Which "Reality"? A reply to Perotti-Sala(2024)[1]

C.A. Favero

October 2024-IMF IEO

[1] I am grateful to Luca Sala and Roberto Perotti for sharing their codes, for many lively interactions and for being great friends and colleagues.

Introduction

- (Perotti and Sala, 2024) study the response of actual (as opposed to announced) discretionary government spending and revenues, and claim to find that the former does not decline in what are classified (on the basis of announcements) as "expenditure-based" consolidations, but does decline by a large amount in "tax-based" consolidations; actual discretionary revenues move little in both regimes.
- **These results depend crucially on their computing impulse responses in a way which is not consistent with the standard definition of the causal effect of a policy intervention.**
- The computation of impulse responses consistent with the standard definition of the causal effect of a policy intervention within the (Perotti and Sala, 2024) specification leads to a different picture of "reality" and does not support their conclusion.
- In addition, some further reflections on the (Perotti and Sala, 2024) specification are offered.

Measuring the Effect of Fiscal Adjustments

- Following (Angrist et al., 2018) define the causal effect of a policy intervention as the unobservable random variable given by the difference
  $(y_{t,h}(d_1) - y_t) - (y_{t,h}(d_0) - y_t)$
  where d_1 measures the (exogenous) policy intervention and d_0 indicates no policy intervention. y_t is only used to benchmark the cumulative change and it is observed at time t.
- Given the specification of a dynamic model for fiscal plans and macroeconomic variables, the impact of fiscal plans on macroeconomic variables can be computed by constructing an impulse response (IR).
- This is obtained as the difference between two dynamic forecasts for the variables of interests: a baseline computed in absence of a policy intervention and an alternative one, computed in presence of policy intervention:

$$IR\left(t,s,d_i\right) = E\left(\mathbf{Y}_{i,t+s} - \mathbf{Y}_{i,t} \mid plans_t;\ I_t\right) - E\left(\mathbf{Y}_{i,t+s} - \mathbf{Y}_{i,t} \mid no\ plans_t;\ I_t\right) \quad s = 0, 1, 2, \ldots$$

The AFG model considered by PS I

- The derivation of IRs requires the specification of an econometric model to predict the variables of interest, (Alesina et al., 2019) consider two type of models in their book: VAR models for the policy variables and the macroeconomic outcome and single equation MA representations that project macroeconomic variables of interests on a finite MA of the policy interventions.
- This second simplified representation suffers of potential omitted variables problems. However, if the omitted shocks are orthogonal to those included (validity of the narrative procedure should deliver this outcome) and if the excluded lags of the intervention are not significant, then the simplified model does still deliver consistent estimates.
- (Perotti and Sala, 2024) focus exclusively on the following reduced form representation of our MA approach (reduced form because plans are collapsed into shocks using correctly our modelling of plans).

The AFG model considered by PS II

- The Econometric Model:

$$\Delta y_{i,t} = \alpha + \lambda_i + \chi_t + \sum_{j=0}^{4} \beta_{j+1} e^{u}_{i,t-j} * TB_{i,t-j} + \sum_{j=0}^{4} \gamma_{j+1} e^{u}_{i,t-j} * EB_{i,t-j} + u_{i,t}$$

- IRs for each type of policy intervention (Tax-Based and Expenditure-Based) are then obtained by simulating the model forward for four-periods (years) under two scenarios: the baseline in absence of the policy and the alternative one in absence of the policy. Importantly presence of the policy means **a shock is present, standardized to be 1 (one per cent of GDP adjustment), and the activation dummy is set to 1**. Effects on the level of variables are then computed by cumulating the effects on differences.
- To illustrate the procedure consider the contemporaneous effect of a TB based adjustment, we then have:

$$\begin{aligned} E\left(\Delta y_{i,t} \mid e^{u}_{i,t-j} = 1,\ TB_{i,t} = 1, EB_{i,t} = 0\right) &= \alpha + \lambda_i + \chi_t + \beta_1 \\ E\left(\Delta y_{i,t} \mid e^{u}_{i,t-j} = 0,\ TB_{i,t} = 0, EB_{i,t} = 0\right) &= \alpha + \lambda_i + \chi_t \\ IR(t, 0, TB) &= \beta_1 \end{aligned}$$

- Panel restrictions imply that IRs are not country-specific.

