Working Paper No. 940

On the "Utilization Controversy": A Rejoinder and Some Comments

by

Michalis Nikiforos*
Levy Economics Institute of Bard College

November 2019

* The author would like to thank Luca Zamparelli for useful comments and suggestions. The usual disclaimer applies.

Levy Economics Institute
P.O. Box 5000
Annandale-on-Hudson, NY 12504-5000
http://www.levyinstitute.org

ISSN 1547-366X

ABSTRACT

The critique by Gahn and González (2019) of the conclusions in Nikiforos (2016) regarding what data should be used to evaluate whether capacity utilization is endogenous to demand is weak for the following reasons: (i) The Federal Reserve Board (FRB) measure of utilization is not appropriate for measuring long-run variations of utilization because of the method and purpose of its construction. Even if its difference from the measures of the average workweek of capital (AWW) were trivial, this would still be the case; if anything, it would show that the AWW is also an inappropriate measure. (ii) Gahn and González choose to ignore the longest available estimate of the AWW produced by Foss, which has a clear long-run trend. (iii) Their econometric results are not robust to more suitable specifications of the unit root tests. Under these specifications, the tests overwhelmingly fail to reject the unit root hypothesis. (iv) Other estimates of the AWW, which were not included in Nikiforos (2016) confirm these conclusions. v) For the comparison between the AWW series and the FRB series, they construct variables that are not meaningful because they subtract series in different units. When the comparison is done correctly, the results confirm that the difference between the AWW series and the FRB series has a unit root. (vi) A stationary utilization rate is not consistent with any theory of the determination of capacity utilization. Even if demand did not play a role, there is no reason to expect that all the other factors that determine utilization would change in a fashion that would keep utilization constant.

KEYWORDS: Capacity Utilization; Workweek of Capital; Stationarity

JEL CLASSIFICATIONS: C22; D24; E11; E23

1 INTRODUCTION

My paper on the "utilization controversy" (Nikiforos 2016) discusses, among other things, what data should be used to evaluate whether capacity utilization is endogenous to demand.[1] I argued that the Federal Reserve Board (FRB) data are not right for that purpose. This becomes clear from the documentation of the data and the related literature. A more appropriate measure of long-run variations in utilization is the average workweek of capital (AWW). Unlike the FRB measure, several measures of the AWW have positive trends over significant periods of time.

Gahn and González (2019) question this argument. They make their point in two related ways. First, they compare the FRB's measure of utilization with the measures of the AWW and show that the difference between the two is stationary using several unit root tests. Second, they examine the stationarity of the AWW measures, and they also find evidence that most of them are stationary as well.

Their argument is problematic for several reasons. First, it contains an obvious logical fallacy. The FRB measure is not appropriate for measuring long-run variations in utilization because of the way it is constructed; its documentation and a large body of other sources confirm this. Even if it was similar to the AWW, this would not prove the appropriateness of the FRB measure, but rather that the AWW is inappropriate.

Second, in Nikiforos (2016) I present four different estimates of the AWW. Gahn and González choose to examine three of them and ignore the longest one (Foss 1984, 1995). Foss's work is the cornerstone of the literature on the empirical estimation of the AWW and shows that there have been very significant increases in the time capital is utilized.

[1] The "utilization controversy" refers to the debate on the endogeneity (or lack thereof) of the long run rate of capacity utilization to demand.

Third, econometric best practices favor modified information criteria or other methods for selecting the lag length of unit root tests (Ng and Perron 1995, 2001). Gahn and González's results hinge on using the conventional model selection criteria. When the proper lag length selection methods are chosen, the results change dramatically.

Fourth, in the present paper I also discuss two more estimates, which were not presented in Nikiforos (2016), and confirm the difference between the FRB measure and the AWW.

Fifth, in the comparison between the FRB's rates and the AWW performed by Gahn and González, the variable that is being tested is constructed by subtracting variables that are expressed in different units. Obviously, this is not a meaningful variable. I provide a comparison by converting all measures into indices and then examining the trajectory of their difference over time. The results confirm that there is a difference between the FRB and the AWW measures.

