LUND UNIVERSITY

The Future of Female is U

A cross-country study of the U-shaped relationship between economic development and female labor force participation during the time period 1991-2016

Lund University School of Management and Economics

Department of Economics

Bachelor's thesis, NEKH02

Author: Irem Ucan

Supervisor: Gunes Gökmen

Date: 25/01/2021

Abstract

The purpose of this paper is to examine the relationship between economic development and female labor force participation within the theoretical framework of Feminization U for 47 countries over the time period 1991 and 2016, and the effects of economic and socioeconomic factors on female labor force participation. The examination is explored through the employment of fixed effects model, and the basis for the statistical analysis is pooled panel data on the countries for the given time period. The data is retrieved from ILOSTAT and World Bank Indicators. The results suggest that the prevalence of a U-shaped relationship depends on the indicator used for economic development. Furthermore, the agricultural sector, male unemployment and primary education seem to have a positive effect on female labor force participation, whereas female unemployment had negative effect of female labor force participation.

Keywords: Female labor force participation, economic development, Feminization U, cross- country analysis

Table of Contents

List of acronyms

FLPR – Female labor force participation rate

ILO – International Labour Organisation

ILOSTAT – International Labour Organisation Department of Statistics

LFS – Labor Force Survey

MENA – Middle East and North Africa

OECD – Organisation of Economic Co-operation and Development

SDGs – Sustainable Development Goals

1. Introduction

1.1 Background

Gender gaps in the labor market is one of the major challenges faced globally (ILO 2017). Over the past 25 years, female labor force participation rate (the proportion of females aged 15 and over to be economically active) has been gradually declining from 51% in 1994 to 47% in 2019 globally (World Bank 2020a) Furthermore, females have only accounted for 39% of the total global work force over the past two decades (World Bank 2020b). However, female labor force participation rate (FLPR) varies greatly between countries and regions, especially between developed and developing countries (Kumari 2018). For example, the Middle East, North Africa and South Asia are regions with the lowest FLPR, as low as 20% in MENA and 24% in South Asia for the year 2019, whereas this number is 50% for Europe and Central Asia during the same year (World Bank 2020c). To understand what accounts for this disparity in female labor force participation between countries (and regions) becomes important, especially since one of the SDGs is to promote full and productive employment for men and women, and another being gender equality.

One body of research suggests that changes in female labor force participation follow a U-shape as an economy becomes more industrialized (also known as Feminization U), whereby FLPR remains high at lower levels of economic development, to later drop as the economy experiences economic growth. After a certain point in economic development, the FLPR increases again and remains high at higher levels of economic development (Sinha 1967). To relate this to current trends in economic development on an international setting, FLPR is high in low- and high-income countries, and relatively low in middle- income countries (exceptions are plausible) (Fatima & Sultana 2009). The implications of this framework would not only allow us to understand current trends in FLPR on a global and country level, but also highlight the need for active labor market policies to enhance the participation of females in the labor market, especially in developing countries to prevent a further drop in FLPR (Luci 2009). To advance FLPR could further accelerate the process of economic development by increasing the aggregate labor input in the economy, which on microlevel could improve the household income, thus increase her consumption of goods and services, and also alleviate the existence of poverty on micro- and macrolevel (Verick 2018).

1.2 Aim of the study

The aim of the study is to test the feminization U hypothesis and identify the determinants of female labor force participation.

1.3 Research questions

- To what extent can the relationship between economic development and female labor force participation be explained within the Feminization U framework?
- What effects do economic and socioeconomic factors have on female labor force participation?

1.4 Scope of the study

The relationship between economic development and female labor force participation is limited to the theoretical framework of Feminization U, and other proposed theories on the relationship will not be regarded. The effect of certain economic and socioeconomic factors will be explored, and these factors are limited to labor market conditions such the sectoral composition of female employment, and unemployment, and demographic factors such as education and fertility. The time period is restricted to 1991-2016 and 47 countries are included in the study. These countries vary in level of economic development and geographical location. The reason for including both developed and developing countries is to examine whether the U-shaped relationship can be traced for both countries that have experienced higher levels of development during the studied time period and countries that are in the process of developing.

1.5 Disposition

In the following section, a closer examination of the Feminization U hypothesis, and the identified determinants of female labor force participation will be presented. After the literature review, the process of data collection and the analysis method will be explained. In section 6, the results will be reported and analyzed. Lastly, in the final part of this paper, a general discussion along with concluding remarks will be presented.

2. Literature review

2.1 Feminization U hypothesis

The feminization U hypothesis or the U-shaped female labor force curve has been constructed to conceptualize the long-term relationship between economic development and female labor force participation (Sinha 1967; Boserup 1970; Durand; 1975; Psacharopoulus & Tzannatos 1989; Goldin 1995). According to Goldin (1995), the dynamic of the female labor market changes as the economy progresses following a U-shape. During the initial stages of economic development, female labor force participation remains at a higher rate due to low income and the prevailing agricultural practices in the country, for personal and/or commercial consumption (ibid.). However, with time as the economy experiences growth in terms of increased income and moves toward industrialization due to technological developments or the expansion of the market, the female labor force participation tends to fall. The reasons for this decline in female labor force participation at this stage, suggested by Goldin (1995), Cagatay and Özler (1995) and Tansel (2001), are the decrease in demand for female labor force in the agricultural sector (due to difference in labor productivity between men and women and the mechanization of the agricultural sector) and the gender segregated demand in the non-agricultural sectors (e.g. the preference for male workers in the manufacturing sector). Furthermore, Cagatay and Özler (1995) emphasize the difficulty for women to balance reproductive activities with industrial labor in an urban setting. The upward slope in U-shape, that is, the increase in the female labor force participation occurs as the economy develops further to become more service-based and the demand for female labor force increases (Lechman & Kaur 2015). The improvement in female education, decline in fertility, commodification of domestic labor and the increased female political advocacy have been identified as some of the drivers of this change in the female labor force participation (Lechman & Kaur 2015; Goldin 1995; Cagatay & Özler 1995).

The advocates of this hypothesis have conducted both cross-country analysis and country-specific case studies over various time periods using both panel and time series data (see Tansel 2011; Fatima & Sultana 2009; Goldin 1995; Kottis 1990; Psacharopoulos & Tzannatos 1989; Boserup 1970). However, more recent studies have re-examined this U-shaped relationship between economic development and female labor force participation (Lechman & Kaur 2015; Tsani et al. 2013; Luci 2009). Although Lechman and Kaur (2015) confirm the Feminization U hypothesis with the analysis of 162 countries over the period 1990-2012, the extent of the relationship is questioned. When the

countries were divided into four income groups; low, lower-middle, higher-middle and high, the results were varied. The U-shape was prevailing in high and higher-middle-income countries, whereas the U-shaped relationship was weak in the lower-middle countries. Furthermore, an inverted U-shape relationship was noted in the low-income group. Thus, the authors concluded that even though the U-shaped relationship may have captured the development of the developed countries in the past, it may not help to explain the developments in current developing countries. Instead, Lechman and Kaur (2015) encourages the consideration of the social conditions of the developing countries to explain the relationship between economic development and female labor force participation (ibid.). Another empirical study that confirms the occurrence of the inverted U-shaped relationship between economic development and female labor force participation focuses on India over the time period 1983-2009 (Lahoti & Swaminathan 2013).

Gaddis & Klasen (2014) argues that the relationship between economic development and female labor participation is more complex than what is implied in the literature by highlighting the importance of the data used (and the methods to acquire such data) in the analysis of the U-shaped relationship. By comparing the results from two different versions of ILO's EAPEP database (please note that this database is no longer accessible), the authors conclude that the U-shaped relationship between economic development and female labor force participation is "feeble and not robust" and a cross-country analysis could be affected by the data used (ibid.).

