



# Working Paper Series

*Spatial distribution of the international food prices: unexpected randomness and heterogeneity*

by

Tiziano Distefano, Guido Chiarotti, Francesco Laio, Luca Ridolfi

**01/2018**

SEEDS is an interuniversity research centre. It develops research and higher education projects in the fields of ecological and environmental economics, with a special focus on the role of policy and innovation. Main fields of action are environmental policy, economics of innovation, energy economics and policy, economic evaluation by stated preference techniques, waste management and policy, climate change and development.

The SEEDS Working Paper Series are indexed in RePEc and Google Scholar.  
Papers can be downloaded free of charge from the following websites:  
<http://www.sustainability-seeds.org/>.

Enquiries: [info@sustainability-seeds.org](mailto:info@sustainability-seeds.org)

SEEDS Working Paper 01/2018

January 2018

by Tiziano Distefano, Guido Chiarotti, Francesco Laio, Luca Ridolfi

The opinions expressed in this working paper do not necessarily reflect the position of SEEDS as a whole.

# Spatial distribution of the international food prices: unexpected randomness and heterogeneity

Tiziano Distefano<sup>a</sup>, Guido Chiarotti<sup>a</sup>, Francesco Laio<sup>a</sup>, Luca Ridolfi<sup>a</sup>

<sup>a</sup>*Department of Environmental, Land and Infrastructure Engineering, Politecnico di Torino, Corso Duca degli  
Abruzzi, 24, 10129 Torino, (IT).*

*Corresponding author's e-mail: tiziano.distefano@polito.it.*

---

## Abstract

Global food prices are typically analysed in a times series framework to assess the causes of volatility and to highlight spikes, that are interpreted as a signal of food crises. We address the spatial dimension of the issue at hand by focusing on the spatial food price dispersion, at the country-scale, in the international food trade network (IFTN) for ten relevant commodities.

We base our study on bilateral trade by focusing on both the “internal” variance, which indicates that an exporter sets *different* prices to different importers for the *same* commodity, and the “external” variance, that is a measure of market price competitiveness. We find that spatial price dispersion is remarkable and persistent over time and that there exists a strict correlation between price spikes (in level) and peaks in spatial price variability. This entails that during price crises the market is more fragmented and a higher spatial price dispersion is found.

Moreover, we implement a *randomness* test on the country-scale price distributions to test whether they can be replicated through a random process of extraction. It results that the actual distribution of prices of several commodities is well described by a random distribution. It follows that the process of data aggregation is not neutral because, in several cases, the information at the micro-level scale (firms' decision) is lost at the macro-scale due to the complexity of the international food trade network (IFTN). We suggest some possible economic explanations of this occurrence and we discuss the main methodological consequences.

*Keywords:* food price, spatial dispersion, international trade, randomness

**JEL:** L66, N50, Q11, Q17, R12, R32

## 1. Introduction

Recently, after two waves of world food price crises (2008 and 2011), the economic literature focused on the aftermaths of price ‘spikes’ to capture the short-run effect (e.g., Piesse and Thirtle, 2009, Bellemare, 2015). In addition, a brunch of the literature has focused on the drivers of *temporal* food price volatility (FPV, see Díaz-Bonilla (2016) for a discussion), relating them to: energy prices (Baffes and Haniotis, 2016), speculation (Gilbert, 2010), real interest rate (Frankel, 2006), weak governance (Wang et al., 2016), extreme weather events (Leblois et al., 2014, Tadasse et al., 2016), stock strategies (Wright, 2011), and biofuels (Headey and Fan, 2008).

In contrast, no attempts have been done to understand *spatial* price heterogeneity, in spite of the fact that in the international food market the bilateral prices show a remarkable range of variation within the same category of food (e.g., wheat) during the same period (e.g., one year). The key assumption behind all the above cited analyses is in fact the existence of a unique representative average price (at net of transaction and transport costs) through which assessing the dynamic of food commodities price worldwide. This would be justified in presence of a Gaussian price distribution.

Theory suggests that each form of market (from monopoly to perfect competition) should generate a unique exchange price. However, price dispersion – namely, a homogeneous product being sold at different prices by different exporters – can emerge, for example, in case of entry barriers. In this case, when the difference in net price offered by a seller at two locations does not equal the difference in transport costs among those locations (Graubner et al., 2011, Grebitus et al., 2013), then a dumping strategy could be exploited. The literature empirically analysed spatial price dispersion at the level of food retailer markets (Anania and Nisticò, 2014) but, to our knowledge, no attempts have been done to consider this issue in the international market. In economics, the concept of spatial distribution has been developed mostly to explain where firms strategically decide where to install their location (e.g., the Hotelling model<sup>1</sup>). However, here we depart from this approach because we focus on country-scale exchange and because we are interested in the analysis of the food price variability.

To the best of our knowledge, this study is the first attempt to empirically assess the spatial price distribution in the IFTN, in order to test whether there is a significant price discrimination. In this perspective, we face the seemingly forgotten question: “is what we speak of as *the* price

---

<sup>1</sup>See Graitson (1982), Pinkse et al. (2002), Vogel (2008) for a discussion and review.

[...] in a particular market really a parameter, or is it merely an approximation of a value that varies erratically over a substantial range?” (Brandow (1973), p.390 original emphasis). Or, in other words, (i) is there a *unique price* in the international food price system? (ii) If not, how the same exporter can set different prices to its importers (internal variance)? (iii) Is there a spatial variability among the average prices set by each exporter (external variance)? (iv) Is raw-food price always a good proxy for firms’ choice, or is the IFTN so complex to generate a seemingly random distribution of choice at country-level scale? To shed light on these questions, we run a data-driven analysis of the distribution of bilateral, importing, and exporting prices. Contrary to the theory, we will show that remarkable differences in prices are present even when we get rid of transport costs (and, consequently, of geographical distance).

Another remarkable result refers to the distribution of prices in the IFTN: for a bunch of commodities the actual price distribution is well replicated through a process of random extraction. This unexpected behaviour is possibly due to the scale of our analysis: we focus on aggregate trade data among countries (macro-scale) while we have no information about firms’ strategies (micro-scale). It appears that – for some commodities – the interaction of many factors (economic decisions, political reforms, international agreements, and so on) generates, at country scale, a randomic price distribution. This fact has relevant consequences in terms of modelling the mechanism of price formation, at least at the macro-scale, and it suggests the necessity to introduce product-specific models given the observed cross-commodity heterogeneity. We discuss the main possible drivers behind these findings – e.g., asymmetric bargaining power, different quality within the same category (aggregation bias), low value of raw-commodities with respect to final product, and the impact of domestic food policy – suggesting potential development for future research.

The current analysis is structured as follows: Section 2 explains the reconstruction of food price and the improvements of the original dataset. Section 3 discusses the methodology and the concepts of spatial and temporal heterogeneity of food prices and it shows the main results. Subsection 4.1 describes the statistical analysis to test the randomness of the price distributions. Subsection 4.2 describes the cross-commodity results from the test of randomness of international food prices, and subsection 4.3 discusses the the main economic factors of influence. Finally, Section 5 draws the main conclusions and it introduces some hints for future research.

## 2. Data

Trade data are taken from the publicly-available Food and Agricultural Organization of the United Nations' online database (FAOSTAT),<sup>2</sup> which reports the trade flow among 254 countries<sup>3</sup> for several commodities, from 1986 to 2013. FAOSTAT provides physical (tonnes) and monetary (US dollar) values of the bilateral trade exchanges from which we build the matrices whose entries are the amount of exchange between any single exporter and importer, both in tonnes ( $\mathbf{F}$ ) and dollars ( $\mathbf{P}$ ). We also use the World Bank Data to recover the global inflation rate ( $d$ ).

We select four basic raw food products (wheat, maize, rice milled, and soy-beans) because they cover more than 50% of the global calories intake (D'Odorico et al., 2014). We also add luxury goods such as coffee green and cocoa, some vegetables and fruits (apple and potatoes), eggs and honey. For the sake of simplicity, in most cases we show only the results for two exemplifying goods – *wheat* and *coffee green* – unless specific cases. All the results for the other commodities are shown in the Supplementary Materials.

It is worth to underline that FAOSTAT – given any importer  $i$  and exporter  $j$  – provides both the importer-reported (e.g.,  $F_{jk}^k$ ) and exporter-reported (e.g.,  $F_{jk}^j$ ) declarations, so that the amount of bilateral flows (either in tonnes or in dollars) may differ depending on who is the declarer. Hence, it may happen that, for the same exchange from exporter  $j$  to importer  $k$ , the two declarations differ ( $F_{jk}^j \neq F_{jk}^k$ ), depending on who is the reporting country. This discrepancy might be due to: transport costs (that usually are burden on the importer),<sup>4</sup> loss of part of the commodities during the shipping, errors in reporting values, and political motivations (e.g., a country may be interested in hiding its trade relation with another country that is not appreciated internationally or within the country itself). We applied an algorithm<sup>5</sup> in order to smooth these inconsistencies in the declarations and to build a unique and coherent matrix of bilateral trade, for each year and product. The algorithm assigns an index of reliability to each country, based on the number of reliable declarations. Thus, in case of conflict between exporter ( $F_{jk}^j$ ) and importer ( $F_{jk}^k$ ) reported value of bilateral trade (either tons or in monetary terms), the algorithm will assign to  $F_{jk}$  (or  $P_{jk}$ )

---

<sup>2</sup>FAO, Statistics Division. FAOSTAT online database. Available at <http://www.fao.org/faostat/en>. Last Update on December 11, 2015.

<sup>3</sup>The number of active countries changed over time due to political reasons. For example, the USSR is active only until 1991.

<sup>4</sup>In FAOSTAT the exporter-reported values are mostly reported as f.o.b. (free on board), while importer-reported values are mostly reported as CIF (cost insurance and freight).

<sup>5</sup>See the Supplementary Materials and Gehlhar (1996) for a complete description.

the value declared by the country with a higher reliability index. Obviously, in case one of the two countries declares nothing, then the algorithm will assign the value declared by the other partner. The step-by-step description of data reconciliation is provided in the Supplementary Materials.

### 3. Spatial price dispersion in the IFTN

The current study focuses on the spatial variability (and its temporal evolution) of product-specific prices set by different countries in the IFTN. Our approach is complementary to the standard analysis of FPV, which considers the time variance of the average (yearly, monthly, or daily) price of a specific commodity. Different from standard economic analysis, we aim at analysing three measures of *spatial price dispersion* – that we observe from the data – when looking at the annual average price of each country-pair exchange. The IFTN is a directed and weighted network composed by  $N_E$  exporters and  $N_M$  importers, where each edge is directed from the exporter to the importer and weighted by total volume (either tons or dollars) of food traded. The set of all trade relationships – weighted and directed links – between exporters and importers generates the distribution of bilateral price from which we compute the *total* variance ( $\sigma_{tot}^2$ ) of prices. Moreover, it is possible to further decompose  $\sigma_{tot}^2$  between: (i) the *internal* variance  $\sigma_{int}^2$ , that is the weighted average of the internal variability of the distribution of prices set by a single  $j$ -th exporter ( $\sigma_{int,j}^2$ ) to all its direct importers, and (ii) the *external* variance ( $\sigma_{ext}^2$ ) evaluated from the distribution of the average prices set by every exporter.

Without loss of generality, we describe our procedure without the time specification since each operation is repeated every year in the time window 1986-2013. Given the matrix of bilateral trade price per ton ( $\mathbf{P}$ )<sup>6</sup> – deflated by global inflation rate ( $d$ ) in order to be comparable over time<sup>7</sup> – and

---

<sup>6</sup>Note that in our case we should speak of yearly ‘unitary average value’ since we recover  $P$  as the ratio between nominal values (\$) and the physical flow (tonnes). For the sake of simplicity, and because the average values we found are close to the average prices of New York and Bremen/Hamburg markets and of the USDA database, we opt to keep the label “price”. For a detailed description on the differences and implications of these two concepts see Gehlhar and Pick (2002). On the distinction between intra- and inter-annual variability of price, see Ott (2014).

<sup>7</sup>Note that given the presence of two partners, the exporter and the importer, the selection of the “right” deflator might be ambiguous. Indeed, if one chooses, for instance, the Consumer (or Producer) Price Index of the exporter then all the values of the importers will be distorted, mostly in the case of a rich exporting country and a poor importing country with different food baskets of consumption (von Braun and Tadesse, 2012). For this reason we opt to select a single yearly *global* deflator to reduce possible distortions.

tonnes ( $\mathbf{F}$ ) for a given product, the yearly-average global price ( $\bar{P}$ ) is computed as

$$\bar{P} = \sum_j^{N_E} \sum_k^{N_M} P_{jk} \cdot \frac{F_{jk}}{F_{tot}}, \quad (1)$$

where  $F_{tot}$  is the overall physical flow traded on the network (in a given year). Note that Eq. (1) corresponds to a weighted average, where the weights are the market shares ( $F_{jk}/F_{tot}$ ). The set of all the trade relationships generates the distribution of bilateral price from which we compute the total weighted variance of bilateral prices as

$$\sigma_{tot}^2 = \sum_j^{N_E} \sum_k^{N_M} (P_{jk} - \bar{P})^2 \cdot \frac{F_{jk}}{F_{tot}}. \quad (2)$$

Let  $\bar{P}_j$  be the weighted (by import share) average price set by each exporter  $j$ , defined as

$$\bar{P}_j = \sum_k^{N_M^j} P_{jk} \cdot \frac{F_{jk}}{F_j}, \quad (3)$$

where  $N_M^j$  is the total number of importers for any generic exporter  $j$  (i.e., the out-degree) and  $F_j = \sum_k^{N_M^j} F_{jk}$  is the total export of  $j$  (in a given year).

