*Journal of Global Economic Analysis*, Volume 09 (2024), No. 1, pp. 62-76.

# Comments on "Estimation of the Value-Added/Intermediate Input Substitution Consistent with the GTAP Data"

BY ALEXANDRE GOHIN[a] AND OBAFEMI PHILIPPE KOUTCHADE[b]

*The first part of the paper written by Ivanic et al. (2023) (hereafter IBN) offers new estimates of the substitution elasticities between inputs for many industries. IBN develop an original econometric framework that solely relies on the GTAP (Global Trade Analysis Project) databases, which lack data on input prices. IBN find, first, that the short run elasticities are often statistically different from zero and of the correct sign; second, that the long-run elasticities are larger than their short-run counterparts. Our comment identifies three concerns with their econometric procedures and results. First, IBN do not acknowledge the Constant Return to Scale (CRS) assumption of the Constant Elasticity of Substitution (CES) function. Second, IBN's statistical inference fails to correct the p-value problems associated with large samples. Third, the dynamic specification developed by IBN, where decisions are a function of price changes and not price levels, lacks theoretical justifications. We propose simple remedies to these three issues and, in some cases, find that many elasticity estimates are no longer statistically different from zero or of the correct sign. Moreover, we do not find different levels of significance between short- and long-run elasticities.*

JEL codes: C13, D21, Q10

Keywords: Substitution elasticities, Econometrics, Crops

## 1. Introduction

Motivated by the short and long run impacts of the European Green Deal that intends to constraint farm input uses, the first part of the paper written by Ivanic et al (2023) (hereafter IBN) offers new estimates of the substitution elasticities between inputs for many industries. A common criticism of much of the general equilibrium modelling literature is the absence of econometrically estimated parameters. The recent literature referred to as the 'New Quantitative Trade Models' (e.g. Costinot et al., 2016) seeks to overcome this challenge, but the cost has been working with highly stylized models that contain insufficient commodity

[a] INRAE, UMR SMART, Institut Agro and CEPII, France (alexandre.gohin@inrae.fr).
[b] INRAE, UMR SMART, Institut Agro, France (obafemi-philippe.koutchade@inrae.fr).

detail for most policy analysis. IBN seek to address the parameterization problem by utilizing the time series/cross section data now available to the Global Trade Analysis Project (GTAP) community. The major challenge confronting the authors is the absence of price observations. Following the earlier consumer demand estimation research by Reimer and Hertel (2004), IBN seek to estimate the supply side parameters by exploiting the tax data in the GTAP Data Base (Aguiar et al., 2023). They focus on agricultural producers' response to prices in the form of constant elasticities of substitution, as these are highly relevant for agricultural and environmental policies such as the recent EU Green Deal. IBN find first that the short run elasticities are often statistically different from zero and of the correct sign, second that the long run elasticities are larger than their short run counterparts. In the second part of the paper, IBN show the significant effects of these elasticity estimates on the market impacts of the EU Green Deal proposal to constraint input uses for crop activities.

We commend IBN efforts to estimate these parameters. However, we take issue with several methodological features of their work. In our approach to addressing these issues, we derive econometric estimates of substitution elasticities that are significantly different to those by IBN in some cases.[1] First, both in their theoretical derivation and econometric implementation, IBN do not acknowledge the Constant Return to Scale (CRS) assumption of their specified Constant Elasticity of Substitution (CES) functions. We propose a simple remedy that minimizes the change made to their estimated equations. We find that imposing CRS changing the values of estimate estimates but overall has a modest impact on IBN's findings. Second, IBN's statistical inference fails to correct the p-value problems associated with large samples (Leamer, 1978). We solve this issue by relying on the Bayesian School with the introduction of external plausible information. We find that many elasticity estimates are no longer statistically different from zero or of the correct sign. Third, the dynamic specification developed by IBN, where decisions are functions of price changes and not price levels, lacks theoretical justifications. We slightly extend their dynamic specification to make it consistent with the large literature on expectations and dynamic behavior (Nerlove and Bessler, 2001). In the augmented framework, we do not find different levels of significance between short and long run elasticities (as IBN do).