Not only does the data used in the studies affect the results and the conclusions that can be drawn regarding the U-shaped relationship between economic development and female labor force participation, but also the variables used in the models to determine the relationship (Schultz 1990; Pampel & Tanaka 1986). GDP per capita has been used as the solemn measure of economic development in the studies conducted by Goldin (1995) and Psacharopoulos and Tzannatos (1989), however as noted by Pampel and Tanaka (1986) this measure could be an unreliable indicator for economic development due to the existence of economies with high GDP per capita (e.g., oil-producing countries) and low economic development. The authors propose the use of energy consumption per capita as a more appropriate measure of economic development in relation to female labor participation. According to Pampel and Tanaka (1986), this measurement captures both the increased use of energy both in the household and the large-scale production associated with industrialization.

2.2 Determinants of female labor force participation

This section aims to present the economic and socioeconomic factors that could influence the female labor force participation in an economy, which could help to explain the relationship between economic development and female labor force participation within the framework of Feminization U.

2.2.1 Labor market conditions

As presented in the Feminization U hypothesis, the labor market experiences a structural change as a country becomes more industrialized and this section details the labor market conditions in relation to female labor force participation.

2.2.1.1 Sector of economic activity

As explained in the previous section, an economy's sectoral composition changes as it becomes more industrialized during economic development. The three primary sectors mentioned in the literature are agriculture, industry/manufacturing and service with regard to the U-shaped relationship between economic development and female labor force participation (Shultz 1990; Goldin 1995; Cagatay & Özler 2011). According to Fatima and Sultana (2009), the size of the agricultural sector has a positive effect on FLPR since it has been, and still is, the sector which has the highest FLPR. The effects of the manufacturing and service sectors on FLPR depend on which stage in economic development the economy is in and could be either positive or negative (ibid.). Based on a cross-country analysis of 117 countries, Besamusca et al. (2015) found a positive effect of the size of the agricultural and the service sector on FLPR, however they found no evidence that the increase in the size of the manufacturing sector has a negative effect on female labor force participation, which contradicts the notion of the downward sloping part of the U-shape in Feminization U hypothesis. Gaddis and Klasen (2014) came to same conclusion as Besamusca et al. (2015) regarding the effect of the manufacturing sector on female labor force participation.

2.2.1.2 Unemployment

The unemployment rate in an economy not only influences a women's ability to enter the labor force, but also her decision to search for a job in the labor market due to the associated economical and psychological cost of looking for a job (Kottis 1990). The higher the unemployment rate in an economy, the less likely is the woman to find and search for a job (ibid.). Kottis (1990) uses the term

" discouraged- worker effect" to describe the negative effect of unemployment rate on female labor force participation. However, this discouraged-worker effect could be countered if the male unemployment rate increases in the economy, to which women would be more inclined to look for a job to compensate for the family income loss (if male member(s) is/are present in the household) (Fatima & Sultana 2009).

2.2.2 Demographic factors

2.2.2.1 Education

Abramo and Valenzuela (2005) have denoted a strong positive impact of educational attainment on FLPR in 18 countries in Latin America. Higher level of education increases the probability of the woman to acquire the different skills demanded in the labor force and thus expands her job prospects (ibid.) According to the authors educational attainment increases the household income through skilled employment, and thus the educated women could outsource some of the domestic labor to uneducated or lower educated women, which further increases FLPR (ibid.). Furthermore, as noted by Besamusca et al. (2015) the opportunity cost to leave the labor force increases, for example for reproductive activities, the higher the level of educational attainment she acquires. Tumsarp and Pholphirul (2020) have deduced similar findings in Thailand, where each additional year of schooling increased FLPR in the country.

Psacharopoulos & Tzannatos (1989) also confirms the positive effect of education on FLPR however regards the duration of her employment as well. Through her education, she might be able to find skilled employment, and earn a higher wage compared to less educated or uneducated women. This could lead to the decision to consume more leisure and work less hours or exit the labor market at some point when she reaches her income target (ibid.). Nevertheless, the authors note that educated women are more likely to remain in the labor market than less educated or uneducated women for she has undertaken an investment by attaining a higher education and thus the opportunity cost to exist increases (ibid.).

Contrary to other researchers (Tumsarp & Pholphirul 2020; Besamusca et al. 2105; Abramo & Valenzuela 2005; Psacharopoulos & Tzannatos 1989), Kottis (1990) found that education had a negative effect on FLPR in Greece. Based on a cross-section analysis conducted for the years 1971

and 1981, the author denotes the importance of the relationship between the supply and demand for educated and uneducated women in the labor force (ibid.). The shortage of employment opportunities for educated women, especially in non-urban areas of the country, and the abundant supply of, and thereof demand for, labor of uneducated women, could result in negative effect of education on FLPR. However, despite the negative effect found in the case of Greece the author concludes that the effect of education on FLPR depends on the level of the development of the economy, where this effect is expected to be negative during the early stages of development and positive after a certain point in development (ibid.).

2.2.2.2 Fertility

Two causations have been highlighted in the literature regarding the relationship between fertility and FLPR (Mishra & Smyth 2010; Engelhardt & Prskawetz 2004). The first being that fertility effects FLPR and the second that FLPR effects fertility (ibid.). In the case when causation runs from fertility to FLPR, two outcomes have been identified. Fertility could have a positive effect on FLPR since additional income would be needed to provide for the child or children, which could increase her willingness to find a job (ibid.). Furthermore, the reproduction of additional children could alleviate some of the mother's childcare and household duties, since the older child could help to take care of the younger children and possibly carry out certain household chores (Psacharopoulos & Tzannatos 1989).

On the other hand, fertility could also have a negative effect on FLPR. Mishra & Smyth (2010) have identified three reasons for the negative effect of fertility on FLPR; increased time spent on childcare, emotional attachment to the small child (or children) and the increased economic burden of childcare for each additional child if she would enter the labor market and won't experience an improvement in her salary to sustain the increased cost of childcare. Based on an OLS-regression analysis, the authors found evidence that fertility had a negative effect on FLPR in 28 OECD countries during the time period 1995-2005 (ibid.). (For other studies on the negative effect of fertility on FLPR see Spierings et al. 2010; Hartani et al. 2015)

In the latter causation (from FLPR to fertility), FLPR could have a negative effect on fertility due to the opportunity cost of having children while in employment (Mishra & Smyth 2010). The break in

her employment for reproductive activities could not only result in income loss but could also stagnate her career opportunities (ibid.).

2.2.2.3 Marital status

Lee et al. (2008) have found strong evidence that marriage has a negative effect on FLPR in Korea. According to the authors, married women are, on average 40-60%, less likely to participate in the labor market compared to unmarried women (ibid.). Factors driving this disparity in the FLPR between married and unmarried women lie both on the supply and demand side of the labor market (ibid.). The hinder of human capital formation among married women due to the social norm of unpaid domestic labor put upon her, as well as the husband's employment status and educational attainment (women whose husbands had a formal employment were less likely to enter the labor market), affects the supply of labor provided by married women. On the demand side, discriminatory practices by employers against married women, such as offering lower wages compared to other employees (males or unmarried women) and/or hindering her career development in the private sector, are identified to be significant factors in the stagnation in FLPR among married women (ibid.). The negative effect of marriage on FLPR has also been confirmed for Pakistani women by Fatima and Sultana (2009). (For other empirical studies on the negative impact of marriage on FLPR see Tong & Chui 2017; Chen at al. 2014; Francis 2011; Sasaki 2002)

On the other hand, there is also evidence that marriage could have a positive effect on FLPR, according to an empirical study conducted in Thailand by Tumsarp and Pholphirul (2020). Based on a cross-sectional regression analysis for the year 2016, the authors found that married Thai women were 15.9% more likely to participate in the labor market compared to unmarried women (ibid.). The highest FLPR among married women were those who were younger, less educated, not head of the household and had a smaller family size (ibid.). It is likely that these married women come from poorer households and need to contribute to the family income (ibid.).

