

IZA DP No. 10179

Hours Worked in Europe and the US: New Data, New Answers

Alexander Bick
Bettina Brüggemann
Nicola Fuchs-Schündeln

August 2016

Hours Worked in Europe and the US: New Data, New Answers

Alexander Bick
Arizona State University

Bettina Brüggemann
McMaster University

Nicola Fuchs-Schündeln
*Goethe University Frankfurt,
CEPR, CESifo, CFS and IZA*

Discussion Paper No. 10179
August 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Hours Worked in Europe and the US: New Data, New Answers*

We use national labor force surveys from 1983 through 2011 to construct hours worked per person on the aggregate level and for different demographic groups for 18 European countries and the US. We find that Europeans work 19% fewer hours than US citizens. Differences in weeks worked and in the educational composition each account for one third to one half of this gap. Lower hours per person than in the US are in addition driven by lower weekly hours worked in Scandinavia and Western Europe, but by lower employment rates in Eastern and Southern Europe.

JEL Classification: E24, J21, J22

Keywords: labor supply, employment, hours worked, Europe-US hours gap, demographic structure

Corresponding author:

Nicola Fuchs-Schündeln
Goethe University Frankfurt
House of Finance
60323 Frankfurt
Germany
E-mail: fuchs@wiwi.uni-frankfurt.de

* This paper is based on material included in previous work of us which was circulated under the title "Labor Supply Along the Extensive and Intensive Margin: Cross-Country Facts and Time Trends by Gender". We thank Domenico Ferraro, Berthold Herrendorf, Bart Hobijn, Roozbeh Hosseini, Richard Rogerson, Luigi Pistaferri, Valerie Ramey, B Ravikumar, Todd Schoellman, seminar participants at Aarhus University, the Federal Reserve Bank of San Francisco, Free University Berlin, and conference participants at the 2012 and 2015 Annual Meetings of the Society for Economic Dynamics, the Federal Reserve Bank of Atlanta Workshop on "Household Production and Time Allocation", and the NBER Summer Institute Macro Perspectives Workshop 2014 and the NBER Summer Institute Conference on Research in Income and Wealth 2016 for helpful comments and suggestions. The authors gratefully acknowledge financial support from the Cluster of Excellence "Formation of Normative Orders" at Goethe University and the European Research Council under Starting Grant No. 262116. All errors are ours.

1 Introduction

An active recent literature has documented large differences in the levels and trends of aggregate hours worked per person across OECD countries, and specifically lower aggregate hours worked per person in Europe than in the US. This literature traces these lower hours in Europe back to, amongst others, labor income taxation (e.g., Prescott (2004), Rogerson (2006), Olovsson (2009), McDaniel (2011), to name a few), institutions (Alesina et al. (2005)), and social security systems (Erosa et al. (2012), Wallenius (2014), Alonso-Ortiz (2014)). One basic step to understand the causes of the large differences in labor supply is to analyze, first, whether these differences exist for all margins of labor supply, and, second, how much they are driven by different characteristics of the population. Are fewer Europeans working, or do they work fewer hours per work week, or do they simply enjoy more vacation days? Are the aggregate differences driven by different sectoral compositions across countries, different age compositions, or different educational compositions? These questions have been raised early on in Rogerson (2006) as an important avenue for advancing the research agenda on understanding differences in hours worked across countries, but could only be answered to a limited extent based on existing data readily available to researchers.

This paper documents differences in hours worked per person between the US and 18 European countries based on “new data” compiled from national labor force surveys.¹ Constructing internationally comparable hours estimates based on labor force surveys allows us to decompose differences in hours per person into the different components of labor supply, namely employment rates, weeks worked per year, and weekly hours worked per employed, and to analyze to which extent these differences are driven by different demographic and sectoral compositions. This paper thus provides a careful statistical decomposition of the hours differences between Europe and the US. The resulting facts open up new avenues for research. We use three different labor force surveys, namely the European Labor Force Survey, the US Current Population Survey, and the German Microcensus, covering the time period 1983 to 2011. First, we describe how we measure employment rates and hours worked per employed in a consistent way that leads to comparable hours worked per person estimates across countries. Next, we document that differences in *aggregate* hours worked per person between Europe and the US are substantially larger in our data

¹Ohanian and Raffo (2012) also add new evidence on the measurement of hours worked across OECD countries by constructing hours measures at the quarterly frequency. Bick et al. (2016) measure hours worked for 81 countries in the early 2000s focusing on the differences between low-, middle- and high-income countries rather than the differences among the high-income countries as we do here. Burda et al. (2008), Burda et al. (2013), Fang and McDaniel (2016), and Bridgman et al. (2016) construct cross-country hours worked measures using time-use surveys. The advantage of time use surveys is the precise measurement of time spent over an entire work day or weekend day. The disadvantage – particularly in comparison to labor force surveys – is the much lower frequency and much smaller sample size.

than in the National Income and Product Accounts (NIPA) data. We analyze the reasons for this discrepancy, and show implications for the measurement of labor productivity differences and for the performance of macroeconomic models that explain the cross-country differences through taxation (Prescott (2004)). We then address the core question of this paper and show that differences in weeks worked over the year account for one third to one half of the Europe-US hours gap, and a further one third to one half are explained by differences in the educational composition across countries. Lower hours per person in Scandinavia and Western Europe than in the US are in addition driven by lower weekly hours worked, whereas lower hours per person in Eastern and Southern Europe are in addition driven by lower employment rates. Hence, the “new answers” we provide refer both to different evidence on *aggregate* hours worked per person than the one provided by NIPA data, and more importantly to the *disaggregate* evidence that cannot be obtained from aggregate NIPA data. We make the data available on our webpage to the research community.

The measurement of hours worked per person, employment rates, and hours worked per employed from labor force surveys is not trivial. First, the main difference between the several surveys, at a given point in time across countries or for a given country over time, regards the sampling of reference weeks over the year. The reference weeks range from one specific week only, to a quarter, to the sampling of all 52 weeks of the year. We document that this is not a major problem for the consistent measurement of employment rates, but that without any further adjustment hours worked per employed, which exhibit a fair amount of seasonality, are not measured consistently. Second, we show that even if all 52 weeks of the year are covered, vacation days are severely underestimated in the labor force surveys, partly because not all weeks are sampled with equal probability. We implement a measurement strategy of hours worked per employed that is able to deal with both these issues by adjusting for vacation days, i.e. annual leave and public holidays, obtained from external data sources. It turns out that this also drastically decreases the problem of seasonality in hours worked per employed, which is mostly driven by vacation days.

As is common practice when working with micro data sets, see e.g. the 2010 Review of Economic Dynamics special issue on “Cross Sectional Facts for Macroeconomists” (Krueger et al. (2010)), we compare our (LFS) aggregate statistics on hours worked per person with those from the National Income and Product Accounts (NIPA), which are provided by the OECD “National Accounts Database” and the Conference Board’s “Total Economy Database” (TED). For most countries and years, the numbers by the OECD and TED coincide, but the time-series available from TED is longer than the one from the OECD. While time trends for hours are similar in LFS and NIPA data, cross-country estimates show substantial level differences between both data sets. LFS data indicate that Europeans work 19% fewer hours than Americans in the recent pre-crisis

years 2005-2007, while NIPA data only imply a difference of 7%. We proceed by investigating the sources of these large differences between LFS and NIPA data. We find that these mainly arise from the fact that the national statistical agencies use vastly different kinds of data sets and methods to construct hours estimates. As a consequence, the OECD itself cautions that their data are not suitable for cross-country comparisons. In contrast, we use always the same kind of data source, namely labor force surveys, to construct hours worked estimates, and apply exactly the same algorithm to all data sets, thereby maximizing the comparability across countries. A further advantage of the LFS data are that they are not subject to major revisions like the NIPA data.

We find that it matters for macroeconomic analyses whether one relies on LFS or NIPA data. First, we analyze implications for labor productivity differences between Europe and the US. Based on LFS data, Europe looks with 86% of the US level closer to the US in terms of labor productivity than with 78% based on NIPA data. This is mostly driven by the narrowing of the rather small labor productivity gap to the US for Western Europe by three fourths, and the large gap for Southern Europe by one third in LFS data compared to NIPA data, while differences in productivity measurement between the different data sets are smaller for Eastern Europe and Scandinavia. Second, we show, relying on the work by [Prescott \(2004\)](#), that through the lens of a neo-classical growth model average tax rate differences and differences in the consumption/output ratios can account for 52 percent of the differences in hours worked between Europe and the US if these are measured based on LFS data, as opposed to more than 100 percent if they are measured by NIPA data. Thus, the explanatory power of taxes for the Europe-US hours gap is smaller when relying on LFS data than when relying on NIPA data.

After documenting these differences in *aggregate* hours worked per person between LFS data and NIPA data, we turn to the core question of the paper and provide a decomposition of hours worked differences between Europe and the US for the recent pre-crisis years 2005 to 2007, focusing on individuals aged 15 to 64. We first document that employment rates and weekly hours worked are strongly negatively correlated across countries. By contrast, the number of work weeks during the year is uncorrelated with both hours worked per week and employment rates. All European regions (Scandinavia, Western Europe, Eastern Europe, and Southern Europe) have on average similar hours that are 16 to 19% lower than in the US. While weeks worked are uniformly substantially lower in Europe than in the US, in Southern and Eastern Europe lower hours are additionally driven by lower employment rates, but in Western Europe and Scandinavia by fewer hours worked per work week. The other component of labor supply always points in the opposite direction: individuals in Southern and Eastern Europe work longer weekly hours than US citizens, and individuals in Scandinavia and to some extent Western Europe are more likely to be employed

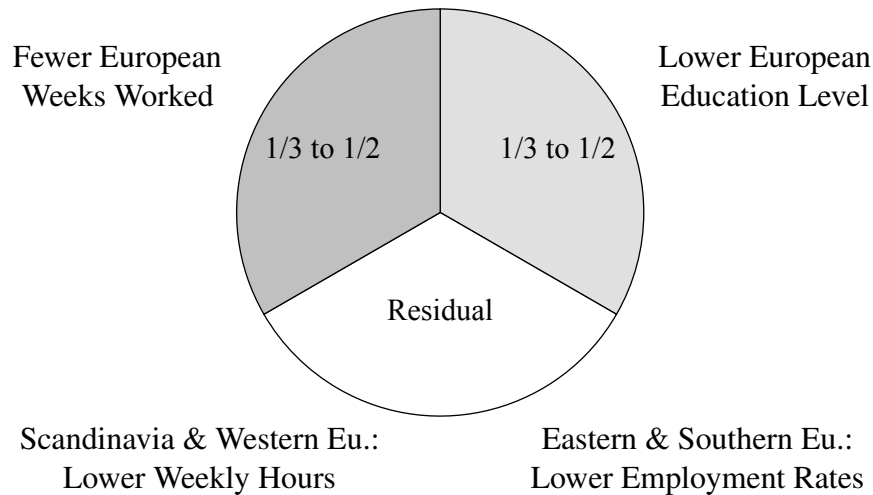
than US citizens.

Exploiting the rich micro-structure of our data, we then ask which role differences in the demographic and sectoral composition between the US and different European countries play in accounting for these differences. For instance, if older people worked on average fewer hours than younger people in all countries, and European countries had an on average older population, this could partly account for the lower hours in Europe. We investigate differences in the composition by gender, age, education, and sectors. Any of these factors can play a role in accounting for hours differences across countries only if both the composition across countries is different *and* different groups exhibit different labor supply behavior within a country. Since the age and gender compositions across countries are in fact very similar, these two factors turn out not to play a role. By contrast, the educational compositions differ vastly across countries. In general, Europe has a higher share of low- or medium-educated individuals and a lower share of high-educated individuals than the US, especially in Eastern and Southern Europe. While weekly hours worked are similar for the three educational groups, employment rates increase substantially by education in all countries. Therefore, we find that differences in the educational composition between the US and the European countries are very important for understanding the hours differences via their impact on employment rather than on weekly hours worked. Differences in the sectoral structure turn out to matter only minimally in the statistical decomposition because differences in weekly hours per worker between sectors are small, though larger than between education groups. Since we cannot assign non-working individuals to a sector, our analysis does not capture the connection between sectoral shifts and employment rates, see [Rogerson \(2008\)](#).

Figure 1 summarizes our results: weeks worked account for between one third and one half of the Europe-US hours differences, and the educational composition for another one third to one half. However, because of the strong negative cross-country correlation between employment rates and weekly hours worked, this does not imply that these two components alone account for almost all of the differences. Lower weekly hours worked per person account for 75 and roughly 40 percent of the lower hours in Scandinavia and Western Europe, respectively, than in the US, with employment rates in Scandinavia predicting significantly higher hours there than in the US. By contrast, for Eastern Europe and Southern Europe (excluding Portugal), lower employment rates account for roughly one half and one quarter, respectively, of the lower hours than in the US, with weekly hours worked in Eastern Europe predicting higher hours there than in the US. Lower hours in Portugal are driven exclusively by the lower number of weeks worked and the educational composition.

What do our results imply for the study of hours worked differences between Europe and the

Figure 1: Decomposing the Europe-US Hours Gap



US? We arrive at three main conclusions. First, differences in vacation days play a quantitatively significant role in accounting for hours worked per person differences across countries. Should we think of the number of vacation days as individual choice variables, or are they rather pre-determined from an individual perspective? We document an essentially zero correlation between weeks worked per year and either employment rates or weekly hours worked per employed, which could be indicating the latter. In that case, the question arises which institutional factors drive international differences in vacation weeks. Moreover, to analyze the driving forces (e.g. taxation) of the hours worked decision from an individual perspective, one might want to use a model that takes vacation days as exogenous, similarly to e.g. the model used in [Kaplan \(2012\)](#). In this model, involuntary unemployment (in our case vacation weeks) determines the number of weeks worked per year, and agents only choose hours worked per work week optimally. Assuming separability in the utility function between weeks worked and hours worked per week could lead to a zero correlation between weeks worked per year and either employment rates or weekly hours worked per employed. On the other hand, models with positive externalities of sharing leisure time together or with a (negative) signalling effect of taking annual leave could lead to multiple equilibria with either high or low vacation days as equilibrium outcomes, and factors like taxes could serve as coordination devices on an equilibrium.

