

The Updated Okun Method for Estimation of Potential Output with Broad Measures of Labor Underutilization: An Empirical Analysis*

Claudia Fontanari, Antonella Palumbo, Chiara Salvatori**

Working Paper No. 158

May 11th, 2021

ABSTRACT

This paper extends to different indicators of labor underutilization the Updated Okun Method (UOM) for estimation of potential output proposed in Fontanari et al (2020), which, from a demand-led growth perspective, regards potential output as an empirical approximation to full-employment output, as in A.M.Okun's (1962) original method. Based on the apparent incapability of the official rate of unemployment to fully account for labor underutilization, in this paper we offer estimates of Okun's law both with broad unemployment indicators and with an indicator of 'standardized hours worked' which we propose as a novel measure of the labor input. The paper reflects on the possible different empirical measures of full employment. The various measures of potential output that we extract from our analysis show greater output gaps than those produced by standard methods, thus highlighting a systematic tendency of the latter to underestimate potential output. Output gaps that underestimate the size of the output loss or that tend to close too soon during recovery, may produce a bias towards untimely restriction.

<https://doi.org/10.36687/inetwp158>

JEL Codes: E60; E24; O40; E12; C30

Keywords: potential output; alternative unemployment indicators; Okun's law; demand-led growth

* We would like to thank Thomas Ferguson, Jie Chen and Alessia Naccarato for precious comments on an early draft. Financial support of the Institute for New Economic Thinking (INET – New York) is gratefully acknowledged.

** Roma Tre University.

1. Introduction

Assessing the size of the gap between actual and potential output that the pandemic has opened in various economies is a very complex exercise, given the peculiar characteristics of the crisis and the violence of the ups and downs most economies are experiencing. During 2020, amidst the huge uncertainty about the full extent of the losses, the ever-changing restrictions associated with new waves of infection and the impact of relief policy measures, even the most intrepid international institutions have presented their estimates with extreme caution (IMF 2020a, p.136; ECB 2020b). Also the figures produced in early 2021, the same agencies warn, are subject to significant uncertainty (see IMF 2021, p. 12 and 122).

Estimates are vary widely across the different institutions. The 2020 output gap for the USA, for example, has been around 3 % for the International Monetary Fund and almost double according to the OECD, while the US Congressional Budget Office (CBO, 2021) sets it at 4.4%. Even wider discrepancies are registered for other countries (for the UK, for example, the estimated gaps were -4.4% for the IMF, -8.5 for the EC and -8.9 for the OECD).¹ In addition, quite naturally, each agency revises its estimates while events unfold. While some of these variability and uncertainty are rooted in the peculiarity of the situation, yet potential output estimates have a long history of unreliability of their own. These famously show in the excessive pro-cyclicality of the estimates, frequent revisions over time, and excessive variability of the signs and size of the output gaps across estimates based on the same model and differing only by minor auxiliary hypotheses (Bundesbank 2014, Heimberger 2020). Non-negligible differences in output gaps, for example, are also observable in the estimates of the same institutions for 2018 or 2019, before the pandemic hit, even though they employ the same theoretical conception of potential output and use broadly the same methods of estimation.

The urgency of the pandemic and the long history of poor performance of potential GDP estimates reinforces the conviction that it is time to rethink profoundly both the definition and the measurement of potential output, as we advocated in Fontanari et al (2020). A better assessment of the gulf that has opened between actual and potential output may prove especially important during the phase of recovery. In the worst of the pandemic recession, the consensus on the need for a massive fiscal effort went undisputed, and there is little doubt that the size of the fiscal stimulus packages adopted in most countries have not been determined on the basis of traditionally estimated output gaps, but rest rather on experienced losses in actual output and employment and the urgency of immediate relief needed by large parts of societies. However, if mainstream economic theory recognizes the need for demand management to fight recessions, and specifically for fiscal policy in the current context of very low interest rates that weakens the effectiveness of

¹ The figures are taken from the World Economic Outlook of the International Monetary Fund (IMF), issues of October 2020 and April 2021, the Autumn forecast of the European Commission (EC), November 2020, and the Economic Outlook of the Organization for Economic Co-operation and Development (OECD) of December 2020. Estimates will likely be revised shortly.

monetary policy, it nonetheless generally denies any role for it beyond cyclical stabilization. For this reason, the measurement of output gaps is crucial in assessing how much of the great fiscal effort associated to the pandemic has to be regarded as structural rather than purely cyclical. As shown by recent debates,² the idea that the repeated waves of massive fiscal stimulus may overshoot the output gap has raised concerns for their possible inflationary effects. Yet, output gaps that underestimate the size of the output loss or that, as has repeatedly happened in the past, tend to close too soon following a deep recession and a timid recovery, may produce, in the near future, serious bias towards untimely restriction. This would not only produce a gratuitous lengthening of the phase of slack, but would also permanently affect the economy: as the demand-led growth perspective has pointed out repeatedly and much mainstream literature has finally come to acknowledge, recessions and slow growth have permanent effects on the growth prospects of an economy by reducing investment, the capital stock and labor participation and depressing labor skills (IMF 2020a, ECB 2020a).

It is thus very relevant to reconsider the mainstream idea that the level of potential output and its growth are to be defined as those compatible with constant inflation, corresponding to the unique non accelerating inflation rate of unemployment (NAIRU) postulated in the accelerationist Phillips-curve models.³ Given that the output gaps in this theoretical account should be strongly correlated with inflation changes, both the historical irregularity of the inflation-unemployment relationship in actual data – including the even more puzzling ‘disconnect’ between inflation and unemployment observed in recent decades (see Del Negro et al 2020) – and the fact that the estimated output gaps are usually poor predictors of inflation (Banbura and Bobeica, 2020) point to deep flaws in the standard conception of potential output. In actual fact, inflation data both offer little information and play a comparatively minor role in the procedures for empirical identification of the NAIRU (see the critical analysis in Brooks and Fortun, 2020). The latter ends up being defined in practice as any level of unemployment that happens to persist on average for a sufficient length of time, based on the theoretical idea, built into the estimation methods, that deviations of actual from equilibrium unemployment only occur on a temporary basis. As we have shown in Fontanari et al (2020), this confers a tautological quality to estimates of both the NAIRU and potential output, which in the medium-long period tend to reflect actual trends of unemployment and output. This explains why recessions sediment into estimates of potential output, drag them down, leading to hasty judgements that output gaps have closed, when in fact they have not.

² Based on the above-mentioned CBO’s estimates of February 2021, some (see for example Summers 2021) have raised misgivings about the size of the additional fiscal stimulus enacted by the US administration, calculating that it would exceed several times the estimated output gap thus giving rise to serious inflationary risks. On the other hand, due to the very flat slope of the currently estimated Phillips curves, Krugman (2021) has noted that in the current circumstances “even a very hot economy only leads to modest inflationary overheating”.

³ For the relevance of the Phillips curve framework in policy making, both at central banks and other institutions such as the European Commission, see for example Williams (2019), Hooper et al (2019), Eser et al (2020), Havik et al (2014).

It is interesting to note, in this respect, that the sizeable negative output gaps that the European Commission estimated in 2020 (both in its Spring and Autumn Forecast) for many European countries were generally much bigger than those estimated by the IMF: compare for example -8.3 vs -4.6% for France, or -9.5 vs -5.7% for Italy. These have been obtained by the EC by ‘forcing’ a correction into the standard procedure so as to prevent the cyclical downturn from affecting the potential output path, which in the EC’s 2020 estimates is almost consistent with its pre-pandemic trajectory.⁴ That such a result is obtained only by introducing judgmental elements and discretionary corrections in the procedure of estimation points once more to the general unreliability of standard estimates of potential output.

The failure of estimation models based on the Phillips curve echoes in the change in the monetary policy strategy that the US Federal Reserve System announced in August 2020, a change which clearly downplays, for the time being, the policy relevance of output gaps for setting the interest rates. The new strategy amounts to abandoning strict inflation targeting in favor of an approach that will allow the rate of inflation to stay above target for a protracted period, so that the inflation target is attained only on average over a lengthy period. This reflects not only the above-mentioned observed disconnect between unemployment and inflation implying the practical irrelevance of estimated output gaps in predicting inflation, but also the shift in policy priorities in favor of a direct targeting of unemployment,⁵ possibly coupled with the need for keeping the interest rates very low given the growing size of both public and private debt (see the discussion in Lubik and Schwartzman 2020).⁶

All of the above highlight the fact that, in order to correctly assess the extent of the slack and have genuine policy relevance, the theoretical definition and practical estimation of potential output needs to be both immune to the obvious flaws of the standard notion but also consistent with the disconnect between inflation and unemployment so clearly borne out by empirical evidence in the last decades. In Fontanari et al (2020), we have proposed such an alternative conception and measurement of potential output. By starting from a demand-led growth perspective and rescuing A.M. Okun’s (1962) original method of estimation, we calculate potential output as an empirical approximation to a traditional Keynesian notion of full-employment output, understood as a sort of ‘maximum’ of production attainable in any given situation. Our notion reflects the theoretical

⁴ For the details about corrections, see EC (2020c, p.54) and Italian Ministry of Finance (2020), illustrating the indications of the Output Gaps Working Group (OGWG) of the European Council. In (EC 2020c) the European Commission considers alternative scenarios, some implying long-term damages of the pandemic crisis on potential output.

⁵ See for example J. Smialek, How Full Employment Became Washington’s Creed, New York Times, Jan. 18 2021, available at <https://www.nytimes.com/2021/01/18/business/economy/full-unemployment-fiscal-policy.html>

⁶ For discussion of similar steps in Europe, see for example Lengwiler and Orphanides (2020) and the interview with Philip Lane, 2 December 2020, available at <https://www.ecb.europa.eu/press/inter/date/2020/html/ecb.in201202~d24616dd00.en.html>). See also Blanchard et al (2020), proposing a reform of the Stability and Growth Pact aimed to downplay the role of estimated output gaps.

idea, typical of the demand-side perspective, that no automatic mechanisms ensure the tendency of actual output towards potential, and thus allows the opening up of sizeable, non-self-correcting divergences, in phases of slack, between actual and potential output.

In Fontanari et al (2020) we reviewed the advantages that this revised notion of potential output has over standard methods of estimation. In the first place it refrains from any tautological definition of potential output as the trend component of actual output. In the second place, it is a relatively simple procedure, with a very intelligible meaning. The gaps calculated through the updated Okun method may vary in size, but this depends straightforwardly on the conventionality of the chosen policy target. Finally, our method entirely refrains from the use of inflation data, which implies that the deeply problematic Phillips relationship plays no role in it.

However, the procedure relies, for its validity, on the hypothesis that the unemployment rate is a good indicator of labor underutilization and thus constitutes the basis on which the amount of slack can be properly gauged. This can be viewed as a possible shortcoming of our method, especially in a historical phase in which – first in the prosperity of the late 2010s, then in the current pandemic crisis – the rate of unemployment is apparently unable to account properly for the situation of the labor market. We will thus, in this paper, extend our original analysis by testing the validity of Okun’s law with different indicators that in our opinion – and according to the literature – should be better able to grasp the actual extent of labor underutilization (a preliminary analysis in this direction has been proposed in Fontanari et al 2019).

We will extract from our analysis a series of alternative evaluations of potential output and the output gaps, focusing both on the analysis of the last part of the long expansionary phase of the 2010s and on the assessment of the extent of slack caused by the pandemic in the second and third quarters of 2020, notwithstanding the tentative character of this exercise. We will try to show that full employment has a necessarily conventional definition and may be differently measured in terms of various indicators. The measures of potential output we extract from our analysis will vary accordingly. Yet, in general our estimates consistently show greater output gaps compared with the standard methods, thus once again highlighting a systematic tendency of the latter to underestimate potential output.

The structure of the paper is as follows. Section 2 assesses the ability of the rate of unemployment to correctly approximate labor underutilization, on the basis of both the experience of the late 2010s and that of 2020. It also reviews possible alternative indicators, especially the ‘broad’ indicators of unemployment that take into account such phenomena as discouragement in participation and part-time employment for lack of full-time opportunities.

In sections 3 and 4 we test Okun’s law with such ‘broad’ unemployment indicators, something that, to our knowledge, has never been attempted in other works. Section 3 is devoted to a cross-country analysis of Okun’s law. The estimates confirm the general validity of the relationship, both across countries and with different indicators. In section 4, we focus our attention on the

United States. Using the different unemployment indicators, we test and confirm the hypothesis we have originally put forward in Fontanari et al 2020, namely that the labor input tends to react more to output changes when it is massively underutilized, while, when the economy approaches high levels of employment, output increases tend to translate in a greater proportion in productivity increases.

But while we think that these measures of labor underutilization are superior to official unemployment rates, we also suspect that these wider measures fail to properly account for all the important features of slackness. Some literature, which we review in section 4.2, tries indeed to measure unemployment/inactivity not only in terms of numbers of persons but in terms of hours that people would be prepared to work but do not work. On this basis, we propose in section 4.3 an indicator of the labor input based on hours worked, which we regard as a more accurate expression of the actual extent of labor utilization in the economy, and test Okun's law with such an indicator, finding again confirmation of the validity of the law.

Section 5 is devoted to the application of our Updated Okun Method, enhanced by the array of indicators we have reviewed, to the calculation of potential output and the output gaps for the US economy in the last part of the 2010s and in 2020. Section 6 sums up the results and concludes.

2. Unemployment and labor underutilization

2.1. The dubious reliability of the official unemployment rate

As is known, the Okun method for potential output estimation relies on the existence of a sufficiently regular empirical relationship between unemployment changes and output growth, the so-called Okun's law. First detected by A.M. Okun (1962) on US data (on the 1947-1960 period), such a law finds general confirmation in the subsequent empirical literature (see Fontanari et al 2020 for a review), especially – perhaps surprisingly – over long periods, while showing some variability in the short period. The independent estimation that we conducted in Fontanari et al (2020) on a 59-year timespan on US data once again confirms the validity of the relationship and finds that its fit improves when allowing the Okun coefficient to take different values for different ranges of the unemployment rate.

