Working Papers in Economics

The ARDL Test of Gender Kuznets Curve for G7 Countries

Dilara Kılınç*
* Department of Economics, Izmir University of Economics, 35330, Izmir, Turkey

Esra Onater**
** Department of Economics, Izmir University of Economics, 35330, Izmir, Turkey

Hakan Yetkiner***
** Department of Economics, Izmir University of Economics, 35330, Izmir, Turkey

Working Paper # 13/05

July 2013

Izmir University of Economics

Department of Economics

Sakarya Cad. No:156

35330 Balçova Izmir

Turkey

Working Paper # 13/05

* Corresponding author. Email. kilinc.dilara@std.ieu.edu.tr
** Email esra.onater@ieu.edu.tr.
*** Email Hakan.Yetkiner@ieu.edu.tr.

The ARDL Test of Gender Kuznets Curve for G7 Countries

Abstract

The Gender Kuznets Curve (GKC) hypothesis argues that economic development has a non-linear effect on the female share of workers. There is, however, growing debate on the exact shape of this non-linear relationship. The aim of this paper is to test the GKC hypothesis in order to determine whether data supports a quadratic or a cubic GKC for each G7 countries in the long run. The ARDL bounds testing approach of cointegration yields evidence for the following: Canada, United Kingdom and United States have an inverted U-shaped GKC; Japan has an S-shaped GKC and France has an inverted-S shaped GKC; and finally that Italy and Germany have no long run GKC relationship in the respective periods of countries considered. We conclude that gender equality is not a direct result of development, and therefore policy makers having a gender equalization policy need to subsidize the employment of female workers in periods of fall.

Keywords: Gender Kuznets Curve; Economic Development; ARDL.
JEL classification: J16, O47, C32

The ARDL Test of Gender Kuznets Curve

1. Introduction

Although female participation in the labor force has increased in the past few decades, the gender inequality still exists. Males have greater employment opportunities than females in even the most developed countries. Let us take the most advanced economies in income and social transformation in the world, namely G7 countries.[1] According to the Global Gender Gap Report (2012), the female-to-male ratio of labor force participation is 0.90 for Canada, 0.87 for both France and Germany, 0.70 for Italy, 0.73 for Japan, 0.84 for UK, and 0.85 for the USA in 2012. Then, the succeeding question is how the female labor force participation (interchangeably, feminization or gender equality) changes during the development stages of these advanced economies. One conceivable approach is to use the 'Kuznets Curve of Income Inequality' idea, originated by Kuznets (1955), which was the first showing the non-linear relationship between income inequality and the level of development. This is because in the early stages of economic development, a relatively small section of labor force is employed in high productivity sectors, leading to income inequality. As economic development progresses, more are employed in the high productivity sectors, leading to a reduction in the income gap. Correspondingly, if the female labor force participation is expected to respond to economic development in such a non-linear way, the 'Gender Kuznets Curve' hypothesis, shortly GKC, provide an explanation for this. Thus, it is intuitively expected that gender equality (gender gap) increases (decreases) in the initial stages of development, but starts to decline (increase) as development progresses.

A general impression can be gained by examining the scatter plot of female-to-male ratio of labor force participation against income per worker, which is shown in figure 1 below.[2,3] Raw data from G7 countries, the most advanced economies in income and social transformation in the world, shows that there is perhaps a non-linear relationship between the gender equality and level of development but it is not necessarily quadratic for all. In particular, instead of the expected U-shaped relationship between gender equality and level of development, which would be a direct translation of original Kuznets hypothesis, gender equality may show some cubic response to development. An example of a possible response is an increase in gender equality due to social and political developments, caused by economic growth, followed by a plateau or even decrease, due to, for example, the rise of discriminatory institutions, and finally a rise caused by the evolution of new norms and institutions favoring gender equalization in employment. This pattern is known as 'S-shaped'.[4]

[1] Canada, France, Germany, Italy, Japan, United Kingdom (UK), the United States of America (USA).
[2] The time series data covers the period 1955-2010. The initial years concerned are not 1955 for all countries due to limitations in the availability of data. The initial year is 1956 for France, 1957 for Japan, 1980 for UK and 1970 for Germany and Italy. Authors can provide the data on request.
[3] Please note that the correct interpretation of vertical axis is gender equality in our case, as ultimate aim is to reach (at least) equal female-to-male ratio of labor force participation.
[4] cf., Eastin and Parakash (2013).

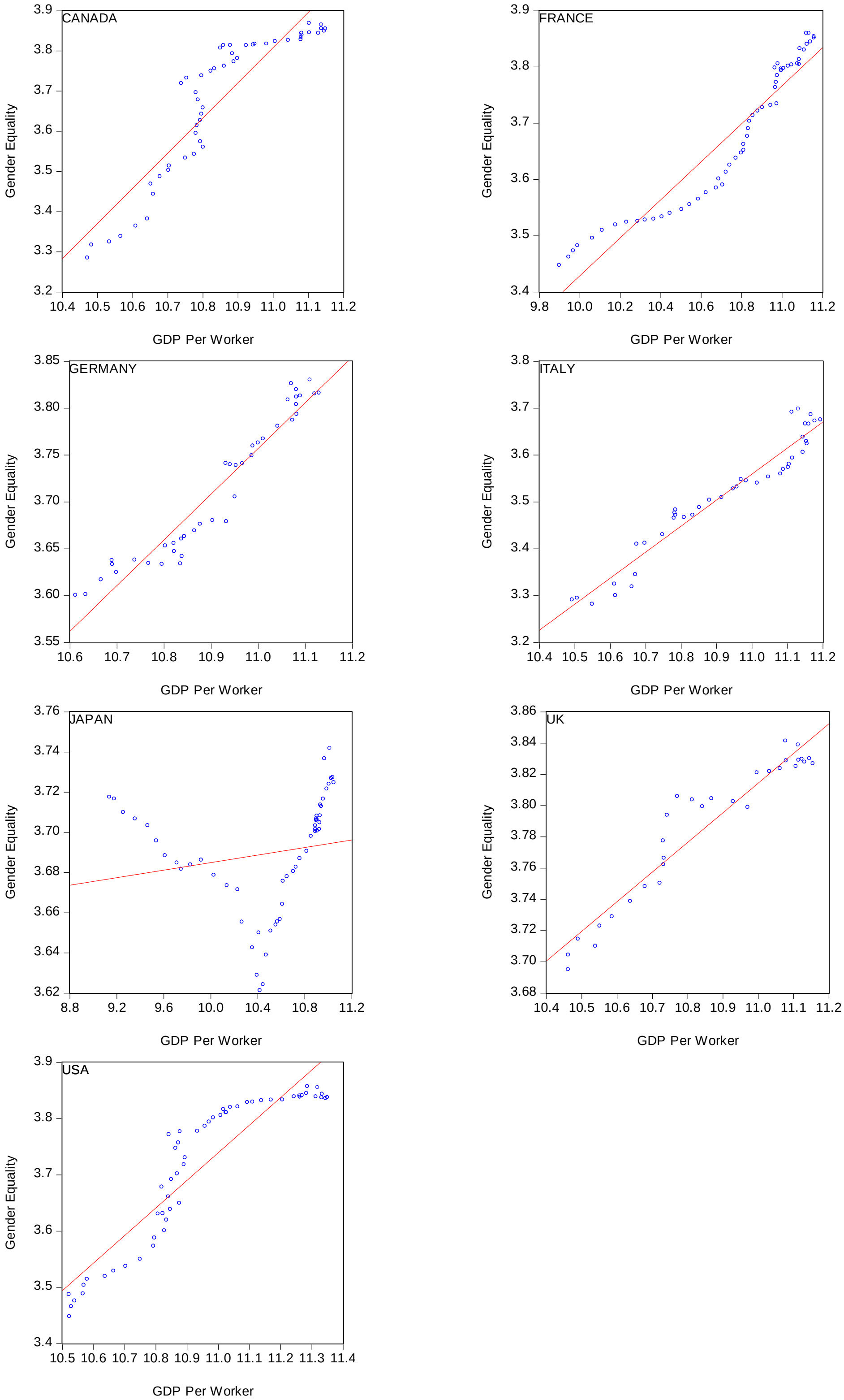


Figure 1. Share of Female Workers vs. Income per Worker
Source: International Labor Statistics and Heston et al. (2012)