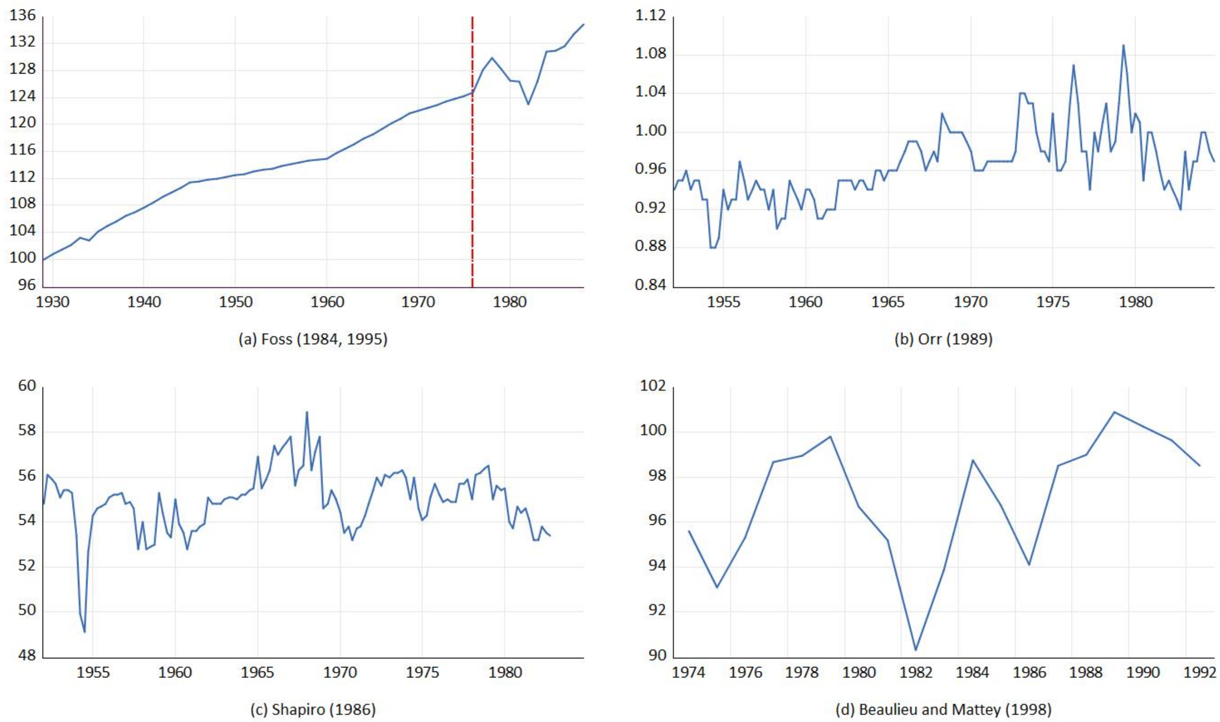
Finally, a rate of capacity utilization that is relatively stable over time is not consistent with any theory for the determination of capacity utilization. Even if one believes that demand does not play a role and that only technological, cost, or other factors enter in its determination, there is no reason to expect that all these factors will change in such a way to keep utilization constant. This is another—logical—reason why the FRB data are not the appropriate.

For reasons of space constraints, the present rejoinder will discuss only the third, fourth, and fifth issues above, which are directly related to Gahn and Gonzalez's note. The reader can refer to Nikiforos (2016) for the of the FRB measure's method of construction (sections 4.1 and 4.2) and Foss's estimates (447–448). In Nikiforos (2019) I provide a more detailed discussion of the FRB measure that confirms my original conclusions, and also analyze the sixth point above, namely why stationary utilization is not consistent with any theory of utilization.

2 THE AVERAGE WORKWEEK OF CAPITAL

In Nikiforos (2016) I presented data on the AWW from four different sources: i) several studies by Foss (1963, 1981b, 1981a, 1984, 1995); ii) Orr (1989), who closely follows the methodology of Taubman and Gottschalk (1971); iii) Shapiro (1986); and iv) Beaulieu and Mattey (1998). These estimates are presented in figure 1.

Figure 1: Estimates of the Average Workweek of Capital



The pioneer of the workweek of capital studies was Foss. He was the first to show that a significant portion of capital lies idle most of the time and that over significant stretches of time there is a pronounced upward trend in the capital's utilization (figure 1a). Gahn and González (2019) completely ignore his estimates. They only examine the other three estimates with the use of the usual unit roots tests: (i) the Augmented Dickey-Fuller (ADF) test; (ii) the Dickey-Fuller

GLS (DFGLS) test; (iii) the Phillips-Perron (PP) test; (iv) the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test; (v) the Elliott-Rothenberg-Stock (ERS) test; and (vi) the Ng-Perron (NP) test. For the selection of the lag order of the tests, “whenever possible” they use the Bayesian (or Schwarz) information criterion (SIC), “due to its consistency properties.”

Gahn and González’s results are misleading and are not robust to different—and in fact more appropriate—specifications of the unit root tests. It is well-known that unit root tests are sensitive to the number of lags. Therefore the criterion used for lag selection can potentially be crucial for the test results. It is also well-known that conventional information criteria, such as the Akaike Information Criterion (AIC) and the SIC, can be problematic. For that reason, Ng and Perron (1995) suggest that sequential testing is superior, while Ng and Perron (2001), show that the *modified* AIC (MAIC) performs better compared to conventional tests. More generally, and for different reasons, the AIC and SIC do not perform well in small samples and all the samples examined here are very small ones. Therefore, it is not clear what “properties” of the SIC Gahn and González refer to.

To understand the importance of this issue, I ran the six different unit root tests for the series by Orr (1989), Shapiro (1986), and Beaulieu and Mattey (1998) using five different lag selection criteria: the conventional AIC and SIC, the MAIC, the modified SIC (MSIC), and sequential testing (t-test). The results are reported in table 1 (all the tables are placed at the end of the paper). In the upper panel of the table we see that in the case of the series by Orr (1989)—which is the longest series and the one I used in my econometric exercise (Nikiforos 2016, section 7)—the tests point unequivocally toward the existence of a unit root. The only exception is the SIC, which is the one Gahn and González chose to report in their paper. However, as explained above, the SIC is inferior to the modified criteria and sequential testing.

In the case of the series by Shapiro (1986) and Beaulieu and Mattey (1998), the lag selection is not sensitive to the criterion chosen. In the case of Shapiro there is a rejection of the unit root hypothesis, while in Beaulieu and Mattey, 's case the results are mixed and inconclusive. I will come back to Shapiro's data in the next section and show that the rejection of the unit root is due to a structural break in the data construction, which needs to be taken into account.