[1] A minor issue comes from the format of Tables 2 to 4 (due to a missing sort instruction in their R code), where rows are not correctly labelled. For instance, with Table 2, the results for the wheat sector (wht) are reported in the row attached to business services (obs). While Table 2 suggests that for 7 out of 8 crop sectors the substitution elasticities are statistically significantly different from zero, the correct number is 5 (see below).

## 2. On the omission of the constant returns to scale

IBN start from the microeconomic specification of the GTAP model implemented in the GEMPACK software. They thus start with the linearized representation of behavioral equations and then integrate them to obtain level equations for their econometric analysis. In our presentation below, we start from the level representation, clarifying the fixed effects introduced in the econometric analysis. In this static computable general equilibrium (CGE) model, producers are assumed to minimize their production costs by optimally choosing their inputs subject to market prices, taxes on these inputs, technological constraints. Production opportunities are represented by nested functions that are homogeneous of degree one. In the lower nest, the primary factors of production (labor, capital and land) are combined in a value-added aggregate. In the upper nest, the value-added aggregate and the variable inputs are combined in another CES function that defines the level of production. The nested production structure with linearly homogenous CES functions allows solving the cost minimization program in two steps: first the optimal levels of the primary factors of production for a given value-added aggregate, second the optimal levels of the variable inputs and of the value-added aggregate for a given production level. The paper by IBN focuses on estimating the parameters of the upper level CES function and accordingly focus on the optimal decisions of the second step. Adopting IBN's notation, the second step for a given sector (j) in region (r) is:

$$min_{QF_{i,j,r}} C_{j,r} = \sum_i WF_{i,r}.TF_{i,j,r}.QF_{i,j,r} \tag{1}$$

Subject to

$$QO_{j,r} = \delta o_{j,r}.\left(\sum_i \delta f_{i,j,r}.QF_{i,j,r}^{-\rho_{j,r}}\right)^{-1/\rho_{j,r}} \tag{2}$$

with $C$ is the cost function, $WF$ the market price of input $i$ (we include the value-added aggregate in the set of $i$), $TF$ one plus the ad valorem input tax paid by industry, $QF$ the input quantity, $QO$ the output quantity, $\delta o$ the efficiency parameter, $\delta f$ the distribution parameters and $\rho$ the substitution parameter of the CES function.

Solving this program gives the Hicksian input demands:

$$QF_{i,j,r} = \frac{QO_{j,r}}{\delta o_{j,r}}.\left(\frac{\delta f_{i,j,r}}{WF_{i,r}.TF_{i,j,r}}\right)^{\sigma_{j,r}}\left(\sum_i \delta f_{i,j,r}^{\sigma_{j,r}}.\left(WF_{i,r}.TF_{i,j,r}\right)^{1-\sigma_{j,r}}\right)^{1/\rho_{j,r}} \tag{3}$$

with $\sigma_{j,r} = \frac{1}{1+\rho_{j,r}}$. The cost function is then:

$$C_{j,r} = \frac{QO_{j,r}}{\delta o_{j,r}} . \left( \sum_i \delta f_{i,j,r}^{\sigma_{j,r}} . \left( WF_{i,r} . TF_{i,j,r} \right)^{1-\sigma_{j,r}} \right)^{1/(1-\sigma_{j,r})} \tag{4}$$

Due to the CRS assumption, this cost function depends linearly on the production level. The output level is implicitly determined by the following profit maximization problem:

$$max_{QO_{j,r}} \frac{PM_{j,r}}{TO_{j,r}} QO_{j,r} - C_{j,r} \tag{5}$$

with $PM$ is the market price of output, $TO$ capturing the output tax. Solving this problem leads to:

$$\frac{PM_{j,r}}{TO_{j,r}} = \frac{1}{\delta o_{j,r}} \left( \sum_i \delta f_{i,j,r}^{\sigma_{j,r}} . \left( WF_{i,r} . TF_{i,j,r} \right)^{1-\sigma_{j,r}} \right)^{1/(1-\sigma_{j,r})} \tag{6}$$

Substituting equation (6) into the input demand equation (3) gives:

$$QF_{i,j,r} = QO_{j,r} \delta o_{j,r}^{\sigma_{j,r}-1} . \left( \frac{\delta f_{i,j,r} . \frac{PM_{j,r}}{TO_{j,r}}}{WF_{i,r} . TF_{i,j,r}} \right)^{\sigma_{j,r}} \tag{7}$$

