1. Some insights about IV specification

1.1. A note on prices

We have three prices: the (US) consumer price, say p_{ik}^c, the "import" or cif price (at the entrance of the US) p_{ik} and the fas price (at the export gate in the origin country) $\widetilde{p}_{ik}$, denoting i the origin country, k the product at the 8 or 5 digit level, and reasoning on a yearly basis, and $s(k)$ the 3-digit classification the product k belongs to. Further, if we denote τ_{ik}^d the duty tax rate paid when the good crosses the US border, then we have:

$$p_{ik}^c = (1 + \tau_{is(k)}^d)p_{ik}$$
$$with \qquad p_{ik} = \tau_{is(k)}\widetilde{p}_{ik} + t_{is(k)}$$

1.2. Specification of the first stage equation

The question we deal with, is the functional form of the first stage equation, where we instrument the fas price with duties as suggested by Referee 1. The idea is that firms might react to changes in duty tax rates which have nothing to do with transport costs changes. In this respect, considering the predicted part of the fas price related to tax duty is likely to solve potential endogeneity biases.

Denoting i the origin country, k the product at the 8/5 digit level, and reasoning on a yearly basis, if we assume that the fas price $\widetilde{p}_{ik}$ decomposes in two components, say $\bar{p}_{ik}$ the "firm-specific" price (related to its cost and pricing strategy) and τ_{ik}^d the tax duty (out of the firm's hands) according to :

$$\widetilde{p}_{ikt} = (1 + \tau_{is(k)t}^d)^\alpha \, (\bar{p}_{ikt})^\beta \tag{1}$$

with $s(k)$ the 3-digit classification (the disaggregation level for tax duties in the database) as a function of the HS8 (or 5 digit) product classification

1.2.1. First-stage equation in first difference

One option to specify the functional form of the first stage is to take the total differential around some reference point at time $t-1$, with Δ the difference operator ($\Delta\widetilde{p}_{ikt} = \widetilde{p}_{ikt} - \widetilde{p}_{ikt-1}$ and so on):

$$\Delta\widetilde{p}_{ikt} = \beta\bar{p}_{ikt-1}^{\beta-1}(1 + \tau_{is(k)t-1}^d)\Delta\bar{p}_{ikt} + \alpha\bar{p}_{ikt-1}^{\beta}(1 + \tau_{ikt-1}^d)^{\alpha-1}\Delta\tau_{is(k)t}^d$$

$$\Leftrightarrow \quad \frac{\Delta\widetilde{p}_{ikt}}{\widetilde{p}_{ikt-1}} = \beta\frac{\Delta\bar{p}_{ikt}}{\widetilde{p}_{ikt-1}}\frac{(1 + \tau_{is(k)t-1}^d)^\alpha\bar{p}_{ikt-1}^\beta}{\bar{p}_{ikt-1}^\beta(1 + \tau_{is(k)t-1}^d)^\alpha} + \alpha\frac{\Delta\tau_{is(k)t-1}^d}{1 + \tau_{is(k)t-1}^d}\frac{(1 + \tau_{is(k)t-1}^d)^\alpha\bar{p}_{ikt-1}^\beta}{\bar{p}_{ikt-1}^\beta(1 + \tau_{is(k)t-1}^d)^\alpha}$$

leading to:

$$\frac{\Delta\widetilde{p}_{ikt}}{\widetilde{p}_{ikt-1}} = \beta\frac{\Delta\bar{p}_{ikt}}{\bar{p}_{ikt-1}} + \alpha\frac{\Delta\tau_{is(k)t}^d}{1 + \tau_{ikt-1}^d} \tag{2}$$

The intuition behind the Referee 1's endogeneity concerns, is that we need to eliminate from the fas price, any endogenous component that might me related to transport costs. To do so, the referee suggests that we instrument the export price by tariff rates. Put it in plain words, this suggests to run the first-stage regression based on Equation (3) according to:

$$\frac{\Delta\widetilde{p}_{ikt}}{\widetilde{p}_{ikt-1}} = \alpha\frac{\Delta\tau_{is(k)t}^d}{1 + \tau_{is(k)t-1}^d} + \gamma_i + \gamma_k + \epsilon_{ik}$$

Or equivalently, taking logs:

$$\Delta \log \widetilde{p}_{ikt} = \alpha \frac{\Delta \tau^d_{is(k)t}}{1 + \tau^d_{is(k)t-1}} + \gamma_i + \gamma_k + \epsilon_{ikt}$$

with the LHS being the growth rate of the fas price (between t and $t-1$), the first term in the RHS the change in duty tax rates, the second and third terms fixed effects to capture changes in the "firm-specific price" $\bar{p}_{ik}$, ϵ_{ik} being the residual. Notice though that the structure of fixed effects should remain consistent between the first and the second stages. This implies to rather consider the following functional form of the first-stage equation:

$$\Delta \log \widetilde{p}_{ikt} = \alpha \frac{\Delta \tau^d_{is(k)t}}{1 + \tau^d_{is(k)t-1}} + \gamma_i + \gamma_s + \epsilon_{ikt} \tag{3}$$

If this reasoning is correct,

- we should take as predicted value only the predicted price WITHOUT the fixed effects:

$$\frac{\widehat{\Delta \widetilde{p}_{ikt}}}{\widetilde{p}_{ikt-1}} = \widehat{\alpha} \frac{\Delta \tau^d_{is(k)t}}{1 + \tau^d_{is(k)t-1}}$$

- and we should expect a value of the coefficient $\widehat{\alpha}$ between $\Delta \log \widetilde{p}_{ikt}$ and $\frac{\Delta \tau^d_{is(k)t}}{1 + \tau^d_{is(k)t-1}}$ between -1 and 0, depending on the degree of "pricing-to-market" of firms.

