School of Arts & Sciences Theses
Hunter College

Spring 5-5-2020

Female Labor Force Participation and Economic Growth: Granger Causality Test

Zulakha Rahman
CUNY Hunter College

Female Labor Force Participation and Economic Growth: Granger Causality Test

by

Zulakha Rahman

Submitted in partial fulfillment
of the requirements for the degree of
Master of Arts in Economics, Hunter College
The City University of New York

2020

05/10/2020	Randall Filer
Date	Thesis Sponsor
05/10/2020	Johnathan Conning
Date	Second Reader

Abstract:

Does rising female labor force participation precede economic growth, or does the relationship run in the other direction? This paper addresses this question by performing a Granger causality test to examine the relationship between economic growth and female labor force participation worldwide. The economic growth –is measured in two ways: using Gross Domestic Product, Purchasing Power Parity (constant 2011 international dollar) and Gross Domestic Product using local currency units. The test uses a completely balanced panel dataset spanning over 28 years from 1990 to 2018 and using 151 countries for Purchasing Power Parity dataset and 154 countries in GDP in local currency unit dataset.

Table of Contents

I) Introduction

The economics literature has made major contributions toward understanding how worldwide movements in female labor force participation rates have played a significant role in the economic development of nations. One theory of the connection between these two variables was the feminization U-hypothesis, developed in the 1960s (Sinha 1967). Using this theory as a foundation, many studies found evidence that female labor force participation rates at first declined and then later rose with economic development in a U-shape way (Sinha 1967; Boserup 1970; Durand 1975; Pampel and Tanaka 1986; Psacharopoulos and Tzannatos 1989; Goldin 1995; Cağatay and Özler 1995; Mammen and Paxson 2000; Luci 2009; Tam 2011). The explanation on how female labor force participation could increase together with income happens overtime. The stylized argument of this theory is that when a country is poor, women work out of necessity, mostly in subsistence agriculture or home-based production. However, as a country develops, economic activity shifts from agriculture to industry which benefits more men than women which is portrayed by the declining line in the U-hypothesis. As the country reaches a higher level of economic development, female education levels rise, fertility rates fall, and social stigmas weaken which enable women to take advantage of new jobs emerging in the service sector that are more family-friendly and accessible. As a result, this leads to the rising curve of the U-shape (Sinha 1967; Boserup 1970; Durand 1975; Pampel and Tanaka 1986; Psacharopoulos and Tzannatos 1989; Goldin 1995).

While this pattern holds true worldwide in broad terms, not all countries follow this U-shape. The relationship between female labor force participation and economic growth is in fact more complex, since many social and cultural factors may affect female labor participation (Standing 1978; Steel 1981; Antecol 2003; Rahman and Islam 2013; Klasen and Gaddes 2013; Lechman and Kaur 2015; Koyuncu and Ozen 2017; Klasen 2019). The explored possible factors why in developing countries may not conform with the U-hypothesis is due to the limiting economic opportunities for women in employment sector (Verick 2018; Klasen 2019). In developing countries, the quality of employment and opportunities for better jobs continue to be unequally distributed between men and women, sometimes even in countries where there is an equal labor force participation rate. Often, in most developing countries, when women work, they tend to earn less, work in less productive jobs, and be overrepresented in unpaid family work and other forms of vulnerable work (Verick 2018; Klasen 2019). This limitation within the labor market means that social stigmas may not weaken allowing there to be the rising curve hypothesized for U-shape. Another reason why countries may not follow the typical U-shape hypothesis is if there might be a possible bi-directional movement. This could mean that changes in economic growth lead to changes in female labor force participation and female labor force participation in return influence changes in economic development. Historically, the worldwide trend of female labor force participation is shown to suggest that it follows the suggested U-shape pattern where female labor force participation is the highest in some of the poorest and richest countries while female labor force participation is lower in countries with average national income. Meanwhile, while looking at individual countries and its timeseries patterns at the World Bank Graphs, they seem to each follow their own unique pattern which may or may not reflect the U-shape hypothesis (World Bank). Therefore, given these possible reasons, it may be useful to further explore how these two variables move together and whether we can observe any ‘causal’ relationship between the two variables.

Except for a few empirical studies examining the Granger causality relationship between economic growth and female labor force participation, the literature on the causal relationship between these two variables is underdeveloped. The causality tests between the variables have been conducted in Asia, Pakistan, and Bangladesh (Su et al, 2019; Mujahid and Zafar 2012; Haque et al, 2019).

A Granger causality test may show whether a possible ‘causal’ relationship or precedence ordering between economic growth and female labor force participation exists for both uni and bi-directionality. Since the question of true causality is much more intricate and the assumption of that one thing preceding another can be used as a proof of causation, the Granger test is mostly useful to assess predictive causality (Granger, 1969). Therefore, this causality test will reveal whether one can influence changes in other, both variables influence changes in each other, or neither variables can explain changes in another. This test is done using recent developments adapted for short time series and panel data, where these methods will be applied to the causality testing leading to results with more accurate estimates (Hurlin and Venet, 2001; Dumitrescu and Hurlin, 2012).

