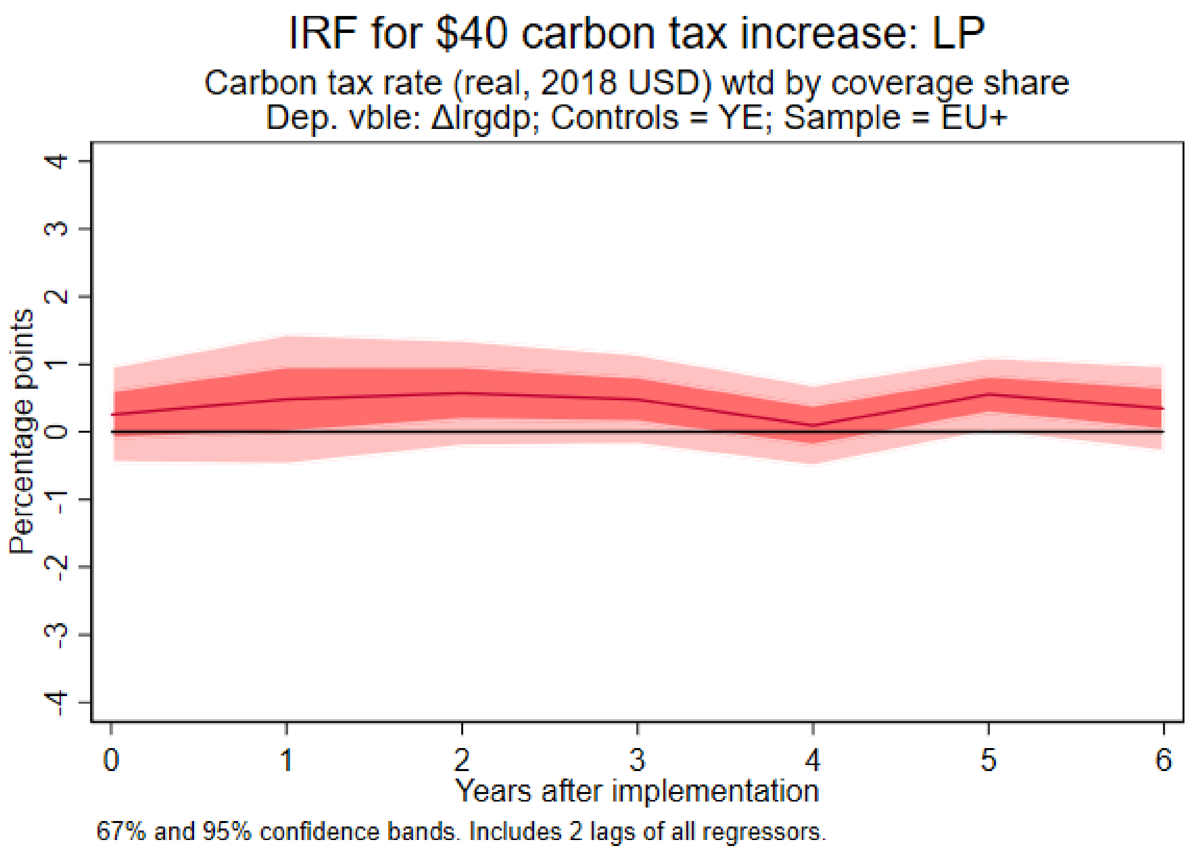
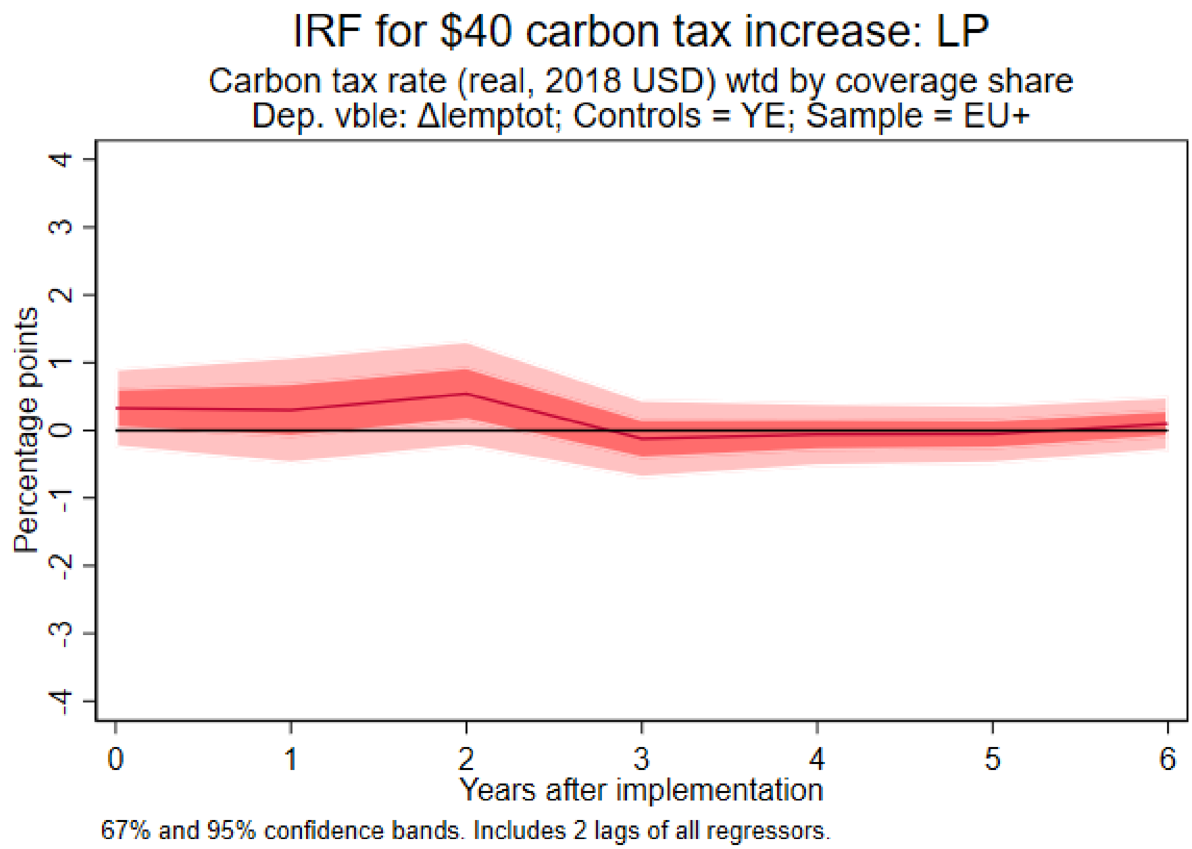
January 12, 2020