We noticed however a less good fit of the relationship in recent years, especially following the 2009 structural break. In parallel, we reviewed some literature (Erceg, Levin 2014; Aaronson et al. 2014; Summers 2017; Storm 2018) highlighting the characteristics of the US economy in the 2010s, when progressively lower unemployment rates went together with the virtual absence of any appreciable inflationary pressures and comparatively low participation rates. Indeed, the unemployment rate remained steadily at near 4 percent (or even below) for 10 consecutive quarters, from 2017Q3 (when it was 4.1%) to 2020Q1, with no signs of excessive growth of wages nor any other symptoms of an overheating economy. The lack of inflationary pressures and the disappointing growth performance induced the Federal Reserve, in July 2019, to cut interest rates, a move that should have produced spiraling inflation if the economy had been anywhere near (or

beyond) full employment. The participation rate in the USA shows a declining trend in recent years (in the pre-pandemic period) beyond the effect of cyclical factors, in contrast to the contrary trend of most of other OECD countries (see Pérez-Arce and Prados, 2020 and the literature there quoted). Such a trend may be only partially explained by population aging and other structural factors such as higher enrollment rates in the education of young adults. Focusing on the crucial 25-54 age range, indeed, the participation rate in January 2020 (the highest after the Great Recession) was fully 1.6 percentage point lower than the peak value of the late 1990s. In addition, the very low unemployment rates possibly also masked some underutilization of labor in terms of hours worked.

Similar considerations may be made for some other countries that recorded equally low rates of unemployment in the late 2010s. In Germany, for example, the unemployment rate has been steadily below 4 percent since 2017, reaching a value of 3.4% in 2018 and a still lower value in 2019 (3.1%). This latter reduction, however, has been accompanied by a substantial increase (23 percent in one year) in the number of short-time workers. Even the much higher rates of unemployment recorded in other European economies in the end of the 2010s, for that matter, probably seriously underestimated the actual extent of labor underutilization.⁷

If doubts could already be cast on the reliability of the unemployment rate as an indicator of labor underutilization in the long recovery of the 2010s (see our discussion in Fontanari et al 2019), the dramatic recession of 2020 associated with the coronavirus pandemic and the subsequent upheavals in the labor market have only made such doubts stronger. In the first place, it is apparent that the contrasting cross-national policy responses to the crisis, which in most European countries took the form of short-time work schemes (full or partial furloughs heavily subsidized by the state, see ECB 2020b), has meant that the (possibly analogous) strong reduction in the labor input associated with the crisis has had enormously different effects on the unemployment rates. In Europe, these appeared for the second and third quarter of 2020 almost stable with respect to the pre-pandemic period,⁸ while undergoing wild up and downs in the USA. In the latter country, the fact that the rate of unemployment exploded to 14.8 percent in April 2020 while steadily decreasing in the following months (for example it was 6.9 in October and 6.7 in November, according to BLS data), has soon induced some authors (Hall and Kudlyak, 2020; Gallant et al, 2020) to maintain that the dramatic increase in unemployment consisted mainly of temporary lay-offs and was to be soon re-absorbed almost entirely.⁹ However, the second quarter of 2020 has been also characterized by sharp falls in participation rates and generalized reductions in working hours. In the USA the participation rate, which dropped in April 2020 by more than 3 percentage points with respect to February, has recovered by only 1.5 points until April 2021; while in European countries,

⁷ In Italy, for example, the unemployment rate was 9.9 percent in the fourth quarter of 2019, corresponding to about 2.5 million unemployed, but an equally great number of people (2.6 million) were computed by the national institute of statistics (ISTAT) as potentially attached to the labor force, though not looking actively for work.

⁸ As an example consider France, where the rate of unemployment was 8.1% in the fourth quarter of 2019, and respectively 7.5 and 8.6 in the second and third quarters of 2020.

⁹ The US rate of unemployment has since decreased further, although being in April 2021 still much above the pre-pandemic level.

notwithstanding the short-time work schemes, participation rates have decreased.¹⁰ Thus, on the whole, data on participation and working time point to a much more substantial amount of slack than that recorded by the unemployment rates.¹¹

2.2. Alternative indicators: the ‘broad’ unemployment rates

Even before the pandemic crisis, the limitations of the official unemployment rate had been pointed out in several works in the wake of the Great Recession. Several authors have called for the use of more accurate indicators of labor slack (see Komlos 2019b and the literature there quoted), especially due to the apparent complete disconnect between unemployment and inflation. Yellen (2019) pointed out the possibility “that labor market slack is not appropriately measured by the civilian unemployment rate. Perhaps broader measures of slack including, for example, individuals involuntarily working part-time or some who are considered to be out of the labor force entirely are relevant to wage and price inflation”.

We thus consider it natural to turn our attention to the ‘alternative’ indicators of unemployment that the BLS produces and releases monthly, an array of six different indicators meant to highlight different dimensions of labor underutilization. The current definitions of such indicators were introduced by the BLS in 1994; they replace a previous set of seven indicators that had been introduced in 1976.¹²

The definitions currently in use are the following (the six indicators represent increasingly broad definitions of unemployment; the first two indicators, U1 and U2, represent more restrictive definitions of unemployment than U3, the official unemployment rate; see the BLS website):

- U-1, persons unemployed 15 weeks or longer, as a percent of the civilian labor force;
- U-2, job losers and persons who completed temporary jobs, as a percent of the civilian labor force;
- U-3, total unemployed, as a percent of the civilian labor force (this is the definition used for the official unemployment rate);

¹⁰ In France for example, the participation rate fell by 2.2 percentage points in the second quarter of 2020; on average, in the Eurozone, it fell by 1.6 points.

¹¹ “The coronavirus pandemic has put millions more workers across the developed world out of jobs than official unemployment statistics suggest, according to economists’ estimates — threatening economic recovery prospects in the months ahead... But economists say the true number of people who have lost work because of the pandemic is far higher, after taking account of those whose jobs are temporarily protected by state-subsidized furlough schemes, those who have dropped out of the labor force and those who cannot work as many hours a week as they would like. Estimates of the exact number of people whose livelihoods have been damaged vary wildly. But governments across the developed world all face the same dilemma: the immediate economic shock of the initial lockdown measures has eased, but lingering joblessness is likely to cause widespread hardship, hold down wages and slow the pace of a consumer-driven recovery.” (Delphine Strauss on the Financial Times, 7 October 2020).

¹² Originally, the introduction of multiple unemployment indicators is especially due to Julius Shiskin, former Commissioner of the BLS. For the meaning and characteristics of the original indicators and the reasons for their updating see Bregger and Haugen, 1995.

- U-4, total unemployed plus discouraged workers, as a percent of the civilian labor force plus discouraged workers;
- U-5, total unemployed, plus discouraged workers, plus all other marginally attached workers, as a percent of the civilian labor force plus all marginally attached workers; and
- U-6, total unemployed, plus all marginally attached workers, plus total employed part time for economic reasons, as a percent of the civilian labor force plus all marginally attached workers.

As specified by the BLS, the marginally attached workers are ‘those who currently are neither working nor looking for work but indicate that they want and are available for a job and have looked for work sometime in the past 12 months’ (it is worth recalling that in order to be regarded as officially unemployed, persons must have conducted a job search in the four weeks preceding the interview). Discouraged workers, a subset of the marginally attached, are persons that have given a job-market related reason for not currently looking for work. In other words, they are prepared to work but did not look recently for a job because they believe finding one was unlikely or impossible. Persons employed part time for economic reasons are those who want and are available for full-time work but have had to settle for a part-time schedule for lack of full-time jobs.

We thus combine the literature on the limitations of the official unemployment rate as an indicator of labor slack and the growing attention to the alternative indicators of unemployment with the suggestion found in some literature on Okun’s law (particularly in Sogner and Stiassny 2002 and in Busetta and Corso 2012) that Okun’s law may hold, or even be improved, by using indicators of the labor market other than the rate of unemployment. These authors experiment with such indicators as the employment rate or the annual working units. Our analysis, by contrast, focuses on the above-mentioned alternative unemployment indicators, particularly on the broadest ones, U5 and U6.

3. Okun’s Law with alternative indicators of unemployment: a cross-country analysis

As a first step, we test Okun’s law with alternative indicators of unemployment in a panel of OECD countries. The availability of only short time series of the components of the alternative unemployment indicators leads us to build two different datasets. The first dataset, which is extracted from the AMECO database of the European Commission, contains longer series but no data on U5 and U6 and will be used to test the traditional version of Okun’s law.¹³ It covers annual data for a panel of 19 OECD countries over a period of 36 years from 1983 to 2019.¹⁴ The second

¹³ We will make use later of this dataset also to test Okun’s law with an indicator of hours worked; see section 4.3.1 below.

¹⁴ The reference period has been chosen with the aim of including as many countries as possible over a fairly long period of time. The choice of data frequency is essentially due to data availability.

dataset, which is extracted from the Labor Force Statistic database of the OECD, contains data on U5 and U6 for 18 OECD countries but covers a shorter time-span (2000-2018).¹⁵

Using the first dataset, we thus begin by estimating the traditional version of Okun's Law, which relates unemployment changes to output changes by using the official rate of unemployment (U3).

3.1. Okun's law with U3 in a panel of countries

Cross-country analyses of Okun's law have been performed in several works (Paldam 1987; Moosa 1997; Lee 2000; Freeman 2001; Virén 2001; Mayes and Virén 2002; Sogner and Stiassny 2002; Perman and Tavera 2005; Malley and Moolana 2008; Stock and Vogler-Ludwig 2010; Cazes and Verick 2011; Ball et al. 2017; Garavan 2017); generally on samples of OECD countries. Some authors estimate Okun's law individually for each country, applying either an Error Correction Model (ECM; see Lee 2000, Viren 2001, Mayes and Viren 2002) or the Seemingly Unrelated Regressions (SUR) technique (Moosa 1997; Freeman 2001; Perman and Tavera 2005, Ball et al 2017). Some analysts also estimate the relationship in a pooled sample (Freeman 2001 on 10 OECD countries; Garavan 2017 on 19 European countries). Overall, the literature finds both confirmation of the law and differences in the Okun coefficient between countries.

Based on our analysis in Fontanari et al 2020, we use the difference specification of Okun's law in its dynamic form (Ball et al., 2017; Knotek, 2007; Silvapulle et al., 2004). The difference specification extracts the Okun coefficient directly from observed data, without requiring additional hypotheses about the potential values of the variables; while the dynamic form allows consideration of lags in the adjustment of employment and labor force participation when output changes (Ball et al. 2017; Knotek 2007; Lee 2000).¹⁶ Based on our demand-led growth approach and following Okun's (1962) original analysis, we model changes in unemployment as the dependent variable and output growth as the independent one. The estimation model is thus the following:¹⁷

$$\Delta u_t = \alpha - \beta_1 gy_t - \beta_2 gy_{t-1} \quad (1)$$

where Δu_t is the first difference of the unemployment rate and the gy_i 's the contemporaneous and lagged values of the rate of growth of output. For all variables, we check stationarity using the Levin-Lin-Chu test for unit roots and find that stationarity is confirmed (see Table A2 in the

¹⁵ In both datasets, gross domestic product is valued at constant 2015 prices; the unemployment rate is expressed as a percentage of the civilian labor force and for non-EU countries it is defined according to the OECD definition (see Table A1 in the Appendix).

¹⁶ A fuller discussion of the different possible specifications of Okun's law and their meaning is contained in Fontanari et al (2020).

¹⁷ Schwarz's Bayesian, Akaike, Hannan and Quinn information criterion have been used to choose the optimal number of lags of the independent variable. The result varies by country, indicating as optimal for some countries only the contemporaneous effect and for most countries up to one lag. For the sake of cross-country comparison, we therefore decide to include the one-period lagged term in the estimates.

Appendix). We control for other potential sources of bias and find autocorrelation of residuals, heteroscedasticity and correlation across panels in both datasets.¹⁸

Equation (1) is estimated both individually and in a pooled sample. The individual country regressions cannot control for time-varying effects but, on the other hand, they permit assessing how the Okun coefficient varies between countries and the different extent to which the relationship explains the variability of the unemployment rate in the different countries. Based on Freeman, 2001; Moosa, 1997; Ball et al, 2017 and Perman and Tavera, 2005, we apply the seemingly unrelated regressions (SUR) technique to our DL model in order to obtain individual estimates while controlling for cross-country correlations.¹⁹ Results are shown in Table 1. The fit of the model is acceptable for most countries: the R-squared coefficient exceeds 0.4 for all apart from Germany, Italy, Norway and Japan (where it equals respectively 0.17, 0.25, 0.25 and 0.32). In four countries (Canada, Denmark, Finland, Sweden), the model explains more than 60% of unemployment changes. For all countries the F-tests on the whole model reject the null hypothesis of all coefficients different from zero.

The estimated coefficients vary considerably across countries: most cumulated coefficients range from -0.21 to -0.43, but in two cases these are smaller in absolute value (Japan, -0.08; Iceland - 0.17), thus indicating lesser sensitivity of unemployment to output changes, while in four countries sensitivity is rather high (Belgium, Denmark, Sweden, Spain, where the coefficients range from -0.51 to -0.66).

Table 1. Okun's law with U3. SUR estimates by country					
Country	Variable	Coefficient	Standard Errors	Cumulated coeff.	R ²
Belgium	gy_t	-0.27***	(.0395)	0.54	0.53
	gy_{t-1}	-0.27***	(.0390)		
	$const$	0.92***	(.1327)		
Canada	gy_t	-0.31***	(.0239)	0.39	0.77
	gy_{t-1}	-0.08***	(.0224)		
	$const$	0.77***	(.0843)		
Denmark	gy_t	-0.39***	(.0347)	0.53	0.74
	gy_{t-1}	-0.14***	(.0337)		
	$const$	0.85***	(.1051)		

¹⁸ Specifically, we use the Wooldridge test for panel data to detect the presence of autocorrelation in the residuals, the modified Wald test for within group heteroscedasticity and the Breusch-Pagan LM test of independence to check for cross-section correlation.

¹⁹ SUR is a GLS-type estimation that takes account of cross-equation correlations to get more efficiency, but it does not correct standard errors for heteroscedasticity or autocorrelation.

