DISCUSSION PAPER SERIES

DP15118
(v. 3)

Real-Time Inequality and the Welfare State in Motion: Evidence from COVID-19 in Spain

Oriol Aspachs, Ruben Durante, Alberto Graziano, Josep Mestres, Jose G Montalvo and Marta Reynal-Querol

PUBLIC ECONOMICS

ISSN 0265-8003

Real-Time Inequality and the Welfare State in Motion: Evidence from COVID-19 in Spain

Oriol Aspachs, Ruben Durante, Alberto Graziano, Josep Mestres, Jose G Montalvo and Marta Reynal-Querol

Discussion Paper DP15118
First Published 31 July 2020
This Revision 31 January 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Real-Time Inequality and the Welfare State in Motion: Evidence from COVID-19 in Spain

Abstract

Official statistics on economic inequality are only available at low frequency and with considerable delay. This makes it challenging to assess the impact on inequality of fast-unfolding crises like the COVID-19 pandemic, and to rapidly evaluate and tailor policy responses. We propose a new methodology to track income inequality at high frequency using anonymized data from bank records for over three million account holders in Spain. Using this approach, we analyze how inequality evolved between February and November 2020 (compared to the same months of 2019). We first show that the wage distribution in our data matches very closely that from official labor surveys. We then document that, in the absence of government intervention, inequality would have increased dramatically, mainly due to job losses and wage cuts experienced by low-wage workers. The increase in pre-transfer inequality was especially pronounced among the young and the foreign-born, and in regions more dependent on services. Public transfers and unemployment insurance schemes were effective at providing a safety net to the most affected segments of the population and at offsetting most of the increase in inequality. Increased inequality is primarily driven by differential changes in employment rate. Indeed, using individual-level regressions, we find that, over the course of the pandemic, the probability of being employed decreased drastically for workers in the lower part of the pre-COVID wage distribution, young cohorts, and foreign-born.

JEL Classification: C81, D63, E24, J31

Keywords: Inequality, COVID-19, Administrative data, High Frequency Data

Oriol Aspachs - oaspachs@caixabank.com
Caixabank Research

Ruben Durante - ruben.durante@upf.edu
ICREA, Universitat Pompeu Fabra, IPEG, Barcelona GSE, CEPR and CEPR

Alberto Graziano - agraziano@caixabank.com
Caixabank Research

Josep Mestres - jmestres@caixabank.com
Caixabank Research

Jose G Montalvo - jose.garcia-montalvo@upf.edu
Universitat Pompeu Fabra

Marta Reynal-Querol - marta.reynal@upf.edu
Universitat Pompeu Fabra and CEPR

Acknowledgements
We thank seminar participants at the IMF, the Bank of Italy, the Bank of France, and participants in the 2021 European Economic Association Meeting for helpful comments. We also thank the editors, an anonymous reviewer, and the discussants of the paper in the 73rd Economic Policy Panel Meeting - Roland Rathelot and Thiemo Fetzer - for helpful feedback. We thank La Caixa Foundation (Project SR20-00608, Social Research Call 2020) and the Departament de Recerca i Universitats of the Generalitat de Catalunya (Project 00090, Pandemies 2020) for financial support. Daniele Alimonti and Marco Lo Faso provided excellent research assistance. Corresponding author: Ruben Durante (ruben.durante@upf.edu).

Real-Time Inequality and the Welfare State in Motion: Evidence from COVID-19 in Spain*

Oriol Aspachs† Ruben Durante‡ Alberto Graziano†
Josep Mestres† Jose G. Montalvo§ Marta Reynal-Querol‡

September 2021

Abstract

Official statistics on economic inequality are only available at low frequency and with considerable delay. This makes it challenging to assess the impact on inequality of fast-unfolding crises like the COVID-19 pandemic, and to rapidly evaluate and tailor policy responses. We propose a new methodology to track income inequality at high frequency using anonymized data from bank records for over three million account holders in Spain. Using this approach, we analyze how inequality evolved between February and November 2020 (compared to the same months of 2019). We first show that the wage distribution in our data matches very closely that from official labor surveys. We then document that, in the absence of government intervention, inequality would have increased dramatically, mainly due to job losses and wage cuts experienced by low-wage workers. The increase in pre-transfer inequality was especially pronounced among the young and the foreign-born, and in regions more dependent on services. Public transfers and unemployment insurance schemes were effective at providing a safety net to the most affected segments of the population and at offsetting most of the increase in inequality. Increased inequality is primarily driven by differential changes in employment rate. Indeed, using individual-level regressions, we find that, over the course of the pandemic, the probability of being employed decreased drastically for workers in the lower part of the pre-COVID wage distribution, young cohorts, and foreign-born.

Keywords: inequality, COVID-19, administrative data, high frequency

JEL codes: D63, E24, J31

*We thank seminar participants at the IMF, the Bank of Italy, the Bank of France, and participants in the 2021 European Economic Association Meeting for helpful comments. We also thank the editors, an anonymous reviewer, and the discussants of the paper in the 73rd Economic Policy Panel Meeting - Roland Rathelot and Thiemo Fetzer - for helpful feedback. We thank La Caixa Foundation (Project SR20-00608, Social Research Call 2020) and the Departament de Recerca i Universitats of the Generalitat de Catalunya (Project 00090, Pandemies 2020) for financial support. Daniele Alimonti and Marco Lo Faso provided excellent research assistance. Corresponding author: Ruben Durante (ruben.durante@upf.edu)

†Caixabank Research

‡ICREA-UPF, BSE, IPEG, CESifo, and CEPR

§UPF, IPEG, and BSE

1. Introduction

The spread of COVID-19 has taken a heavy toll on the economy of most countries. GDP in the Euro area shrunk by 6.7% in 2020 according to Eurostat. The Spanish economy, which our study focuses on, was among the most affected by the pandemic and GDP declined even further, -10.8%. The pandemic and the lockdown measures implemented to stop it have also had a considerable impact on the labor market. In Spain, between mid-March and the end of April 2020, almost one million workers lost their job and another 4.6 million forcefully transitioned to temporary lay-off schemes.

A crucial concern is that the economic impact of the pandemic may disproportionately hit the most vulnerable segments of the population, leading to a surge in economic inequality (Wade, 2020; Scheidel, 2017). Increasing inequality, in turn, can pose serious risks for political stability, since it erodes social cohesion and can spur support for populist or even undemocratic views (Inglehart and Norris, 2016).

Governments around the world have tried to mitigate the economic consequences of the pandemic investing vast resources in a combination of family income support and credit facilities for firms and self-employed workers (IMF, 2020; ILO, 2020). In Spain in 2020 additional spending measures accounted for 4.1% of GDP (most for workers' unemployment benefits), and liquidity support for 14.4% of GDP. Yet, how appropriate and effective these policies are remains unclear, mainly due to a lack of reliable indicators allowing to track economic activity at a fine temporal and spatial resolution. Indeed, most official macroeconomic statistics are available only at quarterly or yearly frequency and often with long delays, limiting the ability of policymakers to rapidly adjust their responses (Gourinchas, 2020).[1]

This paper aims to fill this gap by proposing a novel methodology that allows to track the evolution of income inequality with a high degree of temporal granularity. Our approach relies on the use and analysis of comprehensive anonymized data from bank records, including information on both the wages and government transfers paid to account holders. Compared to other recent studies that use high-frequency data from credit/debit card transactions to look at the evolution of spending, our paper is novel in that it focuses on wages and salaries to investigate the distributional impact of the pandemic. In addition, we propose a novel approach to evaluate the effectiveness of government programs to attenuate the effect of the pandemic using data on government transfers perceived by account holders.[2]

Our analysis focuses on Spain and uses data from CaixaBank – Spain's second-largest bank by total assets and first by direct payroll deposits - which cover over three million retail

[1] For example, the most recent official measures of inequality available for Spain refer to 2018. A similar lag, when not longer, applies to other countries in and outside the EU.

[2] See Aspachs et al. (2021) for an illustration of this approach.

depositors. Our sample includes all active account holders receiving payroll payments from a private or public employer and/or any government transfers as of February 2020.[3] Given the high level of financial inclusion - 97.6% of Spaniards aged 15 or more hold a bank account - bank records provide a very accurate picture of the Spanish working population.[4] Using this approach, we examine how wage earnings evolved in the months before and after the outbreak of the pandemic for individuals in our sample. In particular, we are able to identify how the pandemic affected the earnings of people in each (pre-pandemic) wage level, and who experienced total or partial decline due to job losses or wage cuts. To control for fluctuations in wages due to seasonality and unrelated to the pandemic, we always use the same months of 2019 as benchmark. Crucially, the availability of comprehensive information about the public transfers received by each account holder provides us with a unique opportunity to evaluate how government intervention contributed to alleviate the effect of the pandemic on earnings and, ultimately, on wage inequality.

We first confirm that our sample is highly representative of the Spanish working population. To this end, we compare the distribution of wages in our sample with the distribution of wages from the latest wave of the Wage Structure Survey (ESS) conducted by the Spanish National Institute of Statistics (INE) in 2018, and find the two are remarkably similar.

We then analyze how income inequality evolved over the course of the crisis and after the end of the lockdown. We document that while the Gini index in February 2020 was virtually the same as in February 2019, income inequality increased sharply in March, and even more so in April and May when the Gini index was about 11 points higher than in February, a 25% increase in just two months. Inequality declined somewhat following the reopening of the economy in the months of June and July. Government intervention was quite effective at containing the spike in inequality. Indeed, while the post-transfer Gini index is usually lower than pre-transfer Gini index by about 5 points, this difference reached 13 points in both April and May 2020, offsetting most, though not all, the increase in pre-transfer inequality. This was less the case in March, when the post-transfer Gini index reached its peak, arguably due to the delay in the disbursement of subsidy and unemployment benefit pay-outs in the early stage of the crisis.

We also look at how the effect of the pandemic on wage income inequality varied across different groups and areas of the country. We find that, while the increase in inequality was similar among men and women, it was more pronounced among young people, and

[3] We exclude self-employed workers for whom, given the irregularity of earnings, computing monthly income is very challenging.

[4] Though our data only cover wage payments, given the overwhelming importance of wages as the main source of income in Spain, our approach allows to shed some light on the effect of the pandemic on overall income inequality. Indeed, as discussed below, the vast majority of the Spanish working population is composed by wage earners, and measures of inequalities for wage income and total income are virtually identical.

among foreign-born individuals, particularly those from lower-income countries. Regarding geographic differences, we document that pre-transfer inequality increased particularly in areas where mobility declined the most relative to the pre-pandemic period, and that rely more heavily on retail, hospitality and other service activities that were most affected by the crisis. We find that government transfers were effective at alleviating at least some of these differences.

We also document that the evolution of inequality in the different phases of the pandemic was mainly driven by workers transitioning from employment to unemployment (and vice versa) rather than by wage changes for workers that remained employed. Using individual-level regressions, we show that the probability of being employed decreased disproportionately for individuals in the lower part of the pre-pandemic wage distribution, for younger cohorts, and for foreign-born.

Finally, looking at two regional elections held in the summer of 2020, we explore how the ability of the government to contain the increase in inequality triggered by the pandemic influenced voting intentions. We find that areas where inequality decreased the most due to government intervention display higher turnout and higher support for left-wing independentist parties, but lower support for the incumbent.

Taken together, our findings indicate that the COVID-19 pandemic, and the confinement and social distance measures adopted to face it, led to a massive increase in earnings inequality in Spain, primarily driven by job losses and wage cuts for low-income workers. They also indicate that, despite an initial delay, government transfers were effective at containing such increase, reducing inequality to levels not too distant from the pre-pandemic ones. Though reassuring of the ability of the welfare state to cope with such extreme situations, this finding generates some concerns for how things may evolve should the intensity of government intervention decline due to budgetary reasons before the health emergency has ceased. From a methodological perspective, our analysis provides the first example of how banking data can be used to track income inequality at a high frequency, and to evaluate and guide government policies as they unfold.

Our paper relates to a number of new initiatives, triggered by the COVID-19 crisis, which use big data to track economic activity in real-time. Perhaps the most impressive effort in this direction is that of Chetty et al. (2020), who build a tracker of economic activity at a high-frequency granular level using data from credit card processors, payroll firms, job posting aggregators, and financial services firms. Other prominent examples include Cicala (2020), who uses data on electricity usage to follow changes in industrial activity in Europe, and by Bick and Blandin (2020), who use the Real Time Population Survey (RPS) to construct high frequency estimates of employment, hours worked and earnings. In this regard, our project is the first to propose the use of big administrative data from private companies to construct a high-frequency indicator of wage income inequality, that overcomes the limitations of existing low frequency macroeconomic statistics and of high-frequency measures based on surveys.

Our paper also relates to previous contributions using Spanish bank data on credit cards usage at POS (point of sale), online transactions, and cash withdrawals to study the effect of the lockdown measures on different categories of spending during the COVID-19 pandemic (Carvalho et al., 2020), and the evolution of consumption by level of income during the epidemics (Montalvo and Reynal-Querol, 2020).[5] In this regard, our paper is novel in that it shows that other bank data - e.g., on payroll and government transfers - can be harnessed to study income inequality and to evaluate and guide government policies.

The remainder of the paper is organized as follows. Section 2 provides background information on the government response to the economic consequences of the COVID-19 pandemic in Spain. Section 3 presents the data, explains the construction of the sample, and discusses its representativeness. Section 4 illustrates the evolution of wage income inequality in the entire sample and differences across demographic groups and areas. Section 5 documents how most of the variation in wage income inequality can be attributed to changes in employment status, and what individual and local factors mediate this relationship. Section 6 discusses the political ramifications of the change in inequality and of government intervention to limit it. Section 7 concludes.

