Charles University in Prague
Faculty of Social Sciences
Institute of Economic Studies

Home assignment 2

Marek Chadim, Jakub Strašlipka

Econometrics II – JEB110
Winter semester 2022/2023

Problem 2

Introduction and Theoretical Background

In this problem, we seek to explore the relationship between government expenditures and household consumption. Keynesian macroeconomic theory suggests that as government increases its spending, the households should follow by an increase in consumption. This stems from the definition of consumption

$$C = c_0 + c(Y - \overline{T}), \tag{1}$$

where $c_0 > 0$ is constant in time, $c \in (0, 1)$ is a marginal propensity to consume, and $(Y - \overline{T})$ is a disposable income. Since government expenditures G contribute to the disposable income, to the Y term in particular, their increase should, ceteris paribus, induce an increase in C. More specifically, the model of fiscal multiplier directly links the change in household expenditures to the change in government expenditures (see for instance Mankiw (2010, p. 293-294)) via a simple linear equation

$$\Delta C = \frac{c}{1-c}\Delta G. \tag{2}$$

Because equation (2) theoretically assumes infinite number of time periods, let us discuss the time factor. In general, it is unlikely that the change in government expenditures has an immediate impact on consumption; on the other hand, it would be unreasonable to expect a significant impact after more than a few years, so for the purpose of this report, our expectations are that government expenditures have effect in the year in which they are budgeted and in one following year.

Other variables that we expect to be statistically significant in forecasting consumption are (a) real interest rates (negative relationship – the higher the IR, the more the households are incentivized to save instead of spending), (b) real GDP (positive relationship).

Data

To test whether the positive relationship exists also empirically, we use the World Bank data for all 38 current OECD member countries. We opt to use the real growth variables per capita (in percent), which account for both the price level changes and population changes. The main variables of interest are the following:

- real household consumption (C) – obtained from dataset *Households and NPISHs Final consumption expenditure per capita growth (annual %)*; NPISH = Non-Profit Institutions Serving Households;
- real government expenditures (G) – obtained from dataset *General government final consumption expenditure (annual % growth)*; we use the current value and the first lags of change government expenditures to allow for a delay in consumption response.

Other data we got from the World Bank are real absolute GDP (Y) and real interest rates[1] (r).

All data are yearly observations from 2012–2019. We choose this time period for its relative economic stability as most countries had largely recovered from the Great Recession by 2012 and 2019 was the last pre-pandemic year; we exclude years 2020 and 2021 due to the uniqueness and severity of the exogenous covid shock that, we suspect, could distort the results.

Model Specification and Assumptions

We will estimate the following panel regression

$$C_{it} = \beta_0 + \beta_1 G_{it} + \beta_2 G_{i,t-1} + \beta_3 Y_{it} + \beta_4 r_{it} + a_i + u_{it}, \tag{3}$$

where variables are those described in the previous section.

[1]the most difficult data to acquire; dataset of real interest rates from World Bank contained RIR for only a half of our countries, the rest were calculated from nominal IR available at OECD and country CPI available at World Bank

Our panel data is balanced with no missing value. As we include one lag, we will have 7 time periods for all 38 OECD countries, thus 266 individual observations.

We will estimate the the panel regression using both the fixed effects and random effects method, then perform a Hausman test, test the better model (as determined by the Hausman test) for heteroskedasticity, and finally, in case of heteroskedasticity, perform a regression with robust standard errors.

Regarding the FE, RE assumptions:

(a) The model is linear, where β_j are estimated parameters and a_i is an unobserved effect.

(b) Since we use the entire population of OECD countries, we fulfill the random sample assumption. An alternative could be to eliminate every second (or third or fourth) country in alphabetical order.

(c) We have no constant and no perfect linear relationships among the explanatory variables.

(d) We must rely on assumption that the idiosyncratic error given the explanatory variables in all time periods and the unobserved effect is zero.

(e) Similarly, we assume that the expected value of a_i given all explanatory variables is constant.

(f) We will test for heteroskedasticity.

Results

Results for all three estimated model are reported in the table.

	Fixed effects	Random effects	Random effects (robust SE)
Intercept	–	0.7621***	0.7621**
		(0.2115)	(0.3772)
G_t	0.0776	0.0564	0.0564
	(0.0562)	(0.0535)	(0.0594)
G_{t-1}	0.1454***	0.1192**	0.1192**
	(0.0552)	(0.522)	(0.0515)
Y_t	0.2800***	0.2931***	0.2931
	(0.0474)	(0.0447)	(0.1844)
i_t	0.1622***	0.1207**	0.1207***
	(0.0606)	(0.0487)	(0.0444)
Note:	$^{***}p < 0.1$;	$^{**}p < 0.05$;	$^{***}p < 0.01$

We fail to reject the null hypothesis in the Hausman test ($p = 0.3499$), thus we proceed by using the random effects model.

