

Analysis of Public Employment Rate

David Pham

2016-01-13

Contents

1	Summary	1
2	Model and Data	2
2.1	Data	2
2.2	Model	2
2.3	Results	3
2.4	Note on robustness	4
3	Challenges with the data	4
3.1	Note about the data	4
4	Note	6
5	Bibliography	6

1 Summary

Using the 2015 OECD Economic Outlook database, factors explaining public employment rate within countries were uncovered. The analysis with annual data emphasizes *GDP growth* as the main contributor of the public employment rate, whereas with quarterly data, the *unemployment rate* contributed the most.

Neither fiscal transparency nor income inequality were retained as significant variables in the analysis (with annual and quarterly data). Results were mixed with the variables *political direction of the executive government* and *the period left until the next official elections*.

Missing data (variables and frequency) made the analysis challenging and using smoother and adaptive analysis might lead to over-interpretation: The adjustment of the period the analysis might considerably impact the regression coefficients.

2 Model and Data

2.1 Data

Following [?], the next variable were selected from the Economic Outlook database to explain the public employment rate.

Table 1: Main variables for the regression

Variable	OECD code
Country	country
Government employment	eg
Total employment	et
GDP growth	gdpv_annpct
Unemployment rate	unr
Government spending	ypgtq
adult population	pop1574
Net household disposable income	ydrh

Additionally, the below time series were retrieved to compare the results from [?] and [?]:

Table 2: Additional variables for the regression

Variable	Source
GPD per capita	[?]
Gini index	[?]
Gini index	[?]
Fiscal transparency	IMF, [?]
Laassen fiscal transparency Score	[?]
Left or right side government	[?]
Year until next elections	[?]

Unfortunately, these data had an annual frequency only. Note that the IMF fiscal transparency data provides time series for the last ten years. For the analysis described here, only data from 1990 up to 2012 (included) were considered.

2.2 Model

[?] explains the level of public employment with gdp growth with an additional nonlinear effect provided by the fiscal transparency score of a country, namely

$$Y_{i,t} = \alpha + \beta_1 G_{i,t} + \beta_2 G_{i,t} T_i + \beta_3 X_{i,t} + \eta_i + \tau_t + \varepsilon_{i,t},$$

where the variables are

- $Y_{i,t}$, the public employment;

- $G_{i,t}$, GDP growth;
- T_i , fiscal transparency;
- X is the vector of control variables;
- η_i , country fixed effect;
- τ_t , year fixed effect.

However, with the data from the OECD, the diagnostic plots of this model are quite disturbing: residuals show a clear correlation and tails of the distribution clearly deviate from the Gaussian assumption. Moreover, in my opinion, it is quite controversial to add a nonlinear effect ($\beta_2 G_{it} T_i$) without looking at the linear effect of T_i .

After several attempts to find an acceptable model, the fixed effect with auto-regressive components looked promising. In this model, the difference of the public employment are explained:

$$\Delta Y_{i,t} = Y_{i,t} - Y_{i,t-1} = \alpha + \rho Y_{i,t-1} + \beta_1 G_{it} + \beta_2 X_{i,t} + \beta_3 X_{i,t-1} + \tau_t + \varepsilon_{it},$$

where the variable are same as previously. Controls variables (in X_{it}) were:

- Unemployment rate;
- Log of adult population (aged between 15-74 years old);
- Government expenses;
- Net household disposable income in percent of GDP.

Robustness has been studied by adding and deleting the following variables from the predictors.

- Total population;
- GDP per capita;
- Gini coefficient;
- Fiscal transparency scores (IMF time series, Lassen scores and IMF scores);
- Government political side;
- Years until next official elections.

Detailed results are found in appendix.

2.3 Results

With annual data, gdp growth was definitively the main contributor for explaining changes in public employment rate. Government expenditure and unemployment rate were not always significant in the regression. It highly depends on which the period of data included in the analysis. More precisely, their regression coefficient often lies around the 10%-level of rejection ¹. Political side of the government and the period until the next official election were not significant with the period of the analysis (1990-2012). However, when including all variables from 1960 onward, government political direction is significant at the 1%-level and years left until election at the 10%-level.

Concerning quarterly data, only unemployment rate was consistently statistically significant, whereas GDP growth almost never (probably seasonal effect were taken adjusted in GDP growth). Political side of government and years left until the next election were statistically significant (at the 10%-level) depending on the period of the analysis.

¹Nevertheless, with so many predictors and regression, one should consider adjusting the significance threshold.

Note that with quarterly data some predictors are interpolated, hence the model can hardly predict shocks. Neither income inequality (Gini coefficient) nor fiscal transparency affected the main coefficient of the regression.

Comparing with [?], we get similar results, excluding effect of fiscal transparency.

2.4 Note on robustness

The previous analysis should be repeated with more robust methods, e.g. by scaling the variable, using cross-validation and bootstrapped estimators of the errors, using other statistical methods (CART, LASSO or GLM). Nevertheless, using these alternative render interpretation more difficult. Additionally, missing methods (as in [?]) could be applied have better predictors.

If all the data from the OECD database are included in the regression, that some variables are then statistically significant, one should study this effect.

In robustness analysis, as complete-case analysis is performed, the number of observation varies significantly.

3 Challenges with the data

The data present some challenges: missingness among variable and among frequency.

As observations are not always present for all variables for a country at a fixed year, the common approach in regression analysis is to keep only complete-case. This can obviously lead to some bias. In this study, when doing robustness analysis, it is possible that the size of the data set vary considerably between models.

Concerning quarterly data, most of control variable are not reported on quarterly basis in the oecd data set ². A spline smoother has been applied to interpolated between points and obviously this affect the quality of the fits with quarterly data as the independent variable is too smooth.

3.1 Note about the data

Fiscal transparency is interacted with gdp growth as it should contribute as a good factor. Here are the Lassen Scores as in [?]

Table 3: Fiscal transparency scores from [?].

Country	Fiscal Score
Greece	1
Italy	2
Norway	2
Belgium	3
Denmark	3
Germany	3
Ireland	3
Spain	3
Switzerland	3

²The data set in Stata format seemed to code quarters and date unconventionally.

Country	Fiscal Score
Austria	4
France	4
Portugal	4
Canada	5
Finland	5
Netherlands	5
Sweden	5
Australia	6
United Kingdom	8
United States	9
New Zealand	10

Here is the list of countries belonging in the OECD database. A cross tells if public employment data were available.

Country	Annually	Quarterly
Australia		
Austria		
Belgium	x	x
Brazil		
Canada	x	x
Chile		
China (People's Republic of)		
Colombia		
Czech Republic	x	x
Denmark	x	x
Estonia	x	x
Finland	x	x
France	x	x
Germany		
Greece		
Hungary	x	x
Iceland		
India		
Indonesia		
Ireland	x	x
Israel	x	
Italy	x	
Japan	x	x
Korea		
Latvia		
Luxembourg	x	x
Mexico		
Netherlands	x	x
New Zealand		
Norway	x	x
Poland	x	x
Portugal	x	
Russia		

Country	Annually	Quarterly
Slovak Republic	x	
Slovenia		
South Africa		
Spain	x	
Sweden	x	x
Switzerland		
Turkey	x	
United Kingdom	x	x
United States	x	x

4 Note

The original variables were selected because of the prior distribution we had on them. Advanced statistical methods should be tried to discover if any other predictors could be better. For example, in quarterly data, Lasso method selected the variable *Net exports, contribution to growth* as a potential variable.

5 Bibliography