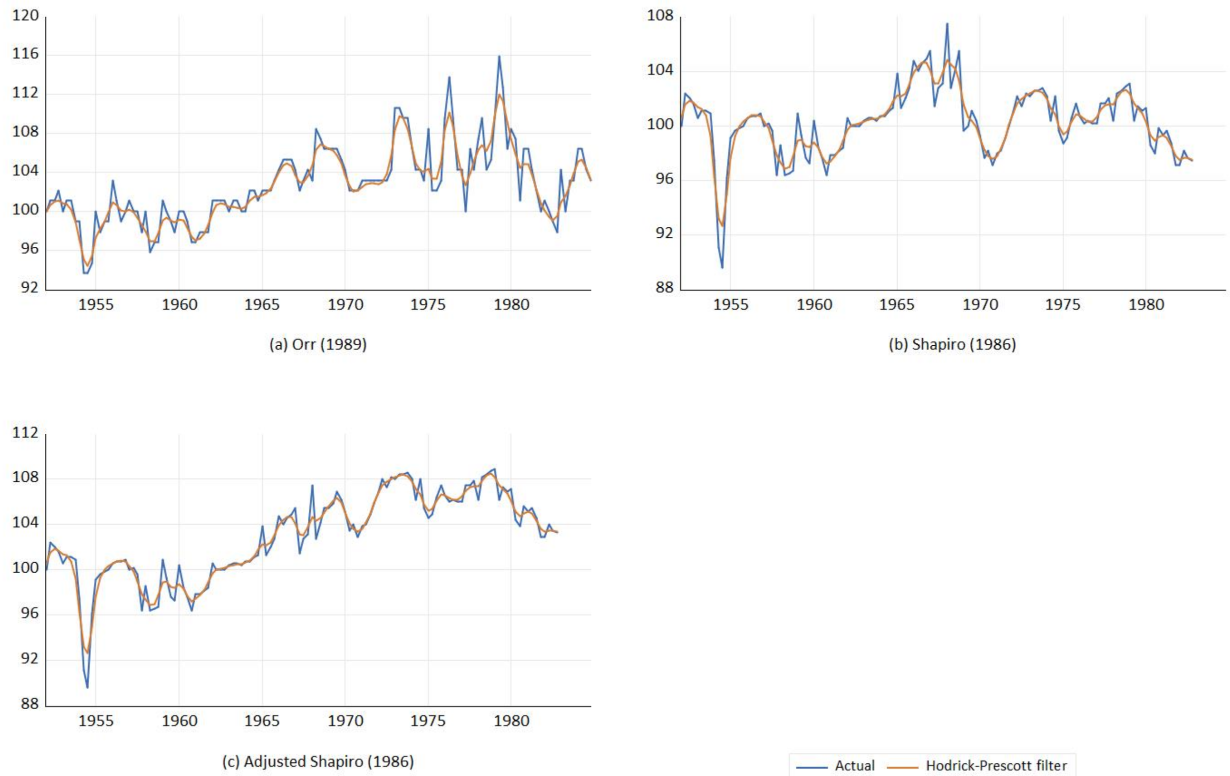
It is worth mentioning that Beaulieu and Mattey (1998) provide estimates of the AWW for the various two-digit industries within manufacturing. These estimates show that in some industries the increase over time is quite strong, and the unit root tests point toward nonstationarity. On the other hand, in the continuous industries the AWW is relatively stable.[2] This is important because if utilization was stationary, this stability should be the case in all industries.

Related to that, Gahn and González (2019, fn. 5) write that they performed a panel unit root test that strongly rejects the unit root null hypothesis. For the reason explained in the previous paragraph, they miss the point. Because of the heterogeneity of the various industries' adjustment margins, it is to be expected that in some industries the AWW will be relatively stable over time. Therefore, panel unit root tests are inappropriate.

Finally, a close look at the data in quarterly frequency by Orr (1989) and Shapiro (1986) reveals that they contain a lot of noise and they are very volatile at high frequencies (see figures 1b and 1c). To a large extent this is due to the fact that the data come directly from surveys, and also that they are not seasonally adjusted. Orr (1989, 92) writes that "the series does have a tendency to peak in the first quarter each year." At the same time there are many instances of consecutive quarters with exactly the same value for the AWW. In figure 1b we can see that there are many horizontal segments in the graph. This is also due to a lack of sufficient information from the surveys. This series' behavior comes in stark contrast to the FRB series, which is very smooth at

[2] Continuous industries are those with high startup and shutdown costs, which tend to adjust utilization through changes in the speed of operation (for a discussion, see Mattey and Strongin [1997]).

Figure 2: Estimates of the AWW by Orr (1989), Shapiro (1986), and Shapiro (1986) Adjusted for the Break in 1969q1



Note: The filtered series were created with the Hodrick-Prescott filter, with $\lambda = 1$. The base period for all actual series is 1952q1.

high frequencies—this is also due to the way the FRB series is constructed—and therefore is important for the comparison between them.

I examine what happens if we try to eliminate some of this noise, as well as consecutive quarters with exactly the same rate of utilization. Toward that purpose, I applied the Hodrick-Prescott (HP) filter on the two series with its smoothing parameter set as small as possible ($\lambda = 1$). The produced series are presented in figures 2a and 2b. The smoothed series are very close to the original estimates and are far from an HP trend in the underlying series.