The impact of economic development on gender equality has been subject to extensive research in the literature with very different conclusions. Some studies assert that there has been a linear relationship among gender inequality and economic development. For example, Weiss et al. (1976) examine the impact of economic structures on female labor force participation in a cross-national study and find that gender inequalities decrease in correspondence with increasing levels of industrialization and economic growth. Clark et al. (1991) show that the female share in labor force is more likely to increase in developed than in developing countries. Some of the cross-national studies find an adverse relationship between economic growth and gender inequality, including Clark (1991) and Charles (1992). On the other hand, several studies argue that the relationship is a quadratic one between gender equality and economic development. For example, adapting Kuznets' income inequality theory to the gender inequality case, Boserup (1970) argues that, while the gender equality decreases at the initial stages of economic development, it then increases when the country develops beyond a certain threshold. This finding coincides with the inverted-U shaped pattern of income inequality first revealed by Kuznets (1955). Goldin (1994) shows the existence of a quadratic relationship between female labor force participation and economic development via cross sectional data for more than 100 countries. Goldin (1994) concludes that the labor force participation of married females first declines, and then rises after a threshold of economic development is reached. Pampel and Tanaka (1986) estimate models using 70 nations at two different time points, 1965 and 1970, to determine how the economic development affects gender equality. They find that at the initial levels, development actually excludes females from the labor force, but at more advanced levels, increasing development causes an expansion of female labor force participation, which decreases gender inequality. Kottis (1990) concludes that the decrease in women's activity rates in Greece is explained by the U-shaped effect of economic development. Psacharopoulos and Tzannatos (1989) observe the implications of feminization of labor force for development for 136 countries in the early 1980s. The findings reveal that the female labor force participation initially decreases, and then increases after a certain level of economic development is achieved in the period of transition from agrarian subsistence economy. A test of the impact of economic growth on gender inequality with time-series data for a set of Asian countries by Lancitan et al. (1996) supports Kuznets's theory that income growth leads to an eventual decrease in gender inequality in the long run. Forsythe et al. (2000) find a longitudinal evidence of a curvilinear relationship between economic growth and gender inequality. They state that the economic development should aim to increase gender equality, since the increased status of women creates more social integration, and also improves investment in human capital. Tam (2011) makes dynamic panel data estimation for 130 countries over 31 years, and the results suggest that the U-shaped pattern in the relationship between female labor force participation and economic growth seems to hold.

Nevertheless, not all the studies agree on the U-shaped pattern of feminization of labor force during economic development. Durand (1975) claims that the U-shaped model of female labor force participation cannot be generalized for developing economies. Steel (1981) also finds no evidence of the U-shaped pattern in female labor force participation for Ghana in 1960s. Although the rapid initial growth in manufacturing leads to increases in the employment of women, Cagatay and Ozler (1995) find evidence that this early growth of women's employment declines after a threshold is reached, based on cross-country data for 1985 and 1990. They term the inverted U-shaped GKC as 'Feminization U'. Eastin and Prakash (2013) estimate a model including the cubic specification of income per capita to determine the behavior of gender equality in the subsequent stages of economic development for 146 countries for the period 1980-2005. The evidence suggests a curvilinear relationship of gender equality and economic development in the form of S-shaped GKC, in which the second and third phases coincide with the U-shaped pattern of gender equality.

The aim of this work is to test the long-run relationship between gender equality in employment and GDP per worker empirically for G7 countries, and to identify whether the data supports Kuznets curve that is quadratic (e.g., U-shaped), or a cubic (e.g., S-shaped). To this end, we use the recently developed Auto-Regressive Distributed Lag (ARDL) bounds testing approach to determine the long-run equilibrium cointegration relationship between gender equality (female share in employment) and economic development for each G7 country in the period 1955-2010. The advantage of this approach is that a long-run cointegration analysis is enabled, irrespective of whether the variables are trend or first difference stationary. As our GKC model includes share and level variables, the ARDL approach fits better than any other cointegration technique. The aim of this study is to consider the implications of the data ex post for the relationship between gender equality and economic development. To this end, we tested both quadratic and cubic forms of GKC, and found that the GKC is S-shaped (+,-,+) for Japan, inverted-S shaped (-,+,-) for France, and inverted-U shaped (+,-) for Canada, United Kingdom, and United States, á la Cagatay and Ozler (1995). These results have important implications for policy. Note that a curvilinear GKC has two implications. First, economic development does not directly involve gender equalization; whether cubic or quadratic, any country will experience periods of fall in gender equality, which require female employment to be subsidized. Second, and perhaps more serious, it is possible that economic development may ultimately result in lower levels of gender equality (e.g., the inverted U-shaped GKC for United States implies that gender equality declines as GDP per worker rises in the period covered). Hence, if a curvilinear GKC is the true specification, it is essential to subsidize gender equality in periods when it is falling.

Our results require careful interpretation in one respect: by history, the data intervals correspond to different development stages of G7 countries. For example, the 1955-2010 period perhaps corresponds to a different stage of economic development for USA, compared to, say, Japan or Italy. In that sense, though USA may show quadratic GKC in that period, its true GKC could be, for example, S-shaped, if a longer time period were considered. That is, over the period of considered, the G7 countries are not homogenous in terms of the stage of development. We therefore interpret our results with some reservation.

The organization of the paper is as follows. Section 2 introduces the data and the methodology. Section 3 is reserved for empirical analysis. Our results confirm that G7 countries show varying GKC patterns, leading to different policy implications. In particular, those with inverted U-shaped or inverted S-shaped GKC need more demanding policy intervention towards gender equalization in the further stages of development. Section 4 concludes the paper.

2. Data and Methodology

This paper studies G7 countries over the period 1955-2010. The dependent variable is the share of female workers in the age range of 15-64 in total employment, which is a direct measure of the feminization of labor force. The time-series data are drawn from various sources. In particular, total employment is compiled from The Conference Board of Total Economy Database,[5] the employed female population is obtained from OECD Stats and the ILO (International Labor Organization) Databases and GDP per worker (at 2005 international dollars) are obtained from the Heston et al. (2012) dataset. All the variables are expressed in their natural logarithmic levels. Table 1 presents the descriptive statistics of the feminization and level of development data for G7 countries.

Table 1. Descriptive Statistics of Data

		Share of Females in Employment (%)					GDP per worker (2005 I$/worker)				
Country	**# of Obs./ Period Incl.**	**Mean**	**Median**	**Max.**	**Min.**	**SD**	**Mean**	**Median**	**Max.**	**Min.**	**SD**
Canada	51 (1960-2010)	3.68	3.75	3.87	3.28	0.18	10.85	10.82	11.15	10.47	0.18
France	55 (1956-2010)	3.68	3.69	3.86	3.45	0.13	10.74	10.83	11.16	9.90	0.37
Germany	41 (1970-2010)	3.71	3.71	3.83	3.60	0.08	10.91	10.93	11.13	10.61	0.15
Italy	41 (1970-2010)	3.52	3.53	3.70	3.28	0.12	10.92	10.96	11.19	10.49	0.21
Japan	54 (1957-2010)	3.69	3.70	3.74	3.62	0.03	10.48	10.63	11.05	9.14	0.56
United Kingdom	31 (1980-2010)	3.79	3.80	3.84	3.70	0.05	10.85	10.84	11.15	10.46	0.23
United States	56 (1955-2010)	3.71	3.77	3.86	3.45	0.13	10.95	10.89	10.35	10.52	0.25

Note: Max., Min. and SD denote maximum, minimum and standard deviation, respectively. The series are in their natural logarithmic levels.