The value of input demand at the agent prices is then:

$$\begin{aligned} VF_{i,j,r} &= WF_{i,r} . TF_{i,j,r} . QF_{i,j,r} \\ &= VO_{j,r} . \delta o_{j,r}^{\sigma_{j,r}-1} . \delta f_{i,j,r}^{\sigma_{j,r}-1} . \left( \frac{\frac{PM_{j,r}}{TO_{j,r}}}{WF_{i,r} . TF_{i,j,r}} \right)^{\sigma_{j,r}-1} \end{aligned} \tag{8}$$

Assuming no productivity changes, total differentiation of this equation (8) leads to:

$$vf_{i,j,r} = vo_{j,r} + \left(1 - \sigma_{j,r}\right) . \left( wf_{i,r} + tf_{i,j,r} + to_{j,r} - pm_{j,r} \right) \tag{9}$$

where $vf$ stands for the percent change of the value of input demand and similarly for the other notations. This is similar to equation (4) of IBN.

One crucial issue is that input prices are not available in the GTAP database. IBN solve this issue by assuming that input and output market prices do not change when taxes change. This leads to their equation (6):

$$vf_{i,j,r} = vo_{j,r} + \left(1 - \sigma_{j,r}\right).\left(tf_{i,j,r} + to_{j,r}\right) \tag{10}$$

This exogeneity assumption of market prices problem is however theoretically problematic because nothing ensures that the zero-profit condition is satisfied. As detailed above, the input equation has been obtained thanks to the condition (equation 6) that the output market price equals the marginal cost. Full differentiation of this condition leads to (Gohin and Hertel, 2003):

$$pm_{j,r} - to_{j,r} = \sum_i sf_{i,j,r}.\left(wf_{i,r} + tf_{i,j,r}\right) \tag{11}$$

where $sf$ measures the initial share of input values in production costs. The exogeneity assumption of market prices thus implies that input and output taxes are not independent. They are linked by:

$$-to_{j,r} = \sum_i sf_{i,j,r}.tf_{i,j,r} \tag{12}$$

However, these taxes are assumed exogenous and thus independent in both the CGE model and the econometric procedures used by IBN. As IBN later discuss, the assumption of exogenous market prices is likely to hold, absent monopsony or monopoly market power. These situations are less likely for sectors with many agents like the farm and food sectors.

IBN do not directly estimate their input demand equations expressed in relative terms. They transform them in absolute terms, leading to our level equation above. In this equation, input prices are explicit. IBN get around the problem of unavailable regional input prices by introducing input and country fixed effects. This practice of fixed effects is standard in the trade literature estimating the Armington elasticities (Fontagné et al., 2022). Indeed, we now show that this is consistent if we assume the same production technologies in all countries and that the regional input prices differ from “world” input prices due to regional trade costs that we denote $TC_r$. That is, if we assume that $\delta f_{i,j,r} = \delta f_{i,j}$, $\sigma_{j,r} = \sigma_j$, $WF_{i,r} = W_i + TC_r$, then the value of input demand equation becomes:

$$VF_{i,j,r} = VO_{j,r}.\delta o_{j,r}^{\sigma_j-1}.\delta f_{i,j}^{\sigma_j-1}.\left(\frac{\frac{PM_{j,r}}{TO_{j,r}}}{WF_i.TC_r.TF_{i,j,r}}\right)^{\sigma_j-1} \tag{13}$$

This nonlinear equation becomes linear after log transforming both sides and introducing fixed effects:

$$\begin{aligned}\log(VF_{i,j,r}) - \log(VO_{j,r}) \\ &= (1-\sigma_j).\left(\log(TF_{i,j,r}) + \log(TO_{j,r})\right) \\ &+ F_{i,j} + G_{r,j}\end{aligned} \tag{14}$$

with $F_{i,j} = (\sigma_j - 1).log\left(\frac{\delta f_{i,j}}{WF_i}\right)$ and $G_{r,j} = (\sigma_j - 1)log\left(\delta o_{j,r}.\frac{PM_{j,r}}{TC_r}\right)$