II) Background & Method

The Granger causality test, proposed by Clive Granger in 1969, is a statistical test to determine whether one time series is useful in forecasting another. Granger's work claims that the causality between variables in economics could be tested by measuring the ability to predict future values of a time series using prior values of another time series. In a bivariate framework, the first variable is said to cause the second variable in the Granger sense if the forecast for the second variable improves when lagged values for the first variable are taken into account. Using this proposed econometric method will allow us to see if there is any relationship between the two variables - female labor force participation rate and economic growth - in a cross-country sample.

The Granger framework for a bivariate panel data model for a stationarity time series is a two-step model. The first uses the idea that when the time series X Granger causes time series Y, the patterns in X are repeated in Y after some time lag which allows us to use the past values of X for the prediction of future values of Y. The first step of the model is computing an autoregressive model to find the optimal lag length. The autoregressive model is represented as:

$$(a) \qquad y_{i,t} = \alpha_i + \sum_{k=1}^{K} \gamma_{ik} y_{i,t-k} + \varepsilon_{i,t}$$

$$(b) \qquad x_{i,t} = \alpha_i + \sum_{k=1}^{K} \gamma_{ik} x_{i,t-k} + \varepsilon_{i,t}$$

Following this, the autoregressive distributed lag model is created with the dependent variable. The equation is shown as:

$$(a) \qquad y_{i,t} = \alpha_i + \sum_{k=1}^{K} \gamma_{ik} y_{i,t-k} + \sum_{k=1}^{K} \beta_{ik} x_{i,t-k} + \varepsilon_{i,t}$$

$$(b) \qquad x_{i,t} = \alpha_i + \sum_{k=1}^{K} \gamma_{ik} x_{i,t-k} + \sum_{k=1}^{K} \beta_{ik} y_{i,t-k} + \varepsilon_{i,t}$$

where $i = 1, \ldots, N,\ t = 1, \ldots, T$, and $\varepsilon_{i,t}$ are $i.i.d.(0,\ \sigma_{\varepsilon}^{2})$. The autoregressive distributed lag model allows us to assess whether the dependent variable is significant to the model from its test statistics result.

Following this, the second step is computing a joint F-test. If this null hypothesis is rejected for this F-statistics, there is evidence of Granger causality. The null hypothesis for a panel data for non-causality is defined as:

$$H_0: \beta_{i1} = \cdots = \beta_{ik} = 0 \qquad \forall i = 1, \ldots, N$$
$$H_1: \beta_{i1} \neq \cdots \neq \beta_{ik} = 0 \qquad \forall i = N_1 + 1, \ldots, N$$

And the joint F-test which is computed as:

$$F = \frac{\left[\frac{RSS_r - RSS_u}{N_k}\right]}{\left[\frac{RSS_u}{SN - N - NK - K}\right]}$$

where RSS_r denotes the restricted sum of squared residuals obtained under the null hypothesis, RSS_u is the unrestricted sum of squared residuals computed from time-stationary results from equations (a) or (b), and SN denotes the total number of observations within K groups and N observation for overall sample size of the panel dataset (Dumitrescu and Hurlin, 2012).

The null hypothesis that *x* does not Granger-cause *y*, for equation (a), is accepted if and only if no lagged values of *x* add explanatory power in the regression and fails the joint F-test. Similarly, for equation (b), *y* does not Granger-cause *x*, is accepted if and only if no lagged values of *y* add explanatory power in the regression and fails the joint F-test. On the other hand, results may imply that neither variable Granger-causes the other, or that each of the two variables Granger-causes the other.

III) <u>Literature Review</u>

Causality tests between female labor force participation and economic growth have been conducted in Asia, Pakistan, and Bangladesh (Su et al, 2019; Mujahid and Zafar 2012; Haque et al, 2019). Su et al (2019) analyze data from Asia for the following set of countries: China, Japan, Korea, Indonesia, Malaysia, the Philippines, Singapore, Thailand, Vietnam and India. To study how the variables interact with each other, the authors use the method of panel bootstrap Granger causality to investigate between the two variables from 1990 - 2016. The authors use a panel bootstrap method since the sample may have some cross-sectional dependence and causal interaction due to the fact that a shock in one country might affect other countries since they are close together regionally (Su et al, 2019). Since the current paper explores worldwide effects where there is less concern about cross-sectional dependence in a region and on overall movement, an autoregressive model will be used to explore the causality rather than a panel bootstrap method. The author's estimations result for China, Korea, and Indonesia show that female labor force participation rate Granger causes economic growth, implying that greater female labor force participation rate has been conducive to economic development. However, when the test for economic growth Granger causes female labor force participation is conducted, it reveals that it is significant for more than half the countries considered in the sample which are: Korea, Malaysia, Singapore, Thailand, Vietnam, and India. The results from the sample suggest that Malaysia, Singapore, Thailand, Vietnam, and India follow the U-shaped model between FLFPR and economic development. Meanwhile, Korea shows bidirectional causality and the Philippines show no causal relationship (Su et al, 2019).

A second empirical analysis has been investigated in Pakistan between female labor force participation and economic growth (Mujahid and Zafar 2012). The authors used a sample of Pakistan on an annual data from 1980 to 2010. The authors use an autoregressive distributed lag model (shown above) to evaluate the Granger causality between the two variables. The Granger test is applied to reveal that there is a unidirectional causality between the two variables. The authors find that the economic growth Granger causes female labor force participation, but not the other way around. This allows the authors to conclude that the results support that there is a U-shaped association between the two variables (Mujahid and Zafar 2012).