We will employ the following two equations to test whether there is a cubic or quadratic relationship between the feminization of labor force and GDP per worker in the long run:

$$Ln\ (\theta_{F,t}) = \lambda_{c0} + \lambda_{c1} Ln\ (y_t) + \lambda_{c2}\left(Ln\ (y_t)\right)^2 + \lambda_{c3}\left(Ln\ (y_t)\right)^3 + \epsilon_{ct} \quad (1a)$$

[5] http://www.conference-board.org/data/economydatabase/. We prefer the Conference Board instead of OECD or ILO for total employment data due to its comprehensiveness.

$$Ln\ (\theta_{F,t}) = \lambda_{q0} + \lambda_{q1} Ln\ (y_t) + \lambda_{q2}\ (Ln\ (y_t))^2 + \epsilon_{qt} \quad (1b)$$

where the coefficients λ_{cj}, $j = 1,2,3$ are the long-run elasticity estimations of female share in employment ($\theta_{F,t}$) with respect to GDP per worker (y_t), to the square of GDP per worker, and to the cube of GDP per worker, respectively. The subscripts c and q refer to cubic and quadratic, respectively. If the data does not support a cubic relationship, or if no statistically significant evidence is found via (1a), we run (1b) to test quadratic pattern of GKC. In (1b), the coefficients λ_{qi}, $i = 1,2$ stand for the long-run elasticity estimations of $\theta_{F,t}$ with respect to y_t and to the square of it. In addition, ϵ_t and the subscript t denote the error term and the time period index, respectively.

3. Empirical Analysis

3.1 Unit Root Tests

The stationary features of the variables are analyzed via unit root tests. We need to make sure that the order of integration of series is either I(0) or I(1) for the applicability of critical bounds test of Pesaran et al. (2001) or Narayan (2005). In this paper, for the sake of robustness, we used 2 different unit root tests, namely, augmented Dickey–Fuller (ADF) and Phillips–Perron (PP).[6] We run Schwarz Information Criterion (SIC) and Akaike Information Criterion (AIC) as lag selection criteria.[7] The bandwidths for PP test are selected with the Newey–West Bartlett kernel method. Tables 2a-2g illustrate the unit root test results of the series for seven countries. The tests cover an intercept and a linear trend in the levels, and an intercept in first difference.

---

[6] We also run Ng–Perron MZa (NP) test but did not present it, as the test yielded similar results. We refrained from using the Elliot–Rothenberg–Stock Dickey–Fuller GLS detrended (DF-GLS) test because the minimum critical value of the test (with intercept and trend) is for 50 observations, which Germany, Italy, and UK fails to satisfy.

[7] Here, we represent the results obtained via SIC since both of the lag selection criteria yield similar results.

Table 2a. Unit Root Test Results for Canada

Variables	ADF-test	PP-test
Panel A: Level (Intercept and Trend)		
$Ln\theta_F$	-1.318(0)	-1.309(3)
Lny	-3.185(1)*	-2.275(0)
Lny^2	-3.101(1)	-2.206(0)
Lny^3	-3.017(1)	-2.139(0)
Panel B: First difference (Intercept, no trend)		
$Ln\theta_F$	-2.224(1)	-4.321(4)***
Lny	-4.640(0)***	-4.446(5)***
Lny^2	-4.657(0)***	-4.457(5)***
Lny^3	-4.673(0)***	-4.466(5)***

Table 2b. Unit Root Test Results for France

Variables	ADF-test	PP-test
Panel A: Level (Intercept and Trend)		
$Ln\theta_F$	-0.551(0)	0.985(3)
Lny	-1.724(0)	-1.724(0)
Lny^2	-1.498(0)	-1.498(0)
Lny^3	-1.291(0)	-1.291(0)
Panel B: First difference (Intercept, no trend)		
$Ln\theta_F$	-5.548(0)***	-5.565(2)***
Lny	-3.956(0)***	-3.944(3)***
Lny^2	-4.087(0)***	-4.095(3)***
Lny^3	-4.226(0)***	-4.250(3)***

Table 2c. Unit Root Test Results for Germany

Variables	ADF-test	PP-test
Panel A: Level (Intercept and Trend)		
$Ln\theta_F$	-1.848(0)	-2.025(2)
Lny	-2.620(0)	-2.403(5)
Lny^2	-2.656(0)	-2.438(5)
Lny^3	-2.696(0)	-2.475(5)
Panel B: First difference (Intercept, no trend)		
$Ln\theta_F$	-5.101(0)***	-5.001(6)***
Lny	-5.841(0)***	-6.051(9)***
Lny^2	-5.887(0)***	-6.140(9)***
Lny^3	-5.934(0)***	-6.322(10)***

Table 2d. Unit Root Test Results for Italy

Variables	ADF-test	PP-test
Panel A: Level (Intercept and Trend)		
$Ln\theta_F$	-3.492(3)*	-1.914(3)
Lny	0.099(2)	0.109(8)
Lny^2	0.152(2)	0.159(8)
Lny^3	0.194(2)	0.198(8)
Panel B: First difference (Intercept, no trend)		
$Ln\theta_F$	-3.363(5)**	-5.876(3)***
Lny	-4.035(1)***	-5.391(2)***
Lny^2	-4.051(1)***	-5.367(2)***
Lny^3	-4.067(1)***	-5.341(2)***

Table 2e. Unit Root Test Results for Japan

Variables	ADF-test	PP-test
Panel A: Level (Intercept and Trend)		
$Ln\theta_F$	-2.170(1)	-1.864(3)
Lny	-2.029(0)	-1.920(2)
Lny^2	-1.608(0)	-1.551(2)
Lny^3	-1.243(0)	-1.231(2)
Panel B: First difference (Intercept, no trend)		
$Ln\theta_F$	-4.578(0)***	-4.590(3)***
Lny	-3.241(0)**	-3.127(3)**
Lny^2	-3.464(0)**	-3.387(3)**
Lny^3	-3.696(0)***	-3.654(3)***

Table 2f. Unit Root Test Results for United Kingdom

Variables	ADF-test	PP-test
Panel A: Level (Intercept and Trend)		
$Ln\theta_F$	-1.287(0)	-1.222(1)
Lny	-2.054(3)	-0.422(2)
Lny^2	-2.124(3)	-0.490(2)
Lny^3	-2.188(3)	-0.562(2)
Panel B: First difference (Intercept, no trend)		
$Ln\theta_F$	-5.364(0)***	-5.371(3)***
Lny	-1.342(2)	-3.986(2)***
Lny^2	-1.335(2)	-4.024(2)***
Lny^3	-1.328(2)	-4.061(2)***

Table 2g. Unit Root Test Results for United States

Variables	ADF-test	PP-test
Panel A: Level (Intercept and Trend)		
$Ln\theta_F$	0.311(0)	-1.556(3)
Lny	-2.307(1)	-2.013(2)
Lny^2	-2.295(1)	-2.011(2)

Lny^3	-2.278(1)	-2.003(2)
Panel B: First difference (Intercept, no trend)		
$Ln\theta_F$	-2.562(1)	-4.127(4)***
Lny	-5.842(0)***	-5.853(1)***
Lny^2	-5.818(0)***	-5.792(2)***
Lny^3	-5.793(0)***	-5.766(2)***

Notes: The null hypothesis is the existence of unit root for ADF and PP tests. In the tables, superscripts ***, **, * in bold denote the rejection of the null hypothesis at 1%, 5% and 10% significance levels, respectively. ADF and PP critical values are due to MacKinnon (1996). Lag lengths for ADF test and bandwidths for PP test are in parentheses.

