Piero De Dominicis
Nova School of Business and Economics
pierodedominicis1995@gmail.com
DOI: https://doi.org/10.14195/2183-203X_51_7

Routinization and Covid-19: A Comparison Between the United States and Portugal

Rotinização e COVID-19: Uma comparação entre os EUA e Portugal

Piero De Dominicis

Received for publication: December 20, 2019
Revision accepted for publication: May 22, 2020

ABSTRACT
The purpose of this study is to identify the role of automatization in increasing wage inequality, by comparing the United States to Portugal. Using the PSID and Quadros de Pessoal (Personnel Records), we find that labor income dynamics are strongly determined by the variance of the individual fixed component. This effect is drastically reduced by adding information on workers' occupational tasks, confirming that a decreasing price of capital and the consequent replacement of routine manual workers have deepened wage inequality. During the current crisis, we find that the ability to keep working is strongly related with the kind of occupation. As such, we foster 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 expense of manual workers.
Keywords: Routinization; wage inequality; Covid-19; income processes; teleworking.

JEL Classification: E21; E24; J24.

RESUMO
O objetivo deste estudo é identificar qual o papel da automatização no aumento da desigualdade salarial, fazendo uma comparação entre os Estados Unidos e Portugal. Usando PSID e Quadros de Pessoal, constate-se que a dinâmica dos rendimentos de trabalho é fortemente determinada pela variância da componente fixa individual. Este efeito é drasticamente reduzido ao adicionar informação sobre as tarefas ocupacionais dos trabalhadores, confirmando que a diminuição do preço do capital e a consequente substituição de trabalhadores manuais que executam tarefas rotineiras aprofundaram a desigualdade salarial. Durante a crise atual, constatamos que a capacidade de continuar a trabalhar está fortemente relacionada com o

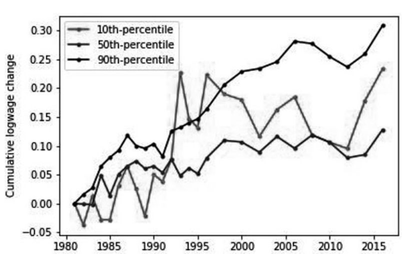
tipo de ocupação. Como tal, simulamos o impacto de um choque de procura permanente usando um modelo de gerações sobrepostas com mercados incompletos e agentes heterogéneos para prever quantitativamente o impacto da Covid-19 e das medidas de bloqueio no prémio salarial e na desigualdade de rendimentos. Conclui-se que que os prémios salariais e a dispersão dos rendimentos aumentam, sugerindo que a desigualdade de rendimentos aumentará em detrimento dos trabalhadores manuais.
Palavras-chave: Rotinização; desigualdade salarial; Covid-19; teletrabalho.

Acknowledgement: 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 Joao 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.

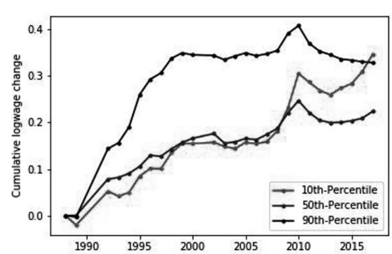
1. Introduction

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.

Figure 1: Real wage increase per percentiles



a) Portugal: 57.354 Observations

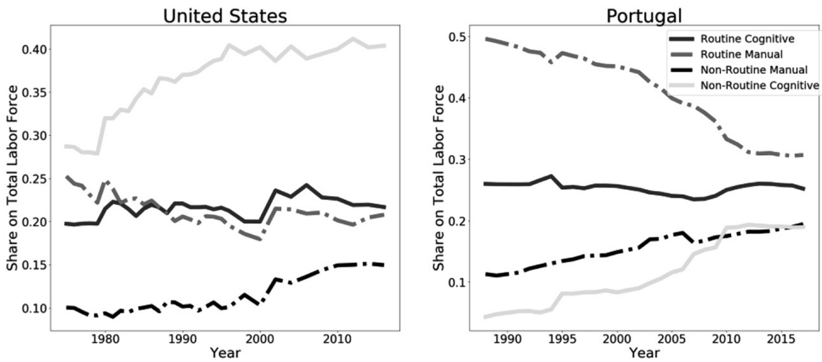


b) US: 24316 Survey-Weighted Observations

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 limit 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 expect the wage premium of non-routine workers to increase, following the drop in investment price and the decrease in tax progressivity[1], 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.