3. Data

The unavailability of data restricted the selection of countries for this study to 47 and many countries had to be removed during the initial sorting. The aggregate female labor participation rates for the period 1991-2016 were collected from the ILOSTAT database (ILOSTAT 2020a). Other variables obtained from this database were female employment in agriculture, industry and services, unemployment rate (female, male) (ibid.). All the figures obtained from the ILOSTAT database are ILO modelled estimates, meaning that ILO has combined nationally- observed data with imputed data for countries with missing data based on econometric models constructed by the organization. The purpose of these estimations is to obtain balanced panel data for global and regional comparability and trends on various labor market indicators (ILOSTAT 2020b).

The remaining explanatory variables were obtained from the World Bank database, namely GDP per capita PPP (constant 2011 International $), electric power consumption, female school enrollment (primary, secondary), and fertility rate (World Bank 2020d). Missing data was also an issue in this database and is evident even for certain selected countries. How this will be treated will be presented in the model section of this paper along with the definitions of the variables obtained from the databases (see pp. 16-19).

The available time period for the data from the ILOSTAT database was 1990-2020, whereas it was limited to 1990-2016 in the World Bank database (with the exception of electric power consumption which was only available up until 2014 for all countries). However, the year 1990 was exempt from the study due to the extensive set of missing data for most countries in this study. For this reason, the time period was restricted to 1991-2016 due to the inaccessibility of the data for previous years on both of the databases.

4. Methodology

4.1 Selection of countries

47 countries are included in the analysis of the relationship between economic development and female labor force participation, and the factors affecting female labor force participation, in this study (see p. 31 for the list of countries). Two important factors that affected the selection of countries for this study are a) data availability b) the geographical location of the country. Countries with no or few datasets in the explanatory variables were excluded from the analysis in order to acquire as balanced panel dataset as possible. Furthermore, to avoid a cluster of countries to a specific geographic location, an attempt was made to include countries from various geographical locations such as Europe, the Americas (North, South, Central), Africa (South, North, East), Asia (South, Southeast, West) and the Middle East with the aim to increase the probability of capturing a variation in economic development between the countries. However, no prior analysis of individual countries in the study (or the countries eliminated from the study) was made to determine the level of economic development of the country or its relationship to FLPR.

4.2 Selection of econometric method

Various methods have been used to determine the relationship between economic development and female labor force participation, and the factors affecting female labor force participation, in the literature. Time series, cross-sectional estimations, and panel data estimations are among the methods used. Studies using time series data observe for a single entity (e.g. a person, firm or a country) over multiple time periods (e.g. monthly, quarterly or annually), whereas if the basis for the analysis is cross-sectional data, then multiple entities are observed for a single time period (Stock & Watson 2015). On the other hand, panel, or longitudinal, data combines multiple entities over multiple time periods (ibid.).

Various factors need to be considered for the selection of method, as the methods have both advantages and disadvantages for an econometric analysis. A time series analysis could give information about the evolution of a variable and could be used to make forecasts (Stock & Watson 2015). However, the problem of (perfect) multicollinearity can be prevalent in time-series regressions i.e. when explanatory variables are in perfect linear correlation with each other and thus not independent (ibid.). This would also mean that the errors are correlated across observations which would question the validity of the method (ibid.). A cross-section estimation, on the other hand, can

be used to determine the relationship between different variables by studying the differences (and similarities) across different entities at a particular time period (ibid.). Heteroskedasticity could be an issue when cross-section data is used for the basis of the analysis (it may also be an issue for time-series regressions, but it is less common), which occurs when the variance of the error term is not constant (ibid.). Lastly, panel data can be used to explore the relationship between different variables by studying the differences across different entities and the evolution of the variables over multiple time periods. Multicollinearity could also be an issue for panel data estimations (ibid.).

In the literature, Psacharopoulus and Tzannatos (1989) uses a simple pooled cross-sectional estimation for 136 countries in the year of 1987, and thus confirm the U feminization hypothesis with GDP per capita and female labor force participation rate as the solemn variables. Goldin (1995) combines a cross-sectional regression analysis for 100 countries in the year 1980 with a case study on the United States from a historical perspective (time-series) as support for the feminization U hypothesis. Cagatay and Özler (1995) has a similar approach to Psacharopoulus and Tzannatos (1989), however use two time points (1985 and 1990) instead of one to also confirm the hypothesis for 165 countries.

The more contemporary papers in the literature have examined the relationship using panel data estimations. Lechman and Kaur (2015) uses various different panel data estimations and confirm the Feminization U hypothesis for 162 countries over the time period 1990-2012. Another empirical study conducted by Luci (2009), has a similar approach to Lechman and Kaur (2015) based on panel data for 184 countries from the years 1965 to 2004, and also confirms the hypothesis. (For other studies that use cross-country panel data regressions see Besamusca et al. (2015), Tsani et al. (2013), Spierings et al. (2010)).

After having evaluated the methods, fixed effects model will be used in this study, which is a type of panel data regression model. The advantage of fixed effects regression model is the ability to control the omitted variables that vary between entities (Stock & Watson 2015). In this model, both cross-section (or in this case country) and time fixed effects can be added. Both country and time fixed effects will be employed in this study in order to control for country-specific variation and time-specific variation (e.g. economic shocks).

The fixed effect regression models will be run on the statistical software program EViews (11th version). Furthermore, to test for multicollinearity between the regressors the method "Variance Inflation Factor" (VIF) will be used, where a value of 10 or less is considered acceptable.

4.3 Limitation to the methodology

One limitation of the methodology regards the panel dataset, more specifically the lack of balanced panel data. Once the models are run in EViews, the program will produce an estimation based on the imputed values in the dataset. The variables of concern are mainly the ones that regard female education. However, no visible trend was detected in the missing dataset and the gaps appear sporadic (and thus it is assumed that the data is not missing systematically). Furthermore, only 8 of 47 countries had a complete set of data on female primary and secondary education (both variables), whereas the remaining countries missed at least one data point in one or both variables. Another unbalanced dataset is on the variable *electric power consumption*, whereby data is missing for all countries for the time period 2014-2016.

There is high probability of multicollinearity in this study since many of the variables such as electric power consumption, fertility and education could be correlated with GDP per capita. However, GDP per capita can't be removed from fixed effects regression models since this variable is needed to answer one of the research questions.

5. Model

5.1 Descriptive statistics

5.1.1 Dependent variable

The female labor force participation rate (*FLPR*) is the only dependent variable in this study. Female labor force participation rate is the proportion of females of the age 15 and over to be economically active of the total female population of the age 15 and older. Economically active are those considered to supply labor for the production of goods and services during a specific time period (ILOSTAT 2020c).

5.1.2 Explanatory variables

- *GDP per capita, PPP (constant 2011 International $)*: Gross Domestic Product per capita adjusted to Purchasing Power Parity and converted to international dollars
- *Electric Power Consumption (kWh per capita)*: a country's production and consumption of electricity divided by its midyear population
- *Female employment in agriculture (%)*: The proportion of females employed in the agricultural sector (of the total economically active females). The agricultural sector includes activities in agriculture, hunting, forestry and fishing
- *Female employment in industry (%):* The proportion of females employed in the industry sector (of the total economically active females). The industry sector includes activities in mining and quarrying, manufacturing, construction, and public utilities (electricity, gas, and water)
- *Female employment in services (%):* The proportion of females employed in the service sector (of the total economically active females). The service sector includes activities in wholesale, retail trade, restaurants, hotels, transport, storage, communications, financing, insurance, real estate, business services, community, social, and personal services
- *Female unemployment (%)*: The proportion of females unemployed of the total female labor force (aged 15 and over). Unemployed are those considered who are without work but are available for work and are seeking for employment
- *Male unemployment (%):* The proportion of males unemployed of the total male labor force (aged 15 and over). Unemployed are those considered who are without work but are available for work and are seeking for employment
- *Female primary school enrollment (% gross)*: The proportion of females enrolled in primary education of the total female population of the official primary education age
- *Female secondary school enrollment (% gross):* The proportion of females enrolled in secondary education of the total female population of the official secondary education age
- *Total fertility rate (births per woman):* the number of children that would be born to a woman if she were to live to the end of her childbearing years and bear children in accordance with age-specific fertility rates of the specified year

5.1.2.1 Examination of the explanatory variables

Six different models are estimated in this study, based on the variables above. A more detailed explanation of the models will be presented in the following pages of this paper however this information is relevant for the explanation of the variables used in this study.