We ensure that all variables are either I(0) or I(1) and that satisfy the underlying conditions of the ARDL bounds testing approach of cointegration with ADF and PP unit root tests. The results reveal that, for both tests, almost all the series are non-stationary at levels and stationary in first differences, though at different significance levels. Hence, in general, all variables are integrated of order one.

Next, we run Zivot and Andrews (1992), hereafter ZA, unit root test to determine the structural breaks and their dates. Perron (1989) argues that most of the structural breaks in macroeconomic time series are actually around a deterministic trend if a potential change in intercept or slope is allowed. However, the traditional unit root tests do not concern structural changes in testing the stationarity of series. The ZA testing procedure allows the determination of an endogenous structural break at an estimated point, either in the intercept, in the linear trend, or in both, and the determination of the order of integration of a series simultaneously. Tables 3a-3g present the results of ZA unit root tests for the series in the study for seven countries.

Table 3a. ZA Unit Root Test Results for Canada

Panel A:Level	Intercept	Trend	Both
$Ln\theta_F$	-3.97(4)[1979]	-4.23*(4)[1986]	-4.42(4)[1982]
Lny	-4.81*(1)[1980]	-3.96(1)[1992]	-4.23(1)[1980]
Lny^2	-4.76*(1) [1980]	-3.95(1) [1992]	-4.23(1)[1980]
Lny^3	-4.70* (1) [1980]	-3.94(1) [1992]	-4.22(1)[1980]
Panel B:First Difference	Intercept	Trend	Both
$Ln\theta_F$	-6.20***(3)[2002]	-6.40***(3)[1998]	-6.47***(3)[1993]
Lny	-5.35***(0)[1992]	-5.17***(0)[1976]	-6.01***(0)[1983]
Lny^2	-5.35***(0)[1992]	-5.15***(0)[1976]	-5.97***(0)[1983]
Lny^3	-5.34***(0)[1992]	-5.13***(0)[1976]	-5.92***(0)[1983]

Table 3b. ZA Unit Root Test Results for France

Panel A:Level	Intercept	Trend	Both
$Ln\theta_F$	-2.34(1)[1995]	-2.80(1)[1993]	-3.60(1)[1990]
Lny	-2.10(0)[1990]	-3.92(0)[1972]	-3.62(0)[1969]
Lny^2	-1.95(0)[1990]	-3.87(0)[1972]	-3.55(0)[1969]
Lny^3	-	-3.80(0)[1972]	-3.47(0)[1969]
Panel B:First Difference	Intercept	Trend	Both
$Ln\theta_F$	-7.01***(0)[1995]	-6.06***(0)[1994]	-7.14***(0)[1995]
Lny	-6.91***(0)[1975]	-6.18***(0)[1982]	-6.83***(0)[1975]
Lny^2	-6.86***(0)[1975]	-6.15***(0)[1982]	-6.79***(0)[1975]
Lny^3	-6.79***(0)[1975]	-6.11***(0)[1982]	-6.75***(0)[1975]

Table 3c. ZA Unit Root Test Results for Germany

Panel A:Level	Intercept	Trend	Both
$Ln\theta_F$	-4.73*(1)[1990]	-3.19(1)[1979]	-5.91***(1)[1990]
Lny	-3.82(1)[2004]	-4.00(1)[2002]	-4.03(1)[2000]
Lny^2	-3.85(1)[2004]	-4.03(1)[2002]	-4.07(1)[2000]
Lny^3	-3.87(1)[2004]	-4.04(1)[2002]	-4.11(1)[2000]
Panel B:First Difference	Intercept	Trend	Both
$Ln\theta_F$	-6.02***(1)[1990]	-5.46***(1)[1992]	-6.41***(1)[1990]
Lny	-5.64***(1)[1980]	-5.30***(1)[1983]	-5.76***(1)[1980]
Lny^2	-5.62***(1)[1980]	-5.31***(1)[1983]	-5.74***(1)[1980]
Lny^3	-5.62***(1)[1980]	-5.31***(1)[1983]	-5.73***(1)[1980]

Table 3d. ZA Unit Root Test Results for Italy

Panel A:Level	Intercept	Trend	Both
$Ln\theta_F$	-5.60***(0)[1977]	-4.44**(0)[1980]	-
Lny	0.72(4)[2003]	-2.26(4)[2001]	-2.09(4)[2000]
Lny^2	0.68(4)[2003]	-2.31(4)[2001]	-2.11(4)[2000]
Lny^3	0.64(4)[2003]	-2.33(4)[2001]	-2.12(4)[2000]
Panel B:First Difference	Intercept	Trend	Both
$Ln\theta_F$	-6.67***(0)[1981]	-5.12***(1)[1986]	-7.70***(0)[1981]
Lny	-6.52***(4)[1988]	-4.90***(4)[1998]	-6.42***(4)[1988]
Lny^2	-6.39***(4)[1988]	-4.91***(4)[1998]	-6.27***(4)[1988]
Lny^3	-6.24***(4)[1988]	-4.92***(4)[1998]	-6.11***(4)[1988]

Table 3e. ZA Unit Root Test Results for Japan

Panel A:Level	Intercept	Trend	Both

Table 3f. ZA Unit Root Test Results for United Kingdom

Panel A:Level	Intercept	Trend	Both

$Ln\theta_F$	-3.70(1)[1968]	-3.78(1)[1975]	-4.49(1)[1971]
Lny	-3.18(0)[1966]	-4.12*(0)[1970]	-4.04(0)[1967]
Lny^2	-2.96(0)[1966]	-3.82(0)[1970]	-3.74(0)[1967]
Lny^3	-2.74(0)[1966]	-3.52(0)[1970	-3.45(0)[1967]
Panel B:First Difference	Intercept	Trend	Both
$Ln\theta_F$	-7.46***(0)[1976]	-5.67***(0)[1983]	-7.48***(0)[1976]
Lny	-6.81***(0)[1974]	-5.92***(0)[1978]	-6.61***(0)[1973]
Lny^2	-6.72***(0)[1974]	-5.83***(0)[1978]	-6.53***(0)[1973]
Lny^3	-6.59***(0)[1974]	-5.76*** (0)[1978)	-6.44***(0)[1973]

$Ln\theta_F$	-2.61(0)[1989]	-4.25*(0)[1994]	-4.82*(0)[1992]
Lny	-2.42(3)[2006]	-3.60(3)[2005]	-3.54(3)[2005]
Lny^2	-2.42(3)[2006]	-3.50(3)[2006]	-3.45(3)[2006]
Lny^3	-2.41(3)[2006]	-3.40(3)[2006]	-3.36(3)[2006]
Panel B:First Difference	Intercept	Trend	Both
$Ln\theta_F$	-7.08***(0)[1994]	-7.08***(0)[1987]	-7.23***(0)[1994]
Lny	-5.49***(0)[1993]	-4.91***(0)[2001]	-5.39**(0)[1997]
Lny^2	-5.47*** (0)[1993]	-4.96*** (0)[2001]	-5.43**(0)[1997]
Lny^3	-3.34(2))[1994]	-2.44(2)[2000]	-3.27(2)[1994]

Table 3g. ZA Unit Root Test Results for United States

Panel A:Level	Intercept	Trend	Both
$Ln\theta_F$	-1.83(1)[1966]	-4.54**(1)[1983]	-4.01(1)[1978]
Lny	-3.38(1)[1974]	-	-3.84(1)[1974]
Lny^2	-3.47(1)[1974]	-2.54(1)[1983]	-3.84(1)[1974]
Lny^3	-3.55(1)[1974]	-2.59(1)[1983]	-3.83(1)[1974]
Panel B:First Difference	Intercept	Trend	Both
$Ln\theta_F$	-6.60***(0)[1983]	-6.28***(0)[1967]	-7.01***(0)[1983]
Lny	-6.63***(0)[1983]	-5.91***(0)[2000]	-6.59***(0)[1983]
Lny^2	-6.59***(0)[1983]	-5.88***(0)[2000]	-6.54***(0)[1983]
Lny^3	-6.55***(0)[1983]	-5.86***(0)[2000]	-6.50***(0)[1983]

Notes: The null hypothesis is the existence of unit root with a structural break in intercept, trend, or both. The critical values are due to Zivot and Andrews (1992). Superscripts ***, **, * denote the stationarity for the ZA unit root test at 1%, 5% and 10% critical levels, respectively. Lag lengths are in parenthesis and the date in square brackets denote the time of the structural change. "-" represents cases in which a result cannot be obtained due to data problem.