The PS model

- The PS Econometric Model:

$$\begin{aligned}\Delta y_{i,t} &= \alpha + \lambda_i + \chi_t + \sum_{j=0}^{4}\left(\psi_{j+1}^{TB} * TB_{i,t-j} + \beta_{j+1} e_{i,t-j}^{u} * TB_{i,t-j}\right) \\ &+ \sum_{j=0}^{4}\left(\psi_{j+1}^{EB} * EB_{i,t-j} + \gamma_{j+1} e_{i,t-j}^{u} * EB_{i,t-j}\right) + u_{i,t}\end{aligned}$$

  so they allow for a (regime but not country specific shift in the intercept when policies are active).
- Then they state **the estimates of the coefficients β_j and γ_j provide an immediate estimate of the impulse response of the variable of interest to the TB and EB shocks.**
- Think again, for the issue of simplicity, of the contemporaneous effect of a TB- based adjustment: this would be estimated by PS at β_1.
- **β_1 is different from what it would be obtained by applying the standard approach to measure the effect of the policy intervention.**
- Given the PS model, we have:

$$\begin{aligned}E\left(\Delta y_{i,t} \,\middle|\, e_{i,t-j}^{u} = 1,\ TB_{i,t} = 1, EB_{i,t} = 0\right) &= \alpha + \lambda_i + \chi_t + \psi_1^{TB} + \beta_1 \\ E\left(\Delta y_{i,t} \,\middle|\, e_{i,t-j}^{u} = 0,\ TB_{i,t} = 0, EB_{i,t} = 0\right) &= \alpha + \lambda_i + \chi_t \\ IR(t, 0, TB) &= \psi_1^{TB} + \beta_1\end{aligned}$$

Which Reality?

Table 1 here uses the specification in PS:

$$\Delta y_{i,t} = \alpha + \lambda_i + \chi_t + \sum_{j=0}^{4} \left(\psi_{j+1}^{TB} * TB_{i,t-j} + \beta_{j+1} e_{i,t-j}^{u} * TB_{i,t-j} \right)$$
$$+ \sum_{j=0}^{4} \left(\psi_{j+1}^{EB} * EB_{i,t-j} + \gamma_{j+1} e_{i,t-j}^{u} * EB_{i,t-j} \right) + u_{i,t}$$

to report both their "impulse responses"

$$IR(t,k,TB)^{PS} = \sum_{j=0}^{k} \beta_{j+1} \qquad k = 0,1,2,3,4$$
$$IR(t,k,EB)^{PS} = \sum_{j=0}^{k} \gamma_{j+1} \qquad k = 0,1,2,3,4$$

and the impulse responses obtained by using the standard approach on their model specification:

$$IR(t,k,TB)^{AFG} = \sum_{j=0}^{k} (\beta_{j+1} + \psi_{j+1}^{TB}) \qquad k = 0,1,2,3,4$$
$$IR(t,k,EB)^{AFG} = \sum_{j=0}^{k} (\gamma_{j+1} + \psi_{j+1}^{EB}) \qquad k = 0,1,2,3,4$$

Which Reality?