We define the country-side internal variance ( $\sigma_{int,j}^2$ ) – of the prices set by each exporter  $j$  to its importers – as the weighted (by import share) average of the quadratic distance of the bilateral price ( $P_{jk}$ ), set to each importer, from  $\bar{P}_j$ ; namely,

$$\sigma_{int,j}^2 = \sum_k^{N_M^j} \frac{(P_{jk} - \bar{P}_j)^2 \cdot F_{jk}}{F_j}. \quad (4)$$

The aggregate *internal* variance ( $\sigma_{int}^2$ ) is the weighted (by tonnes exported)<sup>8</sup> sum of any internal variance ( $\sigma_{int,j}^2$ ), while the *external* variance is a measure of the quadratic distance of the average price fixed by each exporter ( $\bar{P}_j$ ) with respect to the average global price ( $\bar{P}$ ). Namely,

---

<sup>8</sup>We focus only on exporter side for two reasons: first, importers typically have only few partners (3 on average); second, the IFTN is usually dominated by few big exporters that cover most of the overall trade.

$$\sigma_{int}^2 = \sum_j^{N_E} \sigma_{int,j}^2 \cdot \frac{F_j}{F_{tot}} \quad (5)$$

$$\sigma_{ext}^2 = \sum_j^{N_E} (P_j - \bar{P})^2 \cdot \frac{F_j}{F_{tot}}. \quad (6)$$

Finally, the *temporal* variability, of the yearly average price set by every exporter  $j$ , is computed as the weighted standard deviation from the yearly average price ( $\bar{P}_j(t)$ ):

$$\sigma_{T,j}^2 = \sum_t^T (\bar{P}_j(t) - \bar{P}_{T,j})^2 \cdot \frac{F_j(t)}{F_{tot,j}}, \quad (7)$$

where  $T$  is the overall time span (viz. 28 years),  $\bar{P}_{T,j}$  is the average price of the  $j$ -th exporter during the time span, and  $F_{tot,j}$  is the overall amount of exports from  $j$  during the entire period. The distribution of  $\sigma_{T,j}^2$  provides further information about the external variability and on market competitiveness. Indeed, if most of the  $\sigma_{T,j}^2$  values are concentrated within a narrow range, it follows that most of the exporters set a price close to each other over time.

### 3.1. Empirical Evidence

Left and central panels of Figure 1 compare, for wheat and coffee green, the time series of the weighted average price  $\bar{P}_j$  (red lines) for top exporters,<sup>9</sup> where the red band width is proportional to the internal standard deviation ( $\sigma_{int,j}$ , see Eq. (4)). Several features that characterise and differentiate the wheat and coffee green market emerge. Firstly, the average price is much higher for coffee green, which ranges between  $\sim 1000\$$  to  $\sim 4000\$$  per ton (with a peak of  $\sim 6000\$/\text{ton}$  for Colombia during the 2011 crisis), while the wheat price is within the interval  $\sim 100\$$  —  $\sim 300\$$  per ton. Secondly, the prices set by the exporters are remarkably more spatially correlated in case of coffee green than for wheat. For instance, Brazil and Colombia exhibit a similar trend – with a coefficient of correlation of about 0.47 – although the average prices are not the same; differently, the time series of  $\bar{P}_j$  of wheat shows that the USA and France are negligibly correlated ( $-0.05$ ),

---

<sup>9</sup>Note that in both cases (wheat and coffee green) the two countries selected cover jointly about the 33% of market sales.

mostly before 1994 when the prices of the latter were more stable and higher.<sup>10</sup> Interestingly, the wheat and coffee green markets reacted differently to the two recent world price food crises: wheat showed two remarkable spikes in those years, while coffee green was not affected by the crisis in 2008 but only by that occurred in 2011.<sup>11</sup>

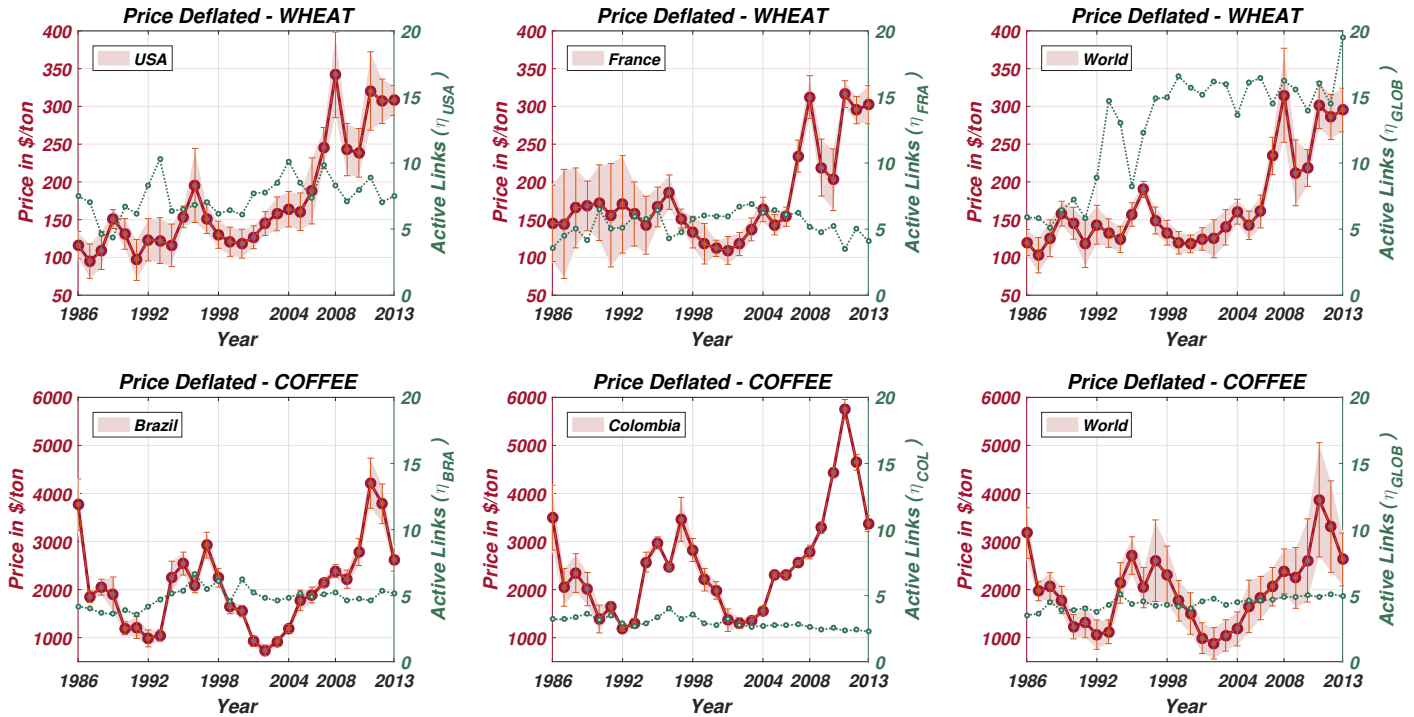


Figure 1: Empirical time series – from 1986 to 2013 – of the the average price of top two exporters,  $\bar{P}_j$ , and global price,  $\bar{P}$ , for wheat (top) and coffee green (bottom). Red band width, around the red line, is proportional to the internal standard deviation of each country ( $\sigma_{int,j}$ ), in case of single exporter, and to the external standard deviation in the global case ( $\sigma_{ext}$ ). Green lines (right y-axes) are the scaled out-degree ( $\eta$ ).

Thirdly, the internal standard deviation (red band width) suggests a first clue of the possible dumping strategy in the wheat market that has a greater internal variance than coffee; in fact, exporters (mostly France until 1996) showed, over time, a remarkable variation of prices around their mean (viz. internal variance). Fourthly, the world time series (right panels) confirms the previous findings: the trend of the global coffee price is strictly correlated with those of the main exporters, while in case of wheat the linkage is weak; moreover, the range of variation of exporter

<sup>10</sup>Note that the same considerations hold even when we include more exporters. See the Supplementary Materials for a description with more traders and for all the commodities.

<sup>11</sup>See Luttinger and Dicum (2011) for a description of the coffee market.

prices (i.e.  $\sigma_{ext}$ , red band width) is higher in case of coffee green.

The green line in figure 1 shows the “scaled out-degree” of each exporter  $j$ , that is the weighted number of active links, computed as (neglecting the time specification):

$$\eta_j = \sum_k^{N_M^j} \eta_{jk} = \sum_k^{N_M^j} \frac{F_{jk}}{F_{jk}^{max}} \quad (8)$$

where  $F_{jk}^{max}$  is the maximum amount of single trade from exporter  $j$  to its top importer  $k$ , which then has  $\eta_{jk} = 1$ . Note that in case of the world,  $\eta$  is computed over the whole set of importers in the IFTN. Formula (8) puts a minor weight to importers with a little share of imports, so that the indicator  $\eta_j$  is always lower than the simple out-degree (viz.  $\eta_j < N_M^j$ ), unless all the importers have the same share. In this way, we identify the number of main players in each market; indeed, if  $\eta_j \simeq m$ , independently of the actual number of active links of  $j$ , it entails that most of the bilateral trade from  $j$  is concentrated toward  $m$  big importers. The main difference among the two commodities, shown in Fig. 1, appears in the global values: wheat is experiencing a remarkable increase in the number of relevant agents that goes up to almost 20 in 2013; while in the case of coffee the market appears smaller, with few countries ( $\sim 5$ ) that dominate the international trade.

At this point it is worth to investigate the three types of variances defined in Section 3. Figure 2 (left panel) shows the time series of  $\sigma_{tot}$  (associated to bilateral trade, see Eq. (2)) for both commodities (red lines). In case of wheat,  $\sigma_{tot}$  is rather stable from 1986 to 2006 when it sharply increased with two peaks in 2008 and in 2011; after that, it is again stable but on a higher level, showing a higher  $\sigma_{ext}$  than in the pre-shock period. This outcome strengthens the common observation for which a sudden rise in price level increases the temporal volatility (von Braun and Tadesse, 2012), suggesting that during the crises the spatial variance increased as well. In line with the literature of temporal volatility (Maurice and Davis, 2011), we observe that spatial price dispersion of the coffee green market was higher with two peaks in 1997 (due to an increase in energy, raw materials and payroll costs (Talbot, 2004)) and in 2011. Over time it showed many ups and downs. A further confirmation of the higher variability of coffee green market is given by the average (over time) coefficient of variation of bilateral prices ( $\overline{CV}_{T,jk}$ ) – where the yearly coefficient of variation is given by the ratio between the standard deviation ( $\sigma_{tot}$ ) and the mean ( $\overline{P}$ ) of the considered year – which is higher for coffee (0.32) than for wheat (0.25).

When looking at wheat, we observe that  $\sigma_{int} > \sigma_{ext}$  always (except for 2002), meaning that the

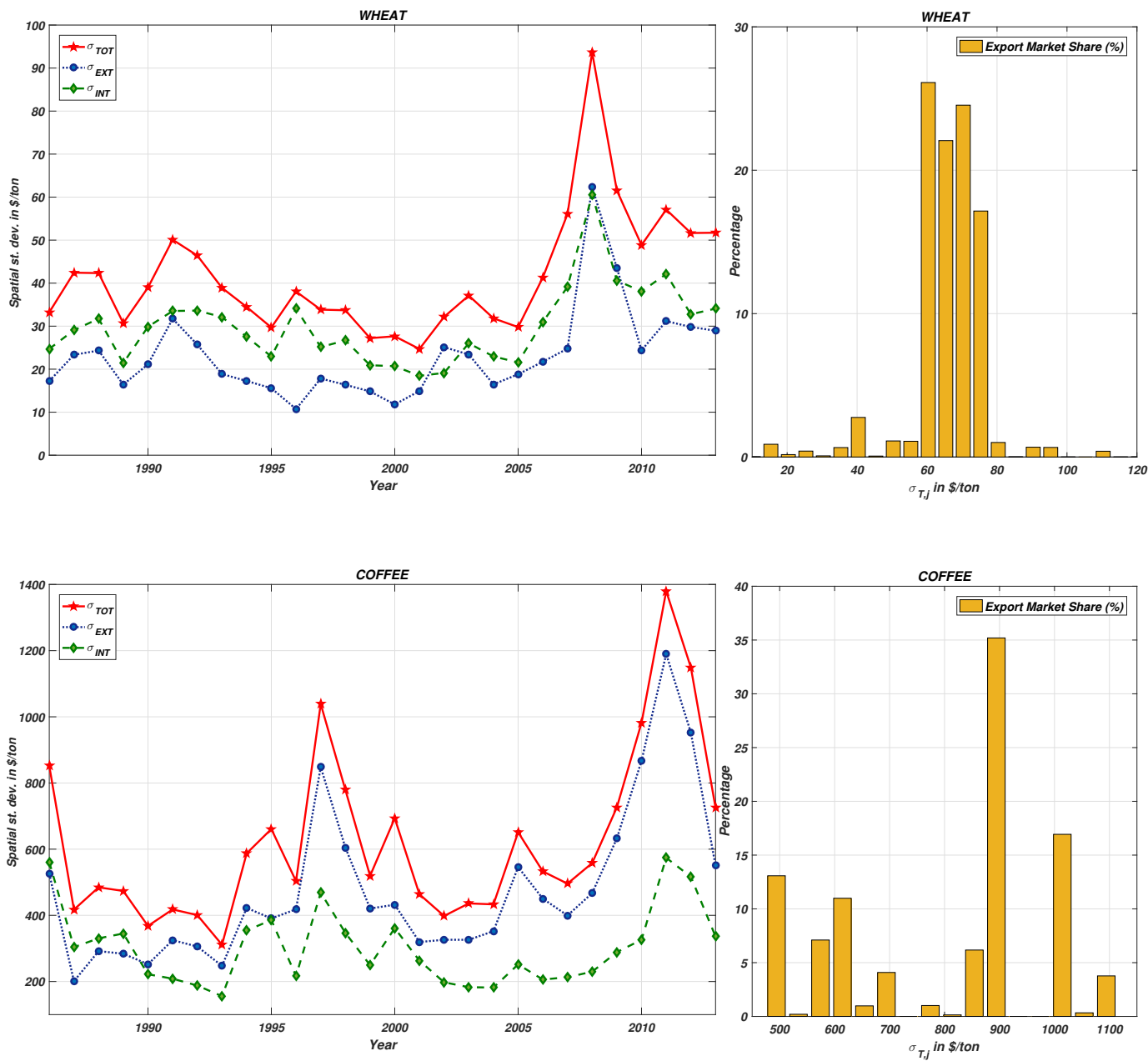


Figure 2: Empirical time series of the square roots of  $\sigma_{tot}^2$  (red line),  $\sigma_{ext}^2$  (blue line), and  $\sigma_{int}^2$  (green line). The right panels show the empirical distribution of the weighted standard deviation in the average price set by each exporter in the whole time span (i.e.,  $\sigma_{T,j}$ ). The importance of the exporter (i.e., total export  $F_j$ ) is given by its market share. The two commodities are wheat (top) and coffee green (bottom).

exporting countries have very similar competitive power, and that each exporter set, on average, different prices to its importers (viz. dumping strategy). In case of coffee green the opposite con-

sideration holds ( $\sigma_{int} < \sigma_{ext}$ ), hence a dumping strategy seems more difficult since each exporter<sup>12</sup> sets similar prices to its trading partners (low  $\sigma_{int}$ ); however, the exporters set different prices from competitors (high  $\sigma_{ext}$ ) probably through different market niches (differentiated by quality). This observation is supported by the average (over time) coefficient of variation: the  $\overline{CV}_{T,int}$  of wheat is higher than for coffee (0.20 and 0.15, respectively); while the opposite holds for the  $\overline{CV}_{T,ext}$  (0.14 and 0.25, respectively).

