Title: Supplementary material for “Market structure, power and the Unfair Trading Practices Directive in the EU food sector: a review of indicators”

Authors: Kjersti Nes; Liesbeth Colen; Pavel Ciaian

Date: October 18, 2024

Note: The material contained herein is supplementary to the article named in the title and published in the *Agricultural and Resource Economics Review*.

Selection of Member States

In the data section, we selected Member States to be included in the paper. In addition to the exclusion criteria mentioned in section 2.1, Czechia and Hungary were dropped from estimating markups due to missing information on input costs[1]. Bulgaria was dropped due to concerns about the quality of the input cost data. The final set of 7 Member States included for estimating the markups is as follows: Spain, France, Italy, Portugal, Romania, Finland and Sweden.

Coverage of food sector by the Orbis database

The information in Table A1 is from 2016 for all selected EU Member States except France. France has a lower level of data availability for the more recent years than the rest of the Member States considered. For this reason the data for France are from 2014. As expected, for most cases, the coverage of the number of firms is lower than the coverage of the total turnover, which is the result of smaller firms not reporting turnover data or not being included at all in the Orbis database. This pattern is particularly prominent in the retail sector. For example, only 8 % and 12 % of food retail firms in Spain and Czechia, respectively, have turnover reported in the Orbis database. In terms of the total turnover in food retailing, coverage in Orbis in both of these Member States is still quite substantial, at 76 % and 86 %, respectively.

As the data availability differs by year and country, the time periods used for the markup estimations differed. Table A3 shows the years available and the total observations count by country. Similar years are available for the HHI calculations. The exceptions are Sweden and France, where the data extend to 2017 for Sweden and 2014 for France. Although the time periods used for the estimations differ slightly, we do not think this difference will affect our cross-country comparison, as the median markups and distributions did not vary much

[1] For instance, in Hungary, more than 47 % of firms have turnover information; however, less than 9 % of the manufacturing firms have input information, which is needed for estimating markups. Hungary is therefore excluded from the markup estimations.

Table A1: Share of food sector covered by the Orbis database, by sector and Member State

	Number of firms			Percentage of sector turnover		
	Manufacturing	Wholesale	Retail	Manufacturing	Wholesale	Retail
(a) *Turnover and HHI index calculations*						
BG	97%	97%	98%	92%	96%	101%
CZ	137%	50%	12%	103%	93%	86%
ES	48%	42%	8%	87%	85%	76%
FI	109%	105%	101%	118%	79%	89%
FR	39%	61%	42%	81%	91%	84%
HU	68%	83%	47%	101%	98%	87%
IT	57%	72%	56%	96%	87%	84%
PT	49%	55%	25%	94%	94%	122%
RO	122%	123%	117%	102%	96%	98%
SE	58%	56%	57%	96%	96%	96%
(b) *Markup estimations*						
ES	45%	37%	8%	87%	82%	75%
FI	51%	53%	42%	111%	76%	78%
FR	26%	40%	29%	75%	79%	76%
IT	20%	30%	7%	92%	79%	66%
PT	43%	46%	23%	93%	92%	122%
RO	78%	66%	56%	102%	94%	96%
SE	39%	29%	28%	58%	44%	28%

Source: Orbis 2019 and Eurostat. Note: The percentage indicates the share of the sector, in total number of firms (left), or in total sector turnover (right), that is covered by firms included in the Orbis database, relative to the total economy of the sector as covered by Eurostat, for the year 2016 for all countries, except France. Due to data quality the year covered for France is 2014. Part (a) of the table indicates the share of firms covered by the Orbis database for which turnover information is available, allowing for the calculation of HHI; Part (b) indicates the share for which more detailed firm-level information is available, allowing for markup estimations. Some of the numbers in the Table exceed 100%, meaning that the Orbis database contains more firms or covers a greater aggregated turnover than Eurostat. This may be due to differences in reporting or deflating of the turnover variable.

during the time period.