[1] Ferriere, and Navarro (2018), and Nóbrega (2020).

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 bigger decreasing 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 an upward sloping trend, steeper for cognitive occupations. The differences between US and Portugal are evident in terms of share of composition of the labor force as for 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 Portugal is experiencing labor market polarization but is lagging behind the United States in the adoption of computer capital. Fonseca et al. (2018) claims that routinization is the main cause of this shift in labor force composition in Portugal.[2]

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, and Restrepo (2017) 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, and Restrepo (2018) discuss the impact of increasing demand

[2] 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 18-20 in the Appendix B.

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, and 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, and Manning (2007) argue that the "routinization" hypothesis is the driving factor of the increase in highest and lowest wage occupations in the 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. In the spirit of Fonseca et al. (2018) we replicated figure 7: it shows that wage inequality is mainly determined by skills level but, more importantly, the increase in minimum wage had a positive impact for Portugal on the 10^{th} percentile as it may have impacted the wage convergence observed and the growth in wages for manual workers. For U.S. we cannot argue the same as the difference in wage still is clearly not impacted by the raise in minimum wage. Krusell et al. (2000) and Karabarbounis and Neiman (2014) argues that the more recent decline in relative price of investment has been triggered by the investment-specific technological change. Eden and Gaggl (2018) shows that 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 of the drop in labor share.

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

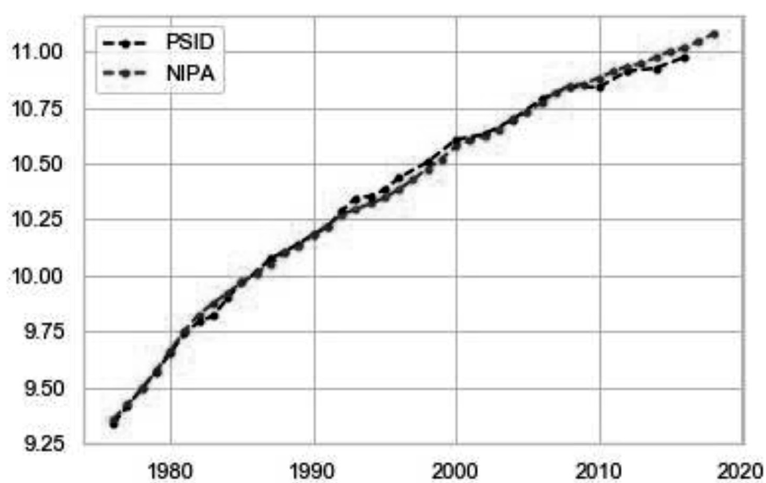
This database is a matched employer-employeee dataset created by the Portuguese Ministry of Labor in the 1980s, it includes Portuguese firms with at least one employee and does not take into account self-employed workers. The dataset covers the time period going from 1987 to 2017. The original occupations 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 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[3]. The sample is made of only heads and spouses of the families where the greatest level of accuracy in the data is guaranteed.

Observations with a wage lower than half of the minimum wage[4] have been dropped, also individual working less than 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.

Figure 3: PSID and NIPA in comparison



[3] Series for National Income and Product Account have been obtained from Bureau of Economic Analysis website. The series is obtained 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.