The right panels in Fig. 2 show the distribution of the weighted (by market shares) average of  $\sigma_{T,j}$  (see Eq. 7). They confirm that the wheat and coffee markets are remarkably different: the former shows that the average exporting price has not changed significantly over time and that most of the exporters have followed a similar path. Indeed, most of the variation is concentrated around 60-70\$/ton ( $\sim 80\%$  of traders). Instead, in the case of coffee green, the variations are much higher and less concentrated, with a peak of  $\sim 35\%$  of traders on an average temporal variation of 900\$/ton, and all the other ranging from 500 to 1100 dollars per ton. This is coherent with what observed for  $\sigma_{ext}$ .

Table 1: Cross-commodity summary of the key findings about spatial price dispersion.  $\overline{P}$  stands for the weighted (by total yearly flow) average global price (\$/ton) in the whole time span.  $\overline{CV}_T$  is the average coefficient of variation in the whole period computed on total ( $jk$ ), external ( $ext$ ), internal ( $int$ ), and temporal ( $TIME$ ) variances. For the latter in brackets it is specified whether the distribution is concentrated around the mean ('con.') or not ('spread').  $\eta$  is the average (over time) scaled degree – as computed in Eq. (8) – for the overall number of (relevant): links ( $\eta_{tot}$ ), exporters ( $\eta_{exp}$ ), and importers ( $\eta_{imp}$ ).

Item	$\overline{P}$	$\overline{CV}_{T,jk}$	$\overline{CV}_{T,ext}$	$\overline{CV}_{T,int}$	$\eta_{tot}$	$\eta_{exp}$	$\eta_{imp}$	$\overline{CV}_{TIME}$
<i>maize</i>	159	0.42	0.26	0.27	6.9	2.1	5.7	0.37 (con.)
<i>wheat</i>	178	0.25	0.14	0.18	23.2	4.3	13.3	0.38 (con.)
<i>soy-beans</i>	320	0.16	0.08	0.11	6.6	2.1	3.4	0.40 (con.)
<i>rice</i>	401	0.49	0.32	0.32	20.2	3.3	13.5	0.39 (con.)
<i>apples</i>	614	0.46	0.34	0.28	19.9	6.9	6.8	0.30 (spread)
<i>potatoes</i>	906	0.16	0.07	0.10	10.7	4.8	6.2	0.07 (con.)
<i>honey</i>	1277	0.58	0.37	0.33	10.5	4.2	3.8	0.43 (spread)
<i>eggs</i>	1304	0.69	0.45	0.42	5.4	3.6	4	0.31 (spread)
<i>cocoa beans</i>	1632	0.20	0.12	0.12	10.1	2.9	4.4	0.38 (con.)
<i>coffee green</i>	2017	0.33	0.25	0.16	19.2	4.3	4.5	0.39 (spread)

<sup>12</sup>See Hernandez et al. (2017) for a description of the buying (cartel) strategy of importers in the coffee market.

Table 1 summarises the main outcomes for the other commodities (all the corresponding Figures are shown in the Supplementary Materials). The basket of food commodities investigated shows a remarkable range of variation for several features. Almost half of them have a relatively low unitary price ( $< 500\$/\text{ton}$ ); however, in every cases (but potatoes) the average global price increased substantially over time (with a  $\overline{CV}_{TIME}$  between 0.3 and 0.43).<sup>13</sup> Beside this common pattern, the market competitiveness was different: in four cases (apple, honey, eggs, and coffee green) the distribution of the price variation over time, for the overall exporters, was rather spread, suggesting the presence of hidden quality. This is further confirmed by observing that each of these products has  $\overline{CV}_{T,ext} > \overline{CV}_{T,int}$  that, following the interpretation done before for coffee green, would suggest the possibility of hidden quality (not represented at current categorical aggregation) which can allow the emergence of market niches differentiated by quality.

Additionally, based on Eq. (8), we compute the number of ‘relevant’ traders from both the exporter- ( $\eta_{exp}$ ) and importer-side ( $\eta_{imp}$ ). It results that  $\eta_{exp} < \eta_{imp}$  always, with the exceptions of apples and honey. In general, these indicators are extremely low, meaning that only few countries (around 3-4) are dominating the IFTN, with the exception of wheat and rice that show higher values ( $\sim 13$ ) from the importer-side. Moreover, the ‘scaled’ number of edges ( $\eta_{tot}$ ) shows remarkable differences among the products: in some cases most of the trade is concentrated in few links ( $\sim 6$ , for soy-beans, eggs, maize), while in other cases we find higher values (apples, coffee green, rice, and wheat,  $\sim 20$ ).

Last but not least, each category of food showed a persistent and not decreasing high spatial price heterogeneity over time ( $\overline{CV}_{T,jk}$ ). It entails that the average global price is not a representative indicator of the value of bilateral exchanges. If so, what is the distribution of prices? Does the macro-scale preserve the micro-scale characteristics and choices (allegedly rational) or does new properties (e.g., randomness) emerge in the IFTN at country-scale due to high complexity? To answer to these questions, we investigate the price formation mechanism through a statistical perspective, as described below.

---

<sup>13</sup>See the Supplementary Material for the graphical representation of the distribution of  $\sigma_{T,j}$  for each commodity.

#### 4. Food price formation: a statistical approach

In this Section we attempt to reconstruct the statistical distribution from which the observed prices (bilateral, import, and export average) are possibly generated. The main idea is to test whether the actual prices distribution may or not be replicated from a random process of extractions. In a simplified world, one would expect that each agent tries to buy (sell) at the lowest (highest) possible price, if the quality of the commodity is homogeneous, and that this feature is preserved at the country scale. Hence, once we get rid of transaction costs<sup>14</sup> we should either observe either a unique global price (equal to the marginal cost if we assume a supply-driven market under perfect competition) or a distribution of prices that would reflect the bargaining power of each partner (mostly dependent on market share, trade barriers, and economic power). Over time, one would guess that arbitrage and/or the process of globalisation (increasing number of countries and transactions) should smooth the range of variation because of a higher possibility to access the international market. However, as observed above, the IFTN shows a persistent and remarkable range of variation of bilateral prices within the same year for the same kind of commodity.<sup>15</sup>

Before presenting the mathematical details, we briefly discuss the assumptions behind the process of random extractions that we test against real data. Our approach is based on two key assumptions: (i) there exists a unique price distribution from which the prices are drawn; (ii) each country-level trade relation ( $F_{jk}$ ) can be decomposed into homogeneous blocks of the same amount (in tons), representing the trade among firms. Indeed, the bilateral flow ( $F_{jk}$ ), from an exporter (e.g., the USA) to an importer (e.g., Italy), is the overall sum of all the transactions occurred among the *firms* of the two countries. For example, if the USA sells 30000 tons of wheat to Italy, then we assume that this bilateral flow is formed by, say, 30 single blocks, ‘as if’ 30 Italian firms are buying 1000 tons each from the American ones.<sup>16</sup>

The logic behind our procedure is that the random sampling of the price distribution fixes the null hypothesis of the random versus causality price determination of goods trading: if the actual distribution (empirically measured) is similar to that of the null hypothesis we can conclude that - at the country level - there is no cross-correlation among the decisions of buyers and/or sellers, while for significant deviations of the measured distribution from the null hypothesis, signs of causality

---

<sup>14</sup>See Appendix A.1 for a detailed description of the data analysis and of the treatment of transaction costs.

<sup>15</sup>See the Supplementary Materials for a description of spatial price distribution for the other commodities.

<sup>16</sup>It can be shown that the block size does not alter the outcome from the random extraction.

persist also at the country level. In the following we will show that both conditions occur in the IFTN depending on the kind of the traded commodity.

#### 4.1. The Global Price Distribution

In this subsection we investigate the probabilistic behaviour of the food price at trade. The available data include the bilateral price ( $P_{jk}$ , in \$/ton) paid by the importer country  $k$  to the exporting country  $j$ . Clearly,  $P_{jk}$ , as well as the aggregate average values  $\bar{P}_k$  and  $\bar{P}_j$  (see Eq. (3)), will assume different values, depending on the specific countries considered. However, these empirical differences in price could be interpreted or as the outcome of specific price-formation mechanisms, or as the outcome of the sample variability of price. In what follows, we try to clarify what of these two issues is suitable for the different commodities. We are therefore interested in defining a procedure to test the null hypothesis  $H_0$  that the  $P_{jk}$  values are obtained by randomly sampling the price from a unique probability distribution, against the alternative hypothesis  $H_1$  that heterogeneities exist in the sampling procedure, to say that the  $P_{jk}$  values have been sampled from different parent distributions. If the null hypothesis turns out to be true, the price formation mechanism is a simple random sampling; otherwise, endogenous variables have an influence on the price formation (for example, export prices of a country are systematically larger than others because the cost of production in the country is higher than elsewhere).

Our understanding of the spatio-temporal price dynamics at country-scale and our capacity to reproduce them are thus crucially determined by the test we are performing. However, formalizing a testing procedure is quite complicated, because price data refer to highly heterogeneous fluxes, which in turn impacts the statistical characterization of the price. For example, we expect a much higher sample variability of the price of a  $10^2$  tons flux compared to a  $10^6$  tons flux, because the latter is likely made up of a large number of smaller exchanges among the firms of the two countries (i.e., blocks), with random price fluctuations compensating with one another, thus reducing the variability of the aggregated price. We thus build up our testing procedure starting from the following ancillary assumptions:

- (i) each edge  $F_{jk}$  is composed by a number  $N_{jk}$  of homogeneous blocks of size  $f$  (e.g.,  $f = 1000$  tons), representing the typical amount of food exchanged in a single economic transaction between two firms. In formulas,  $F_{jk} = f \cdot N_{jk}$  (the flux size is approximated to the closer multiple of  $f$ );

- (ii) the exchange price  $p$  of the block  $f$  is a random variable with a distribution  $g_{jk}(p)$ . The distribution is the same for all blocks belonging to  $F_{jk}$ , but might be different for different fluxes;
- (iii) the price of each block is independent of the price of the other blocks in  $F_{jk}$ ;
- (iv) the distribution  $g_{jk}(p)$  is a Gamma distribution, with parameters  $\theta_{jk}$  and  $\lambda_{jk}$ :

$$g_{jk}(p; \theta_{jk}, \lambda_{jk}) = \frac{1}{\theta_{jk}^{\lambda_{jk}} \cdot \Gamma(\lambda_{jk})} \cdot p^{\lambda_{jk}-1} e^{-\frac{p}{\theta_{jk}}}. \quad (9)$$

The assumptions (ii) and (iii) correspond to the hypothesis that price-formation mechanisms at the country scale loses the causality present at the firm scale, assuming a rational behaviour in the decision process. Note that price data are typically only available at the country scale, and thus these assumptions cannot be profitably verified with the available data. Verification of the null hypothesis  $H_0$ , however, entails an indirect verification of these ancillary assumptions too.

Under the assumptions (i)–(iii) the probability distribution of  $P_{jk}$  is obtained as the distribution of the average of  $N_{jk}$  independent random variables with common distribution  $g_{jk}(p)$ . Using the assumption (iv), and the fact that the sum of independent Gamma variables is again Gamma-distributed, one obtains:  $P_{jk} \stackrel{d}{\sim} \text{Gamma}[\theta_{jk}/N_{jk}, N_{jk} \cdot \lambda_{jk}]$ , where  $\stackrel{d}{\sim}$  means ‘distributed as’. The null hypothesis of complete randomness of the price-formation mechanism can now be formalized as follows: the hypothesis  $H_0$  is  $P_{jk} \stackrel{d}{\sim} \text{Gamma}[\hat{\theta}/N_{jk}, N_{jk} \cdot \hat{\lambda}]$ , namely  $p \stackrel{d}{\sim} \text{Gamma}[\hat{\theta}, \hat{\lambda}]$ , where the subscripts have been dropped from the estimated parameters  $\hat{\theta}$  and  $\hat{\lambda}$  because we assume that they do remain the same for any considered couple of countries. Under  $H_0$ , the global parameters  $\hat{\theta}$  and  $\hat{\lambda}$  characterize, together with the size of the fluxes, the probability distribution of each and any of the food price at trade.