Methodology for mark-ups estimations

The estimation of the markups follows the method developed by De Loecker and Warzynski (2012), who derive the relationship between mark-ups (μ), expenditures on input x (α^x),

Table A2: Overview years included in the estimation, by country

Country	Years	Total observations
Spain	2006-2017	455,427
Finland	2007-2017	23,868
France	2006-2013	400,504
Italy	2008-2017	281,779
Portugal	2007-2017	174,073
Romania	2009-2014	174,570
Sweden	2008-2015	31,746

and the output elasticity with respect to input x (θ^x) from a standard cost minimisation problem, as follows:

$$\mu = \theta^x (\alpha^x)^{-1} \tag{1}$$

The output elasticity, θ^x, is estimated using the method suggested by Ackerberg, Caves and Frazer (2015), which is referred to as ACF. Similarly to the De Loecker and Warzynski (2012) method, a translog production function is assumed, which is defined as:

$$y_{it} = \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{lk} k_{it} l_{it} + \beta_{km} m_{it} k_{it} + \beta_{ml} l_{it} m_{it} + \omega_{it} + \epsilon_{it} \tag{2}$$

Here, y_{it} is the log of output, k_{it} is the log of capital input, m_{it} is the log of material input, and l_{it} is the log labour input. ϵ_{it} is a random production shock unobserved by the firm. Note here that material cost, m_{it}, is a flexible input where the input decision is taken at time t. The labour decision is assumed to be decided either at the same time as m_{it} or before. k_{it} is considered to be inflexible, and the firm decides how much capital, k_{it}, to use at time $t-1$. ω_{it} is unobserved in the data, but predicted or observed by the firm (for instance, a productivity shock). As such, ω_{it} is correlated with the input choices.

Following the ACF method, ω_{it} and k_{it} are known to by the firm when making the input decision about m_{it} and the decision on l_{it} is taken either before or the same time as m_{it}.

Thus, m_{it} can be written as an expression of l_{it}, ω_{it}, and k_{it}, as:

$$m_{it} = f(l_{it}, k_{it}, \omega_{it}) \tag{3}$$

Assuming strict monotonicity, equation (A.3) can be inverted as:

$$\omega_{it} = f^{-1}(l_{it}, k_{it}, m_{it}) \tag{4}$$

The production function can now be re-written as:

$$\begin{aligned} y_{it} = \beta_l l_{it} + \beta_k k_{it} + \beta_m m_{it} + \beta_{ll} l_{it}^2 + \beta_{kk} k_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{lk} k_{it} l_{it} + \beta_{km} m_{it} k_{it} + \\ \beta_{ml} l_{it} m_{it} + f^{-1}(l_{it}, k_{it}, m_{it}) + \epsilon_{it} \end{aligned} \tag{5}$$

Note here that the functional form of f^{-1} is unknown, and we instead use a third-order polynomial expansion of f^{-1} with respect to l_{it}, m_{it}, and k_{it}. As argued in ACF, as f^{-1} is also an expression of polynomials of the input choices, it is not possible to separate out the effect of the coefficients in the the production $(\beta_k, \beta_l, \beta_m, \beta_{ll}, \beta_{mm}, \beta_{kk}, \beta_{kl}, \beta_{lm}, \beta_{km})$ from coefficients in $f^{-1}(l_{it}, k_{it}, m_{it})$. Instead ACF proposed that due to to the law of motion of productivity, ω_{it} can also be written as an expression of $\omega_{i,t-1}$ and γ_{it}, where γ_{it} is innovation of productivity:

$$\omega_{it} = g(\omega_{i,t-1}) + \gamma_{it} \tag{6}$$

Based on pervious assumptions, γ_{it} is independent of the input choices and the moment conditions can be written as:

$$\frac{1}{N}\frac{1}{T}\sum_i \sum_t [\hat{\gamma_{it}}(\hat{\beta}) k_{it} \hat{\gamma_{it}}(\hat{\beta}) l_{it} \quad \hat{\gamma_{it}}(\hat{\beta}) m_{i,t-1} \hat{\gamma_{it}}(\hat{\beta}) k_{it} k_{it} \quad \hat{\gamma_{it}}(\hat{\beta}) l_{it} l_{it} \hat{\gamma_{it}}(\hat{\beta}) m_{i,t-1}^2 \quad \hat{\gamma_{it}}(\hat{\beta}) l_{it} k_{it} \hat{\gamma_{it}}(\hat{\beta}) k_{it} m_{i,t-1} \\ \hat{\gamma_{it}}(\hat{\beta}) l_{it} m_{i,t-1}]^T = 0 \tag{7}$$

From the Orbis database, we used operating turnover as a indicator of turnover, material cost as an indicator of material cost, tangible fixed assets as an indicator of capital, and cost of employment for the labour cost variable. The variables — output, labour, capital, and material — are deflated, based on the sources given in Table A3.

