

ESTIMATION OF IMPACT OF PARTICIPATION IN PUBLIC GREEN PROCUREMENT OF FOOD ON FARM NUMBERS IN RURAL LATVIAN COUNTIES

Juris Hazners¹, Andra Zvirbule²

¹Institute of Agricultural Resources and Economics, Latvia;

²Latvia University of Life Sciences and Technologies, Latvia
juris.hazners@arei.lv, andra.zvirbule@lbtu.lv

Abstract. Green Public Procurement (GPP) of food in Latvia is a tool that directly provides the increase in sales of local food products via purchases by local communities and state institutions in sectors such as education, defence, interior, healthcare and welfare. Public procurement within the framework of GPP ensures that the purchase of the food products or catering services minimizes the environmental impact while having positive social consequences. The share of local farms in the total directly provided product volumes within GPP is minor and the bulk of the farm produce goes through wholesale, food processing and public catering companies. The Latvian agriculture is experiencing continuous structural changes with the consolidation and decrease in farm numbers. The rate of the farm “disappearance” in rural counties with companies participating in the food GPP is similar to the counties without such companies. This could lead to the expectations that changes in the farm numbers are not affected by participation in GPP. While the previous research on the impact of GPP in the European countries has extensively focused on the environmental, economic and social issues, the causal inference with respect to changes in farm numbers has not been addressed. Coarsened exact matching (CEM) is the method for the estimation of the causal effects that has been widely applied to observational data. The advantages of CEM are associated with the reduction of the differences between multivariate distributions of covariates in treatment and control groups by pruning the data with splitting into pre-determined bins. The research objective is to evaluate the net impact of the participation in GPP by a rural county on the number of farms in this county. Considering the participation in GPP as a treatment variable in the rural county panel data, the results of CEM reveal positive net impact on the farm numbers in the treated counties. Albeit the direct farm participation in food GPP is negligible, the participation of other companies in food GPP has positive net direct impact on farm numbers. Participation in a food GPP over the period from 2018 to 2020 has contributed to sustainability of 2179 farms. Considering that only about 21% of rural communities have direct participation in food GPP, a modification of the National regulation of GPP towards an increased preferences for local suppliers would improve the resilience of the farming sector.

Keywords: Green Public Procurement, food, coarsened exact matching, rural communities, farm numbers.

Introduction

Since the initiation in 2017, the food Green Public Procurement (GPP) in value terms in Latvia has been unstable. According to the publicly available information by the Procurement Supervision Office (IUB), the average annual food GPP stands at around EUR 145 million [1]. The food GPP in Latvia can be divided into two broad categories - procurement of food and procurement of catering services. On the average, the share of catering services procurement fluctuates around 60% of total food GPP. Specialized catering companies provide around 85% of these services. Food wholesalers and processors provide around 50% and 35%, respectively, of food procurement. Direct participation of farms in food GPP either in sole or combined contracts is negligible, as only 46 farms have been engaged in contracts over the period from 2018 to 2020. On the demand side, educational establishments have 73% share in total food GPP followed by the armed forces with 17% share.

The role of GPP in Latvian food market often is overestimated. In a study commissioned by a major local retail network [2], in 2019 the total retail food market size is estimated at EUR 3.30 billion. The annual turnover of the public catering sector in 2019 adds EUR 0.54 billion [3]. Total food GPP reached EUR 68 million in 2019. Adding these three numbers yields the total Latvian food market size in 2019 at EUR 3.912 billion. The total food GPP in 2019 amounted to EUR 152 million. Hence, the share of food and catering GPP in total the Latvian food market stands at about mere 3.9%. Moreover, the market growth in the segment is limited by mostly unchanged allowances for daily meals in kindergartens and primary schools, hospitals and care homes, prisons and armed forces.

The previous rather scarce research on the Latvian food GPP is mostly observational and it lacks the estimation of possible causal inferences between the participation in food GPP and selected socio-economic variables. Simanovska et.al. [4] regard the existing national legal framework regulating on GPP as rather vague with a broad range of equally rated criteria. Hence, often less environmentally beneficial options are chosen. Only in a fraction of tenders so called higher quality products are preferred

including organic produce or products from integrated farming. National quality schemes are also shunned. Zvaigzne et.al. (2018) [5] consider GPP an instrument contributing to an increase in local food sales, jobs and business opportunities for small and medium enterprises. At the same time, in some areas food GPP is provided almost entirely by large established wholesalers. Analysing GPP in general without emphasis on food procurement, Pelša [6] stresses the importance of municipal purchases as promoting a sustainable consumption. However, such suggestion has its caveats. Municipalities tend to neglect small farm suppliers in favour of mid-term large contracts either in food or catering procurement thus creating certain advantages for established large wholesalers and caterers. In other countries, the advantages of an opposite approach are described by Mensah and Karriem (2021) [7]. South African school feeding programmes with home-grown supplies assume decentralised catering model with schools purchasing and preparing meals via budgets deposited into their accounts. This makes a vital contribution to sustainable rural development. Cervantes-Zapana et.al. (2020) [8] regard institutional or public food procurement programs from family farming (PP-FF) as beneficial to increase in income, productivity, providing price support and market inclusion in Latin American and Caribbean countries.