2. The policy response to the COVID-19 crisis

Two types of benefits have been primarily used to support workers' income during the COVID-19 pandemic: unemployment benefits, and furlough schemes known in Spain as ERTE (acronym for *Expedientes de Regulación Temporal de Empleo*).

Unemployment benefits in Spain normally require to have worked for at least 360 days in the previous six years. Normally, benefits are proportional to the time previously worked, and can be received for a maximum of 18 months.[6] During the COVID-19 emergency, the Spanish government made access to these benefits easier and extended its coverage, creating special unemployment benefits for those who would otherwise be no longer eligible.

Furlough schemes are temporary layoff mechanisms that maintain the employer-employee relationship, as workers remain affiliated to the Social Security, and that allow employees to

[5] The use of bank information for the analysis of spending in other countries during the COVID-19 pandemic includes, among others, Hacioglu et al. (2020) and Crawford et al. (2020) for the UK, Baker et al. (2020) for the US and Sheridan et al. (2020) for Denmark and Sweden.

[6] These conditions apply to all workers independently of their country of birth, so no distinction exist between native and foreign-born individuals.

receive unemployment benefits while they are not working (or experience a reduction in the number of hours/days worked). These benefits are proportional to the time employees are not working, and are compatible with part-time employment.

Unlike other European countries like Germany with its *Kurzarbeit* programme, furlough schemes were not widely used in Spain prior to this crisis. To encourage its use, during the pandemic the Spanish government subsidized part of the employer Social Security contributions for those workers on furlough schemes. In addition, the benefits received did not reduce future unemployment benefit entitlements.

These schemes ensured workers received an income stream while unemployed or while their contract was suspended, though of a lower amount than their regular salary. Indeed, in both schemes, benefits amount to 70% of the Social Security tax base. This tax base is often smaller than the wage received since it is capped. This implies that the actual benefits can be less than the 70% of the wage. In addition, in the case of furlough schemes, the amount of benefits received is proportional to the time the individual has not worked and post-benefits income can include both wages and benefits. Nevertheless, as depicted in Figure 1, the change in post-benefits wage income for benefits recipients display an abnormal concentration at -30%.

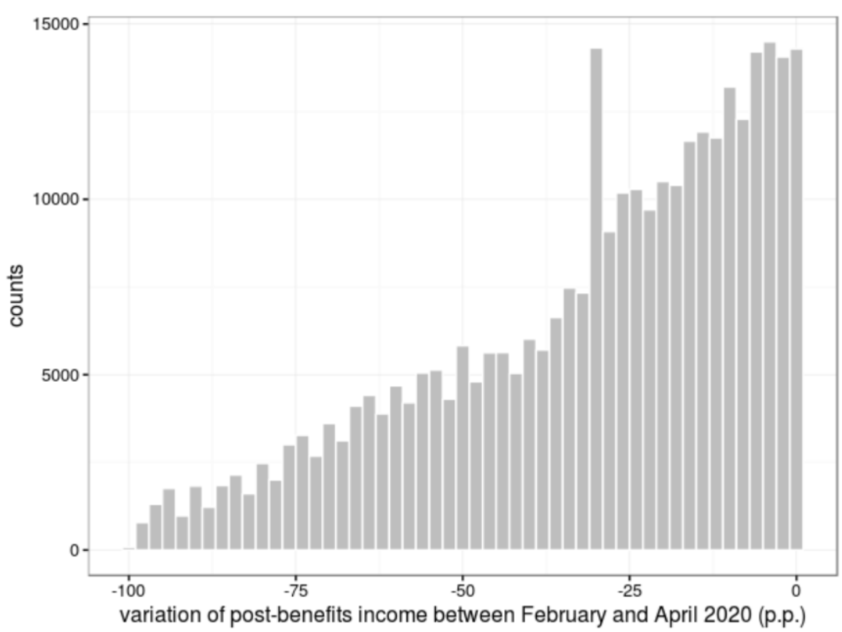


FIGURE 1: VARIATION OF POST-BENEFITS INCOME FOR THOSE RECEIVING BENEFITS

As a result, these public benefit schemes partially compensated wage losses and had a significant take-up. However, not all workers affected by the economic consequences of the pandemic were entitled to these benefits. In particular, workers that were already unemployed before the pandemic or workers with temporary contracts that expired might have been eligible for reduced unemployment benefits or none at all. This issue is of particular relevance in a country like Spain where about a quarter of all workers - particularly young and foreign-born - have temporary contracts. An additional problem was associated with the considerable delay in the processing of the public benefits which caused a large number of beneficiaries to receive their benefits only after several months. In light of all these issues, a rigorous analysis of the effectiveness of government support programs in alleviating income inequities is highly relevant.

3. Data

In this section we describe the procedure employed to construct the sample used in our analysis and discuss how representative it is of the Spanish working population.

3.1. Sample construction

Our base sample comes from anonymized personal bank accounts of CaixaBank and includes all payroll income and labor market related benefits received in that account each month. CaixaBank is the second largest Spanish bank and has the highest market share of direct payroll deposits (27.1%). The sample considers individuals aged 16 to 64 years old. Every month we have around 3 million individuals. Payroll income is precisely identified as a specific type of bank transfer, distinguished from any other type of bank account movements. All payrolls paid in a specific bank account are included, independently of the amount or the frequency in which they are paid.[7] Salaries in month t are then defined as the sum of all wage payments received in a particular bank account from the 16th of that month until the 15th of the following month. Labor market related benefits are uniquely identified as a particular type of transfer paid by the Social Security for those workers in unemployment and furlough schemes (ERTE). Public benefits in month t refer to all benefits perceived due to unemployment or furlough in month t (covering the same time period as salaries), and they are usually paid around 10-12 days after the end of month t. In addition, we know for each client her gender, age, province of residence and country of birth.

We focus on accounts with one account holder, or with multiple co-holders but only one employer paying-in wages. This way, we ensure that the payrolls or transfers recorded correspond to only one individual and avoid recording multiple payrolls or transfers from multiple

[7] In Spain, employees' payments are usually deposited monthly, towards the end of the month, but they could less frequently be paid fortnightly, weekly, etc. and they are very rarely paid in cheques.

account holders.[8] In addition, we exclude the account of those individuals who died during our period of analysis, and of those that did not use the account for their usual financial transactions over the same time.[9] Finally, to ensure the stability of the sample, we only include individuals who received wage or benefit payments in the two months prior to the period of analysis, which starts in February 2020. We follow those individuals in the following months starting in February 2020 (lockdown measures started on March 14th 2020 when the state of alarm was declared in Spain) until November 2020 (lockdown measures started to be lifted in May) and we observe all wage income and unemployment benefits received or the absence of any of those. We do not filter the sample by any other variable including by income level, neither minimum nor maximum. Obviously, we have to restrict the sample to wage earners when comparing our data with other datasets to assess the representativeness of our sample (see next section).

3.2. Sample representativeness

Since the data we employ are based on bank records, the representativeness of our sample crucially depends on what share of the Spanish active population holds a bank account. According to data from the Global Findex, the index of financial inclusion produced by the World Bank, 97.6% of Spanish people aged 15 or more hold a bank account, compared to an average of 93.7% in high-income countries. There are no significant gender differences in financial inclusion, as the share of individuals with a bank account is exactly the same for men and women. As mentioned above, we restrict our sample to people that either work or receive some type of government transfer related with their job market participation.

We exclude the self-employed since they are more likely to use multiple bank accounts for their personal and professional needs, which makes it much more difficult to track their income from information about a single account. That said, it is important to note that, according to the latest Labor Force Survey (Encuesta Población Actva, EPA) by the Spanish Statistical Office, as of the first quarter of 2020 the share of wage earners in the Spanish working population was 84.4%.[10]

Since most of the individuals in our sample are workers, to assess how representative it is

[8] Only 5% of individual bank accounts have multiple holders receiving two or more payrolls and were hence excluded from the sample.

[9] In particular, we identify each month those clients who are actively using their bank account and perform at least two non-automatic transactions (e.g., payments, etc.) during the last two months. Those that do not satisfy this criterion represent about 0.7% of the sample.

[10] The relevance of wages as the main source of income is also attested by the similarity of the inequality measures using income and gross wages respectively. Indeed, for the last period for which both measures are available, the Gini coefficient was 34.5 for income inequality and 34.5 for wage inequality. Considering the entire period 2008-2016, the average absolute difference between the Gini index for income and wages was a mere 0.3.

of the wage distribution in Spain, we compare it with data from the Spanish Wage Survey (Encuesta de Estructura Salarial, EES) conducted every four years by the Spanish Statistical Office. To this end, we consider all individuals in our sample who were working in February 2020.

One concern could be that some people may have multiple jobs and therefore could receive salaries in different bank accounts. However only 2.1% of the employees have more than one salaried job, one of the lowest proportion among the European countries (Addeco, 2015). Additionally, the official wage survey (EES) only considers the wage of each employee in one company even if they have a second salaried job. Therefore, the EES uses the same criterion that we use to select our sample.

The salaries deposited into workers' bank accounts are net of tax withholdings and social security contributions. Hence, to facilitate the comparison with our sample, we compute the distribution of net salaries by estimating and subtracting from the gross salaries reported in the EES data both tax withholdings and social security contributions.[11]

Therefore, in order to compare our data with the official EES, we have calculated the distribution of net salaries transforming the gross salaries (GW) of the EES into net salaries (NW) as

$$GW_i = BW_i + OC_i + B_i \tag{1}$$

$$NW_i = GW_i - SS_i - Taxes_i \tag{2}$$

where BW is the base salary, OC is the overtime compensation; B is the bonus or any other extraordinary payments, SS are Social Security contributions and taxes are withheld taxes.

Figure 2 shows the distribution of monthly wages in our sample for February of 2020, compared with the distribution of monthly net salaries from the EES. The latest available microdata from the ESS refer to the year 2018. To account for this time gap, we adjust the entire distribution by the increase in the average salary in 2019 (2.4%).[12] As depicted in Figure 2, the histogram of net wages in our sample matches remarkably well the distribution of net salaries in the ESS.

[11] Gross salaries also include extraordinary payments and payments for extra hours.

[12] In our previous work, Aspachs et al. (2021) we could only compare with the wage survey of 2014 which was the latest available at that time. The assumption of proportional movement of the whole distribution with the average wage inflation was more questionable than in the current comparison where we only consider one year after the last available official wage survey.

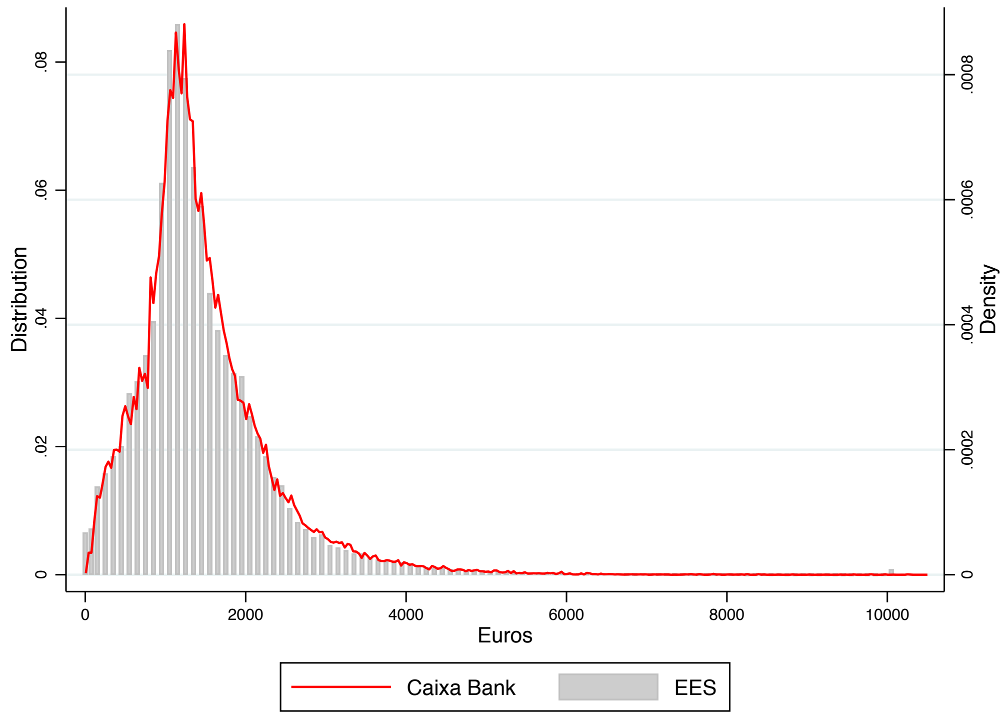


FIGURE 2: DISTRIBUTION OF MONTHLY NET SALARIES:
CAIXABANK SAMPLE VS. ESS SAMPLE (2018 UPDATED)

TABLE 1: QUANTILE RATIOS OF THE DISTRIBUTION OF NET WAGES:
CAIXABANK SAMPLE VS. ESS SAMPLE

	Our sample (CBK)	EES 2018
P90/P10	4.24	4.14
P90/P50	1.88	1.86
P10/P50	0.44	0.46
P75/P25	1.85	1.82

Note: The table reports the following ratios of percentiles for the distribution of net wages separately for the Caixabank sample (left column) and the ESS sample (right): $90^{th}/10^{th}$, $90^{th}/50^{th}$, $10^{th}/50^{th}$, and $75^{th}/25^{th}$. The ESS (*Encuesta de Estructura Salarial*) is the Spanish Wage Survey conducted every four years by the Spanish Statistical Office. We consider the latest one from 2018 with salaries updated using the growth of wages since 2018.