Two variables are used as proxy for economic development; GDP per capita and electric power consumption. Electric power consumption was added to the model based on the argument put forth by Pampel and Tanaka (1986) about the problem of using GDP per capita as an indicator for economic development. Models 1 to 3 uses electric power consumption and GDP per capita as indicators for economic development, whereas GDP per capita was the solemn variable for economic development in models 4 to 6. The reason for this design is to test if there is a U-shaped relationship between economic development and FLPR even if electric power consumption is used as one of the proxies for economic development. The values of GDP per capita and electric power consumption will be transformed to logs since this captures the growth rates in the indicators and allows for a more normal distribution which makes the calculations for the regression easier and simpler (income data is often skewed). Furthermore, a quadratic term is introduced for the economic development indicators in order to determine whether there is a U-shape in FLPR.

To consider the effect of the labor market conditions, the industry mix and the unemployment rates are introduced. Model 1 and 4 examine the effect of the agricultural sector on FLPR, whereby the female employment in agricultural sector is used as the variable. The variable is expected to be positive since it is the sector with the highest FLPR according to the literature. The effect of the industry sector is examined in model 2 and 5. The effect of this sector is ambiguous according to the literature. And lastly, the effect of the services sector is expected to be positive since the demand for female labor force increases as the country becomes more service-based (Goldin 1995; Cagatay & Özler 1995). The choice of variables for sectoral composition was based on the variables used in the empirical studies by Fatima and Sultana (2009), Spierings et al. (2010), Tumsarp and Pholphirul (2020). On the other hand, Besamusca et al. (2015) uses the value added of the sectors (% of GDP) as proxy for the size of the sectors. The variables used in this study with regard to the employment structure of the different sectors could indicate the level of development as mentioned in the literature. Furthermore, these variables were balanced whereas the datasets on value added of the sectors were unbalanced.

The examine the effect of unemployment on FLPR, both female and male unemployment are added to all of the models. The reason for including both unemployment rates is to test both the discouraged-worker effect and if male unemployment has a positive effect on FLPR.

Various variables have been used as proxy for female education in the literature; school enrollment rates, mean years of schooling, literacy rate and illiteracy rate. For this study, only primary and secondary school enrollment will be included since these variables are most commonly used in the literature.

On the note of fertility, it is not clear whether fertility impacts FLPR or if FLPR impacts fertility, or a combination of both. It is possible that women will be less inclined to produce children due to the opportunity cost of exiting the labor market, in which FLPR impact fertility negatively. Nevertheless, the impact of fertility on FLPR will be examined, as this is one of the most common variables used in the literature. The impact of fertility on FLPR could be positive or negative, however it is expected to be negative.

5.1.3 Excluded variables

The data on female tertiary school enrollment and literacy rate was unavailable for most countries in the study for the time period 1991-2016, and thus these variables were excluded. The data for female tertiary school enrollment was to be included to determine if the increase in level of education would have a positive impact on FLPR.

The impact of marital status on FLPR will not be added to the regression models in this study due to the complications in data accessibility. First of all, the observations on FLPR by marital status are based on national LFS conducted at certain years (and often at different years), which means that data on this variable is only available on that particular year in which the surveys were conducted. Second of all, the dataset on FLPR by marital status is available for more recent years for most countries in the study i.e. only 5 countries have reported the marital status of employed females during the time period 1991-1999 (no common year during the time period was able to be found for these 5 countries), whereas most countries have reported data from 2007 and onward (also at different years).

5.2 The estimation model

The general model for this is study is the following:

$$FLPR_{it} = \alpha_i + \beta_1 DEV_{it} + \beta_2 DEV_{it}^2 + \beta_3 SEC_{it} + \beta_4 UF_{it} + \beta_5 UM_{it} + \beta_6 EDp_{it} + \beta_7 EDs_{it} + \beta_8 F_{it} + \delta_t + u_{it}$$

Where i denotes for a country and t for time. α_i is the country-fixed effect and δ_t is the time-fixed effect.

DEV: Economic development variable

SEC: Sector

UF: Female unemployment

UM: Male unemployment

EDp: Female primary school enrollment

EDs: Female secondary school enrollment

F: Fertility rate (total)

6. Empirical findings

Before the empirical results are reported and interpreted, the reader should be aware that the existence of multicollinearity has been detected which means that the precision of the coefficients could be impacted, and thus the empirical results should be viewed with caution. The VIF test results can be found in the Appendix. The collinearity between the sectors was somewhat treated by only including one sector per model (and this can be seen by the VIF results for the sector in each model).

A note on the explanatory variables: *EPC* denotes electric power consumption, *GDPPC* is GDP per capita, *Empl_ - agr, ind, serv* refers to the shares of females employed in the three sectors, *Unempl_ - fem, male* are the unemployment rates for females and males, whereas *Ed_ - prim, sec* are the female enrollment in primary and secondary education.

Table 1. Fixed effects regression results

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
log(EPC)	-5.230 (-1.444)	-2.842 (-0.749)	-7.543 (-2.054)**			
$\log(EPC)^2$	0.720 (-1.444)***	0.489 (1.718)*	0.868 (3.139)***			
log(GDPPC)	-37.028 (-4.289)***	-43.177 (-4.834)***	-40.420 (-4.740)***	-34.353 (-6.723)***	-38.786 (-7.166)***	-40.597 (-8.117)***
$\log(GDPPC)^2$	1.609 (3.260)***	1.813 (3.532)***	1.881 (3.848)***	1.670 (5.507)***	1.755 (5.428)***	2.087 (6.920)***
Empl_agr	0.221 (7.579)***			0.195 (7.036)***		
Empl_ind		-0.069 (-1.320)			-0.048 (-0.959)	
Empl_serv			-0.294 (-8.314)***			-0.268 (-7.978)***
Unempl_fem	-0.099 (-1.590)	-0.112 (-1.715)*	-0.139 (-2.243)**	-0.135 (-2.206)**	-0.142 (-2.252)**	-0.170 (-2.810)***
Unempl_male	0.189 (2.760)***	0.137 (1.849)*	0.317 (4.504)***	0.240 (3.587)***	0.188 (0.008)***	-0.170 (-2.810)***
Ed_prim	0.033 (1.557)	0.071 (3.316)***	0.048 (2.342)**	0.005 (2.359)**	0.083 (4.188)***	0.053 (2.761)***
Ed_sec	-0.0009 (-0.076)	-0.015 (-1.219)**	-0.004 (-0.334)	0.009 (0.782)	-0.007 (-0.606)	0.008 (0.498)
Fertility	0.417 (0.827)	-0.254 (-0.491)	0.217 (0.437)	-0.561 (-1.213)	-0.954 (-2.021)**	-0.726 (0.112)
F-statistic	265.404	247.325	275.315	270.977	256.079	275.315
N_obs	871	871	871	950	950	950

***p<0.01, ** p<0.05, * p <0.1. The t-statistic values in parathesis.

Based on the fixed effect regression results, there is some evidence for the U-shaped relationship between economic development and female labor force participation. The negative coefficients of *log(EPC)* and *log(GDPPC)*, and the positive coefficients of $log(EPC)^2$ and $log(GDPPC)^2$ suggests a U-shaped relationship between economic development and FLPR. The downward sloping part of the U shape could be explained by the negative coefficient of the *log(EPC)* and *log(GDPPC)*, whereby FLPR decreases as ECP and GDP increases until the vertex point of the parabola, and after this point the relationship becomes positive (as noted by the positive coefficients of $log(EPC)^2$ and $log(GDPPC)^2$).