Finland	gy_t	-0.23***	(.0359)	0.41	0.66
	gy_{t-1}	-0.18***	(.0356)		
	$const$	0.85***	(.1637)		
France	gy_t	-0.24***	(.0356)	0.34	0.49
	gy_{t-1}	-0.10***	(.0353)		
	$const$	0.63***	(.0981)		
Germany	gy_t	-0.15***	(.0365)	0.21	0.17
	gy_{t-1}	-0.06*	(.0381)		
	$const$	0.26	(.1375)		
Greece	gy_t	-0.21***	(.0404)	0.38	0.52
	gy_{t-1}	-0.17***	(.0403)		
	$const$	0.67***	(.2148)		
Iceland	gy_t	-0.17***	(.0283)	0.17	0.47
	gy_{t-1}	0.004	(.0279)		
	$const$	0.53**	(.1516)		
Ireland	gy_t	-0.18***	(.0216)	0.28	0.57
	gy_{t-1}	-0.10*	(.0218)		
	$const$	1.15	(.2124)		
Italy	gy_t	-0.23***	(.0419)	0.23	0.25
	gy_{t-1}	-0.06	(.0404)		
	$const$	0.37***	(.1200)		
Japan	gy_t	-0.08***	(.0106)	0.08	0.25
	gy_{t-1}	-0.002	(.0105)		
	$const$	0.13**	(.0559)		
Netherlands	gy_t	-0.21***	(.0335)	0.38	0.57
	gy_{t-1}	-0.17***	(.0325)		
	$const$	0.68***	(.1102)		
Norway	gy_t	-0.15***	(.0401)	0.25	0.32
	gy_{t-1}	-0.10**	(.0381)		
	$const$	0.60***	(.1113)		
Portugal	gy_t	-0.25***	(.0485)	0.25	0.44
	gy_{t-1}	-0.06	(.0464)		
	$const$	0.57**	(.1754)		
Spain	gy_t	-0.66***	(.1143)	0.66	0.55
	gy_{t-1}	-0.008	(.1101)		
	$const$	1.5***	(.3252)		
Sweden	gy_t	-0.28***	(.0354)	0.51	0.67
	gy_{t-1}	-0.23***	(.0324)		
	$const$	1.2***	(.1385)		
Switzerland	gy_t	-0.16***	(.0308)	0.28	0.51
	gy_{t-1}	-0.12***	(.0315)		
	$const$	0.62***	(.0907)		
United Kingdom	gy_t	-0.23***	(.0394)	0.33	0.58

	gy_{t-1}	-0.10**	(.0388)		
	$const$	0.51***	(.1224)		
United States	gy_t	-0.32***	(.0359)	0.43	0.52
	gy_{t-1}	-0.09***	(.0328)		
	$const$	0.97***	(.1379)		
Legend: *p<.1; **p<.05; ***p<.001					

We next perform a panel estimate on the same dataset. By removing time-invariant country-specific effects, panel estimates allow assessing the net effect of the predictors within each country (i.e. considering only time-varying effects). We also introduce time fixed effects in order to control for unexpected variations or special events that may affect the outcome variable.²⁰ We thus obtain an estimate of the Okun coefficient across countries that is corrected by the individual characteristics and time-specific factors that may possibly distort individual country estimates (Freeman, 2001; Garavan, 2017).

Table 2 reports the results of the panel estimation of Okun's Law using a two-way fixed effects model. Due to the presence of serial correlation, heteroscedasticity and cross-section correlation in the data, we decided to run a Prais-Winsten regression, which calculates correlated-panel corrected standard error (PCSE) estimates for linear cross-sectional time-series models that are assumed to follow a panel-specific first-order autoregressive process.

Table 2. Panel estimates of Okun's Law in 19 OECD countries	
Variables	Coefficients
gy_t	-0.18*** (.0201)
gy_{t-1}	-0.11*** (.0199)
$const$	0.58*** (.1442)
<i>cumulated coefficient</i>	-0.29
R^2	0.51
N	665
Legend: *p<.1; **p<.05; ***p<.001	

²⁰ We check the correct specification of the model by running the Hausman specification test. We also run a Wald test on time effects and reject the null hypothesis that the coefficients for all years are jointly equal to zero.

The cumulated coefficient is -0.29. Overall, the relationship is robust and explains about half of observed unemployment changes on average across countries. The value of the Okun coefficient is in line with that observed in the literature.

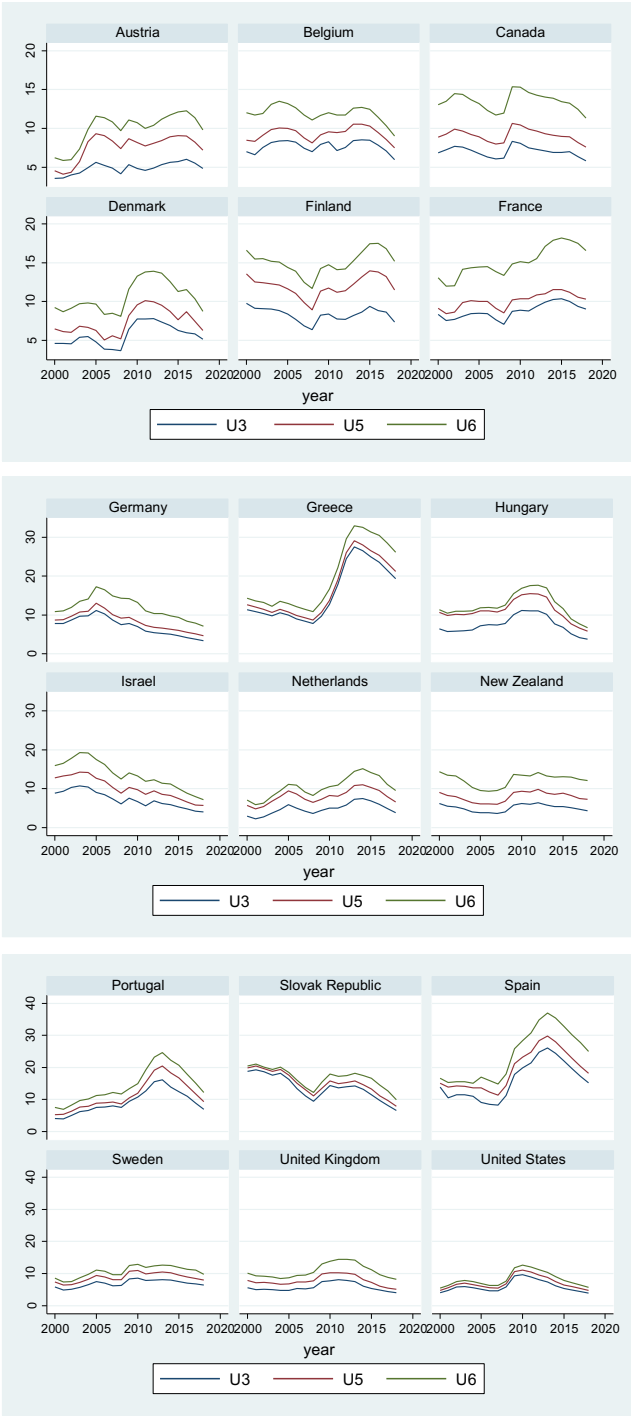
3.2. A panel estimate of Okun's law with U5 and U6

We now test Okun's law on some of the broader indicators of unemployment described in section 2.2 above. We select the two broadest unemployment rates, i.e. U5, which includes in the definition of unemployment discouraged and marginally attached workers, and U6, which also includes workers that work a less-than-their desired number of hours for lack of full-time jobs. For this estimate we use the second of our above-described datasets, which includes the time series of U5 and U6. We build U5 and U6 for the various countries on the basis of the definition and methodology used by the US Bureau of Labor Statistics (BLS); however, availability of data and the need to use homogeneous indicators across countries, together with the attempt to keep our panel as large as possible, imply that we have to introduce some adjustments. Particularly, in building the U6 indicator, we are forced to include only involuntary part-time workers and exclude short-time workers (which are computed separately by the OECD, but with data that end in 2015). This implies that our measure of U6 in this panel is underestimated.²¹

Figure 1 shows, for each country, the two broad indicators U5 and U6 compared with the official unemployment rate U3.

²¹ As is clearly shown by comparison of data for the USA of the two different sources, BLS and OECD, the number of part-time workers for economic reasons (source BLS) is greater than the number of involuntary part-time workers (source OECD), and the difference is represented by economic short-time workers (OECD source).

Figure 1. Official and broad unemployment indicators by country.



Source: our calculation on OECD data

The estimation model is the same as equation (1), where the broader indicators of unemployment are substituted, in turn, for the official unemployment rate. Table 3 presents the panel estimates

for each indicator. As before, we estimate a two-way fixed effects model through a Prais-Winsten panel regression in order to control for contemporaneous correlation.

Variables	Coefficients for U3	Coefficients for U5	Coefficients for U6
gy_t	-0.21*** (.0349)	-0.23*** (.0372)	-0.25*** (.0406)
gy_{t-1}	-0.17*** (.0343)	-0.20*** (.0365)	-0.26*** (.0399)
<i>const</i>	1.1*** (.2029)	1.2*** (.3158)	1.4*** (.3743)
Cumulated coefficient	-0.38	-0.43	-0.51
R^2	0.61	0.62	0.64
N	306	306	306

Legend: *p<.1; **p<.05; ***p<.001

The model explains about 60% of the variability of the dependent variable for each of the unemployment indicators. The cumulated coefficients, on average across the 18 countries in the panel, are -0.38, -0.43 and -0.51 for U3, U5 and U6 respectively. The results show clearly that Okun's law also holds when tested on the broad unemployment rates.

Table A3 in the Appendix reports country-by-country SUR estimates of the same equation.²² Notwithstanding the low number of observations for each country, the fit of the model is on average acceptable. The estimated coefficients vary considerably across countries. In the period 2000 to 2018, the same countries that show the lowest sensitivity of U3 to output growth, also show the lowest sensitivity of U5 and U6 (this applies to Austria, Germany and Sweden). Portugal and Spain show the highest sensitivity regardless of the indicator used.

4. Okun's Law in the United States with alternative indicators of labor underutilization

Our analysis in the previous section shows that Okun's law is a robust relationship that holds across different countries and different indicators of labor utilization. However, it also shows that the relationship has a certain cross-country variability, including with respect to proper specification. A deeper analysis of Okun's law thus requires focusing on single-country estimates, also with the aim of testing, using the alternative above-described unemployment indicators, the hypothesis we have originally put forward in Fontanari et al (2020), namely that the fit of Okun's law improves

²² Due to the limited number of observations for each country, in order to perform a SUR estimate we had to modify the model, inserting only the contemporaneous effect, and removing one country (we chose to remove the USA, for which we run separate estimates in section 4).

when allowing the Okun coefficient to assume different values for different ranges of unemployment. In this section we thus concentrate our attention on the United States.

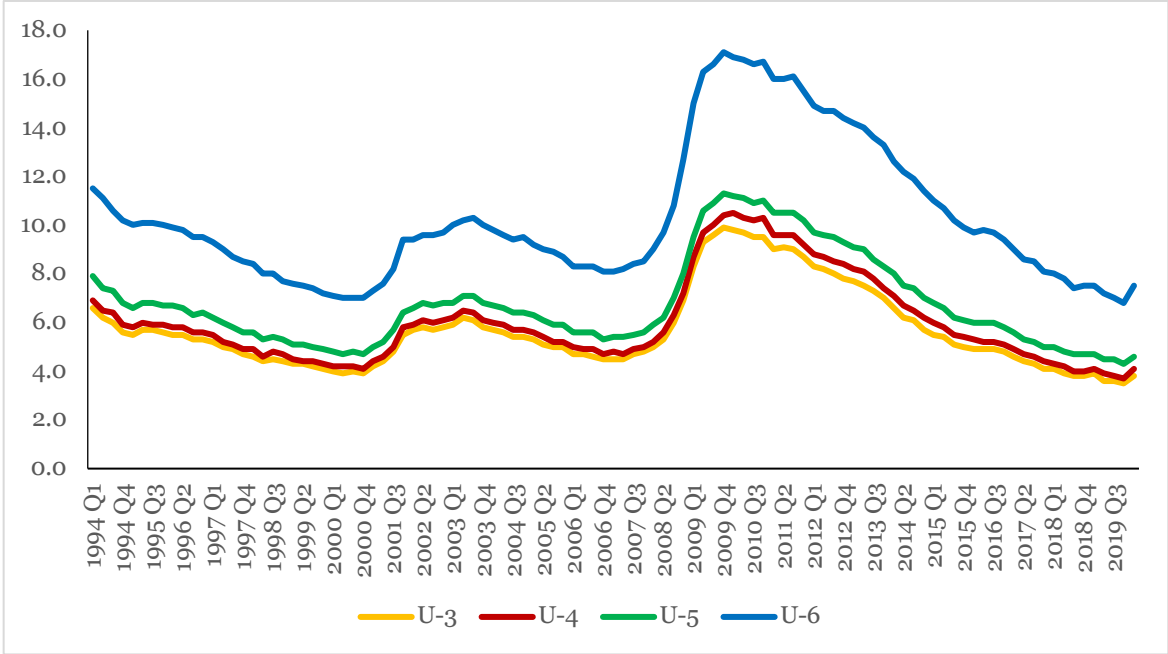
4.1. The broad unemployment rates and the different reactivity of unemployment according to its level

The current definitions of the various alternative rates of unemployment described in section 2.2 have been adopted by the US Bureau of Labor Statistics since 1994, which is the starting year of the time series of these indicators. These are available both at annual and infra-annual frequency, which allows us to test Okun's law on quarterly data over the 1994Q1-2020Q1 period.

It is worth, in the first place, to have a look at the time series of the three indicators provided by the BLS that are broader than the official unemployment rate, i.e. U4, which includes the discouraged workers, U5, which adds to the latter other marginally attached workers, and U6, the broadest indicator, which includes all persons employed part time because they cannot find as much work as they seek for economic reasons. Figure 2 shows that, similarly to the official unemployment rate, those rates rose to historically high values during the Great Recession (10, 11 and 17 percent respectively in 2009Q4), afterwards decreasing steadily until 2019Q4. In the last quarter of 2019, the three indicators reached the minimum level recorded in their available time series (3.7, 4.3 and 6.8 percent respectively). It is interesting to note that the sizeable difference between U5 and U6 is entirely made of the (economic) part-time workers, since the two indicators have the same denominator while their numerator only differs for this category of workers. This shows clearly the relevance of the phenomenon of part time work for economic reasons and its wide swings in the different phases of the cycle.²³

²³ It is worth noting that the indicator U6 that we use in this analysis on the USA, taken from the BLS database, has the correct definition and measurement and is not underestimated as in the case of our analysis in section 3.