Another empirical analysis has been conducted in Bangladesh (Haque et al, 2019). For the sample used in Bangladesh, the authors investigate the relationship between the labor force participation rate for both male and female, gross fixed capital formation, and economic growth using annual time series data from 1991 to 2017. Since the focus of interest of this paper is on female labor force participation and economic growth, the results on only those variables will be mentioned. The authors here also use the autoregressive distributed lag model to evaluate the Granger causality between the two variables. The authors use a co-integration test and Granger causality model to test the variables, and find that there is a unidirectional association between the variables. The results of the output show that the female labor force participation rate Granger causes economic growth, but growth does not Granger cause female labor force participation (Haque et al, 2019).

IV) Data

Female labor force participation rates (FLFP) by country and year were downloaded from the World Bank Development Indicators and International Labour Organization (ILO). FLFP is measured by percentage of female population ages 15 and older that are economically active and supply labor for the production of goods and services during a specified period. The data for female labor force participation are derived from ILO estimates. The ILO has designed and actively maintains a database of values of different labor market indicators taken from national estimates which may be composed of labor force survey, a household survey, or a population census. However, sometimes, there are years that the country-level data may be unavailable. In order to resolve this issue to recreate a robust dataset, the ILO has a series of econometric models that produce overall estimates using different national labor survey samples reported within the country (International Labour Organization).

Growth data results have been shown to differ drastically depending on the data source and how they adjust for changes in relative prices across countries (Hanousek et al, 2008). Hanousek et al (2008) explores the sensitivity of growth determinant by replicating several recent studies using different definitions of growth. Their finding shows how the results differ with the choice of data since growth rates calculated from different data sets conceptually measure different things. Although, it would seem that conversions of local currencies are converted using a single exchange rate and/or base year, there may be some other adjustments made during data revision. Therefore, the results show that it is beneficial to use local currency units rather than using a growth determinant dataset that has already been adjusted to relative prices across countries for comparability. Given this, the authors suggest avoiding using growth data that has been adjusted to create comparability across countries for a particular year to calculate growth over time within a given country since this may not lead to an accurate estimate for cross-country regressions (Hanousek et al, 2008).

Due to the sensitivity of growth data as discussed above, the Granger test will be computed using Gross Domestic Product, Purchasing Power Parity (constant 2011 international dollar) and Gross Domestic Product in local currency. The Gross Domestic Product, Purchasing Power Parity (constant 2011 international dollar) will be used for the first part of the analysis. The gross domestic product (GDP) is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. In this dataset, the GDP has been converted to international dollars using purchasing power parity rates. This allows an international dollar to have the same purchasing power as the U.S. dollar in the United States over GDP. The World Bank Development Indicators have robust data sets for the period 1990 to 2018 for economic growth and female labor force participation. Data on the growth of Gross Domestic Product, Purchasing Power Parity adjusted (constant 2011 international dollars) are compiled from the World Bank, International Comparison Program database. The test between the two variables will be conducted for a cross-sectional time series dataset consisting of 151 countries spanning over 28 years for the period 1990 to 2018.

Gross Domestic Product in local currency units will be used for the second part of the analysis. This dataset for growth is downloaded from the World Bank, International Comparison Program database which includes creating estimates using World Bank's data and estimates from OECD National Accounts database. The World Bank defines the gross domestic product (GDP) in local currency unit as the purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. The value in this dataset is measured in local currency. The World Bank Development Indicators have compiled robust data sets for the period 1990 to 2018 for economic growth and female labor force participation used for this paper. This data will be conducted for a cross-sectional time series dataset consisting of 154 countries spanning over 28 years for the period 1990 to 2018. Using this advice, the Granger test will be computed using Gross Domestic Product, Purchasing Power Parity (constant 2011 international dollar) and Gross Domestic Product in local currency. Using the author's advice, these two definitions may explore the sensitivity of different determinant of growth data. The Gross Domestic Product, Purchasing Power Parity is a growth definition where is the international dollar have been adjusted to have the same purchasing power as the U.S. dollar across all countries, while Gross Domestic Product in local currency units have been unadjusted for comparability across all countries.

Using the World Bank datasets, Figure 1a shows the list of countries included in the sample for this paper. The list of the countries has been organized by World Bank identified seven regions and four income levels. GDP Purchasing Power Parity and GDP in local currency unit have the same 151 countries with the local currency unit dataset having an additional 3 countries indicated by the asterisk next to the country name. The numerical values next to the country indicate their identified income level: 1) Low Income, 2) Lower Middle Income, 3) Upper Middle Income, and 4) High Income. In addition, Figure 1b provides the overview of movement of female labor force participation rate based on income group and region. Looking at the graph of female labor force participation rate based on income group, it shows a couple of significant trends. Based on the trends, we see that the trend of female labor force participation rate is the highest in low income countries and is decreasing over the years while for high income countries it is rising over the span of 28 years. This could be the reflection of the rising and declining curve of the U-shape hypothesis. Additionally, we see that Upper Middle Income has a lower female participation rate than Lower Middle Income countries. These overall trend movements can reflect the U-shape hypothesis where low income countries have the highest female labor force participation rate where it steadily declines through lower and upper middle income. Following this, countries see a rising female labor force participation rate once higher income level is achieved.