This is similar to equation (10) of IBN. IBN then add error terms $\epsilon$ to account for omitted exogenous variables (and an additional time fixed effect discussed later on the dynamic specification):

$$\begin{aligned}\log(VF_{i,j,r}) - \log(VO_{j,r}) \\ &= (1-\sigma_j).\left(\log(TF_{i,j,r}) + \log(TO_{j,r})\right) \\ &+ F_{i,j} + G_{r,j} + \epsilon_{i,j,r}\end{aligned} \tag{15}$$

IBN estimate this equation (15) with the ordinary least squares method using all inputs and regions in the database. However, this estimation approach fails to recognize that the input demands are not independent in each region. For each region, summing over all inputs the left side of the level equation (13) gives the value of output, again due to the zero profit condition. This implies that the errors terms added to these equations are correlated (even after the log transformation):

$$\sum_i\left(\delta o_{j,r}^{\sigma_j-1}.\delta f_{i,j}^{\sigma_j-1}.\left(\frac{\frac{PM_{j,r}}{TO_{j,r}}}{WF_i.TC_r.TF_{i,j,r}}\right)^{\sigma_j-1}\right).e^{\varepsilon_{i,j,r}} = 1 \tag{16}$$

The equation-by-equation estimation procedure delivers consistent but not efficient parameter estimates. Several approaches are possible to deal with this issue, such as performing Seemingly Unrelated Regression (SUR) of the full demand system or estimating ratios of input demands. To be as close as possible to the econometric approach of IBN, we propose to estimate the following equation:

$$\log\left(VF_{i,j,r}\right) = \left(1-\sigma_j\right).\log\left(TF_{i,j,r}\right) + F_{i,j} + G_{r,j} + \epsilon_{i,j,r} \tag{17}$$

With the regional fixed effects now given by:

$$G_{r,j} = \left(\sigma_j - 1\right)log\left(\delta o_{j,r}.\frac{PM_{j,r}}{TC_r}\right) + \log\left(VO_{j,r}\right) + \left(1-\sigma_j\right).\log\left(TO_{j,r}\right)$$

Note that the estimates of the substitution elasticity obtained from equation (17) equal those from the following equation:

$$\log\left(WF_{i,r}QF_{i,j,r}\right) = -\sigma_j \log\left(TF_{i,j,r}\right) + F_{i,j} + G_{r,j} + \epsilon_{i,j,r} \tag{18}$$

Equation (18), which is equivalent to that derived by Fontagné et al. (2022), clarifies the absence of correlation among residuals, since the sum over all inputs of expenditures before tax is no longer equal to the production value. This equation can be estimated by the ordinary least squares method without loss of efficiency because their error terms are not correlated.

Table 1 reports static estimates for substitution elasticities among inputs for alternative estimation procedures. The first column reproduces the estimates of IBN. The second column of Table 1 display our estimates using equation (18). It appears that they do not differ to a great extent from those obtained by IBN. The largest differences are observed for the oilseed and other crops sectors, with a decrease of their estimate by 40 per cent. At the same time, the standard errors for those parameters increase, thus lowering their significance level. The number of elasticity estimates that are statistically different from zero remains the same (five out of eight crops).