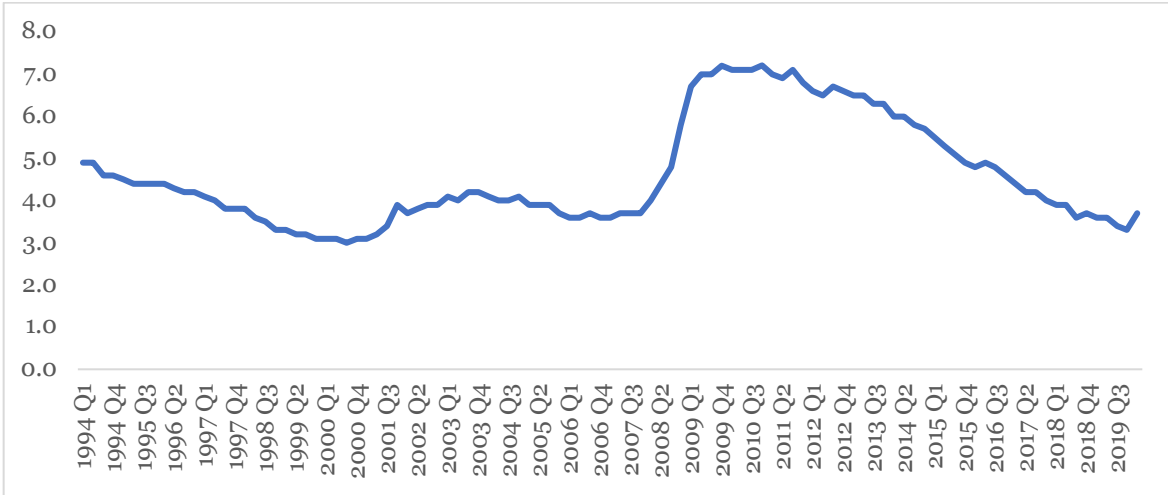
Figure 2. The official and the broad unemployment rates. USA, 1994Q1-2020Q1



Source: BLS data

An additional indicator, that may give an idea of how ‘hidden’ unemployment relates to the ‘open’ one, is the difference between U6 and U3 (see Fontanari et al 2019, Komlos 2019b). As shown in figure 3, this difference generally increases during recessions, and increased sharply during the Great Recession, to decline steadily in the following years at pre-crisis levels. At 3.3 percentage points in 2019Q4, this difference has remained slightly above the lowest value recorded in the peak of activity of 1999-2000 (3 percentage points in 2000Q3).

Figure 3. Difference between U6 and U3. USA, 1994Q1-2020Q1



Source: our calculation on BLS data.

As we did in our cross-country analysis in section 3, we test Okun’s law with both the U5 and U6 unemployment rates and compare the results with the traditional version of Okun’s law with the official rate of unemployment U3, tested on the same time-period. As in section 3 and in our previous work (Fontanari et al 2020), and for the same reasons, we use the difference specification, treating in each regression the various unemployment indicators as dependent variables and output growth as the independent one, and allowing for contemporaneous as well as lagged effects of output growth on the rates of unemployment. We select the optimal number of lags for both dependent and independent variables through various information criteria and find that two lagged terms have to be introduced for output growth.

We test here the same hypothesis advanced in Fontanari et al. (2020), that the effect of output growth on unemployment varies in intensity according to the level of the latter. We apply this hypothesis to each of the (narrow and broad) unemployment rates identifying in each case three ranges of values, which we label as ‘low’, ‘medium’ and ‘high’ values of the indicator.²⁴ The intuition is that changes in unemployment are more pronounced in response to output growth, when underutilization is high, while for very low rates of unemployment (high levels of activity) much of output growth is accommodated by productivity growth. Table 4 shows the limit values of the ranges for the three indicators considered.

Table 4. The unemployment ranges			
Ranges	Limit values for U3	Limit values for U5	Limit values for U6
u_{LOW}	$(u_t \leq 4.8)$	$(u_t \leq 5.7)$	$(u_t < 8.9)$
u_{MEDIUM}	$(4.8 < u_t \leq 5.7)$	$(5.8 \leq u_t \leq 6.8)$	$(8.9 \leq u_t \leq 10.3)$
u_{HIGH}	$(u_t \geq 5.8)$	$(u_t \geq 6.9)$	$(u_t > 10.3)$

Having checked both the dependent and independent variables for stationarity, we take into account the presence of a structural break whose date varies according to the indicator. When the break intercept dummy is fitted into the model together with the ranges of unemployment, the residuals are homoscedastic and not auto-correlated. We therefore run an OLS estimation. Only in the case of U6 do we use robust errors for heteroscedasticity.

The model equation is the following:

²⁴ The limit values for the three ranges have been chosen on the criterion that each of the three ranges contains one third of the observed values.

$$\Delta u_t = \delta_l L + \delta_m M + \delta_h H + \delta_{break} + \delta_{l1} g y_t^L + \delta_{m1} g y_t^M + \delta_{h1} g y_t^H + \delta_{l2} g y_{t-1}^L + \delta_{m2} g y_{t-1}^M + \delta_{h2} g y_{t-1}^H + \delta_{l3} g y_{t-2}^L + \delta_{m3} g y_{t-2}^M + \delta_{h3} g y_{t-2}^H + e_t$$

Table 5 reports the results of the regressions with U3, U5 and U6.

Table 5. Okun's law with different unemployment indicators. USA, 1994Q1-2020Q1									
	Dependent variable: U3			Dependent variable: U5			Dependent variable: U6		
	u_{LOW}	u_{MEDIUM}	u_{HIGH}	u_{LOW}	u_{MEDIUM}	u_{HIGH}	u_{LOW}	u_{MEDIUM}	u_{HIGH}
gy_t	-0.16*** (0.0481)	-0.12* (0.0622)	-0.16*** (0.0410)	-0.19*** (0.0532)	-0.16** (0.0679)	-0.22*** (0.0452)	-0.32*** (0.0827)	-0.19** (0.0832)	-0.41*** (0.1008)
gy_{t-1}	-0.09 (0.0526)	-0.23*** (0.0594)	-0.24*** (0.0447)	-0.09 (0.0574)	-0.23*** (0.0618)	-0.22*** (0.0511)	-0.16*** (0.0517)	-0.41*** (0.1193)	-0.38*** (0.0836)
gy_{t-2}	-0.13** (0.0562)	-0.10 (0.0601)	-0.18*** (0.0444)	-0.11* (0.0621)	-0.09 (0.0657)	-0.21*** (0.0478)	-0.04 (0.0684)	-0.13 (0.0982)	-0.18** (0.0796)
Constant terms	0.31*** (0.0500)		0.06 (0.0561)	0.32*** (0.0534)		0.08 (0.0611)	0.41*** (0.0780)	0.10 (0.1330)	0.19* (0.1027)
Cumulated coefficient	-0.28	-0.35	-0.58	-0.30	-0.39	-0.66	-0.48	-0.60	-0.97
<i>Break</i>	-0.23*** (0.0340)			-0.26*** (0.0380)			-0.36*** (0.0539)		
Obs.	102			102			102		
<i>Adj - R²</i>	0.72			0.72			0.75		

The fit of the three regressions is very good, with the R-squared coefficients greater than 0.7 in all cases. Compared to our precedent estimation with U3 on a different time-span (see Fontanari et al, 2020), all our previous results are qualitatively confirmed: the signs of the coefficients, the significance of the three ranges of unemployment and their ordering are as expected. Our intuition that higher levels of unemployment imply greater unemployment changes in response to output growth is confirmed on the broader indicators.²⁵

There is little doubt that the alternative unemployment indicators are better suited to represent the extent of labor underutilization than the official unemployment rate U3. The fact that they show a reliable and fairly stable relation to output growth, and also the fact that they are produced on a regular basis by the BLS, also makes them a suitable basis for estimating potential output, an exercise we will undertake in section 5. Before proceeding to that, however, we should note that these indicators, even though they are more accurate than U3, might still not be entirely adequate for assessing labor slack. This suspicion arises from inspection of the data referring to the late

²⁵ We run a Wald test to check for significance of the differences in coefficients both for contemporaneous and each lagged term separately and for the cumulated coefficients. According to our tests, the difference between the coefficients of the three groups is statistically significant only for the cumulated effect. We also run LR-test and F-test in order to compare models and we reject the null hypothesis that the restricted model is nested in the non-restricted one (see tables A4 in the appendix for details).

2010s. As noticed above, in the last quarter of 2019 all such indicators reached historically minimum levels in the USA, lower than those recorded in the peak of activity of the late 1990s, actually the lowest in the available time series. But as we remarked in section 2.1 and in Fontanari et al 2019, the growth in activity during the long recovery of the 2010s has been comparatively disappointing by comparison with the strong expansion of 1997-2000. One can infer from the very slow growth of wages that the US economy was likely still below the full employment ceiling during the last part of the 2010s, notwithstanding the very low official unemployment rates and the historically low values of U5 and U6.

Two different explanations of this are possible: the first one, that a condition of nearly full employment was actually realized, but with peculiar characteristics, in terms of a high polarization of the labor market, widespread presence of low-wage and low-productivity jobs (Storm 2017), and generalized conditions of weakness in the bargaining position of workers, which could account for the muted response of wages to labor market tightness; the second one, that – without downplaying the role of these factors – underutilization of labor was actually still there, although in a form not easily detected even by the broad unemployment indicators. In the next section, we thus explore the possibility that different indicators may better account for labor underutilization.

4.2. Underemployment and time-related unemployment

One limit of the broad unemployment indicators is that, being based on headcounts, they abstract from variations in hours worked, and may thus fail to capture correctly the extent of actual labor utilization/underutilization (Faberman et al. 2020). We review here some alternative indicators proposed in the literature, which focus on the idea that labor underutilization should be properly defined and measured in terms of working time. Such indicators, in other words, try to measure time-related unemployment and focus on the notion of ‘underemployment’, i.e. the condition of those employed workers that work fewer hours than desired (a condition also normally associated with less than sufficient labor incomes).

The first of such indicators is a time-related unemployment measure proposed by Bell and Blanchflower in a series of works starting from Bell and Blanchflower (2011). In Bell and Blanchflower (2018) the indicator, originally proposed for the UK, is applied to other 25 European countries.²⁶ The authors criticize the standard neoclassical account of the labor market, on the ground that workers are rarely free to choose the optimal amount of hours since working times are essentially a choice of the employers. The indicator aims to capture the idea that many of the part-time workers, including some of those that *voluntarily* chose part-time jobs, would be willing to increase the number of hours worked, while, on the other hand, some of the full-time workers would prefer to reduce their working time. The indicator is built first by converting the official unemployment rate into a measure expressed in terms of hours rather than persons, by multiplying

²⁶ The two authors originally built the indicator based on data of the UK Labour Force Survey (UKLFS); in Bell and Blanchflower (2018) this is extended to other European countries through data of the European Labor Force Surveys (EULFS). The kind of data the authors use are lacking for the USA.

each argument of the unemployment rate by the average number of hours worked annually by the employed,²⁷ then adding the overall number of additional hours that the part-time workers would be prepared to work, and then subtracting the overall desired reduction in the number of hours (which measures ‘overemployment’).

The index is thus defined as follows:

$$u_v = \frac{U\bar{h} + \sum_k \tilde{h}_k^U - \sum_j \tilde{h}_j^O}{U\bar{h} + \sum_i^N h_i}$$

Where \bar{h} is the average number of hours worked, h_k^U is the sum of preferred additional hours and h_j^O is the preferred reduction in hours. The denominator is the amount of ‘unemployed hours’ plus the number of total hours worked by employees.

The index is thus a sort of time-related index of workers’ dissatisfaction. Empirically, it reveals the presence of underemployment also for those European countries that showed very low unemployment rates in the late 2010s, and, according to the authors, it is better correlated with the behavior of wages.

In the same vein, Faberman et al (2020) develop a measure of labor underutilization (that they call the Aggregate Hours Gap) for the USA, based on the notion that desired work hours reflect the potential labor supply of individuals, so that any gap between desired and actual hours worked reflects some degree of labor slack. They combine the information of the Current Population Survey of the Bureau of Labor Statistics with information derived from the Job Search supplement to the Survey of Consumer Expectations (SCE) administered by the Federal Reserve Bank of New York. This is a survey conducted annually on a sample of about 1200 individuals, which offers information on a broad set of variables related to labor status and demographic characteristics, thus allowing a detailed analysis of the variability of the hours gap across different groups of the population. The authors note that their measure “exhibits notable deviations from the unemployment rate following the Great Recession and (...) performs relatively well in accounting for wage fluctuations over time” (Faberman et al 2020, p. 2)

A different strategy is used by Komlos (2019a), who builds an indicator which is not based on dedicated surveys on hours worked but on interpretation and correction of the US official unemployment rate produced by the BLS, with the aim of transforming it into a more accurate measure of underutilization. His indicator, which he labels the ‘real’ rate of unemployment, is obtained first of all by taking into account the military (both in employment and in the labor force),

²⁷ Bell’s and Blanchflower’s indicator is based on the assumption that the unemployed desire a job the working time of which equals the average economy-wide working time of the employees. The authors justify this assumption on the basis of the average forecast on the part of the unemployed about hours worked in future job.

which the BLS excludes from such aggregates, then by adding to the unemployed all those who declare they want to work but have not looked for a job during the previous month,²⁸ and finally by taking into account the phenomenon of underemployment. In this latter respect, Komlos (2019a) calculates the ratio between the average hours worked by part-time workers and the average hours worked by full-time workers, finding it equal to 62.5 percent in March 2019. Each voluntary part-time worker is then considered as the equivalent of 0.625 of a full-time worker and as participating in the same proportion to the labor force, while involuntary part-time workers are regarded as entirely participating to the labor force and as employed in the proportion of 0.625 each. Thus, 62.5% of the involuntary part-time workers are included in the labor force as employed and 37.5% are included in the labor force as unemployed. With these corrections, the so-calculated ‘real’ unemployment rate proves higher than the official unemployment rate, on average, by 4.4 percentage points over the 1994-2019 period (Komlos, 2019a, p. 11).