V) Empirical Analysis and Results

In order to properly estimate equation (a) and (b), as shown above, the variables for growth and female labor force participation need to be time-stationary so as to not lead to spurious results (Granger and Newbold 1974). The equations (a) and (b) for the first Granger step are re-written as:

$$(a) \qquad Growth_{i,t} = \alpha_i + \sum_{k=1}^{p} \gamma_{ik} \, Growth_{i,t-k} + \sum_{k=1}^{p} \beta_{ik} \, FLFP_{i,t-k} + \varepsilon_{i,t}$$

$$(b) \qquad FLFP_{i,t} = \alpha_i + \sum_{k=1}^{p} \gamma_{ik} \, FLFP_{i,t-k} + \sum_{k=1}^{p} \beta_{ik} \, Growth_{i,t-k} + \varepsilon_{i,t}$$

To make sure the variables are covariance stationary, an Augmented Dickey-Fuller test has been calculated for unit root both growth variables. The Augmented Dickey-Fuller test reveals that neither the variables growth nor female labor force participation are time-stationary, so the first differences are taken to make them stationary. The results in Table 1 show the values of the original and first differences for Augmented Dickey-Fuller test. The test results show that the p-value without the first-differenced variables are not statistically significant so, therefore, the null hypothesis of non-stationarity cannot be rejected. However, once the first differences are taken, the p-values are statistically significant and can be rejected at a 1% confidence level. Given this result, the first differences of these variables are used to compute the autoregressive models.

Using these stationary variables, we are able to compute Autoregressive Distributed Lag model (a) and (b) in a two-step approach. The first step is to determine the optimal lag for the autoregressive model and then fit the independent variable (female labor force participation in equation (a) and growth in equation (b)) to the correct autoregressive model. Since the autoregressive models are very sensitive to the number of lags included in the regression, both the Akaike (AIC) and Schwarz Information Criteria (BIC) have been used to find the appropriate number of lags. The model selection criterion test, shown in Table 2, reveals that the optimal lag length for both growth variables are at a lag of 2 and for female labor force participation is at a lag of 5.

Results for the Autoregressive Distributed Lag model (a) and (b) results are shown in Table 3. Using the growth data at Purchasing Power Parity (PPP), it shows for equation (a), the test statistic for female labor force participation is not statistically significant at any of the confidence intervals. Meanwhile, when growth is added to equation (b), the test statistics for the variable is shown to be significant at a 10% confidence level. Based on the test statistics results, it potentially reveals that for equation (a), female labor force participation does not appear to Granger cause growth since the test statistic is not significant (b) growth may Granger cause female labor force participation since the test statistic is significant for the lagged value of growth. However, when calculating using growth data in the local currency unit, there is a different outcome. For equations (a) and (b), it shows that the test statistic for neither female labor force participation nor growth are not statistically significant at any of the confidence intervals. Based on the test statistics results, it potentially reveals that for equation (a), female labor force participation may not Granger cause growth since the test statistic is not significant for the lag for female labor force participation rate, while for equation (b) growth may not Granger cause female labor force participation since the test statistics are also not significant. These differences reveal how sensitive results could be based on growth definitions which may lead to different economic conclusions.

The next step is to examine Granger causality is through computing the joint significance F-test between the two variables. In order to examine possible Granger causality between the variables, the last test is done using Dumitrescu and Hurlin's (2012) procedure for detecting joint significance in panel datasets. This procedure is used due Stata's original Granger test command – vargranger - is only compatible for exclusively working with a time series dataset. Since this is a cross-country time variable sample, Stata's time series test does not run. However, the Dumitrescu and Hurlin (2012) procedure has updated Stata's Granger test to be compatible with panel datasets by using the command xtgcause. The Dumitrescu and Hurlin (2012) procedure emphasizes that the test is designed to detect causality at the panel-level and rejecting does not exclude non-causality for some individuals (Lopez and Weber, 2017).

The rewritten and modified null hypothesis and alternative hypothesis for this panel Granger test for the joint F-test for what was shown in Section II is:

$$H_0: \beta_{i1} = \cdots = \beta_{ik} = 0 \qquad \forall i = 1, \ldots, N$$

$$H_1: \beta_{i1} = \cdots = \beta_{ik} = 0 \qquad \forall i = 1, \ldots, N_1$$
$$\beta_{i1} \neq \cdots \neq \beta_{ik} = 0 \qquad \forall i = N_1 + 1, \ldots, N$$

which reflects that Dumitrescu and Hurlin (2012) there can be causality for some individuals in the panel but not necessarily for all.

The procedure follows three steps to derive the joint test statistic value: 1) it runs the individual regressions implicitly implied (a) or (b); 2) then performs the F-test on the linear hypotheses $H_0: \beta_{i1} = \cdots = \beta_{ik} = 0$ (shown above) to retrieve individual Wald statistics (W_i) to compute the average Wald statistics ($\overline{W}$); and lastly, 3) uses the assumption that Wald statistics are normally distributed to give two different standardized Z statistics that are computed at different N and T dimensions. One of the Z-statistics, $\bar{Z}$, is to be considered for large N and T panel datasets, while $\tilde{z}$ (Z-tilde) is favored for large N but relatively small T datasets. If these are larger than standard critical value, then the null hypothesis H_0 should be rejected and conclude that Granger causality exists. As a result, Dumitrescu and Hurlin (2012) procedure is designed and optimized to detect causality at the panel-level for conclusion where the authors use Monte Carlo simulations to show that average Wald statistics are asymptotically well-behaved and can genuinely be used to investigate panel causality. Therefore, this procedure will allow us to test causality for both economic growth and female labor force participation rate in a panel data.