Outliers based on input shares were dropped before estimating the markups. We dropped observations where the labor input share or material input share was greater than 100%.[2] In addition, we also dropped observations where the labor input share or the material input share was less than 0.5%.

For each Member States and 3-digit sector, the markups were estimated following a 2-stage approach:

- **First stage:** The equation (A.5) is estimate using OLS and third-order polynomial expansion of f^{-1} with respect to l_{it}, m_{it}, and k_{it}. Define equation (A.5) as $y = \phi + \epsilon_{it}$. An expression for the unobserved productivity shock can now be deduced as $\hat{\omega_{it}} = \hat{\phi} - \beta_k k_{it} + \beta_l l_{it} + \beta_{kk} k_{it}^2 + \beta_{ll} l_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{kl} k l_{it} + \beta_{lm} l m_{it} + \beta_{km} k m_{it}$.

- **Second stage:** Utilising the estimated values on $\hat{\omega_{it}}$ from stage 1, equation (A.4) is estimated as $\hat{\omega_{it}} = \alpha_1 \hat{\omega_{it}} + \alpha_2 \hat{\omega_{it}}^2 + \alpha_3 \hat{\omega_{it}}^3 + \gamma_{it}$. From this expression, the expression γ_{it} is obtained. This expression is used to solve for the moment condition expressed in equation (A.7).

In deducing the $\hat{\omega_{it}}$ in the first stage, $\beta_k, \beta_l, \beta_m, \beta_{ll}, \beta_{mm}, \beta_{kk}, \beta_{kl}, \beta_{lm}, and \beta_{km}$, are not yet identified. In order to do so, a minimising program is run — starting from the OLS estimates of the production function — which finds the combination of $\beta_k, \beta_l, \beta_m, \beta_{ll}, \beta_{mm}, \beta_{kk}, \beta_{kl}, \beta_{lm}$, and β_{km} that minimise the moment conditions in the second stage.

[2]Note that labour input share is defined as (labour cost)/(turnover)*100 and material input share is defined as (material cost)/turnover*100.

Table A3: Overview of deflator sources

Source	Aggregation level	Variables	Countries
OECD STAN	NACE 2-digit	Material deflator: Intermediate inputs, deflators	FR, FI, SE, IT, PT
	NACE 2-digit	Capital deflator: Gross Fixed Capital Formation, deflator	FR, FI, SE, IT, PT, ES
	NACE 2-digit	Output deflator: Production (Gross Output), deflators	FR, FI, SE, IT, PT
European Central Bank	NACE 2-digit	Labour deflator: Unit Labour Cost (based on hours worked), deflator	FR, FI, SE, IT, PT, ES
	Total economy	Labour deflator: Unit Labour Cost (based on hours worked), deflator	RO
Eurostat	NACE 2-digit for manufacturing	Output deflator manufacturing and input deflator for retail: Output deflator for manufacturing	ES, RO
Eurostat	NACE 2-digit for agriculture	Input deflator manufacturing: Used output deflator from agriculture	ES, RO
World Bank	Retail sector	Output deflator: CPI for food products	ES, RO

The markups are calculated from the prior estimations output elasticity with respect to material. In particular, this output elasticity can be derived from Equation (A.2) as:

$$\theta_{it}^{m} = \hat{\beta}_m + 2\hat{\beta}_{mm} m_{it} + \hat{\beta}_{lm} l_{it} + \hat{\beta}_{km} k_{it} \tag{8}$$

Using the calculated output elasticities, the final markups are calculated as:

$$\hat{\mu}_{it} = \hat{\theta}_{it}^{m} \left(\frac{m_i t}{\frac{\hat{y}_{it}}{exp(\epsilon_{it})}} \right)^{-1} \tag{9}$$