Building on these suggestions on the literature, we try to construct alternative measures of time-related unemployment on the basis of BLS data. We build an approximate measure of ‘desired hours’ based on the difference between total average hours and average hours of part-time workers for economic reasons²⁹ (using Komlos’s suggestion). We focus exclusively on underemployment while neglecting the opposite phenomenon of overemployment, lacking data on the latter.

We thus build the following index, which we label u_{DH} because it includes desired working hours:

$$u_{DH} = \frac{U\bar{h} + \sum h^D}{U\bar{h} + E\bar{h}}$$

where \bar{h} represents average hours worked by the employed, U and E are respectively the number of unemployed and employed, and h^D is the above-defined number of desired additional working hours. The numerator represents a broader measure of unemployment because it also takes into account underemployment, in the form of the extra hours that some of the employed would work if economic conditions allowed it.

A second index we obtain from this one. Following Komlos’s (2019a) suggestion, we add to both numerator and denominator the number of those who want a job but have not looked for work during the previous month, again expressed in terms of potential hours of work. Accordingly, we label the index u_{WJ} .

$$u_{WJ} = \frac{U\bar{h} + \sum h^D + WJ\bar{h}}{U\bar{h} + E\bar{h} + WJ\bar{h}}$$

²⁸ According to Komlos (2019a), the BLS’s definition of ‘active search’ is unduly restrictive by international standards since it does not consider as actively unemployed those people that look for jobs “simply reading want ads”.

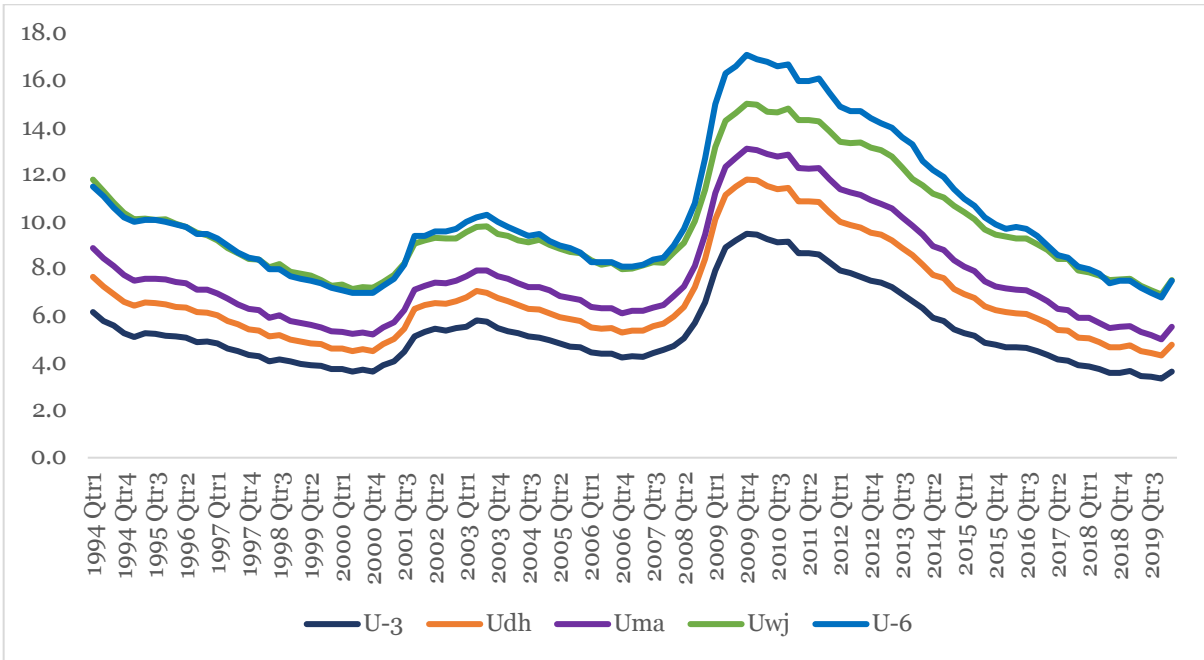
²⁹ To do this we annualize weekly hours of involuntary part-time workers multiplying the annualized average hours by the ratio between part-time weekly hours and total weekly hours (all data come from BLS).

Finally, a third indicator focuses, within the ‘want a job’ category, more narrowly on those that have looked for work in the last 12 months, which in the BLS statistics are defined as ‘marginally attached workers’ (see section 2.2 above). Our third indicator u_{MA} has thus the following formula:

$$u_{MA} = \frac{U\bar{h} + \sum h^D + MA\bar{h}}{U\bar{h} + E\bar{h} + MA\bar{h}}$$

Figure 4 plots the time-series of the three indicators so constructed and compares them to U3 and U6.

Figure 4. U3, U6 and three alternative indicators of time-related unemployment. USA, 1994Q1-2020Q1



Source: our calculation on BLS data

As is apparent, in the period considered U3 and U6 form a sort of corridor of possible labor underutilization measures within which the other proposed indicators capture various aspects of underutilization. For the way it is defined and measured, the U6 rate remains the broadest measure of underemployment because it counts involuntary part-time workers as totally unemployed (in the broad sense), instead of partially unemployed for a fraction of their full (potential) work time.

4.3. Employment and standardized hours worked

In our view, the most important suggestion that comes from the type of analyses reviewed in the previous section concerns the relevance of the phenomenon of underemployment, understood as a

situation in which the potential working time of both the non-employed and of some the employed is underutilized. Based on this kind of intuition, we look for a better indicator of labor underutilization based on hours rather than headcount. However, instead of computing labor underutilization in terms of forgone hours worked, with all the problems of what categories exactly should be considered in the computation and with what potential working time, we propose the more straightforward route of calculating an index of labor utilization on the basis of the total number of hours actually worked in the economy. Hours worked are in fact the most direct measure of the quantitative use of the labor input and are directly related to the amount of production. An indicator based on hours worked is likely to register correctly both the diminished utilization of the labor input associated with lower participation and the changes in the actual average working time of the employed.

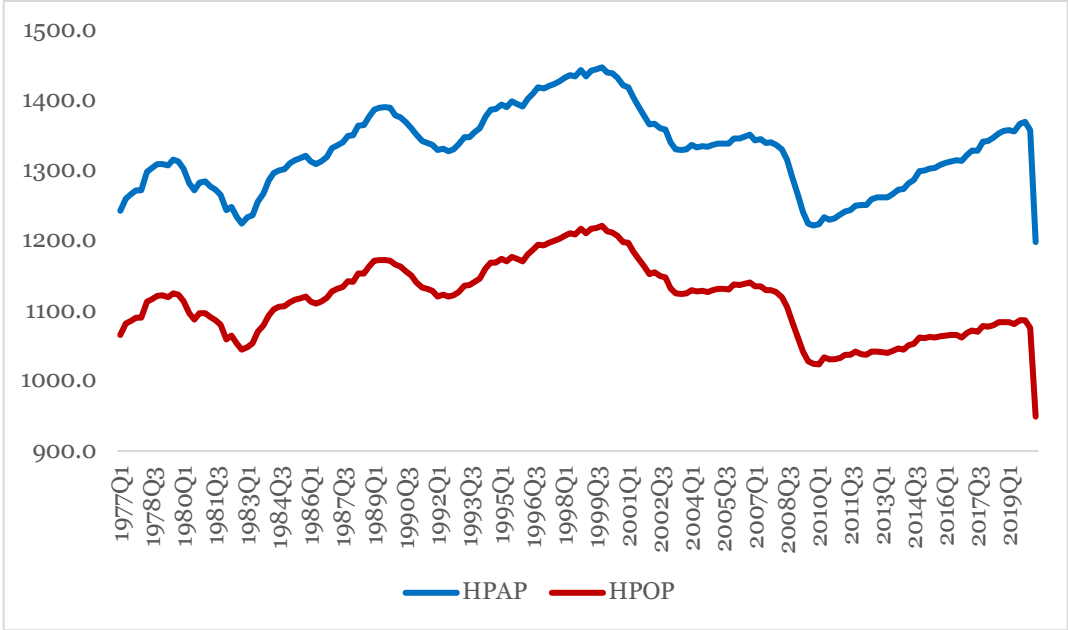
To make intertemporal comparisons possible, we build a ‘standardized’ indicator of hours worked by expressing the total number of hours worked in the economy as a ratio of the working-age population, so that the indicator is not influenced by the changing size of the economy. The working age population may be defined, alternatively, either as the population between the ages of 16-64 or as the whole population of 16 years or more. Since there are many workers in the 65+ range and employment in this age range has been increasing steadily in the last couple of decades, the second definition looks more appropriate. However, given that there has been some change in the age composition of the population in favor of older cohorts, and since the latter on average work less or fewer hours, the indicator might be biased. So we opt for using both and build two different indicators, i.e. (annualized) hours worked divided by “prime working age” population (HPAP) and (annualized) hours worked divided by the population with 16 years or more (HPOP).

$$HPAP = \frac{\text{total hours worked}}{\text{population 16 – 64}}$$

$$HPOP = \frac{\text{total hours worked}}{\text{population 16 +}}$$

Figure 5 depicts the evolution of the two indicators over the period from 1977 to the second quarter of 2020:

Figure 5. Standardized hours worked in the USA (1977Q1-2020Q2)



Source: our calculation on BLS and FRED data. See text for definition of the indicators

Even without mentioning the outstanding datum of the second quarter of 2020, the distance is striking between the recent levels of the indicators (those recorded in the last quarters of 2019 or in 2020Q1) and their peak values (recorded in 1999Q4 for both series), despite the steady growth in hours worked during the 2010s. In 2020Q1, the annualized hours worked divided by prime working age population (HPAP) are 90 less than at 1999Q4, which amounts to a 6 percent gap. As regards the other indicator, i.e. the annualized hours worked divided by the whole population with 16 years and more (HPOP), the gap with respect to its own peak value was even bigger: 146 in absolute terms, amounting to a 12 percent gap.

It is thus apparent that the *quantitative* use of the labor input in the US economy was not at its maximum even at the end of the long recovery of the 2010s, notwithstanding the fact that the various unemployment rates have reached such low levels as those reported above.

4.3.1 Okun’s law with standardized hours worked: estimates for the USA and in a panel of countries

We now test Okun’s law with our indicator³⁰. Although we have focused our analysis in this section on the United States, we think it worth to try our indicator, preliminarily, on the same panel of countries that we have used in section 3.1 to estimate the traditional version of Okun’s law. We thus refer to our first dataset extracted from AMECO, which covers annual data for 19 OECD countries over the period 1983-2019 (see section 3.1 for details on the dataset). As dependent

³⁰ For a test of Okun’s law with an indicator of labor input that also takes into account working time, see Busetta and Corso (2012), who use (for Italy) a measure of employment in terms of full-time equivalent working units.

variable, we use our indicator HPAP, i.e. hours worked divided by the working population of 16-64 age. Although this indicator may imply a more conservative assessment of labor underutilization than the other one, it has the advantage of eliminating any effects of population aging. We regress the first difference in this indicator on output growth, considering both contemporaneous and lagged effects.³¹

As in section 3.1, we estimate the relationship both individually by country and in a pooled sample. Table A5 in the Appendix reports the SUR estimates by country. The fit of the model is generally good and, on average, it explains 50% of the variability of the change in HPAP (a result similar to that obtained in section 3.1 for unemployment).

The cumulated coefficients range between 0.6 and 0.98, but three are lower (France 0.47, Japan 0.29, Germany 0.44), and two higher (Spain 1.01, United States 1.15). Spain, Canada, Denmark and the United States are among the countries with larger coefficients (taken in absolute value) both in these estimates and in the traditional version of Okun's Law (see section 3.1), thus showing a greater sensitivity of labor utilization/underutilization to output changes. In contrast, the coefficients for Germany and Japan are among the lowest in both models.

The panel estimate is presented in Table 6. We use the same estimator as in section 3.1, i.e. a Prais-Winsten regression of a two-way fixed effect model. The sign and significance of coefficients and the general good fit of the model confirm that our indicator of standardized hours worked is at least as strongly correlated to output growth as the unemployment rate.

Table 6. Okun's law with standardized hours worked in a panel of countries	
Variables	Coefficients
gy_t	0.48*** (.0502)
gy_{t-1}	0.17*** (.0494)
<i>const</i>	-0.01*** (.2477)
<i>cumulated coefficient</i>	0.65
R^2	0.55
N	665
Legend: *p<.1; **p<.05; ***p<.001	

³¹ We check and confirm stationarity of all variables using the Levin-Lin-Chu test for unit roots (see also section 3.1 above). We also find autocorrelation of residuals, heteroscedasticity and correlation across panels, by means of, respectively, the Wooldridge test for panel data, the modified Wald test for within group heteroscedasticity and the Breusch-Pagan LM test of independence. The optimal number of lags according to information criteria varies across countries between 0 and 2; however, for most countries the optimal number is 1 and this we have introduced in our model.

In search of the best country-specific model of estimation, we focus now on the estimate of Okun’s law with our indicator for the United States.

For the USA, our dataset covers quarterly data over the period 1977Q1-2020Q1. The data sources are the Bureau of Economic Analysis (BEA) for output, the BLS for hours worked, and the FRED database of the Federal Reserve Bank of St. Louis for population. Our dependent variable is the first difference in HPAP. We regress it on the rate of growth of output, contemporaneous and lagged up to two periods. Both dependent and independent variables are stationary. The disturbances follow an autoregressive process; therefore, we estimate an ARMAX (5,0) model which includes the contemporaneous and lagged independent variable (only significant lagged terms are included according to the parsimony criterion). The selection of the best model and the optimal number of lags for both variables has been made by comparing different information criteria. We also introduce an intercept dummy variable to take into account the presence of a structural break in the first quarter of 2010. Results are shown in Table 7.

Table 7. Okun’s law with standardized hours worked (HPAP) in the USA		
Model Equation: $\Delta hpap_t = \alpha + \gamma_1 \Delta hpap_{t-2} + \gamma_2 \Delta hpap_{t-5} + \beta_0 gy_t + \beta_1 gy_{t-1} + \beta_2 gy_{t-2} + \varepsilon_t$		
Variables	Coefficients	Standard errors
<i>gy_t</i>	0.64***	(.0535)
<i>gy_{t-1}</i>	0.16***	(.0523)
<i>gy_{t-2}</i>	0.19***	(.0509)
<i>Δhpap_{t-2}</i>	0.21***	(.0753)
<i>Δhpap_{t-5}</i>	0.32***	(.0743)
Constant	-0.73***	(.1001)
Break 2010Q1	0.57***	(.1799)
Cumulated coeff.	0.99	
Observations	170	
<i>AdjR²</i>	0.62	
Legend: *p<.1; **p<.05; ***p<.001		

The coefficients are significant, and the fit of the model is generally good (according to the R-squared, our model predicts about 60 percent of the variability of the indicator). Due to its property of representing more accurately, with respect to the unemployment rate, the size of the labor input,

the standardized hours worked indicator can thus be used both in the assessment of labor slack and as a basis for the calculation of potential output.