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

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

	Dropped	Remaining
Initial Sample 1969-2017		453,969
Hourly Wage ≤ 0.5 x min. wage	10,784	443,185
Age 25-64	126,072	317,113
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 1969-2017		76,555,445
Missing Age	441,822	76,113,629
Age 25-64	11,550,875	64,562,754
Miscoded Infos/Wage = 0	6,156,393	57,212,865
Praticante/Ajudante/Estagiario	1,524,276	55,688,589
Monthly Hours ≤ 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

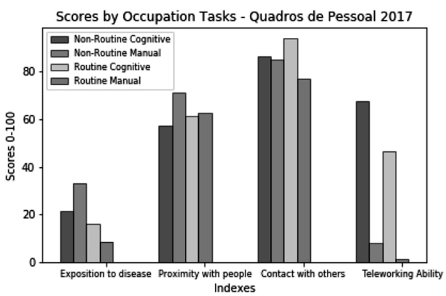
Impact of Covid-19

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 and Neiman (2020)[5] 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

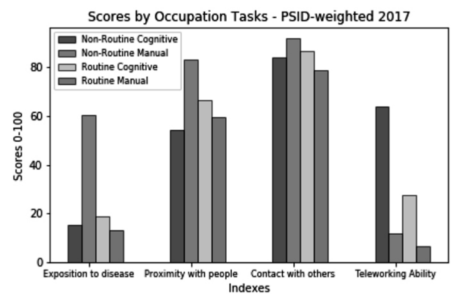
[5] They propose a mapping 6-digits code SOCs to 2-digit ISCOs and work with 2-digits occupational data for countries other than US using country-level data from ILOSTAT.

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[6].

Figure 4: Scores by Occupation for both surveys



a) Portugal: 57,354,268



b) US: 245,316 Survey-Weighted Observations

For both Portugal and the U.S. we observe large differences with respect to the possibility of working from home across types of occupations. This difference motivates our choice to delve into the sources of inequality generated by skill-biased technological change[7]. Within cognitive 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 40$p.p.$ 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 14 shows that for Portugal teleworking feasibility of tasks is increasing with wage, this is not the same for U.S., table 15, where there is no clear correlation between wage and teleworking ability[8]. The effects of restriction measures are not symmetric across sectors, figure 11 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. (in Figure 12), 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

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

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

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

bottom of the teleworking scale, earning inequality is very likely to increase. Susceptibility index[9] 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 $D'_t\xi_i$:

$$w_{i,t} = e^{y_1 j + y_2 j^2 + y_3 j^3 + D'_t \xi_i + u_{i,t}}$$

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

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

where $\alpha \sim N\left(0, \sigma_\alpha^2\right)$ represents the individual permanent ability and $\epsilon_{i,t} \sim N\left(0, \sigma_\epsilon^2\right)$ 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.[10] 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:

$$\ln(w_{it}) = D'_t\xi_i + y_1 j + y_2 j^2 + y_3 j^3 + u_{i,t}$$

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:

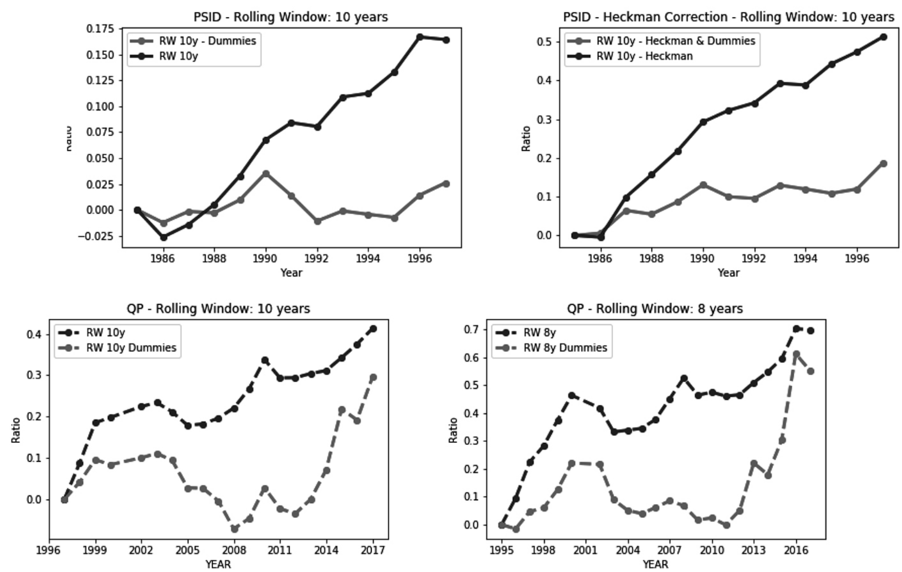
$$\ln(w_{it}) = D'_t\xi_i + y_1 j + y_2 j^2 + y_3 j^3 + NRM_{it} + NRC_{it} + RC_{it} + RM_{it} + u_{i,t}$$

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

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

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 Heckman (1977). 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.