3.2 ARDL Cointegration Analysis

ARDL is a dynamic single equation regression model developed by Pesaran and Shin (1999) and Pesaran et al. (2001) in which the lagged values of the dependent variable and the current and lagged values of the regressors are included in order to directly estimate short-run elasticities, and indirectly, the long-run equilibrium relationship (Wang et al., 2011). Compared to both the residual-based Engle and Granger (1987) test, and the maximum likelihood test of Johansen (1988) and Johansen and Juselius (1990), ARDL specification for cointegration analysis has the advantage that there is no necessity for the variables to be of the same order of integration. The series can be either I(0) or I(1), that is, the technique allows for the series in the system to have different optimal lag orders. It provides efficient estimates irrespective of small sample sizes, and endogenous explanatory variables. The model corrects the endogeneity problem of explanatory variables even in small samples (Menyah and Wolde-Rufael, 2010).

The optimal lag length must be first determined, as the ARDL results are relatively sensitive to lag lengths. In this paper, the selection of the optimal lag is based on Schwarz – Bayesian information criterion (SBIC).[8] The ARDL model is employed in three steps once the order of integration of the system is determined. The first step is to test for the existence of cointegration among the variables via the bounds testing approach. The following equations are employed in order to examine the cubic and quadratic non-linear relationship between the feminization of the labor force and economic development in the long run:

$$\Delta Ln\,\theta_{F,t} = \alpha_{c1} + \sum_{i=1}^{p}\beta_{c1i}\,\Delta Ln\,\theta_{F,t-i} + \sum_{j=0}^{r}\gamma_{c1j}\,\Delta Lny_{t-j} + \sum_{k=0}^{s}\varphi_{c1k}\,\Delta Lny_{t-k}^{2} + \sum_{l=0}^{z}\rho_{c1l}\,\Delta Lny_{t-l}^{3} + \zeta_{c1}Ln\theta_{F,t-1} + \zeta_{c2}Lny_{t-1} + \zeta_{c3}Lny_{t-1}^{2} + \zeta_{c4}Lny_{t-1}^{3} + \eta_{c1t} \qquad (2a)$$

$$\Delta Ln\,\theta_{F,t} = \alpha_{q1} + \sum_{i=1}^{p}\beta_{q1i}\,\Delta Ln\,\theta_{F,t-i} + \sum_{j=0}^{r}\gamma_{q1j}\,\Delta Lny_{t-j} + \sum_{k=0}^{s}\varphi_{q1k}\,\Delta Lny_{t-k}^{2} + \zeta_{q1}Ln\theta_{F,t-1} + \zeta_{q2}Lny_{t-1} + \zeta_{q3}Lny_{t-1}^{2} + \eta_{q1t} \qquad (2b)$$

where η_{c1t} and η_{q1t} denote the white noise error terms for cubic and quadratic forms of the model respectively, and Δ is the first difference operator. The

[8] Pesaran and Shin (1999) state that SBIC is more consistent than Akaike information criterion (AIC) and Hannan – Quinn information criterion (HQ). In addition, Monte Carlo evidence shows that SBIC and AIC determines reliable lag order (Panopoulou and Pittis, 2004; Emran et al., 2007).

parameters β, γ, and φ are the short-run coefficients, and ζ_{ci}, $i = 1,2,3,4$ and ζ_{qj}, $j = 1,2,3$ are the long-run coefficients of the ARDL model. The bounds testing approach is based on the joint F or Wald statistics, testing the significance of the lagged levels of the variables via the null hypothesis of no cointegration, $\mathrm{H}_0\colon \zeta_{\mathrm{c1,c2,c3,c4}} = 0$ against the alternative of the existence of cointegration, $\mathrm{H}_1\colon \zeta_{\mathrm{c1,c2,c3,c4}} \neq 0$. If the cointegrating relation is not found for cubic specification, the same procedure is applied for quadratic specification, e.g., the null hypothesis of no cointegration, $\mathrm{H}_0\colon \zeta_{\mathrm{q1,q2,q3}} = 0$ against the alternative of the existence of cointegration, $\mathrm{H}_1\colon \zeta_{\mathrm{q1,q2,q3}} \neq 0$. The asymptotic distributions of two sets are in Pesaran et al. (2001), and its modified version for small samples, ranging from 30 to 80, are presented in Narayan (2005). This study employs the critical values of Narayan (2005) for the bounds F-statistics due to the limited annual time series data on employed female population, employed population, and GDP per worker for the seven countries. The results of the bounds F-test for cointegration, together with critical values are reported in Table 4.

Table 4. The bounds F-test for cointegration for the estimated ARDL Specification

Cubic Specification				**Quadratic Specification**			
Country	**Period**	**Model**	**F-Statistics**	**Country**	**Period**	**Model**	**F-Statistics**
France	**1956-2010**	**1,1,1,0**	**4.827****	**Canada**	**1960-2010**	**1,1,0**	**9.199*****
Italy	1970-2010	1,0,0,0	3.119	Germany	1970-2010	1,1,0	2.076
Japan	**1957-2010**	**2,0,0,0**	**6.953*****	Italy	1970-2010	1,0,0	1.584
United Kingdom	1980-2010	1,1,0,0	2.794	**United Kingdom**	**1980-2010**	**1,1,0**	**3.787***
United States	1955-2010	1,0,0,1	4.759**	**United States**	**1955-2010**	**1,0,1**	**5.516****
		I(0)	I(1)			I(0)	I(1)
Critical values at 1%		4.614	5.966	Critical values at 1%		5.155	6.265
Critical values at 5%		3.272	4.306	Critical values at 5%		3.538	4.428
Critical values at 10%		2.676	3.586	Critical values at 10%		2.915	3.695

Notes: F-statistics are obtained from the ARDL cointegration test. The critical values for the lower I(0) and upper I(1) are due to Narayan (2005): see Case II in appendix for n=30 and k=3 for cubic relationship, k=2 for quadratic relationship. The superscripts *******, ******, ***** in bold denote significance levels at 1%, 5%, and 10%, respectively. The bounds F-test for cointegration do not work in cubic specification for Canada and Germany due to the data problems. The USA is run for both cubic and quadratic specifications, as the former specification does not imply any significant regressor, though it passes the bounds F-test.

The F-test has a non-standard distribution that depends on the number of independent variables, whether the variables in the system are I(0) or I(1) and whether the model includes an intercept and/or a trend (Narayan, 2005). The upper value supposes that variables are I(1) and lower value supposes that the variables in the system are I(0) in nature (Pesaran et al., 2001). If the computed

F-statistics is above the upper bound, the null hypothesis is rejected at the significance level of the concerning bound. This indicates the existence of cointegration among variables. If the computed F-statistics is below the lower bound, the null hypothesis fails to be rejected, implying the absence of cointegration among variables. The cointegration test is inconclusive if the computed F-statistics falls between the critical values. The bounds F-test for cointegration analysis yields evidence of a cubic GKC for France, United States, and Japan and a quadratic GKC for Canada and United Kingdom. No long-run relationship is found for Germany and Italy, irrespective of the GKC specification (see table 4). We run also a quadratic GKC for USA, as none of the regressors are found statistically significant in the cubic specification (see table 5).