  - $\alpha = 0$ corresponds to the case where the firm does not adjust its fas price to the change in tax duty, that would cancel out the influence of the tax duty change on the price paid by the US consumers; this rather corresponds to small firms which do not have enough market power to manipulate their prices following changes in international competition;

  - $\alpha < 0$ corresponds to the degree of "pricing-to-market" as the firm offsets the impact of the tax change of the price paid by the final consumer by adjusting her producer price in the opposite direction of the tax change,

  - with $\alpha = -1$ being the extreme case of "full PTM" where the firm fully compensates the tax duty change. As shown by Berman, Martin, Mayer (REStats 2012), this is more likely larger firms.

Once this is done, we rebuilt the instrumented fas price at time t:

$$\widehat{\widetilde{p}}_{ikt} = \widetilde{p}_{ikt-1} + \frac{\widehat{\Delta \widetilde{p}_{ikt}}}{\widetilde{p}_{ikt-1}}$$

with $\widehat{\widetilde{p}}_{ikt}$ the instrumented fas price at time t, for product k from country i; $\widetilde{p}_{ikt-1}$ the observed lagged fas price for the same i, k couple; and $\frac{\widehat{\Delta \widetilde{p}_{ikt}}}{\widetilde{p}_{ikt-1}}$ the predicted growth rate of the fas price based on duty changes.

1.3. First-stage equation in level

Alternatively, one can specify the first-stage equation in log levels. Taking Equation (1) into log:

$$\log \widetilde{p}_{ikt} = \alpha \log(1 + \tau^d_{is(k)t}) + \beta\,(\bar{p}_{ikt})$$

Leading to the following functional form of the first-stage equation (to be run on a yearly basis, by transport mode)

$$\log \widetilde{p}_{ikt} = \alpha \log(1 + \tau^d_{is(k)t}) + \gamma_i + \gamma_s + \epsilon_{ikt} \tag{4}$$

If we rather adopt a panel approach, the first-stage equation is modified by adding a year fixed effect, according to:

$$\log \widetilde{p}_{ikt} = \alpha \log(1 + \tau^d_{is(k)t}) + \gamma_i + \gamma_s + \gamma_t + \epsilon_{ikt} \tag{5}$$

1.3.1. Introducing price inertia

As suggested by the referee, one cannot exclude price inertia, such that current prices depend on past prices; the "inertia" part of the fas price being exogenous to current transport costs changes, it could be also considered in the predicted price at the first stage. One can introduce price inertia in Equation (1) by considering that the price component depends on the past price $\widetilde{p}_{ikt-1}$ and the "firm-specific" component, according to:

$$\widetilde{p}_{ikt} = (1 + \tau^d_{is(k)t})^\alpha ([\bar{p}_{ikt-1}(\widetilde{p}_{ikt-1}, \omega_{ikt})]^\beta$$

Specifically, using a Cobb-Douglas specification

$$\widetilde{p}_{ikt} = \left[1 + \tau^d_{is(k)t}\right]^\alpha \widetilde{p}^{\beta_1}_{ikt-1} \omega^{\beta_2}_{ikt} \tag{6}$$

Taking Equation (6) into logs:

$$\log \widetilde{p}_{ikt} = \alpha \log\left(1 + \tau^d_{is(k)t}\right) + \beta_1 \log \widetilde{p}_{ikt-1} + \beta_2 \log \omega_{ikt}$$

leading to the following specification of the first-stage equation (in cross section and in panel respectively):

$$\log \widetilde{p}_{ikt} = \alpha \log\left(1 + \tau^d_{is(k)t}\right) + \beta_1 \log \widetilde{p}_{ikt-1} + \gamma_i + \gamma_s + \epsilon_{ikt} \tag{7}$$

$$\log \widetilde{p}_{ikt} = \alpha \log\left(1 + \tau^d_{is(k)t}\right) + \beta_1 \log \widetilde{p}_{ikt-1} + \gamma_i + \gamma_s + \gamma_t + \epsilon_{ikt} \tag{8}$$

1.3.2. Obtaining the instrumented fas price

As last step, we take as instrument the part of the fas price that is predicted by the sum of the tax duty and the past prices components (if considered):

$$\widehat{\log \widetilde{p}}_{ikt} = \widehat{\alpha} \log\left(1 + \tau^d_{is(k)t}\right) + \widehat{\beta_1}\,\log \widetilde{p}_{ikt-1}$$

with $\widehat{\alpha}$ and possibly $\widehat{\beta_1}$ the estimated coefficients at the first stage (depending on the specification considered, with or without price inertia, in panel or in cross-section).

In contrast with the differential approach (Section 1.2.1), this specification has the advantage to directly yield the instrumented price in log, from which we can deduce the fas price in level (which intervenes in the second-stage equation) by taking the exponential.

1.4. To sum up

What would be my "preferred" first-stage equation: Equation (7) because i) in log level, so do not need to rebuild the level from the first difference + consistent with our

second stage in log level as well; ii) in cross-section, run on a yearly basis, here again consistent with our second-stage equation. But need to be pragmatic as well...