Likewise, the log(TFP) can be calculated from the estimated results, given by equation:

$$log(TFP) = \hat{y_{it}} - \hat{\beta}_l l_{it} - \hat{\beta}_k k_{it} - \hat{\beta}_m m_{it} - \hat{\beta}_{ll} l_{it}^2 - \hat{\beta_{kk}} k_{it}^2 - \hat{\beta}_{mm} m_{it}^2 - \hat{\beta}_{lk} k_{it} l_{it} - \hat{\beta}_{km} m_{it} k_{it} - \hat{\beta}_{ml} l_{it} m_{it} \tag{10}$$

Since various subsectors are expected to differ in their production function, the production function is estimated at the most detailed level of disaggregation possible while taking into consideration data limitations. Subsectors corresponding to four-digit NACE codes for the manufacturing of beverages were considered, with some adjustments: due to the limited number of observations, NACE codes the manufacturing of beer (11.05) and manufacturing of malt (11.06) have been merged, and manufacturing of wine (11.02) has not been considered for Finland and Sweden. For all other manufacturing subsectors, and for the retail and wholesale sectors, the production functions are estimated at the level of three-digit NACE codes. Markups are not calculated for the manufacturing of tobacco (code 12.0) due to an insufficient number of data points.

Estimated parameters for markups estimation

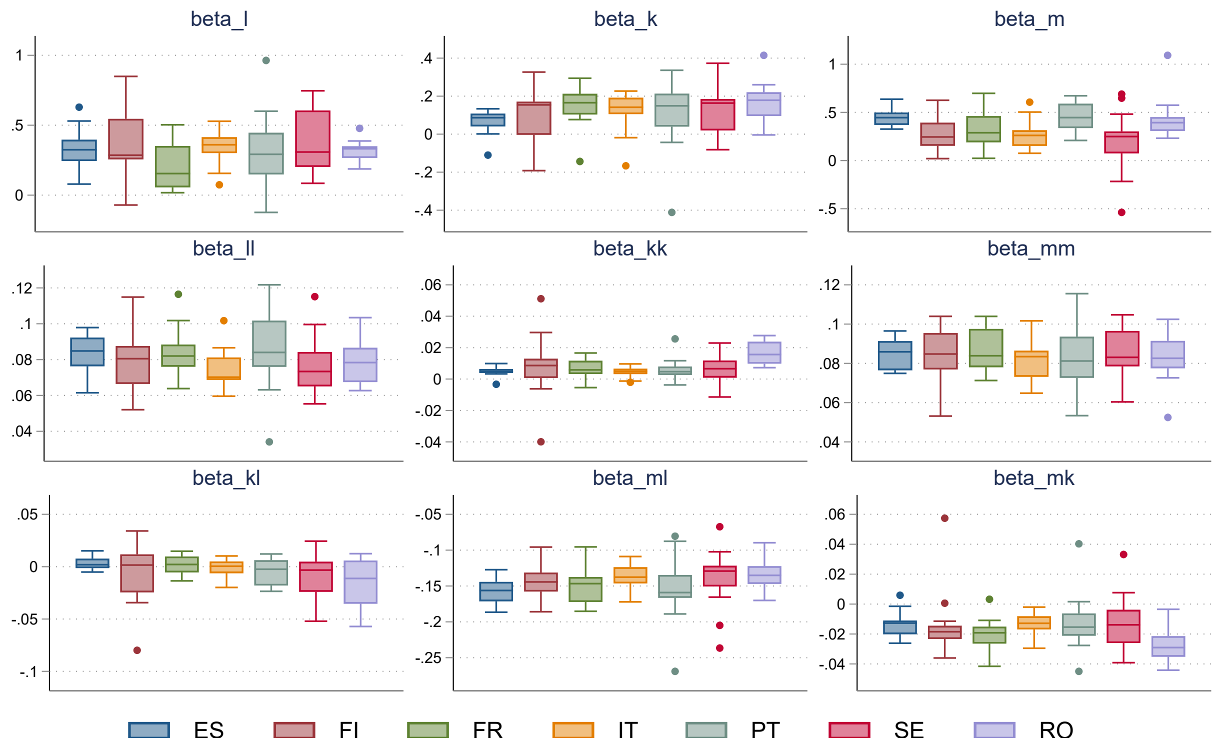
The estimated parameters from the translog production function following the approach in AW are provided in Figure A1. As different technology may entail different production functions, ACF method is estimated for each country and subsector. Figure A1 shows the box plot of estimated coefficients, averaged over subsector and country. Except for a few exceptions, the coefficients on β_k, β_l and β_m — as shown in the upper panel of subfigure A of Figure A1 — are positive and less than 0.5. Likewise, in the middle panel, the coefficients on β_{kk}, β_{ll} and β_{mm} mostly vary between 0 and 0.1, whereas β_{kl}, β_{ml} and β_{mk} are negative, except for some cases of β_{kl}.