5. Potential output and the output gaps in the USA according to the Updated Okun Method

5.1. The late 2010s and the hidden slack

We will now use our results in sections 4.1 and 4.3.1 to estimate potential output and the output gaps. We focus exclusively on the United States, because calculation of potential output with the Updated Okun Method requires a country-specific analysis of Okun's law, which we did, in this paper, only for the USA. We believe however that the significance and potential fruitfulness of our method is much wider. Our panel analyses, after all, represent general confirmation of the fit of the law across different countries and different indicators, on which basis we will undertake detailed country-specific estimates in future work.

The meaning and characteristics of our Updated Okun Method are extensively treated in Fontanari et al (2020), to which we refer. We will only recall here that the method consists generally in measuring the amount of labor slack through an appropriate indicator (traditionally, the unemployment rate), in establishing a fairly reliable empirical relationship between the changes in such indicator and output growth (Okun's law), and in setting an appropriate target value for the indicator, so as to calculate, in each given situation, the amount of output change that would be required to bring the indicator at target value, and to infer from the procedure the level of potential output. This is therefore calculated as the output that would be obtained in the situation of full (or almost full) use of the labor input. This implies that by construction the Okun method only deals with the question of the *quantitative* use of the labor input, without considering the different question of the possibility of identifying further margins of output expansion through a more efficient use of the labor input.

We saw that Okun's law is quite a robust relation, which holds also with alternative indicators of labor underutilization and on different time spans. We will first calculate potential output by relying on our estimates of Okun's law with the U5 and U6 indicators and compare it to that obtained through U3. Afterwards, we will try the same exercise with our indicator of standardized hours worked.

Regarding U5 and U6, the important issue is which levels of U5 and U6 to select as desirable targets. Actually, such a problem also arises in connection with U3, the official unemployment rate, since there is no general consensus on which specific rate of unemployment can be regarded as representing 'full employment' nor any general rule to calculate how high is, in any given situation, the supposedly 'irreducible' unemployment corresponding to unavoidable frictions. As already remarked, the recent experience of the late 2010s has confirmed the difficulty of such

definition. One possible criterion is to target the minimum recorded level of unemployment in the chosen time series, which in our case, starting from 1994Q1, is the 3.5 percent unemployment rate realized in the last quarter of 2019. Yet, as recalled above, the period in which such a minimum was recorded cannot be properly defined as a period of full employment, judging by the comparatively low rate of labor force participation and the sluggish growth of wages. We thus set two different targets for U3, i.e. both 3.5 percent and a lower target, 3.2 percent, which may perhaps better approximate full employment, although this is likely to be a conservative measure too.³²

As regards U6 and U5, the historical record (from 1994Q1 on) gives, respectively, 6.8 and 4.3 percent as the lowest observed values (again in 2019Q4). Even more than in the U3 case, neither objective nor conventional criteria exist to determine the level of those rates that is socially desirable and may be taken as approximately corresponding to full employment. Moreover, the same caveat as for U3 applies also to U5 and U6: the historically low levels of 2019Q4 in all probability did not represent full employment. The time series of our indicator of the standardized hours worked confirms that much. To such minimum-value targets determined historically, we also add, as in the case of U3, more ambitious targets, hopefully closer – although still imperfectly – to a better definition of full employment.

One possibility to determine such targets is to rely on the difference between the broader indicators and U3, and define as a target a minimum for this difference (since such difference historically shrinks in high activity phases). The minimum differences between U6-U3 and U5-U3 observed in our sample are 3 and 0.8 percentage points respectively. It is worth noting that the presence of some individuals that find themselves in the condition of being ‘marginally attached’ to the labor market (or, more likely, a bigger gross share of individuals fluctuating between the different statuses in or out the workforce) – a category that accounts for the difference between U5 and U3 – is hardly removable in its entirety, even in the most prosperous periods. On the contrary, we may suppose that a true full employment target level would be characterized by a drastic reduction in the numbers of those who work part time involuntarily, for lack of full-time jobs. These workers, who account fully for the difference between U6 and U5 and constitute a big share of the U6-U3 difference, are already active participants to the labor market and may easily be converted into full-time workers in periods of authentic prosperity. Even if to eliminate completely this form of underutilization may be regarded as unrealistic, still a sizeable reduction of it seems to be a straightforward attribute of the notion of full employment.

In setting ambitious targets for U5 and U6, therefore, we assume these as constituted, in the case of U5, by the ‘ambitious’ target for U3 augmented with the minimum historically observed

³² The literature of the last decades has given little or no attention to the question of the exact measurement of full employment, given the theoretical and practical dominance of the economics of the natural rate. In what represents one of the few exceptions, Robert Eisner (1998, p. 422) noted that the Humphrey–Hawkins Full Employment Act of 1978 was perfectly sensible in “setting a goal of 3 per cent ‘adult’ unemployment, corresponding to minimal friction and search unemployment, as the full employment target”.

difference between the two indicators (0.8 percentage points), while, in the case of U6, we also target the U6-U3 difference, with the objective to reduce it to 2.5 percentage points. In sum, we calculate potential output and output gaps first in terms of U3, targeting either 3.5 or 3.2 percent, then in terms of U5, targeting either 4.3 or 4.0, and finally in terms of U6, targeting either 6.8 or 5.7 percent.

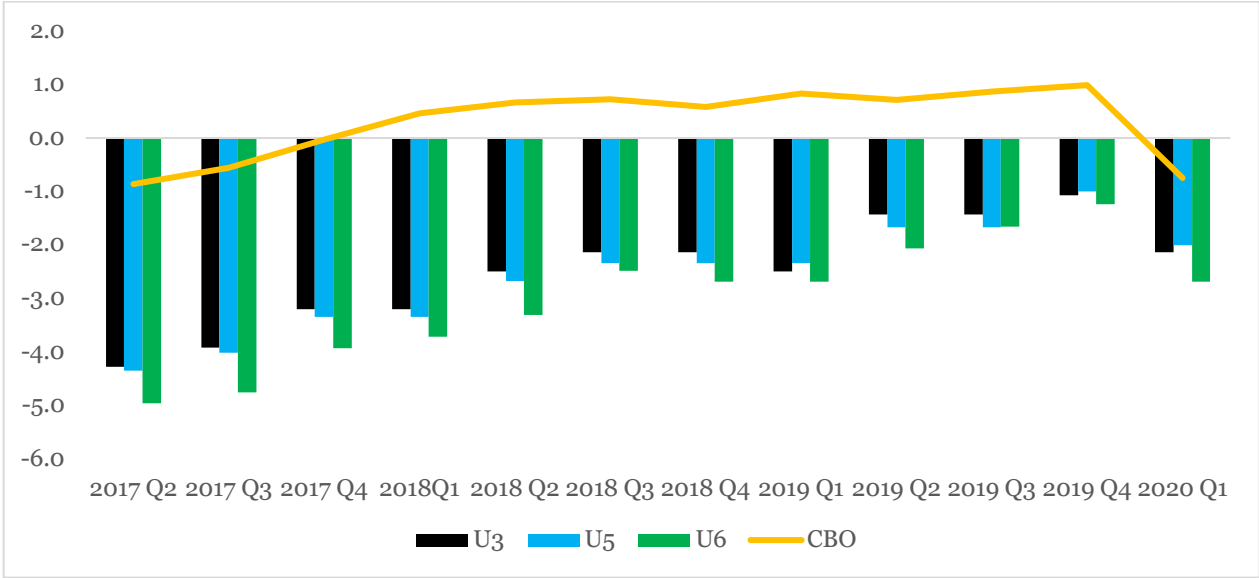
Table 8 shows the output gaps for the last three quarters of 2019 and the first quarter of 2020, which result from application of these different targets together with the Okun coefficients that we have found in the estimates reported in section 4.1 above.

Table 8. Output gaps obtained with different unemployment rates and different targets						
	Minimum-value targets			Minimum-difference targets		
	U3	U5	U6	U3	U5	U6
2019 Q2	-0,4	-0,7	-0,8	-1,4	-1,7	-3,1
2019 Q3	-0,4	-0,7	-0,4	-1,4	-1,7	-2,7
2019 Q4	0,0	0,0	0,0	-1,1	-1,0	-2,3
2020 Q1	-1,1	-1,0	-1,4	-2,1	-2,0	-3,7

It is quite obvious that the value of the target crucially affects the estimates of the gaps. This, however, depends essentially on the fact that no objective and univocal empirical definition of full-employment exists, but only conventional ones, so that multiple targets can be considered. One advantage of our method, we maintain, is the very strict (and very intelligible) link between the level of the socially chosen target and the quantitative assessment of the current position of the economy.

Figure 6 shows the time series of the output gaps over the last 3 years, calculated using the ‘minimum-difference’ targets. For comparison, the series of the output gaps estimated by the US Congressional Budget Office (CBO) is also shown in the figure.

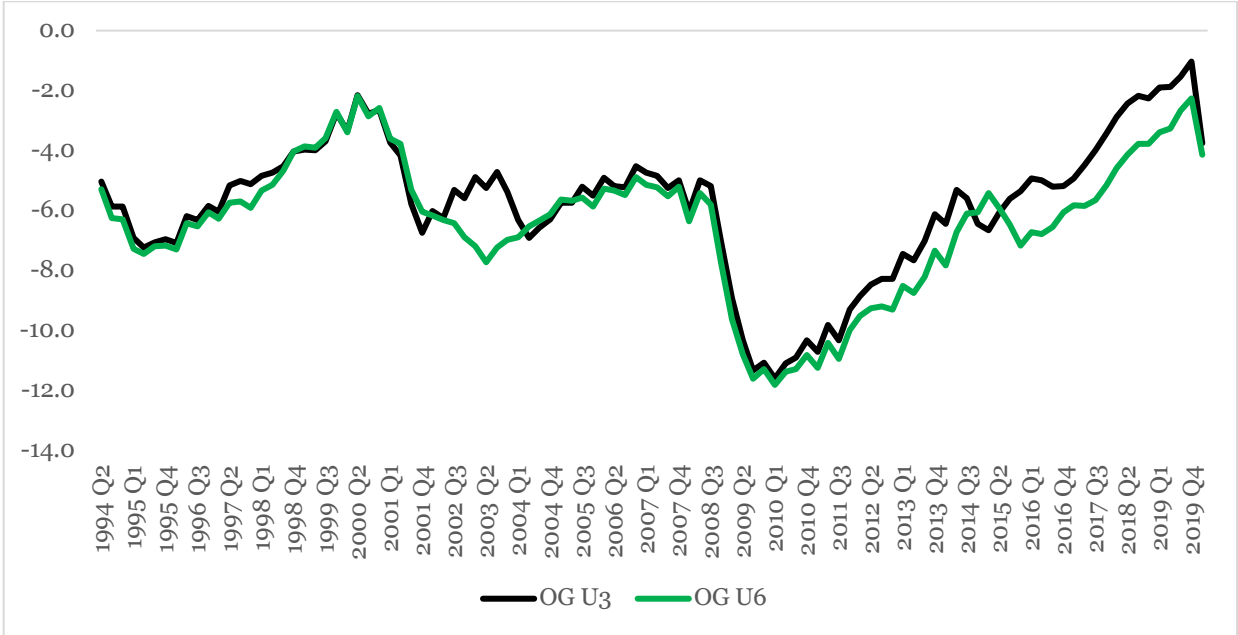
Figure 6. Output gaps obtained with the Updated Okun Method with different unemployment indicators (minimum-difference targets) compared with the output gaps estimated by CBO, 2017Q2-2020Q1



Source: our calculation on BEA and BLS data; CBO (2021)

Figure 7 extends our calculation of the output gaps to the whole 1994Q1-2020Q1 sample, focusing only on U3 and U6, with gaps (obtained through the Updated Okun Method) calculated on the minimum-difference targets. The output gaps in this figure are derived from smoothing the respective series of potential output (through a three-periods moving average).

Figure 7. Output gaps obtained with the Updated Okun Method with U3 and U6, minimum-difference targets, 1994Q1-2020Q1



Source: our calculation on BEA and BLS data.

The figure illustrates well, in our view, the fact that the 5.7 target that we selected for U6 is not unduly ambitious, if compared to the 3.2 target set for U3. Apart from a number of quarters in the early 2000s, the figure shows clearly that the output gaps obtained with either target (3.2 for U3 and 5.7 for U6) almost coincide in all the periods from 1994 up to the Great Recession and immediately after, while diverging only during the 2010s. To set the ‘ambitious’ 5.7 target for U6 seems thus a way to highlight, at least in part, the persistence of slack in the labor market that has characterized the recovery of the 2010s notwithstanding the historically low unemployment rates that have been reached at the end of the decade.

Finally, we calculate the output gaps also using our indicator of standardized hours worked. If in the case of the (narrow and broad) rates of unemployment we have the choice to set alternative targets, in this case the straightforward route seems to be that of calculating how much output the US economy would have produced at the end of the 2010s if the quantitative use of the labor input, in terms of our HPAP indicator, had been the same as that realized at the peak of the late 1990s. Thus we set as target the maximum historical level of our indicator, corresponding to 1999Q4, and use the result of the estimate of Okun’s law with standardized hours worked illustrated in section 4.3. Results for the 8 quarters from 2018Q2 to 2020Q1 are shown in Table 9.

Table 9. Output Gaps obtained with the Standardized Hours Worked indicator (HPAP)	
	Maximum value target
2018Q2	-10,2
2018Q3	-9,4
2018Q4	-9,1
2019Q1	-9,0
2019Q2	-9,2
2019Q3	-8,1
2019Q4	-7,8
2020Q1	-9,1

As is apparent, the calculations based on hours worked instead of the unemployment indicators give a quite different picture of the situation of the US economy in the late 2010s. When judged from the point of view of hours worked, the extent of the use of the labor input in that period is considerably less impressive, and a sizeable gap appears with respect to other historical peaks of activity.