I then applied the usual unit root test on the new, minimally smoothed series. The results are presented in the upper two panels of table 2. As one would expect, the results have moved slightly toward failing to reject unit root. As for the series by Orr (1989), even under the SIC most tests point to that direction. This is also the case in Shapiro's (1986) for the ADF and PP tests under the modified criteria, although the overall picture remains one of stationarity.[3]

3 MORE ON THE ESTIMATES BY SHAPIRO

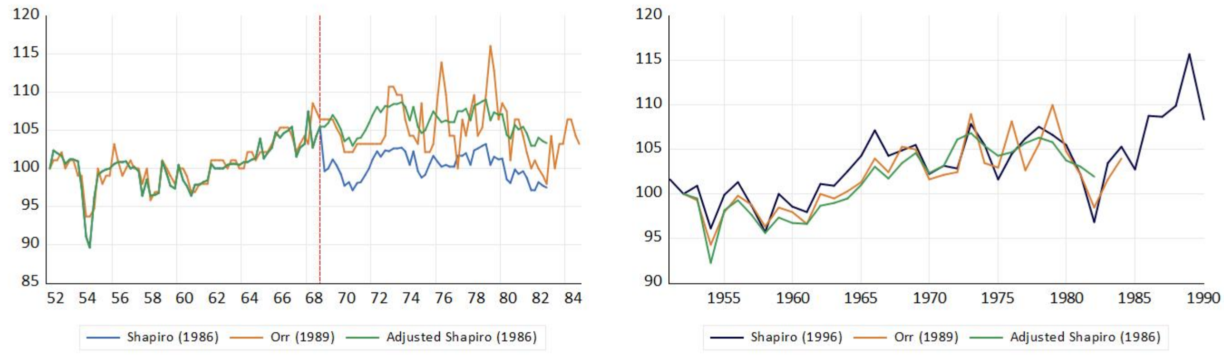
The estimates by Shapiro (1986) need some more discussion. Shapiro provides quarterly estimates for the period 1952–82. Up until 1968 he follows Taubman and Gottschalk(1971), using information from the US Bureau of Labor Statistics Area Wage Survey program in standard metropolitan statistical areas (SMSA) to produce estimates at the SMSA level and then aggregate them to the national level. Orr (1989) follows this methodology to produce his estimates that extend Taubman and Gottschalk's estimates until 1984. After 1969, Shapiro uses national-level data on shift-work, which he then interpolates to the quarterly frequency.[4]

Figure 3a shows that Shapiro's estimates exhibit a major dip in 1969q1. The dip is more than 5.5 percent. One should be suspicious of this kind of rapid shift that coincides with changes in the construction methodology of the series. It is also noteworthy that the economic conditions of the time do not support such a dip. In 1969q1, the FRB utilization rate increases, while the Orr's

[3] The HP filter has recently been criticized for producing spurious dynamics and on other grounds (Hamilton 2018). The results presented here are robust to using other smoothing methods. I chose the HP filter because it is easier to clarify and quantify how much—or how little—smoothing I applied to the original series, so that my results can be easily replicated.

[4] Another important difference is that Orr's data are much more volatile than Shapiro's. This is also related to Shapiro's use of national aggregates. Most of the major metropolitan areas were surveyed every two or three years. For the years that they were not surveyed, the BLS reported the numbers of the last available survey. On the other hand, Orr builds his index based on the surveys available each year. In that sense, Shapiro's series is close to a three-year moving average of Orr's series (Beaulieu and Mattey 1998, 208).

Figure 3: Estimates of the AWW by Shapiro (1986), and Shapiro (1996)



(a) Shapiro (1986)

(b) Shapiro (1996)

Note: The base period is 1952q1 for figure (a) and 1952 for figure (b).

measure remains the same relative to 1968q4. It is therefore likely that the drop is due to the change in the construction method.

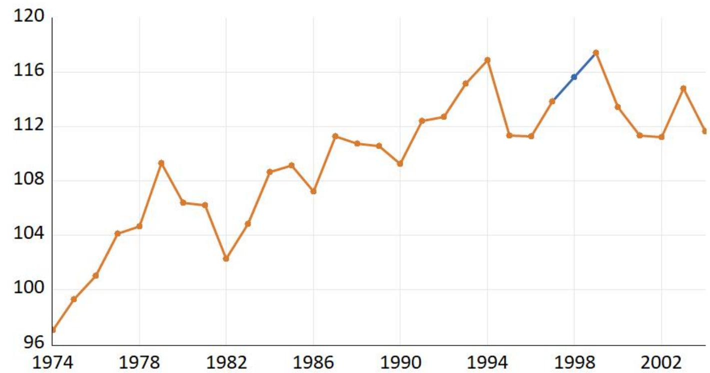
To examine this hypothesis, I produced an *adjusted* Shapiro (1986) series. In this series I assumed that there was no change in utilization in 1969q1 compared to 1968q4. Figure 3a shows that this adjusted Shapiro series now has the same trajectory of Orr's series, albeit much less volatile.

Ten years later, Shapiro (1996) produced an updated AWW series, now for the period 1951–90 and at annual frequency.[5] Figure 3b shows that Shapiro's (1996) estimates do not exhibit the big drop in 1969 that was present in Shapiro (1986)—in fact there is a small increase. It also shows that his 1996 estimates have similar trajectories to Orr's and the adjusted Shapiro (1986) series.[6] It is thus reasonable to conclude that the drop in 1969q1 is related to the change in methodology and does not represent a real shift in the AWW.

[5] These estimates were based on some then-recent estimates by Mayshar and Solon (1993) on shift-work.

[6] Shapiro (1996, , figure 2) plots the series but does not report the numbers. I extracted them from the figure electronically with a specialized software. Some measurement error is unavoidable here, but I do not think it significantly affects the results.

Figure 4: Estimate of the Average Workweek of Capital by Gorodnichenko and Shapiro (2011)



Note: The value for 1998 was produced by linearly interpolating the values for 1997 and 1999.

How does this affect the results of the unit root tests? The upper two panels of table 3 present the results of applying the unit root tests on the adjusted Shapiro (1986) series and those of Shapiro (1996). In both cases there is overwhelming support for the existence of a unit root. Moreover, the lower panel of table 2 presents the results of the unit root tests for the adjusted Shapiro (1986) series corrected for seasonal fluctuations and excessive volatility. In this case, there is not a single result that goes against the unit root hypothesis.