One can therefore calculate the probability value  $q_{jk} = \gamma(\hat{\lambda}, \frac{P_{jk}}{\hat{\theta}})$  by calculating in  $P_{jk}$  the Gamma cumulative probability distribution with parameters  $\hat{\theta}/N_{jk}$  and  $N_{jk} \cdot \hat{\lambda}$ . Under  $H_0$  the  $q_{jk}$  values follow a uniform distribution, so that  $q_{jk} \stackrel{d}{\sim} \text{Uniform}(0,1)$ . Verification of  $H_0$  can thus be performed through a standard uniform probability plot. If the points lie close to the bisector of the plot, the data are likely to be sampled from a uniform, which in turns implies that the  $P_{jk}$  values are obtained by randomly sampling the single-block price from a unique global probability distribution, i.e.  $p \sim \text{Gamma}(\hat{\theta}, \hat{\lambda})$ . We repeat the same procedure at a higher scale by including

the average importing and exporting prices (See the Appendix A for the mathematical details).

In summary, our randomness test is based on the following steps: (i) estimating the spatial price dispersion associated to each block ( $N_{jk}$ ) and the aggregate spatial variance of the unique distribution of bilateral prices ( $\hat{\sigma}_{tot}^2$ ), (ii) estimating the cumulative probability ( $q$ ) as if the bilateral prices were randomly picked from the unique global distribution, (iii) repeating the same analysis for the aggregate level of import ( $\bar{P}_k$ ) and export ( $\bar{P}_j$ ), and (iv) performing a cross-commodity comparison of the actual distribution of prices with the one emerging from a random extraction. Note that step (iii) allows us to verify the presence of asymmetric information among buyers and sellers.

#### *4.2. Cross-commodity comparison*

Here we discuss the graphical results of testing randomness against causality in the distribution of prices per ton for all the ten commodities. Figure 3 shows the cumulative probability of  $q$  (computed for bilateral (left), importing (center), and exporting (right) prices) against the cumulative market share, that is the weight of each edge (composed by  $N_{jk}$  blocks). Given the stability of results over time, and in order to obtain a larger sample, we pool the results of all years together. Note that the procedure described in Section 4.1 is mean-preserving because we compute the  $q$  probability for each observed price. What differs is the total spatial variance that provides us with the information about the randomisation of the prices observed, under  $H_0$ . Obviously, a result distant from the unique Gamma distribution might also be due to the fact that the real distribution is not well represented by a Gamma one, and not only by a different total spatial variance.

We recall that if the observed prices were actually extracted from the unique Gamma distribution, then it must be that the cumulative distribution of  $q$  is a Uniform(0,1). In this case, the cumulative distribution of  $q$  lies on the bisector (black line). This consideration entails that if the actual cumulative distribution of  $q$  – of a given commodity (coloured lines) – is close to the bisector, then its distribution (at macro-scale) can be replicated by a random process of extraction. This implies that the information at micro-scale (firms’ decision) is lost in the process of data aggregation. For the sake of clearness, we split the ten commodities in two categories: the distributions that are well described by a random process of extraction (top) and the others (bottom). This distinction is based on bilateral prices, although in some cases the results from importer and exporter side might be different.

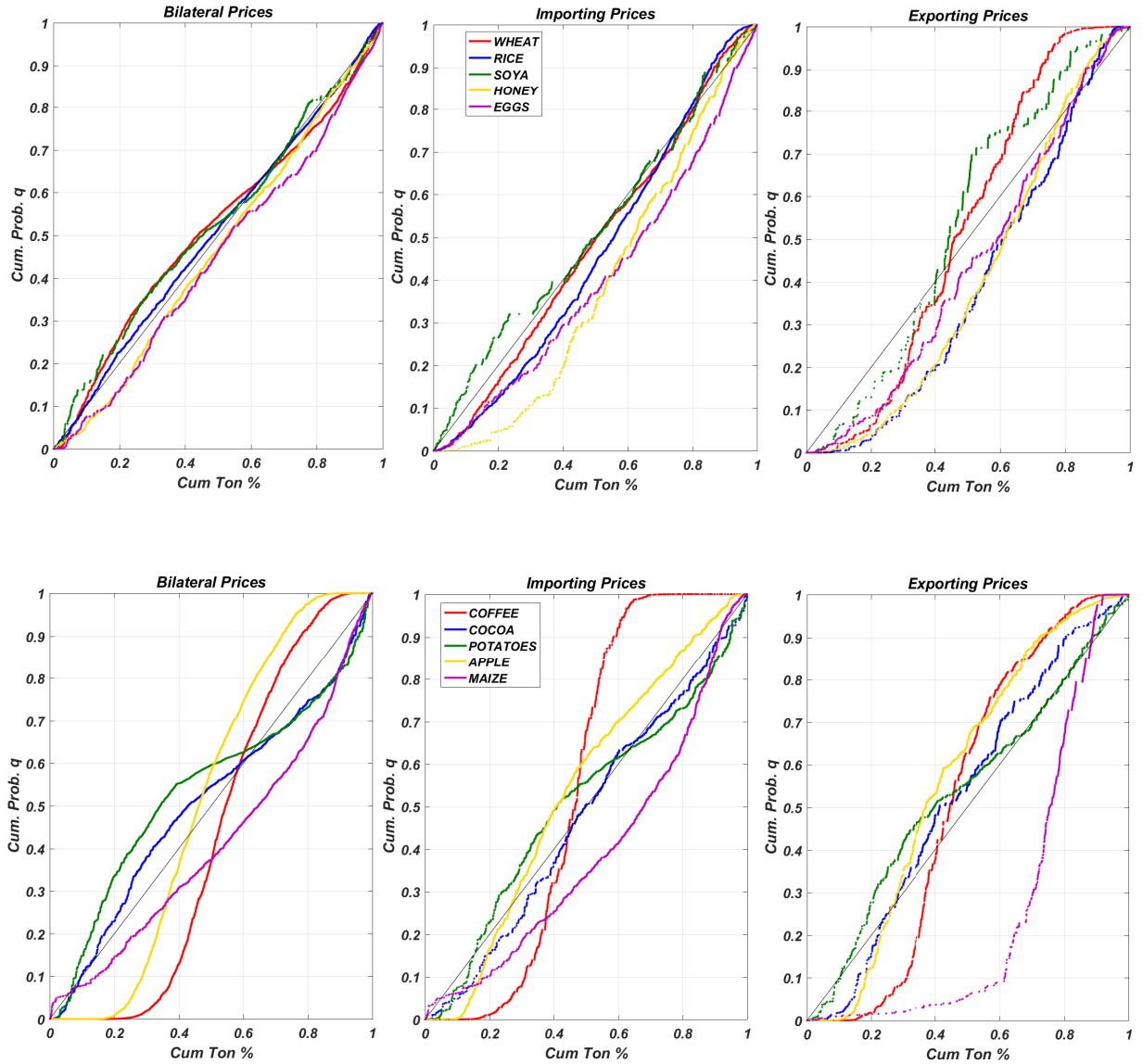


Figure 3: Empirical weighted (by market shares) cumulative distribution against theoretical Gamma distribution, for all years pooled together, for all commodities. Left, center, and right panels refer to bilateral, aggregate import, and aggregate export, respectively.

Based on the cross-commodity comparison, we found that in half of the cases (wheat, rice, soy-bean, honey, and eggs) the distribution of *bilateral* prices is well replicated by the process of random extraction from the global distribution since the curves are close to the bisector. In the other cases we obtain a S-shaped curve, with either an under-estimation (cocoa beans and potatoes) or an over-estimation (coffee green and apples) of the variance. When looking at the *importer* side, we observe that the price distributions of the first item category is still well approximated by the

random process in three cases (wheat, soy-beans, and rice), while honey shows a fatter lower tail and eggs have a distorted (non-randomic) estimation. In the second item category, all the commodities continue to follow a systematic distorted (non-randomic) behaviour, with the exception of cocoa beans, whose distribution is well approximated by a random extraction. From the *exporter* side the picture changes and, in general, the prices do not seem to adhere in a satisfying way to the null hypothesis of a unique distribution. The differences observed between the importer and exporter side might be due the fact that the firms (micro-level) have less information or a minor chance to properly compare all the alternatives (due to cost of research, time constraints, geographical and cultural distance, and so on), then showing a random behaviour at the country-level scale (macro).

The differences observed in the price distribution should not be simply reduced to the specific network properties of each item, because the commodities that fall in the first (random) and second category share common network features (see Table 1). Hence, in what follows we discuss economic factors that shed light in the interpretation of these outcomes.

#### 4.3. Discussion

Here, we outlines some relevant factors that could explain, at least partially, the randomness of price distribution at the country scale.

(I) **Scale of analysis.** Our study, as done in most of the literature, is based on country level data that aggregate all the information related to the firms' behaviour, which we are "mimicking" by decomposing the weight of each edge in many (identical) blocks of trade. Complex systems, as IFTN is, show the emergence of different properties when they are observed at different scales (Georgescu-Roegen, 1993, Sawyer, 2005, Malghan, 2010); in our case, we have 'macro' (country) but not 'micro' (firms) level information. In the current study, the randomness is an emergent property of the network – at least for half of the examined products – meaning that the different (expected rational and deterministic) strategies and actions taken at the micro-level interact in a way that generate a random distribution at the macro-level. On the other hand, a bunch of commodities do not show this feature meaning that the micro-level information is not lost after the aggregation. This consideration entails that the scale of analysis at hand (country-level) is at the frontier between randomness and causality (depending on the commodity one is focusing on), to say that only product-specific analysis might ensure when one of these two cases occur. These results have relevant consequences in terms of modelisation and data collection: when the

micro-level structure is preserved, then a macro model (based on average global price) might still be meaningful to understand deterministic aspects of the market dynamics and data aggregation will not be biased, while the opposite interpretation holds in case of random behaviour at aggregate scale. These considerations suggest a case-specific analysis on the possibility of the micro-foundation of the macro structure and dynamics of a specific market.

(II) **Market structure<sup>17</sup> and vertical price transmission** from farmers to retailers. Recent studies showed the relatively minor role played by food price in the world trade growth of food commodities (Serrano and Pinilla, 2010). In some cases, the initial commodity is such a small ingredient in the end product that any linkage between the farm and retail price could be completely lost (Kim and Ward, 2013). In our context, farm-to-retail gap and imperfect vertical price transmission can explain the different results, in terms of random distribution, at least from the importer side. Indeed, if the farm-to-retail gap is high and the vertical price transmission adjustment is low and asymmetric (sticky retail prices), then the importers (mostly if owners of most of the supply chain, i.e. big companies) can decide to care less about the raw food prices (Vavra and Goodwin, 2005). In addition to these reasons, the presence of (seemingly) random choice might also be due to the absolute low price of the commodity that supersedes the consumers' search costs allowing exporters to apply spatial price discrimination. This, together with collusive behaviours, might generate different pricing decisions by sellers and buyers (Anania and Nisticò, 2014).

(III) **Counter-cyclical policy** intervention to stabilize the domestic food price (importer strategy) or to grasp higher price (exporter strategy). This implies that exporters may impose restrictions to obtain higher prices to importers, and conversely importers may limit the impact of low prices on their economy by applying tariffs. The asymmetric nature of the distributions of food commodity prices, with more prices below than above the mean but with occasional spikes (Gouel, 2016), might generate a series of un-coordinated country-level policies whose effects cancel out at the aggregate level, showing (in some cases) a random process of allocation of food prices. As a matter of example the presence of low  $\sigma_{ext}$  in the coffee market might be explained by the fact that control in coffee markets came via a combination of buffer stocks and quota limitation of exports with the aim of maintaining prices within target bands that were agreed between consumer and producer nations (Gilbert et al., 2010). Moreover, the asymmetric price transmission of global food

---

<sup>17</sup>See Distefano et al. (2017) for a detailed analysis of international food market structure.

prices to domestic prices (in particular in Developing Countries) might be imperfect and slow, then justifying the acceptance of different prices from different exporters (Zorya et al., 2012).

(IV) **Hidden quality heterogeneity**: in the IFTN the quality dimension plays an important role in determining countries' export success, due to the increasing consumers' requirements in terms of food safety and nutrition (Curzi and Pacca, 2015). Indeed, by assuming that higher price corresponds to higher quality when comparing traded food products, we risk incurring in an imperfect identification, since the gap in prices may be also due to other reasons, such as different export strategies or different production costs. This suggests that including quality is fundamental if we want to get a complete picture of how barriers to import influence countries' trade patterns. In our case, quality is hidden due to the categorical aggregation, but the remarkable difference between wheat and coffee green (when looking at  $\sigma_{int}$ ) suggests that the latter category is composed by several types of coffee beans with significant different qualities. It is also conceivable that a picture including different levels of aggregation of the commodities might show the observed random emergence at different geographical levels (e.g., at scales lower than that of country-scale), or conversely that, once a differentiation in quality is implemented, randomness emerges at the country-scale also for those commodities for which is presently not observed.

## 5. Conclusion

The current study focused on two relevant issues (i.e., spatial price dispersion and possible random price formation) that are, at best of our knowledge, unexplored in the available food price literature, but that provide a novel approach to the analysis of food price volatility. The first part was devoted to the evaluation of the *spatial price dispersion*, that is the variability of the price set by each exporter to its trading partners. The key findings are:

- (i) the spatial bilateral price dispersion ( $\sigma_{tot}$ ) is remarkable and persistent over time;
- (ii) the distinction between the internal and external variability allows one to get insights about the market competitiveness and hidden quality (market niches), respectively;
- (iii) there is a strict correlation between price spikes and peaks in spatial  $\sigma_{tot}$  in most of the cases (e.g., wheat price crises in 2008 corresponded to the maximum value of  $\sigma_{tot}$ ). It entails that during price crises the market is more fragmented and more opportunities for dumping strategy may emerge.