Using the Dumitrescu and Hurlin (2012) panel data procedure, the Granger causality for model (a) and (b) is computed. The output results are shown in Table 4 where the results are shown for model (a) and (b) separately for both variables. When computing output at Purchasing Power Parity, we see that when equation (a) is computed both of the Z statistic values are smaller than standard critical values at a 5% confidence level and the p-values reported are larger than 1%. From these results, we can conclude that we cannot reject the null hypothesis that female labor force participation does not Granger cause growth. On the other hand, when equation (b) is computed the output shows that both Z statistics are statistically significant at the 1% confidence level. This allows us to reject the null hypothesis and conclude that growth may Granger cause female labor force participation. However, from the output for GDP in local currency, we see that when equation (a) is computed both of the Z statistic values are statistically significant at 1%. From these results, we can reject the null hypothesis that female labor force participation does not Granger cause growth. Similarly, when equation (b) is computed the output shows that both Z statistics are statistically significant at the 1% confidence level. This also allows us to reject the null hypothesis and conclude that growth may Granger cause female labor force participation.

VI) Conclusion

Testing for Granger causality between growth and female labor force participation in a cross-country sample shows us how these two variables can be used to explain the causal relationship between each other. Using Purchasing Power Parity data, we can make two conclusions: 1) female labor force participation does not Granger cause growth and 2) growth Granger causes female labor force participation. For equation (a), we see that both the test statistics were not significant, since no lagged values of female labor force participation are retained in the regression concluding that there is no directional causality. However, for equation (b), the statistical significance of both the test statistics and Granger panel F-test confirm that there is a unidirectional causality. From this analysis we can make two major economic conclusions: 1) female labor force participation does not play a major contribution towards economic growth and 2) economic growth can be influential in the changes in the female labor force participation rate. The causality results suggest that it may follow the U-hypothesis which would indicate that there is a strong link between economic growth and female labor force participation. It should be noted that the Dumitrescu and Hurlin (2012) procedure concludes this result to describe the overall trend for the majority of countries, and the Granger test results may not hold true for all the individual countries in the panel.

For growth data measured in local currency, we see a different outcome as suggested by Hanousek et al (2008). The results reveal that: 1) female labor force participation Granger causes growth and 2) growth Granger causes female labor force participation. For equation (a) and (b), the statistical significance of Granger panel F-test confirms that there is a possible bi-directional causality. The contrasting causal results from the two different growth datasets reveal the sensitivity of results based on the definition of growth determinant. The results for both growth definition reveal that economic growth is very important in the changes of the female labor force participation rate. However, for this analysis, we can also conclude there may be potential implications that female labor force participation rate may have some possible influential changes on economic growth. Using the results of this analysis, we can conclude that it suggests that economic growth is strongly linked with increasing women's participation in the labor force. In addition, greater female participation in the labor market will result in improving their relative economic status and overall economic efficiency which will benefit the country by influencing economic development. In order to promote higher female labor force participation, countries should focus on effective policies to eliminate barriers for women within the labor force and create more opportunities to join the labor force.

VII) References:

Boserup, E. (1970), Women's role in economic development, St.Martin, New York.

Çağatay, N., Özler, Ş. (1995), Feminization of the labor force: the effects of long-term development and structural adjustment, World development, 23(11), pp.1883-1894.

Dickey, D. A. and W. A. Fuller (1981), Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, Econometrica 49, 1057-1072.

Dickey, D. and W. A. Fuller (1979), Distribution of the Estimates for Autoregressive Time Series with Unit Root, Journal of the American Statistical Association 74, 427-431.

Durand, J. (1975), The Labour Force in Economic Development: A Comparison of International Census Data, 1946-1966.

Hanousek, Durand, J, Dana Hajkova, and Randall Filer (2008), A Rise by Any Other Name? Sensitivity of Growth Regressions to Data Source. Journal of Macroeconomics, vol. 30, no. 3, 1188-1206.

Gaddis, Isis, and Stephan Klasen (2013), Economic Development, Structural Change, and Women's Labor Force Participation: Journal of Population Economics, vol. 27, no. 3, pp. 639–681.

Goldin, C. (1995), The U-Shaped Female Labour Force Function in Economic Development and Economic History. National Bureau of Economic Research: Working Paper Series (Working Paper No. 4707).

Granger, C. W. J. (1969), Investigating Causal Relations by Econometric Models and Cross-spectral Methods, Econometrica 37, 24-36.

Granger, C.W.J and P. Newbold (1974), Spurious Regressions in Econometrics, Journal of Econometrics, vol. 2, no. 2, 215-238.

Haque, Ahasan Ul, et al. (2019), Labor Force Participation Rate and Economic Growth: Observations for Bangladesh, International Journal of Economics and Financial Research, no. 59, 2019, pp. 209–213.

Hurlin, Christophe, and B. Venet (2001), Granger Causality Tests in Panel Data Models with Fixed Coefficients, Working Paper, EURIsCO, Universite Paris IX Dauphin.

Klasen, Stephan (2019), What Explains Uneven Female Labor Force Participation Levels and Trends in Developing Countries?, *The World Bank Research Observer*, vol. 34, no. 2, 2019, pp. 161–197.