TABLE 2: GENDER AND AGE DISTRIBUTION:
CAIXABANK SAMPLE VS. OTHER SOURCES

	Our sample (CBK)	EES 2018	EPA4T19	EPA1T20
N	3,028,204	216,726	≈200,000	≈200,000
Gender				
Male	0.54	0.53	0.52	0.52
Female	0.46	0.47	0.48	0.48
Age				
$\leq$ 19	0.01	0.00	0.008	0.007
20-29	0.18	0.14	0.145	0.142
30-39	0.25	0.25	0.246	0.243
40-49	0.28	0.30	0.305	0.304
50-59	0.21	0.23	0.233	0.237
$\geq$ 60	0.07	0.07	0.060	0.063

Note: The table reports the distribution of individuals by gender and age separately for four different samples: Caixabank (CBK), EES 2018, EPA last quarter of 2019, and EPA first quarter of 2020.

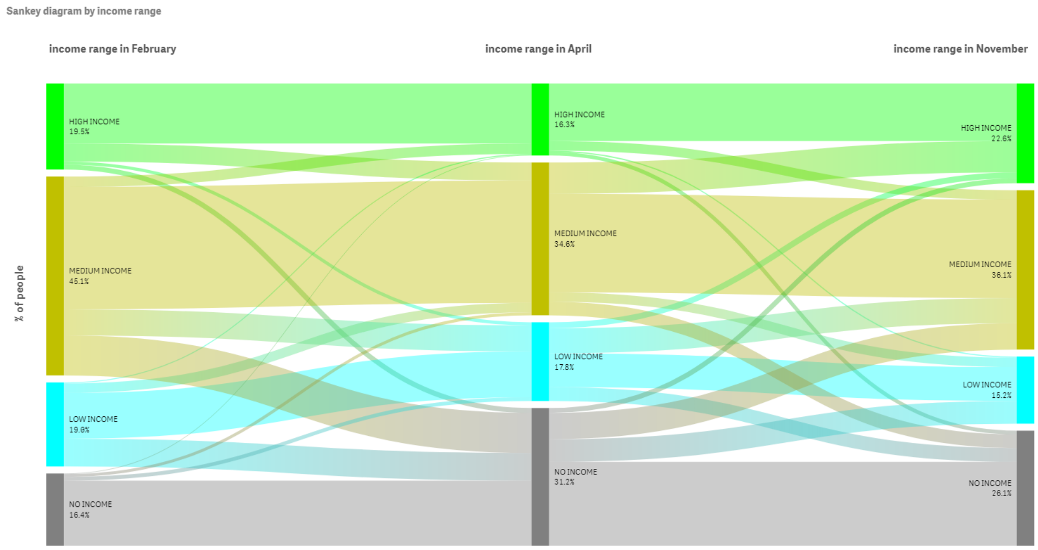
That the bank data we employ provide an accurate picture of the general wage distribution in the Spanish working population is further confirmed by comparing several quantile ratios that are generally used to measure inequality, which again, as depicted in Table 1, are very similar between the two samples.

To further document the representativeness of our bank data, in Table 2 we compare the distribution of individuals by gender and age between the Caixabank sample and some official sources. In addition to the 2018 ESS used above,[13] to compare with more recent estimates we also use data from the two latest wages of the Spanish Labor Force Survey (Encuesta de Población Activa, EPA) which refer to the last quarter of 2019 and the first quarter of 2020. Regarding the gender distribution, the share of males in our sample is 54% which is very similar to the 52-53% recorded in the other sources. A similar picture emerges from the distribution by age groups, which is very much consistent with that in the official surveys.

4. COVID-19 AND INEQUALITY

The pandemic has produced significant changes in the distribution of wages. Since one of the goals of our analysis is to examine to what extent government intervention alleviated the economic consequences of the pandemic, in all the exercises discussed below we present the results before and after taking into account public transfers (pre-benefits and post-benefits scenarios). Specifically, in the pre-benefits scenario, we consider net wages but exclude

[13] This information was not available in Aspachs et al. (2021).



Note: "low income" groups refers to wages below 1,000 euros/month; "middle income", between 1,000 and 2,000 euros/month, and "high income", more than 2,000 euros/month.

FIGURE 3: DISTRIBUTION OF WAGES IN FEBRUARY, APRIL AND NOVEMBER OF 2020 PRE-BENEFITS.

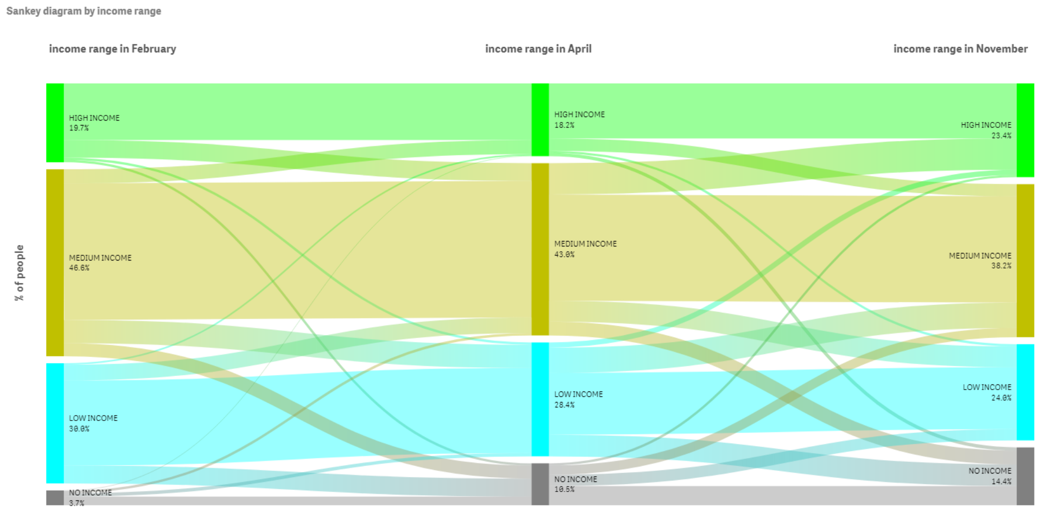
unemployment benefits and furlough schemes, while in the post-benefit scenario we consider all these items together.

Figure 3 provides a first look to the movement of individuals across different levels of income in three moments of time: February (before the beginning of the pandemic); April (end of the first wave); and November (end of the second wave). This figure shows that, before accounting for public benefits, a large proportion of low and middle income workers moving to the no income situation. Only part of this movement was reversed in November.

Figure 4 represents the transitions between different wage groups after taking into account public benefits (unemployment benefits and furlough schemes). In this case, logically, the movement to the no income group is much smaller.

Our analysis also aims to understand how the COVID-19 crisis affected income inequality in Spain and to what extent government intervention was able to mitigate this effect. Some recent research has emphasized the relevance of top income shares in countries like the United States. However, the importance of the evolution of top income shares is less relevant in continental Europe. In addition, our work deals with wage inequality. For these reasons we are interested in describing the movements of the entire distribution and, particularly, of the bottom wages. The Gini index is the most common and intuitive measure of inequality for these purposes.

$$G = \frac{1}{2N^2\bar{y}} \sum_{j=1}^{m} \sum_{i=1}^{m} n_j n_i |y_j - y_i| \tag{3}$$



Note: "low income" groups refers to wages or transfers below 1,000 euros/month; "middle income", between 1,000 and 2,000 euros/month, and "high income", more than 2,000 euros/month.

FIGURE 4: DISTRIBUTION OF WAGES IN FEBRUARY, APRIL AND NOVEMBER OF 2020 POST-BENEFITS.

where N is the number of people, m is the number of income classes, n_j is the number of individuals in class j, $\bar{y}$ is the mean of income, and y_i is the average income in class i.
In this section we consider the evolution of the Gini coefficient during the pandemic, the evolution of between and within groups inequality, and the changes in within group inequality.

4.1. THE EVOLUTION OF INEQUALITY DURING THE PANDEMIC

Figure 5 shows the evolution of the Gini index corresponding to the distribution of wages in our sample between February and November 2020 (right panel) and over the same period of 2019 (left). The brown line refers to the pre-transfer distribution while the blue line to the post-transfer distribution. We have included the confidence intervals around the points. We calculate the confidence interval for the Gini index to show the statistical significance of the indices reported in the text. There are two basic procedures to perform this calculation: using a Jackknife or a Weighted Least Squares (WLS) estimator. It is well known that both procedure produce the same estimators for large sample size which is our case.[14] The WLS estimator is calculated by estimating the following equation:

$$i = \theta + u_i \tag{4}$$

[14] See, for instance, Giles (2004).

where u_i is a heteroskedatic error with variance equal to σ^2/y_i. This implies that the previous regression can be transformed into a regression with an homoskedastic error,

$$\sqrt{y_i}i = \theta\sqrt{y_i} + \varepsilon_i \tag{5}$$

Therefore, the standard error of the Gini index is

$$std(Gini) = \frac{2std(\hat{\theta})}{N} \tag{6}$$

where N is the number of observations. The standard error of the Gini indices calculated in this paper are so tiny that it is difficult to discern them from the value of the Gini index, given the enormous size of our sample.

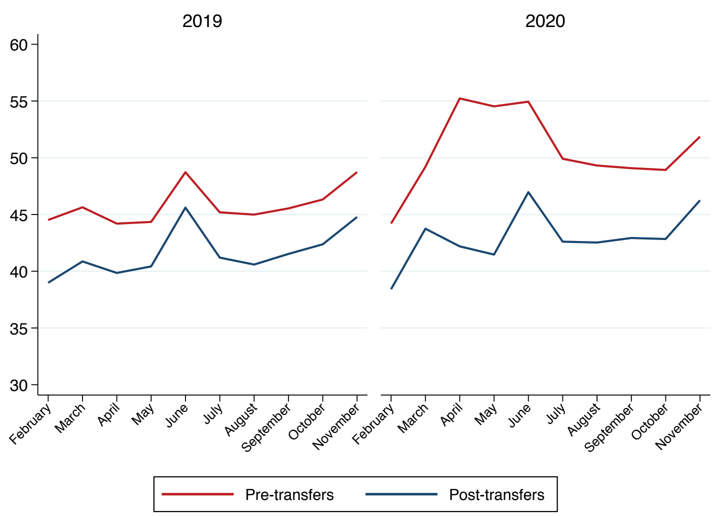


FIGURE 5: EVOLUTION OF PRE- AND POST-TRANSFER GINI INDEX BETWEEN FEBRUARY AND NOVEMBER 2019 AND FEBRUARY AND NOVEMBER 2020

Comparing February 2020 with February 2019 it is evident that the Gini index for both the pre-transfer and the post-transfer distribution is exactly the same between the two years (a difference of 0.3 and 0.5 points respectively on a scale of 100 points). The situation changes abruptly starting from March 2020, when pre-transfer inequality starts to rise considerably beyond 2019 levels. The Gini index displays a value of 49.2 in March 2020, 55.2 in April, 54.5 in May, 54.9 in June, 49.9 in July, 49.3 in August, 40.1 in September, 48.9 in October, and 51.8 in November. These values are considerably higher than those recorded in the same months of the previous year, with the difference reaching 11 and 10 points respectively in April and May, at the pick of pandemic. This represents an unprecedented increase of about

25% in just two months, roughly corresponding to the difference between Germany and the US as of 2016 (The World Bank, 2020).

Focusing on post-transfer inequality, it is clear how government intervention through unemployment insurance benefits and other transfers was instrumental at containing the abrupt spike in wage inequality. Indeed, while the post-transfer Gini index is usually lower than the pre-transfer Gini index by about 5 points, this difference reached 13 points in both April and May 2020, offsetting most, though not all, of the increase in pre-transfer inequality. Government action was somewhat less effective in March, when the post-transfer Gini index reached its pick, about 3 points higher than in the same month of 2019. This was arguably due to the delay in the disbursement of subsidy and unemployment benefit pay-outs in the early stage of the crisis, which temporarily left some of the most vulnerable workers without a safety net. Finally, both pre-transfer and post-transfer inequality remained relatively high or further increased in June, before decreasing sharply in the following months as the economy reopened.

To better visualize the impact of the pandemic on inequality net of seasonality, in Figure 6 we plot the difference in the Gini index between each month from March to November 2020 and February 2020 net of the difference between the same months of 2019 (difference-in-difference), separately for the pre-transfer (blue) and the post-transfer distribution (brown). The graph confirms that while the difference in pre-transfer inequality reached its pick in April and May, post-transfer inequality was the highest in March relative to the previous year. This difference-in-difference calculation relies on a parallel trends assumption for the months of 2019. We show in Appendix figure A.1 that the difference in the Gini index between each month in 2019 net of the difference between the same months of 2018 was almost zero in all months, that is, that pre-trend income inequality was basically flat. To give a better sense of what parts of the wage distribution are driving the change in inequality during the initial months of the pandemic, when most of the action was concentrated, we can analyze the Lorenz curves. Appendix figure A.2 shows the Lorenz curve corresponding to each month between February and July 2020 for pre- and post-transfer distributions. It is apparent how, compared to February, the following months witness a considerable increase in the share of individuals recording zero pre-transfer earnings (up to a staggering 35% in both April and May), and how government transfers effectively contribute to make the Lorenz curve for post-pandemic months more similar to that of February.

Finally, we verify that the results discussed so far are robust to using alternative measures of inequality. In particular, we consider the Theil index, a commonly used measure of inequality which relates to the concept of entropy and to the Shannon's diversity index and, unlike the

Gini, provides an additively decomposable measure of inequality.[15]

Appendix figures A.3 and A.4 replicate figures 5 and 6 for the Theil index. The picture that emerges largely confirms what observed for the Gini index: an increase in pre-transfers inequality in March which becomes much more pronounced in April, May, and June, but that is tamed by the phasing in of government transfers and subsidies.

