02	48.54	50.33	50.79	51.30	51.98

Tax Function

Given the tax function

$$ya = \theta_0 y^{1-\theta_1}$$

which we employ, the average tax rate is defined as

$$ya = (1 - \tau(y))y$$

thus,

$$\theta_0 y^{1-\theta_1} = (1 - \tau(y))y$$

which implies:

$$(1 - \tau(y)) = \theta_0 y^{-\theta_1}$$

$$\tau(y) = 1 - \theta_0 y^{-\theta_1}$$

$$T(y) = \tau(y)y = y - \theta_0 y^{1-\theta_1}$$

$$T'(y) = 1 - (1 - \theta_1)\theta_0 y^{-\theta_1}$$

In this way, the tax wedge for any two incomes $(y_1;y_2)$ is given by:

$$1 - \frac{1 - \tau(y_2}{1 - \tau(y_1} = 1 - (\frac{y_2}{y_1})^{-\theta_1}$$

and therefore independent of the scaling parameter θ_0. In this manner, one can raise average taxes by lowering θ_0 and not the progressivity of the tax code, since the progressivity is uniquely determined by the parameter θ_1.

Information on O-NET Surveys

Exposition to diseases or infections This survey is based on the question "How often does this job require exposure to disease/infections?" and it is calculated as follows:

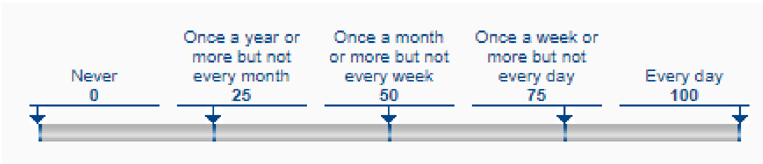


Figure 10: Source:O-NET online

Physical Proximity This survey is based on the question "To what extent does this job require the worker to perform job tasks in close physical proximity to other people?" and it is calculated as follows:

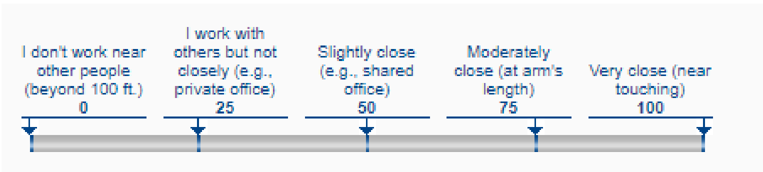


Figure 11: Source:O-NET online

Contact with others This survey is based on the question "How much does this job require the worker to be in contact with others (face-to-face, by telephone, or otherwise) in order to perform it?" and it is calculated as follows:

Figure 12: Source:O-NET online

Mapping indexes from O-NET surveys to Quadros de Pessoal

As previously underlined, between 4-digits ISCO and 6-digits SOCs there is not a one-to-one mapping and when it is the case the value from the O-NET index it is directly mapped to ISCO. The problem before was solved by maintaining the multiple matching and counted the occurrence of every occupation category within the same ISCO code. That solution was needed as the division is on a discrete scale. For O-NET surveys scores, the scale is continuous[26] so that when there are multiple matching we can "smooth" the division.

Following Dingel & Neiman (2020) and using U.S. employment data[27] we allocate the SOC's U.S. employment weight across the ISCOs according to the ISCO's employment share in Quadros de Pessoal. For example, if a particular SOC has 1000 U.S. employees and is associated with two ISCOs that count respectively 6000 and 2000 workers in Portugal, we allocate 3/4 of the employees (750) to the larger ISCO and 1/4 (250) to the smaller one with their respective scores. Once the process is done for whole SOCs we compute the weighted mean for each ISCO code using the U.S. employees share for each occupation.