Koyuncu, Cuneyt and Eda Özen (2017), Religious, Ethnic, Linguistic and Cultural Diversity and Female Labor Force Participation, Journal of Economics Bibliography, KSP Journals, vol. 4(1), pages 87-93.

Lechman, Ewa, and Harleen Kaur (2015), Economic Growth and Female Labor Force Participation – Verifying the U-Feminization Hypothesis. New Evidence for 162 Countries over the Period 1990-2012, Economics & Sociology, vol. 8, no. 1, pp. 246–257.

Lopez, Luciano, and Sylvain Weber Testing for Granger Causality in Panel Data (2017), The Stata Journal: Promoting Communications on Statistics and Stata, vol. 17, no. 4, pp. 972–984.

Luci, Angela (2009), Female Labour Market Participation and Economic Growth, International Journal of Innovation and Sustainable Development, vol. 4, no. 2/3, p. 97.

Mammen, K., Paxson, C. (2000), Women's work and economic development, The Journal of Economic Perspectives, pp.141-164.

Mujahid, Nooreen, and Naeem Uz Zafar (2012), Economic Growth-Female Labour Force Participation Nexus: An Empirical Evidence for Pakistan, The Pakistan Development Review, vol. 51, no. 4II, pp. 565–586.

Rahman, R. I., and R. Islam (2013), Female Labour Force Participation in Bangladesh: Trends, Drivers and Barriers, ILO Asia-Pacific Working Paper.

Sinha, J. N. (1967), Dynamics of female participation in economic activity in a developing economy. In: United Nations department of economic and social affairs, In Proceedings of the World Population Conference, Belgrade 1965 UN Publications. pp. 336–37.

Pampel, F. C. and K. Tanaka (1986), Economic Development and Female Labour Force Participation: A Re-consideration, Social Forces 64:3, 599-619.

Psacharopoulos, G. and Z. Tzannatos (1989), Female Labour Force Participation: An International Perspective, World Bank Research Observer 4:2, 187-201.

Standing, G. (1978), Labour Force Participation and Development, Geneva: International Labour Office.

Steel, W. F. (1981), Female and Small Scale Employment Under Modernisation in Ghana, Economic Development and Cultural Change 30, 153-167.

Su, Chi-Wei, Zheng-Zheng Li, Ran Tao and Oana-Ramona Lobont (2018), Can Economic Development Boost the Active Female Labor Force? Quality & Quantity, vol. 53, no. 2, pp. 1021–1036.

Tam, Henry (2011), U-Shaped Female Labor Participation with Economic Development: Some Panel Data Evidence, Economics Letters, vol. 110, no. 2, pp. 140–142.

Verick, Sher (2018), Female Labor Force and Development, IZA World of Labor, doi: 10.15185/izawol.87.v2

VIII) <u>Tables & Figures</u>

Figure 1a:

Countries in the sample for GDP PPP and GDP Local Currency Unit by region and income level:

East Asia & Pacific		**Latin America & Caribbean**		**Europe & Central Asia**		**Sub-Saharan Africa**	
Australia	4	Argentina	3	Albania	3	Angola	2
Brunei Darussalam	4	Belize	3	Armenia	3	Benin	1
China	3	Bolivia	2	Austria	4	Botswana	3
Fiji	3	Brazil	3	Azerbaijan	3	Burkina Faso	1
Hong Kong SAR, China	4	Chile	4	Belarus	3	Burundi	1
Indonesia	2	Colombia	3	Belgium	4	Cabo Verde	2
Japan	4	Costa Rica	3	Bulgaria	3	Cameroon	2
Korea, Rep.	4	Dominican Republic	3	Cyprus	4	Central African Republic	1
Lao PDR	2	Ecuador	3	Czech Republic	4	Chad	1
Macao SAR, China	4	El Salvador	2	Denmark	4	Comoros	2
Malaysia	3	Guatemala	3	Finland	4	Congo, Dem. Rep.	1
Mongolia	2	Guyana	3	France	4	Congo, Rep.	2
Myanmar	2	Haiti	1	Georgia	3	Cote d'Ivoire	2
New Zealand	4	Honduras	2	Germany	4	Equatorial Guinea	3
Papua New Guinea	2	Jamaica	3	Greece	4	Eswatini	2
Philippines	2	Mexico	3	Iceland	4	Ethiopia	1
Samoa	3	Nicaragua	2	Ireland	4	Gabon	3
Singapore	4	Panama	4	Italy	4	The Gambia	1
Solomon Islands	2	Paraguay	3	Kazakhstan	3	Ghana	2
Thailand	3	Peru	3	Kyrgyz Republic	2	Guinea	1
Tonga	3	Puerto Rico	4	Luxembourg	4	Guinea-Bissau	1
Vanuatu	2	St. Lucia	3	Netherlands	4	Kenya	2
Vietnam	2	St. Vincent and the G	3	North Macedonia	3	Lesotho	2
		Suriname	3	Norway	4	Madagascar	1
Middle East & North Africa		Trinidad and Tobago	4	Poland	4	Malawi	1
Algeria	3	Uruguay	4	Portugal	4	Mali	1
Bahrain	4	The Bahamas* (LCU)	4	Romania	3	Mauritania	2
Egypt, Arab Rep.	2	Barbados* (LCU)	4	Russian Federation	3	Mauritius	3
Iraq	3	Cuba (LCU)	3	Slovenia	4	Mozambique	1
Israel	4			Spain	4	Namibia	3
Jordan	3	**North America**		Sweden	4	Niger	1
Lebanon	3	Canada	4	Switzerland	4	Nigeria	2
Malta	4	United States	4	Tajikistan	1	Rwanda	1
Morocco	**2**			Turkey	3	Senegal	2
Oman	4			Turkmenistan	3	Sierra Leone	1
Saudi Arabia	4	**South Asia**		Ukraine	2	South Africa	3
Tunisia	2	Bangladesh	2	United Kingdom	4	Sudan	2
United Arab Emirates	4	Bhutan	2	Uzbekistan	2	Tanzania	1
Yemen, Rep.	1	India	2			Togo	1
		Nepal	1			Uganda	1
		Pakistan	2			Zambia	2
		Sri Lanka	3			Zimbabwe	2