Figure 5: PSID and QP over time



Note: 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 logchange 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.

This change in wage dispersion determinants is originated by different dynamics for U.S. and Portugal. For the U.S. (in 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 a 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, and Restrepo (2017). Including dummies for different tasks, the increase in individual ability dispersion is much lower meaning that different occupation categories can explain two thirds of the total increase in the relative variance of labor income.

For Portugal (in Tables 9 and 10), the same increase in the ratio is driven by different dynamics[11] as now the noisy component dispersion is decreasing more than individual ability variance, the persistency of the residual increases across years. The impact of individual experience increases particularly from . 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 (Brinca et al., 2019) and the drop in the relative price of investment plays in explaining increases in wage premia and consequently income and earnings inequality.

3. **Model and Calibration**

This paper uses the model 2 as introduced in the introduction chapter. However, the households are segmented into the two groups Cognitive and Manuel rather than Skilled and Non-Skilled. 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, and Tirole (2002) for both US and Portugal, and 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, and Uhlig (2011).

Parameter calibration using SMM

We use simulated methods of moments to calibrate parameters that do not have an empirical counterpart. This method is used to estimate ψ, β_1, β_2, β_3, β_4, h, χ, T_C, T_M, σ_C and

[11] We capture dynamics from 1987 for Portugal, period for which U.S. estimates are different.

σ_M minimizing the loss function between moments from the model and moments observed in the data:

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

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
β_1, β_2, β_3, β_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

4. **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 drop in the long run. Brinca et al. (2020) separate between demand and supply shocks, finding evidence 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 (in Figure 7) are quite strong although we expect that once the restrictions measures are relieved the effects become 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[12], so the demand shock will be

[12] Calculated on the shock estimated from data.

–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 0.518 to 1.83 for the U.S. and from 0.624 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.

5. **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 whose tasks are complementary[13] with capital tend to observe an increase in demand, whereas workers whose main tasks are substitutable[14], 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 countries, 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 is increasing over time. 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

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

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

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.

References

Acemoglu, D.; Restrepo, P. (2020) Robots and jobs: Evidence from US labor markets. *Journal of Political Economy*, 128(6), 2188-2244.

Acemoglu, D.; Restrepo, P. (2018) The race between man and machine: implications of technology for growth, factor shares, and employment. *American Economic Review*, *108*(6), 1488-1542.

Autor, D. H.; Katz, L. F.; Kearney, M. S. (2006) The polarization of the U.S. labor market. *American Economic Review*, *96*(2), 189-194.

Autor, D. H.; Levy, F.; Murnane, R. J. (2003) The skill content of recent technological change: an empirical exploration. *Quarterly Journal of Economics*, *118*(4), 1279-1333.

Bénabou, R.; Tirole, J. (2002) Self-confidence and personal motivation. *Quarterly Journal of Economics*, *117*(3), 871-915.

Brinca, P.; Duarte, J. B.; Faria-e-Castro, M. (2020) Measuring sectoral supply and demand shocks during COVID-19. Federal Reserve Bank of St. Louis, 2020-011.

Brinca, P.; Duarte, J. B.; Holter, H. A.; Oliveira, J. G. (2019) *Investment-Specific Technological Change, Taxation and Inequality in the U.S,* University Library of Munich, 91960.

Brinca, P.; Holter, H. A.; Krusell, P.; Malafry, L. (2016) Fiscal multipliers in the 21st century. *Journal of Monetary Economics*, *77*, 53-69.

Chakraborty, I.; Holter, H. A.; Stepanchuk, S. (2015) Marriage stability, taxation and aggregate labor supply in the U.S. vs. Europe. *Journal of Monetary Economics*, *72*, 1-20.