However, these results should be viewed with caution since there is a variation in the significance of the estimations. *Log(GDPPC*) and $log(GDPPC)^2$ are significant at 1% level in all models, and could give strong evidence for the U-shaped relation between economic development and FLPR. On the other hand, when another variable for economic development was introduced (energy power consumption), the Feminization U hypothesis could not be confirmed. Although the negative coefficient of *log(EPC)* and the positive coefficient of $log(EPC)^2$ suggests a U-shape, only in model 3 can we see a significance at 5% level for *log(EPC),* and $log(EPC)^2$ was significant at 1% level in model 1 and 3 but at 10% in model 2. It is possible that the lack of data for the years 2014-2016 on EPC could have affected the strength of the results, especially if an increase in this variable was experienced during those years for the selected countries.

In model 1 and 4, employment in the agricultural sector had a strong positive effect on FLPR (1% significance level), which gives evidence that countries with large agricultural sectors also have high female labor force participation rates. Despite the negative coefficient of the industry sector (-0.069 and -0.048), which would give evidence for the negative impact of this sector on FLPR, the results are nonsignificant. Contrary to the literature, employment in the services sector had a negative effect on FLPR (at 1% significance level in both model 3 and 6). One plausible explanation for these results regarding the services sector could be that many countries in this study could still remain in early stages of economic development and have not moved to become more service-based. Another reason could be that the variables used for the sectoral composition are not fit to capture the structural change.

There is also evidence for the discouraged-worker effect noted by the negative coefficient of *Unempl_fem* in the models. The negative impact of female unemployment on FLPR could be explained by the economical and psychological costs associated with seeking and finding a job. However, these

results should be viewed with caution as well, seeing that there is a variation in the significance levels of this variable. In models 4 to 6, where GDP per capita was used as the solemn proxy for economic development, the results for female unemployment were significant at 1% and 5% levels. On the other hand, in models 1 to 3, where both GDP per capita and electric power consumption were used to represent economic development, female unemployment still had a negative effect on FLPR however only the result in model 3 was significant at 5%, whereas model 2 was significant at 10% and the result from model 1 was nonsignificant.

It appears that male unemployment has strong positive effect on FLPR in all models, except for in model 2 where it was significant at 10% level. The positive effect could be explained by the inclination of females to find a job in order to compensate for the income loss experienced due to the economic inactivity of the male counterpart. However, the family structure needs to be considered when these results are interpreted since male presence may not always be evident in relation to the female. It is also possible that male unemployment has a positive effect on FLPR since the jobs could be occupied by females which could results in male unemployment. This could to some extent depend on the type of job opportunities available in the economy and the level of gender discrimination in the labor market.

Female primary school enrollment seems to have a positive effect on FLPR. The results were statistically significant at 1% and 5% except for in model 1 which was nonsignificant yet positive. On the other hand, the effect of female secondary school enrollment on FLPR appears to be ambiguous and the results were nonsignificant expect for in model 2 in which secondary educational enrollment had a negative impact on FLPR (at 5% significance level). As noted by Kottis (1990), the effect of education on FLPR might depend on the level of economic development, whereby the negative coefficient could indicate that the many of the countries in this empirical study are at a lower stage of economic development. Another possibility is that the variable *female secondary school enrollment* is sticky, as indicated by the low t-values, by which the fixed effects model would not be able to capture its relationship to FLPR that is statistically significant. Overall, it can't be confirmed that increasing the level of education could have a positive impact on FLPR or in other words that each additional level of schooling increases FLPR. Moreover, the problem of multicollinearity could influence the results with regard to the effect of education of FLPR since it is likely that female education is correlated with GDP per capita. The existence of multicollinearity could also explain the variation in signs of

the coefficients of female secondary school enrollment (and the variation in the significance levels in primary school enrollment in the models).

Another variable that produced ambiguous results was fertility. In model 5, fertility had a statistically significant (at 5% level) negative impact on FLPR however the remaining results were nonsignificant. The presence of a child or children could discourage her to enter or remain in the labor force. However, the result from model 5 could be exaggerated due to using FLPR as the dependent variable, since it observes economically active females from the age 15 and older (ILO has not limited to a certain age but uses national working-age population estimations which varies from 64 to 70 as the ceiling) whereas the fertility rate remains low or closer to zero for females over a certain age (female reproductive lifespan). Fertility could also be another variable that is sticky in this study i.e. there is not much variation in the fertility rate during this time period. Furthermore, the variation in the signs of the coefficient could also be explained by the existence of multicollinearity.

7. Discussion and conclusion

In this empirical study, the relationship between economic development and female labor force participation was examined using the Feminization U hypothesis and the effect of various economic and socioeconomic factors on female labor force participation was explored for 47 countries during the time period 1991-2016. Before any conclusions are drawn, the existence of multicollinearity in the panel data regression models should be addressed. The endogeneity of the explanatory variables was unable to be controlled for using the fixed effects model, which could affect the validity of the analysis presented in this study. The variation in the signs of the coefficients of certain variables (female secondary school enrollment and fertility) is an example where the issue of multicollinearity is evident.

Interestingly, fixed effect models have been used by various authors in the literature, all who have conducted similar research to this one. Some studies have only examined the existence of the U-shape, with FLPR and GDP per capita as the solemn variables, whereas other have also accounted for the socioeconomic factors such education, fertility, marital status etc. The issue of endogeneity between the various regressors (both between the socioeconomic factors and their relation to aggregate income level) have not been addressed in all empirical studies.

Two different indicators were used as proxy for economic development: GDP per capita and electric power consumption. The U-shaped relationship seems to be apparent when GDP per capita was the solemn indicator for economic development, which has been the method employed by the advocates for this hypothesis. However, when electric power consumption was added to the model, the U-shape was no longer as obvious compared to the GDP per capita case.

The strength and validity of the U-shaped relationship between economic development and female labor force participation may be influenced by the selection and number of countries studied. Although the Feminization U hypothesis could be confirmed using GDP per capita as the only indicator for economic development in this study, a generalization can't be made since it is unclear if this relationship would hold or other results would have been obtained if other countries were to be selected or if more countries were added to the model (the problem of multicollinearity is another reason why a generalization can't be made). Furthermore, the effect of economic and socioeconomic factors on FLPR could vary greatly between countries. Although both developed and developing countries were included in this study, a sample of 47 countries may not be enough to draw a conclusion regarding the development of current developing countries.

Data availability and accessibility could also influence the Feminization U hypothesis and its validity. As mentioned in this paper, the lack of data in certain variables (mainly on female education) hindered the inclusion of many countries in this empirical study. Furthermore, many socioeconomic factors that could influence female labor force participation were not examined due to data unavailability. For example, the impact of female tertiary school enrollment was unable to be tested due to extensive gaps in the dataset. Data on female tertiary school enrollement was only available for a few countries, often developed, and if this data was to be used in this study, the validity of the result obtained from this variable would have been questioned. Moreover, to conduct LFSs (and various other surveys) on a national level is a costly process, which could lead to underestimation of the FLPR in the country and to stagnate the identification of the significance of various socioeconomic factors on FLPR. Furthermore, the basis of the analysis in this paper relied on ILO estimates, and it is possible that these estimates are, to some extent, unable to capture the actual labor conditions in the countries. The importance and call for data on developing countries becomes more evident as research on economic development is conducted.

Country-specific variation in data collection and processing should also be considered (especially on the labor force data) which could influence the attempt for a cross-country analysis of, not only, the relationship between economic development and female labor force participation (and its determinants) but also for economic development in general. For its estimations, ILO relies on the national LFSs conducted by the countries. However, the countries may have own interpretations of what is meant by labor, the age limits for economically active population (as mentioned in the results section of this paper), unemployment and, most importantly, the extent of inclusion of data on informal employment. The treatment of this data on a national level could influence the economic analysis conducted in this paper. For example, if data on informal employment is not included in the female labor force participation rate, it is likely that the analysis to reflect the situation of female employment in the countries, and its relationship to economic development would be limited.