The study of the spatial price dispersion has noteworthy consequences on the understanding of shock propagation. Indeed, it tells us that there is not a unique price of exchange in the IFTN; therefore, basing political decision simply on the average global price might lead to misleading results. Methodologically, our approach is complementary with the current food price literature that defines a food crisis in correspondence of a price spike. Indeed, including information about the price spatial dispersion – in addition to the global average price – can improve our understanding of how developed and developing countries are affected by price spikes. Let us provide a couple of examples to clarify how our approach might help in the identification of food price shocks, to understand which countries will actually be affected. During the wheat price crisis of 2008 – where the average global price ( $\bar{P}$ ) was  $\sim 320\$/\text{ton}$  – Egypt bought from the USA more than 2 millions of tonnes of wheat, representing about the 24% of the country’s importation of that year, at an average price of  $\sim 270\$/\text{ton}$  (15% less than  $\bar{P}$ ). Then, in this case the global price spike had no effect on Egypt. Conversely, during the 1999-2002 time window the average global price of wheat was stable around  $130\$/\text{ton}$ ; however, Algeria bought almost 300000 tonnes of wheat from the USA at  $180\$/\text{ton}$  (+30%). It follows that if one had looked only at the global average price one would have not identified a number of ‘local’ food price crises due to the presence of high price spatial heterogeneity. This observation might support the political decision of developing countries that are more sensible to price variations, since they show high income and price elasticities for staple foods (Cornelsen et al., 2015). Indeed, there is no one-fits-all solution for all the countries; rather, future research might benefit from the inclusion of price spatial dispersion when studying the price shock propagation, mostly in the identification of which countries are actually experiencing price spikes.

The second part of the current study was related to the issue of *randomness* against causality in food price distributions at the country-scale. One may expect rational and random decisions generate very different distributions. In contrast, in its seminal paper Becker (1962) pointed out that demand-side rational behaviour can be obtained on average even if consumers choose in a random way (Moscati and Tubaro, 2011). In our case we observe the reverse, at the macro-scale (with no info about firms’ strategy) the distribution of bilateral and importing prices (in half of the cases under assessment) are well replicated by a process of random extraction, while this does not hold true in case of exporters’ price distribution. This result has important consequences: first, the random distributions (when observed) include both rich and poor countries, meaning that

bargaining power, at least from the importing side, is not related to the level of affluence. Second, future models of price formation (at the macro-scale) should be product-specific and they should distinguish between those commodities that preserve the (micro-level) structure after aggregation (e.g., coffee green) from those that show the emergent property of randomness (e.g., wheat). In the latter cases, since our results are time independent, the presence of randomness entails that, predicting the future values of global mean and spatial variability, one can reconstruct the future distribution of bilateral prices.

Finally, the high spatial variability of food prices and the presence of randomness at the country-scale, suggest that price signals might not be always reliable. If so, the consequences after price shocks will be less clear and global average food price might be a weak tool for policy actions. To conclude, some questions arise: why the arbitrage is not effective in levelling (spatially) the prices? Is it due to market imperfection or to specific failures of the IFTN market (commodities storability, stagionality, trade barriers, agricultural policy reforms, and so forth)? Or is it due to the high complexity of this network? At what level of aggregation (both geographical and of categorisation of commodities) do randomness emerge in food price distribution? These questions are not fully answered in this article; they are the subject of ongoing research.

## **Acknowledgments**

Thanks are due to Stefania Tamea who provided us with the algorithm to reconcile the data on bilateral trade. Data and results of our work are available on request from the corresponding author. Authors acknowledge ERC funding for the project: Coping with water scarcity in a globalized world (ERC-2014-CoG, project 647473).

## **References**

- Anania, G., Nisticò, R., 2014. Price dispersion and seller heterogeneity in retail food markets. *Food Policy* 44, 190–201.
- Baffes, J., Hanjotis, T., 2016. What explains agricultural price movements? *Journal of Agricultural Economics* 67 (3), 706–721.
- Becker, G. S., 1962. Irrational behavior and economic theory. *Journal of political economy* 70 (1), 1–13.

- Bellemare, M. F., 2015. Rising food prices, food price volatility, and social unrest. *American Journal of Agricultural Economics* 97 (1), 1–21.
- Brandow, G., 1973. The food price problem. *American Journal of Agricultural Economics* 55 (3), 385–390.
- Cornelsen, L., Green, R., Turner, R., Dangour, A. D., Shankar, B., Mazzocchi, M., Smith, R. D., 2015. What happens to patterns of food consumption when food prices change? evidence from a systematic review and meta-analysis of food price elasticities globally. *Health economics* 24 (12), 1548–1559.
- Curzi, D., Pacca, L., 2015. Price, quality and trade costs in the food sector. *Food Policy* 55, 147–158.
- Díaz-Bonilla, E., 2016. Volatile volatility: Conceptual and measurement issues related to price trends and volatility. In: *Food Price Volatility and Its Implications for Food Security and Policy*. Springer, pp. 35–57.
- Distefano, T., Laio, F., Ridolfi, L., Schiavo, S., September 2017. Shock transmission in the international food trade network. a data-driven analysis, sEEDS Working Paper Series 06/2017. URL <http://www.sustainability-seeds.org/papers/RePec/srt/wpaper/0617.pdf>
- D’Odorico, P., Carr, J. A., Laio, F., Ridolfi, L., Vandoni, S., 2014. Feeding humanity through global food trade. *Earth’s Future* 2, 458–469, doi:10.1002/2014EF000250.
- Frankel, J. A., 2006. The effect of monetary policy on real commodity prices. Tech. rep., National Bureau of Economic Research.
- Gehlhar, M., 1996. Reconciling bilateral trade data for use in gtap. *GTAP Technical Papers*, 11.
- Gehlhar, M. J., Pick, D. H., 2002. Food trade balances and unit values: What can they reveal about price competition? *Agribusiness* 18 (1), 61–79.
- Georgescu-Roegen, N., 1993. The entropy law and the economic problem. *Valuing the earth: Economics, ecology, ethics*, 75–88.
- Gilbert, C. L., 2010. How to understand high food prices. *Journal of Agricultural Economics* 61(2), 398–425.

- Gilbert, C. L., Morgan, C. W., et al., 2010. Has food price volatility risen? In: Technological Studies Workshop on Methods to Analyse Price Volatility. Seville. pp. 28–29.
- Gouel, C., 2016. Trade policy coordination and food price volatility. *American Journal of Agricultural Economics* 98 (4), 1018–1037.
- Graitson, D., 1982. Spatial competition a la hotelling: A selective survey. *The Journal of Industrial Economics*, 11–25.
- Graubner, M., Balmann, A., Sexton, R. J., 2011. Spatial price discrimination in agricultural product procurement markets: a computational economics approach. *American Journal of Agricultural Economics* 93 (4), 949–967.
- Grebitus, C., Lusk, J. L., Nayga, R. M., 2013. Effect of distance of transportation on willingness to pay for food. *Ecological Economics* 88, 67–75.
- Headey, D., Fan, S., 2008. Anatomy of a crisis: the causes and consequences of surging food prices. *Agricultural Economics* 39 (s1), 375–391.
- Hernandez, M. A., Rashid, S., Lemma, S., Kuma, T., 2017. Market institutions and price relationships: The case of coffee in the ethiopian commodity exchange. *American Journal of Agricultural Economics* 99 (3), 683–704.
- Kim, H., Ward, R. W., 2013. Price transmission across the us food distribution system. *Food policy* 41, 226–236.
- Leblois, A., Quirion, P., Sultan, B., 2014. Price vs. weather shock hedging for cash crops: ex ante evaluation for cotton producers in cameroon. *Ecological Economics* 101, 67–80.
- Luttinger, N., Dicum, G., 2011. *The coffee book: Anatomy of an industry from crop to the last drop*. The New Press.
- Malghan, D., 2010. On the relationship between scale, allocation, and distribution. *Ecological Economics* 69 (11), 2261–2270.
- Maurice, N., Davis, J., 2011. Unravelling the underlying causes of price volatility in world coffee and cocoa commodity markets. MPRA.

- Moscato, I., Tubaro, P., 2011. Becker random behavior and the as-if defense of rational choice theory in demand analysis. *Journal of Economic Methodology* 18 (2), 107–128.
- Ott, H., 2014. Volatility in cereal prices: Intra-versus inter-annual volatility. *Journal of agricultural economics* 65 (3), 557–578.
- Piesse, J., Thirtle, C., 2009. Three bubbles and a panic: An explanatory review of recent food commodity price events. *Food policy* 34 (2), 119–129.
- Pinkse, J., Slade, M. E., Brett, C., 2002. Spatial price competition: a semiparametric approach. *Econometrica* 70 (3), 1111–1153.
- Sawyer, R. K., 2005. *Social emergence: Societies as complex systems*. Cambridge University Press.
- Serrano, R., Pinilla, V., 2010. Causes of world trade growth in agricultural and food products, 1951–2000: a demand function approach. *Applied Economics* 42 (27), 3503–3518.
- Tadasse, G., Algieri, B., Kalkuhl, M., von Braun, J., 2016. Drivers and triggers of international food price spikes and volatility. In: *Food Price Volatility and Its Implications for Food Security and Policy*. Springer, pp. 59–82.
- Talbot, J. M., 2004. *Grounds for agreement: The political economy of the coffee commodity chain*. Rowman & Littlefield Publishers.
- UNCTAD, 2015. *Review of maritime transport 2015*. Tech. rep., UNCTAD.  
URL <http://unctad.org/en/pages/PublicationWebflyer.aspx?publicationid=1374>
- Vavra, P., Goodwin, B. K., 2005. *Analysis of price transmission along the food chain*. OECD, iLibrary.
- Vogel, J., 2008. Spatial competition with heterogeneous firms. *Journal of Political Economy* 116 (3), 423–466.
- von Braun, J., Tadesse, G., 2012. *Global food price volatility and spikes: an overview of costs, causes, and solutions*. Center for Development Research, Bonn.
- Wang, X., Biewald, A., Dietrich, J. P., Schmitz, C., Lotze-Campen, H., Humpeöder, F., Bodirsky, B. L., Popp, A., 2016. Taking account of governance: Implications for land-use dynamics, food prices, and trade patterns. *Ecological Economics* 122, 12–24.

Wright, B. D., 2011. The economics of grain price volatility. *Applied Economic Perspectives and Policy* 33 (1), 32-58.

Zorya, S., Townsend, R., Delgado, C., 2012. Transmission of global food prices to domestic prices in developing countries: Why it matters, how it works, and why it should be enhanced (unclassified paper: A contribution of world bank to g20). Washington DC: The World Bank.

## Appendix A. Mathematical details of the randomness test

Here we describe the mathematical details that stand behind the construction of the randomness test, as explained in Section 4. In particular, we show how we estimate: (i) the parameters of the unique global Gamma distribution  $(\hat{\theta}, \hat{\lambda})$ , (ii) the moments of the distribution of each trade block (between firms), to say the mean  $(\psi_p)$  and the variance  $(s_{\psi_{jk}}^2)$ , (iii) the moments of the unique global distribution, to say the mean  $(\psi)$  and the variance  $(s^2)$ , and (iv) the probability density  $(q)$  associated to each price extraction.

The relations between parameters and moments for the Gamma distribution are:

$$\theta = \frac{\sigma_p^2}{\mu_p} \qquad \lambda = \frac{\mu_p^2}{\sigma_p^2} \qquad (\text{A.1})$$

where  $\mu_p$  is the average price of a block and  $\sigma_p^2$  is its variance. Estimators of  $\mu_p$  and  $\sigma_p^2$  are obtained by considering that we may imagine the available data as belonging to a unique global sample of  $N_{tot}$  values of  $p$ , where

$$N_{tot} = \sum_j^{N_E} \sum_k^{N_M} \frac{F_{tot}}{f} \qquad (\text{A.2})$$

where  $F_{tot}$  is the total quantity traded in a year worldwide. The global sample is in turn made up of sub-samples of different size, where each sub-sample contains the same value,  $P_{jk}$ , repeated for  $N_{jk}$  times.