**Table 1.** Static estimates of substitution elasticities among inputs

| | IBN (Eq15) | CRS (Eq18) | Key inputs (Eq15) | Key inputs (Eq18) |
|---|---|---|---|---|
| **Paddy rice** | | | | |
| Estimate | 0.556 | 0.584 | -0.098 | -0.115 |
| Standard error | 0.121 | 0.127 | 0.137 | 0.156 |
| Significance | 0.000 | 0.000 | 0.476 | 0.460 |
| **Wheat** | | | | |
| Estimate | 0.189 | 0.190 | 0.026 | 0.049 |
| Standard error | 0.079 | 0.083 | 0.106 | 0.117 |
| Significance | 0.016 | 0.022 | 0.804 | 0.674 |
| **Coarse grains** | | | | |
| Estimate | 0.101 | 0.131 | -0.314 | -0.274 |
| Standard error | 0.084 | 0.087 | 0.110 | 0.118 |
| Significance | 0.230 | 0.134 | 0.004 | 0.020 |
| **Vegetable and fruits** | | | | |
| Estimate | 0.164 | 0.145 | 0.224 | 0.220 |
| Standard error | 0.091 | 0.093 | 0.130 | 0.136 |
| Significance | 0.072 | 0.120 | 0.084 | 0.106 |
| **Oilseeds** | | | | |
| Estimate | 0.374 | 0.230 | 0.109 | 0.109 |
| Standard error | 0.088 | 0.095 | 0.103 | 0.112 |
| Significance | 0.000 | 0.016 | 0.290 | 0.330 |
| **Sugar crops** | | | | |
| Estimate | 0.074 | 0.043 | -0.352 | -0.480 |
| Standard error | 0.080 | 0.086 | 0.124 | 0.141 |
| Significance | 0.356 | 0.622 | 0.004 | 0.000 |
| **Plant-based fibers** | | | | |
| Estimate | 0.489 | 0.506 | 0.738 | 0.756 |
| Standard error | 0.054 | 0.057 | 0.065 | 0.072 |
| Significance | 0.000 | 0.000 | 0.000 | 0.000 |
| **Other crops** | | | | |
| Estimate | 0.454 | 0.276 | 0.025 | -0.124 |
| Standard error | 0.067 | 0.079 | 0.103 | 0.121 |
| Significance | 0.000 | 0.000 | 0.808 | 0.306 |

*Source:* Author calculations

### 3. On the p-values with large samples

IBN make use of the five reference years (2004, 2007, 2011, 2014 and 2017) represented in Version 11 of the GTAP Data Base. The database gathers economic flows for 141 countries and 65 industries. In particular, they include 65*65 input-output matrices measuring the values of input uses by each industry at agent and market prices (in each country for five reference years). This allows IBN to estimate their equations using a large number of data points. The exact numbers of observations and parameters depend on the industry as some inputs may not be used by some industries or regions. The degree of freedom averaged over the 8 crop industries is 21,391. This affects the significance results obtained by IBN (Leamer, 1978): 61 out of 65 elasticity estimates are reported in their Table 2 as statistically different from zero at the usual 5 per cent level of significance.[2]

One solution to the p-value issue in large samples is to restrict the data set. One possibility, mentioned by IBN, is to assume that substitution elasticities vary per group of countries and are not the same for all countries. The problem is that this will contradict the assumption made to obtain the estimated equation (15) as explained above. Another solution to reduce the size of the sample is to focus on certain key inputs. It is unlikely that all crop production requires some inputs (e.g. rice seed to make wheat is unlikely, similarly meat or dairy products for all crops). With this in mind, we focus on the agricultural literature that estimates the parameters of production functions with a limited number of inputs. For instance, Rosas et al. (2018) consider three inputs for cropping activities, which are fertilizers, an aggregate of pesticides and energy products and finally another aggregate of other purchased intermediate inputs (seeds, contract labor services, custom machine services, machine and building maintenance and repairs and irrigation). Guided by this approach, we estimate equations (15) and (18) on a limited number of inputs while excluding the other inputs: energy products (codes

[2]The so-called p-value problem associated with large samples is not new among statisticians but gains new importance in recent years with the availability of large database and computational power to manage them. Let's briefly explain this problem. Let $(Y_1, Y_2, \ldots, Y_n)$ be a sample of $n$ independent draws from a normal random variable with an unknown mean μ and a known variance of one. We are interested in obtaining the true value of μ. From this sample, we can compute the empirical average $\bar{Y}$ whose distribution is $\mathcal{N}(\mu, 1/n)$. We want to test if the true mean μ is equal to zero ($H_0: \mu = 0$). A two-sided test approach involves using the following test statistic $Z = \sqrt{n}(\bar{Y} - \mu)$ which follows $\mathcal{N}(0,1)$ , and leads us to reject $H_0$ if $abs(z)$ is greater than a critical value $c$ that is related to the risk of Type 1 error (usually denoted $\alpha$) and where $z$ is the realization of $Z$ at $H_0$. A large $n$ will produce a very large $abs(z)$ and will lead to the rejection of $H_0$ when this critical value $c$ (or $\alpha$) is not chosen as a function of n. In the classical approach with an assumed normal distribution, the risk of Type 1 error is usually fixed at 5 per cent ($\alpha = 0.05$) which implies $c = 1.96$ . In this case we are more likely to reject $H_0$ when n is very large.