More recently, Gorodnichenko and Shapiro (2011) published annual AWW estimates for the period 1974–2004. The estimates are based on the Survey of Plant Capacity (SPC) and are presented in figure 4. The observation for 1998 is missing from the estimates due to missing and miscoded information in the surveys. In figure 4, the value for 1998 has been calculated by linearly interpolating the observations for 1997 and 1999, and is depicted with a different color. There is a clear upward trend during the first two decades of the data and relative stability afterwards. The results of the unit root test on this series are reported in the lower panel of table 3. These results also overwhelmingly fail to reject the unit root null hypothesis.[7]

Overall this section's discussion and the results of the statistical tests presented in tables 1–3 provide clear support in favor of the hypothesis that the AWW series has a unit root.

4 IS THE AWW DIFFERENT THAN THE FRB MEASURE?

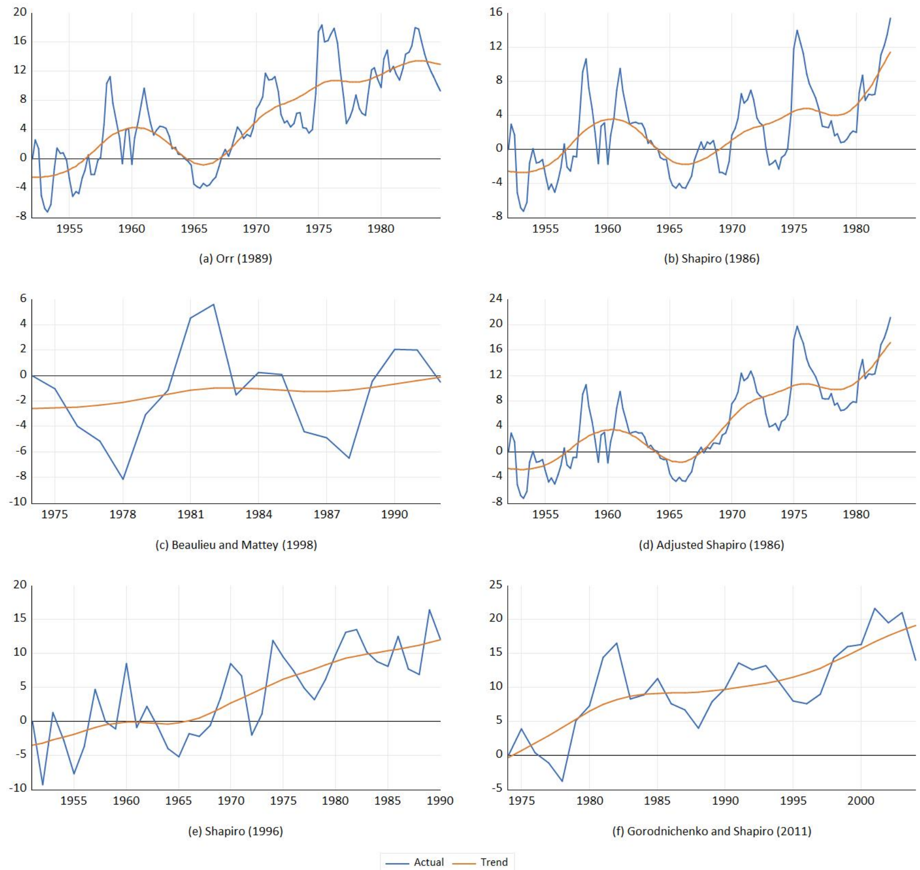
A last pending question is whether the AWW measures are different than the FRB measure. Gahn and González (2019) answer in the negative. They approach the question from the point of view of measurement error and argue that the difference between the AWW and the FRB's utilization measure represents such an error.

Ironically, the only measurement error here is committed by Gahn and González themselves. The FRB's utilization unit of measure is percent of full capacity output (as was defined above). On the other hand, Shapiro's (1986) unit of measure is hours per week, while Orr's (1989) measure is expressed as an index with 1968 as the base year. Gahn and González calculate the differences (or the measurement errors) by subtracting the FRB measure from the two AWW measures in these units. Obviously such a variable is not meaningful.

After sending a previous draft of this reply to Gahn and González, they came back with a new measure that assumes that the average of Orr's (1968) estimates were equal to Shapiro's hours for the same year. These series were then divided by 168 to reach a percentage of full capacity, which was subtracted from the FRB measure. The results of their unit root tests for Shapiro's series are mixed, while those for Orr's point toward stationarity.

[7] The tests were applied to the whole 1974–2004 period using the interpolated value for 1998. Tests for the period 1974–97 are even more supportive of the unit root hypothesis.

Figure 5: Difference between the Federal Reserve Measure of Capacity Utilization and Estimates of the AWW



These series suffer from the same problem, albeit in a more subtle way. The unit of the AWW series is percentage of engineering capacity, while the unit for the FRB series is percentage of "economic" capacity. Moreover, the unit roots tests are also sensitive to the lag selection method, with the appropriate lag selection methods pointing toward non stationarity, as in the previous section. Finally, the results of Orr's series are related to its method of construction and the resulting volatility (see footnote 4); controlling for this excess volatility completely changes the results.

In order to compare the various AWW series with the FRB's utilization measure, I created an index for each of the variables with the base time period the first period with available data for each of them. I also created similar indexes for the FRB's utilization measure—one FRB index for each measure of the AWW. I then subtracted the FRB indices from the respective AWW measures.[8]

The results are presented in figure 5. The blue line shows the actual difference between the two measures, while the orange line shows the trend calculated with the HP filter.[9] A visual inspection of the graph reveals that all series, except for Beaulieu and Mattey (1998), have a significant trend.