Second, lower hours worked in Scandinavia and Western Europe are driven by lower weekly hours per employed, and in Eastern and Southern Europe by lower employment rates. These facts are robust to controlling for differences in the demographic and sectoral composition across countries. What can explain these differences? Generating the negative correlation between em-

ployment rates and weekly hours worked in a quantitative macro model is not trivial, since most factors driving the individual labor supply decision tend to influence both components in the same direction. Analyzing the sources of the negative correlation is thus a fascinating new research agenda. In [Bick and Fuchs-Schündeln \(2016\)](#), we present some suggestive evidence that international differences in part-time regulation could drive the correlation.

Third, the educational composition plays a significant role in accounting for employment rate differences across countries. This calls for models that endogenize human capital accumulation decisions in order to understand why college graduation rates are much higher in some countries than in others. What are the policies that drive different educational compositions across countries? For example, [Guvenen et al. \(2014\)](#), while not distinguishing between employment and hours per employed, show that progressive taxation distorts the incentives to invest in human capital, and that these tax differences can explain a sizable part of the differences in the Europe-US hours per person gap for men. A further interesting new question that arises is why employment rates uniformly increase by education, whereas hours per employed do not vary much by education.

The remainder of the paper is structured as follows: Section 2 describes the micro data sets and explains how we measure hours worked per person, employment rates, and hours worked per employed. Section 3 compares aggregate hours worked from LFS data to those reported by NIPA. Section 4 reports the Europe-US hours difference for LFS and NIPA data, and analyzes implications of using LFS data instead of NIPA data for measuring productivity differences, as well as for models that postulate taxation as the major driving force of international hours differences. Section 5 then comes to the core analysis and decomposes aggregate hours worked differences between Europe and the US in its different components, and analyzes the role of different demographic subgroups. Finally, Section 6 concludes.

2 Data and Measurement

In this section, we first introduce the three labor force surveys, followed by the key questions from these surveys that we rely on to measure hours worked per person. Section 2.3 then presents the two main challenges that we face in measuring hours worked consistently across countries, namely the sampling of reference weeks and the underreporting of vacation days. Only the former affects the measurement of the employment rate (by definition), but we show that this is not a major problem. In contrast, we document that both of these challenges play a potentially large role for the measurement of hours worked per employed. Section 2.4 explains how we construct hours worked per employed to overcome these two issues.

At the start of this section, let us define the main variables used throughout the paper. H refers to average annual hours worked per person, and H^E are average annual hours worked per employed. e is the employment rate, h are hours worked per employed in a non-vacation week (a term which we use interchangeably with weekly hours worked), and w is the average number of non-vacation weeks in a year (which we use interchangeably with weeks worked), such that

$$H^E = h \times w,$$

and

$$H = e \times H^E = e \times h \times w.$$

We call the employment rate e and hours worked per employed H^E the two margins of labor supply, as usual in the literature, while we talk about components when we further divide hours worked per employed into weekly hours in a non-vacation week and weeks worked.

2.1 Data Sets

We construct hours worked from three different micro data sets, namely the European Labor Force Survey, the Current Population Survey, and the German Microcensus.

European Labor Force Survey The European Labor Force Survey (ELFS) is a collection of annual labor force surveys from different European countries. We use the yearly surveys, since the quarterly ones do not provide information on education. The ELFS covers Belgium, Denmark, France, Greece, Italy, Ireland, the Netherlands, and the UK from 1983 on, Portugal and Spain starting in 1986, Austria, Norway, and Sweden starting in 1995, Hungary and Switzerland starting in 1996, the Czech Republic and Poland starting in 1997, and Germany starting in 2002.² The sample size of the ELFS varies across countries and also within a country over time, but is always of considerable magnitude. The minimum annual sample size in the original data we use is 15,400 for Denmark, a country with roughly 5.5 million inhabitants.

Current Population Survey For the US, we use the Current Population Survey (CPS), which is a monthly survey of around 60,000 households. Specifically, we work with the CPS Merged Outgoing Rotation Groups data provided by the National Bureau of Economic Research (see [http:](http://)

²For the Netherlands, we have information from 1983, 1985, and annually from 1987 on. The ELFS also covers Finland from 1995 on. However, the Finnish data have large numbers of missing observations for several years, which implies that we could only use data from 1997 to 2002 for our analysis. The ELFS covers more transition countries, which we also exclude from the analysis because of a shorter time series and a lack of information on vacation days.

[//www.nber.org/data/morg.html](http://www.nber.org/data/morg.html)). This data set includes only those interviews in which the households are asked about actual and usual hours worked, namely the fourth and eighth interview of each household. The data cover around 300,000 individuals per year.³

German Microcensus Until Germany enters the European Labor Force Survey in 2002, we use data from the German Microcensus. The German Microcensus covers a one percent random sample of the population of Germany and is an administrative survey. We use the scientific use files, which are a 70 percent random subsample of the original sample. This leaves us with a sample size of between 400,000 and 500,000 individuals per year. The scientific use files are available biannually from 1985 on, and annually from 1995 on. East Germans are included in the sample from 1991 onwards.

2.2 Key Survey Questions and Sample Selection

To measure employment rates and hours worked per employed person, we rely on five survey questions. We construct the employment rate based on the self-reported employment status in the reference week, which is usually the week prior to the interview. Employed individuals include, in addition to employees, the self-employed and unpaid family workers. Moreover, being employed does not require that the individual works positive hours in the reference week. To estimate hours worked for employed individuals, we rely on questions about actual hours worked in the main job in the reference week, actual hours worked in additional jobs in the reference week, usual hours worked in the main job in a working week, and reasons for having worked more or less hours than usual in the reference week. Web Appendix Section W.1.1 discusses how we deal with missing values for these questions, and explains why we have to drop three country-year pairs (1983 for Denmark, 2001 for the UK, and 2005 for Spain). We include in the final sample all observations of individuals aged 15 to 64. Web Appendix Section W.1.1 also provides an overview of the annual final sample size per country, which ranges from 10,000 to 450,000 observations, with an average of 111,000 observations.

2.3 Two Challenges

We face two key challenges when measuring hours worked per person across countries: first, the sampling of reference weeks differs across countries and within European countries over time,

³While it is well known that wages are imputed for many individuals in the CPS, this is not the case for hours, of which less than 1% are imputed.

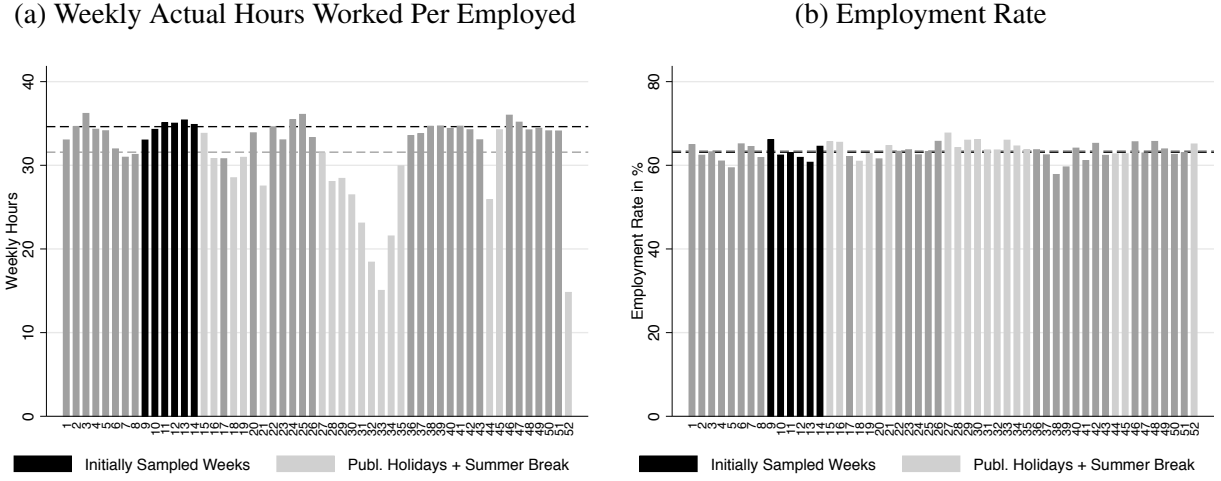
potentially inhibiting the comparability of measured hours. Secondly, we find that vacation days are systematically underreported in labor force surveys.

2.3.1 Sampling of Reference Weeks

The main difference between the several surveys, at a given point in time across all countries or for a given country over time, regards the sampling of reference weeks over the year. The CPS samples all 12 months of the year, but uses as a reference week always the week including the 12th day of a month. Therefore, most major US public holidays are not captured by the CPS (e.g. Memorial Day, 4th of July, Thanksgiving). The reference weeks in the national labor force surveys of the European countries initially fell only into country-specific periods ranging from one single reference week to the sampling of half a year. From 2005 onwards, each week of a year is sampled in the European Labor Force Survey. Eurostat, in its efforts to harmonize the different surveys as much as possible, treated the changes in reference weeks in a two-step procedure. First, if the change to sampling all weeks occurred in a country before 2005, the ELFS micro data reflect this by changing from sampling only single weeks to sampling the second quarter of the calendar year (April to June) from then on until 2004. Exceptions to this rule are detailed in Web Appendix [W.1.2](#). In a second step from 2005 onwards, when the majority of countries included in the ELFS had switched to continuous surveying, Eurostat makes all 52 weeks of the year available. The only exceptions to this second step rule are the UK (continuous surveying from 2008 on), Ireland (from 2009 on), and Switzerland (from 2010 on).

The differences in the sampling of reference weeks raise the question in how far estimates of the employment rate and hours worked per employed are comparable across countries in a given year, and within a country over time. Since there was no change for the US, the latter question is only relevant for the European countries. To give a concrete example, Figure 2 shows weekly average actual hours worked per employed in the left panel, and weekly employment rates in the right panel, for France in 2006 by reference week. The black weeks are the weeks that were initially sampled in France until 2002 (weeks 9 to 14). The light gray weeks are weeks that include a public holiday, plus the summer vacation weeks. As one can see in the left panel of Figure 2, weekly hours worked vary substantially over the 52 weeks of the year, and are on average significantly lower in vacation weeks. The sampling of reference weeks thus matters for the aggregate: the grey horizontal bar indicates average weekly hours worked over all 52 weeks, while the black horizontal bar indicates average weekly hours worked over the initially sampled black weeks: the difference amounts to 3.1 weekly hours. By contrast, the right panel of Figure 2 shows that the employment rate is roughly constant over all 52 weeks of the year.

Figure 2: Avg. Weekly Actual Hours Worked Per Employed and Employment Rate (France, 2006)



To analyze the extent of seasonality in measures of the employment rate and hours worked per employed more generally, we exploit (as in the example above for France) the fact that the ELFS from 2005 onwards samples all weeks of the year for most countries, and also states the reference week for each interviewed person. Thus, for these years, we analyze how employment rates and hours worked per employed differ if they are constructed using information from only the initially sampled weeks, i.e. only those weeks which served as country-specific reference weeks prior to the change to continuous surveying (e.g. the black weeks in Figure 2 for France), and using all weeks of the year as reference weeks. We use all available country-year pairs from 2005 to 2011, which are in total 113 observations (for 15 of our 18 European countries we can do the comparison in all years starting 2005, for the UK, Ireland, and Switzerland only starting later).⁴

Hours Worked per Employed We calculate a measure of “raw” annual hours worked per employed as the sum of actual weekly hours worked of all employed individuals multiplied by 52 weeks in a year:

$$H^{E,raw} = 52 \times \frac{1}{N^e} \sum_{i=1}^N \varepsilon_i \eta_i^{raw},$$

where ε_i is the self-reported employment status of individual i , which takes the value 1 for anyone reporting to be employed (including self-employment or being an unpaid family employee), and 0 otherwise; η_i^{raw} are the actual hours worked in the reference week in all jobs of individual i ; N is the

⁴In Web Appendix W.1.2, we show corresponding evidence when the second quarter is used as reference period, as done for the interim period in the European countries. Differences to using all weeks of the year as reference period are somewhat smaller than under the initially sampled weeks, but still substantial.

total number of individuals; and N^e is the number of employed individuals, i.e. $N^e = \sum_{i=1}^N \varepsilon_i$. All measures are calculated using the individual survey weights, which we omit for ease of exposition.

Table 1 shows in the first row the absolute percentage deviation of annual hours worked per employed (aged 15 to 64) when using only the initially sampled reference weeks from using all weeks of a year, which amounts on average to 3.5%. At the 90th percentile, the deviation already reaches 7.0%, with a maximum deviation of 12.6% in Sweden for the year 2008. Thus, the sampling of the reference weeks matters substantially for the measurement of hours worked per employed. With only few exceptions, the hours when only specific survey weeks are sampled are larger than when all weeks of the year are sampled.⁵

Employment Rate The employment rate (e) is simply given by the number of employed individuals divided by all individuals in the sample

$$e = \frac{N^e}{N}. \quad (1)$$

Equation (1) is the employment to population ratio for the population 15-64, but we henceforth refer to it as the employment rate.⁶ The second row of Table 1 shows that the employment rate exhibits significantly less seasonality than the raw hours worked per employed measure. The average absolute percentage point deviation between constructing the employment rate only based on specific survey weeks relative to the entire year amounts to only 0.7 percentage points. At the 90th percentiles, the difference is still relatively small with 1.3 percentage points. Moreover, employment rates are not systematically higher (or lower) if only specific weeks are sampled than when the entire year is sampled. There are only a few observations that are worrisome with deviations over 5 percentage points. The five country-year observations above the 95th percentile come from Germany (four times) and Belgium. Both countries used only one specific week as reference week (Germany until 2004, Belgium until 1998), suggesting that the corresponding time-series of the employment rates have to be used with some caution.