IR	EB IRF PS	TB IRF PS	EB IRF ST	TB IRF ST
GDP Responses to Fiscal Adjustments				
h = 0	0.14	-1.04	-0.34	-0.74
	(0.29)	(0.00)		
h = 1	-0.19	-1.66	-0.37	-1.59
	(0.22)	(0.00)		
h = 2	-0.06	-1.60	-0.08	-1.98
	(0.72)	(0.01)		
h = 3	0.29	-1.70	0.18	-2.23
	(0.16)	(0.03)		
h = 4	0.31	-1.62	0.11	-2.11
	(0.27)	(0.11)		
Gov Purchases Responses to Fiscal Adjustments				
h = 0	-0.11	-0.41	-0.38	-0.28
	(0.03)	(0.00)		
h = 1	-0.13	-0.71	-0.53	-0.52
	(0.05)	(0.00)		
h = 2	-0.08	-0.64	-0.55	-0.58
	(0.33)	(0.00)		
h = 3	0.03	-0.76	-0.56	-0.64
	(0.73)	(0.00)		
h = 4	0.12	-0.88	-0.52	-0.74
	(0.19)	(0.00)		
Gov Revenues Responses to Fiscal Adjustments				
h = 0	0.38	-0.01	0.14	0.51
	(0.00)	(0.95)		
h = 1	0.20	-0.45	0.21	0.2
	(0.08)	(0.02)		
h = 2	0.17	-0.48	0.34	-0.04
	(0.20)	(0.08)		
h = 3	0.33	-0.36	0.51	0.07
	(0.09)	(0.33)		
h = 4	0.35	-0.34	0.51	0.07
	(0.02)	(0.46)		

Which Reality?

- The Table above shows that computing the impulse responses in the standard way strengthens the similarity between the output responses to consolidations plans in PS and AFG, while the divergence between "reality" and the classification of episodes in the narrative datasets is remarkably weakened.
- The ψ_{j+1} estimated in the GDP equation are very small and not significantly different from zero.
- Instead the ψ_{j+1} in the E and G equations are significantly different from zero, capturing the downard shift in E during EB adjustments and the upward shift in T during TB adjustments.
- Omitting these effects from the IR biased creates a "virtual reality".

Some reflections on the PS specification

There are three relevant aspects of the PS specification:

- collinearity,
- truncation,
- the scale problem in measuring the response of the policy variables to the policy interventions.

Collinearity

Take the PS specification as a typical panel-data specification. As the specification, correctly, contains a fixed effect for the constant one is left to wonder why the same logic is not applied to the regime specific constant. In this case the specification would be:

$$\begin{aligned}\Delta y_{i,t} &= \alpha + \lambda_i + \chi_t + \sum_{j=0}^{4} \left(\psi_{i,j+1}^{TB} * TB_{i,t-j} + \beta_{j+1} e_{i,t-j}^{u} * TB_{i,t-j} \right) \\ &+ \sum_{j=0}^{4} \left(\psi_{i,j+1}^{EB} * EB_{i,t-j} + \gamma_{j+1} e_{i,t-j}^{u} * EB_{i,t-j} \right) + u_{i,t}\end{aligned}$$

- It is clear that when $\psi_{i,j+1}^{TB} and \psi_{i,j+1}^{EB}$ are estimated instead of $\psi_{j+1}^{TB} and \psi_{j+1}^{EB}$ the specification will contain a measure of the exogenous adjustments $e_{i,t}^{u}$ together with an indicator that takes the value of 1 when $e_{i,t}^{u}$¿0 and zero otherwise.
- The specifications therefore includes the measure for the treatment twice: with a time-varying dose and a constant dose.
- This generate a collinearity problem, notice that if the doses of the treatment were constant over time there would be perfect collinearity. Only the time variation in the fiscal adjustments used in the estimation allows to pin down the estimates.

Truncation

- (Perotti and Sala, 2024) introduce the additional terms in their model as they try to address two (related) potential problems in evaluating the output effect of positive (restrictive) fiscal adjustments identified narratively as exogenous because they are "deficit-driven" and not affected by output fluctuations:
  - truncation, which occurs when a variables takes values different from zero when it is recorded as a zero in the database ,
  - omitted variables, which occurs when there are other exogenous fiscal adjustments that have been implemented alongwith the "deficit-driven" that are not included in the model and that are correlated with the deficit-driven adjustments.
- (Alesina et al., 2019)(AFG) argue that
  - there is no truncation problem in deficit driven fiscal adjustments because the presence of this problem would require "fiscal expansions motivated by too low deficits" and these do not exist.
  - There are indeed other possibly exogenous fiscal adjustments motivated by long-run output growth, but these are orthogonal to the deficit-driven fiscal adjustments.