We perform the Lagrange multiplier test on the random effects model. As $p < 0.01$, we reject the null hypothesis of homoskedasticity and continue by regressing the random effects model with heteroskedasticity robust standard errors (as reported in the results table).

Discussion

We find a statistically significant positive relationship between the change in government expenditures and change in household consumption. The results suggest that change in gov. spending from the previous year affect current consumption more that current change in gov. spending. The estimated effect is that 1% increase in government expenditures induces 0.05% increase in household consumption for the year in which the change was budgeted (not a statistically significant realtionship), and a 0.12% increase in consumption in the following time period (statistically significant).

Interestingly, when using heteroskedasticity robust standard errors, we do not find a statistically significant effect of GDP growth on household consumption. Nevertheless, in line with the predictions of the ordinary fixed and random effects models, our belief is that the positive relationship exists and that a dataset with more observations would yield a statistically significant result.

Finally, to our surprise, we obtain a significant positive relationship between real interest rates and consumption. We must beware ourselves from concluding that increase in RIR causes a consumption increase. Our explanation of the relationship is that since the interest rates were low (near zero) between 2012 and 2019 in nearly all OECD countries, consumers were not incentivized enough to save and were indifferent to 25 bps changes in interest rates (e.g. from 0.25% to 0.5%). Furthermore, we suspect that central banks tended to increas interest rates only ex-post, i.e. only after consumption increased (thus affecting the GDP, inflation), not ex-ante as contemporary monetary policy should, therefore giving rise to the positive relationship between the interest rates and household consumption (consumption$\nearrow \Rightarrow$ IR$\nearrow$).

References

Mankiw, N. G. (2010). Macroeconomics (7th ed.). Worth Publishers.

Problem 1

Consider a fixed effects model:

$$y_{it} = \beta_0 + \beta_1 x1_{it} + \beta_2 x2_{it} + a_i + u_{it}.$$

Define the composite error $v_{it} = a_i + u_{it}$, where a_i is uncorrelated with u_{it} and the u_{it} have constant variance $\sigma_u{}^2$ and are serially uncorrelated.

(a) Derive $Corr[v_{it}, v_{is}]$ for $t \neq s$. What does it mean for pooled OLS estimators of model coefficients?

Since the model includes an intercept, we can suppose that a_1, u_{i1} and u_{i2} have zero means without losing generality.

Thus $E(v_{it}) - E(a_i - u_{it}) = E(a_i) + E(u_{it}) = 0$. Similarly, $E(v_{is}) = 0$

Therefore, the covariance between v_{it} and v_{is} is $E(v_{it} v_{is}) - E(v_{it}) E(v_{is}) =$
$= E[(a_i + u_{it})(a_i + u_{is})] = E(a_i^2) + E(a_i u_{it}) + E(a_i u_{is}) + E(u_{it} u_{is}) = E(a_i^2)$, because all of the covariance terms are zero by assumption. But $E(a_i^2) = Var(a_i)$, because $E(a_i) = 0$.

This causes positive serial correlation across time within each i, unless $a_i = 0$.

Further, the variance of v_{it} is $E(v_{it}^2) - [E(v_{it})]^2 = E(a_i^2) + Var(u_{it}) - [E(a_i)]^2 =$
$= Var(a_i) + Var(u_{it})$. Analogously, the variance of v_{is} is $Var(a_i) + Var(u_{is})$.

Finally, under the random effects assumptions,

$$Corr(v_{it}, v_{is}) = \frac{Cov(v_{it}, v_{is})}{\sqrt{Var(v_{it})} \cdot \sqrt{Var(v_{is})}} = \sigma_a^2 / (\sigma_a^2 + \sigma_u^2), \quad t \neq s,$$

where $\sigma_a^2 = Var(a_i)$ and $\sigma_u^2 = Var(u_{it})$.

This (necessarily) positive serial correlation in the error term can be substantial, and, because the usual pooled OLS standard errors ignore this correlation, they will be incorrect, as will the usual test statistics.

(b) Does your answer in point (a) above change when $a_i = 0$, i.e. when there are no unobservable time-constant characteristics of the cross-sectional units? Explain.

When $a_i = 0$, $Corr[v_{it}, v_{is}] = Corr[u_{it}, u_{is}] = 0$, for all $t \neq S$ and the composite errors no longer suffer from seriall correlation.

Since the variance of the OLS slope estimator now corresponds to the OLS variance under Gauss-Markov assumptions the pooled OLS slope estimators have lower variance than in point (a).

(c) Consider now the random effects transformation. Define $e_{it} = v_{it} - \lambda\overline{v_i}$, where λ is as defined during the lecture.

(i) Derive $E[e_{it}]$.

(ii) Derive $Var[e_{it}]$.

(iii) Derive $Corr[e_{it}, e_{is}]$ for $t \neq s$.

(i) $E(e_{it}) = E(v_{it} - \lambda\overline{v_i}) = E(v_{it}) - \lambda E(\overline{v_i}) = 0$ because $E(v_{it}) = 0$ for all t.