⁵This is the case for all countries in all but at most one year. The only exception are the Netherlands, where hours are lower in the initially sampled weeks, with the exception of two years.

⁶Web Appendix Section W.1.3 reports two alternative measures of the employment rate. The first alternative measure relies on defining employment based on usually working positive hours. This leaves the employment rate virtually unchanged. The second alternative measure involves a different definition of employment for women on maternity leave. This has only a modest impact on female employment rates, and is far too small to drive international differences in female employment rates. Note that by construction, hours worked per person remain always the same no matter what definition is used for the employment rate, because an increase (decrease) in the alternative employment rate relative to our baseline definition is always offset by a decrease (increase) in the corresponding measure of hours worked per employed.

Table 1: Absolute %-Deviations* of $H^{E,raw}$, e , and H^E Using Only Initially Sampled Weeks from Using All 52 Reference Weeks

	Mean	75th	90th	95th	99th	Max
$H^{E,raw}$	3.5	4.7	7.0	7.9	11.4	12.6
e	0.7	0.8	1.3	1.9	5.3	5.9
H^E	0.8	1.1	1.5	2.1	3.7	5.0

Note: For $H^{E,raw}$ and H^E we report deviations in percentages, whereas for e we use percentage points.

2.3.2 Underreporting of Vacation Weeks

For most European countries, all weeks of the year are sampled from 2005 onwards. If each week would be sampled with equal probability and cover a representative sample of the population week by week (rather than only over the entire year), then average vacation weeks should be captured accurately in the labor force surveys. Calculating the average number of vacation weeks from the micro data yields an average of 3.3 “self-reported” vacation weeks, where we define vacation as the sum of public holidays and annual leave.⁷ Given that public holidays alone in all countries sum up to 1.5 to 2.5 weeks, these self-reported weeks of annual leave and public holidays seem implausibly low. In fact, based on external data sources, the average weeks of vacation over the same time period for our sample of European countries amount to 6.8 weeks per year.⁸

To get a better understanding of the sources of this discrepancy, we investigate in more detail the case of Germany, which features the largest difference among all countries, amounting to 5.5 weeks. While we provide all details in Web Appendix Section W.2.2, our results of this analysis can be summarized as follows. First, using the German Socio-Economic Panel, Schnitzlein (2011) reports that on average 3 of the average 31 entitled days of annual leave per year go unused. Thus, the underusage of entitled leave can explain only a very small portion of the discrepancy of 5.5 weeks between self-reports and official vacation days. Second, not all weeks are sampled with the same probability in the Microcensus, see Figure W.1 in Web Appendix Section W.2.2. For example, the reference week which contains Christmas day is on average sampled with a probability of 1.3%, significantly below the 1.9% that would be implied by equal sampling. Third, the German Statistical Office also has some evidence that respondents might dislike using a vacation week as a

⁷Table W.10 in Web Appendix W.2.2 reports the results for each country as well as the description of how we construct the self-reported vacation weeks.

⁸Web Appendix W.2.1 lists the external data sources and explains further assumptions in the construction of the number of vacation weeks. It also shows that the time series variation in vacation weeks within each country is small, with the notable exception of Denmark. For the US, the discrepancy between self-reported and external vacation weeks is somewhat smaller but still substantial, with 1.4 vs. 3.5 weeks. We expect a discrepancy for the US, since weeks with major public holidays are not sampled.

reference week, either because they are too busy the first week after a vacation to fill out the questionnaire, or because they perceive it as “inappropriate” to use a vacation week as reference week when in fact they are generally hard working. Summarizing, at least for Germany there exists evidence of underreporting of days of annual leave and public holidays in the Microcensus even after the introduction of continuous sampling over the entire year. It seems at least not implausible that these factors can explain the discrepancies between the self-reported and external vacation weeks for the remaining countries as well. Therefore, it is important to adjust hours worked per employed for vacation using external data in the European countries even when analyzing the recent cross-section from 2005 onwards.

2.4 Measurement of Hours Worked per Employed

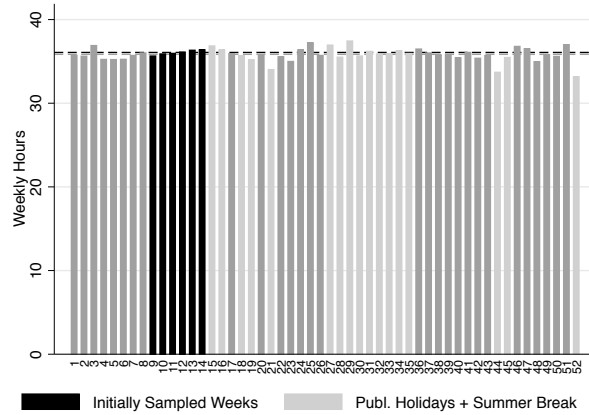
As documented so far, the raw measure of hours worked per employed suffers from two weaknesses: first, seasonality in hours worked per employed impedes the comparability of hours worked across countries and within European countries over time because of differences in the sampled reference weeks; secondly, vacation weeks are underreported in the labor force surveys even if all weeks of a year are sampled. We therefore apply the following adjustment to construct a measure of hours worked per employed that overcomes both weaknesses: we first obtain individual hours worked in a *non-vacation* week, i.e. a typical work week without a reduction in work time because of a public holiday or annual leave, and then require *consistency* of annual leave and public holidays in the micro data with the country-wide average. This closely follows the procedure in [Pilat \(2003\)](#).

To calculate hours worked in a *non-vacation* week η_i for each employed individual, we use as a baseline actual hours worked in all jobs in the reference week. However, if a respondent indicates that he/she worked less hours than usual in the main job in the reference week, and states as the main reason for doing so public holidays or annual leave, we replace weekly hours by usual hours in the main job plus actual hours worked in all additional jobs.⁹ Thus,

$$\eta_i = \begin{cases} \text{usual hours in main job} & \text{if actual hours in main job} < \text{usual hours in main job} \\ + \text{ actual hours in all additional jobs} & \text{because of annual leave or public holiday} \\ \text{actual hours in all jobs } \eta_i^{raw} & \text{otherwise.} \end{cases}$$

⁹Note that respondents can only indicate one reason for working different hours than usual, and for additional jobs we do not have information on usual hours. In Web Appendix [W.1.4](#) we discuss some differences between the CPS and ELFS questionnaire regarding the construction of our hours measure, which however have virtually no impact on the statistics presented in the paper.

Figure 3: Average Hours Worked in a Non-Vacation Week h (France, 2006)



We refer to these hours η_i as “hours worked in a non-vacation week”. Averaging over our population of interest yields mean weekly hours worked in a non-vacation week h , i.e.

$$h = \frac{1}{N^e} \sum_{i=1}^N \varepsilon_i \eta_i. \quad (2)$$

If people work less hours than usual for other reasons than vacations, e.g. because of sickness, or more hours than usual because of overtime, this is captured by this measure. Figure 3 shows this measure of hours worked in a non-vacation week for France in 2006. It does not show any seasonality anymore, in contrast to the raw measure of average weekly hours worked in the left panel of Figure 2.

In order to establish consistency of annual leave and public holidays in the micro data with the country-wide average from external data sources, we multiply our mean weekly hours worked in a non-vacation week h with the number of non-vacation weeks w . We obtain our estimate of w by subtracting the country-wide average weeks of annual leave and public holidays reported in external data sources from the 52 weeks of a year, i.e.

$$w = 52 - \frac{\text{country-wide average days of annual leave \& public holidays}}{5}. \quad (3)$$

The disadvantage of this procedure is that we cannot account for heterogeneity across the population in terms of days of annual leave and public holidays. In Web Appendix Section W.2.5 we allow for heterogeneous vacation weeks. Specifically, we assume that the heterogeneity in self-reported vacation weeks reflects the true degree of heterogeneity, but we still impose that the average num-

ber of vacation weeks corresponds to the one from external sources. The implied differences to the case of homogeneous vacation weeks are small.

The third row of Table 1 shows the deviation of the annual hours worked per employed measure ($H^E = h \times w$) if only the initially sampled survey weeks are used as reference weeks, compared to all weeks of the year. While for the “raw” measure of hours worked per employed in the first row the mean absolute percentage deviation amounts to 3.5%, the mean absolute percentage deviation of our measure of hours worked per employed is with 0.8% substantially smaller. Thus, externally adjusting for vacation is successful in insuring comparability of the data over time. Only at the top end of the distribution we still see a non-negligible effect of the sampling of the reference weeks. The three countries with deviations in the top five percent are Belgium, Norway, and Poland.

Finally, while we document the need to adjust for hours lost due to public holidays and annual leave, the same could in principal apply to working less hours than usual because of sick leave or other reasons, and also to working more hours than usual (overtime and additional jobs). There are two dimensions to this issue. First, in how far do such differences between usual and actual hours vary with the set of reference weeks available? In Web Appendix Section W.3, we provide evidence that none of the other reasons for working more or less time than usual vary with the set of reference weeks, thus seasonality is not important for them. Second, is there systematic under- or overreporting of these categories? We collect some suggestive evidence for sick days for a subset of countries and years. Comparing sick days from external data sources to the reported sick days in our data, we find that the self-reported number of sick days are smaller than sick days from external data sources, a similar finding as the one for vacation days. However, the discrepancy is on average smaller than for vacation days and public holidays, with on average 1.2 weeks in Europe and 0.3 weeks in the US, see Web Appendix Section W.3.

3 Comparing LFS Hours with NIPA Hours

We use data from labor force surveys to analyze the role of different margins and the demographic composition for international hours worked differences. This raises the natural question how the aggregate hours constructed from labor force surveys (henceforth referred to as LFS hours) compare with hours in the National Income and Product Accounts (NIPA). In this section, we show that differences between aggregate hours in LFS and NIPA are large, and argue that they are not driven by conceptual differences between both, but rather by the use of vastly different kinds of data sources and methods in the construction of NIPA data.

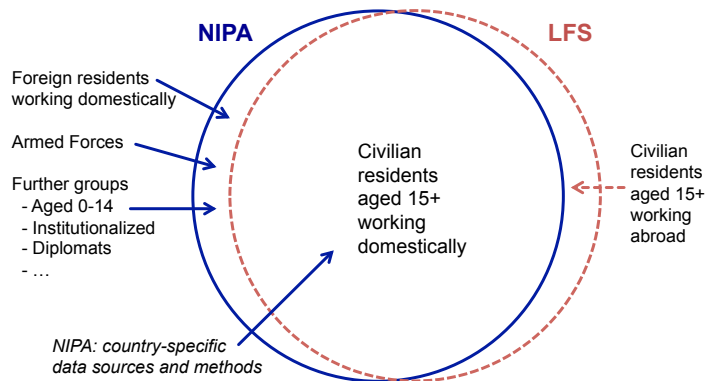
We rely on two commonly used sources for international NIPA data. The OECD gathers in

its “National Accounts Database” (stats.oecd.org) total hours worked and total employment at an annual frequency provided by the national statistical agencies. The measurement of both variables follows the most recent System of National Accounts (SNA) guidelines, see Chapter 19 in [European Commission et al. \(2009\)](#) and Chapter 11 in [Eurostat \(2013\)](#). Secondly, the Conference Board provides total hours worked and total employment in its “Total Economy Database” (TED) (<http://www.conference-board.org/data/economydatabase>). Both OECD and TED data have been widely used in the literature studying driving forces and implications of aggregate hours differences across countries. Since for the majority of countries and years in our sample the data from the OECD and the TED are exactly the same, we focus on the comparison of LFS data to NIPA data obtained from the OECD’s “National Accounts Database”. Note that the five main papers explaining hours worked differences between Europe and the US with quantitative macro models ([Prescott \(2004\)](#), [Rogerson \(2006\)](#), [Ohanian et al. \(2008\)](#), [McDaniel \(2011\)](#), and [Ragan \(2013\)](#)) all use slightly different approaches to calculate hours worked per person.¹⁰ With the exception of [McDaniel \(2011\)](#), these papers multiply a measure of total hours worked per employed, either from the OECD or from the TED, with variants of the civilian employment rate based on OECD data (see Web Appendix Section [W.4.4](#) and [Bick et al. \(2016\)](#) for details). We opt to work with NIPA data, since employment then refers to total employment in both hours worked per employed and the employment rate, thus canceling out when calculating hours worked per person.

The main conceptual difference between LFS and NIPA estimates of employment and hours worked relates to the covered population. LFS data cover civilian, non-institutionalized residents of a country aged 15 and older. NIPA data deviate from this in two dimensions. [Figure 4](#) illustrates these differences in the coverage. First, employment and hours estimates in NIPA are based on the domestic concept and constructed such that the labor inputs are consistent with the measurement of gross domestic output (GDP), while LFS follow the national concept based on residence of the worker. Thus, NIPA exclude employment and hours worked of civilian, non-institutionalized residents who work in a foreign country (e.g., from the German perspective a resident of Germany working in Switzerland). On the other hand, NIPA data include domestic employment and hours worked even if they originate from individuals residing in a foreign country (e.g., from the German perspective also hours worked by residents of Switzerland working in Germany). In addition, NIPA data do not restrict the covered population to the non-institutionalized civilians aged 15 and

¹⁰The OECD also provides information on average annual hours actually worked per worker in its “Labour Database” under the category “Labour Force Statistics”. For most countries in our sample and for most years, this estimate is equal to total hours worked divided by total employment from the “National Account Database”. Moreover, the OECD provides under the category “Labour Force Statistics” two different measures of civilian employment, one under the “Annual Labor Force Statistics”, another one under “LFS by sex and age”, which generally do not coincide, although the differences are usually small.