According to the information provided by the Eurostat [9], the number of farms in Latvia has declined from 128.6 thousand in 2005 to 69.1 thousand in 2016. Hence, the average annual decline in farm numbers over the 2005-2016 period stands at about 4%. Breaking farms in two sets by their economic size yields about 5% average annual decline in farm numbers in the smaller farm set while for larger farms there is an 11% increase. The trend changes over the period from 2016 to 2020, when total farm numbers are virtually unchanged. Nevertheless, the structure in the smaller farm set has changed recording an increase in non-commercial subsistence and hobby farm numbers at the expense of commercial farms.

Coming to an end of the continuous decline in farm numbers coincides with the period when GPP in Latvia was introduced in 2017. This raises a question on the possibility to establish the causal inference between the participation in food GPP and farm numbers.

Materials and methods

The experimental set of local data provided by the National Data Website [10] is used to retrieve the information on rural communities in 2017. The information on the farm numbers in 2018 and 2001 is sourced from the Agricultural Census [11]. The data on participation in the food GPP are collected from publicly available website of the National Procurement monitoring office (IUB).

One of the most popular methods used for causal analysis in observational studies is propensity score matching (PSM) devised by Rosenbaum and Rubin [12]. They recommend to include as many as possible pre-treatment covariates for calculations of propensity score. However, King and Nielsen [13] argue that PSM often increases imbalance, inefficiency, model dependence and bias. Coarsened exact matching (CEM) was proposed by Iacus, King and Porro [14]. Iacus et. al. [15] insist that CEM generates matching solutions that are better balanced with estimates of the causal quantity of interest having lower root mean square error than methods based on propensity scores. Nevertheless, Black et.al. [16] state that CEM can produce very different results than the other methods and can produce full sample results that are inconsistent with subsample results. While it provides reasonable covariate balance, this comes at the cost of much smaller retained samples than other methods, and thus lower precision. Based on randomized clinical trials Guy et.al. [17] compare the performance of PSM and CEM. They find both methods leading to increased balance in pre-treatment baseline covariates, while retaining a majority of the original data. When there are few pre-treatment variables, CEM yields satisfactory results. As there are only seven pre-treatment variables available in the experimental set of local data, CEM was selected for the research purposes.

The participation by the data panel unit in food GPP is considered a binary treatment variable. As the participation in the scheme is not randomly assigned, there are differences in pre-treatment covariates between treatment and control groups. The measure for overall imbalance was proposed by Iacus, King and Porro (2008) [18]. It is based on the difference between the multidimensional histogram of pre-treatment covariates in the treatment and control groups. The first step in preparing the data is coarsening the pre-treatment covariates into bins. The bin sizes can be selected either by user or, in the case automated software is used, applied by functions enclosed in the software. One of the methods for determining bin sizes predominantly used in automated software is the Scott break method [19]. The

Scott's rule to choose bin sizes is based on the standard deviation of the data. The formula for establishing the bin length is:

$$S = 3.49\sigma n^{-1/3}, \quad (1)$$

where S – bin length in units,
 σ – standard deviation of the covariate,
 n – sample size.

The Rice's rule is simpler, with bin size calculated as doubled cubed root from the sample size.

The second step after coarsening the variables is forming the strata over the full data sample of coarsened variables. Each stratum contains the units having equal coarsened values for all k variables. Only strata with at least one unit in treatment group and control group are kept for further calculations. The other strata are dropped. The third step is calculations of overall multivariate imbalance, univariate imbalance for coarsened data samples, as well as difference in means for original data samples for all variables. The justification for further calculations is confirmed by decrease in multivariate imbalance, univariate imbalance and difference in means.

The overall imbalance for coarsened data panel is calculated as:

$$\mathcal{L}_1(f, g) = \frac{1}{2} \sum_{l_1, \dots, l_k} |f_{l_1, \dots, l_k} - g_{l_1, \dots, l_k}|, \quad (2)$$

where \mathcal{L}_1 – multivariate imbalance,
 f_{l_1, \dots, l_k} – k -dimensional relative frequencies for the treated units,
 g_{l_1, \dots, l_k} – k -dimensional relative frequencies for the control units.