(ii) $Var(v_{it} - \lambda\overline{v_i}) = Var(v_{it}) + \lambda^2 Var(\overline{v_i}) - 2\lambda \cdot Cov(v_{it}, \overline{v_i}) = \sigma_v^2 + \lambda^2 E(\overline{v_i}^2) - 2\lambda \cdot E(v_{it}\overline{v_i})$.

Now, $\sigma_v^2 = E(v_{it}^2) = \sigma_a^2 + \sigma_u^2$ and

$E(v_{it}\overline{v_i}) = T^{-1}\sum_{s=1}^{T} E(v_{it}v_{is}) = T^{-1}[\sigma_a^2 + \sigma_a^2 + \ldots + (\sigma_a^2 + \sigma_u^2) + \ldots + \sigma_a^2] = \sigma_a^2 + \sigma_u^2/T$.

Therefore, $E(\overline{v_i}^2) = T^{-1}\sum_{t=1} E(v_{it}\overline{v_i}) = \sigma_a^2 + \sigma_u^2/T$. Now, we can collect terms:

$$Var(v_{it} - \lambda\overline{v_i}) = (\sigma_a^2 + \sigma_u^2) + \lambda^2(\sigma_a^2 + \sigma_u^2/T) - 2\lambda(\sigma_a^2 + \sigma_u^2/T).$$

Now, it is convenient to write $\lambda = 1 - \sqrt{\eta}/\sqrt{\gamma}$, where $\eta \equiv \sigma_u^2/T$ and $\gamma \equiv \sigma_a^2 + \sigma_u^2/T$.

Then
$$\begin{aligned}
Var(v_{it} - \lambda\overline{v_i}) &= (\sigma_a^2 + \sigma^2) - 2\lambda(\sigma_a^2 + \sigma_u^2/T) + \lambda^2(\sigma_a^2 + \sigma_u^2/T) \\
&= (\sigma_a^2 + \sigma_u^2) - 2(1 - \sqrt{\eta}/\sqrt{\gamma})\gamma + (1 - \sqrt{\eta}/\sqrt{\gamma})^2\gamma \\
&= (\sigma_a^2 + \sigma_u^2) - 2\gamma + 2\sqrt{\eta}\cdot\sqrt{\gamma} + (1 - 2\sqrt{\eta}/\sqrt{\gamma} + \eta/\gamma)\gamma \\
&= (\sigma_a^2 + \sigma_u^2) - 2\gamma + 2\sqrt{\eta}\cdot\sqrt{\gamma} + \gamma - 2\sqrt{\eta}\cdot\sqrt{\gamma} + \eta \\
&= (\sigma_a^2 + \sigma_u^2) + \eta - \gamma = \sigma_u^2, \quad t = 1, \ldots, T.
\end{aligned}$$

(iii) Let us show that $E(e_{it}e_{is}) = 0$ for $t \neq s$.

Now
$$\begin{aligned}
E(e_{it}e_{is}) &= E[(v_{it} - \lambda\overline{v_i})(v_{is} - \lambda\overline{v_i})] = E(v_{it}v_{is}) - \lambda E(\overline{v_i}v_{is}) - \lambda E(v_{it}\overline{v_i}) + \lambda^2 E(\overline{v_i}^2) \\
&= \sigma_a^2 - 2\lambda(\sigma_a^2 + \sigma_u^2/T) + \lambda^2 E(\overline{v_i}^2) = \sigma_a^2 - 2\lambda(\sigma_a^2 + \sigma_u^2/T) + \lambda^2(\sigma_a^2 + \sigma_u^2/T).
\end{aligned}$$

The rest of the proof is very similar to part (ii):
$$\begin{aligned}
E(e_{it}e_{is}) &= \sigma_a^2 - 2\lambda(\sigma_a^2 + \sigma_u^2/T) + \lambda^2(\sigma_a^2 + \sigma_u^2/T) \\
&= \sigma_a^2 - 2(1 - \sqrt{\eta}/\sqrt{\gamma})\gamma + (1 - \sqrt{\eta}/\sqrt{\gamma})^2\gamma \\
&= \sigma_a^2 - 2\gamma + 2\sqrt{\eta}\cdot\sqrt{\gamma} + (1 - 2\sqrt{\eta}/\sqrt{\gamma} + \eta/\gamma)\gamma \\
&= \sigma_a^2 - 2\gamma + 2\sqrt{\eta}\cdot\sqrt{\gamma} + \gamma - 2\sqrt{\eta}\cdot\sqrt{\gamma} + \eta \\
&= \sigma_a^2 + \eta - \gamma = 0.
\end{aligned}$$

Thus, for $t \neq s$:

$$E(e_{it}e_{is}) = 0 \Rightarrow Cov(e_{it}, e_{is}) = 0 \Rightarrow Corr(e_{it}, e_{is}) = 0 \quad \square.$$