Figure 4: Population Coverage in NIPA and LFS



older, but include employment and hours worked from all individuals contributing to GDP: non-civilians (members of the armed forces) and further groups, including individuals younger than 15, institutionalized individuals (e.g., prisoners), and others as e.g., diplomats (e.g., from the German perspective the German ambassador in the US) or students abroad. Secondly, besides these differences related to the covered population, LFS and NIPA data on hours worked and employment can differ for the same covered population (i.e. civilian, non-institutionalized residents of a country aged 15 and older that do not work in a foreign country; the intersection in Figure 4). The national statistical agencies use very different country-specific kinds of data sources and methods in the construction of NIPA data (see Web Appendix W.4.1), which makes the cross-country comparison of hours worked difficult, despite the efforts to harmonize measurement through the Systems of National Accounts, see also Fleck (2009). These data sources comprise administrative data, social security data, employer surveys, labor force surveys, census data, etc. In fact, the OECD remarks on its website that “The [hours worked] data are intended for comparisons of trends over time; they are unsuitable for comparisons of the level of average annual hours of work for a given year, because of differences in their sources” and recommends using employment rates based on labor force surveys for cross-country comparisons: “National Labor Force Surveys are the best way to capture unemployment and employment according to the ILO guidelines that define the criteria for a person to be considered as unemployed or employed... data from LFS make international comparisons easier compared to a mixture of survey and registration data ...”.¹¹

Having discussed the conceptual differences between LFS and NIPA hours, we now turn to quantifying the observed differences. As commonly done in cross-country comparisons, we focus

¹¹Both quotes are from the OECD’s website: <http://stats.oecd.org/Index.aspx?DataSetCode=ANHRS> and <http://www.oecd.org/els/emp/basicstatisticalconceptsemploymentunemploymentandactivityinlabourforcesurveys.htm>. They were retrieved on August 12, 2016.

Figure 5: Percentage Deviations of Hours Worked and Employment in NIPA from LFS



on the comparison of hours per person rather than total hours worked. We obtain a NIPA measure of hours worked per person by dividing NIPA total hours by the population aged 15 to 64 obtained from the OECD’s “Annual Labour Force Statistics”.¹² For the LFS data, we divide LFS hours by the population aged 15 to 64 directly taken from the LFS, i.e. by the civilian, non-institutionalized population aged 15 to 64.

Figure 5a shows the average country-specific percentage deviation of NIPA estimates for hours worked per person from LFS estimates for the recent pre-crisis years 2005 to 2007. The NIPA estimate is larger than the LFS estimate for the majority of countries (on average by 7.2%). Only in 5 out of 19 countries are the NIPA estimates below the LFS estimates (on average by -4.2%).

To gain a better understanding of the sources of these differences, we turn to the two margins of labor supply. Figure 5b shows for the employment rate (black bars) and hours worked per employed (grey bars) the percentage deviations of the NIPA estimate from the respective LFS estimate. The NIPA employment rate estimate exceeds the LFS estimate in the majority of countries (on average by 4.0%, or 2.2 percentage points). Italy stands out with the difference exceeding 10%, followed by Switzerland, Germany, and Spain with differences of over 5%. In 5 countries is the NIPA employment rate lower than the LFS one, however only in the US and Poland in an economically significant way. Relative to the employment rate differences, the hours per employed differences are less systematic. In fact, the NIPA hours per employed estimate is lower than the

¹²These data can be found in the OECD’s “Labour Database” under the category “Labour Force Statistics”, and cover all residents of a country. The OECD National Accounts data do not provide any population figures broken down by age groups. Population figures are generally only available for the national concept and including everyone independent of status (civilian vs. non-civilian, institutionalized vs. non-institutionalized).

LFS estimate for 10 countries (on average by -4.1%) and higher in 9 countries (on average by 7.8%). There is also no systematic relationship between NIPA-LFS differences for the employment rate and hours per employed. The correlation between the two differences is -0.13.

What drives these differences between NIPA and LFS estimates? We can partially analyze this for employment rates, based on additional data from the OECD's "National Accounts Database" and "Annual Labor Force Statistics". In Web Appendix [W.4.2](#), we provide a detailed discussion of this investigation. In a nutshell, differences in employment stemming from the location of employment, i.e. the net difference in domestic residents working abroad and foreign residents working domestically, are for most countries negligible, with the exceptions of Switzerland and Hungary. The exclusion of non-civilian employment, i.e. the armed forces, in the LFS is also essentially irrelevant for the differences (even in the US). The "further groups" in Figure 4 are generally too small to rationalize the remaining employment rate differences. This leaves as the major driving force of the differences the use of different kinds of data sets and different methods by the national statistical agencies.

Unfortunately, for hours worked per employed we cannot conduct such a comparison, since the OECD does not provide hours for the armed forces or the total population using the national concept as for employment. The documentations provided by the national statistical agencies on how hours are estimated are also not helpful in getting quick insights into the differences. Web Appendix [W.4.1](#) shows in the left column the country-specific main data sources to construct employment rates, and in the right column the main sources to construct hours worked per employed. Both differ widely across countries. For Germany, for example, in addition to the two main data sources mentioned in the table, the hours estimate relies on 60 different data sets and is estimated with various econometric models. This makes it impossible to investigate what exactly generates the differences relative to the German LFS data.

For the US in turn, it is well known and documented that the hours estimates based on the CPS differ from NIPA estimates, for which the Current Employment Statistics (CES) by the BLS are the main data source. Through 2005, the CES data cover only production workers in goods-producing industries, and non-supervisory workers in services-providing industries, both within the private non-agricultural sector. These groups tend to have lower hours than the non-covered groups. For non-production and supervisory workers, hours are imputed based on data from 1978 assuming the same trends as for the covered groups, though data based on the CPS are not supporting the common trend assumption, see [Eldridge et al. \(2004\)](#). Only from 2006 onwards, all employees are covered. For all years, the NIPA estimate is supplemented by hours of self-employed and unpaid family workers from the CPS. Our data from the CPS in turn always cover all employees, next to

the self-employed and unpaid family workers. Another difference is that the CES measure hours per job, while the CPS measures hours per person in all jobs. In the presence of multiple job holders (about 5 to 6% of the population), the CES estimate will therefore be downward biased. Next, the CES measure hours paid, which are adjusted for paid vacations but not off-the-clock work, whereas the CPS measures hours worked. Last, there are differences in the reference period. The CES report hours for the pay period that includes the 12th of a month and can be weekly, biweekly or monthly, whereas the CPS reports hours for the week which includes the 12th of a month. [Frazis and Stewart \(2010\)](#) conclude that these differences can largely account for the level differences between CPS and NIPA hours estimates. They further document that hours in the CPS line up well with those from the American Time Use Survey, disproving an often mentioned criticism that hours in labor force surveys are overreported. [Ramey \(2012\)](#) shows that the data source matters for the measurement of the trend and cyclical behavior of labor productivity, and suggests the use of the CPS hours series.

While we have documented that in terms of levels NIPA and LFS hours estimates differ substantially for most countries, the time trends are much more in line with each other, although there are some exceptions. Figures [W.22](#) to [W.40](#) in Web Appendix [W.4.3](#) show for each country the time-series of hours per person, employment rates and hours per employed based on LFS and NIPA data from the OECD, and in addition from TED. As mentioned before, for most years and countries, the OECD and TED data series coincide, while the TED data usually go further back in time than the OECD data (and the LFS data).

Which data source is better suited for international comparisons remains an open question. We do not want to discount by any means the important and careful work by the national statistical agencies to measure labor inputs and output in each country in a consistent way. We want to emphasize however that the advantage of the labor force surveys is that we use a consistent data source and apply exactly the same methodology across countries, while NIPA estimates differ substantially in country-specific ways in both the kind of data sources used and the methodology applied. On top of this, NIPA data both by the OECD and the TED are subject to substantial revisions over time that also apply to previous years, which affect quantitative statements dramatically, see Web Appendix [W.4.4](#) and [Bick et al. \(2016\)](#). As argued in [Johnson et al. \(2013\)](#) and [Pinkovskiy and Sala-i-Martin \(2016\)](#) for data revisions in the Penn World Tables, it is not clear that these revisions lead to improvements of the data, and the opposite might actually be true. As mentioned above, the OECD itself cautions that their data are not intended for cross-country comparisons of levels, and suggests the use of labor force surveys instead. Before exploiting the key advantage of our data, namely the possibility to analyze the importance of different components

Table 2: Avg. Hours Worked per Person Relative to the US (in %) - Summary Statistics 2005-2007

Country	LFS	NIPA
Mean	-19.0	-7.1
Standard Deviation	6.6	9.3
Min Abs Deviation	5.1	3.3
Max Abs Deviation	30.2	23.6

and the demographic composition for total hours worked per person, we discuss in the next section the quantitative differences in the estimates of the Europe-US hours gap in NIPA and LFS data.

4 Hours Differences between Europe and the US: Measurement and Implications

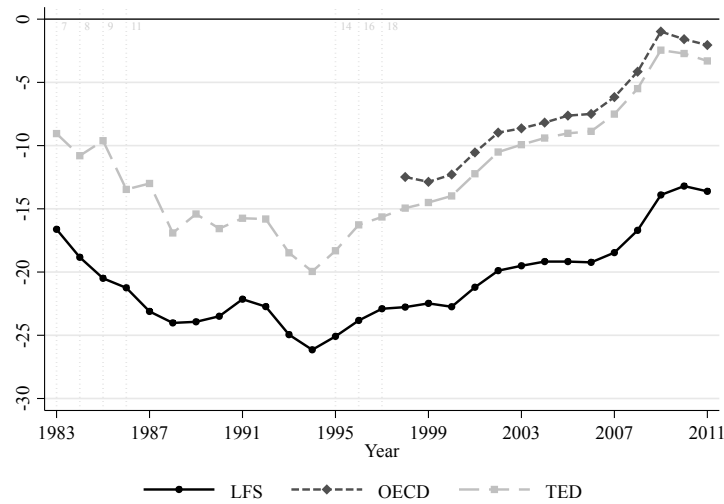
In this section, we first document hours worked per person differences between Europe and the US in the recent pre-crisis period 2005 to 2007 based on LFS data and NIPA data. We show that the differences are much larger based on LFS data. This is driven by both higher hours for the US and on average lower hours for Europe in LFS data. We then analyze two implications of the different measurements of the Europe-US hours gap, namely first for the measurement of productivity differences, and secondly for the analysis of models that postulate taxation as the major driving force of international hours differences.

4.1 Documenting Hours Differences between Europe and the US

Table 2 reports the difference in average hours worked per person between the 18 European sample countries and the US in LFS and NIPA data in the period 2005 to 2007. Based on LFS data Europeans work on average 19% fewer hours than US citizens, compared to only 7% based on NIPA data.¹³ Thus, the differences between Europe and the US are less than half as large in NIPA data than in LFS data, a stunning difference. The larger differences between European and US hours in LFS data relative to NIPA data are driven by both the higher hours worked per person estimates for the US in LFS data and the lower ones for Europe. US hours worked per person are with 1397 hours in LFS data 120 hours higher than in NIPA data with 1277 hours. By

¹³If we use the average vacation weeks from LFS data rather than from external data sources for our adjustment, the Europe-US hours gap is only mildly reduced to -16.1%. The population weighted Europe-US hours gap is -13.5% for LFS data and -6.7% for NIPA data.

Figure 6: Average Hours Worked per Person Relative to the US (in %), 1983-2011



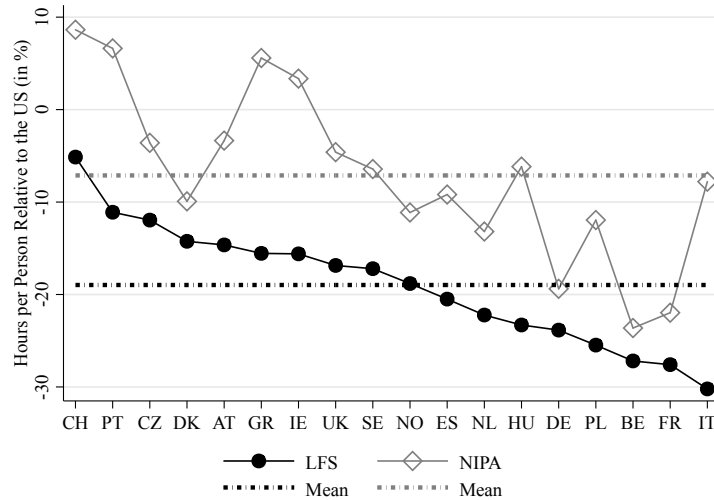
Note: The vertical lines indicate the years in which new countries enter the LFS sample, and the numbers next to them indicate the number of European countries in the sample from then on.

contrast, average hours worked per person over all European countries are with 1132 hours 55 hours lower in LFS data than in NIPA data. Therefore, around two thirds of the larger Europe-US hours difference in LFS data than in NIPA data come from higher US hours, and around one third from lower European hours. Finally, in terms of the cross-sectional variation within Europe, our data display a smaller standard deviation.

One might wonder whether the large discrepancies in the Europe-US hours difference between LFS data and NIPA data are an artifact of the pre-crisis cross-section, or have been present for some time. Figure 6 shows the average Europe-US hours difference in LFS data vs. OECD and TED data starting in the earliest sample year from the LFS data set, namely 1983.¹⁴ The vertical lines indicate the years in which new countries enter the LFS sample, and the numbers next to them indicate the number of European countries in the sample from then on. Since the number of European sample countries expands over time, the graph does not give a meaningful indication of the time-series development of the overall Europe-US hours difference before 1997. However, in any given year the sample countries are always the same in all three data sources. LFS data always indicate substantially larger Europe-US differences than either OECD or TED data, with the average absolute gap between the LFS and the OECD (TED) difference amounting to 11 (10) percentage points after 1997.