The distributions of output elasticities with respect to material calculated from the estimated parameters are seen in subfigure b of Figure A1. The distributions from manufacturing firms are shown to the left of the Figure, while the distributions of retail and wholesale are seen to the right of the Figure. As expected, the calculated elasticities fall between 0 and 1, with some few exceptions that are greater than 1 or less than 0. The output elasticities in both panels follow a normal distribution. For most countries, the peaks of the distributions imply a higher output elasticity in the retail and wholesale sector compared with the manufacturing sector.

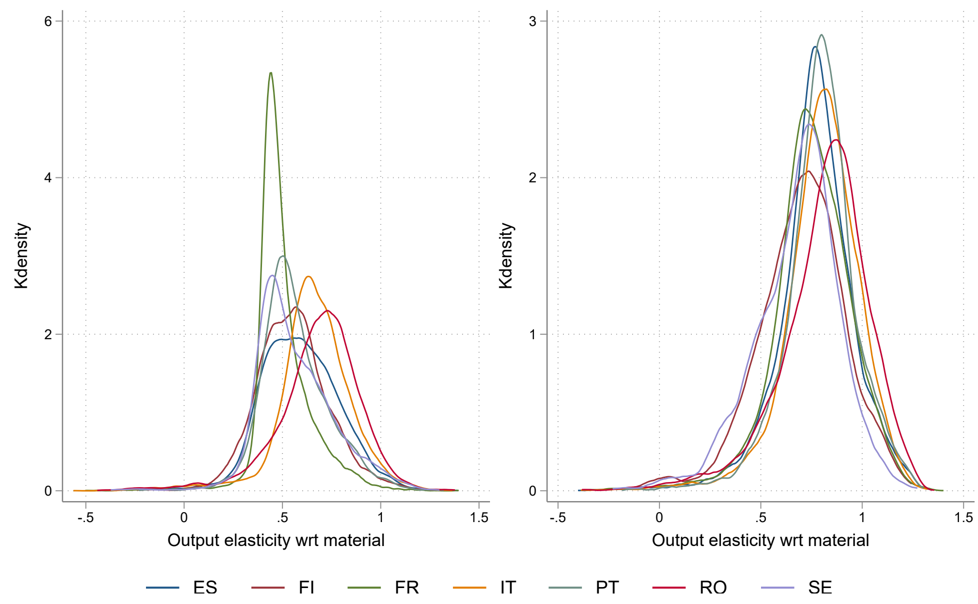
The distributions also vary across countries, within industries. For instance, in the manufacturing subsectors, France has a more narrow distribution compared with the other countries, implying a lower level of variation of output elasticities. The Romanian manufacturing distribution is skewed to right compared with the other countries, implying a greater average output elasticity compared with the other countries.

All-in-all, these results imply industry differences in output elasticities comparing manufacturing to retail and wholesale, with some country-level differences in the distribution of the elasticities.

We perform two tests comparing the distributions of the markups. First, we perform a two-sample variance-comparison test to examine if the variances of the distributions vary by industry. In this test, the null hypothesis is whether the ratio of the standard deviations of the two industries is equal to one. The p-value of this F-test is given in the upper part of Table A4. Column *Manufacturing-Wholesale* tests whether the ratio of the standard deviation of the manufacturing and wholesale industries is equal to 1. Likewise, the columns



(a) Estimated parameters from the production function



(b) Distribution of output elasticities wrt material

Figure A1: Parameters estimated from the ACF production function. *Source*: Authors' calculations, based on Orbis database (BvD, 2019).