The unit root tests that were performed confirm this (table 4). In the case of Beaulieu and Mattey (1998), the results are mixed, but for all other series there is clear support in favor of the unit root hypothesis.

5 CONCLUSION

In this paper I outlined six reasons why I disagree with the argument put forward by Gahn and González (2019), which I think further establish that the FRB measure is not appropriate for the evaluation of the desired utilization's long-run behavior, and that the AWW—which is not stationary—is a more appropriate measure. A more interesting question for future research is whether demand contributes to this non stationarity. I provided a first answer in my econometric analysis in Nikiforos (2016, section 7), but more could be done.

[8] Two technical details are worth mentioning. For the series with quarterly observations, I used their smoothed transformation, which removes noise, seasonality, and consecutive quarters with exactly the same level of AWW (see the discussion around figure 2, above). The reason for that is that the FRB measure is smoothed at several stages between the survey and the final published result (Morin and Stevens 2004; Federal Reserve 2019). Hence some adjustment of the AWW measure—albeit much smaller—was also required for proper comparison. The basic results of the tests performed here (table 4) are robust to different specifications. Second, in the case of the series with annual frequency, I used the FRB end-of-period series, since the annual SPC surveys that serve as the basis for the measures of the AWW are also conducted at the end of the year.

[9] The trends in figure 5 were created using the usual values of the smoothing parameter λ: 1,600 for quarterly series, and 100 for the annual series.

REFERENCES

Beaulieu, J. Joseph, and Joe Mattey. 1998. "The Workweek of Capital and Capital Utilization in Manufacturing." *Journal of Productivity Analysis* 10 (2): 199–223.

Federal Reserve. 2019. "Capacity and Capacity Utilization Methodology." Federal Reserve System. Available at: https://www.federalreserve.gov/releases/g17/Meth/MethCap.htm.

Foss, Murray F. 1963. "The Utilization of Capital Equipment: Postwar Compared With Prewar." *Survey of Current Business* 43 (6): 8–17.

———. 1981a. *Changes in the Workweek of Fixed Capital: U.S. Manufacturing, 1929 to 1976.* Washington, DC: American Enterprise Institute for Public Policy Research.

———. 1981b. "Long-Run Changes in the Workweek of Fixed Capital." *The American Economic Review* 71 (2): 58–63.

———. 1984. *Changing Utilization of Fixed Capital: An Element in Long-Term Growth.* Washington, DC: American Enterprise Institute for Public Policy Research.

———. 1995. "Operating Hours of the US Manufacturing Plants, 1976-1988, and their Significance for Productivity Change." In *Work Patterns and Capital Utilisation: An International Comparative Study,* edited by Dominique Anxo, Gerhard Bosch, Gilbert Cette, Thomas Sterner, and Dominique Taddei. London: Kluwer Academic Publishers.

Gahn, Santiago J., and Alejandro González. 2019. "On the 'utilisation controversy': a comment." Online first, *Cambridge Journal of Economics.*

Gorodnichenko, Yuriy, and Matthew D. Shapiro. 2011. "Using the survey of plant capacity to measure capital utilization." US Census Bureau Center for Economic Studies Paper No. CES-WP-11-19. Washington, DC: US Census Bureau.

Hamilton, James D. 2018. "Why you should never use the Hodrick-Prescott filter." *Review of Economics and Statistics* 100 (5): 831–43.

Mattey, Joe, and Steven Strongin. 1997. "Factor utilization and margins for adjusting output: evidence from manufacturing plants." *Economic Review:* 3–17.

Mayshar, Joram, and Gary Solon. 1993. "Shift work and the business cycle." *The American Economic Review* 83 (2): 224–8.

Morin, Norman, and John Stevens. 2004. "Estimating capacity utilization from survey data." Working Paper 2004-49. Finance and Economics Discussion Series. Divisions of Research & Statistics and Monetary Affairs. Washington, DC: Federal Reserve Board.

Ng, Serena, and Pierre Perron. 1995. "Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag." *Journal of the American Statistical Association* 90 (429): 268–81.

———. 2001. "Lag length selection and the construction of unit root tests with good size and power." *Econometrica* 69 (6): 1519–54.

Nikiforos, Michalis. 2016. "On the 'utilisation controversy': a theoretical and empirical discussion of the Kaleckian model of growth and distribution." *Cambridge Journal of Economics* 40 (2): 437–67.

———. 2019. "Notes on accumulation and utilization of capital: some empirical issues." Levy Economics Institute Working Paper. Annandale-on-Hudson, NY: Levy Economics Institute of Bard College.

Orr, James A. 1989. "The Average Workweek of Capital in Manufacturing, 1952–1984." *Journal of the American Statistical Association* 84 (405): 88–94.

Shapiro, Matthew D. 1986. "Capital utilization and capital accumulation: Theory and evidence." *Journal of Applied Econometrics* 1 (3): 211–34.

———. 1996. "Macroeconomic implications of variation in the workweek of capital." *Brookings Papers on Economic Activity* 1996 (2): 79–133.

Taubman, Paul, and Peter Gottschalk. 1971. "The Average Workweek of Capital in Manufacturing." *Journal of the American Statistical Association* 66 (335): 448–55.