Table 19: Transition matrix PSID U.S. 1969 – 2017

From ↓ To →	Non-Routine Cognitive	Non-Routine Manual	Routine Cognitive	Routine Manual
Non-Routine Cognitive	85.70	2.78	7.95	3.55
Non-Routine Manual	7.28	80.72	5.40	6.58
Routine Cognitive	13.25	3.50	78.64	4.59
Routine Manual	5.21	3.84	4.33	86.59

Table 20: Transition matrix (headcount) PSID U.S. 1969 – 2017

From ↓ To →	Non-Routine Cognitive	Non-Routine Manual	Routine Cognitive	Routine Manual
Non-Routine Cognitive	54.991	1.784	5.107	2.284
Non-Routine Manual	2.018	22.368	1.498	1.825
Routine Cognitive	5.654	1.495	33.545	1.962
Routine Manual	2.525	1.863	2.100	41.905

Table 21: Transition matrix Quadros de Pessoal 1987 – 2017

From ↓ To →	Non-Routine Cognitive	Non-Routine Manual	Routine Cognitive	Routine Manual
Non-Routine Cognitive	89.32	1.60	6.28	2.78
Non-Routine Manual	1.89	86.98	3.49	7.61
Routine Cognitive	4.63	2.23	90.44	2.67
Routine Manual	1.39	3.08	1.81	93.70

[26] Originally on a scale [0,100] or [0,1]. We scaled everything to [0,100].

[27] Occupational Employment Statistics.

Table 22: Transition matrix (headcount) Quadros de Pessoal 1987 – 2017

From ↓ To →	Non-Routine Cognitive	Non-Routine Manual	Routine Cognitive	Routine Manual
Non-Routine Cognitive	6,427,630	115,666	452,182	200,577
Non-Routine Manual	114,958	5,279,883	212,411	462,381
Routine Cognitive	469,993	226,660	9,162,156	270,846
Routine Manual	235,839	519,792	306,187	15,793,224

Characteristics of PSID

"Head" and "Spouse" For each family, the head component represents the person with the most financial responsibility in the household unit and has at least 16 years old. The head can also be female, and it is the case when she is married and her husband is present in the financial unit, also if she has a boyfriend and they are living together for at least one year. When the head of a family die, become incapacitated, or simply move out a new head is selected for the next surveys. Also, if the family splits then a new head is chosen and a new family unit is created, with the respective new head.

Heads are defined in the panel by using the sequence number 1, meaning that they represent the reference person in the household, in combination with the variable "Relation to Head" equal to 1 before the survey wave of 1983 and 10 after. Spouses have sequence number 2, and relation to head 2 before 1983 and 20 or 22 after (The latter indicates female cohabitors who have lived with Head for 12 months or more or who was mover-out nonresponse by the time of the interview)

File structure and data quality of the PSID Data have been retriwed from PSID website, where both family-level series and individual-level series have been used to import or generate time consistent series for different variables. Information from household variables have been disentangled to match only the relative individual to which they were referred to, and mainly all the variables used are from this source. The only variables imported from individual-level data were "Relation to Head" and "Interview Number 1968". By setting panel observations at individual level we did not have to create a matching between family unit and person ID, as frequently done in the literature.

Variables to be imported are designed with two different format, VRxxxx and ERxxxxx, where the former represent *final release* variables, the latter *early release* variables. Anyway, in the most recent years, all the variables have been updated and PSID decided to keep using ER format even if the variables where in final version. Moreover the different files that contains all the information about household income that before were contained the the *Hours of Work and Wage Files* have been unified in the family-level data (source: PSID Help center personal email).

Latino Sample This sample comprises approximately 2000 Latino households that have been added to the PSID In 1990, and they represented families from Mexico, Puerto Rico and Cuba. However after 1995 it was dropped because missing of an important part of the after 1968 immigrants, as Asians for example, and lack of sufficient funding. Many observations of this sample are miscoded in important characteristics, as wages and salaries, for this reason we decided to drop them from our panel.

Variable Definitions Most of the series contained in the family-level data are consistent and can be directly used, however some of them have been changed over the years, in these cases specific amendments have to be done. A specific description of all the variables modified follows here:

- **Education**: Total grades completed by the individual at the moment of the interview, before 1984 a unique variable included all type of education independently of whether it was college or high-school, after that the series has missing years and restarts only after 10 years, to overcome this issue we used the combination of two other series specifying respectively the years of education before college and years of college achieved.

- **Wage and Income from Labor - Head**: Total income from wages and salaries plus overtime, bonuses, commissions and other job-related income, which are unified till 1993, after that all extra-wages source of income are splitted in different series.

- **Wage and Income from Labor - Spouse**: Total income from labor, in 1984 any income from farming, business, market gardening, or roomers and boarders, labor-asset has been added to the series. The respective series with these amount have been used to clear and obtain only income from labor.

- **Sex of Spouse**: This variable has been imputed using combination of Sex of Head, Relation to Head and sequence number.