The variables used in this study, and in the literature, could affect how the relationship between economic development and female labor force participation is viewed, and the socioeconomic factors mentioned in the literature to be of significance to the female labor force participation. The empirical results of this study suggests that the relationship between economic development and female labor force participation may be influenced by the indicator used for economic development, and the strength of the U-shape is amplified when GDP per capita is used as proxy for economic development. Furthermore, the empirical results obtained for the sectoral composition (female employment in agriculture, industry and services), may had not been sufficient to capture the structural change that occurs in an economy (under the assumption that this change occurs the way explained in the literature), and another variable were to be used to represent sector size (such as *value added of the sectors (% of GDP))* could have yielded different results (or if both variables had been used i.e. female employment and value added).

Another factor to consider is the time period studied. It is unclear whether the selected time period or the length of the time period is sufficient to capture the change in economic development of a country to determine its relationship to female labor force participation. It is possible that different results could have been obtained if a longer period was studied.

An assumption made in the Feminization U hypothesis is that countries experience a similar structural change that developed countries experienced from a historical perspective, that is as the country becomes more industrialized it moves from a predominant agricultural economy to a more service-based economy. According to the literature, FLPR drops initially as the country becomes more industrialized since female employment is unfavored in the manufacturing industry, however with the trends in globalization and trade liberalization, in many developing countries, females may be favored in certain industries such as the garment industry. Furthermore, no significant evidence was found in this study for this structural change suggested by the literature. It is possible that the results might have been influenced by the time period studied, in which the structural change has not occurred for some or many countries in the study.

The notion of economic development also needs to be considered. Economic growth and development have been used synonymously in the literature, however economic development also concerns with the improvement of the living standards of people. For example, the analysis in this study (and in the literature of Feminization U) has not regarded the working conditions of female employment, more specifically the extent of decent working conditions. A country may experience economic growth and the female labor force participation may remain high in the country, however the quality of the employment should be considered in the development aspect of the relationship.

In sum, the extent of the U-shaped relationship between economic development and female labor force participation seems depend on the indicator for economic development. On the note of economic and socioeconomic factors, male unemployment had a positive effect on FLPR, as did primary school enrollment and the agricultural sector, whereas female unemployment had a negative effect on FLPR. However, the socioeconomic factors may vary from country to country, and their effect henceforth, thus it would be more suitable to study country-specific conditions to explore the relationship between economic development and female labor force participation (and the determinants of female labor force participation) and adapt macroeconomic policies in such manner that accounts for these current conditions rather than historical economic developments.

References

Abramo, L. & Valenzuela, M. E. (2005). Women's labour force participation rates in Latin America. *International Labour Review*, vol. 144, No. 4, pp. 369-399

Besamusca, J., Tjidens, K., Kuene, M., Steinmets, S. (2015). Working Women Worldwide. Age effects in female labor force participation in 117 countries. *World Development*, Vol. 74, pp. 123-141

Boserup, E. (1970). Women's role in economic development. London: Allen & Unwin

Cagatay, Ö. & Özler, S. (1995). Feminization of the labor force: The effects of long-term development and structural adjustment. *World Development*, Vol. 23, No.11, pp. 1883-1894

Chen, J., Shao, X., Murtaza, G., Zhao, Z. (2014). Factors that influence female labor force supply in China. *Economic Modelling*, vol. 37, pp. 485-491

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

Engelhardt, H. & Prskawetz, A. (2004). On the changing correlation between fertility and female employment over space and time. *European Journal of Population*, Vol. 20, pp. 35-62

Fatima, A. & Sultana, H. (2009). Tracing out the U-shape relationship between female labor force participation rate and economic development for Pakistan. *International Journal of Social Science and Economics*, vol. 36, pp. 182-194

Francis, A. M. (2011). Sex ratios and the red dragon: using the Chinese Communist Revolution to explore the effect of the sex ratio on women and children in Taiwan. *Journal of Population Economics*, vol. 24, No. 3, pp. 813–837

Gaddis, I. & Klasen, S. (2014). Economic development, structural change, and women's labor force participation: a reexamination of the feminization U hypothesis. *Journal of Population Economics*, vol. 27, pp. 639-681

Goldin, C. (1995). The U-shaped female labor force function in economic development and economic history. *National Bureau of economic research working paper series*. No. 4707

Hartani, N. H., Bakar, N.A, Haseeb, M. (2015). The nexus between female labor force participation and female total fertility rate in selected ASEAN countries: panel cointegration approach. *Modern Applied Science*, vol.9, No.8, pp. 29-39

ILO (2017). World employment and social outlook: trends for women 2017. Report downloaded from ILO website URL: https://www.ilo.org/global/research/global-reports/weso/trends-for-women2017/lang--en/index.htm (accessed 19/12/2020)

ILOSTAT (2020a). ILOSTAT data. URL: https://ilostat.ilo.org/data/ (accessed: 27/12/2020)

ILOSTAT. (2020b). ILO modelled estimates and projections. URL: https://ilostat.ilo.org/resources/concepts-and-definitions/ilo-modelled-estimates/ (Accessed: 22/12/2020)

ILOSTAT. (2020c). Indicator description: Labor force participation. URL: https://ilostat.ilo.org/resources/concepts-and-definitions/description-labour-force-participation-rate/ (Accessed: 27/12/2020)

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*, vol. 33, pp. 117-132

Kumari, R. (2018). Economic growth, disparity, and determinants of female labor force participation: A research agenda. *World Journal of Entrepreneurship, Management and Sustainable Development*, Vol. 14, No. 2, pp. 138-152

Lahoti, R. & Swaminathan, H. (2016). Economic growth and female labour force participation in India. *Feminist Economics*, vol. 22, No. 2, pp. 168-195

Lechman, E. & Kaur, H. (2015). Economic growth and female labor force participation – verifying the U-feminization hypothesis. New evidence for 162 countries over the period 1990-2012. *Economics and Sociology*, vol. 8, No 1, pp. 246-257

Lee, B. S., Jang, S., Sarkar, J. (2008). Women's labor force participation and marriage: the case of Korea. *Journal of Asian Economics*, vol. 19, pp. 138-154

Luci, A. (2009). Female labour market participation and economic growth. *International Journal of Innovation and Sustainable Development*, vol. 4, No. 2, pp. 1-12

Mishra, V. & Smyth, R. (2010). Female labor force participation and total fertility rates in the OECD: New evidence from panel cointegration and Granger causality. *Journal of Economics and Business,* pp. 48-64.

Pampel, F.C. & Tanaka, K. (1986). Economic development and female labor force participation: a reconsideration. *Social Forces*, vol. 64, No. 3, pp. 599-619

Psacharopoulos, G. & Tzannatos, Z. (1989). Female labor force participation: an international perspective. *The World Bank Research Observer*, vol. 4, No. 2, pp. 187-201

Sasaki, M. (2002). The casual effect of family structure on labor force participation among Japanese married woman. *The Journal of Human Resources*, pp. 429-440

Sinha, J.N. (1967). Dynamics of Female Participation in Economic Activity, *Proceedings of the World Population Conference, Belgrade*, Vol. 4, pp. 336-337

Spierings, N., Smits, J., Verloo, M. (2010). Micro- and macrolevel determinants of women's employment in six Arab countries. *Journal of Marriage and Family review*, vol. 72, pp. 1391 - 1407

Stock, J. H. & Watson, M. W. (2015). Introduction to econometrics. 3rd ed. Pearson Education Inc.