Coelho, J. (2020) Universal Basic Income and Skill-Biased Technological Change, Germany, University Library of Munich, MPRA Paper 99195.

Cortes, G. M.; Jaimovich, N.; Nekarda, C. J.; Siu, H. E (2014) *The Micro and Macro of Disappearing Routine Jobs: A Flows Approach*, NBER Working Paper 20307.

Dingel, J. I.; Neiman, B. (2020) *How Many Jobs Can Be Done at Home?,* NBER Working Paper 26948.

Eden, M.; Gaggl, P. (2020) Do poor countries really need more IT? *The World Bank Economic Review*, 34(1), 48-62.

Ferreira, A. M. (2019) *Skill-biased technological change and inequality in the US,* Germany, MPRA Paper, University Library of Munich, 93914.

Ferriere, A.; Navarro, G. (2018) *The Heterogeneous effects of government spending: It's all about taxes*, FRB International Finance Discussion Paper 1237.

Fonseca, T.; Lima, F.; Pereira, S. C (2018) Job polarization, technological change and routinization: evidence for Portugal. *Labour Economics, 51*, 317-339.

Goos, M.; Manning, A. (2007) Lousy and lovely jobs: the rising polarization of work in Britain. *Review of Economics and Statistics, 89*(1), 118-133.

Goos, M., Manning, A.; Salomons, A. (2009) Job polarization in Europe. *American Economic Review*, *99*(2), 58-63.

Guerreiro, J.;, Rebelo, S.; Teles, P. (2017) *Should robots be taxed?,* NBER Working Paper 23806.

Heathcote, J.; Perri, F.; Violante, G. 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, J.; Storesletten, K.; Violante, G. L. (2017) Optimal tax progressivity: An analytical framework. *Quarterly Journal of Economics*, *132*(4), 1693-1754.

Heckman, J.J. (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement,* 5(4), 475-92.

Heckman, J. J. (1977) *Sample Selection Bias as a Specification Error (with an Application to the Estimation of Labor Supply Functions*, NEBR Working Paper 0172.

Karabarbounis, L.; Neiman, B. (2014) The global decline of the labor share. *Quarterly Journal of Economics, 129*(1), 61-103.

Krusell, P.; Ohanian, L.; Ríos-Rull, J.; Violante, G. (2000) Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica, 68*(5), 1029-1053.

Nóbrega, V. (2020) Optimal taxation and investment-specific technological change. University Library of Munich, Germany, 98917.

Trabandt, M.; Uhlig, H. (2011) The Laffer curve revisited. *Journal of Monetary Economics*, *58*(4), 305-327.