5.2. A tentative assessment of the ‘pandemic’ gaps

We try now to apply our method to the assessment of the output gaps opened up by the pandemic recession. For the reasons already illustrated, such assessment is of course tentative. We only offer estimates for the second and third quarters of 2020. Our strategy is to regard the pandemic recession as a probable outlier in the estimation of Okun’s law, given its peculiar characteristics. We thus apply the coefficients we have obtained in the estimates that exclude 2020Q2 and 2020Q3 to the calculation of the output gaps for the last two quarters.³³ Results are shown in Table 10, which also reports, for comparison, the output gaps for the same quarters as estimated by the CBO in its February 2021 Economic Projections.

Table 10. Output gaps in 2020Q2 and 2020Q3 with different indicators and different targets								
	Output gaps estimated by the CBO	Output gaps estimated with the Updated Okun Method						
		Minimum-value targets			Minimum-difference targets			Maximum-value target
		U3	U5	U6	U3	U5	U6	HPAP
2020 Q2	-10.1	-16.4	-15.2	-14.4	-16.9	-15.7	-15.5	-25.1
2020 Q3	-3.8	-9.2	-8.5	-8.0	-9.7	-9.0	-9.1	-16.9

³³ For consistency with our method and results, we use the coefficients of the ‘high unemployment’ range for each indicator.

It is worth noting that the CBO estimates a wide output gap for 2020Q2, actually wider than any of the gaps estimated by the same agency in the worst phases of the previous recessions, including the Great Recession. This depends both on the violence of the downturn caused by the pandemic and the assumption, implicit in the CBO's projections, that it has little affected the path of potential output. Yet our estimates reveal still wider gaps, though varying according to the indicator and the selected target (as already noted for the late 2010s, our indicator based on standardized hours worked is the one that reveals the biggest gaps). What is more relevant, if the slack should persist or be newly aggravated, this would not immediately pass on to our measures of potential output and would thus not produce the too early closure of the gaps that usually characterizes standard estimates.

6. Conclusions

Our analysis once again confirms the general validity of Okun's Law, which holds both across countries and with all the different indicators of labor utilization/underutilization that we have tested. The cross-country analyses reveal differences across countries and prompt detailed country-specific estimates, something that we have done in this paper for the United States.

Using the broad unemployment indicators U5 and U6, we have confirmed here the hypothesis first put forward in Fontanari et al (2020) that the fit of Okun's law improves when allowing the Okun coefficient to assume different values for different ranges of unemployment (however broadly understood). As an additional original contribution of this paper, we have proposed an indicator of labor utilization that we have labeled 'standardized hours worked', i.e. the total number of hours worked in the economy divided by the working age population. This indicator, which seems able to measure as accurately as possible the actual extent of (quantitative) labor utilization at different moments of time, is also well correlated to output changes, as our analysis of Okun's law with this indicator shows.

The output gaps we have calculated on the basis of our estimates of Okun's law vary across the different indicators and the different target values we have set for them, but on the whole reveal more slack than that calculated with standard methods. This is relevant both for a better retrospective evaluation of the expansionary phase of the 2010s, with all the puzzles it has proposed to interpreters and commentators, and for a more accurate assessment of the extent of the slack associated to the pandemic. Our method, we believe, is an antidote against the standard idea that potential output should be defined as the trend of actual output, an idea which produces systematic underestimation of the output gaps and premature closing of the gaps when a timid recovery follows a deep recession – which also implies, too often, a serious bias in favor of restrictive policies that may suffocate the recovery and create long-term damage to growth.

As we have tried to argue above, the variability in the measure of the gaps that our method produces is almost entirely due to the possibility of choosing different indicators to measure labor

slack and of setting different values for them as desirable policy targets. We regard this as an advantage of our method, since it transparently links the extent of the gap to the policy objective and allows for a focusing of the public debate on the merits of the various indicators and their ability to correctly measure the quantitative dimension of labor underutilization. A further advantage of our approach is the lack of any mechanical link between labor slack and inflation. As recent history has shown, this link is too erratic in the data to base on it harmful policies that purport to control inflation by curbing economic activity and producing unemployment. The disconnect observed in the data between unemployment and inflation should produce, in our view, not only a deep theoretical reconsideration of the Phillips-curve models, but also their abandonment as foundation for policy, where they showed clear failure.

If our Updated Okun Method thus offers a more correct assessment of the output gaps and a more effective basis for policy making, it by no means exhausts the question of the potential margins for output expansion, especially in a medium-long term perspective. By focusing exclusively on the quantitative dimension of labor underutilization, it is worth noting, it does not address the question of the *qualitative* dimension of underutilization, associated with the share of labor employed in low-productivity sectors and occupations and the general trend of labor productivity in recent times. Consideration of the productivity issue would however reinforce the idea that even at the end of the long expansionary phase of the 2010s the US economy had still sizeable margins for further expansion, based on a fuller quantitative and qualitative use of labor.

References

- Aaronson, S., Cajner, T., Fallick, B., Galbis-Reig, F., Smith, C., Wascher, W. (2014). Labor force participation: recent developments and future prospects. *Brookings Papers on Economic Activity*, No. 2, 197–275.
- Ball, L., Leigh, D., Loungani, P. (2017). Okun’s law: Fit at 50. *Journal of Money, Credit and Banking*, 49(7), 1413–1441.
- Bañbura, M., Bobeica, E. (2020). *Does the Phillips curve help to forecast euro area inflation?* European Central Bank, No. 2471
- Bell, D. N., Blanchflower, D. G. (2011). Underemployment in the UK in the Great Recession. *National Institute Economic Review*, 215(1), R23-R33.
- Bell, D. N., Blanchflower, D. G. (2018). Underemployment in the US and Europe. *National Bureau of Economic Research*, No. w24927.
- Blanchard, O., Leando, A., Zettelmeyer, J. (2020). Redesigning EU Fiscal Rules: From Rules to Standards. *72nd Economic Policy Panel Meeting*.
- Bregger, J. E., Haugen, S. E. (1995). BLS introduces new range of alternative unemployment measures. *Monthly Labor Review*, 118(10), 19-26.
- Brooks, R., Fortun, J. (2020). Eurozone Output Gaps and the COVID-19 Shock. *Intereconomics*, 5, 291-296.
- Bundesbank (2014). On the reliability of international organisations’ estimates of the output gap. *Deutsche Bundesbank Monthly Report (April)*, 13-35.
- Busetta, G., Corso, D. (2012). Labor Productivity and Okun’s Law: An Empirical Application to Italian Regional Panel Data. *Rivista italiana degli economisti*, 17 (2), 279–298.
- Cazes, S., Verick, S. (2011). What has happened to Okun’s law in the United States and Europe? Insights from the global financial and economic crisis and long-term trends. The global crisis: Causes, responses and challenges. *International Labour Office*.
- CBO (2020). An Update to the Economic Outlook: 2020 to 2030, July. *Congressional Budget Office*.
- CBO (2021). An Overview of the Economic Outlook: 2021 to 2031, February. *Congressional Budget Office*.
- Del Negro, M., Lenza, M., Primiceri, G. E., Tambalotti, A. (2020). *What’s up with the Phillips Curve?* National Bureau of Economic Research, No. 27003.

- EC, (2020a). European Economic Forecast: Spring 2020. *European Commission Institutional Paper*, No. 125.
- EC (2020b). European Economic Forecast: Autumn 2020. *European Commission Institutional Paper*, No. 136.
- EC (2020c). The 2021 Ageing Report. Underlying Assumptions & Projection Methodologies. *European Commission Institutional Paper*, No. 142.
- ECB (2020a). Short-time work schemes and their effects on wages and disposable income. *Economic Bulletin, European Central Bank*, Issue 4.
- ECB (2020b). The impact of COVID-19 on potential output in the euro area. *European Central Bank, Economic Bulletin*, Issue 7.
- Eisner, R. (1998). *Investment, National Income and Economic Policy*. Edward Elgar, United Kingdom
- Erceg, C. J., Levin, A. T. (2014). Labor force participation and monetary policy in the wake of the Great Recession. *Journal of Money, Credit and Banking* 46 (S2), 3–49.
- Eser, F., Karadi, P., Lane, P. R., Moretti, L., Osbat, C. (2020). The Phillips Curve at the ECB (No. 2400). *European Central Bank Working Paper Series*, No. 2400.
- Faberman, R. J., Mueller, A. I., Şahin, A., Topa, G. (2020). The Shadow Margins of Labor Market Slack. *National Bureau of Economic Research*, No. w26852.
- Fontanari, C., Palumbo, A., Salvatori, C. (2019). Is it Really “Full Employment”? Margins for Expansion in the US Economy in the Middle of 2019. *Institute for New Economic Thinking Blog*, available at <https://www.ineteconomics.org/perspectives/blog/is-it-really-full-employment>.
- Fontanari, C., Palumbo, A., Salvatori, C. (2020). Potential Output in Theory and Practice: A Revision and Update of Okun's Original Method. *Structural Change and Economic Dynamics*, 54, 247-266.
- Freeman, D. G. (2001). Panel tests of Okun’s law for ten industrial countries. *Economic Inquiry*, 39(4), 511–523.
- Gallant, J., Kroft, K., Lange, F., Notowidigdo, M. J. (2020). Temporary unemployment and labor market dynamics during the COVID-19 recession. *National Bureau of Economic Research*, No. w27924.
- Garavan, S. (2017). Okun’s law: an empirical investigation into Eurozone growth and unemployment. *The Student Economic Review* XXXI.

Hall, R. E., Kudlyak, M. (2020). Unemployed With Jobs and Without Jobs. *National Bureau of Economic Research*, No. w27886.

Havik, K., Mc Morrow, K., Orlandi, F., Planas, C., Raciborski, R., Röger, W., Rossi, A., Thum Thysen, A., Vandermeulen, V. (2014). The production function methodology for calculating potential growth rates & output gaps. *Directorate General Economic and Financial Affairs (DG ECFIN), European Commission*, No. 535.

Heimberger, P. (2020). Potential Output, EU Fiscal Surveillance and the COVID-19 Shock. *Intereconomics*, 55, 167-174.

Hooper, P., Mishkin, F. S., Sufi, A. (2019). Prospects for Inflation in a High Pressure Economy: Is the Phillips Curve Dead or is It Just Hibernating? *National Bureau of Economic Research Working Paper Series*, No. 25792.

IMF (2020a). World Economic Outlook, October: A Long and Difficult Ascent. *International Monetary Fund*.

IMF (2020b). Fiscal Monitor Report, October: Policies for the Recovery. *International Monetary Fund*.

IMF (2021). World Economic Outlook, April: Managing Divergent Recoveries. *International Monetary Fund*.

Knotek II, E. S. (2007). How useful is Okun's law? *Economic Review Federal Reserve Bank of Kansas City*, 92(4), 73.

Komlos, J. (2019a). The Real US Unemployment Rate Is Twice the Official Rate, and the Phillips Curve. *CESifo Working Paper*, No. 7859.

Komlos, J. (2019b). Estimating Labor Market Slack, US 1994-2019. *CESifo Working Paper*, No. 7941.

Krugman, P. (2021). Stagflation revisited, Feb. 5 2021, available at <https://paulkrugman.substack.com/p/stagflation-revisited>

Lee, J. (2000). The robustness of Okun's law: Evidence from OECD countries. *Journal of macroeconomics*, 22(2), 331–356.

Lengwiler, Y., Orphanides, A. (2020). Options for the ECB's Monetary Policy Strategy Review. Study for the Committee on Economic and Monetary Affairs, Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, Luxembourg.

Lubik, T., Schwartzman, F. (2020). Public and Private Debt after the Pandemic and Policy Normalization. *Richmond Fed Economic Brief*, No. 20-06.

Malley, J., Molana, H. (2008). Output, unemployment and Okun's law: Some evidence from the G7. *Economics Letters*, 101(2), 113–115.

Mayes, D. G., Viren, M. (2002). Asymmetry and the Problem of Aggregation in the Euro Area. *Empirica*, 29(1), 47–73.

Ministero dell'Economia e delle Finanze (2020). Documento di Economia e Finanza 2020, Nota di Aggiornamento (*Italian Ministry of Finance*).

Moosa, I. A. (1997). A cross-country comparison of Okun's coefficient. *Journal of comparative economics*, 24(3), 335–356.

OECD (2020). OECD Economic Outlook, Volume 2020 Issue 2: Preliminary version. *OECD Publishing*, No. 108, Paris.

Okun, A. (1962). Potential GNP: Its measurement and significance. In: *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, Washington. ASA, pp. 98–104.

Paldman, M. (1987). How much does one percent of growth change the unemployment rate. *European Economic Review*, 31, 306–313.

Perez-Arce, F., Prados, M. J. (2020). The Decline in the US Labor Force Participation Rate: A Literature Review. *Journal of Economic Surveys*, No. 2020-005.

Perman, R., Tavera, C. (2005). A cross-country analysis of the Okun's law coefficient convergence in Europe. *Applied Economics*, 37(21), 2501–2513.

Silvapulle, P., Moosa, I. A., Silvapulle, M. J. (2004). Asymmetry in Okun's Law. *Canadian Journal of Economics*, 37(2), 353–374.

Sögner, L., Stiassny, A. (2002). An analysis on the structural stability of Okun's law a cross-country study. *Applied Economics*, 34(14), 1775–1787.

Stock, L., Vogler-Ludwig, K. (2010). NAIRU and Okun's law—The macro-economy in a nutshell. Thematic Paper for the European Commission Directorate General for Employment, *Social Affairs and Equal Opportunities, European Employment Observatory*. Munich.

Storm, S. (2017). The New Normal: Demand, Secular Stagnation and the Vanishing Middle-Class. *Institute for New Economic Thinking*, Working Paper No. 55, available at https://www.ineteconomics.org/uploads/papers/WP_55-Storm-The-New-Normal.pdf.

Storm, S. (2018). With Official Unemployment This Low, Why Are Wages Rising So Slowly? *Institute for New Economic Thinking Blog*, available at <https://www.ineteconomics.org/perspectives/blog/with-official-unemployment-this-low-why-are-wages-rising-so-slowly>.