Once the cointegration among the variables is confirmed, the subsequent procedure is to estimate the long-run coefficients (Equations 3a, 3b), and the short-run coefficients (Equations 4a, 4b) via the ARDL approach and the Error-Correction Model (ECM) for the associated ARDL:

$$Ln\,\theta_{F,t} = \alpha_{c2} + \sum_{i=1}^{p} \beta_{c2i}\, Ln\,\theta_{F,t-i} + \sum_{j=0}^{r} \gamma_{c2j}\, Lny_{t-j} + \sum_{k=0}^{s} \varphi_{c2k}\, Lny_{t-k}^{2} + \sum_{l=0}^{z} \rho_{c2l}\, Lny_{t-l}^{3} + \eta_{c2t} \tag{3a}$$

$$Ln\,\theta_{F,t} = \alpha_{q2} + \sum_{i=1}^{p} \beta_{q2i}\, Ln\,\theta_{F,t-i} + \sum_{j=0}^{r} \gamma_{q2j}\, Lny_{t-j} + \sum_{k=0}^{s} \varphi_{q2k}\, Lny_{t-k}^{2} + \eta_{q2t} \tag{3b}$$

$$Ln\,\theta_{F,t} = \alpha_{c3} + \sum_{i=1}^{p} \beta_{c3i}\, \Delta Ln\,\theta_{F,t-i} + \sum_{j=0}^{r} \gamma_{c3j}\, \Delta Lny_{t-j} + \sum_{k=0}^{s} \varphi_{c3k}\, \Delta Lny_{t-k}^{2} + \sum_{l=0}^{z} \rho_{c2l}\, \Delta Lny_{t-l}^{3} + \mu\, ECT_{t-1} + \eta_{c3t} \tag{4a}$$

$$Ln\,\theta_{F,t} = \alpha_{q3} + \sum_{i=1}^{p} \beta_{q3i}\, \Delta Ln\,\theta_{F,t-i} + \sum_{j=0}^{r} \gamma_{q3j}\, \Delta Lny_{t-j} + \sum_{k=0}^{s} \varphi_{q3k}\, \Delta Lny_{t-k}^{2} + \mu\, ECT_{t-1} + \eta_{q3t} \tag{4b}$$

where μ is the coefficient of the Error-Correction term (hereafter ECT), and it should be statistically significant and negative. ECT determines the speed of convergence of the variables to the equilibrium.[9] The long run and short run coefficients are presented in table 5 below.

[9] The ECT for cubic and quadratic specifications is defined as follows:
$ECT_t = Ln\,\theta_{F,t} - \sum_{i=1}^{p} \beta_{c2i}\, Ln\,\theta_{F,t-i} - \sum_{j=0}^{r} \gamma_{c2j}\, Lny_{t-j} - \sum_{k=0}^{s} \varphi_{c2k}\, Lny_{t-k}^{2} - \sum_{l=0}^{z} \rho_{c2l}\, Lny_{t-l}^{3}$
$ECT_t = Ln\,\theta_{F,t} - \sum_{i=1}^{p} \beta_{q2i}\, Ln\,\theta_{F,t-i} - \sum_{j=0}^{r} \gamma_{q2j}\, Lny_{t-j} - \sum_{k=0}^{s} \varphi_{q2k}\, Lny_{t-k}^{2}$

Table 5. The Estimated Long run and Short run coefficients

Indicator/Country	Canada	France	Japan	United Kingdom	United States	United States
Structural Break	1992	1975	1974	1993	1983	1983
Estimated Long-run Elasticities						
Lny	27.1288(2.23)**	-142.9744(-1.94)*	19.5237(2.51)**	6.1261(2.73)***	37.8931(0.12)	21.3004(4.60)***
Lny^2	-1.2346(-2.22)**	13.3700(1.91)*	-1.9998(-2.63)***	-0.2759(-2.66)***	-2.4746(-0.09)	-0.9634(-4.51)***
Lny^3	-	-0.4153(-1.88)*	0.0681(2.74)***	-	0.0459(0.05)	-
Constant	-145.0988(-2.18)**	511.5941(1.97)*	-59.6685(-2.26)**	-30.1672(-2.50)**	-174.4783(-0.15)	-113.7837(-4.54)***
Dummy	0.0097(0.09)	-0.0550(-1.1576)	-0.1143(-3.43)***	0.0430(1.93)*	0.0256(0.30)	0.0276(0.34)
Estimated Short-run Elasticities						
$\mathrm{Ln}\theta_{\mathrm{F}}(-2)$	-	-	0.4770(4.53)***	-	-	-
$\mathrm{Lny}(-1)$	1.6504 (1.50)	-13.9963(-2.07)**	4.592(1.901)*	1.5286(2.16)**	1.9334(0.12)	1.1066(4.09)***
$\mathrm{Lny}^2(-1)$	-0.0831(-1.67)*	1.4753(2.24)**	-0.470(-1.95)*	-0.0751(-2.31)**	-0.1263(-0.09)	-0.0554(-4.49)***
$\mathrm{Lny}^3(-1)$	-	-0.0514(-2.40)**	0.016(2.00)**	-	0.0020(0.05)	-
$\mathrm{ECT}(-1)$	-0.0673(-3.43)***	-0.1237(-3.31)***	-0.235(-3.94)***	-0.2723(-2.49)**	-0.0510(-2.85)***	-0.0520(-3.53)***
Constant	-9.7622(-1.63)*	63.2800(2.51)**	-14.035(-1.77)*	-8.2148(-2.22)**	-8.9026(-0.16)	-5.9111(-4.02)***
Dummy	0.0007(0.09)	-0.0068(-1.16)	-0.027(-6.03)***	0.0117(1.95)*	0.0013(0.29)	0.0014(0.33)
ARDL Estimates						
Model	(1,1,0)	(1,1,1,0)	(2,0,0,0)	(1,1,0)	(1,0,0,1)	(1,0,1)
Adjusted R^2	0.9983	0.9984	0.9791	0.9876	0.999	0.9991
RSS	0.0021	0.0012	0.0008	0.0007	0.0008	0.0008

Notes: Year Dummy is the time of the significant structural break in intercept for the series Lny and Lny^2 is obtained from ZA unit root test. (-1) refers one lag of the associated variable. '-'denotes that variable does not take place in the model. t-statistics for coefficients are in paranthesis. RSS is the residual sum of squares. The superscripts *******, ******, ***** in bold denote significance at 1%, 5%, and 10% respectively. USA is included twice because cubic form does not yield statistically significant explanatory variables, though the equation passes F-test.

Given that there is a long run GKC for G7 countries, the next question is the exact form of that relationship in the long run. In particular, a quadratic GKC can either be U-shaped or inverse U-shaped, and a cubic one can be S-shaped or inverted S-shaped, although, in the absence of evidence to the contrary, S- or U-shaped GKC are assumed, in parallel to the Kuznets Curve of economic inequality. Table 5 presents our results. We find that Japan and France both have a cubic but mirror imaged GKC in the long run: Japan has an S-shaped ($\gamma > 0$, $\varphi < 0$ and $\rho > 0$) and France has an inverted S-shaped ($\gamma < 0$, $\varphi > 0$ and $\rho < 0$) GKC. All long run coefficients are significant at 1% level in Japan and at 10% level in France. On the other hand, data for Canada, United Kingdom and United States clearly fit quadratic, namely inverted U-shaped, GKC specification in the long run in the period studied. While long run coefficients are significant at 1% for United Kingdom and United States, they are significant at 5% level in Canada. These results have one very important policy implication; In the range of period covered for G7 countries, gender equality appear to have a nonlinear GKC, that is, falls are as common as increases. Hence, gender equality cannot be considered to be the natural result of development. There is a clear need for policy interventions, in particular, during the downfall periods of female employment in order to maintain levels of gender equality.