APPENDIX A

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

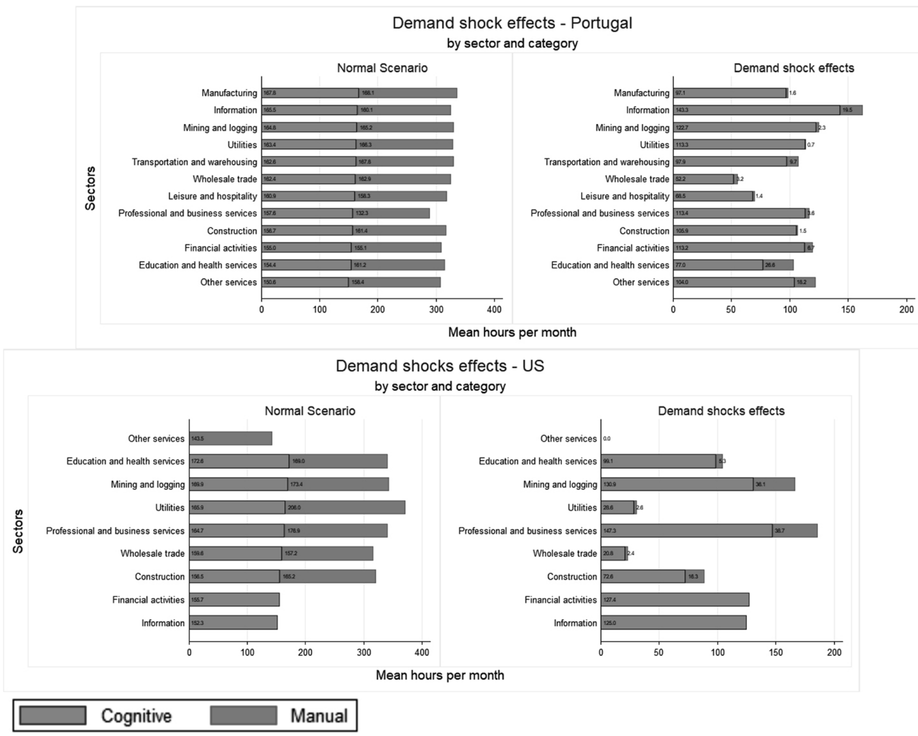


Figure 7: Task wage percentiles and minimum wage

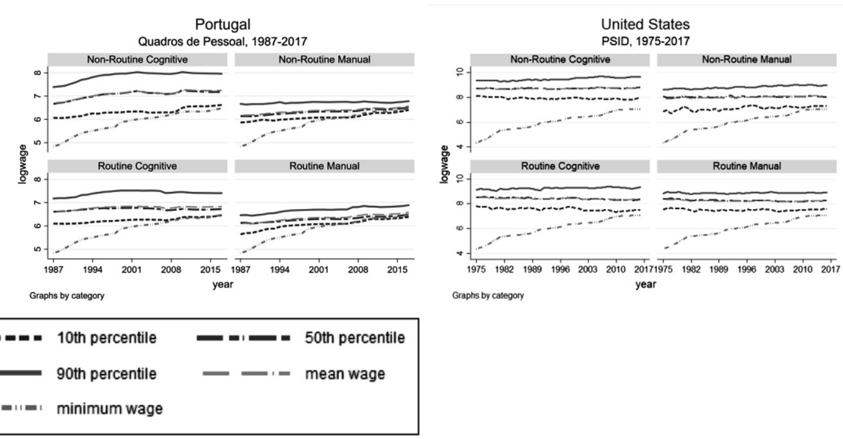


Table 6: U.S. – Heckman

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	0.267	0.267	0.258	0.242	0.246	0.238	0.220	0.215	0.202	0.186	0.165
y_1	0.237	0.213	0.201	0.181	0.155	0.141	0.130	0.112	0.0864	0.066	0.0562	0.048	0.038

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
y_1	0.188	0.162	0.153	0.140	0.133	0.125	0.117	0.105	0.087	0.767	0.075	0.077	0.076

Table 8: 2010 Benchmark calibration for US

Description	Parameter	Value	Source
Preferences			
Inverse Frisch elasticity	η	1.000	Brinca et al. (2016)
Risk aversion parameters	λ	1.000	Brinca et al. (2016)
Labor Productivity			
Depreciation rate equipment	δ_e	0.105	BEA
Depreciation rate structures	δ_s	0.033	BEA
Parameter 1 age profile of wages	y_1	0.236	Authors' Calculations
Parameter 2 age profile of wages	y_2	-0.0012	Authors' Calculations
Parameter 3 age profile of wages	y_3	1.58e-06	Authors' Calculations
Variance of idiosyncratic shock	σ_u	0.330	Authors' Calculations
Persistence of idiosyncratic risk	ρ_u	0.335	Authors' Calculations
Technology			
Share of income which goes to structures	α	0.151	Authors' Calculations
Share of the ICT cap/Cognitive composite	ϕ_1	0.469	Eden and Gaggl (2018)
Share of the ICT cap in the ICT Cognitive composite	ϕ_2	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	Normalizationi
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)
Capital income tax rate	τ_k	0.469	Trabandt and Uhlig (2011)
Tax scale parameter	θ_0	0.85	Implied value from
Tax progressivity parameter	θ_1	0.160	Ferriere and Navarro (2018)
Government debt to GDP	B/Y	0.880	(FRED) Average 2008-2012
Government spending to GDP	G/Y	0.213	FRED
SS tax employees	$\tilde{\tau}_{ss}$	0.077	OECD Tax Data
SS tax employers	τ_{ss}	0.078	OECD Tax Data