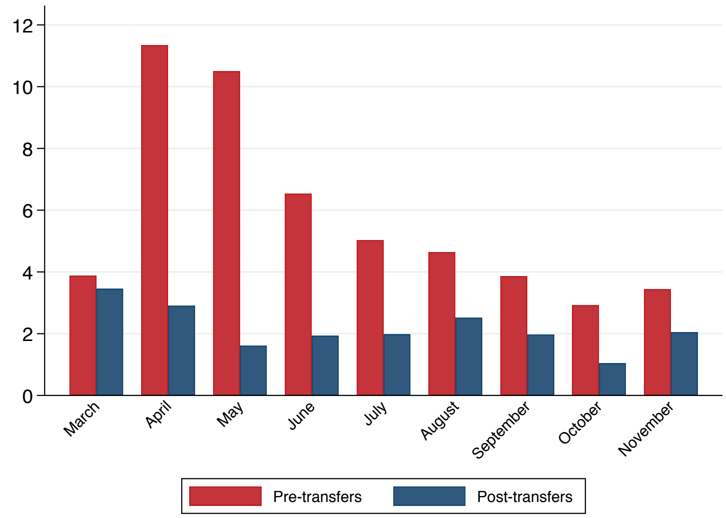


FIGURE 6: DIFFERENCES-IN-DIFFERENCES IN PRE- AND POST-TRANSFER GINI WITH RESPECT TO FEBRUARY 2020 (RELATIVE TO 2019)

4.2. THE EVOLUTION OF BETWEEN AND WITHIN GROUP INEQUALITY

In this section we analyze the decomposition of inequality between and within groups and regions. For this purpose we use the classical decomposition of the Gini index in a between-group inequality, a within-group inequality and overlap or interaction.[16] Opposite to the case of the Theil index, which is decomposable, the Gini index is not. For this reason there is an overlap component which occurs when there are individuals in different groups who have an income difference between them that is of opposite sign to the average income difference

[15] Starting from the formula of the generalized entropy index:

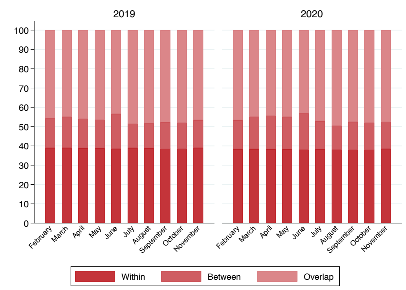
$$GE(\alpha) = \frac{1}{N\alpha(\alpha-1)} \sum_{i=1}^{N} \left[\left(\frac{y_i}{\bar{y}} \right)^{\alpha} - 1 \right] \; for\, \alpha \neq 0, 1 \tag{7}$$

where $\bar{y}$ is the mean of y, the formula for the Theil index is derived by setting $\alpha = 1$:

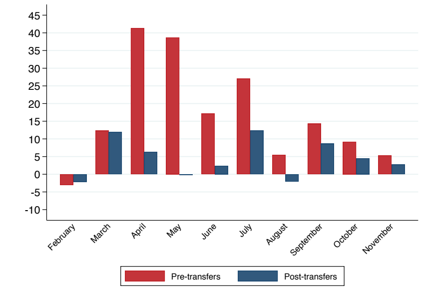
$$GE(1) = Theil = \frac{1}{N} \sum \left(\frac{y_i}{\bar{y}} \right) ln \left(\frac{y_i}{\bar{y}} \right) \tag{8}$$

[16] See Mookherjee and Shorrocks (1982) or Cowell (2000).

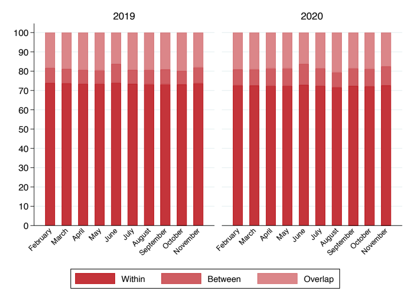
between the groups. This effect is called trans-variation between groups.



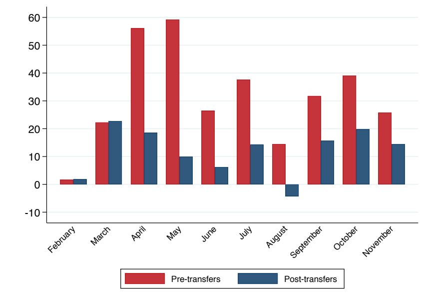
(a) Age decomposition



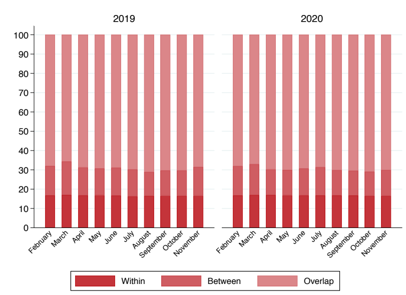
(b) Age diff-in-diffs



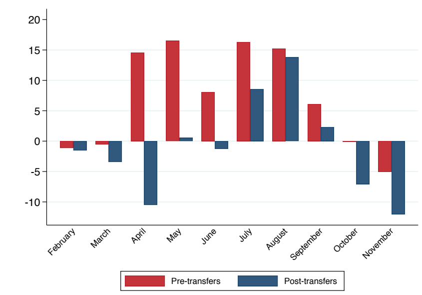
(c) Foreign-born decomposition



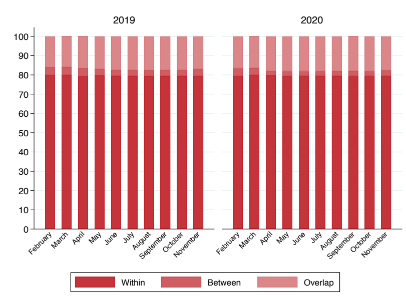
(d) Foreign-born diff-in-diffs



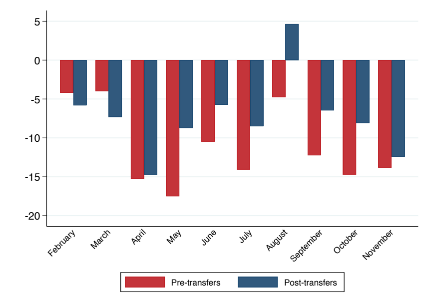
(e) Regional decomposition



(f) Regional diff-in-diffs



(g) Rural/Urban decomposition



(h) Rural/Urban diff-in-diffs

Note: The first column includes the decomposition of pre-subsidies inequality in between inequality, within inequality and the overlap component. The second column represents the percentage change in the between component for each month from 2019 to 2020 for the pre-benefits and the post-benefits situation.

FIGURE 7: DECOMPOSING INEQUALITY BY GROUPS AND REGIONS

Figure 7 depicts the decomposition of inequality in four dimensions of the data.[17] High levels of overlapping indicate that the characteristic analyzed contributes only slightly to inequality. For instance, the decomposition by region[18] and, to a lesser extent, by age groups generate a large overlapping component. However, these are also the dimensions in which the between component is more important, opposite to what happen in the nationality or the rural/urban dimensions.

The second column of Figure 7 depicts the growth rate of the between component between each month of 2020 and the corresponding month of 2019. The between component of age increases significantly in April and May to reduce its importance thereafter. The between inequality between foreign born and native workers increases drastically in March and stays up most of the months. The inequality between regions increases significantly between April and September although, as we argued before, the overlap component is quite large. Finally, the inequality between rural and urban areas goes down significantly during all the months after April.

4.3. COVID-19 AND INEQUALITY: DIFFERENCES WITHIN GROUPS AND REGIONS

In this section we examine how the increase in inequality due to the pandemic affected different groups of the population. Our primary focus is on within-group inequality. In particular, we are interested in understanding whether inequality evolved differently among different age groups, among natives than among foreign-born and, finally, within different regions of Spain. Shedding light on this aspect is crucial to understand what segments of the population were most vulnerable to the economic consequences of the pandemic and less protected by the policies implemented by the government.

This choice is motivated by the fact that our analysis of the decomposition between within-group and between-group inequality indicates that the former accounts for most of the variation in total inequality while the latter explains very little.[19]

We start by examining how inequality evolved among men and among women, both before and after government transfers are taken into account. Figure 8 shows how the Gini index increased in each month between March and November relative to February, net of the difference between the same months in 2019, separately for women (left) and men (right). Though the pattern is very similar for the two groups, the increase in pre-transfers inequality is slightly larger within women in all months, particularly in and after March. The same

[17] The decomposition is performed in each dimension separately. It is also possible to analyze the contribution of several dimensions simultaneously.

[18] Spanish regions correspond to the Autonomous Communities.

[19] The results based on the decomposition of the Theil index lead to similar conclusions and are available upon request.

small difference holds for post-transfer earnings, which suggests that government intervention does not entirely close the gap.

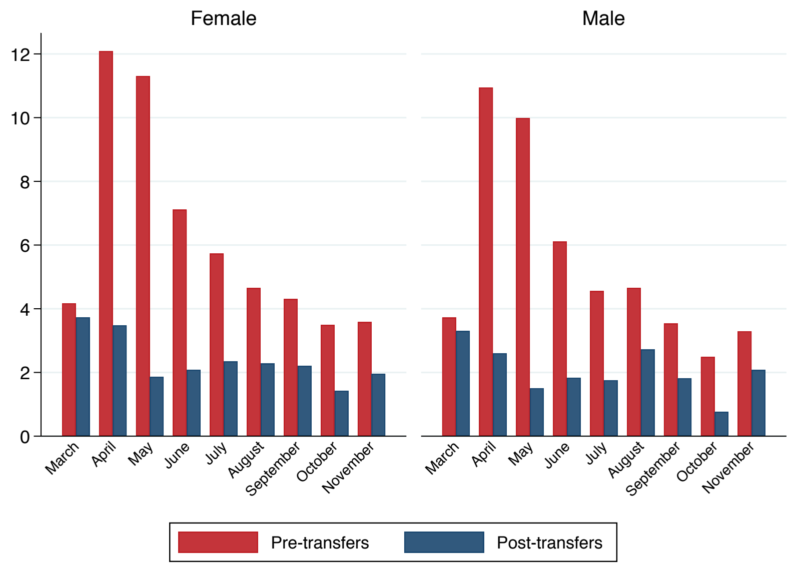


FIGURE 8: DIFFERENCES-IN-DIFFERENCES IN PRE- AND POST-TRANSFER GINI WITH RESPECT TO FEBRUARY 2020 (RELATIVE TO 2019) - BY GENDER

To understand whether raising inequality affected young cohorts more than older ones, we then examine differences across age groups. To this end, in Figure 9 we report the difference-in-difference between each month and February separately for people aged 16 to 29, 30 to 49, and 50 to 64, respectively. The results clearly indicate that pre-transfers inequality, as measured by the Gini index, raised considerably more among the young than among adults and the elderly. This is arguably due to the fact that a disproportionate share of younger workers, particularly low skilled ones, have temporary contracts and are employed in the service sector, especially in industries such as leisure and tourism, that were heavily affected by the crisis. Although government intervention mitigated such increase, the post-transfers inequality for younger cohorts raised more than twice as much than for the other groups.

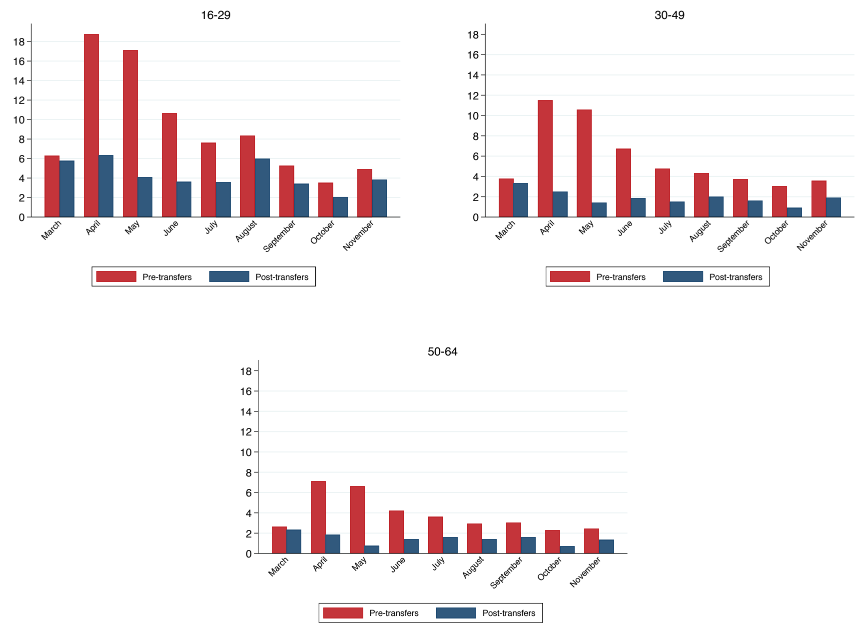


FIGURE 9: DIFFERENCES-IN-DIFFERENCES IN PRE- AND POST-TRANSFER GINI WITH RESPECT TO FEBRUARY 2020 (RELATIVE TO 2019) - BY AGE

The availability of information on each account holder's country of birth, allow us to analyse whether the crisis affected natives and foreign-born in different ways. Shedding light on this aspect is especially important since immigrants can generally count on a thinner support network (e.g., through extended family), and may hence depend more on public support in the face of adversities.

As shown in Figure 10, the pre-benefits Gini index increases significantly more for foreign-born individuals than for native-born. Interestingly, when dividing foreign-born by the level of GDP per capita of the country of origin a big difference emerges between immigrants from richer countries and all the others.[20] As shown in Figure 11, while among the latter pre-transfers inequality rises considerably, among the former the increase is more limited. Government intervention is generally effective at limiting the spike in inequality. However, this is more the case for natives than for immigrants and, among these, even less so for immigrants from poorer countries for whom post-transfers inequality remains very high.