A different estimator of  $\mu_p$  and  $\sigma_p^2$  can be obtained from each sub-sample. The estimator of  $\mu_p$  from the  $(j, k)$  sub-sample is  $\psi_{jk} = P_{jk}$ . A global estimator can be obtained as the weighted average of the sub-sample estimators, where the weight is the size of the sub-sample (larger samples produce more accurate estimators and should be provided with a larger weight):

$$\psi_p = \frac{1}{N_{tot}} \sum_j^{N_E} \sum_k^{N_M} N_{jk} \cdot \psi_{jk} = \frac{1}{N_{tot}} \sum_j^{N_E} \sum_k^{N_M} N_{jk} \cdot P_{jk} = \frac{1}{F_{tot}} \sum_j^{N_E} \sum_k^{N_M} P_{jk} \cdot F_{jk} \qquad (\text{A.3})$$

The variance of  $\psi_{jk}$  about the global average is  $(P_{jk} - \psi_p)^2$  and can be related to the variance of a block in the sub-sample,  $s_{\psi_{jk}}^2$ , by the relation  $s_{\psi_{jk}}^2 = s_{jk}^2 / N_{jk}$ , which holds because  $\psi_{jk}$  is the average

of  $N_{jk}$  independent elements. The estimator of  $\sigma_p^2$  from the  $(j, k)$  sub-sample is thus  $s_{jk}^2 = N_{jk} (P_{jk} - \psi_p)^2$ . A global estimator can be obtained again as the weighted average of the sub-sample estimators,

$$s^2 = \frac{1}{N_{tot}} \sum_j^{N_E} \sum_k^{N_M} N_{jk} \cdot s_{jk}^2 = \frac{1}{N_{tot}} \sum_j^{N_E} \sum_k^{N_M} N_{jk}^2 \cdot (P_{jk} - \psi_p)^2 = \frac{1}{f \cdot F_{tot}} \sum_j^{N_E} \sum_k^{N_M} F_{jk}^2 \cdot (P_{jk} - \psi_p)^2 \quad (\text{A.4})$$

The method-of-moments estimators of  $\theta$  and  $\lambda$ ,  $\hat{\theta}$  and  $\hat{\lambda}$ , are obtained by setting  $\mu_p = \psi_p$  and  $\sigma_p^2 = s_p^2$  in Eq. (12). The last step toward the verification of the hypothesis  $H_0$  entails using the information that, under  $H_0$ ,  $P_{jk} \stackrel{d}{\sim} \text{Gamma}(\hat{\theta}/N_{jk}, N_{jk} \cdot \hat{\lambda})$ . One can therefore calculate the probability value  $q_{jk} = \gamma(\hat{\lambda}, \frac{P_{jk}}{\hat{\theta}})$  by calculating in  $P_{jk}$  the Gamma cumulative probability distribution with parameters  $\hat{\theta}/N_{jk}$  and  $N_{jk} \cdot \hat{\lambda}$ , as

$$q_{jk} = G(P_{jk}; \hat{\theta}, \hat{\lambda}) = \int_0^{P_{jk}} g(P_{jk}; \hat{\theta}, \hat{\lambda}) du \quad (\text{A.5})$$

Note that  $\hat{\theta}$  scales with  $1/f$  and  $\hat{\lambda}$  scales with  $f$ ; as a consequence, both  $\hat{\theta}/N_{jk}$  and  $N_{jk} \cdot \hat{\lambda}$  are independent of  $f$ , which means that the procedure produces the same results for any value of  $f$ .

Under  $H_0$  the  $q_{jk}$  values follow a uniform distribution (because Eq. (A.5) is a probability integral transform), then  $q_{jk} \stackrel{d}{\sim} \text{Uniform}(0,1)$ . Verification of  $H_0$  can thus be performed through a standard uniform probability plot. If the points lie close to the bisector of the plot, the data are likely to be sampled from a uniform, which in turns implies that the  $P_{jk}$  values are obtained by randomly sampling the single-block price from a unique global probability distribution,  $p \sim \text{Gamma}(\hat{\theta}, \hat{\lambda})$ . We repeat the same procedure at a higher scale by including the average importing and exporting price.

To summarise, in the three cases we need to compute the incomplete gamma function as:

(i)  $q_{jk} = G(P_{jk}; \hat{\theta}, \hat{\lambda})$  for bilateral trade, where:

$$\hat{\sigma}_{jk}^2 = \hat{\sigma}_{tot}^2 \cdot \frac{f}{F_{jk}} \qquad \hat{\theta}_{jk} = \frac{\hat{\sigma}_{jk}^2}{\psi_p} \qquad \hat{\lambda}_{jk} = \frac{\psi_p^2}{\hat{\sigma}_{jk}^2} \quad (\text{A.6})$$

(ii)  $q_k = G(P_k; \hat{\theta}_k, \hat{\lambda}_k)$  for the importer side, where  $P_k$  is the average *importing* price of country  $k$

computed as in Eq. (3) (where in this case the sum runs over the trading partner of  $k$ ):

$$\hat{\sigma}_k^2 = \hat{\sigma}_{tot}^2 \cdot \frac{f}{F_k} \quad \hat{\theta}_k = \frac{\hat{\sigma}_k^2}{\psi_p} \quad \hat{\lambda}_k = \frac{\psi_p^2}{\hat{\sigma}_k^2} \quad (\text{A.7})$$

(iii)  $q_j = G(P_j; \hat{\theta}_j, \hat{\lambda}_j)$  for the importer side, where  $\bar{P}_j$  is the average *exporting* price of country  $k$  computed as in Eq. (3):

$$\hat{\sigma}_j^2 = \hat{\sigma}_{tot}^2 \cdot \frac{f}{F_j} \quad \hat{\theta}_j = \frac{\hat{\sigma}_j^2}{\psi_p} \quad \hat{\lambda}_j = \frac{\psi_p^2}{\hat{\sigma}_j^2} \quad (\text{A.8})$$

## Supplementary Materials

### SM.1 Reconciling Bilateral Trade: Tonnes and Unitary Monetary Values (\$)

In what follows we explain the step-by-step procedure for the reconstruction of bilateral monetary and physical flows, as follow:

1. download of original data from FAOSTAT for bilateral trade in \$ and tonnes both reported from exporter and importer side;
2. build the matrices of monetary ( $\tilde{V}$ ) and physical ( $\tilde{F}$ ) Bilateral Trade, for each year, for exporter reported ( $\tilde{V}_E$  and  $\tilde{F}_E$ ) and importer reported ( $\tilde{V}_M$  and  $\tilde{F}_M$ ) data;
3. in case of monetary values, subtract the transport costs<sup>18</sup> from the importer reported declaration (in order to be consistent with exporters' declarations. Indeed, import values are mostly reported as CIF (cost insurance and freight), while export values are declared on f.o.b. (free on board) basis;
4. apply the algorithm<sup>19</sup> in order to obtain a unique consistent matrix of physical ( $\mathbf{F}$ ) and monetary ( $\mathbf{V}$ ) bilateral trade;
5. correct minor bias, such as deleting values on the diagonals;
6. in order to avoid bias and inconsistency, put equal to zero any bilateral flows smaller than a given threshold (in tonnes) depending on the commodity. The threshold is set in order to preserve at least the 99% of total flows and the 50% of links;<sup>20</sup>
7. building the matrix of bilateral price ( $\tilde{P}$ , i.e. unitary values as average \$/ton) as:  $\tilde{P}_{jk} = \frac{V_{jk}}{F_{jk}}$ ;
8. correct unreliable declarations: if  $\tilde{P}_{jk} < \bar{P}_{min} = 0.1 \cdot \bar{P}$  or  $P_{jk} > \bar{P}_{max} = 10 \cdot \bar{P}$  then  $P_{jk} = \gamma \bar{P}_j$ , where  $\bar{P}_j$  is the aggregate average price of exporter  $j$  and  $\gamma$  is the weighted (by tons) average ratio between the price done by  $j$  to all other exporter with respect to  $\bar{P}_j$ . Here  $\bar{P}$  stands for the weighted (by tonnes) global average price. This operation is done in each year and it does

---

<sup>18</sup>See UNCTAD (2015) – Chapter 3 (Figure 3.6) – for a comparison of transport costs among macro-economic regions.

<sup>19</sup>See Gehlhar (1996) for a complete description of the algorithm.

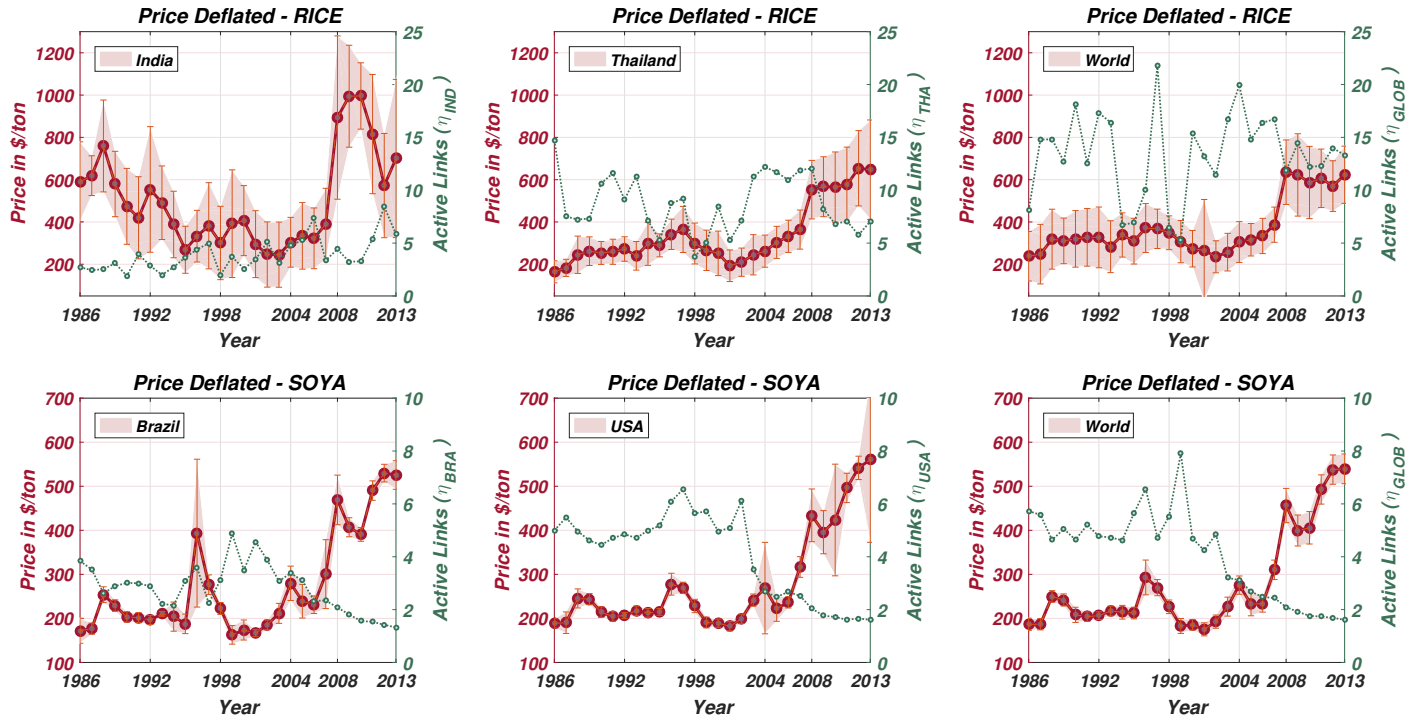
<sup>20</sup>In case of wheat, rice milled, soy-bean, and maize the threshold equals 1000 tonnes; for coffee green and cocoa-beans is 100 tonnes; while for honey, eggs, potatoes, and apple is 10 tonnes.

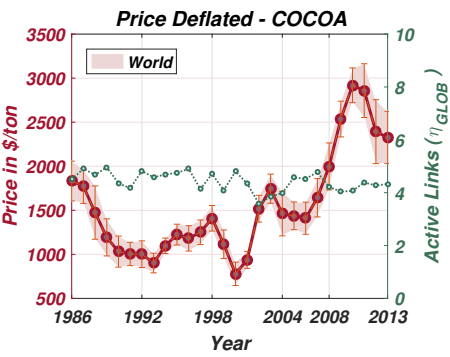
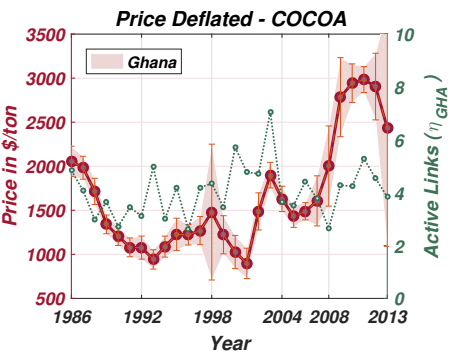
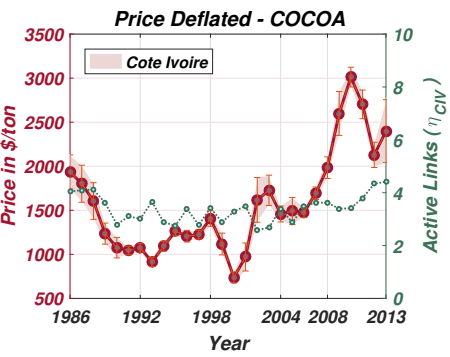
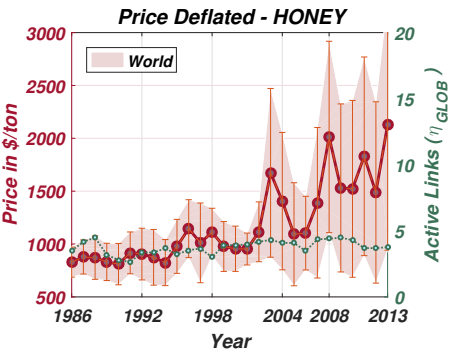
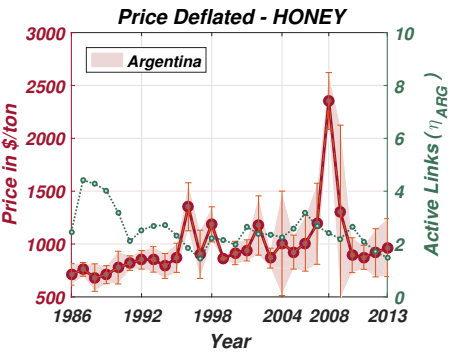
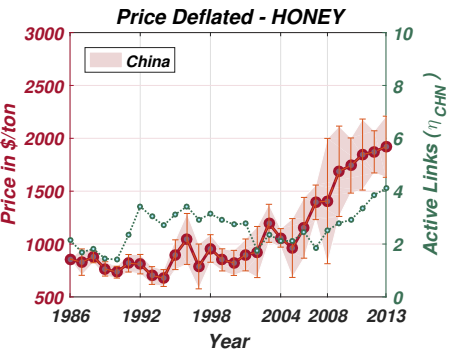
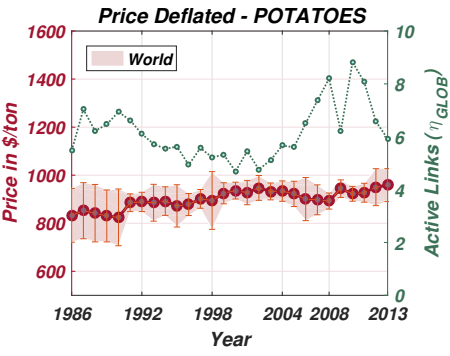
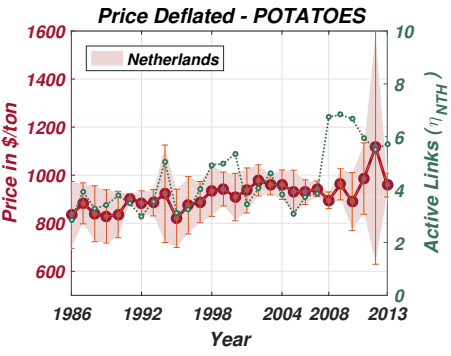
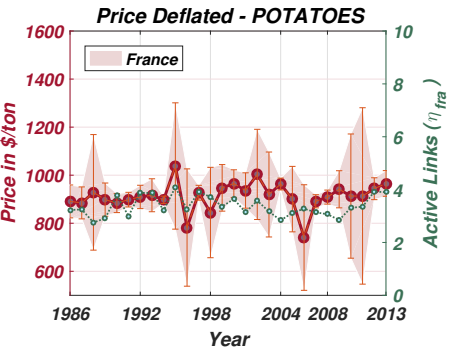
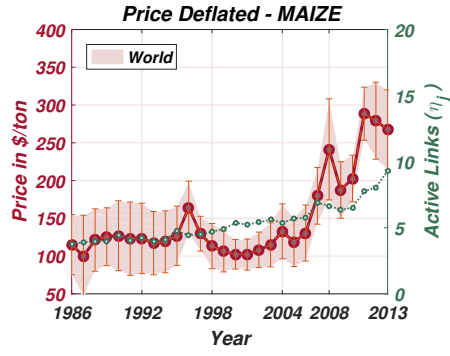
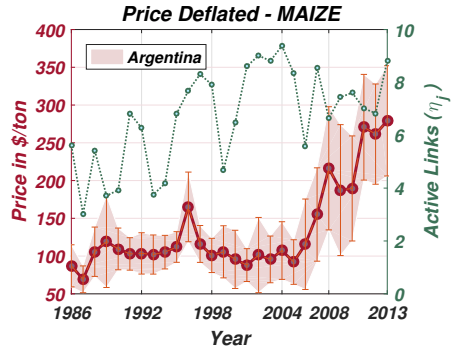
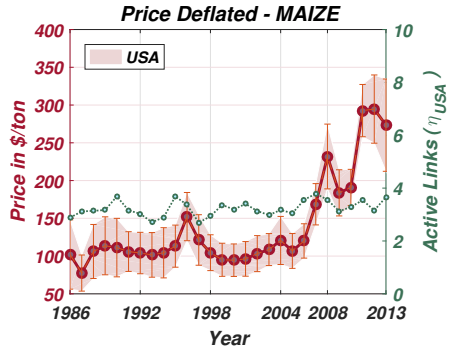
not alter the structure of the network since the whole corrections relate to less than 0.5‰ of total flows;