¹⁴The Europe-US hours gap is slightly larger for TED data as they feature a slightly larger hours per person estimate for the US, while TED and OECD hours per person estimates for Europe are basically the same.

Figure 7: Average Hours Worked per Person Relative to the US by Country, 2005-2007



Note that Prescott (2004), as well as other papers from this literature like Ohanian et al. (2008), Rogerson (2006), and McDaniel (2011), all rely on data from either OECD or TED using three different methods to measure hours worked per person, but find Europe-US hours gaps that are similar to the ones we document for LFS data. The reason is that both OECD and TED substantially revised their measures of hours worked per employed in new releases over the years, each time also affecting all previous years in the data series. Relying exactly on the formula and the data releases used by Ohanian et al. (2008), we find a Europe-US hours gap of 17.6% for 2003 (the latest available year in Ohanian et al. (2008)). Using exactly the same formula and same data sources, but working with the most recent releases from May 2016 (TED) and August 2016 (OECD), instead gives a gap of 11.0% for the year 2003.¹⁵ Web Appendix Section W.4.4 and Bick et al. (2016) give further evidence of the importance of data revisions in OECD and TED data.

Figure 7 plots hours worked per person differences (averaged over the years 2005 to 2007) between each European country and the US for the NIPA data (from the OECD) and LFS data, sorting the countries by decreasing difference to the US based on LFS data. With the exceptions of Switzerland, Portugal, Greece, and Ireland for NIPA data, hours worked per person in Europe are always lower than in the US. Furthermore, LFS data always predict a larger negative difference to the US than NIPA data. The smallest difference to US hours (in absolute value) in LFS data amounts to -5.1% for Switzerland, where NIPA reports higher hours than in the US by 8.6%. The largest difference to US hours in LFS data comes from Italy with -30%, where NIPA indicates

¹⁵Note that the three Eastern European countries in our sample are excluded from this comparison as they are not in the Ohanian et al. (2008) data set.

a difference of only -7.8%. Italy is the country on which the two data sources deviate most (22 percentage points), followed by Greece, Ireland, and Portugal. The smallest discrepancies arise for Belgium, Denmark, and Germany, with less than 5 percentage points. Still, the *ranking* of the countries is mostly preserved between both data sets: the correlation between the LFS hours worked per person measure and the NIPA measure is 0.74.

4.2 Productivity Differences between Europe and the US

Hours worked are a crucial input into the measurement of labor productivity, i.e. GDP divided by total hours worked. NIPA hours are generally the appropriate measure to calculate labor productivity because they are constructed specifically such that the labor inputs are consistent with the measurement of GDP. We (and the OECD itself) however argue above that LFS data might be better suited for international comparisons, since they are built using the same sources and methodology across countries. Given that LFS data provide different estimates of total hours than NIPA data, LFS data also lead to different conclusions regarding productivity differences between Europe and the US. Taking a measure of total GDP from the Penn World Tables (Y_{PWT}^{Total}), and defining total hours worked as H^{Total} , we can define labor productivity in country c as $LP_i^c = \frac{Y_{PWT}^{Total,c}}{H_i^{Total,c}}$ with $i = \{NIPA, LFS\}$. Labor productivity relative to the US is then defined as $\frac{LP_i^c}{LP_i^{US}} = \frac{Y_{PWT}^{Total,c}/Y_{PWT}^{Total,US}}{H_i^{Total,c}/H_i^{Total,US}}$. The larger the ratio between a country's hours and US hours in LFS data relative to NIPA data, the smaller is the labor productivity gap between this country and the US based on LFS data relative to NIPA data.

Table 3 calculates average GDP per hour worked over the years 2005 to 2007. In the first column, total hours worked are constructed from LFS data, in the second column they are taken from NIPA. GDP comes from the Penn World Tables Version 9.0, and is defined as the output-side real GDP at chained PPP in 2011 US-Dollar (*rgdpo*; see [Feenstra et al. \(2015\)](#)). GDP per hour worked in the US is normalized to 100 in both columns. The following rows then show GDP per hour worked in each of the European countries relative to the US. Total hours worked in the US are 10% higher in LFS data than in NIPA data, which leads to a lower estimate of GDP per hour worked in the US of \$49.7 (in 2011 US-Dollar) based on LFS data compared to \$57.9 based on NIPA data. We group the European countries into four regions, namely Scandinavia, Western Europe, Eastern Europe, and Southern Europe.

Given the larger hours difference between Europe and the US in LFS data than in NIPA data, we find smaller productivity differences overall between Europe and the US based on LFS data. While based on NIPA data Europe features 78% of the labor productivity of the US, based on

Table 3: GDP per Hour Worked Relative to the US (in %)

Country	LFS	NIPA
US	100.0	100.0
Europe	86.4	78.0
Denmark	88.3	86.9
Norway	172.0	161.8
Sweden	92.0	84.8
<i>Mean</i>	<i>117.4</i>	<i>111.2</i>
Austria	85.2	78.0
Belgium	91.4	88.7
France	98.3	92.5
Germany	97.0	94.2
Ireland	114.8	96.9
Netherlands	107.2	98.8
Switzerland	94.9	83.8
United Kingdom	92.0	80.3
<i>Mean</i>	<i>97.6</i>	<i>89.1</i>
Czech Republic	50.2	47.6
Hungary	45.2	37.8
Poland	40.1	34.4
<i>Mean</i>	<i>45.1</i>	<i>39.9</i>
Greece	64.0	51.3
Italy	93.7	74.0
Portugal	50.6	44.3
Spain	78.2	68.5
<i>Mean</i>	<i>71.6</i>	<i>59.5</i>

LFS data Europe's labor productivity amounts on average to 86% of the US one. Quantitatively, the differences across data sets are largest in absolute terms for Southern Europe: based on NIPA data, Southern Europe exhibits only 60% of the labor productivity of the US, while based on LFS data Southern Europe's labor productivity amounts to 72% of the US one. Put differently, the productivity gap is reduced by nearly one third relative to NIPA data. The results for Italy are particular striking, where the labor productivity gap nearly disappears for LFS data. Western European labor productivity is at 89% of the US productivity based on NIPA data, but comes very close to the US level with 98% based on LFS data. Hence, for Western Europe we see the

largest reduction in relative terms of the labor productivity gap (by three fourths). Differences in productivity measurement between the different data sets are smallest for Eastern Europe and Scandinavia. In Eastern Europe, labor productivity is with 45% of the US level still higher in LFS data than with 40% in NIPA data. Because of the extraordinary high labor productivity in Norway, on average Scandinavian labor productivity surpasses the one in the US based on both LFS and NIPA data, with the former being about 6 percentage points higher.

Summarizing, in terms of labor productivity Europe looks closer to the US based on LFS data than based on NIPA data. This is mostly driven by the closing of the rather small gap for Western Europe by three fourths and the large gap for Southern Europe by one third in LFS data compared to NIPA data. Scandinavian labor productivity surpasses US labor productivity in both LFS and NIPA data, whereas the difference between Western Europe and the US nearly disappears in LFS data. Southern and Eastern European labor productivity levels are significantly lower than in the US also in LFS data, amounting to three quarters and one half of the US level.

4.3 Implications for Macro Models of Taxation

A series of recent papers in the macroeconomic literature analyzing the driving forces of cross-country differences in *aggregate* hours worked focus on the the role of taxes. How much does it matter for the conclusions which data source one uses for the analysis? For labor productivity measurement all hours counting towards the production of GDP should be included, and thus in principle NIPA data should be better suited than LFS data. However, for the analysis of the role of taxation for determining hours worked one could argue the opposite: if e.g. a large part of the institutionalized population does not have a meaningful choice of how many hours to work (e.g. the incarcerated population), then their hours should not be affected by taxation. In addition, residents working abroad are usually taxed by their home country.

To gain insights into the question in how far differences in tax rates can explain cross-country differences in hours worked per person, we follow the analysis in [Prescott \(2004\)](#) and incorporate taxes and lump-sum transfers into the neoclassical growth model. Each country is inhabited by a representative household that maximizes the present discounted value of future utility, with a utility function that is separable in consumption and leisure:

$$\max_{\{c_t, h_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t (\log c_t + \alpha \log(\bar{h} - h_t)), \quad (4)$$

where c_t is consumption, h_t hours worked, \bar{h} the total time endowment, and α determines the

relative disutility of work. The household budget constraint is given by

$$(1 - \tau_c)c_t + k_{t+1} = (1 - \tau_h)w_t h_t + (1 + r_t)k_t + T_t, \quad (5)$$

where τ_c is the consumption tax rate, τ_h is a linear labor income tax rate, k is capital, and T is a lump-sum government transfer.¹⁶ The government uses tax revenues to finance the lump-sum transfer. Last, there is a representative firm with a Cobb-Douglas production function

$$y_t = A_t k_t^\theta h_t^{1-\theta}. \quad (6)$$

Maximization by the household implies that the marginal rate of substitution between leisure and consumption equals their price ratio, and profit maximization by the firm that prices equal the marginal products. Combining these first order conditions, one can obtain optimal hours as

$$h_t = \frac{1 - \theta}{1 - \theta + \frac{c_t}{y_t} \frac{\alpha}{1 - \tau_t}} \quad (7)$$

with $\tau_t = \frac{\tau_h + \tau_c}{1 + \tau_c}$ being the tax wedge. Rather than solving the full dynamic model, [Prescott \(2004\)](#) interprets Equation (7) as the equilibrium value of hours worked given parameters and values for the consumption-output ratio. The consumption-output ratio, directly taken from the data, captures the dynamic component of the neo-classical growth model.

We calibrate the model and then assess its predictions for the Europe-US hours gap in the recent pre-crisis years 2005 to 2007. We take $\frac{c}{y}$ from the Penn World Tables 9.0 (see [Feenstra et al. \(2015\)](#)), specifically we set $\frac{c}{y} = csh_c + csh_g$,¹⁷ and following [Prescott \(2004\)](#) set the capital share $\theta = 0.3224$, the time endowment $\bar{h} = 52 \times 100$ hours, and the tax rate $\tau_t = \frac{\tau_{ss,t} + 1.6\tau_{inc,t} + \tau_{c,t}}{1 + \tau_{c,t}}$. $\tau_{ss,t}$ is the social security tax rate, and $\tau_{inc,t}$ is the average income tax rate, which is multiplied by 1.6 to obtain an average marginal tax rate. We use the tax rates from [McDaniel \(2012\)](#). Last, we calibrate the disutility of labor α to match average hours worked per person in the US, using the consumption to output ratio and the tax rate for the US. To predict hours worked in European countries, we assume that preferences are the same as in the US, but plug in country-specific values for the consumption to output ratio and the tax rates.

¹⁶We abstract for simplicity from investment and capital interest taxation, as well as capital depreciation, since they do not enter the static first order condition for the optimal hours choice.

¹⁷Note that consumption here refers to consumption by private households as well as by the government. Government consumption is thus treated as a lump sum transfer to households. [Prescott \(2004\)](#) subtracts from government consumption two times the military's share of employment times GDP, as he does not regard military expenditure as a substitute for private consumption. Since we do not have information on the military share of employment for all

Table 4: Europe-US Hours Difference (in %) in Data and Model

	LFS	NIPA
α	1.54	1.75
Data	-19.0	-7.1
Model	-10.1	-10.3
R²	0.16	0.14

Summary results are shown in Table 4, while the country-specific results are presented in Web Appendix Section W.5. The first column presents results when LFS hours are used for the calibration and subsequently for the comparison of model predictions to data from European countries, the second column presents results when we use NIPA hours. The first row shows the calibrated values of the disutility of labor α . Since US hours are higher in LFS data, α is lower when LFS data are used for the calibration. The second row repeats the Europe-US hours gap in the two different data sets. The next row gives the model predictions. It turns out that the different values of α have relatively small effects, such that the model predicts a very similar hours gap between Europe and the US based on both data sources. Through the lens of the neo-classical growth model differences in taxation and consumption to output ratios can thus account for 52% of the Europe-US hours gap when compared to LFS data (-10.1% in the model vs. -19% in LFS data). Evaluated against NIPA data, taxation and the consumption to output ratio account for more than the entire difference (-10.3% in the model vs. -7.1% in NIPA data).

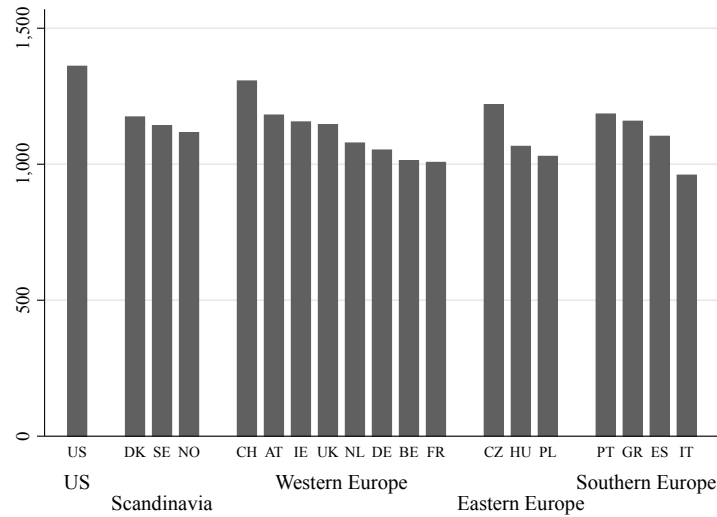
Put differently, while with NIPA hours taxes appear to be more than sufficient in accounting for the large Europe-US hours gap, with LFS hours taxes account for an important fraction of the gap, but leave ample room for other factors. Obviously, with a smaller labor supply elasticity than the one implied by Equation 4, the predicted Europe-US hours worked difference would be smaller, but the model would still perform better contrasted with NIPA data than with LFS data. Note that the model fit of the variation of hours *within* Europe is similar using both data sources as indicated by the R-squared in the last row.