ely, p_c), chemical (chm), relevant crop seeds and the value-added aggregate. These inputs represent the main expenditure for all crops, while the share of other inputs is lower than two per cent. The degree of freedom averaged over the 8 crops is now 2,688. To be consistent with the equation estimated by IBN, we first estimate equation (15) for the inputs noted above. The results of this estimation are reported in the third column of Table 1. Note that these results do not change if we include expenditures on materials (with codes ome, mvh) as another input.

It appears that the estimated results change dramatically. The estimated substitution elasticities are no longer always of the correct (positive) sign or not always significantly different from zero. We end up with only one industry (plant-based fibers) with the correct sign and a statistically significant estimate. The message does not change whether we apply these approaches to the original IBN specification (equation 15) or to our specification that takes returns to scale into account (equation 18).

**4. On the dynamic specification**

IBN argue that the reference specification is likely representing the degree to which industries may switch between inputs in the near term of about one year. This is not justified as their empirical specification only includes time fixed effects to control for exceptional events, not prices from successive years. In the absence of dynamic observations, we instead consider their estimates as prevailing in the steady state, that is when price expectations by producers are stabilized and input adjustments to any shock are completed.

IBN then assume a dynamic specification without clear theoretical justification. They assume that producers react to the changes of taxes across years, while still assuming no changes of market prices. Their estimated equation (12) is:

$$\begin{aligned}\log(VF_{i,j,r,t}) - \log(VO_{j,r,t}) \\ &= (1-\sigma_j).\left(\log(TF_{i,j,r,t}) + \log(TO_{j,r,t})\right. \\ &\quad \left. - \log(TF_{i,j,r,t-1}) - \log(TO_{j,r,t-1})\right) + F_{i,j} + G_{r,j}\end{aligned} \tag{19}$$

They do not refer to the large econometric literature which tries to identify the dynamic behavior of industries, for instance with their investment decisions (see e.g. Mundlak 2001). This literature introduces the possibility that producers face constraints to quickly adjust their capital, which leads to the fact that current investment decisions depend on past investment decisions. By focusing on the substitution between variable inputs and the value-added aggregate, IBN do not directly consider the investment issue.

The dynamic econometric literature also introduces the possibility that producers do not perfectly anticipate market prices. The price expectation issue is for instance related to the fact that there is a lag between all production decisions and the final decision of marketing the output. Many price expectations have been explored in the literature, using mostly public information such as lagged observed prices. A common specification for the expected price of output is (Nerlove, 1958):

$$PM^*_{j,r,t} = PM^*_{j,r,t-1} + \beta.\left(PM_{j,r,t-1} - PM^*_{j,r,t-1}\right) \tag{20}$$

where the $PM^*_{j,r,t}$ denotes the price expected by agents for period t and $\beta$ the coefficient of expectations. It is possible to interpret the dynamic specification of IBN as if:

$$PM^*_{j,r,t} = PM_{j,r,t} - PM_{j,r,t-1} \tag{21}$$

which implies that

$$\beta = -1 \tag{22}$$

and

$$PM^*_{j,r,t-1} = \frac{PM_{j,r,t}}{3} \tag{23}$$

This last equation means that, assuming that price expectations follow the process set out by Nerlove, the price expected by agents for the previous period amounts to one third of current price. This is "more" than the perfect foresight assumption, in the sense that this implies that economic agents know before their eventual announcement the taxes of the next period.