Table A4: Two sample variance-comparison and Kolmogorov-Smirnov p-value test results

	Manufacturing - Wholsale	**Manufacturing - Retail**	**Retail - Wholesale**
		Two-sample variance-comparison	
Total	0.000	0.000	0.002
ES	0.000	0.000	0.083
IT	0.000	0.000	0.000
SE	0.000	0.000	0.000
FR	0.000	0.000	0.000
FI	0.000	0.000	0.000
RO	0.000	0.000	0.000
PT	0.000	0.000	0.091
		Two-sample Kolmogorov-Smirnov	
Total	0.000	0.000	0.000
ES	0.000	0.000	0.000
IT	0.000	0.000	0.000
SE	0.000	0.000	0.175
FR	0.000	0.000	0.000
FI	0.000	0.000	0.000
RO	0.000	0.000	0.000
PT	0.000	0.000	0.000

Manufacturing-Wholesale and *Retail-Wholesale* provide the p-values comparing the variance within these sectors. We reject the null hypothesis that the ratio of the two distributions is equal to 1 for $p<0.05$ for all industry-country combinations. In two cases, retail-wholesale industries in Spain and Portugal, we fail to reject the null hypothesis for $p<0.10$. We can conclude that the test supports that the variance of the distributions is different across countries and industries.

Second, we perform a Kolmogorov-Smirnov test to examine if the distributions of markups differ across industries. The null hypothesis of the Kolmogorov-Smirnov test is that the distribution of markups within two groups – here defined by industries – are the same and the corresponding p-values for the test are provided in the lower part of Table A4. The test results show that the null hypothesis is rejected in all but one case; i.e., we can conclude that the distributions of markups across industries differ in all but one case. The only exception is that the null hypothesis that the distribution of the wholesale and retail industries is the

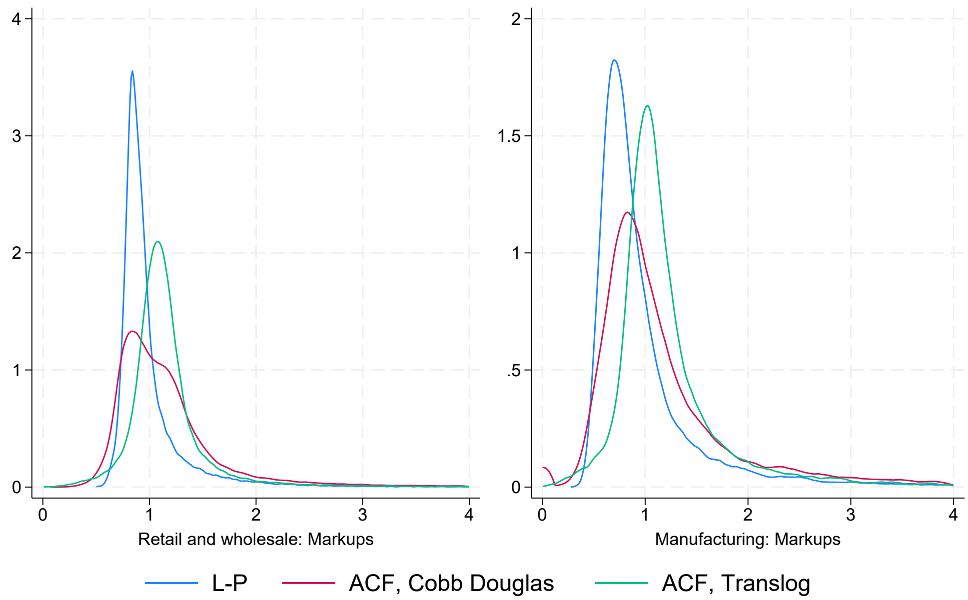


Figure A2: Markups distrubution with alternative production function estimations. *Source*: Authors' calculations, based on Orbis database (BvD, 2019). Note: The year used in the Figure is 2015 for all countries, except for France and Romania, where the estimates are calculated for the years 2014 and 2013, respectively.

same cannot be rejected in Sweden.