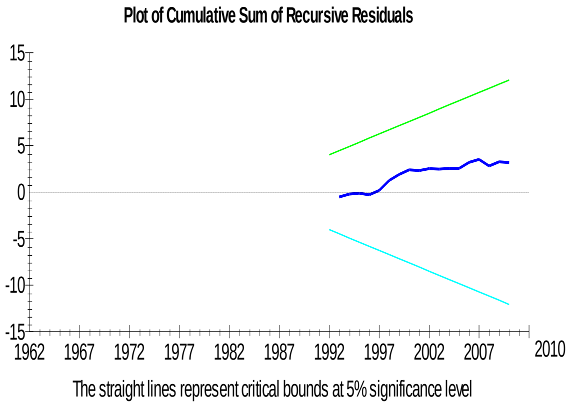
In the ECT model, the coefficient of the lagged residual shows the adjustment speed towards the equilibrium following a shock to the system. As expected, the estimated ECT coefficients are negative and statistically significant at 1% level for Canada, France, Japan, and United States, and at 5% level for United Kingdom. This indicates that the deviations from the long-run equilibrium among variables are corrected to return to the long-run equilibrium level for each period. The error correction terms highlight that, within the cointegration model, there is a correction of the disequilibrium conditions at the following approximate amounts: Canada 7%, France 12%, Japan 24%, United Kingdom 27%, and the USA 5%.

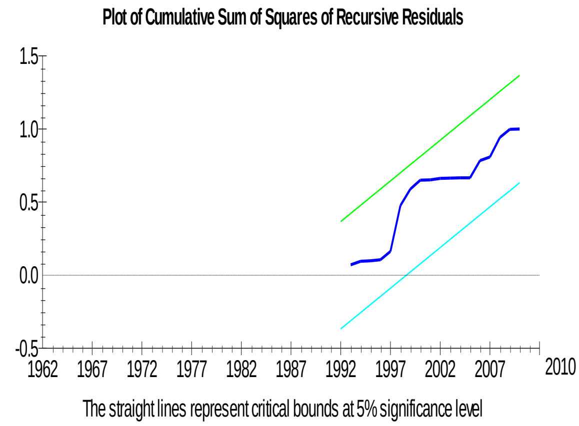
3.3 Stability of Long run and Short run Coefficients

The ZA unit root test results in section 3.1 have shown that there are structural breaks in regressors. The stability of the short run and long run coefficients are checked through the cumulative sum (CUSUM) and cumulative sum squares (CUSUMSQ) tests due to Brown et al. (1975). Figure 2 presents the plot of CUSUM and CUSUMSQ test statistics that fall inside the critical bounds of 5% significance level. This implies that the estimated parameters are stable over the period.