To analyze the robustness of the dynamic analysis performed by IBN, we consider the following type of price expectations:

$$\begin{aligned} PM^*_{j,r,t} &= \beta.PM_{j,r,t} + (1-\beta).PM_{j,r,t-1} \\ &= \beta.\left(PM_{j,r,t} - PM_{j,r,t-1}\right) + PM_{j,r,t-1} \end{aligned} \tag{24}$$

We assume that the expected price is a weighted sum of the true price (perfect foresight as in the steady state approach) and the past price (naïve expectation). This specification ends up with perfect foresight if the coefficient of expectation $\beta$ equals one, and naïve price expectations if this coefficient equals zero. We thus estimate the following equation, with one additional parameter to estimate:

$$
\begin{aligned}
\log\left(VF_{i,j,r,t}\right) - \log\left(VO_{j,r,t}\right) \\
&= \beta.\left(1-\sigma_j\right).\left(\log\left(TF_{i,j,r,t}\right) + \log\left(TO_{j,r,t}\right)\right. \\
&\left. - \log\left(TF_{i,j,r,t-1}\right) - \log\left(TO_{j,r,t-1}\right)\right) \\
&+ \left(1-\sigma_j\right)\left(\log\left(TF_{i,j,r,t-1}\right)\right. \\
&\left. + \log\left(TO_{j,r,t-1}\right)\right) + F_{i,j} + G_{r,j}
\end{aligned}
\tag{25}
$$

The distribution of the substitution elasticity estimate is directly retrieved from the coefficient associated with past taxes. On the other hand, the distribution of the coefficient of expectation is obtained thanks to the delta method.

We first estimate equation (25) ignoring the previous CRS and p-value issues. Contrary to the dynamic results of IBN reported in the first column of our table 2, our dynamic estimates reported in the second column of this table are not very different from the initial results. In particular, we do not obtain much larger substitution elasticities. For instance, the elasticity estimate increases from 0.19 to 0.26 for wheat. At the same time, the standard error increases, implying no statistical difference between the two estimates.

**Table 2.** Dynamic estimates of substitution elasticities among inputs

| | IBN (Eq19) | Our specification (Eq25) | Key inputs (Eq25) |
|---|---|---|---|
| **Paddy rice** | | | |
| Estimate | 1.000 | 0.484 | -0.302 |
| Standard error | 0.156 | 0.156 | 0.185 |
| Significance | 0.000 | 0.002 | 0.102 |
| **Wheat** | | | |
| Estimate | 0.473 | 0.261 | -0.192 |
| Standard error | 0.086 | 0.116 | 0.145 |
| Significance | 0.000 | 0.024 | 0.186 |
| **Coarse grains** | | | |
| Estimate | 0.536 | 0.136 | -0.627 |
| Standard error | 0.093 | 0.117 | 0.146 |
| Significance | 0.000 | 0.246 | 0.000 |
| **Vegetable and fruits** | | | |
| Estimate | 0.402 | 0.249 | 0.159 |
| Standard error | 0.104 | 0.129 | 0.167 |
| Significance | 0.000 | 0.054 | 0.342 |
| **Oilseeds** | | | |
| Estimate | 0.924 | 0.238 | -0.146 |
| Standard error | 0.107 | 0.119 | 0.137 |
| Significance | 0.000 | 0.044 | 0.288 |
| **Sugar crops** | | | |
| Estimate | 0.496 | 0.045 | -0.725 |
| Standard error | 0.081 | 0.120 | 0.171 |
| Significance | 0.000 | 0.708 | 0.000 |
| **Plant-based fibers** | | | |
| Estimate | 0.860 | 0.526 | 0.633 |
| Standard error | 0.055 | 0.087 | 0.094 |
| Significance | 0.000 | 0.000 | 0.000 |
| **Other crops** | | | |
| Estimate | 0.557 | 0.542 | -0.211 |
| Standard error | 0.069 | 0.099 | 0.135 |
| Significance | 0.000 | 0.000 | 0.118 |

*Source:* Author calculations

If we limit the estimation to the relevant inputs, the situation appears even more different (last column of table 2): only one substitution elasticity is statistically significantly different from zero with the correct sign.

## 5. Perspectives

IBN's search for the short run and long run estimates of production elasticities is highly relevant for global economic analysis using the GTAP database and parameters. Their idea to use the GTAP Data Bases without input price information is quite original. Unfortunately, their methodological implementation suffers from three issues that undermine the robustness of their estimates. Not imposing the CRS assumption appears to have marginal impacts on IBN's results. On the other hand, our empirical analysis of both the p-value problem and the dynamic specification critically alters their results.