[20] We classify as "rich" those countries that, according to the the World Bank classification are "High income", while the others include countries classified as "Upper middle income", "Lower middle income", and "Low income".

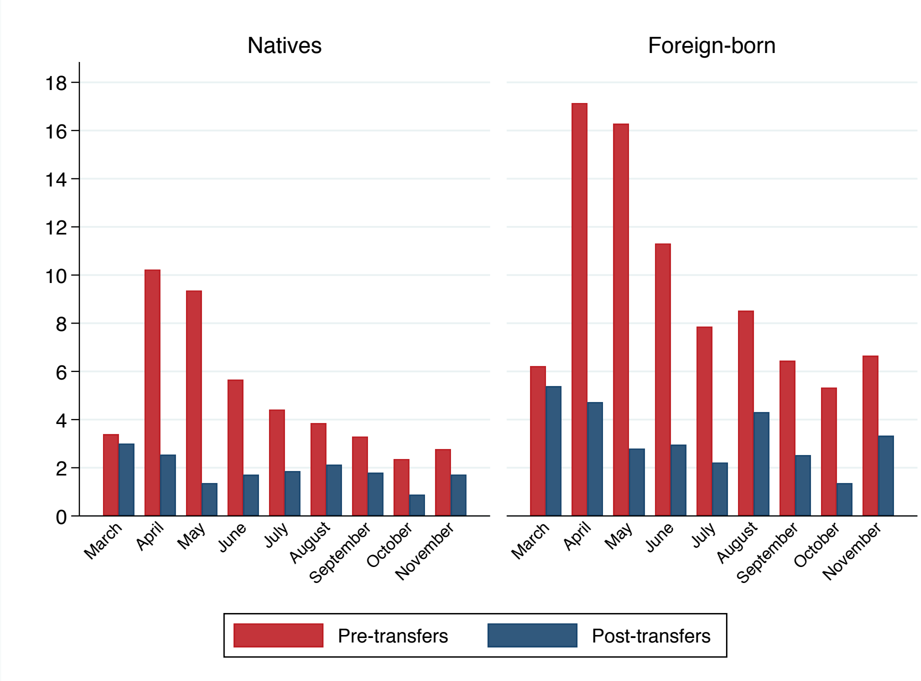


FIGURE 10: DIFFERENCES-IN-DIFFERENCES IN PRE- AND POST-TRANSFER GINI WITH RESPECT TO FEBRUARY 2020 (RELATIVE TO 2019) - BY COUNTRY OF BIRTH

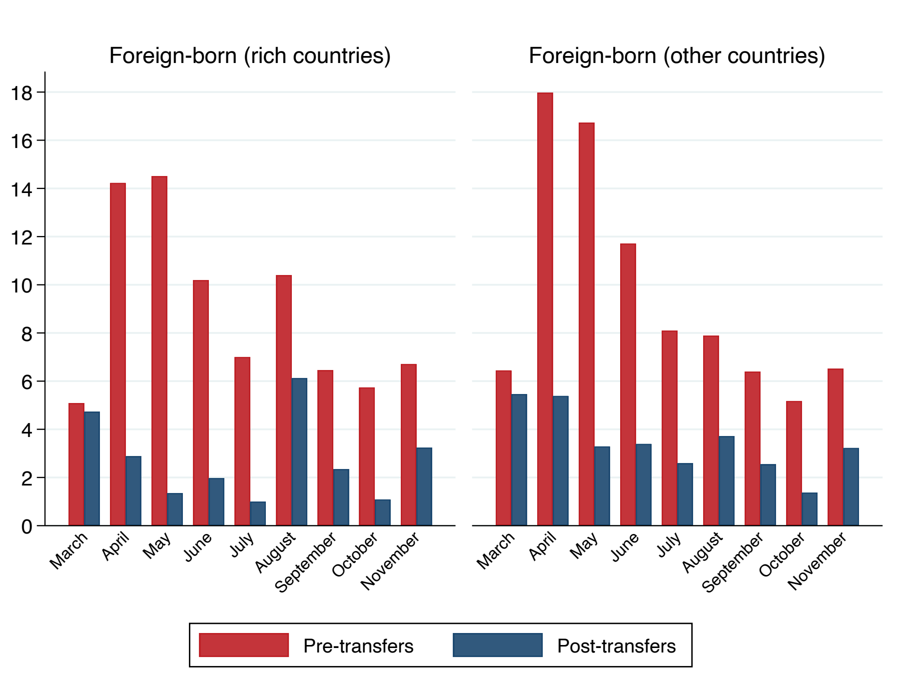


FIGURE 11: DIFFERENCES-IN-DIFFERENCES IN PRE- AND POST-TRANSFER GINI WITH RESPECT TO FEBRUARY 2020 (RELATIVE TO 2019) - FOREIGN-BORN BY COUNTRY OF BIRTH

Finally, the presence of information about account holders' place of residence, allows us to gauge whether changes in earnings in equality were more pronounced in certain areas of the country than in others. Figures 12 and 13 show the increase in pre-transfers and post-transfers Gini in May, July and November of 2020 with respect to February 2020 (relative to the change in 2019) in each region of Spain.

It is clear that in certain regions pre-transfers inequality rose much more than in others (e.g., +18% in the most affected region vs. +7% in the least). The spike was especially pronounced in the Balearic and the Canary Islands, two regions that largely depend on tourism, one of the sectors most affected by the restrictions to national and international mobility adopted by the government in response to the pandemic. Differences across regions are largely offset by government transfers, and even in the regions with the highest increase in post-transfers Gini (such as Navarra and Catalonia) this is never above 5 percentage points.

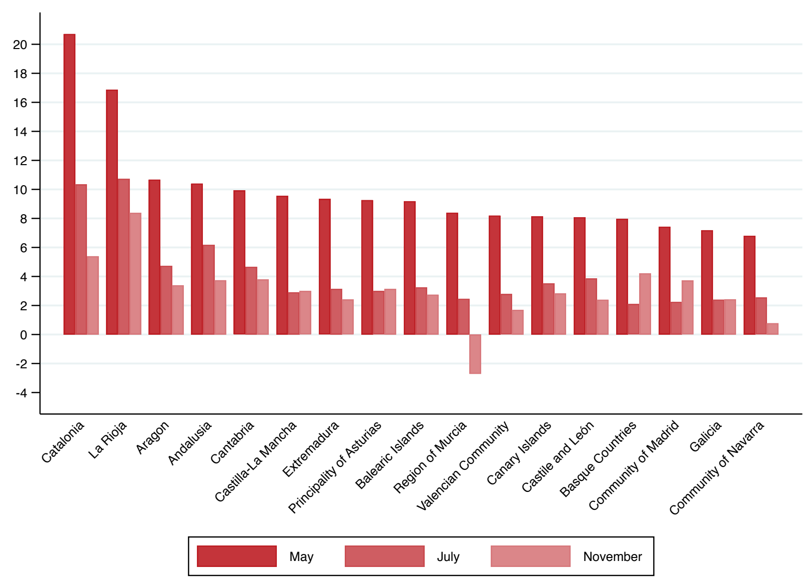


FIGURE 12: DIFFERENCES-IN-DIFFERENCES IN PRE-TRANSFERS GINI
MAY, JULY AND NOVEMBER 2020 WITH RESPECT TO FEBRUARY 2020 (RELATIVE TO 2019) - BY REGION

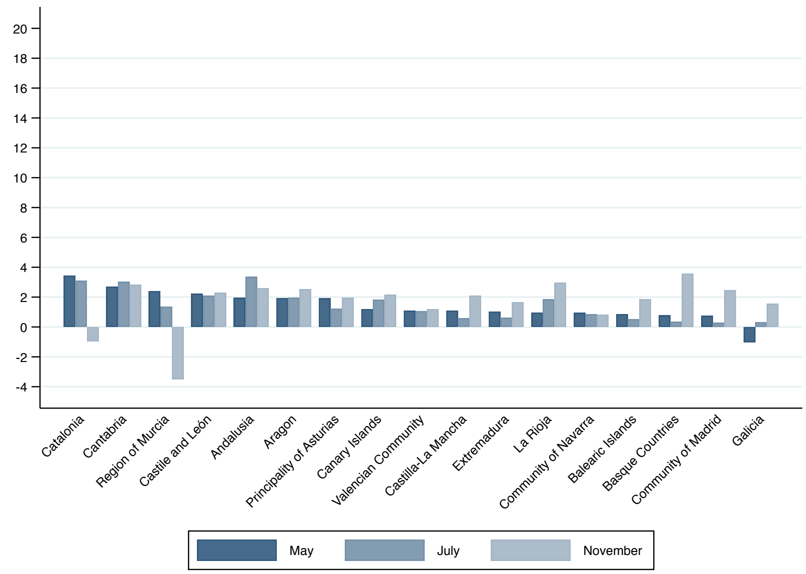


FIGURE 13: DIFFERENCES-IN-DIFFERENCES IN POST-TRANSFERS GINI
MAY, JULY AND NOVEMBER 2020 WITH RESPECT TO FEBRUARY 2020 (RELATIVE TO 2019) - BY REGION

To further examine geographical differences we explore the determinants of the change in inequality across regions.[21] First, to study the determinants of the difference-in-differences change in the regional Gini indices, we estimate the following equation:

$$Ln((Gini_{i,t} - Gini_{i,February2020}) - (Gini_{i,t-12} - Gini_{i,February2019})) = \beta_0 + \sum_{j=1}^{J} \beta_j X_{ijt} + \varepsilon_{it} \quad (9)$$

where the dependent variables is the difference-in-differences of the Gini index in each region (before or after subsidies) in logarithm, and X_{ijt} include a regional mobility index indicator, an indicator of the percentage of population living in rural areas in each region, indicators of the sectoral economic structure in each region and monthly dummies. Mobility is defined as the percentage of mobility in each region with respect to the reference period (14 to 20 of February) using anonymized data of mobile phone usage and localization from

[21] Spain is divided into 17 regions and 50 provinces. Therefore, provinces provide a much finer geographical disaggregation than regions. We have used regional data as the information on the sectoral distribution of activity is more disaggregated than at the province level and it is important to explain the economic impact of the COVID-19. However, we also computed the analysis at the province level and obtained very similar results.

one of the three main operators in Spain[22] and it is expressed as the logarithm of percentage points. The economic structure of each region is defined as the percentage of GDP of each of the following sectors: Agriculture (NACE sector A), Mining, Energy Water Supply (B,D,E), Manufacturing (C), Construction (F), Retail trade, Transport and Hospitality Services (G, H, I), Financial and insurance activities (K), Other Services (J, L, M, N, R, S) and Public Sector (O, P, Q) omitted category), and available from the Spanish National Statistics Institute (Spanish Regional Accounts from the latest year with information available at the regional level, 2019). The percentage of people living in rural areas is computed using information on the municipality of the bank account holder, which is defined as rural if its population is less than 30,000 inhabitants and the population density is less than 100 inhabitants per square kilometer.

The results in Table 3 indicate a statistically significant relationship between the decline in mobility and the increase in pre-transfer wage inequality at the regional level. Interestingly, once government transfers are accounted for, this correlation decreases and becomes statistically insignificant, a result which corroborates the importance of the welfare state in mitigating the increase in inequality associated with lower mobility. The sectoral composition also matters for the increase in wage inequality, with retail and hospitality - the sectors mostly affected by the pandemic - displaying a large positive coefficient, and finance and insurance, and mining, energy and water supply, the least affected, showing a negative coefficient.

In the second part of analysis we explore the determinants of the difference between the pre-benefits and the post-benefit Gini index, which captures the mitigating effect of government intervention on inequality. To this end we estimate the following equation:

$$Ln(GiniPre_{i,t} - GiniPost_{i,t}) = \gamma_0 + \sum_{j=1}^{J} \gamma_j X_{ijt} + \eta_{it} \tag{10}$$

The dependent variable is the log difference between the Gini index before and after government transfers in region i in month t. On the right-hand side we include all the variables in equation (9) plus two variables meant to capture the effectiveness of the public administration in processing unemployment and furlough schemes in a region: i.e., the average processing time (in days) and the percentage of applications processed in under 15 days.

The results, reported in Table 4, indicate that the change in Gini index due to government intervention is related to mobility, regional characteristics and administrative efficiency. In particular, the positive coefficient on administrative efficiency confirms that the timely delivery of government transfers is crucial to limit the increase in wage inequality. Finally, the

[22] This regional mobility index was created by the Spanish Ministry of Transportation, Mobility and Urban Agenda and is based on over 13 million records (https://www.mitma.gob.es/ministerio/covid-19/evolucion-movilidad-big-data/movilidad-ccaa).

coefficients on the monthly dummies confirm that the reduction in inequality due to government transfers is especially sizeable in April and May (compared to March) and decreases afterwards.