Strauss, D. (2020). Hidden joblessness threatens economic recovery in US and Europe. *Financial Times*, available at <https://www.ft.com/content/ec3d88dc-0dc1-4f6e-adf7-37e8f4316a22>

Summers, L. (2017). America needs its unions more than ever. *Financial Times*, available at <https://www.ft.com/content/180127da-8e59-11e7-9580-c651950d3672>.

Summers, L. (2021). The Biden stimulus is admirably ambitious. But it brings some big risks, too. *Washington Post*, Feb. 4 2021, available at <https://www.washingtonpost.com/opinions/2021/02/04/larry-summers-biden-covid-stimulus/>

Virén, M. (2001). The Okun curve is non-linear. *Economics letters*, 70(2), 253–257.

Williams, J. (2019). Discussion of 'Prospects for Inflation in a High Pressure Economy: Is the Phillips Curve Dead or Is It Just Hibernating?' by Peter Hooper, Frederic S. Mishkin, and Amir Sufi: remarks at the US Monetary Policy Forum, New York City. *Federal Reserve Bank of New York*, No. 305.

Yellen, J. (2019). Former Fed Chair Janet Yellen on why the answer to the inflation puzzle matters. Remarks at the Hutchins Center on Fiscal & Monetary Policy at Brookings on October 3, available at <https://brook.gs/2IlxqpJ>

Appendix

TABLE A1. Data definitions and sources			
Variables	Frequency	Unit/measure	Source
Gross domestic product	annual	Chained 2015 prices National currency	AMECO
Unemployment rate	annual	Percentage of total active population	AMECO
Total annual hours worked	annual	Millions of hours	AMECO
Population	annual	Persons 15 to 64 years	AMECO
Gross domestic product	annual	Chained 2015 prices National currency	OECD
Unemployment	annual	Persons, Thousand	OECD
Labor Force	annual	Persons, Thousand	OECD
Marginally attached	annual	Persons, Thousand	OECD
Involuntary part-time	annual	Persons, Thousand	OECD
Gross domestic product	quarterly	Chained 2012 dollars	BEA
Unemployment rate	quarterly	Percentage of the civilian labor force	BLS
U-4	quarterly	total unemployed plus discouraged workers, as a percent of the civilian labor force plus discouraged workers	BLS
U-5	quarterly	total unemployed, plus discouraged workers, plus all other marginally attached workers, as a percent of the civilian labor force plus all marginally attached workers	BLS
U-6	quarterly	total unemployed, plus all marginally attached workers, plus total employed part time for economic reasons, as a percent of the civilian labor force plus all marginally attached workers	BLS
Hours worked	quarterly	Hours worked in total US economy and subsectors	BLS
Population 15-64	quarterly	Working Age Population: Aged 15-64: All Persons for the United States	FRED
Population 15+	quarterly	Working Age Population: Aged 15 and Over: All Persons for the United States	FRED

TABLE A2. Tests for unit root		
Levin-Lin-Chu unit-root test*		
Variables	Adjusted test statistic	p-value
First dataset		
<i>gy</i>	-9.039	0.0000
Δu	-10.979	0.0000
<i>gHAP</i>	-10.102	0.0000
Second dataset		
<i>gy</i>	-6.644	0.0000
Δu	-5.228	0.0000
$\Delta u2$	-4.101	0.0000
$\Delta u5$	-4.952	0.0000
$\Delta u6$	-4.061	0.0000
*we include lagged values up to 3 lags; we report here results only for the first lag		

Table A3. SUR estimates of Okun's Law with different unemployment indicators										
Country	Variable	Okun's Law with U3			Okun's Law with U5			Okun's Law with U6		
		Coefficient	Standard errors	R ²	Coefficient	Standard errors	R ²	Coefficient	Standard errors	R ²
Austria	gy_t $const$	-0.20*** 0.37***	(.0204) (.0997)	0.3 9	-0.21*** 0.47**	(.0500) (.2174)	0.1 3	-0.22*** 0.53**	(.0162) (.2276)	0.1 3
Belgium	gy_t $const$	-0.22*** 0.29*	(.0371) (.1569)	0.1 2	-0.25*** 0.34**	(.0324) (.1478)	0.1 7	-0.16*** 0.08	(.0005) (.1519)	0.1
Canada	gy_t $const$	-0.24*** 0.54***	(.0241) (.1167)	0.5 4	-0.25*** 0.54***	(.0198) (.1295)	0.5 0	-0.34*** 0.78***	(.0028) (.1682)	0.4 7
Denmark	gy_t $const$	-0.34*** 0.45***	(.0139) (.1266)	0.6 0	-0.41*** 0.49**	(.0641) (.2002)	0.4 5	-0.45*** 0.53***	(.0116) (.1928)	0.5 8
Finland	gy_t $const$	-0.19*** 0.13	(.0166) (.1023)	0.6 3	-0.25*** 0.23	(.0219) (.1439)	0.5 9	-0.26*** 0.29*	(.0034) (.1539)	0.5 8
France	gy_t $const$	-0.31*** 0.43***	(.0210) (.1506)	0.5 0	-0.35*** 0.51***	(.0407) (.1213)	0.4 4	-0.37*** 0.66***	(.0015) (.1645)	0.3 1
Germany	gy_t $const$	-0.15*** -0.06	(.0210) (.1506)	0.2 8	-0.23*** 0.07	(.0505) (.1875)	0.2 7	-0.23*** 0.09	(.0133) (.2519)	0.1 8
Greece	gy_t $const$	-0.47*** 0.45	(.0259) (.3470)	0.6 3	-0.47*** 0.49	(.0183) (.3487)	0.6 5	-0.48*** 0.67*	(.0024) (.3620)	0.6 6
Hungary	gy_t $const$	-0.20*** 0.31	(.0079) (.2095)	0.2 7	-0.24*** 0.30	(.0174) (.2547)	0.3 0	-0.28*** 0.42	(.0011) (.2839)	0.3 1
Israel	gy_t $const$	-0.36*** 0.88***	(.0164) (.1352)	0.5 9	-0.33*** 0.65***	(.0218) (.1626)	0.4 9	-0.52*** 1.16***	(.0035) (.1564)	0.6 4
Netherlands	gy_t $const$	-0.30*** 0.45***	(.0185) (.1561)	0.4 0	-0.41*** 0.62***	(.0397) (.2054)	0.3 9	-0.53*** 0.86***	(.0049) (.2345)	0.4 4
New Zealand	gy_t $const$	-0.31*** 0.77***	(.0269) (.1178)	0.5 5	-0.41*** 1.03***	(.0504) (.1856)	0.5 5	-0.53*** 1.3***	(.0047) (.1697)	0.5 2
Portugal	gy_t $const$	-0.54*** 0.50**	(.0163) (.2287)	0.5 6	-0.76*** 0.70**	(.0251) (.2738)	0.6 1	-0.78*** 0.75**	(.0049) (.3288)	0.5 6
Slovak Rep.	gy_t $const$	-0.30*** 0.48*	(.0238) (.2899)	0.3 7	-0.23*** 0.24	(.0268) (.2939)	0.3 2	-0.29*** 0.54*	(.0032) (.3067)	0.3 4
Spain	gy_t $const$	-0.88*** 1.5***	(.0512) (.3253)	0.7 2	-0.84*** 1.5***	(.0399) (.3120)	0.7 2	-0.98*** 2.04***	(.0049) (.3637)	0.6 9
Sweden	gy_t $const$	-0.17*** 0.40***	(.0075) (.1542)	0.2 4	-0.21*** 0.49***	(.0208) (.1857)	0.2 8	-0.22*** 0.53**	(.0035) (.2162)	0.2 2
United Kingdom	gy_t $const$	-0.30*** 0.43***	(.0170) (.1124)	0.5 2	-0.31*** 0.39***	(.0198) (.1364)	0.4 8	-0.41*** 0.61***	(.0021) (.1748)	0.4 6

Legend: *p<.1; **p<.05; ***p<.001

Table A4.1. DL model with range-of-unemployment asymmetries - <i>Wald test on parameters</i>						
Null hypothesis	Dependent variable: U3		Dependent variable: U5		Dependent variable: U6	
	F statistic	p-value	F statistic	p-value	F statistic	p-value
All coefficients (equal zero)	22.20	0.0000	22.48	0.0000	16.16	0.0000
Contemporaneous effect (equal between groups)	0.15	0.8617	0.31	0.7356	1.38	0.2569
First lag effect (equal between groups)	2.96	0.0567	2.17	0.1199	3.69	0.0288
Second lag effect (equal between groups)	0.62	0.5399	1.44	0.2415	0.91	0.4055
Cumulated effect (equal between groups)	3.91	0.0235	5.43	0.0059	6.16	0.0031

TABLE A4.2. DL model with range-of-unemployment asymmetries – Comparing models with different lag orders

Dependent variable: U3		Obs	Ll (model)	df	AIC	BIC
	No-lags model	104	16.37	6	-20.75	-4.88
1-lag model	103	43.36	9	-68.72	-45.00	
2-lag model	102	54.10	12	-84.21	-52.71	
Dependent variable: U5		Obs	Ll (model)	df	AIC	BIC
	No-lags model	104	9.96	6	-7.92	7.95
1-lag model	103	32.63	9	-47.26	-23.55	
2-lag model	102	44.25	12	-64.51	-33.01	
Dependent variable: U6		Obs	Ll (model)	df	AIC	BIC
	No-lags model	104	-24.19	6	60.38	76.25
1-lag model	103	6.10	9	5.80	29.51	
2-lag model	102	11.37	12	1.27	32.77	

TABLE A4.3. DL model with range-of-unemployment asymmetries – *F*-test for comparing models

Dependent variable:		F stat	p-value
U3	<i>H</i> ₀ : no-lags model nested in 1-lag model	20.04	0.0000
	<i>H</i> ₀ : no-lags model nested in 2-lag model	16.01	0.0000
	<i>H</i> ₀ : 1-lag model nested in 2-lags model	7.67	0.0000
	<i>H</i> ₀ : 1-lag model nested in 2-lags model (without contemp. effect)	9.33	0.0000
Dependent variable:		F stat	p-value
U5	<i>H</i> ₀ : no-lags model nested in 1-lag model	16.17	0.0000
	<i>H</i> ₀ : no-lags model nested in 2-lag model	14.11	0.0000
	<i>H</i> ₀ : 1-lag model nested in 2-lags model	8.26	0.0001
	<i>H</i> ₀ : 1-lag model nested in 2-lags model (without contemp. effect)	8.76	0.0276
Dependent variable:		F stat	p-value
U6	<i>H</i> ₀ : no-lags model nested in 1-lag model	23.33	0.0000
	<i>H</i> ₀ : no-lags model nested in 2-lag model	14.18	0.0000
	<i>H</i> ₀ : 1-lag model nested in 2-lags model	3.30	0.0240
	<i>H</i> ₀ : 1-lag model nested in 2-lags model (without contemp. effect)	4.21	0.0077

Table A5 SUR estimates of Okun's Law with Standardized Hours Worked					
Country	Variable	Coefficient	Standard errors	Cumulated coeff.	R ²
Belgium	gy_t	0.60***	(.0902)	0.88	0.37
	gy_{t-1}	0.28***	(.0899)		
	<i>const</i>	-0.01***	(.0029)		
Canada	gy_t	0.77***	(.0855)	0.98	0.61
	gy_{t-1}	0.21***	(.0813)		
	<i>const</i>	-0.2***	(.0029)		
Denmark	gy_t	0.53***	(.0811)	0.89	0.60
	gy_{t-1}	0.36***	(.0781)		
	<i>const</i>	-0.02***	(.0025)		
Finland	gy_t	0.54***	(.0632)	0.77	0.62
	gy_{t-1}	0.23***	(.0617)		
	<i>const</i>	-0.017***	(.0032)		
France	gy_t	0.39***	(.0929)	0.47	0.22
	gy_{t-1}	0.08	(.0925)		
	<i>const</i>	-0.008***	(.0025)		
Germany	gy_t	0.44***	(.0439)	0.44	0.38
	gy_{t-1}	0.02	(.0456)		
	<i>const</i>	-0.006***	(.0019)		
Greece	gy_t	0.19*	(.1112)	0.60	0.46
	gy_{t-1}	0.41***	(.1102)		
	<i>const</i>	-0.04	(.0035)		
Iceland	gy_t	0.94***	(.1348)	0.94	0.49
	gy_{t-1}	-0.20	(.1326)		
	<i>const</i>	-0.03***	(.0065)		
Ireland	gy_t	0.40***	(.0744)	0.71	0.55
	gy_{t-1}	0.31***	(.0743)		
	<i>const</i>	-0.03***	(.0059)		
Italy	gy_t	0.52***	(.0753)	0.84	0.50
	gy_{t-1}	0.32***	(.0739)		
	<i>const</i>	-0.007***	(.0021)		
Japan	gy_t	0.29***	(.0901)	0.29	0.10
	gy_{t-1}	-0.07	(.0898)		
	<i>const</i>	-0.004	(.0036)		
Netherlands	gy_t	0.48***	(.0719)	0.72	0.56
	gy_{t-1}	0.24**	(.0714)		
	<i>const</i>	-0.008**	(.0021)		
Norway	gy_t	0.54***	(.1071)	0.96	0.33
	gy_{t-1}	0.42***	(.1014)		
	<i>const</i>	-0.02***	(.0033)		

Portugal	gy_t	0.86***	(.1076)	0.86	0.50
	gy_{t-1}	-0.06	(.1039)		
	<i>const</i>	-0.01***	(.0036)		
Spain	gy_t	1.01***	(.0853)	1.01	0.80
	gy_{t-1}	0.05	(.0828)		
	<i>const</i>	-0.18***	(.0025)		
Sweden	gy_t	0.63***	(.0615)	0.86	0.61
	gy_{t-1}	0.23***	(.0584)		
	<i>const</i>	-0.02***	(.0026)		
Switzerland	gy_t	0.49***	(.1219)	0.77	0.39
	gy_{t-1}	0.28**	(.1236)		
	<i>const</i>	-0.01***	(.0035)		
United Kingdom	gy_t	0.60***	(.1052)	0.92	0.48
	gy_{t-1}	0.32***	(.1055)		
	<i>const</i>	-0.02***	(.0035)		
United States	gy_t	0.93***	(.1053)	1.15	0.54
	gy_{t-1}	0.22**	(.0944)		
	<i>const</i>	-0.03***	(.0037)		
Legend: *p<.1; **p<.05; ***p<.001					