Figure A2 shows markup distributions based on alternative estimation techniques for the production functions. In the graphs, our preferred specification is given by ACF translog. Two alternative specifications are considered: i) ACF approach with a Cobb-Douglas production function and ii) the approach suggested by Levinsohn and Petrin (2003), referred to in the Figure as LP. In both the retail and manufacturing industries, our preferred approach—ACF, translog—yields distributions with higher average markups. Estimates based on the Cobb-Douglas production function yield smaller markups for both the LP and ACF method. The main difference between these two approaches is that the approach by ACF Cobb-Douglas tends to have greater variance compared to the method suggested by Levinsohn and Petrin (2003).

.1 Robustness regressions

Table A5: Robustness for the regression results

	HHI		Markups	
	Median	Top 10	Median	Top 10
VARIABLES	(1)	(2)	(3)	(4)
Ln(Turnover)	0.0213	-0.2782***	0.00405	0.0205
	(0.0758)	(0.0297)	(0.0121)	(0.0194)
Ln(TFP)			-0.0415***	-0.0349***
			(0.0120)	(0.0115)
Constant	6.4855***	11.6723***	0.218	0.0515
	(0.948)	(0.5288)	(0.178)	(0.355)
Observations	1,934	1,934	1,187	1,165
R-squared	0.169	0.329	0.178	0.127

Standard errors are clustered at MS-subsector level. Fixed effects included are the MS-year interaction.
*** p<0.01, ** p<0.05, * p<0.1

Agricultural subsectors

In this section of the appendix, we provide further details on two of the assumptions about comparing the concentration and turnover with the agricultural subsectors.

First, we assume that most farms have turnover less than EUR 2 million and thus are below the lower threshold for the UTP directive. The Orbis database contains a low level of coverage of agricultural farms. However, Eurostat provides data on the number of farms in various economic sizes. The size categories in Eurostat do not perfectly match the size categories in the UTP directives. The Eurostat category with the largest farms by turnover is farms with more than EUR 0.5 million in turnover, while the UTP directive only applies to firms with more than EUR 2 million. Table A6 shows the percentage of farms with more than EUR 0.5 million in turnover, based on the Eurostat data. The percentage of larger farms depend on the Member State, with the range of 7.27% in the Czech Republic to 0.05% in Romania. Accordingly, we can infer that the percentage of firms with turnover more than

EUR 2 million is small in all the Member States, and that the assumption that most firms are covered by the UTP directive is reasonable.

Table A6: The percentage of farms with standard output (turnover) greater than EUR 0.5 millions in 2016

MS	Percentage	Member States	Percentage
Bulgaria	0.74%	Hungary	0.40%
The Czech Republic	7.27%	Portugal	0.41%
Spain	1.17%	Romania	0.05%
France	4.26%	Finland	1.87%
Italy	1.14%	Sweden	3.53%

Second, we limit the analysis to subsectors buying directly from farmers. The NACE codes corresponding to these subsectors are shown in Table A7. As a robustness check, we aggregate the NACE codes to reflect a more realistic buyers for agricultural goods. For instance, since fruit and vegetables farmers can sell products to both the fresh market (wholesale) or for processing (manufacturing of fruit and vegetables), we assume that these two subsectors both represent buyers for these farmers. Table A7 shows aggregation of the NACE codes into the various subsectors.

Table A7: Aggregated agricultural subsectors

Subsector	NACE
Animal feed manufacturing	109
Beer malt manufacturing	1105; 1106
Dairy manufacturing	105
Edible oils manufacturing	104
Fruit and veg. wholesale/manufacturing	102; 1103; 4631
Fish manufacturing	1020
Flowers plants wholesale	4622
Grain manufacturing	1061
Live animal wholesale	4623
Meat manufacturing	101
Sugar manufacturing	1081
Tobacco manufacturing	120
Wine manufacturing	1102

References

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer. 2015. "Identification properties of recent production function estimators." *Econometrica*, 83(6): 2411–2451.

De Loecker, Jan, and Frederic Warzynski. 2012. "Markups and firm-level export status." *American economic review*, 102(6): 2437–71.

Levinsohn, James, and Amil Petrin. 2003. "Estimating production functions using inputs to control for unobservables." *The Review of Economic Studies*, 70(2): 317–341.