TABLE 3: EXPLAINING THE DYNAMICS OF THE GINI INDEX ACROSS REGIONS

	(1) Pre-benefits	(2) Pre-benefits	(3) Pre-benefits	(4) Post-benefits	(5) Post-benefits	(6) Post-benefits
Log Mobility	-0.056***	-0.057***	-0.019	-0.007*	-0.007**	0.003
	(-9.23)	(-14.11)	(-1.97)	(-2.51)	(-3.11)	(0.43)
Rural (%)		0.053*	0.061***		-0.001	0.001
		(2.12)	(3.90)		(-0.10)	(0.08)
Agriculture		0.102	0.058		0.122*	0.110*
		(0.98)	(0.88)		(1.98)	(2.03)
Mining, Energy & Water Supply		-0.601	-0.713***		-0.222	-0.254
		(-1.86)	(-3.49)		(-1.16)	(-1.52)
Manufacturing		0.086*	0.053*		0.055**	0.046*
		(2.56)	(2.34)		(2.80)	(2.49)
Construction		0.647*	0.374		-0.028	-0.107
		(2.16)	(1.87)		(-0.16)	(-0.65)
Finance & Insurance		-0.650	-0.614*		0.111	0.122
		(-1.49)	(-2.23)		(0.43)	(0.54)
Retail & Hospitality		0.424***	0.426***		0.124**	0.124***
		(6.25)	(9.98)		(3.08)	(3.57)
Other Services		0.099	0.112		0.019	0.023
		(0.66)	(1.19)		(0.22)	(0.30)
April			0.048***			-0.008
			(6.34)			(-1.31)
May			0.048***			-0.019***
			(9.93)			(-4.75)
June			0.021***			-0.018***
			(4.76)			(-5.00)
July			0.008			-0.023***
			(1.30)			(-4.42)
August			0.006			-0.020***
			(0.83)			(-3.64)
September			-0.001			-0.021***
			(-0.20)			(-4.48)
October			-0.011*			-0.025***
			(-2.28)			(-6.54)
November			-0.008*			-0.017***
			(-2.05)			(-5.13)
Constant	0.024***	-0.130***	-0.106***	0.013***	-0.029*	-0.002
	(7.15)	(-6.54)	(-6.52)	(8.57)	(-2.46)	(-0.13)
R-squared	0.335	0.732	0.899	0.036	0.317	0.514
N	171	171	171	171	171	171

* p<0.05, ** p<0.01, *** p<0.001

TABLE 4: EXPLAINING GINI BEFORE AND AFTER POLICY RESPONSE ACROSS REGIONS

	(1)	(2)	(3)	(4)	(5)
Log Mobility	-0.607***	-0.578***	-0.574***	0.098	0.113
	(-10.43)	(-9.41)	(-9.95)	(0.80)	(0.91)
Rural (%)	0.212	0.136	0.179	0.229	0.353
	(0.60)	(0.38)	(0.52)	(1.13)	(1.79)
Agriculture	5.282***	5.503***	5.593***	4.920***	4.667***
	(3.54)	(3.68)	(3.84)	(5.87)	(5.60)
Mining, Energy & Water Supply	-6.632	-7.848	-8.178	-10.715***	-9.642***
	(-1.43)	(-1.67)	(-1.80)	(-4.10)	(-3.74)
Manufacturing	-1.973***	-2.051***	-1.862***	-2.679***	-2.513***
	(-4.11)	(-4.26)	(-3.97)	(-9.53)	(-8.81)
Construction	5.502	4.274	3.639	-1.376	-0.666
	(1.28)	(0.98)	(0.86)	(-0.54)	(-0.26)
Finance & Insurance	-21.852***	-19.853**	-20.851***	-17.685***	-20.566***
	(-3.48)	(-3.10)	(-3.41)	(-4.92)	(-5.97)
Retail & Hospitality	1.437	1.572	1.481	1.697**	1.487**
	(1.47)	(1.61)	(1.56)	(3.17)	(2.78)
Other Services	4.598*	4.301*	4.255*	4.293***	4.624***
	(2.15)	(2.00)	(2.03)	(3.64)	(3.93)
Unem. Claims Av. Processing time		-0.049		-0.086**	
		(-1.41)		(-3.07)	
Unem. Claims % solved cases under 15 days			3.280**		2.006**
			(3.09)		(2.68)
April				0.993***	0.954***
				(10.44)	(10.07)
May				0.935***	0.885***
				(14.92)	(14.71)
June				0.364***	0.304***
				(5.98)	(5.45)
July				0.251**	0.174*
				(2.86)	(2.15)
August				0.219*	0.112
				(2.26)	(1.30)
September				0.150	0.079
				(1.73)	(1.02)
October				0.136	0.060
				(1.74)	(0.91)
November				0.122	0.033
				(1.83)	(0.61)
Constant	-3.690***	-3.527***	-3.405***	-3.132***	-3.140***
	(-12.93)	(-11.49)	(-11.62)	(-15.25)	(-15.16)
R-squared	0.636	0.640	0.656	0.898	0.897
N	171	171	171	171	171

* p<0.05, ** p<0.01, *** p<0.001

5. EXPLAINING INEQUALITY DYNAMICS IN THE LABOR MARKET

In this section we examine different drivers of the changes in inequality documented above. Overall inequality among working-age individuals can be divided into inequality among workers - i.e., due to the wage dispersion - and inequality between employed and unemployed. Hence, the change in the Gini index can be decomposed as follows:

$$\Delta Gini = e_0 \Delta Gini_{emp} - (1 - Gini_{emp,1})\Delta e \tag{11}$$

where e is the proportion of working age population that is employed, and $Gini_{emp}$ is the Gini coefficient for the working population.

Previous research using estimations based on this decomposition identify changes in employment status as the main driver of the evolution of inequality in Spain. OECD (2015) estimates that around 80% of the increase of inequality in Spain after the financial crisis (2007-2011) was due to increases unemployment. Of an overall increase in inequality of 7.5 points, 6 points were due to the decline in employment and the rest to increase in wage inequality among workers. Similar calculations indicate that 87% of the 11-point increase in pre-benefits Gini index recorded between February and April 2020 was due to employment reduction. Hence, investigating changes in employment seems crucial to fully understand the evolution of inequality since the pandemic begun. To this end, using individual level data we estimate the following linear probability model:

$$E_{ijt} = \sum_{k=1}^{5} \alpha_{kt} I(q_k) + \sum_{g=1}^{G} \beta_{gt} X_{ijtg} + \sum_{j=1}^{49} \delta_j I(Prov = j) + \varepsilon_{ijt} \tag{12}$$

E_{ijt} is a dummy variable equal to one if worker i in province j is employed a time t; $I(q_k)$ is an indicator equal to one if worker i's wage prior to the beginning of the pandemic (i.e., February 2020) was in the k quintile of the wage distribution; the vector X_{ijtg} includes a set of individual characteristics (i.e., gender, age group,[23] native/foreign-born) and a provincial measure of mobility. To examine how the effect of these factors on employment varies over time, we interact each variable with month fixed effects.[24] Finally $I(Prov = j)$ represents province fixed effects, which capture all time-invariant factor of a province, including the sectoral composition of the economy.[25]

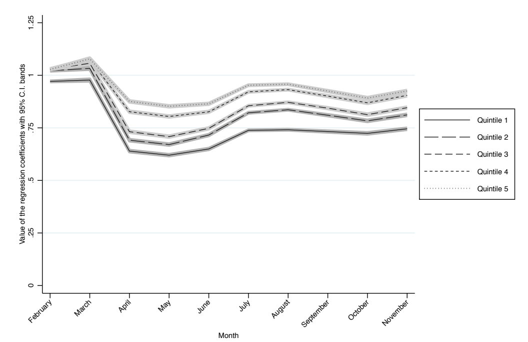
[23] Age groups include: people aged 30 or less, people aged 30 to 54, and people aged 54 or more, which is the omitted category.

[24] All results discussed below are robust to controlling for the interaction between province fixed effects and time dummies.

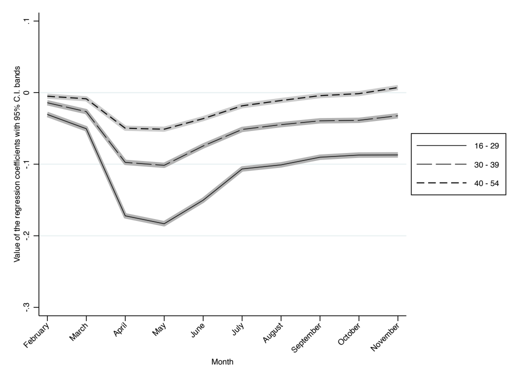
[25] Since the dataset contains more than 30 millions of observations we need an efficient algorithm to estimate the regressions. We are using algorithm AS274 that updates the orthogonal reduction as each new case is added instead of calculating the Cholesky decomposition from the design matrix X. See Strang (2019).

For this exercise, we do not consider furloughed workers as being employed since their activity and contractual relationship with the firm is suspended.

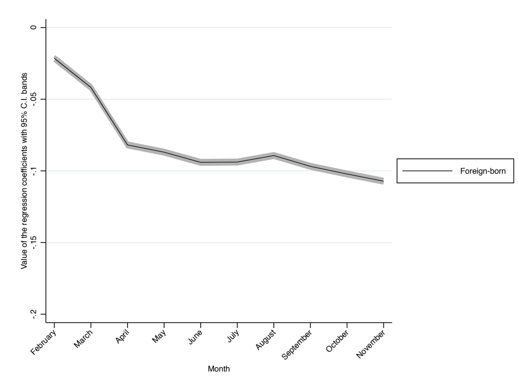
We start by looking at how the probability of being employed evolves over time for workers in different parts of the pre-pandemic wage distribution. In panel a of Figure 14 we plot the coefficients on the interactions between each quintile and month fixed effects, as well as the respective 95% confidence intervals, from a regression including all other controls.[26] The results indicate a positive monotonic relationship between pre-pandemic wage and the probability of being employed after the beginning of the pandemic. Difference across groups are sizeable and persistent. For example, individuals in the bottom quintile of the distribution were 25 percentage point less likely to be employed than those in the top quintile as of April 2020; the difference remained stable until June 2020, and, despite a slight decrease in the following months, was still 17 percentage points in November 2020.



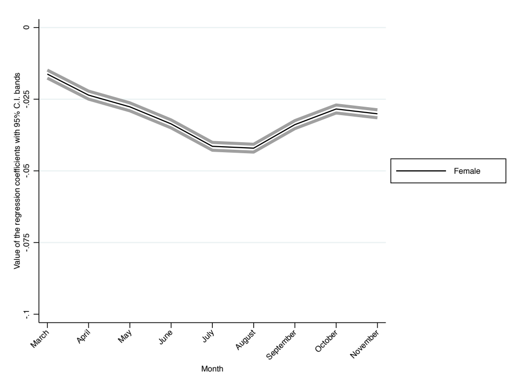
(a) Pre-pandemic wage quintile



(b) Age (reference group: workers aged 55-65)



(c) Country of birth (reference group: natives)



(d) Gender (reference group: males)

FIGURE 14: PROBABILITY OF EMPLOYMENT CONDITIONAL ON OBSERVABLES

Panel b reports the same results by age group (with individuals aged 54 or more as the default

[26] Since the sample size is very large, the confidence intervals are extremely tight around the point estimate.

category). The results clearly document that, during the lockdown months (i.e., March to May), younger cohorts were considerably less likely to be employed than all other age groups (up to 18 p.p. in May 2020), with a gradual though only partial recovery occurring during and after the summer (8 p.p. difference in November 2020).

A different pattern emerges in panel c in which we compare the evolution of employment of foreign-born and natives. While in the early months of the pandemic foreign-born are somewhat less likely than natives to be employed (4 to 9 p.p. between March and May), what is striking is that this difference persists and gradually increases with the reopening of the economy, reaching 11 p.p. in November 2020.

To analyze the probability of employment over time by gender we need to slightly modify the baseline specification:[27]

$$E_{ijt} = \phi_t I(Women_i = 1) + \sum_{g}^{G} \beta'_{gt} X_{ijtg} + \sum_{j=1}^{49} \delta'_j I(Prov = j) + u_{ijt} \tag{13}$$

In panel d of Figure 14 we plot the coefficient on the interaction between the gender dummy and month fixed effects, which captures how the effect of being female on employment evolves over time. The result indicate that, controlling for other observables, women were significantly less likely than men to be employed after the beginning of the pandemic. The gender gap is smaller than the differences by pre-pandemic wage, age, and place of birth, and tends to gradually decrease after August.

Finally, Figure 15 depicts the evolution of the impact of mobility on employment. It is interesting to note how this effect evolves over time. While during the first wave of the pandemic (i.e., March-May) mobility had a large impact on employment, this was much more limited during the second wave (i.e., September-November). This finding supports the view that, as the pandemic persists, economic activity is less sensitive to mobility restrictions, arguably because individuals and firms become more adaptive.

[27] Since women have a significantly lower wage than men, considering the quintiles makes it difficult to interpret the results.

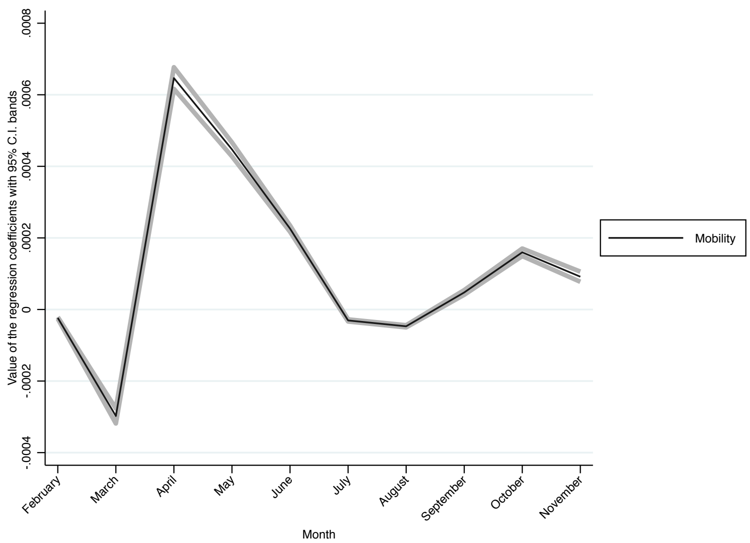


FIGURE 15: PROBABILITY OF EMPLOYMENT BY THE MOBILITY INDEX CONDITIONAL ON OBSERVABLES.