Table 1: Unit Root Tests for Orr (1989), Shapiro (1986), and Beaulieu and Mattey (1998)

		Lags	ADF	DFGLS	PP	KPSS	ERS	NP			
								MZa	*MZt*	*MSB*	*MPt*
Orr (1989)	AIC	10	-1.787	-1.146	-2.439	111.020***	10.469	-2.097	-0.991	0.473	11.386
	SIC	0	-3.930***	-3.473***	-3.93***	14.712***	1.369***	-20.438***	-3.186***	0.156***	1.238***
	MAIC	11	-1.728	-1.044	-2.445	144.106***	13.675	-1.597	-0.857	0.536	14.661
	MSIC	10	-1.787	-1.146	-2.439	111.020***	10.469	-2.097	-0.991	0.473	11.386
	t-stat	10	-1.787	-1.146	-2.439	111.020***	10.469	-2.097	-0.991	0.473	11.386
Shapiro (1986)	AIC	1	-3.790***	-3.810***	-4.033***	2.497	1.031***	-26.22***	-3.586***	0.137***	1.049***
	SIC	1	-3.790***	-3.810***	-4.033***	2.497	1.031***	-26.220***	-3.586***	0.137***	1.049***
	MAIC	1	-3.790***	-3.810***	-4.033***	2.497	1.031***	-26.220***	-3.586***	0.137***	1.049***
	MSIC	1	-3.790***	-3.810***	-4.033***	2.497	1.031***	-26.220***	-3.586***	0.137***	1.049***
	t-stat	0	-3.973***	-3.976***	-3.973***	2.589	1.069***	-25.369***	-3.526***	0.139***	1.083***
Beaulieu and Mattey (1988)	AIC	0	-2.188	-2.237**	-2.190	0.605**	3.800*	-6.253*	-1.745*	0.279	3.988*
	SIC	0	-2.188	-2.237**	-2.190	0.605**	3.800*	-6.253*	-1.745*	0.279	3.988*
	MAIC	0	-2.188	-2.237**	-2.190	0.605**	3.800*	-6.253*	-1.745*	0.279	3.988*
	MSIC	0	-2.188	-2.237**	-2.190	0.605**	3.800*	-6.253*	-1.745*	0.279	3.988*
	t-stat	0	-2.188	-2.237**	-2.190	0.605**	3.800*	-6.253*	-1.745*	0.279	3.988*

Note: *** p<0.01, ** p<0.05, * p<0.1

Table 2: Unit Root Tests for Orr (1989), Shapiro (1986), and Shapiro (1986) Adjusted for the Break in 1969q1, Filtered with the HP filter ($\lambda = 1$).

		Lags	ADF	DFGLS	PP	KPSS	ERS	NP			
								MZa	*MZt*	*MSB*	*MPt*
Orr (1989)	AIC	10	-1.903	-1.341	-1.722	83.889***	6.502	-3.520	-1.291	0.367	6.960
	SIC	3	-2.325	-2.000	-2.466	38.449***	2.980***	-8.835**	-2.079**	0.235**	2.863**
	MAIC	10	-1.903	-1.341	-1.722	83.889***	6.502	-3.520	-1.291	0.367	6.960
	MSIC	2	-1.958	-1.652	-1.974	61.993***	4.805	-5.545	-1.637*	0.295	4.504
	t-stat	10	-1.903	-1.341	-1.722	83.889***	6.502	-3.520	-1.291	0.367	6.960
Shapiro (1986)	AIC	3	-3.283**	-3.267***	-3.826***	2.812***	1.006***	-29.584***	-3.792***	0.128***	0.998***
	SIC	3	-3.283**	-3.267***	-3.826***	2.812***	1.006***	-29.584***	-3.792***	0.128***	0.998***
	MAIC	2	-2.315	-2.293**	-2.305	7.232***	2.589**	-10.975**	-2.257**	0.206**	2.570**
	MSIC	2	-2.315	-2.293**	-2.305	7.232***	2.589**	-10.975**	-2.257**	0.206**	2.570**
	t-stat	3	-3.283**	-3.267***	-3.826***	2.812***	1.006***	-29.584***	-3.792***	0.128***	0.998***
Adjusted Shapiro (1986)	AIC	8	-1.716	-1.024	-0.899	461.649***	25.796	-0.873	-0.618	0.708	25.425
	SIC	3	-1.674	-1.486	-1.791	93.297***	5.213	-4.988	-1.560	0.313	4.962
	MAIC	8	-1.716	-1.024	-0.899	461.649***	25.796	-0.873	-0.618	0.708	25.425
	MSIC	2	-1.316	-1.116	-1.345	173.947***	9.720	-2.639	-1.122	0.425	9.180
	t-stat	12	-1.246	-0.795	-1.063	300.939***	16.816	-1.503	-0.833	0.554	15.589

Note: *** p<0.01, ** p<0.05, * p<0.1

Table 3: Unit Root Tests for Shapiro (1986) Adjusted for the Break in 1969q1, Shapiro (1996), and Gorodnichenko and Shapiro (2011).