Our comments only focus on the methodologies. Future efforts may explore if the different assumptions made to build the GTAP Data Bases affect the properties of the econometric estimates. The availability of five years of observation may be fruitfully exploited to pursue econometric investigation on some groups of countries (e.g. developed vs developing countries). Future efforts may also depart from a single CES among all variable inputs and consider flexible functional forms or nested structures. All these efforts should capitalize on insights from the micro-econometric literature - e.g. Carpentier et Letort (2012) and Frisvold (2019) work on the substitution between chemical inputs and other inputs used in crop productions).

## Acknowledgements

The authors thank the referee and editor for their useful comments. This research has been conducted within the project 'Facilitate public Action to exit from peSTicides (FAST)' as part of the French Priority Research Programme 'Growing and protecting crops differently' of the French National Research Agency (Agence Nationale de la Recherche, ANR-20-PCPA-0005) https://anr.fr/ProjetIA-20-PCPA-0005. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## References

Aguiar, A., Chepeliev, M., Corong, E., & van der Mensbrugghe, D. 2023. "The Global Trade Analysis Project (GTAP) Data Base: Version 11." *Journal of Global Economic Analysis*, 7(2). doi: https://doi.org/10.21642/JGEA.070201AF.

Carpentier, Alain, and Elodie Letort. 2012. "Accounting for Heterogeneity in Multicrop Micro-Econometric Models: Implications for Variable Input Demand Modeling." *American Journal of Agricultural Economics*, 94(1): 209-224. doi: https://doi.org/10.1093/ajae/aar132.

Costinot, Arnaud, Donaldson, Dave, and Cory Smith. 2016. "Evolving Comparative Advantage and the Impact of Climate Change in Agricultural Markets: Evidence from 1.7 Million Fields Around the World." *Journal of Political Economy*, 124(1): 205-248. doi: https://doi.org/10.1086/684719.

Fontagné, Lionel, Houssein Guimbard, and Gianluca Orefice. 2022. "Tariff-Based Product Level Trade Elasticities". *Journal of International Economics*, 137. doi: https://doi.org/10.1016/j.jinteco.2022.103593

Frisvold, Georges. 2019. "How low can you go? Estimating impacts of reduced pesticide use." *Pest management science*, 75: 1223-1233. doi: https://doi.org/10.1002/ps.5249

Gohin, Alexandre, and Thomas W. Hertel. 2003. "A Note of the CES Functional Form and Its Use in the GTAP model". *GTAP Research Memorandum* n°2, 14 p.

Ivanic, Maros, Jayson Beckman, and Noe J. Nava. 2023. "Estimation of the Value-Added/Intermediate Input Substitution Elasticities Consistent with the GTAP Data." *Journal of Global Economic Analysis*, 2: 134-158. doi: https://doi.org/10.21642/JGEA.080204AF.

Leamer, Edward E. 1978. "*Specification Searches: Ad Hoc Inference with Nonexperimental Data*." John Wiley & Sons, 370 p.

Mundlak, Yair. 2001. "Production and Supply". Handbook of Agricultural Economics, Vol1, Chapter 1: 4-85. doi: https://doi.org/10.1016/S1574-0072(01)10004-6.

Nerlove, Marc. 1958. "Distributed Lags and Estimation of Long-Run Supply and Demand Elasticities: Theoretical Considerations". *Journal of Farm Economics*, XL: 301-310. doi: https://doi.org/10.2307/1234920.

Nerlove, Marc, and David A. Bessler. 2001. "Expectations, Information and Dynamics." Handbook of Agricultural Economics, Vol1, Chapter 3: 155-206. Doi: https://doi.org/10.1016/S1574-0072(01)10006-X.

Reimer, Jeff, and Thomas Hertel. 2004. Estimation of International Demand Behavior for Use with Input-Output Based Data. *Economic Systems Research*, 16(4): 347-366. doi: https://doi.org/10.1080/0953531042000304245.

Rosas, Francisco, Sergio H. Lence, and Dermot J. Hayes. 2018. "Crop yield responses to prices: a Bayesian approach to blend experimental and market data". *European Review of Agricultural Economics*, 46(4): 551-577. doi: https://doi.org/10.1093/erae/jby032.