Table 9: QP 10 years RW

Year	σ^2_α	σ^2_ϵ	ρ	y_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	y_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	y_1	0.0638	Authors' Calculations
Parameter 2 age profile of wages	y_2	0.0020	Authors' Calculations
Parameter 3 age profile of wages	y_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 $q1$
Tax progressivity parameter	θ_1	0.136	OECD Tax Data
Government debt to GDP	B/Y	0.447	IMF Data[17]
Government spending to GDP		0.37	OECD

Note: 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:

$$\Phi(participation) = \Phi(Z'_{it}\epsilon + v_{it})$$

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:

$$E[ln(w_{it}) \,|\, X_{it}, workers = 1] = D'_t\xi + y_1 j + y_2 j^2 + y_3 j^3 + NRM_{it} + \rho\sigma_u \lambda(Z'_{it}\epsilon) + u_{i,t}$$

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,

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

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:

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.

Markets clear:

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

$$N^C = \int_{a > a*} n d\phi,$$

$$N^M = \int_{a \le a*} n d\phi,$$

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

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

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

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

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

The government budget balances:

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

The social security system balances:

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

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

$$\Gamma\int w(j)\,d\phi = \int(1-w(j))\,hd\phi.$$

APPENDIX B

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 1994 3d* between 2007 and 2006 and *CNP1994 3d* → *CNP1985 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[15].

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[16]. The method is describe in details in Appendix A. that has an almost unique one-to-one match with Census;[17] 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 and Neiman (2020) and occupations categories are defined by counting how many times ISCO-08 values fall in each category range, according to Cortes et al. (2014), in case of tie the occupation code is defined as 'Ambiguous'.[18]

---

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

[16] 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.

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

[18] These cases represent only a small portion of the workers in the data, on the file sample, this group is made of , representing the of the whole sample.

Matching Occupation across years

To recover previous years mapping in Portugal we then use the crosswalk CPP 2010 to CNP 1994.[19] 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.[20] 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 is multiple matching the same algorithm described before is used.

Teleworking and Susceptibility to Covid-19 by earnings percentiles

Table 12: 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 13: 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%

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

[20] If the match is lower than the occupation is defined as "Ambiguous".

Table 14: 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 15: 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 16: 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 17: 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

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

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:

$$1 - \frac{1 - \tau(y_2)}{1 - \tau(y_1)} = 1 - \left(\frac{y_2}{y_1}\right)^{-\theta 1}$$

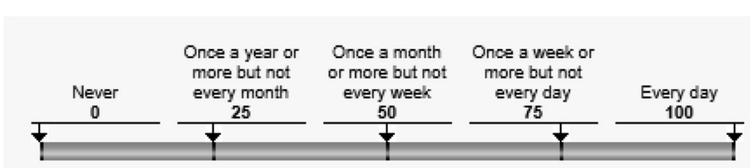
and therefore independent of the scaling parameter θ_0. In this manner, one can raise average taxes by lowering θ_0 and not the progressivity of the tax code, since the progressivity is uniquely determined by the parameter θ_1.

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:

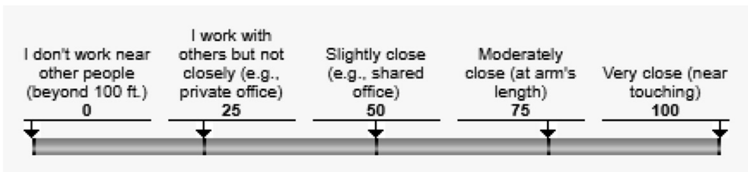
Figure 8: Source: O-NET online



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:

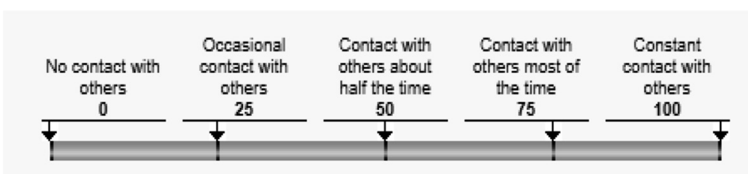
Figure 9: Source: O-NET online



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 10: 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[21] so that when there are multiple matching we can "smooth" the division.