		Lags	ADF	DFGLS	PP	KPSS	ERS	NP			
								MZa	*MZt*	*MSB*	*MPt*
Adjusted Shapiro (1986)	AIC	4	-1.720	-1.435	-2.004	79.562***	5.387	-4.813	-1.528	0.317	5.145
	SIC	1	-2.235	-1.954*	-2.384	52.407***	3.548*	-7.677*	-1.941*	0.253*	3.262*
	MAIC	9	-1.716	-0.915	-1.291	336.086***	22.756	-0.964	-0.647	0.671	22.952
	MSIC	1	-2.235	-1.954*	-2.384	52.407***	3.548*	-7.677*	-1.941*	0.253*	3.262*
	t-stat	9	-1.716	-0.915	-1.291	336.086***	22.756	-0.964	-0.647	0.671	22.952
Shapiro (1996)	AIC	0	-2.479	-2.479**	-2.481	3.569***	2.934*	-9.803**	-2.125**	0.217**	2.841**
	SIC	0	-2.479	-2.479**	-2.481	3.569***	2.934*	-9.803**	-2.125**	0.217**	2.841**
	MAIC	2	-0.966	-0.997	-1.813	7.865***	6.466	-3.923	-1.271	0.324	6.348
	MSIC	2	-0.966	-0.997	-1.813	7.865***	6.466	-3.923	-1.271	0.324	6.348
	t-stat	0	-2.479	-2.479**	-2.481	3.569***	2.934*	-9.803**	-2.125**	0.217**	2.841**
Gorodnichenko and Shapiro (2011)	AIC	0	-2.666*	-1.436	-2.664*	9.998***	21.944	-2.154	-0.943	0.438	10.575
	SIC	0	-2.666*	-1.436	-2.664*	9.998***	21.944	-2.154	-0.943	0.438	10.575
	MAIC	3	-2.194	-0.406	-2.945*	23.434***	51.432	-0.514	-0.369	0.719	28.507
	MSIC	3	-2.194	-0.406	-2.945*	23.434***	51.432	-0.514	-0.369	0.719	28.507
	t-stat	3	-2.194	-0.406	-2.945*	9.998***	21.944	-0.514	-0.369	0.719	28.507

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4: Unit Root Tests for the Difference between the FRB series and Estimates of the AWW.

		Lags	ADF	DFGLS	PP	KPSS	ERS	NP			
								MZa	*MZt*	*MSB*	*MPt*
Orr (1989)	AIC	8	-1.369	-0.733	-1.563	119.005***	9.76	-2.309	-0.968	0.419	9.919
	SIC	1	-2.799*	-2.364**	-2.886**	31.348***	2.571**	-11.709**	-2.366**	0.202**	2.306**
	MAIC	8	-1.369	-0.733	-1.563	119.005***	9.76	-2.309	-0.968	0.419	9.919
	MSIC	2	-2.245	-1.776*	-2.345	48.502***	3.978*	-7.191*	-1.829*	0.254*	3.654*
	t-stat	8	-1.369	-0.733	-1.563	119.005***	9.76	-2.309	-0.968	0.419	9.919
Shapiro (1986)	AIC	8	-1.201	-1.023	-1.188	22.619***	6.798	-3.379	-0.842	0.249	7.073
	SIC	1	-2.889**	-2.806***	-3.163**	7.463***	2.243**	-22.164***	-3.025***	0.136***	2.123***
	MAIC	8	-1.201	-1.023	-1.188	22.619***	6.798	-3.379	-0.842	0.249	7.073
	MSIC	0	-1.666	-1.547	-1.671	16.677***	5.012	-7.352*	-1.499	0.204**	4.736
	t-stat	8	-1.201	-1.023	-1.188	22.619***	6.798	-3.379	-0.842	0.249	7.073
Beaulieu and Mattey (1988)	AIC	0	-2.170	-2.217**	-2.170	0.222	3.645*	-6.201*	-1.759*	0.284	3.956*
	SIC	0	-2.170	-2.217**	-2.170	0.222	3.645*	-6.201*	-1.759*	0.284	3.956*
	MAIC	0	-2.170	-2.217**	-2.170	0.222	3.645*	-6.201*	-1.759*	0.284	3.956*
	MSIC	0	-2.170	-2.217**	-2.170	0.222	3.645*	-6.201*	-1.759*	0.284	3.956*
	t-stat	0	-2.170	-2.217**	-2.170	0.222	3.645*	-6.201*	-1.759*	0.284	3.956*

Note: *** p<0.01, ** p<0.05, * p<0.1

Table 4, Continued

		Lags	ADF	DFGLS	PP	KPSS	ERS	NP			
								MZa	*MZt*	*MSB*	*MPt*
Adjusted Shapiro (1986)	AIC	8	-0.566	-0.094	-0.400	105.51***	15.018	0.605	0.236	0.389	15.636
	SIC	1	-1.702	-1.408	-1.934	34.864***	4.962	-7.481*	-1.568	0.210**	4.534
	MAIC	8	-0.566	-0.094	-0.400	105.51***	15.018	0.605	0.236	0.389	15.636
	MSIC	0	-0.901	-0.531	-0.91	71.07***	10.116	-1.731	-0.516	0.298	9.157
	t-stat	8	-0.566	-0.094	-0.400	105.51***	15.018	0.605	0.236	0.389	15.636
Shapiro (1996)	AIC	0	-2.593	-2.548**	-2.596	4.386***	3.055*	-10.221**	-2.152**	0.211**	2.811**
	SIC	0	-2.593	-2.548**	-2.596	4.386***	3.055*	-10.221**	-2.152**	0.211**	2.811**
	MAIC	4	-1.292	-1.085	-1.660	16.053***	11.180	-1.871	-0.774	0.413	10.834
	MSIC	4	-1.292	-1.085	-1.660	16.053***	11.180	-1.871	-0.774	0.413	10.834
	t-stat	2	-1.124	-1.006	-1.703	14.548***	10.133	-2.275	-0.882	0.387	9.519
Gorodnichenko and Shapiro (2011)	AIC	0	-2.042	-1.596	-2.05	5.227***	9.425	-3.874	-1.308	0.338	6.376
	SIC	0	-2.042	-1.596	-2.05	5.227***	9.425	-3.874	-1.308	0.338	6.376
	MAIC	0	-2.042	-1.596	-2.05	5.227***	9.425	-3.874	-1.308	0.338	6.376
	MSIC	0	-2.042	-1.596	-2.05	5.227***	9.425	-3.874	-1.308	0.338	6.376
	t-stat	0	-2.042	-1.596	-2.05	5.227***	9.425	-3.874	-1.308	0.338	6.376

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$