5 Decomposing the Europe-US Hours Difference

In this section, we turn to the core question of the paper and investigate the role of different components of labor supply and the demographic and sectoral composition for hours worked differences

countries in our sample, we do not implement this adjustment.

Figure 8: Average Hours Worked per Person (2005-2007)



between Europe and the US, using LFS data from the recent pre-crisis cross-section 2005-2007 and covering the age group 15 to 64.¹⁸

Before we do this, we show in Figure 8 hours worked per person in the LFS data on the country level by geographic region. Within each region, we order countries by average hours worked per person. The US and Switzerland stand out with the highest hours worked per person, exceeding 1300, while Italians work less than 1000 hours per year. The mean hours worked per person within each European region are quite similar, ranging from 1102 hours in Southern Europe to 1144 hours in Scandinavia.

5.1 The Three Components of Hours Worked per Person

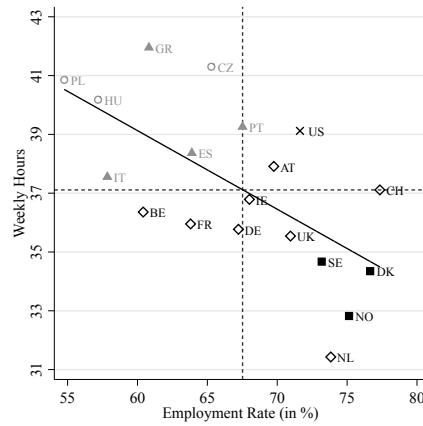
Average hours worked per person are the product of the employment rate and hours worked per employed, with the latter being a product of the weekly hours worked per employed in a non-vacation week and the number of weeks worked per year, i.e. $H = e \times h \times w$. Figure 9 shows how these three components relate to each other and reveals interesting patterns of heterogeneity across regions, despite the similar average hours worked per person across regions. The five regions are marked by different markers, namely the US by an x, Eastern Europe by a circle, Scandinavia by squares, Western Europe by diamonds, and Southern Europe by triangles.

Figure 9a plots weekly hours worked per employed in a non-vacation week (h) against the

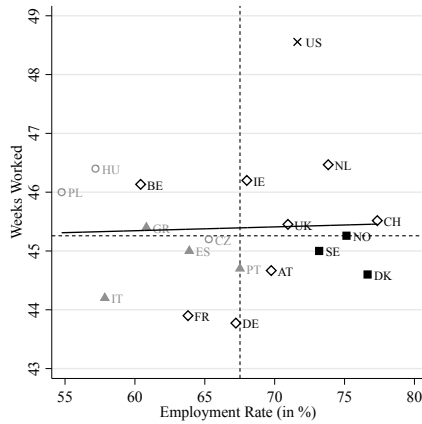
¹⁸In contrast to Sections 3 and 4, we use hours of anyone aged 15 to 64 in this section, not hours of anyone aged 15 or older as the NIPA data does. This accounts for the small differences between Figure 7 and Table 5.

Figure 9: Hours Worked Components

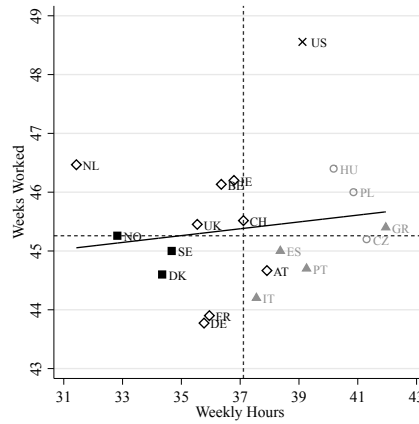
(a) Weekly Hours & Empl. Rate



(b) Weeks Worked & Empl. Rate



(c) Weeks Worked & Weekly Hours



employment rate (e), with the dotted vertical and horizontal lines indicating median values, and the solid line representing an OLS regression line. While hours worked per person do not differ much on average across the four European regions, as shown in Figure 8, a region-specific pattern for the relationship between the employment rate and weekly hours emerges. Eastern and Southern European countries all have above median weekly hours, but below median employment rates, with Portugal being the median country. The Scandinavian countries in turn have below median weekly hours and above median employment rates. Weekly hours worked per employed in the Western European countries are between those two groups of countries. Two exceptions are Austria with above median weekly hours and the Netherlands with the lowest weekly hours. The employment rates range from the fourth lowest in Belgium to the highest in Switzerland. Taken together, a strong negative cross-country correlation between weekly hours worked per employed and the

employment rate of -0.64 arises: countries with high employment rates tend to feature low weekly hours worked per employed in a non-vacation week, and vice versa.

Figure 9b plots the weeks worked (w) in a given country against the employment rate (e). Weeks worked range in Europe from below 44 in Germany and France to 46.5 in the Netherlands, with most of the cross-country variation coming from annual leave rather than public holidays (see Web Appendix W.2.1). The US clearly stand out with more than 48 weeks worked per year. Overall, the correlation between weeks worked and the employment rate is weak, amounting to 0.04 for all countries and -0.10 if the US is excluded. Figure 9c plots weeks worked (w) against weekly hours worked in a non-vacation week (h). Again, the correlation between both variables is rather small with 0.15 and drops further to 0.05 once the US is excluded. Web Appendix Section W.6.1 quantifies the contribution of the three components to explaining hours worked differences between the respective European country and the US, similarly to Alesina et al. (2005).

While we of course cannot make any causal statements based on our data, the results presented in Figure 9a to Figure 9c provide some information on potential driving forces of hours differences and give some guidance for modeling choices. They indicate that countries featuring high employment rates tend to feature low weekly hours worked per employed (and vice versa), while weeks worked are largely uncorrelated with both employment rates and weekly hours. The negative correlation between employment rates and weekly hours worked seems surprising: standard factors that influence labor supply, e.g. taxes, tend to influence both components in the same direction. Establishing this negative correlation in a quantitative model of labor supply is therefore not trivial. The reasons for the negative correlation between employment rates and weekly hours worked could potentially stem from the employee's side, the employer's side, or from regulation. In Bick and Fuchs-Schündeln (2016), we e.g. document a similar negative correlation for married women aged 25 to 54, and provide some suggestive evidence that it might be driven by differences in part-time regulation across countries.

The absence of any correlation between weeks worked and employment rates or hours worked, on the other hand, seems to indicate that more public holidays or annual leave, i.e. less weeks worked per year, neither induce systematically more people to work nor to work longer hours during a non-vacation week. Weeks worked (or the reverse, namely vacation weeks) thus potentially seem to be determined by other forces than those determining employment rates and weekly hours worked. These could e.g. be institutional factors, like historic unionization rates. Moreover, weeks worked do not seem to influence the individual decision whether to enter the labor force and how many hours to work. This would call for modeling weeks worked differently than employment and hours worked per week. One could e.g. follow the model in Kaplan (2012), who introduces a

model in which involuntary unemployment (in our case vacation weeks) determines the number of weeks worked per year, and agents choose only hours worked per work week optimally. Assuming separability in the utility function between weeks worked and hours worked per week in such a model could lead to a zero correlation between weeks worked per year and either employment rates or weekly hours worked per employed.

5.2 Disaggregate Decomposition

Countries in our sample differ not only in the three components of labor supply, but also in their demographic structure, i.e. the composition of the population by demographic characteristics. We take into account three demographic characteristics, namely gender, age, and education, as well as the sectoral composition of employment in a country. Explicitly accounting for these groups, we can write aggregate hours worked per person in country c as

$$H^c = \sum_j f_j^c e_j^c \times \sum_k s_{j,k}^c h_{j,k}^c \times w^c, \quad (8)$$

where j represents a set of demographic characteristics and k the sector of employment. f_j^c is the fraction of individuals with a given set of characteristics j (with $\sum_{j=1}^J f_j^c = 1$), e_j^c is the employment rate of group j , $s_{j,k}^c$ is the share of employed individuals of group j working in sector k (with $\sum_{k=1}^K s_{j,k}^c = 1$), $h_{j,k}^c$ are the weekly hours worked per employed in a non-vacation week by group j in sector k , and w^c are the weeks worked (non-vacation weeks). Each group j is defined by the interaction of one of three age groups (15-24, 25-54, 55-64), three education groups (low, medium, high), and gender. Only for the young age group, we allow a fourth educational category “still enrolled in education”. As a result, there are 20 groups j . Since only employed individuals can be allocated to a sector, the sectoral composition only affects hours worked per employed, not the employment rate. We consider the three sectors services, manufacturing, and agriculture.¹⁹ Last, remember that in the baseline analysis we assume the same number of weeks worked for everyone, an assumption we relax below.

To quantify the contribution of the three components of labor supply, the demographic composition, and the sectoral composition to the difference of the country-specific hours relative to the US, we calculate counterfactual hours by setting for every country one of these five features after the other equal to the corresponding US value. We then measure the incremental fraction of the aggregate hours worked per person difference relative to the US accounted for by this feature.²⁰

¹⁹Defining self-employment as an own “sector” does not affect the results in a substantial way.

²⁰Web Appendix Section W.6.1 describes the procedure in detail focusing only on the three components of labor

The quantitative results depend on the order in which the features are set country specific. With five features, there are 120 different possible orderings, and we report the mean results over all 120 different orderings. This decomposition is somewhat similar in spirit to the one done by [Blundell et al. \(2011\)](#) and [Blundell et al. \(2013\)](#) for time-trends in hours in the France, the UK, and the US.

Demographic groups (or the sectoral composition) matter in the decomposition analysis only if two conditions are met: first, the composition across countries needs to be different. E.g. since in all sample countries and years women make up around 50 percent of the population aged 15 to 64, applying the gender composition of the US to a country will not affect the corresponding hypothetical estimate of aggregate hours worked per person. Secondly, the demographic groups need to exhibit either different employment rates or different weekly hours worked *within* a country. E.g. even if the educational composition is different across countries, if low, medium, and high educated people were to participate in the market to the same degree and exhibit the same weekly hours worked within a country, applying the US educational composition would have no effect on the corresponding hypothetical estimate of aggregate hours worked per person. Note that by applying all possible orderings of the components in the decomposition (i.e. sometimes the demographic structure is set to the US level before hours and/or employment rates, sometimes afterwards), the groups could behave differently either in the US or in the respective European country for the demographic composition to play a role.

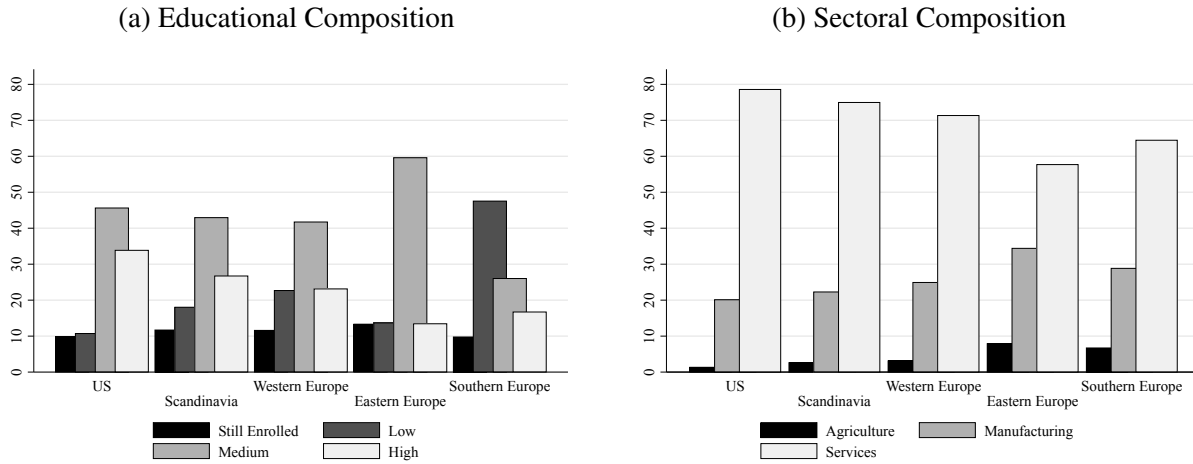
With regard to the demographic composition, not only is the gender composition very similar across countries, but the age composition, referring to individuals aged 15-24, 25-54, and 55-64, is relatively similar as well (as can be seen in Web Appendix Figure [W.41](#)). The same is however not true for the educational composition, which [Figure 10](#) shows in the left panel. The US stands out with the lowest (highest) share of low (high) educated individuals with 11% (above 30%), whereas in Eastern and Southern Europe less than 20% of the population have high education. Eastern Europe exhibits the highest share of medium educated individuals, amounting to almost 60%, while Southern Europe features the highest share of low educated individuals, amounting to almost 50%. The fraction of young adults still enrolled in education is with 10 to 15% rather similar across regions. The different levels of education are defined according to the ISCED classifications, with “low” corresponding to lower secondary education, “medium” to upper secondary education, and “high” comprising any tertiary education degree.²¹

The right panel of [Figure 10](#) shows the sectoral composition across regions. The share of employed people working in services is highest in the US with 79 percent, followed by Scandinavia

supply analyzed in the previous subsection.

²¹For example, for the US “low” corresponds to less than high school education, “medium” to completed high school education, and “high” to having at least an associate’s degree.