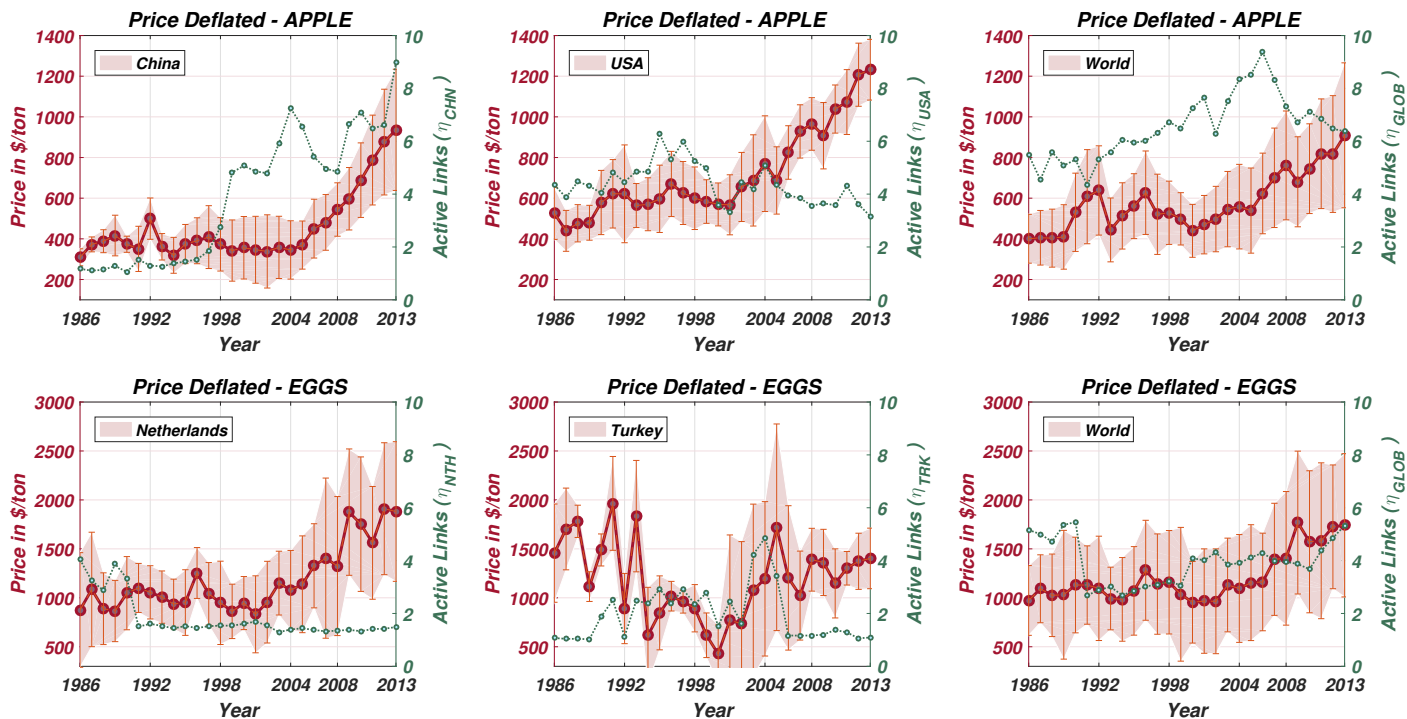
9. build the deflated bilateral price by dividing the nominal values with the yearly global inflation rate, obtaining the final matrix of bilateral price ( $\mathbf{P}$ ).

SM.2 Spatial price dispersion for all considered commodities

Empirical time series – from 1986 to 2013 – of the the average price of top two exporters and global price. Red band width, around the red line, is proportional to the the standard deviation of variation of each year ( $\sigma_{ext}$  or  $\sigma_{int,j}$  in case of single exporter). Green line (right y-axes) is the scaled out-degree ( $\eta$ ).

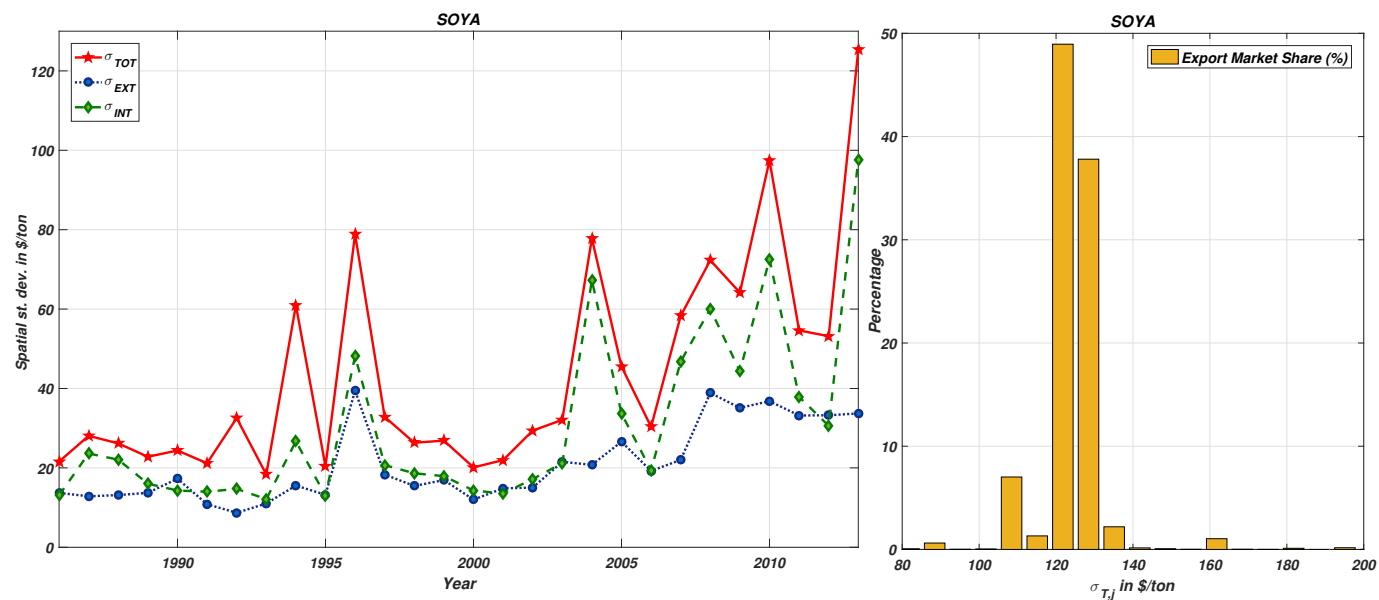
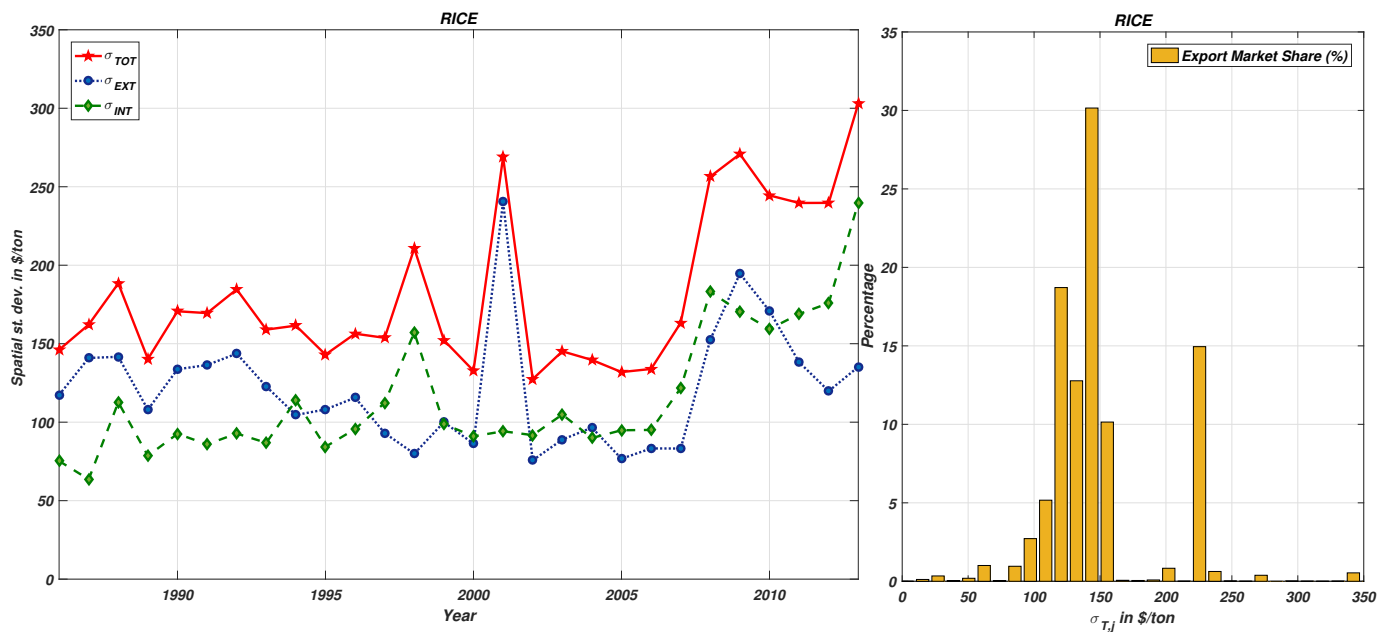