Income Level Codes: 1 = Low Income, 2 = Lower Middle Income, 3 = Upper Middle Income, 4 = High Income
(*) = The 3 additional country in Local Currency Unit dataset

Figure 1b:

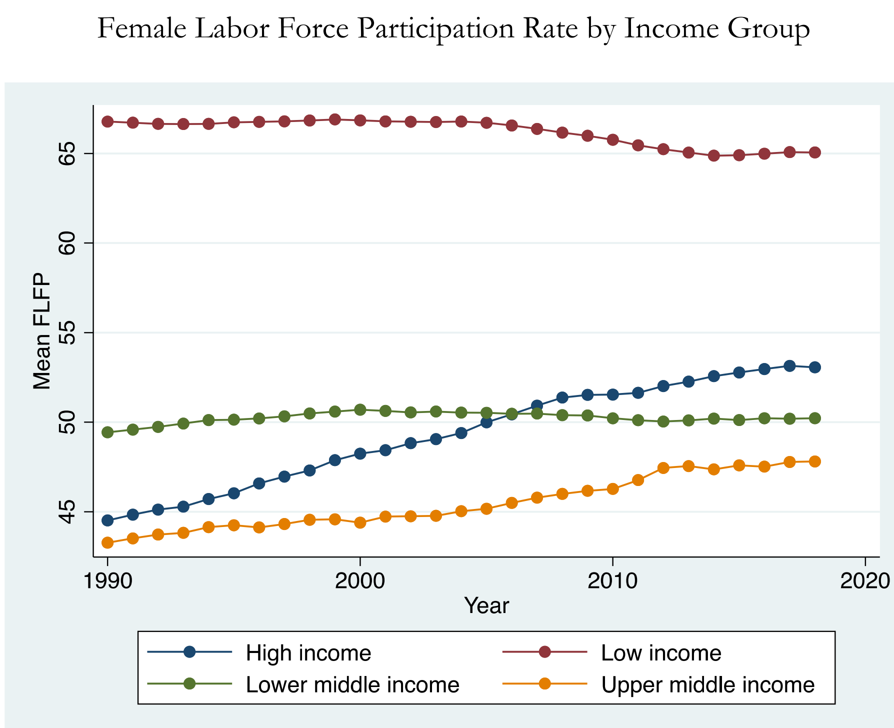
Female Labor Force Participation Rate by Income Group
65
60
55
50
45
Mean FLFP
1990
2000
2010
2020
Year
High income
Low income
Lower middle income
Upper middle income

Table 1:

Augmented Dickey-Fuller Test for Growth

	ADF for GDP PPP				ADF for GDP Local Currency			
	Without First Difference		First Difference		Without First Difference		First Difference	
	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
Inverse chi-squared	46.5353	1.0000	1467.8950	0.0000	56.6729	1.0000	1504.5409	0.0000
Inverse normal	21.1042	1.0000	-26.5211	0.0000	20.7770	1.0000	-26.9175	0.0000
Inverse logit t	23.3514	1.0000	-32.3081	0.0000	23.0670	1.0000	-32.7991	0.0000
Modified inv. chi-squared	-10.3947	1.0000	47.4396	0.0000	-10.1263	1.0000	48.2100	0.0000

Augmented Dickey-Fuller Test for FLFP

	ADF for GDP PPP Data				ADF for GDP Local Currency Data			
	Without First Difference		First Difference		Without First Difference		First Difference	
	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
Inverse chi-squared	403.6879	0.0001	1109.0056	0.0000	383.9483	0.0021	1083.6083	0.0000
Inverse normal	-1.1040	0.1348	-19.2839	0.0000	0.0032	0.5013	-18.6494	0.0000
Inverse logit t	-1.4227	0.0776	-22.9633	0.0000	-0.1630	0.4353	-21.8882	0.0000
Modified inv. chi-squared	4.1376	0.0000	32.8366	0.0000	3.0600	0.0011	31.2501	0.0000

Table 2:

Autoregressive Distributive Lag Model for Growth

		GDP Purchasing Power Parity		GDP Local Currency	
Lag Number	**Observation**	**AIC**	**BIC**	**AIC**	**BIC**
1	4077	-13041.39	-13028.76	-13349.41	-13028.76
2*	3,926	-13211.6**	-13192.77**	-13522.98 **	-13504.1**
3	3,775	-12868.97	-12844.02	-13177.14	-13152.12