6. INEQUALITY AND POLITICAL OUTCOMES

Finally, we explore how the ability of the government to reduce the increase in inequality caused by the pandemic influenced citizens' electoral behavior. To this end, we look at voting outcomes in two regional elections held in Galicia and the Basque Country in early July 2020, i.e., shortly after the first wave of the pandemic.

A large body of work has analyzed the effect of negative economic shocks on voting, finding an ambiguous impact on participation, a positive effect on support for left-wing parties, and a negative effect on support for the incumbent (Margalit, 2019). In this case, the economic shock is due to a global pandemic, so that voters cannot blame the incumbent for the hardship they may experience. Yet, they can react to the measures enacted by the government to alleviate the economic consequences of the pandemic.[28] In light of this, we examine how voters' behavior were affected by the reduction in inequality due to government transfers.

For each municipality in the two regions, we compute the difference in Gini coefficient between pre- and post-government transfers. We then combine this information with data from a post-electoral survey available from the *Centro de Investigación Sociológicas* (CIS), one of the most reputed sources of electoral surveys in Spain for decades. We focus on three types of outcomes: i) abstention rate, ii) voters' ideological position (on a 10-point left-right scale), For the ideological position, we consider the weighted average score of all respondents conditional on the political party they voted in a given municipality. The weight

[28] In this respect, our analysis relates to previous work on the effect of disaster relief on voter behavior (Chen, 2013).

is the proportion of votes for each political party in a given municipality. and iii) support for the largest parties in the region. For all outcomes, we consider the difference between 2020 and the previous regional election held in 2016. The results are reported in Table 5. In all specifications we control for the size of the population (in logs), gender and age distribution (i.e., share of males, share of people aged 50 to 64) and education (share of people with high school diploma or more).

In the first two columns, we pool data for the two elections. Column (1) reports the results for abstention. From a theoretical standpoint, whether government performance at reducing inequality encourages or deter participation is a priori unclear, as it arguably depends on local voters' political preferences.[29] We find that a larger reduction in inequality due to government intervention is associated with a significantly lower abstention rate (compared to 2016), which suggests that, on average, government effectiveness in this domain tends to mobilize voters. Column (2) examines the effect on voters' ideological leaning. We find that where government intervention leads to a larger reduction in inequality, voters tend to report leaning more to the left of the ideological spectrum, though the effect is not statistically significant at standard levels.

In the following columns, we look at the effect on the (change in) vote share for the largest parties separately for Galicia (columns 3-5) and the Basque Country (6-8), respectively. In both regions, the electoral contest is defined over two dimensions: left vs. right, and independentist vs. non-independentist. The situation was further complicated by the difficulty for citizens to distinguish what level of government - i.e., national, regional, or local - was responsible for the various schemes enacted to cope with the economic effect of the pandemic. Our findings suggest that a larger reduction in inequality due to government transfers, was generally associated with increased support for left-wing independentist parties at the expense of national parties both on the right (in Galicia) and on the left (in the Basque Country).

Regarding the impact on support for the incumbent, here too, the issue is complicated by the co-existence of a regional incumbent government and of a national one. Our findings indicate that, in general, a larger drop in inequality due to government intervention is, if anything, associated with lower support for the incumbent, the conservative party in Galicia and the socialist party in the Basque Country, respectively. One interpretation of this finding is that voters were disappointed with the delay with which government income-support mechanisms were deployed, a sentiment which was likely more pronounced in areas where inequality increased the most and such schemes were most needed.

[29] For example, looking at the effect of FEMA hurricane disaster aid delivery in the weeks prior to the 2004 US elections, Chen (2013) finds that it increases participation by Republican voters, supporters of the incumbent party, but depresses it among Democratic voters.

TABLE 5: INEQUALITY AND VOTING

	All regions		Galicia			Basque Country		
	Abstention Diff 2020-2016	Ideology Diff 2020-2016	Conservative PP	Socialist PSOE	Left-independentist BNG	Right-independentist PNV	Left-independentist EH BILDU	Socialist PSOE
$\text{Gini}_{PRE} - Gini_{POST}$	-0.313*	-0.128	-0.494***	0.841	2.904*	0.405	1.783**	-1.149**
	(0.166)	(0.086)	(0.173)	(0.865)	(1.721)	(0.522)	(0.738)	(0.484)
Proportion tertiary education	0.309***	-0.065	0.099	-0.524*	2.183**	0.161	-0.731***	0.051
	(0.062)	(0.046)	(0.088)	(0.298)	(1.013)	(0.098)	(0.223)	(0.180)
log(population)	0.016**	0.007	0.009	0.051*	0.144*	0.023**	0.039**	0.023*
	(0.007)	(0.004)	(0.008)	(0.027)	(0.086)	(0.010)	(0.015)	(0.014)
Proportion population 50-64	0.042	-0.017	-0.218*	0.140	-2.416*	-0.412***	0.999***	0.214
	(0.096)	(0.075)	(0.126)	(0.357)	(1.241)	(0.142)	(0.340)	(0.336)
Proportion of men	-0.038	0.196***	0.170	0.420	0.710	-0.184	-0.677	0.370
	(0.127)	(0.064)	(0.122)	(0.469)	(1.242)	(0.180)	(0.484)	(0.312)
Constant	-0.001	-0.115**	-0.049	-0.691**	0.076	-0.070	-0.027	-0.188
	(0.071)	(0.045)	(0.081)	(0.271)	(0.841)	(0.112)	(0.240)	(0.166)
Observations	195	195	125	125	120	70	69	61

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7. Concluding remarks

Measuring the economic impact of the COVID-19 pandemic, and of the measures adopted by governments to tackle it, is key to evaluate past policy decisions and guide future ones. Yet, given the speed at which the pandemic has unfolded, relying solely on official statistics, which become available with considerable delay, is not a viable option and new creative solutions need to be found.

This paper moves in this direction by proposing the use of large anonymized data from bank records to track the evolution of economic inequality at a high frequency before and during the pandemic. Specifically, we use data on over three million bank accounts from Caixabank, Spain's second-largest bank by total assets and first by direct payroll deposits. Crucially, the availability of data on both payroll payments and government transfers make it possible to assess the impact of the pandemic both before and after government intervention.

We first confirm that bank payroll data are a reliable and valuable source of information to track changes in the wage distribution in an accurate and timely fashion. We then confirm that our sample is highly representative of the Spanish working population and that the bank data closely match the distribution of wages from official wage surveys.

We then document that, in the absence of the large public scheme activated soon after the beginning of the pandemic, earnings inequality would have risen dramatically. This tendency was mainly driven by the severe effect of the pandemic on low-wage workers, many of whom suffered large wage cuts or lost their job altogether. The rise in inequality was especially strong during the lockdown months and gradually weakened following the reopening of the economy. The effect was especially pronounced among young people and immigrants, and in regions more dependent on sectors, like tourism, heavily affected by the mobility restrictions put in place to limit the spread of the virus.

Comparing the inequality index before and after considering the public benefits we show that government transfers and furlough schemes were very effective at mitigating the rise in inequality, and at providing a much needed safety net to the most affected segments of the population.

Finally we show that inequality dynamics is mostly driven by change in the employment rate. Using individual-level regressions we analyze the role of different factors (demographic, geographic, etc.) on the evolution of the employment rate over the pandemic. We show that the probability of employment is lower the lower is the income quintile at which the worker belonged before the beginning of the pandemic. Younger and foreign-born workers exhibit also a lower rate of employment during the whole period, with the situation improving over time for young people but not for foreign-born individuals. The employment rate of women was also lower than that of male compared with the initial situation but the effect was not as strong as in the case of young or foreign-born people.

References

Addeco, *Análisis del pluriempleo en España y Europa*, Addeco, 2015.

Aspachs, O., R. Durante, A. Graziano, J. Mestres, M. Reynal-Querol, and J García-Montalvo, "Tracking the impact of COVID-19 on economic inequality at high frequency," *PLoS ONE*, 2021, *16* (3).

Baker, Scott R, Robert A Farrokhnia, Steffen Meyer, Michaela Pagel, and Constantine Yannelis, "Income, liquidity, and the consumption response to the 2020 economic stimulus payments," *Review of Asset Pricing Studies*, 2020, p. Forthcoming.

Bick, Alexander and Adam Blandin, "Real time labor market estimates during the 2020 coronavirus outbreak," *Unpublished Manuscript, Arizona State University*, 2020.

Carvalho, Vasco M, Stephen Hansen, Álvaro Ortiz, Juan Ramon Garcia, Tomasa Rodrigo, Sevi Rodriguez Mora, and Pep Ruiz de Aguirre, "Tracking the Covid-19 crisis with high-resolution transaction data," 2020.

Chen, J., "Voter partisanship and the effect of distributive spending on political participation," *American Journal of Political Science*, 2013, *57* (1), 200–217.

Chetty, Raj, John N Friedman, Nathaniel Hendren, and Michael Stepner, "Real-Time Economics: A New Platform to Track the Impacts of COVID-19 on People, Businesses, and Communities Using Private Sector Data," Technical Report, NBER Working Paper 27431 2020.

Cicala, Steve, "Early Economic Impacts of COVID-19 in Europe: A View from the Grid," *Energy Policy Institute at the University of Chicago, WP - April*, 2020.

Cowell, F., "Measurement of inequality, In Atkinson and Bourguignon (Eds.)," *Handbook of Income Distribution*, 2000, pp. 87–166.

Crawford, R., J. Davenport, R. Joyce, and P. Levell, "Household spending and coronavirus," *Institute of Fiscal Studies*, 2020, *Note 14,795.*

Giles, D., "Calculating the standard error for the Gini coefficient: some further results," *Oxford Bulletin of Economics and Statistics*, 2004, *66* (3), 425–433.

Gourinchas, Pierre-Olivier, "Flattening the pandemic and recession curves," *Mitigating the COVID Economic Crisis: Act Fast and Do Whatever*, 2020, p. 31.

Hacioglu, S., D. Kanzig, and P. Surico, "The distributional impact of the pandemic," *CEPR Discussion Paper 15101*, 2020.

ILO, "COVID-19 and the World of Work: Country Policy Responses," 2020.

IMF, "Policy responses to COVID-19," 2020.

Inglehart, Ronald and Pippa Norris, "Trump, Brexit, and the Rise of Populism: Economic Have-nots and Cultural Backlash," 2016.

Margalit, Y., "Political response to economic shocks," *Annual Review of Political Science*, 2019, *22*, 277–295.

Montalvo, J. and M. Reynal-Querol, "Distributional effects of COVID-19 on spending: a first look at the evidence from Spain," *BarcelonaGSE WP 1201*, 2020.

Mookherjee, D. and A. Shorrocks, "A decomposition analysis of the trend in UK income inequality," *Economic Journal*, 1982, *368*, 886–992.

OECD, *In it together. Why less Inequality Benefits all*, OECD Press, 2015.

Scheidel, W., *The Great Leverer*, Princeton University Press, 2017.

Sheridan, A., A. Andersen, E. Hansen, and N. Johannesen, "Social distancing laws cause only small losses of economic activity during the COVID-19 pandemic in Scandinavia," *Proceeding of the National Academy of Sciences*, 2020, (August 2020), 1–6.

Strang, G., *Linear algebra and learning from data*, Wellesley Cambridge Press, 2019.

The World Bank, "World Development Indicators," 2020.

Wade, L., "An unequal blow," *Science*, 2020, *368* (6492), 700–703.

APPENDIX

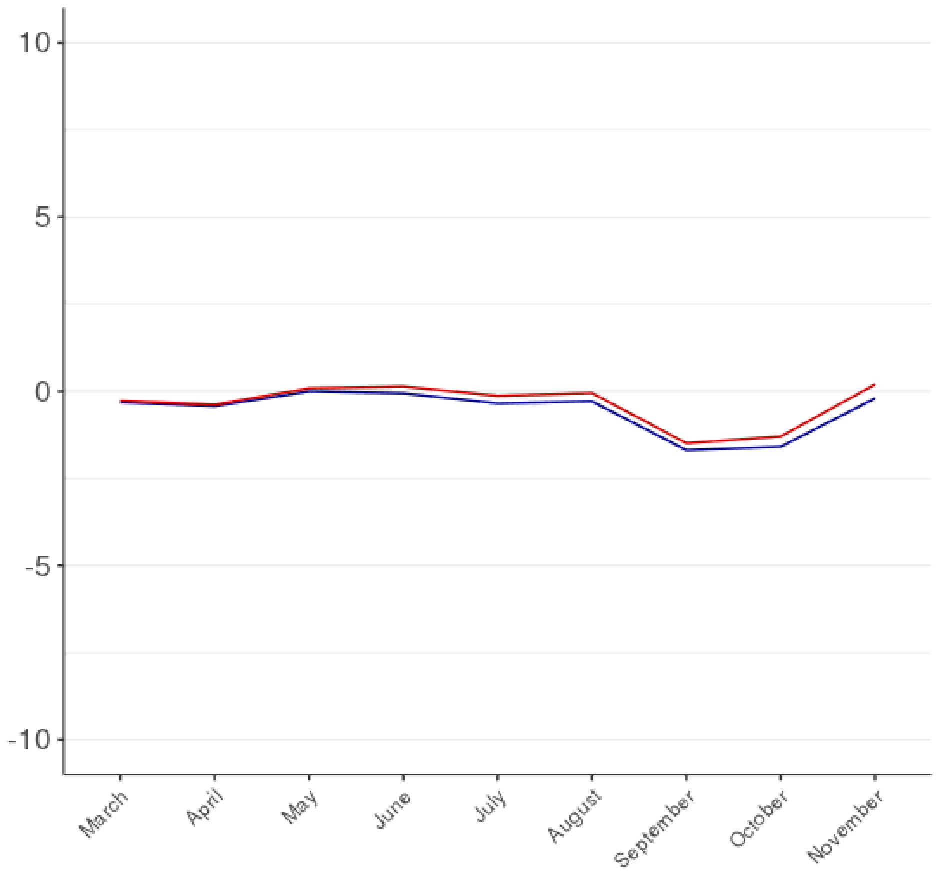
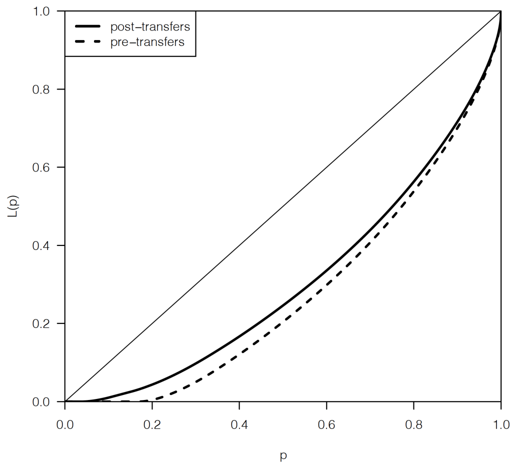
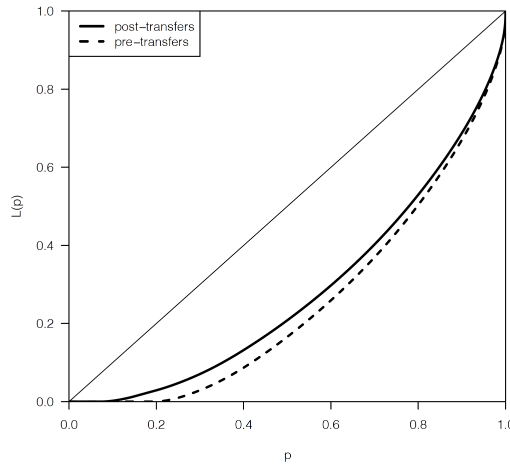