Schultz, T.P. (1990). Women's changing participation in the labor force: a world perspective. *Economic Development and Cultural Change*, vol. 38, pp. 457-488

Tansel, A. (2001). Economic development and female labor force participation in Turkey: Time-series evidence and cross-province estimates. *ERF working paper series,* No. 1024

Tong, Y. & Chiu, S.W. (2017). Women's labor force participation in Hong Kong: 1991-2011. *Chinese Sociological Reviews*, vol. 49, No.1, pp. 35-64

Tsani, S., Paroussos, L., Fragiadakis, C., Charalambidis, I., Capros, P. (2013). Female labour force participation and economic growth in the South Mediterranean countries. *Economics Letters*, vol. 120, pp. 323-328

Tumsarp, P. & Pholphirul, P. (2020). Does marriage discourage female labor force participation? Empirical evidence from Thailand. *Journal of Marriage and Family review*, vol. 56, No. 7, pp. 677-688

Verick, S. (2018). Labor force participation and development. *IZA World of Labor*, Vol. 87, No. 2, pp. 1-11

World Bank (2020a). Labor force participation rate, female (ILO estimates) – World. Accessed: 20/12/2020. URL: https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS

World Bank (2020b). Labor force, female (of total labor force) – World. URL: https://data.worldbank.org/indicator/SL.TLF.TOTL.FE.ZS (Accessed: 20/12/2020)

World Bank (2020c). Labor force participation rate, female (ILO estimates) – Europe & Central Asia, Middle East and North Africa, South Asia. URL: https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS?locations=Z7-ZQ-8S (Accessed: 20/12/2020)

World Bank (2020d). World Development Indicators. URL: https://databank.worldbank.org/source/world-development-indicators (Accessed: 27/12/2020)

Appendix

A1. List of countries

Alphabetical order		
Albania	Germany	Poland
Austria	Greece	Romania
Argentina	Guatemala	Portugal
Armenia	Hungary	Russian Federation
Algeria	Iran, Islamic Rep.	South Africa
Belgium	Indonesia	Spain
Bulgaria	Ireland	Sri Lanka
Botswana	Italy	Sweden
Canada	Mexico	Switzerland
China	Morocco	Tanzania
Colombia	Mozambique	Thailand
Costa Rica	Netherlands	Turkey
Czech Republic	Norway	United Kingdom
Egypt, Arab Rep.	Pakistan	United States
Finland	Peru	Uzbekistan
	Philippines	Zimbabwe

A2. VIF test results

Model 1

Variance Inflation Factors

Sample: 1991 2016
Included observations: 871

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	845.5643	84928.07	NA
LOG(EPC)	13.46732	84255.39	50.54088
LOG(EPC)^2	0.076406	31253.91	42.22592
LOG(GDPPC)	74.51730	692292.3	95.80211
LOG(GDPPC)^2	0.243741	213683.3	89.41289
EMPL_AGR	0.000854	35.17013	1.485103
UNEMPL_FEM	0.003910	51.65549	3.378452
UNEMPL_MALE	0.004705	40.56425	3.511942
ED_PRIM	0.000443	454.4013	1.649136
ED_SEC	0.000153	115.7240	1.312770
FERTILITY	0.255110	126.8113	1.421314

Model 2.

Variance Inflation Factors
Sample: 1991 2016
Included observations: 871

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	882.2918	82800.11	NA
LOG(EPC)	14.40787	84222.96	50.52143
(LOG(EPC))^2	0.081120	31004.04	41.88832
LOG(GDPPC)	79.76569	692408.8	95.81824
LOG(GDPPC)^2	0.263493	215836.5	90.31388
EMPL_IND	0.002693	65.05119	1.238738
UNEMPL_FEM	0.004244	52.39760	3.426989
UNEMPL_MALE	0.005476	44.11282	3.819167
ED_PRIM	0.000464	443.9226	1.611106
ED_SEC	0.000161	113.7350	1.290206
FERTILITY	0.268208	124.5709	1.396204

Model 3.

Variance Inflation Factors
Sample: 1991 2016
Included observations: 871

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	799.2069	81378.04	NA
LOG(EPC)	13.48108	85503.65	51.28966
(LOG(EPC))^2	0.076469	31710.72	42.84309
LOG(GDPPC)	72.78097	685478.2	94.85916
LOG(GDPPC)^2	0.238998	212411.5	88.88073
EMPL_SERV	0.001247	526.6800	1.629003
UNEMPL_FEM	0.003851	51.58171	3.373627
UNEMPL_MALE	0.004967	43.41635	3.758869
ED_PRIM	0.000415	431.4493	1.565838
ED_SEC	0.000149	114.1177	1.294547
FERTILITY	0.246070	124.0032	1.389841

Model 4.

Variance Inflation Factors
Sample: 1991 2016
Included observations: 950

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	471.8258	49754.91	NA
LOG(GDPPC)	26.11182	255964.8	37.91504
(LOG(GDPPC))^2	0.092017	85509.52	38.78866
EMPL_AGR	0.000766	31.90664	1.469139
UNEMPL_FEM	0.003722	50.33789	3.402687
UNEMPL_MALE	0.004466	39.84511	3.517522
ED_PRIM	0.000391	421.0209	1.544095
ED_SEC	0.000126	103.1686	1.191493
FERTILITY	0.213721	109.8841	1.279091

Model 5.

Variance Inflation Factors
Sample: 1991 2016
Included observations: 950

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	490.3673	48978.16	NA
LOG(GDPPC)	29.29297	271977.6	40.28696
LOG(GDPPC)^2	0.104508	91986.63	41.72679
EMPL_IND	0.002474	61.72540	1.195372
UNEMPL_FEM	0.003983	51.02967	3.449450
UNEMPL_MALE	0.005053	42.69413	3.769033
ED_PRIM	0.000394	402.1201	1.474776
ED_SEC	0.000129	99.95135	1.154337
FERTILITY	0.225146	109.6424	1.276277

Model 6.

Variance Inflation Factors
Sample: 1991 2016
Included observations: 950

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	426.2755	45643.01	NA
LOG(GDPPC)	25.01511	248986.1	36.88131
LOG(GDPPC)^2	0.091008	85872.71	38.95341
EMPL_SERV	0.001128	510.9968	1.593008
UNEMPL_FEM	0.003668	50.36945	3.404821
UNEMPL_MALE	0.004724	42.79170	3.777647
ED_PRIM	0.000371	405.2567	1.486279
ED_SEC	0.000123	101.6126	1.173522
FERTILITY	0.207700	108.4307	1.262173

A3. Correlation matrix

	GDPPC	EPC	EMPL_AGR	EMPL_IND	EMPL_SERV	ED_PRIM	ED_SEC	FERTILITY
GDPPC	1.000	0.820	-0.655	-0.273	0.745	0.153	0.598	-0.543
EPC	0.820	1.000	-0.509	-0.280	0.601	0.078	0.548	-0.398
EMPL_AGR	-0.655	-0.509	1.000	-0.161	-0.945	-0.343	-0.669	0.595
EMPL_IND	-0.273	-0.280	-0.161	1.000	-0.172	0.030	-0.074	-0.224
EMPL_SERV	0.745	0.601	-0.945	-0.172	1.000	0.333	0.692	-0.519
ED_PRIM	0.153	0.078	-0.343	0.030	0.333	1.000	0.308	-0.366
ED_SEC	0.598	0.548	-0.669	-0.074	0.692	0.308	1.000	-0.685
FERTILITY	-0.543	-0.398	0.595	-0.224	-0.519	-0.366	-0.685	1.000

As we can see here, GDP per capita and Energy Power Consumption are highly correlated, which is logical since the income level influences how much can be invested in the production of electric power and thereby how much is consumed.

A4. Fixed effects regression results

Model 1.

Dependent Variable: FLPR
Method: Panel Least Squares
Sample (adjusted): 1991 2014
Periods included: 24
Cross-sections included: 47
Total panel (unbalanced) observations: 871

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	240.3650	29.07859	8.266047	0.0000
LOG(EPC)	-5.299945	3.669784	-1.444212	0.1491
(LOG(EPC))^2	0.720459	0.276417	2.606424	0.0093
LOG(GDPPC)	-37.02790	8.632340	-4.289440	0.0000
LOG(GDPPC)^2	1.609259	0.493702	3.259578	0.0012
EMPL_AGR	0.221474	0.029224	7.578588	0.0000
UNEMPL_FEM	-0.099417	0.062527	-1.589990	0.1122
UNEMPL_MALE	0.189231	0.068590	2.758877	0.0059
ED_PRIM	0.032783	0.021059	1.556706	0.1199
ED_SEC	-0.000940	0.012368	-0.075972	0.9395
FERTILITY	0.417488	0.505084	0.826573	0.4087

Effects Specification

Cross-section fixed (dummy variables)
Period fixed (dummy variables)

Root MSE	2.806312	R-squared	0.963645
Mean dependent var	47.82750	Adjusted R-squared	0.960014
S.D. dependent var	14.72670	S.E. of regression	2.944807
Akaike info criterion	5.085316	Sum squared resid	6859.462
Schwarz criterion	5.523400	Log likelihood	-2134.655
Hannan-Quinn criter.	5.252930	F-statistic	265.4036
Durbin-Watson stat	0.141950	Prob(F-statistic)	0.000000

Model 2.