Following Dingel and Neiman (2020) and using U.S. employment data[22] 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

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

[22] Occupational Employment Statistics.

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 18: 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 19: 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 20: 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

Table 21: 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

Table 22: Sector coding

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

Figure 11: Portugal

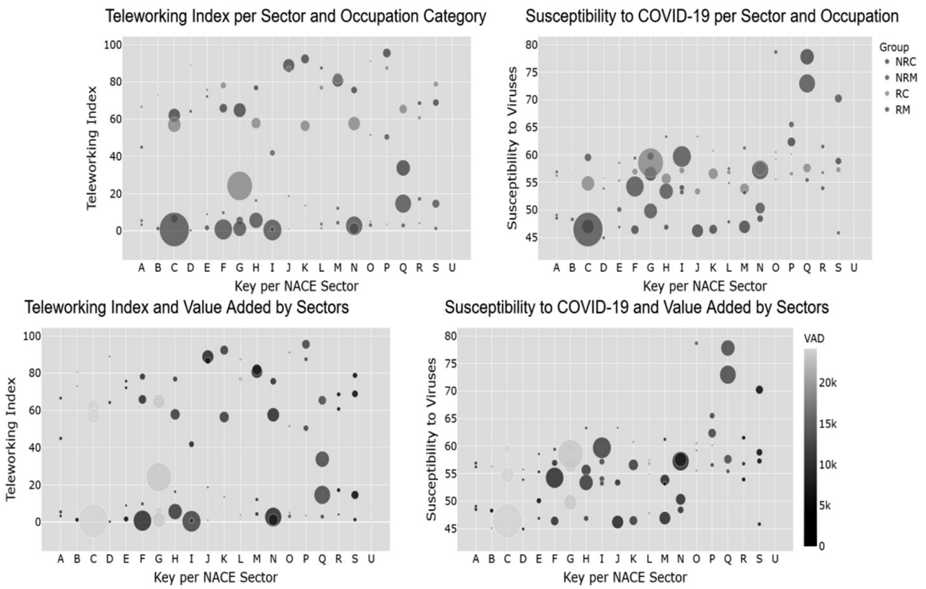
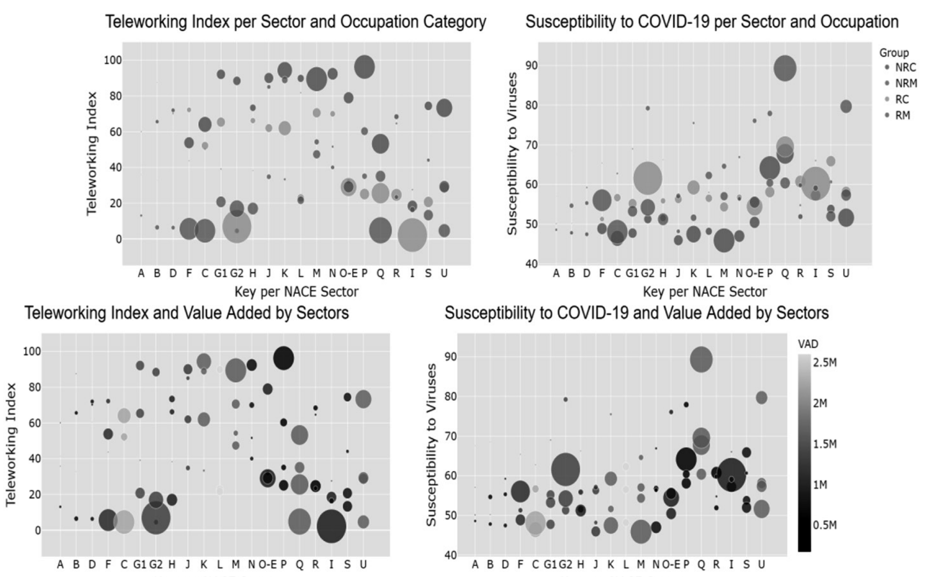


Figure 12: United States



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 retrieved 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 *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 split 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.