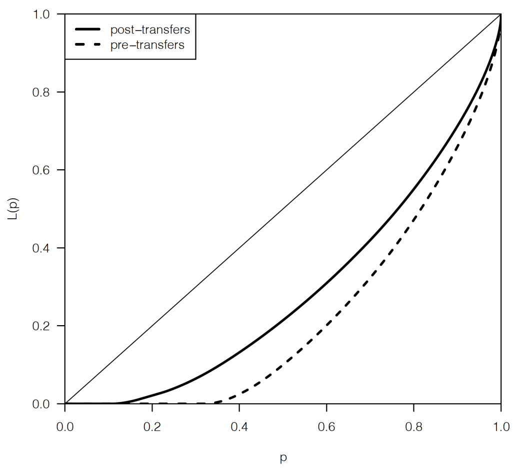
FIGURE A.1: DIFFERENCES-IN-DIFFERENCES IN PRE- AND POST-TRANSFER GINI WITH RESPECT TO FEBRUARY 2019 (RELATIVE TO 2018)



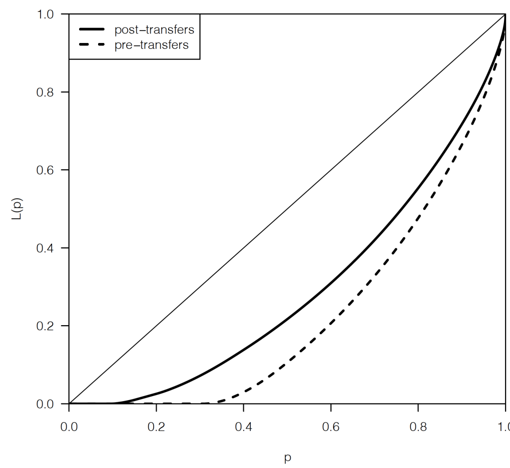
(a) February 2020



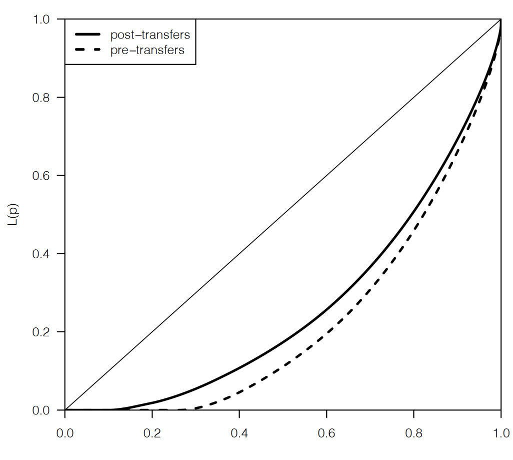
(b) March 2020



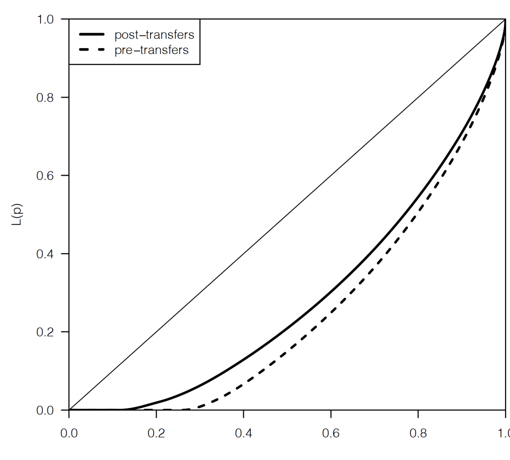
(c) April 2020



(d) May 2020



(e) June 2020



(f) July 2020

FIGURE A.2: LORENZ CURVES CORRESPONDING TO PRE-TRANSFER AND POST-TRANSFER MONTHLY WAGE DISTRIBUTIONS, FEBRUARY THROUGH JULY 2020

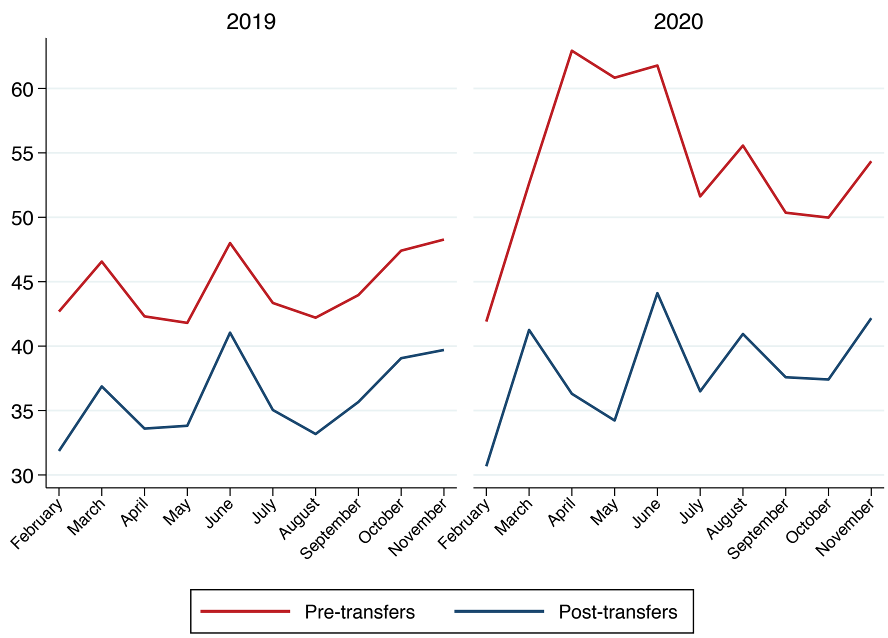


FIGURE A.3: EVOLUTION OF PRE- AND POST-TRANSFER THEIL INDEX BETWEEN FEBRUARY AND NOVEMBER 2019 AND FEBRUARY AND NOVEMBER 2020

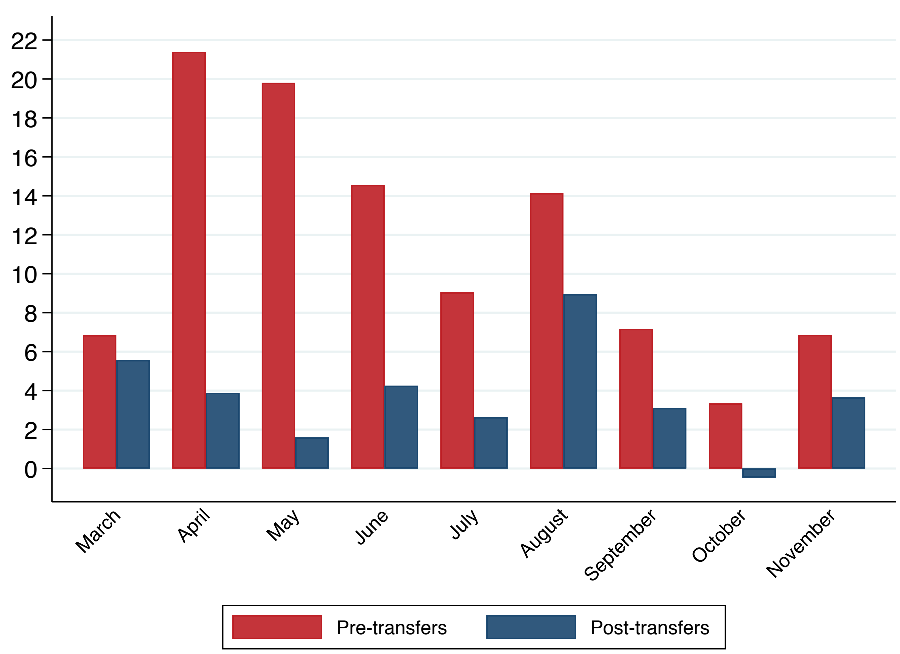


FIGURE A.4: DIFFERENCES-IN-DIFFERENCES IN PRE- AND POST-TRANSFER THEIL INDEX WITH RESPECT TO FEBRUARY 2020 (RELATIVE TO 2019)

DATA CONSTRUCTION, VARIABLE DEFINITIONS, AND DATA SOURCES

The data used to perform the study come from several administrative sources: the dataset including information of the elections of the parliament of Galicia for the years 2016 and 2020 comes from the online platform of “Xunta de Galicia" (the Regional Goverment of Galicia)[30]. Likewise, the dataset containing information of the elections of the parliament of the Basque Country comes from the online platform of “Gobierno Vasco" (the Security Department of the Government of the Basque Country)[31]. Below the list of the variables already included in each dataset and generated for the analysis.

XUNTA DE GALICIA AND GOBIERNO VASCO DATASETS

Both sources of data have the same structure and include the same variables for both the 2016 and 2020 elections of the parliament.

Municipality. The name of each of the municipalities in which the voting took place.

Census. The number, in absolute terms, of citizens officially registered with right to vote.

Voters. The number, in absolute and percentage terms, of citizens officially registered with right to vote who actually voted out of the total. It is the total number of valid ballots.

Abstention. The number, in absolute and percentage terms, of citizens officially registered with right to vote who actually did not voted out of the total.

Valid ballots. The number, in absolute and percentage terms, of valid ballots out of the total.

Null ballots. The number, in absolute and percentage terms, of null ballots out of the total.

Blank ballots. The number, in absolute and percentage terms, of blank ballots out of the total.

First political party. The percentage of votes out of the valid ballots earned by the political party that won in each municipality and its name.

Second political party. The percentage of votes out of the valid ballots earned by the political party that arrived second in each municipality and its name.

Third political party. The percentage of votes out of the valid ballots earned by the political party that arrived third in each municipality and its name.

Fourth political party. The percentage of votes out of the valid ballots earned by the political party that arrived fourth in each municipality and its name.

In order to perform this study, several variables have been created using those already included in the data.

Change in Census. The change in the number of citizens officially registered with right to vote (difference between census 2020 and census 2016 over census 2016).

[30] The link to the results for the 2016 elections: http://resultados2016.xunta.gal/ini11v.htm?lang=gl. The link to the results for the 2020 elections: https://resultados2020.xunta.gal/inicio.

[31] The link to the results for the 2016 and 2020 elections: https://www.euskadi.eus/elecciones/.

Change in Voters. The change in the number of citizens officially registered with right to vote who actually voted out of the total (difference between voter 2020 and voters 2016 over voters 2016).

Change in Abstention. The change in the number of citizens officially registered with right to vote who actually did not voted out of the total (difference between abstention 2020 and abstention 2016 over abstention 2016).

Change in Valid ballots. The change in the number of valid ballots out of the total (difference between valid ballots 2020 and valid ballots 2016 over valid ballots 2016).

Change in Null ballots. The change in the number of null ballots out of the total (difference between null ballots 2020 and null ballots 2016 over null ballots 2016).

Change in Blank ballots. The change in the number of blank ballots out of the total (difference between blank ballots 2020 and blank ballots 2016 over blank ballots 2016).

Change in ideology. The change in the political ideology of voters (difference between ideology 2020 and ideology 2016 over ideology 2016).

Change in votes received by a political party. The change in the votes earned by a political party out of the valid ballots (difference between votes received by a political party 2020 and votes received by a political party 2016 over votes received by a political party 2016). The list of the political parties for which this variable has been created includes: PP, BNG, PSdeG-PSOE, ESCO-RC-OV-M, PODEMOS-ESQUERDA, VOX, EN MAREA, Cs, DO, EAJ-PNV, EH BILDU, PSE-EE/PSOE, PP+Cs, PODEMOS-IU.

Change in Gini index pre and post. The change in the Gini index (difference between Gini pre and Gini post).

Ln of Population. The natural logarithm of the population number in each municipality.