The argument by AFG was motivated by the analysis of the narrative fiscal adjustment built by (Romer and Romer, 2010)

No Truncation

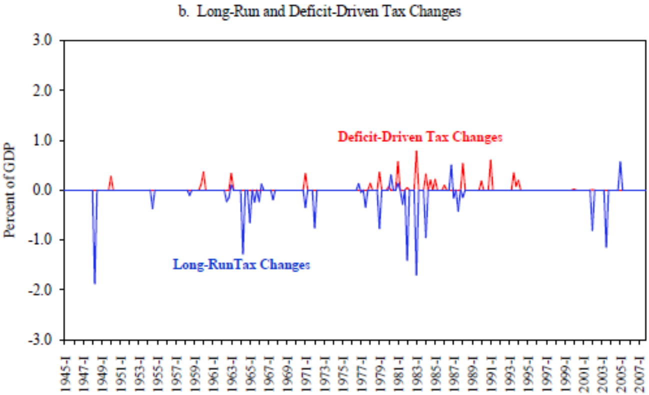

Figure: Narrative Exogenous Tax Changes. Source: (Romer and Romer, 2010)

No Truncation

The Figure illustrates that

- Deficit-Driven Tax changes are all positive, so there is no truncation at the zeros as "fiscal expansion motivated by too low deficits" do not exist.
- There is no significant correlation between Long-Run Tax Changes and Deficit-Driven Tax Changes.
- There would be a truncation problems in the case deficit-driven fiscal adjustments were used a a proxy for all fiscal adjustments but AFG are very clear in stating that their objective is to measure the effect of exogenous deficit-driven stabilization

Scale Problems

- AFG never use the MA representation to assess the effect of the exogenous adjustments on the fiscal variables.
- They instead do so in a panel VAR that, in its most parsimonious specification, includes the growth rate of per capita output ($\Delta y_{i,t}$) as the only macroeconomic variable, the change of tax revenues as a fraction of GDP ($\Delta \tau_{i,t}$) and that of primary government spending, also as a fraction of GDP ($\Delta g_{i,t}$) as the two fiscal variables.
- The estimated coefficients points toward shares of .79, .21 in revenes and expenditure in TB adjustments and a .39, .61 shares in EB adjustments.

Scale Problems

- PS choose to use a MA approach to measure the responses of the (log) levels of expenditure and revenues to our adjustments (expressed as percentage of the GDP). I believe that comparing the PS results with ours is difficult for a number of reasons
- the different and not comparable approach to measuring the effect of Fiscal adjustments used by AFG and PS
- the different specifications: MA and VAR
- the different scales of the policy variables. Here my argument is very simple: a correction of 1 per cent in the revenue (or expenditure) to GDP ratio at the end of the sample has not the same implications for the volumes of expenditure and revenues with a correction of 1 per cent in in the revenue (or expenditure) to GDP ratio at the beginning of the sample.

Bibliography I

Alesina, Alberto, Francesco Giavazzi, and Carlo Favero (2019) "Austerity: When It Works and When It Doesn't."

Angrist, Joshua D, Òscar Jordà, and Guido M Kuersteiner (2018) "Semiparametric estimates of monetary policy effects: string theory revisited," *Journal of Business & Economic Statistics*, 36 (3), 371–387.

Perotti, Roberto and Luca Sala (2024) "Fiscal Consolidations:Annoncements and Reality," discussion papers, Bocconi University.

Romer, Christina D and David H Romer (2010) "The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks," *American economic review*, 100 (3), 763–801.