Figure 10: Demographic Structure



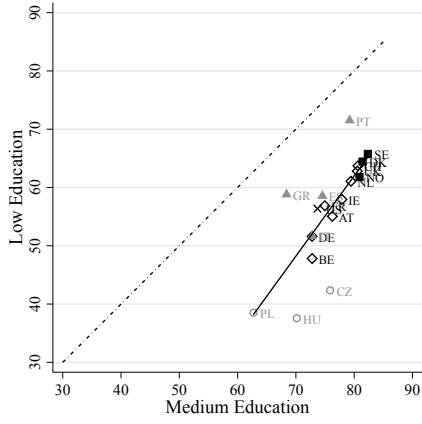
and Western Europe, which still exhibit shares above 70 percent. It is however significantly lower in Southern Europe with 64 percent, and especially in Eastern Europe with 58 percent.²² The low share in services in Eastern and Southern Europe is offset by both a higher share of individuals working in manufacturing, and a slightly higher share working in agriculture. The share of employed people working in agriculture is however very low everywhere, reaching its maximum with 8 percent in Eastern Europe.

While the sectoral and the educational compositions thus differ significantly across countries, this only has an effect in the disaggregate decomposition if different educational groups show different labor market behavior within a country, and individuals in different sectors work different hours. Panels (a) to (d) in Figure 11 show the corresponding evidence for education groups: there are indeed substantial differences in employment rates of low, medium, and high educated individuals within a country for all countries (panels (a) and (b)), but rather small differences in weekly hours worked (panels (c) and (d)). On average across all countries, medium educated individuals have a 20 percentage points higher employment rate than low educated individuals, but a 9 percentage points lower one than high educated individuals. By contrast, the differences in weekly hours are with on average 1.5 hours between low and medium, and 0.6 hours between medium and high educated individuals small. Panels (e) and (f) show that differences in weekly hours between individuals working in different sectors are somewhat larger than between individuals with different education levels: individuals working in the manufacturing sector work on average 3.1 more weekly hours than individuals in the services sector. Individuals working in agriculture work even longer weekly hours, especially in the Western European countries, but their share there is

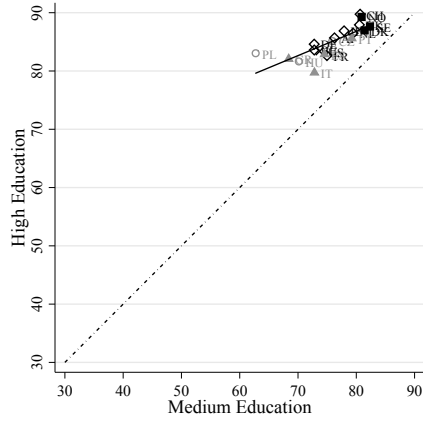
²²Rogerson (2008) documents the historically lower employment share in services in Europe than in the US.

Figure 11: Labor Supply Broken Down by Education or Sector

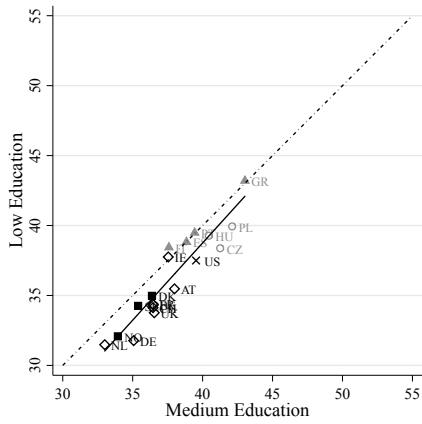
(a) Employment Rate: Low & Medium



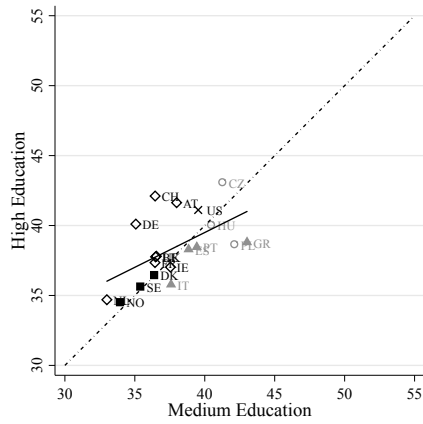
(b) Employment Rate: High & Medium



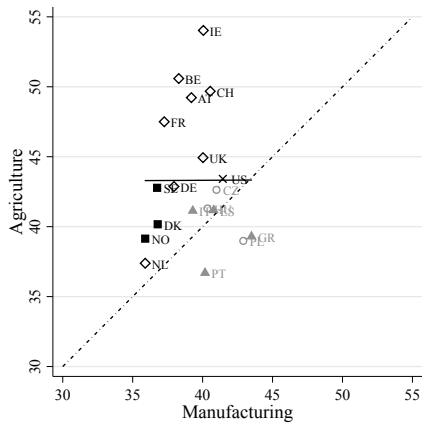
(c) Weekly Hours: Low & Medium



(d) Weekly Hours: High & Medium



(e) Weekly Hours: Agriculture & Manufacturing



(f) Weekly Hours: Services & Manufacturing

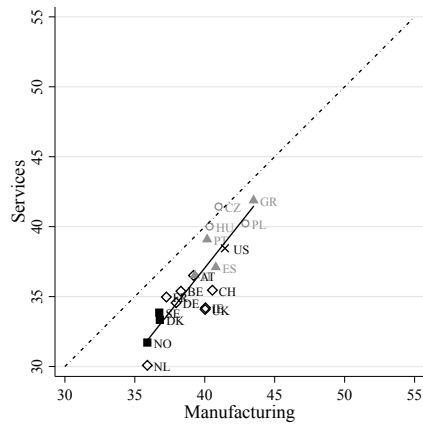
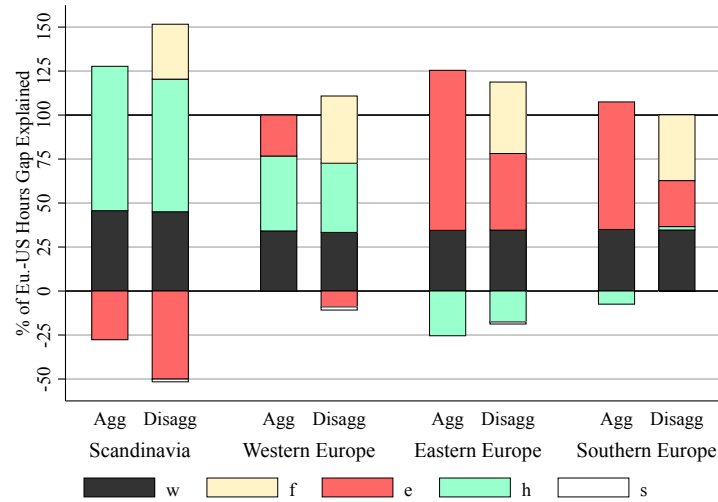


Figure 12: Aggregate and Disaggregate Decomposition



Note: Each part of a bar represents the fraction of the country group-US hours gap explained by the respective component. All parts of a bar add up to 100%. The averages for Western and Southern Europe exclude Switzerland and Portugal, respectively.

too small to affect the aggregate.

The consequences of these findings on the decomposition of cross-country differences in hours worked per person are shown in Figure 12. To facilitate the comparison, we show only averages over the regions (excluding Switzerland from Western Europe and Portugal from Southern Europe for reasons explained below) and also show the corresponding results from an aggregate decomposition that only analyzes the importance of the three components of labor supply, but omits the demographic and sectoral composition, i.e. the results presented in Web Appendix Section W.6.1. In this figure, the difference in hours worked per person with the US is divided into the fraction explained by weeks worked w , employment rates e , weekly hours h , the demographic composition f , and the sectoral composition s . The two latter components only show up in the disaggregate composition, which is presented in the right bar for each country group, but not in the aggregate composition in the left bar. If a factor does not positively contribute to the overall difference, but in fact indicates higher hours in the European region than in the US, this is indicated by a negative entry.

The demographic structure f accounts for a substantial fraction of the hours worked per person differences relative to the US in all European regions. Put differently, the demographic composition in Europe predicts substantially lower hours than in the US. Introducing the demographic structure reduces the importance of the employment rate in all regions significantly. The share

explained by the employment rate even turns from positive to negative in Western Europe. By contrast, there are only minor changes in the share of the Europe-US hours difference explained by weekly hours when the demographic and sectoral structures are introduced in the disaggregate decomposition. These two results are due to the fact that employment rates increase substantially by education in all countries, whereas weekly hours are similar for the three educational groups within each country. Thus, not accounting for the higher share of low- or medium-educated individuals in Europe than in the US attributes too much importance to employment rates.²³ While weekly hours worked differ more by sectors than by education groups, the differences across sectors within each country are still relatively small, and as a result the sectoral composition is a negligible factor in explaining the Europe-US hours gap. Note that in our decomposition the sectoral composition can only affect hours worked per person through differences in weekly hours worked per employed between sectors, since we cannot assign non-employed individuals to a sector. Rogerson (2008) stresses the importance of an underdeveloped service sector in Europe for explaining low hours there. In his model, high tax rates in Europe induce individuals to specialize in home production, leading to a stagnating service sector instead of a growing one as in the US, since services produced in the market are closer substitutes to services produced at home than goods produced in the market to goods produced at home. As far as this channel works through drawing less people into employment, it is not captured in our decomposition which by construction can only handle sectoral differences in weekly hours worked.

The full disaggregate decomposition results are shown in Table 5. It reports in column 1 the hours worked per person difference relative to the US, i.e. $(\frac{H^c - H^{US}}{H^{US}}) \times 100$. In column 2 we report what fraction of the difference in hours worked per person with the US can be accounted for by differences in the number of non-vacation weeks w . Columns 3 to 6 report the same statistic for the demographic composition f , employment rates e , weekly hours worked h , and the sectoral composition s , respectively. The five last columns add up to 1. A positive value indicates that a factor positively accounts for the (negative) difference, while a negative value means that the respective factor does not contribute to the (negative) difference, but would in fact indicate higher hours in the respective European country than in the US. As an example, Scandinavians work on average 16% lower hours worked per person than US Americans. 75% of this difference can be accounted for by lower weekly hours worked, 45% by a lower number of weeks worked, and 31% by the demographic structure. At the same time, employment rates are larger in Scandinavia than in the US (see Figure 9a). This is not entirely accounted for by the demographic structure, and

²³Using only gender and age to define a demographic group yields results very similar to using the aggregate data only (not shown here).

Table 5: Disaggregate Decomposition

Country	$\frac{H^c - H^{US}}{H^{US}}$	Δ^w	Δ^f	Δ^e	Δ^h	Δ^s
Denmark	-13.7	56.5	39.6	-66.1	73.3	-3.3
Sweden	-16.1	43.0	32.1	-43.5	68.4	0.0
Norway	-18.0	35.7	21.9	-40.5	84.3	-1.4
<i>Mean</i>	<i>-15.9</i>	<i>45.1</i>	<i>31.2</i>	<i>-50.1</i>	<i>75.3</i>	<i>-1.6</i>
Switzerland	-4.0	160.9	86.1	-240.9	103.3	-9.4
Austria	-13.2	59.2	39.8	-13.7	21.3	-6.6
Ireland	-15.0	26.2	56.8	-16.1	42.9	-9.8
United Kingdom	-15.8	37.7	51.3	-38.3	50.0	-0.7
Netherlands	-20.8	18.4	35.4	-33.7	71.7	8.3
Germany	-22.6	40.7	23.5	5.7	31.8	-1.8
Belgium	-25.5	17.5	30.0	24.1	28.9	-0.5
France	-26.0	33.5	30.6	8.8	28.6	-1.5
<i>Mean</i>	<i>-17.9</i>	<i>49.3</i>	<i>44.2</i>	<i>-38.0</i>	<i>47.3</i>	<i>-2.8</i>
<i>Mean w/o CH</i>	<i>-19.8</i>	<i>33.3</i>	<i>38.2</i>	<i>-9.0</i>	<i>39.3</i>	<i>-1.8</i>
Czech Republic	-10.4	65.4	50.4	21.9	-36.7	-1.0
Hungary	-21.6	18.7	39.4	49.1	-5.7	-1.6
Poland	-24.4	19.7	32.0	59.6	-10.3	-1.0
<i>Mean</i>	<i>-18.8</i>	<i>34.6</i>	<i>40.6</i>	<i>43.5</i>	<i>-17.6</i>	<i>-1.2</i>
Portugal	-12.9	58.5	89.9	-43.3	-4.6	-0.5
Greece	-14.9	41.5	40.5	45.5	-29.0	1.6
Spain	-19.1	35.7	38.2	11.0	16.4	-1.3
Italy	-29.4	26.8	33.5	21.9	18.4	-0.6
<i>Mean</i>	<i>-19.1</i>	<i>40.6</i>	<i>50.5</i>	<i>8.8</i>	<i>0.3</i>	<i>-0.2</i>
<i>Mean w/o PT</i>	<i>-21.1</i>	<i>34.7</i>	<i>37.4</i>	<i>26.1</i>	<i>1.9</i>	<i>-0.1</i>

Note: Columns 2 to 6 indicate what fraction (in %) of the hours worked per person difference to the US in column 1 is accounted for by the corresponding factor given in the column header. Columns 2 to 6 sum up to 100 for each country.

hence the differences in the employment rate cannot account for why hours worked per person are lower in Scandinavia than in the US, and that causes a negative entry in the table. Scandinavian hours worked per person would actually be 8% (-0.16×-0.50) larger than US hours based on the employment rate differences after controlling for the demographic structure, rather than 16% lower.