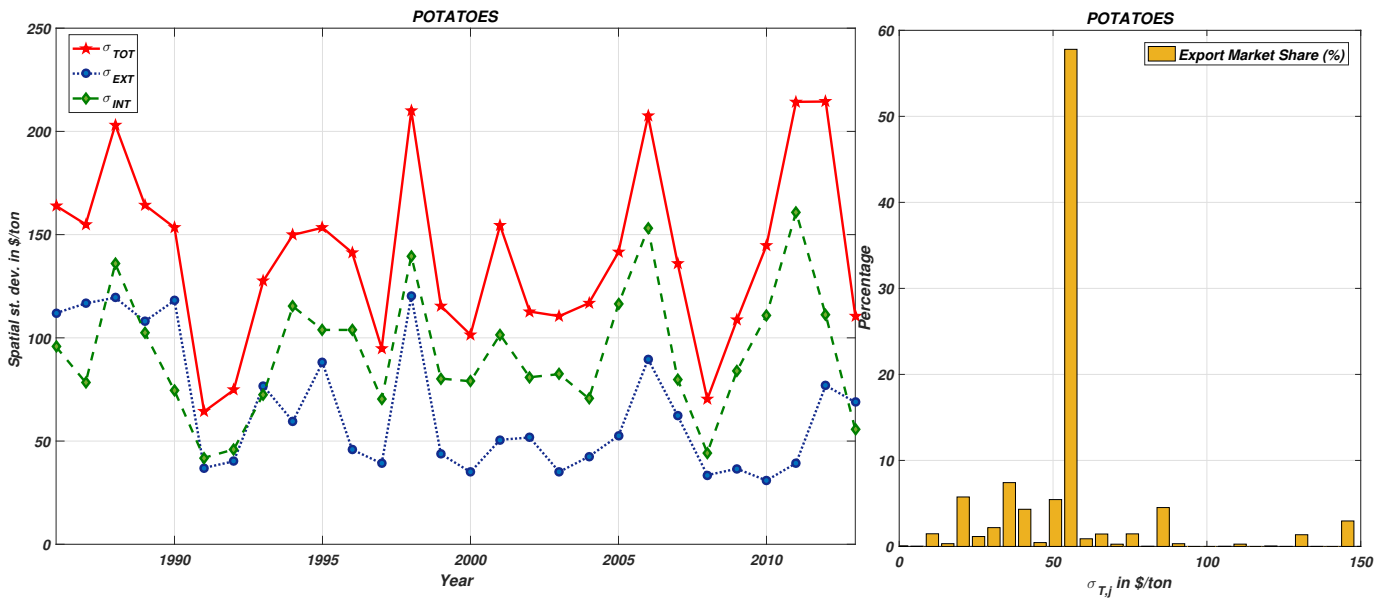
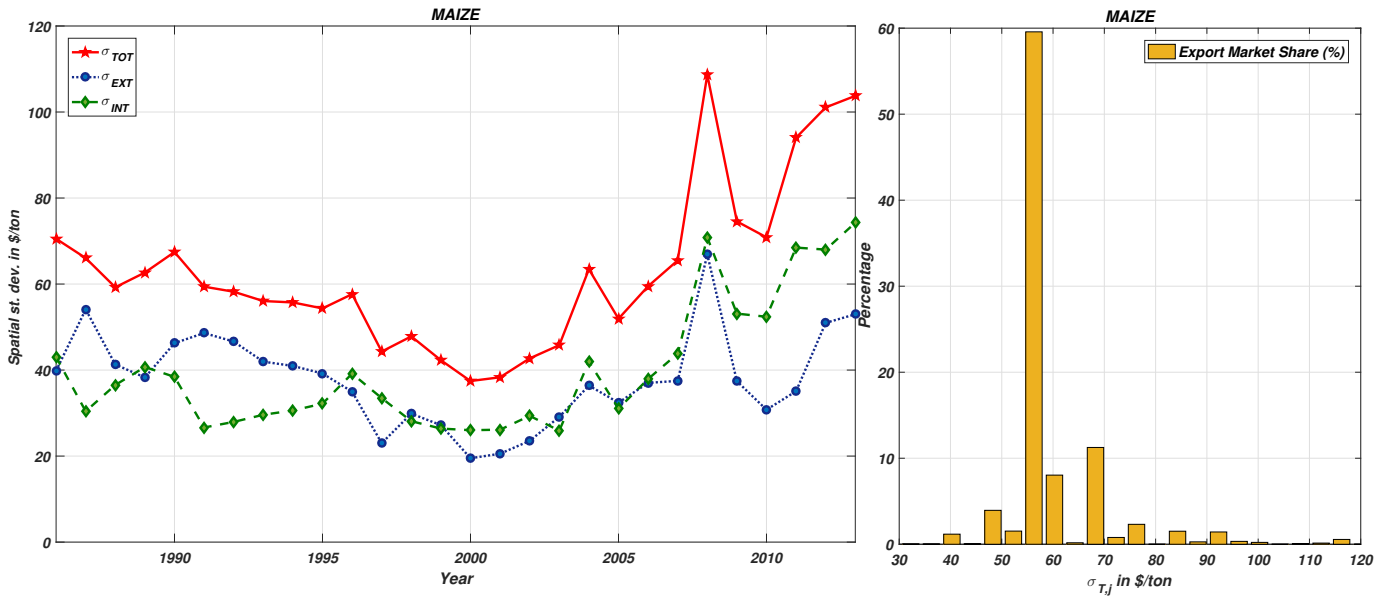




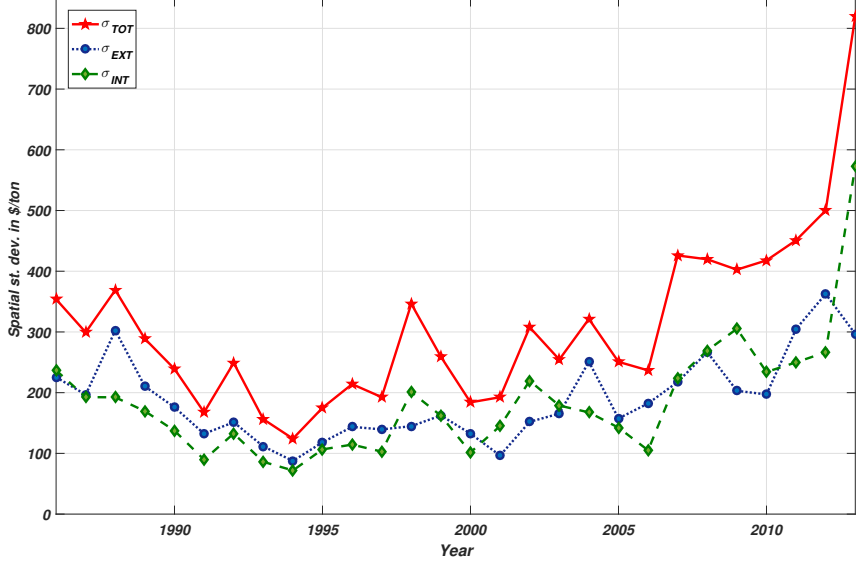
In what follows we show the empirical time series of the square roots of  $\sigma_{tot}^2$  (red line),  $\sigma_{int}^2$  (green line), and  $\sigma_{ext}^2$  (blue line).

The right panels show the empirical distribution of the weighted standard deviation in the average price set by each exporter in the whole time span. The importance of the exporter (i.e., total export  $F_j$ ) is given by its market share. The commodities are: rice, soy-beans, maize, potatoes, apples, eggs, honey, and cocoa beans.

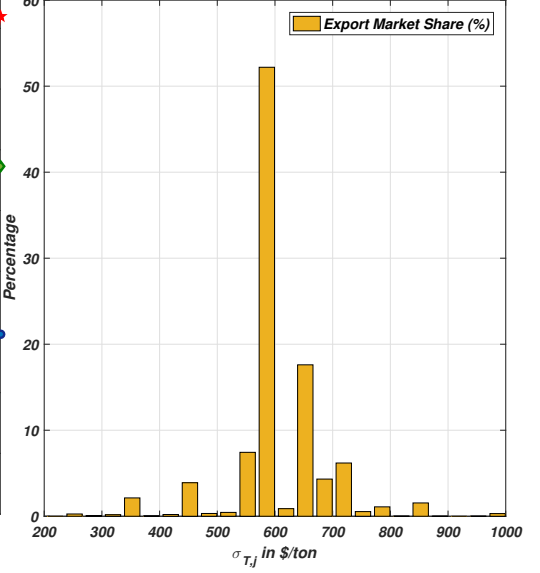




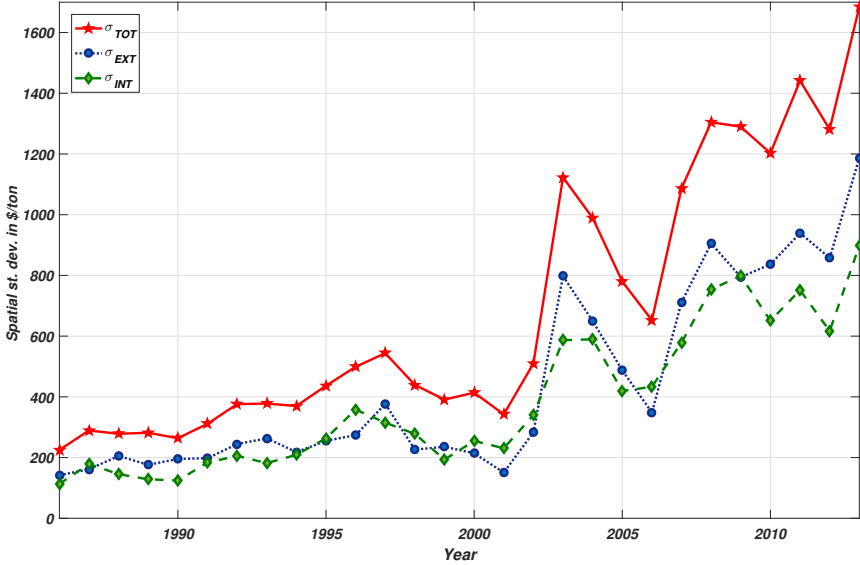
COCOA



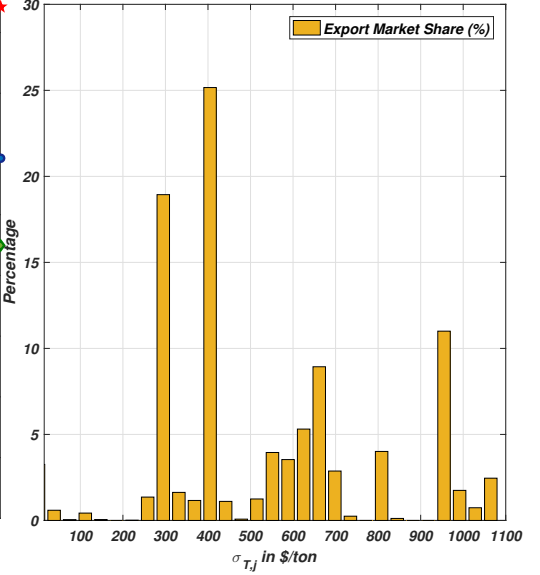
COCOA



HONEY



HONEY



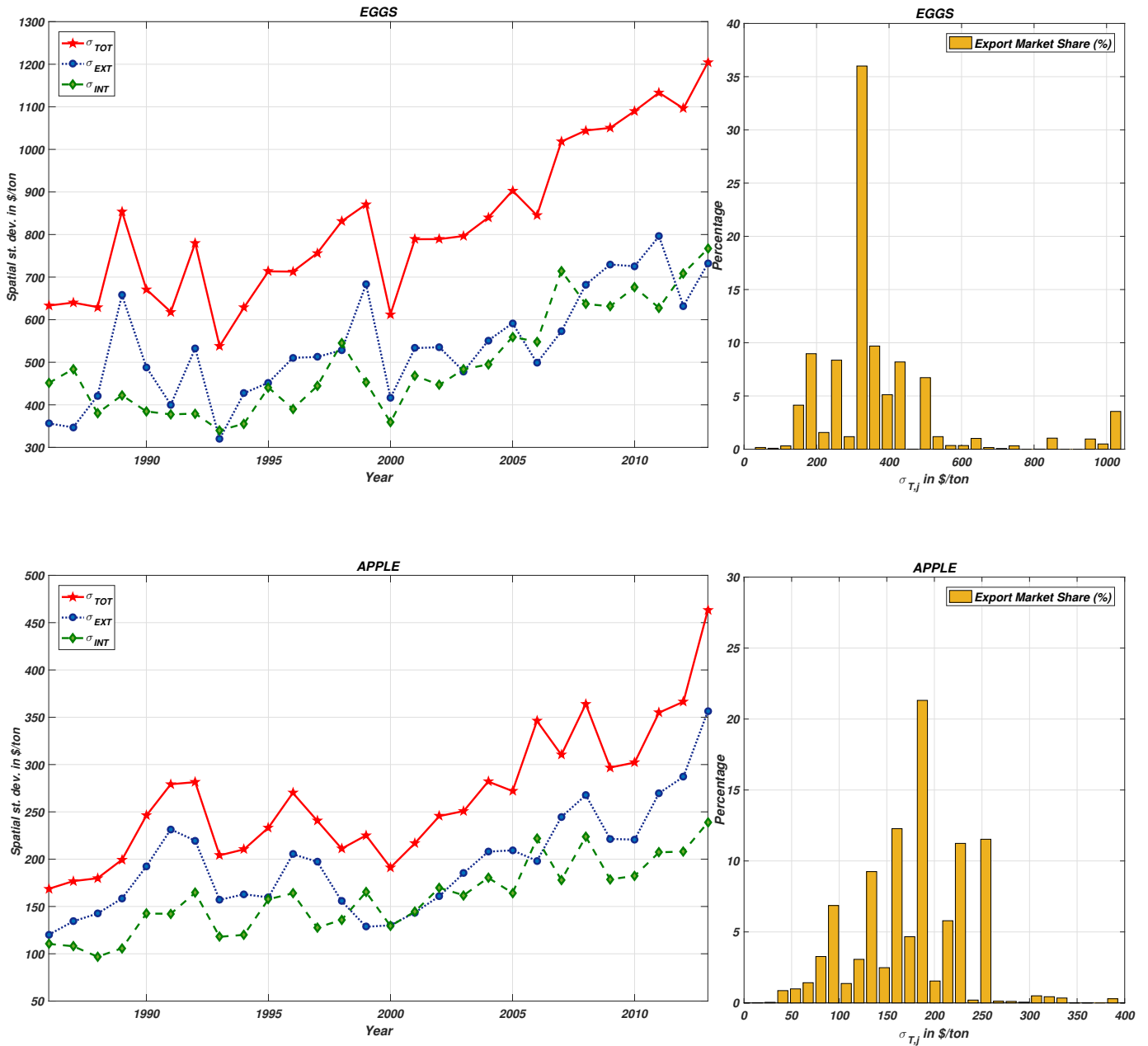


Figure A.4: Empirical time series of the square roots of  $\sigma_{tot}^2$  (red line),  $\sigma_{int}^2$  (green line), and  $\sigma_{ext}^2$  (blue line). The right panels show the empirical distribution of the weighted standard deviation in the average price set by each exporter in the whole time span. The importance of the exporter (i.e., total export  $F_j$ ) is given by its market share. The two commodities are eggs (top) and apple (bottom).