Measuring the Macroeconomic Impact of Carbon Taxes: Supplement
Gilbert Metcalf and James H. Stock

Supplemental Figures

Below we provide plots, with 67% and 95% confidence bands, of the impulse response functions estimated by local projections for the full EU+ sample, presented as rows 2 and 5 in Table 2 in the paper, for horizons of 0, 1,…, 6 years. The point estimates and their standard errors are computed in the supplemental code.





Calculation of IRFs

We are interested in estimating the effect on GDP growth in the h^{th} year after an increase in the tax by $40. The impact effect is the effect in year 0 (the year of the increase). We consider a carbon tax hike that increases from $$\tau_0$ to $$(\tau_0 + 40)$/ton in year 0 and stays at $40/ton. In our linear model this estimate does not depend on τ_0 so for simplicity we set $\tau_0 = 0$. We consider first the distributed lag (DL) estimate, then the local projection (LP) estimate, which is somewhat more complicated. This discussion focuses on the technicalities of lag accounting and computing covariance matrices and assumes identification conditions hold.

Let $y_t = \ln(GDP_t)$ or $\ln(Total\ Employment_t)$ in year t and let x_t = carbon tax rate in year t.

DL

Ignore the intercept and control variables and consider the DL regression,

$$\Delta y_t = \sum_{j=0}^{p} \beta_j x_{t-j} + u_t \,. \tag{1}$$

The effect of a carbon tax path of $x_s = 0$, $s < 0$ and $x_s = 40$, $s \geq 0$, compared to no carbon tax, on GDP growth in year h is,

$$E\left(\Delta y_h \mid x_h = 40, x_{h-1} = 40, ..., x_0 = 40, x_{-s} = 0, s < 0\right) - E\left(\Delta y_h \mid x_h = 0, x_{h-1} = 0, ..., x_0 = 0, x_{-s} = 0, s < 0\right)$$
$$= 40\beta_0 + 40\beta_1 + ... + 40\beta_h = 40\sum_{j=0}^{h}\beta_j$$

The cumulative sum of the β coefficients in (1) and its covariance matrix are conveniently computed directly as the regression coefficients in a rewritten version of (1):

$$\Delta y_t = \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + ... + \gamma_{p-1}\Delta x_{t-p+1} + \gamma_p x_{t-p} + u_t \, , \tag{2}$$

where $\gamma_h = \sum_{j=0}^{h}\beta_j$ (see Stock and Watson (2019, Eq. (16.7)).

These equations are expressed as time series relations, in the paper they are implemented using the panel of countries with fixed effects. Because the error term and the regressor plausibly are correlated, standard errors are clustered by country.

LP

The LP regression is,

$$\Delta y_{t+h} = \Theta_{yx,h} x_t + \phi_{yx}^h(\mathrm{L}) x_{t-1} + \phi_{yy}^h(\mathrm{L})\Delta y_{t-1} + u_t \, , \tag{3}$$

where again we ignore the intercept (fixed effects) and control variables and suppress the *i* subscript over countries. By the population counterpart of the Frisch-Waugh theorem, we can write (3) as,

$$\Delta y_{t+h}^{\perp} = \Theta_{yx,h}\eta_t + u_t \, , \tag{4}$$

where $\Delta y_{t+h}^{\perp} = \Delta y_{t+h} - \mathrm{Proj}\left(\Delta y_{t+h}^{\perp} \mid x_{t-1}, x_{t-2}, ..., y_{t-1}, y_{t-2}\right)$ and $\eta = x_t - \mathrm{Proj}\left(x_t \mid x_{t-1}, x_{t-2}, ..., y_{t-1}, y_{t-2}\right)$. Under the identifying assumption that the tax rate innovation η_t is uncorrelated with other shocks, then $\left\{\Theta_{yx,h}\right\}, h = 0,1,2,...$ is the impulse response function from x shocks to Δy.