Autoregressive Distributive Lag Model for FLFP

		GDP Purchasing Power Parity		GDP Local Currency	
Lag Number	**Observation**	**AIC**	**BIC**	**AIC**	**BIC**
1	4077	9702.644	9715.27	9897.756	9910.421
2	3,926	9288.007	9306.833	9513.008	9531.893
3	3,775	8908.881	8933.825	9139.673	9164.696
4	3,624	8536.483	8567.46	8752.684	8783.759
5*	3,473	8211.267**	8248.183**	8428.97 **	8466.005**
6	3,322	8638.801	8681.56	8546.631	8592.86

Table 3:

GDP Purchasing Power Parity

Equation (a)

Linear regression

dlngrowth	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
L.dlngrowth	0.371	0.014	25.70	0.000	0.343	0.400	***
L2.dlngrowth	0.089	0.013	6.62	0.000	0.063	0.115	***
L.dflfp	-0.001	0.001	-0.61	0.540	-0.002	0.001	
Constant	0.020	0.001	22.42	0.000	0.018	0.022	***
Mean dependent var		0.037	SD dependent var			0.050	
R-squared		0.205	Number of obs			3926.000	
F-test		337.744	Prob > F			0.000	
Akaike crit. (AIC)		-13209.976	Bayesian crit. (BIC)			-13184.874	

*** *p<0.01*, ** *p<0.05*, * *p<0.1*

Equation (b)

Linear regression

dflfp	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
L.dflfp	0.120	0.017	7.19	0.000	0.087	0.153	***
L2.dflfp	0.113	0.017	6.72	0.000	0.080	0.145	***
L3.dflfp	0.084	0.017	4.94	0.000	0.050	0.117	***
L4.dflfp	0.098	0.017	5.69	0.000	0.064	0.132	***
L5.dflfp	0.081	0.017	4.73	0.000	0.047	0.115	***
L.dlngrowth	0.542	0.285	1.90	0.057	-0.016	1.100	*
Constant	0.034	0.018	1.91	0.057	-0.001	0.069	*
Mean dependent var		0.123	SD dependent var			0.826	
R-squared		0.092	Number of obs			3473.000	
F-test		58.229	Prob > F			0.000	
Akaike crit. (AIC)		8209.636	Bayesian crit. (BIC)			8252.705	

*** *p<0.01*, ** *p<0.05*, * *p<0.1*

GDP Local Currency Unit

Equation (a)

Linear regression

dlngrowth	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
L.dlngrowth	0.374	0.014	26.12	0.000	0.346	0.402	***
L2.dlngrowth	0.088	0.013	6.61	0.000	0.062	0.114	***
L.dflfp	0.000	0.001	-0.52	0.600	-0.002	0.001	
Constant	0.020	0.001	22.60	0.000	0.018	0.022	***

Mean dependent var	0.036	SD dependent var	0.050
R-squared	0.208	Number of obs	4004.000
F-test	349.229	Prob > F	0.000
Akaike crit. (AIC)	-13521.256	Bayesian crit. (BIC)	-13496.076

*** *p<0.01, *** *p<0.05, ** *p<0.1*

Equation (b)

Linear regression

dflfp	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
L.dflfp	0.165	0.017	9.82	0.000	0.132	0.198	***
L2.dflfp	0.087	0.017	5.11	0.000	0.053	0.120	***
L3.dflfp	0.070	0.017	4.11	0.000	0.036	0.103	***
L4.dflfp	0.109	0.017	6.31	0.000	0.075	0.143	***
L5.dflfp	0.057	0.017	3.29	0.001	0.023	0.091	***
L.dlngrowth	0.440	0.286	1.54	0.124	-0.120	1.000	
Constant	0.038	0.018	2.17	0.030	0.004	0.073	**

Mean dependent var	0.113	SD dependent var	0.833
R-squared	0.091	Number of obs	3542.000
F-test	59.246	Prob > F	0.000
Akaike crit. (AIC)	8428.597	Bayesian crit. (BIC)	8471.804

*** *p<0.01, *** *p<0.05, ** *p<0.1*

Table 4:

Equation (a)

Granger non-causality test results: GDP PPP	
Lag order: AIC/BIC	
W-bar	1.1732
Z-bar	1.5049 (p-value = 0.1324)
Z-bar tilde	0.6112 (p-value = 0.5411)

Granger non-causality test results: GDP Local Currency	
Lag order: AIC/BIC	
W-bar	1.4273
Z-bar	3.7492 (p-value = 0.0002)
Z-bar tilde	2.5229 (p-value = 0.0116)

H_0: dflfp does not Granger-cause dlngrowth.
H_1: dflfp does Granger-cause dlngrowth

Equation (b)

Granger non-causality test results: GDP PPP	
Lag order: AIC/BIC	
W-bar	1.9672
Z-bar	8.4036 (p-value = 0.0000)
Z-bar tilde	6.5082 (p-value = 0.0000)

Granger non-causality test results: GDP Local Currency	
Lag order: AIC/BIC	
W-bar	2.3790
Z-bar	12.1007 (p-value = 0.0000)
Z-bar tilde	9.6617 (p-value = 0.0000)

H_0: dlngrowth does not Granger-cause dflfp.
H_1: dlngrowth does Granger-cause dflfp