Weeks worked w and the demographic structure f uniformly predict lower hours in Europe than in the US. Weeks worked account on average over the European regions for one third to one half

of the lower hours, and the demographic structure for another one third to one half. Differences in the sectoral composition *s* play no significant role in explaining hours differences to the US in any of the European countries. After controlling for the demographic and sectoral compositions in this decomposition exercise, we still observe marked patterns in the relative importance of employment rates and weekly hours: in Scandinavia and Western Europe, weekly hours worked are the main driver of lower hours per person compared to the US, explaining a share of 75 and 39 percent, respectively, with the employment rate going substantially in the opposite direction by 50 percent in Scandinavia, and slightly by 9 percent in Western Europe. Note that we exclude Switzerland from the Western European average when discussing the results, because it features very high values in the decomposition table because of its small difference in hours worked per person to the US of only -4%.²⁴

While weekly hours universally point to lower hours in all Scandinavian and Western European countries, though by varying degree, the employment rates do not universally indicate higher hours, with Belgium, France, and Germany being exceptions to this rule. In contrast to Scandinavia and Western Europe, employment rates indicate lower hours worked per person in Eastern and Southern Europe except Portugal, explaining on average 44 and 26 percent of the hours difference to the US, respectively. Portugal is the stark exception here, with employment rates indicating substantially higher hours than in the US. Weekly hours go in the opposite direction in Eastern and Southern Europe, indicating higher hours than in the US, with the exception of Spain and Italy. For Spain and Italy, weekly hours and employment rates are both roughly equally important in explaining the hours gap to the US. In Portugal, lower hours are entirely driven by a lower number of work weeks and the demographic decomposition. Overall, the results in Table 5 make clear that the negative cross-country correlation of the two components employment rates and weekly hours is not primarily driven by differences in the demographic or sectoral compositions, but by some more fundamental factors that apply to all groups in the labor market.

By definition, the importance of work weeks is not affected by the demographic or sectoral decomposition, as we impose the same number of work weeks for all subgroups within a country. We relax this assumption in Web Appendix Section W.6.3, where we allow for heterogeneity in weeks worked as described in Web Appendix Section W.2.5 by scaling the reported vacation weeks in the LFS data of each gender \times age \times education \times sector group up by a common multiplicative factor such that the average vacation weeks match the ones from external data sources. This procedure (rather than assuming a common additive factor, or scaling up weeks worked rather than

²⁴Remember that the hours differences to the US in this section are slightly different than the hours differences to the US presented in Figure 7, because in this section we include only hours of individuals aged 15 to 64, not hours of all working individuals aged 15 and above; see footnote 18.

vacation weeks) maximizes the degree of heterogeneity. Web Appendix Table W.18 shows the results. While the results are almost unchanged for Scandinavia and Western Europe, allowing for heterogeneity in weeks worked by demographic group reduces the importance of the demographic composition in Eastern and Southern Europe by 9 and 14 percentage points, respectively, which is nearly completely offset by a corresponding increase in the importance of weeks worked.

Summarizing, there are three main findings from the disaggregate decomposition analysis. First, more vacation weeks alone account uniformly across Europe for between one third to one half of the hours difference to the US. Secondly, there exist large differences across Europe in the relative contribution of employment rates and weekly hours in explaining the hours difference to the US even after controlling for differences in the demographic and sectoral composition. For Scandinavia and Western Europe, lower weekly hours worked explain three quarters and roughly 40 percent, respectively, of the lower hours than in the US, with employment rates indicating higher hours than in the US. For Eastern and Southern Europe (excluding Portugal), by contrast, lower employment rates explain roughly one half and one quarter, respectively, of the lower hours than in the US, and weekly hours worked tend to be higher than in the US. Understanding the roots of these differences constitutes an interesting new research agenda. Third, differences in the educational composition constitute an important driving force of hours worked differences across countries: the different educational composition, with a higher share of low and medium educated people in Europe, accounts for one third to one half of the differences in hours worked per person, similarly to weeks worked. Not taking the educational composition into account would overestimate the importance of the employment rate in explaining lower hours in Eastern and Southern Europe, and would underestimate the degree to which employment rates indicate higher hours in Scandinavia and Western Europe than in the US.

The results point to the importance of studying the educational composition as a driving force of hours worked differences across countries. Which policies or other factors determine the different educational compositions across countries? For example, Restuccia and Vandenbroucke (2014) analyze the role of longevity and economic growth on the educational composition across countries and over time. Note that some of these policies affecting educational choices could affect hours worked directly at the same time. Guvenen et al. (2014), while not distinguishing between employment and hours per employed, show that progressive taxation distorts the incentives to invest in human capital, and that these tax differences can explain a sizable part of the differences in the Europe-US hours per person gap for men. Also, the results point to the importance of understanding why employment rates increase strongly with education within a country, but hours worked per employed do not. Policies that affect the incentives to work differentially by education could e.g.

be welfare systems, or different degrees of flexibility in the labor markets for rich and poor, e.g. due to a minimum wage being binding for the poor. Moreover, the disutility of work could differ in low- and high-education jobs. Yet, one would have to investigate if these factors influence the extensive margin more than the intensive one. In this regard, we want to mention the caveat that categorizing educational groups across countries in a consistent way is far from trivial, given the vastly different educational systems across countries.

6 Conclusion

In this paper, we document the construction of a new data set of aggregate hours worked starting from labor force surveys. This data set allows to analyze which fraction of the Europe-US hours analysis is driven by the different components of labor supply, as well as the role of the demographic and sectoral composition. We rely on data from the US Current Population Survey, the European Labor Force Survey, and the German Microcensus to construct hours worked measures. The main challenges for comparability of the several surveys at a given point in time across countries, or for a given country over time, regard the sampling of reference weeks over the year and the underreporting of vacation weeks. We show that we overcome both problems by obtaining hours worked in a *non-vacation* week, i.e. a work week without a reduction in work time because of a public holiday or annual leave, and requiring consistency of annual leave and public holidays in the micro data with the country-wide average. Information on the latter is collected from external data sources. While time trends in LFS and NIPA data are similar, we document substantial differences in the level of the Europe-US hours gap in both types of data sources. In the LFS data, Europeans work on average 19% fewer hours than Americans in the recent pre-crisis period 2005 to 2007, compared to only 7% in NIPA data. The likely reasons lie less in the slightly different coverage of the population, but more in the country-specific kinds of data sources and methodologies used to construct NIPA hours, which might inhibit the international comparability. The differences between LFS and NIPA data are large enough to matter for the measurement of labor productivity and the explanatory power of taxation for hours differences across countries.

In the main part of the paper, we decompose hours worked differences between Europe and the US into weeks worked, employment rates, and weekly hours worked per employed in a non-vacation week, as well as the role of the demographic and sectoral composition in accounting for these differences. Weeks worked and the demographic composition, driven by the educational composition, both account for between one third and one half of the lower hours in Europe than in the US. Weekly hours worked account for three quarters and roughly 40 percent of the lower hours

in Scandinavia and Western Europe, respectively, where employment rates contribute negatively to explaining the hours gap to the US. By contrast, employment rates account for roughly one half and one quarter, respectively, of the lower hours in Eastern Europe and Southern Europe without Portugal, with weekly hours worked mostly contributing negatively to the difference.

Our results indicate new avenues in the research agenda of the study of hours worked differences across Europe and the US. First, they show that the higher number of vacation weeks in Europe is a significant driver of lower hours, and that longer vacation weeks are not associated with either higher employment rates or longer weekly hours worked. This might indicate that the number of work weeks is determined by other driving factors than those determining employment rates and weekly hours worked, and raises the question whether they should be modeled as a choice variable for the individual. Secondly, we document a negative correlation between weekly hours worked and employment rates across countries that is not entirely driven by differences in the sectoral or educational composition. Modeling this negative correlation is not trivial, and investigating the sources is a promising research agenda. Third, the educational composition matters significantly in accounting for cross-country differences in hours worked through its effect on employment rates. Understanding why different education groups exhibit such different employment rates (but not weekly hours worked), and different countries have such different educational compositions, will thus be a helpful step in explaining hours worked differences across countries.

References

- Alesina, A., E. Glaeser, and B. Sacerdote (2005). Work and Leisure in the United States and Europe: Why So Different? *NBER Macroeconomics Annual* 20, 1–64.
- Alonso-Ortiz, J. (2014). Social Security and Retirement Across the OECD. *Journal of Economic Dynamics & Control* 47, 300–316.
- Bick, A., B. Brüggemann, and N. Fuchs-Schündeln (2016). A Note on the Europe-US Gap in Hours Worked per Person: Data Revisions and Different Formulas. Working Paper.
- Bick, A. and N. Fuchs-Schündeln (2016). Taxation and Labor Supply of Married Couples Across Countries: A Macroeconomic Analysis. Working Paper, Arizona State University and Goethe University Frankfurt.
- Bick, A., N. Fuchs-Schündeln, and D. Lagakos (2016). How Do Average Hours Worked Vary with Development? Cross-Country Evidence and Implications. *NBER Working Paper* 21874.
- Blundell, R., A. Bozio, and G. Laroque (2011). Labour supply and the extensive margin. *American Economic Review Papers and Proceedings*.

- Blundell, R., A. Bozio, and G. Laroque (2013). Extensive and Intensive Margins of Labour Supply: Work and Working Hours in the US, the UK and France. *Fiscal Studies* 34, 1–29.
- Bridgman, B., G. Duernecker, and B. Herrendorf (2016). Structural Transformation, Marketization, and Household Production Around the World. Working Paper.
- Burda, M., D. Hammermesh, and P. Weil (2008). *Working Hours and Job Sharing in the EU and USA*, Chapter The Distribution of Total Work in the EU and the US.
- Burda, M., D. Hammermesh, and P. Weil (2013). Total Work and Gender: Facts and Possible Explanations. *Journal of Population Economics* 26(1), 239–161.
- Eldridge, L. P., M. E. Manser, and P. F. Otto (2004). Alternative measures of supervisory employee hours and productivity growth. *Monthly Labor Review* April, 9–28.
- Erosa, A., L. Fuster, and G. Kambourov (2012). Labor Supply and Government Programs: A Cross-Country Analysis. *Journal of Monetary Economics* 59, 84–107.
- European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, and World Bank (2009). *System of National Accounts 2008*.
- Eurostat (2013). *European System of Accounts ESA 2010*.
- Fang, L. and C. McDaniel (2016). Home Hours in the United States and Europe. *B.E. Journal of Macroeconomics: Advances Forthcoming*.
- Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015). The Next Generation of the Penn World Table. *American Economic Review* 105(10), 31503182.
- Fleck, S. E. (2009, May). International Comparisons of Hours Worked: An Assessment of Statistics. *Monthly Labor Review* 132(5), 3–31.
- Frazis, H. and J. Stewart (2010). Why Do BLS Hours Series Tell Different Stories About Trends in Hours Worked? *BLS Working Paper* 433.
- Güvener, F., B. Kuruscu, and S. Ozkan (2014). Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis. *Review of Economic Studies* 81, 818–850.
- Johnson, S., W. Larson, C. Papageorgiou, and A. Subramanian (2013). Is Newer Better? Penn World Table Revisions and Their Impact on Growth Estimates. *Journal of Monetary Economics* 60(2), 255–274.
- Kaplan, G. (2012). Inequality and the Life Cycle. *Quantitative Economics* 3, 471–525.
- Köerner, T. and K. Marder-Puch (2015). Der Mikrozensus im Vergleich mit anderen Arbeitsmarktstatistiken. *WISTA – Wirtschaft und Statistik* (4).

- Krueger, D., F. Perri, L. Pistaferri, and G. L. Violante (2010). Cross-Sectional Facts for Macroeconomists. *Review of Economic Dynamics* 13(1), 1–14.
- McDaniel, C. (2011). Forces Shaping Hours Worked in the OECD, 1960-2004. *American Economic Journal: Macroeconomics* 3, 27–52.
- McDaniel, C. (2012). Average Tax Rates on Consumption, Investment, Labor and Capital in the OECD 1950-2003. Working Paper.
- Ohanian, L. and A. Raffo (2012). Aggregate Hours Worked in OECD Countries: New Measurement And Implications for Business Cycles. *Journal of Monetary Economics* 59(1), 40–56.
- Ohanian, L., A. Raffo, and R. Rogerson (2008). Long-Term Changes in Labor Supply and Taxes: Evidence from OECD Countries, 1956-2004. *Journal of Monetary Economics* 55, 1353–1362.
- Olovsson, C. (2009). Why Do Europeans Work So Little? *International Economic Review* 50(1), 39–61.
- Pilat, D. (2003). Measuring Hours Worked in OECD Countries. DSTI/EAS/IND/SWP(2003)10, OECD.
- Pinkovskiy, M. and X. Sala-i-Martin (2016). Newer Need Not Be Better: Evaluating the Penn World Tables and the World Development Indicators Using Nighttime Lights. *NBER Working Paper* 22216.
- Prescott, E. C. (2004). Why Do Americans Work So Much More Than Europeans? *Federal Reserve Bank of Minneapolis Quarterly Review* 28, 2–13.
- Ragan, K. (2013). Taxes, Transfers and Time Use: Fiscal Policy in a Household Production Model. *American Economic Journal: Macroeconomics* 5(1), 168–192.
- Ramey, V. A. (2012). The Impact of Hours Measures on the Trend and Cyclical Behavior of U.S. Labor Productivity. Working paper, University of California San Diego.
- Restuccia, D. and G. Vandenbroucke (2014). Explaining Educational Attainment across Countries and over Time. *Review of Economic Dynamics* 17(4), 824–841.
- Rogerson, R. (2006). Understanding Differences in Hours Worked. *Review of Economic Dynamics* 9, 365–409.
- Rogerson, R. (2008). Structural Transformation and the Deterioration of European Labor Market Outcomes. *Journal of Political Economy* 116, 235–259.
- Schnitzlein, D. D. (2011). Umfang und Folgen der Nichtinanspruchnahme von Urlaub in Deutschland. *DIW Wochenbericht* 51+52, 14–20.
- Wallenius, J. (2014). Social Security and Cross-Country Differences in Hours: A General Equilibrium Analysis. *Journal of Economic Dynamics and Control* 37, 2466–2482.