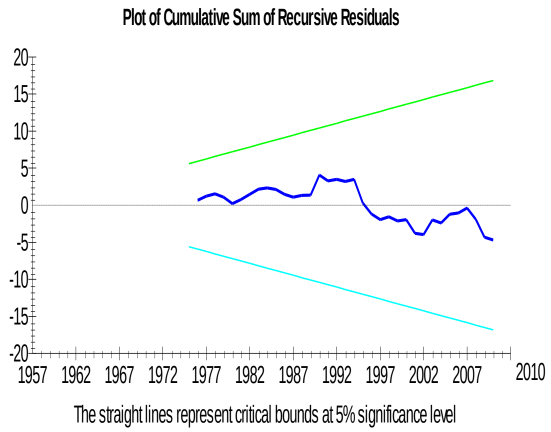
The ARDL Test of Gender Kuznets Curve

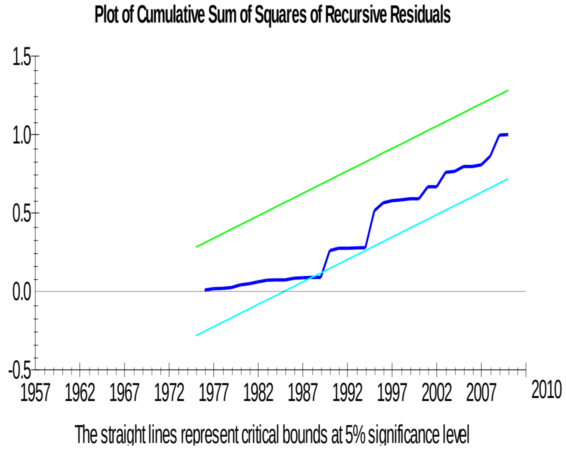
Canada



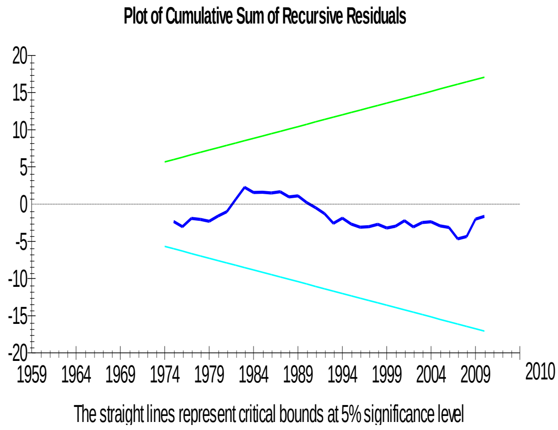


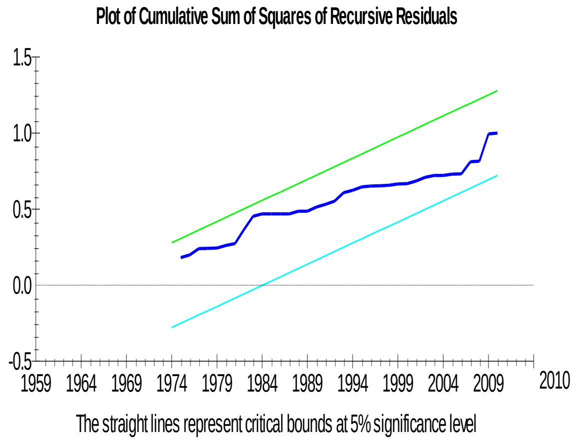
France



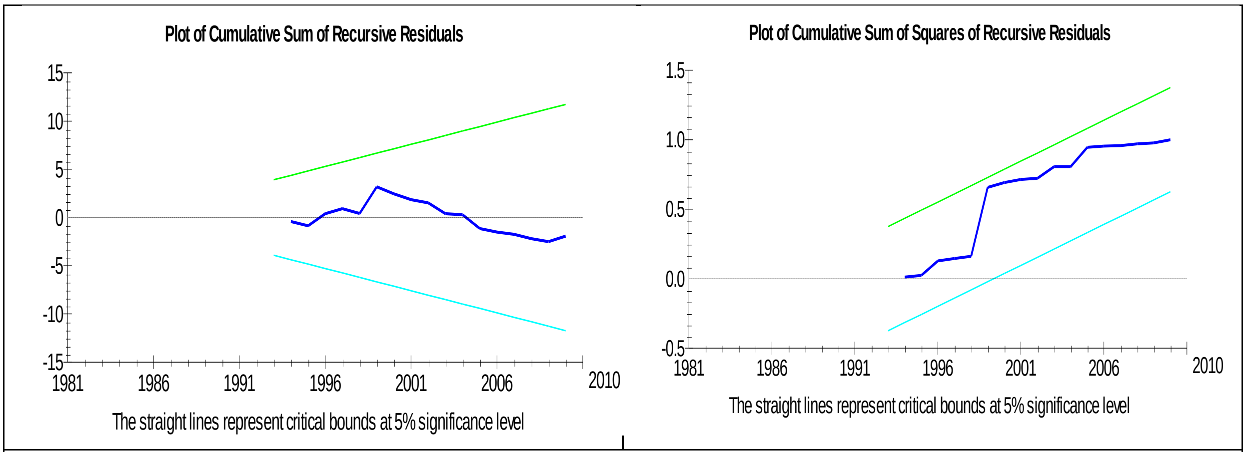


Japan





United Kingdom



United States

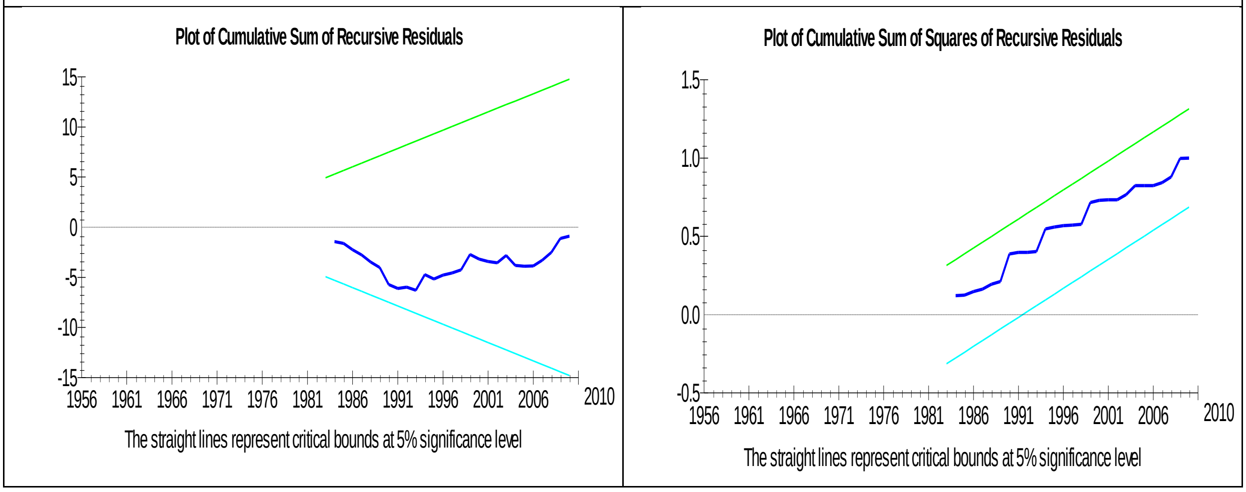


Figure 2. Plot of CUSUM and CUSUMSQ tests for the parameter stability

4. Concluding Remarks

This paper tests the form of long run Gender Kuznets Curve (GKC) for G7 countries. The bounds F-test for cointegration yields evidence that there is a cubic GKC relationship for France (inverted S-shaped) and Japan (S-shaped) and a quadratic GKC relationship for Canada, United Kingdom and United States (all inverted U-shaped). Hence, we argue that gender equalization needs to be subsidized, especially during the downfall periods. In addition to this, as economic development does not guarantee gender equalization, it is the duty of policy makers to ensure that subsidies and other incentive mechanisms are in order, when they are needed.

References

Brown, R.L., Durbin J., Evans J.M. (1975), "Techniques for testing the consistency ofregression relations over time". Journal of the Royal Statistical Society, 37(2), 149-192

Boserup, E. (1970), "Women's Role in Economic Development". St. Martin's Press, New York.

Cagatay, N. and Ozler, S. (1995), "Feminization of the labor force: The effects of long-term development and structural adjustment". World Development, 23(11), 1883-1894.

Charles, M. (1992), "Cross-National Variation in Occupational Sex Segregation". American Sociological Review , 57(4), 483-502.

Clark, R. (1991), "Contrasting Perspectives on Women's Access to Prestigious Occupations: A Cross-National Investigation". Social Science Quarterly, 72, 20-32.

Clark, R., Ramsbey, T. W. and Adler, E. S. (1991), "Culture, Gender, and Labor Force Participation: A Cross-National Study". Gender and Society, 5(1), 47-66.

Durand, J.D. (1975), "The Labor Force in Economic Development". Princeton: Princeton University Press.

Eastin, J. and Prakash, A. (2013), "Economic Development and Gender Equality: Is there a Gender Kuznets Curve?". World Politics, 65(1), 156-186.

Emran, M.H., Shilpi, F., Alam, M.I. (2007), " Economic liberalization and price response of aggregate private investment, time series evidence from India". Canadian Journal of Economics 40(3), 914–934.

Engle, R.F., Granger, C.W.J. (1987), "Co-Integration and error correction: representation, estimation, and testing". Econometrica, 55(2), 251–276.

Forsythe, N., Korzeniewicz, R. P. and Durrant, V. (2000), "Gender Inequalities and Economic Growth: A Longitudinal Evaluation". Economic Development and Cultural Change, 48(3), 573-617.

Goldin, C. (1994), “The U-shaped female labor force function in economic development and economic history”. NBER Working Paper Series, No. 4707.

Heston, Alan, Summers, Robert and Aten, Bettina, Penn World Table Version 7.1, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, Nov 2012.

ILO Database on Conditions of Work and Employment Laws. ILO, Geneva. Available at: http://www.ilo.org/dyn/travail/ (accessed December 2012).

Johansen, S. (1988), “Statistical analysis of cointegration vectors”. Journal of Economic Dynamics and Control ,12 (2-3), 231–254.

Johansen, S., Juselius, K. (1990), “Maximum likelihood estimation and inferences on cointegration with applications to the demand for money”. Oxford Bulletin of Economics and Statistics, 52(2), 169–210.

Kottis, A. P. (1990), “Shifts Over Time and Regional Variation in Women’s Labor Force Participation Rates in a Developing Economy: The Case of Greece”. Journal of Development Economics, 33(1), 117-132.

Kuznets, Simon (1955), “Economic Growth and Income Inequality”. The American Economic Review, 45(1), 1-28.

Lantican, C. P., Gladwin, C. H., Seale, J. L., Jr. (1996), “Income and Gender Inequalities in Asia: Testing Alternative Theories of Development”. Economic Development and Cultural Change, 44(2), 235-263.

Menyah, K. and Wolde-Rufael,Y. (2010), “Energy consumption, pollutant emissions and economic growth in South Africa”. Energy Economics, 32(6), 1374-1382.

Narayan, P.K. (2005) “The saving and investment nexus for China: evidence from cointegration tests”. Applied Economics, 37(17), 1979-1990.

OECD (2012), OECD.Stat, (database). http://stats.oecd.org/ (accessed December 2012)

Pampel, F. C. and Tanaka, K. (1986), “Economic Development and Female Labor Force Participation: A Reconsideration”. Social Forces, 64(3), 599–619.

Panopoulou, E. and Pittis, N. (2004), "A comparison of autoregressive distributed lag and dynamic OLS cointegration estimators in the case of a serially correlated cointegration error". The Econometrics Journal, 7(2), 585–617

Perron, P. (1989), "The Great Crash, the oil price shock and the unit root hypothesis". Econometrica, 57(6), 1361–1401.

Pesaran, M.H., Shin, Y. (1999), "An autoregressive distributed lag modelling approach to cointegration analysis". In: Strom, S. (Ed.), Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium. Cambridge University Press, Cambridge.

Pesaran, M.H., Shin, Y., Smith, R.J. (2001), "Bounds Testing Approaches to the Analysis of Level Relationships". Journal of Applied Econometrics, 16(3), 289-326.

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

Steel, W. F. (1981), "Female and Small-Scale Employment under Modernization in Ghana". Economic Development and Cultural Change, 30(1), 153-167.

Tam, H. (2011), "U-shaped female labor participation with economic development: Some panel data evidence". Economics Letters, 110(2), 140-142.

The Conference Board Total Economy Database, Output, Labor and Labor Productivity Country Details 1950-2011, January (2012). http://www.conference-board.org/data/economydatabase/ (accessed November 2012)

The Global Gender Gap Report 2012, published by the World Economic Forum. http://www.weforum.org/issues/global-gender-gap (accessed April 2013)

Weiss, J. A., Ramirez, F. O. and Tracy, T. (1976), "Female Participation in the Occupational System: A Comparative Institutional Analysis". Social Problems, 23(5), 593-608.

Zivot, E. and Andrews, D. (1992), ‘‘Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis’’. Journal of Business and Economic Statistics, American Statistical Association, 10(3), 251–270.