We are interested not in the impulse response of Δy to a shock to x, but in the response of Δy to a change in the path of x from no tax to a \$40 tax imposed in year 0. To compute the response of Δy to this one-time increase of \$40, we use the device of Sims (1986) to compute the shocks necessary for x to follow the specified path. The Sims device is normally implemented for VARs so we adapt it to LP. Using the same notation as above, the response of x_t to a sequence of shocks to x, $\{\varepsilon_s\}$, is $x_t = \Theta_{xx}(\mathrm{L})\varepsilon_t$. We estimate the IRF $\Theta_{xx}(\mathrm{L})$ by LP using (3), except however that the dependent variable is x_{t+h} and the corresponding coefficient on x_t is $\Theta_{xx,h}$. Note that this calculation automatically imposes the unit effect normalization $\Theta_{xx,h} = 1$ (e.g. Stock and Watson (2018)). With this normalization, for a given specified path of x, the requisite shocks can be computed by the recursion, $\varepsilon_h^* = x_h - \sum_{j=1}^{h}\Theta_{xx,j}\varepsilon_{h-j}$, where $\varepsilon_s = 0$ for $s < 0$.

These shock and impulse response calculations are conveniently written as matrix equations. Let IRF_h denote the effect on Δy_h of a \$40 carbon tax imposed at date 0. Then,

$$IRF = A\varepsilon^* \text{ and} \tag{5}$$
$$\varepsilon^* = B^{-1}x \tag{6}$$

where ε^* is the sequence of shocks that delivers the desired carbon tax path rate x and

$$IRF = \begin{pmatrix} IRF_0 \\ IRF_1 \\ \vdots \\ IRF_h \end{pmatrix},\ x = \begin{pmatrix} 40 \\ 40 \\ \vdots \\ 40 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \Theta_{xx,1} & 1 & 0 & 0 \\ \vdots & \cdots & \ddots & \vdots \\ \Theta_{xx,h} & \Theta_{xx,h-1} & \cdots & 1 \end{pmatrix}, \text{ and } A = \begin{pmatrix} \Theta_{yx,0} & 0 & 0 & 0 \\ \Theta_{yx,1} & \Theta_{yx,0} & 0 & 0 \\ \vdots & \cdots & \ddots & \vdots \\ \Theta_{yx,h} & \Theta_{yx,h-1} & \cdots & \Theta_{yx,0} \end{pmatrix}$$

where the expression for B uses the unit effect normalization. Equation (6) follows from $x = B\varepsilon^*$.

To compute standard errors, note that equation (5) can be rewritten,

$$IRF = \begin{pmatrix} \varepsilon_0^* & 0 & 0 & 0 \\ \varepsilon_1^* & \varepsilon_0^* & 0 & 0 \\ \vdots & \cdots & \ddots & \vdots \\ \varepsilon_h^* & \varepsilon_{h-1}^* & \cdots & \varepsilon_0^* \end{pmatrix} \begin{pmatrix} \Theta_{yx,0} \\ \Theta_{yx,1} \\ \vdots \\ \Theta_{yx,h} \end{pmatrix} = \Lambda\Theta_{yx}, \text{ where } \Theta_{yx} = \begin{pmatrix} \Theta_{yx,0} \\ \Theta_{yx,1} \\ \vdots \\ \Theta_{yx,h} \end{pmatrix}. \tag{7}$$

Thus $IRF = \Lambda\hat{\Theta}_{yx}$ and the variance matrix of the IRF, treating ε^* as fixed, is $\hat{V}_{IRF} = \Lambda\hat{V}_{\hat{\Theta}_{yx}}$.

The variance matrix of $\hat{\Theta}_{yx}$, $\hat{V}_{\hat{\Theta}_{yx}}$, is not directly available from the h+1 separate LP regressions. If the regressions were computed over the same sample they could be computed using a Seemingly Unrelated Regression (SURE) procedure which would deliver $\hat{V}_{\hat{\Theta}_{yx}}$. However, we compute them over different samples so that all the data are used (the samples are longer at shorter horizons). Computation of the heteroskedasticity-robust covariance matrix over the different samples is a straightforward calculation that takes advantage of the fact that HAC SEs are not needed either in the time series or panel context (Montiel Olea and Plagborg-Møller (2019)).

Additional References

Stock, James H. and Mark W. Watson (2019), *Introduction to Econometrics*, 4th *Edition*. Pearson.

Stock, James H. and Mark W. Watson (2018), "Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments," *Economic Journal* 128, 917-948.