Dependent Variable: FLPR
Method: Panel Least Squares
Sample (adjusted): 1991 2014
Periods included: 24
Cross-sections included: 47
Total panel (unbalanced) observations: 871

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	280.7543	29.70340	9.451926	0.0000
LOG(EPC)	-2.841931	3.795770	-0.748710	0.4543
LOG(EPC)^2	0.489283	0.284816	1.717893	0.0862
LOG(GDPPC)	-43.17734	8.931164	-4.834458	0.0000
LOG(GDPPC)^2	1.812914	0.513316	3.531769	0.0004
EMPL_IND	-0.068527	0.051898	-1.320411	0.1871
UNEMPL_FEM	-0.111751	0.065149	-1.715327	0.0867
UNEMPL_MALE	0.136854	0.073997	1.849455	0.0648
ED_PRIM	0.071400	0.021534	3.315725	0.0010
ED_SEC	-0.015465	0.012685	-1.219129	0.2232
FERTILITY	-0.254235	0.517888	-0.490908	0.6236

Effects Specification

Cross-section fixed (dummy variables)
Period fixed (dummy variables)

Root MSE	2.903213	R-squared	0.961091
Mean dependent var	47.82750	Adjusted R-squared	0.957205
S.D. dependent var	14.72670	S.E. of regression	3.046490
Akaike info criterion	5.153210	Sum squared resid	7341.350
Schwarz criterion	5.591294	Log likelihood	-2164.223
Hannan-Quinn criter.	5.320824	F-statistic	247.3252
Durbin-Watson stat	0.122506	Prob(F-statistic)	0.000000

.

Model 3.

Dependent Variable: FLPR
Method: Panel Least Squares
Sample (adjusted): 1991 2014
Periods included: 24
Cross-sections included: 47
Total panel (unbalanced) observations: 871

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	277.6522	28.27025	9.821358	0.0000
LOG(EPC)	-7.543154	3.671659	-2.054427	0.0403
LOG(EPC)^2	0.868048	0.276531	3.139065	0.0018
LOG(GDPPC)	-40.41891	8.531177	-4.737788	0.0000
LOG(GDPPC)^2	1.881122	0.488874	3.847868	0.0001
EMPL_SERV	-0.293645	0.035318	-8.314196	0.0000
UNEMPL_FEM	-0.139247	0.062056	-2.243893	0.0251
UNEMPL_MALE	0.317412	0.070476	4.503823	0.0000
ED_PRIM	0.047725	0.020381	2.341697	0.0194
ED_SEC	-0.004076	0.012199	-0.334151	0.7384
FERTILITY	0.216744	0.496054	0.436936	0.6623

Effects Specification

Cross-section fixed (dummy variables)
Period fixed (dummy variables)

Root MSE	2.787175	R-squared	0.964139
Mean dependent var	47.82750	Adjusted R-squared	0.960558
S.D. dependent var	14.72670	S.E. of regression	2.924725
Akaike info criterion	5.071631	Sum squared resid	6766.228
Schwarz criterion	5.509715	Log likelihood	-2128.695
Hannan-Quinn criter.	5.239245	F-statistic	269.1986
Durbin-Watson stat	0.160483	Prob(F-statistic)	0.000000

Model 4.

Dependent Variable: FLPR
Method: Panel Least Squares
Sample: 1991 2016
Periods included: 26
Cross-sections included: 47
Total panel (unbalanced) observations: 950

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	213.8311	21.72155	9.844190	0.0000
LOG(GDPPC)	-34.35331	5.109973	-6.722798	0.0000
LOG(GDPPC)^2	1.670420	0.303343	5.506706	0.0000
EMPL_AGR	0.194748	0.027680	7.035597	0.0000
UNEMPL_FEM	-0.134550	0.061005	-2.205561	0.0277
UNEMPL_MALE	0.239723	0.066831	3.587035	0.0004
ED_PRIM	0.046647	0.019773	2.359182	0.0185
ED_SEC	0.008791	0.011244	0.781856	0.4345
FERTILITY	-0.560608	0.462300	-1.212649	0.2256

Effects Specification

Cross-section fixed (dummy variables)
Period fixed (dummy variables)

Root MSE	2.872318	R-squared	0.960947
Mean dependent var	48.06945	Adjusted R-squared	0.957400
S.D. dependent var	14.54226	S.E. of regression	3.001475
Akaike info criterion	5.116537	Sum squared resid	7837.700
Schwarz criterion	5.525502	Log likelihood	-2350.355
Hannan-Quinn criter.	5.272359	F-statistic	270.9771
Durbin-Watson stat	0.127436	Prob(F-statistic)	0.000000

Model 5.

Dependent Variable: FLPR
Method: Panel Least Squares
Sample: 1991 2016
Periods included: 26
Cross-sections included: 47
Total panel (unbalanced) observations: 950

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	252.2717	22.14424	11.39220	0.0000
LOG(GDPPC)	-38.78585	5.412298	-7.166243	0.0000
LOG(GDPPC)^2	1.754713	0.323278	5.427884	0.0000
EMPL_IND	-0.047713	0.049744	-0.959169	0.3377
UNEMPL_FEM	-0.142151	0.063112	-2.252347	0.0245
UNEMPL_MALE	0.187782	0.071082	2.641773	0.0084
ED_PRIM	0.083155	0.019855	4.188051	0.0000
ED_SEC	-0.006893	0.011372	-0.606104	0.5446
FERTILITY	-0.954050	0.474496	-2.010660	0.0447

Effects Specification

Cross-section fixed (dummy variables)
Period fixed (dummy variables)

Root MSE	2.951339	R-squared	0.958768
Mean dependent var	48.06945	Adjusted R-squared	0.955024
S.D. dependent var	14.54226	S.E. of regression	3.084049
Akaike info criterion	5.170816	Sum squared resid	8274.884
Schwarz criterion	5.579782	Log likelihood	-2376.138
Hannan-Quinn criter.	5.326638	F-statistic	256.0788
Durbin-Watson stat	0.115021	Prob(F-statistic)	0.000000

Model 6.

Dependent Variable: FLPR
Method: Panel Least Squares
Sample: 1991 2016
Periods included: 26
Cross-sections included: 47
Total panel (unbalanced) observations: 950

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	255.2624	20.64644	12.36351	0.0000
LOG(GDPPC)	-40.59731	5.001511	-8.117009	0.0000
LOG(GDPPC)^2	2.087492	0.301675	6.919672	0.0000
EMPL_SERV	-0.267944	0.033587	-7.977701	0.0000
UNEMPL_FEM	-0.170227	0.060560	-2.810887	0.0051
UNEMPL_MALE	0.354815	0.068731	5.162376	0.0000
ED_PRIM	0.053161	0.019251	2.761413	0.0059
ED_SEC	0.007509	0.011074	0.678088	0.4979
FERTILITY	-0.725863	0.455741	-1.592711	0.1116

Effects Specification

Cross-section fixed (dummy variables)
Period fixed (dummy variables)

Root MSE	2.850478	R-squared	0.961538
Mean dependent var	48.06945	Adjusted R-squared	0.958046
S.D. dependent var	14.54226	S.E. of regression	2.978653
Akaike info criterion	5.101272	Sum squared resid	7718.964
Schwarz criterion	5.510237	Log likelihood	-2343.104
Hannan-Quinn criter.	5.257093	F-statistic	275.3148
Durbin-Watson stat	0.143127	Prob(F-statistic)	0.000000