The overall imbalance shows the similarity (or dissimilarity) between the distributions of selected pre-treatment covariates among treatment units and control units. All unique values for all covariates are put into the rows of the contingency table. For every row, in two columns frequencies for treatment and control groups are calculated. After that, the absolute values of the differences between frequencies in treatment and control columns are summed up. The calculated sum is divided by two.

The fourth step is the creation of weights for every stratum in the matched data panel. As the research goal is to evaluate the average treatment effect on treated (ATT), treatment units are unweighted getting the weight of 1 while control units get weighted:

$$w_t = 1, \quad (3)$$

$$w_c = \frac{n_t}{n_c} \times \frac{N_c}{N_t} \quad (4)$$

where w_t – weights for the treated units,
 w_c – weights for the untreated units,
 n_t – number of treated units within the stratum,
 n_c – number of untreated units within the stratum,
 N_t – number of treated units in the matched data sample,
 N_c – number of untreated units in the matched data sample.

The fifth step is the estimation of causal inference with the weighted ordinary least squares (Weighted OLS) regression. In a matrix form, the weighted OLS regression is performed by regressing the outcome variable of interest, Y on treatment variable T with some modifications in variable matrices:

$$B = (T^T W T)^{-1} T^T W Y, \quad (5)$$

where B (2×1) – matrix containing intercept β_1 and regression slope β_2 ,
 T ($n \times 2$) – matrix which first column is set to 1, the second column contains the binary treatment variable,
 W ($n \times n$) – matrix which diagonal contains the calculated weights, the other cells are set to zero,
 Y ($n \times 1$) – matrix containing the values of outcome variable,
 n – size of the matched data sample.

The variable of interest, β_2 shows the value of the ATT effect.

Results and discussion

As the automated software is not used in calculations, instead of the Scott break method (formula 1) a rather convenient and simplified approach is used for coarsening the pre-treatment covariates into bins. The variables are coarsened into quartiles. Thus, the coarsened variables can take values from 1 to 4. The original data panel of 504 units is divided into 203 strata containing 106 treated units and 398 untreated units. It means that in only about 21% of all rural communities there are companies participating in food GPP. After dropping the unsuitable strata, the pruned data panel contains 44 strata with 83 treated units and 163 untreated units. The calculated overall multivariate imbalance (formula 2), univariate imbalance for coarsened data samples, as well as the difference in means for original data samples for all variables are shown in Table 1.

Table 1

Multivariate \mathcal{L}_1 distance, univariate distances and differences in means before and after matching

Multivariate L1 distance	Before		After	
	0.238		0.103	
Pre-treatment variables	Univariate imbalance			
	L1	mean	L1	mean
Farms	0.152	30	0.056	18
Area	0.078	14	0.053	11
Population	0.270	1046	0.134	940
Employment	0.294	607	0.129	600
Vacancies	0.284	586	0.129	578
Production	0.301	6310998	0.108	2820293
Value added	0.289	5899096	0.108	5256744

As seen from the table, the multivariate imbalance, univariate imbalances in pre-treatment variables before and after matching along with the average differences in means of variables all have decreased. Hence, the research can proceed for the intended estimation of causal inference using the coarsened data panel. The weights for every panel unit using formulae (3, 4) are calculated and formed into a diagonal matrix. After that, the regression using formula (5) is performed. The regression results are shown in Table 2.

Table 2

Results of the weighted OLS regression

Variables	<i>B</i>	SE	<i>t</i>	<i>p</i>
Treatment	15.42	9.280	1.66	0.098*
Constant	134.45	5.390	24.94	0.000***
R2	0.011			
Adjusted R2	0.007			
F(1,261)	2.76			

Note: *** $p \leq 0.001$, ** $p \leq 0.05$, * $p \leq 0.1$.

As the statistical significance has been reached, the variable of interest, the coefficient for treatment variable at the value 15.42 shows the changes in farm numbers in a rural community caused solely by participation at least of one company located in this community in food GPP over the period from 2018 to 2020.

The estimated ATT effect at 15.4 farms shows the direct net impact on the changes in farm numbers over the three year period from 2018 to 2020. By dividing this number to three, the average expected annual change in farm numbers with the participation in food GPP is obtained - 5.1 farms.

The gross annual ATT effect at the national level is calculated by multiplying the estimated ATT annual effect by the number of communities that have participated in GPP from the original data panel – 106 and it equals about 545 farms. In other words, without the participation in GPP in every year from

2018 to 2020 there would be 545 less farms in Latvia. By multiplying the gross annual ATT to four, the gross ATT effect over the period from 2016 to 2020 is obtained (Fig. 1).

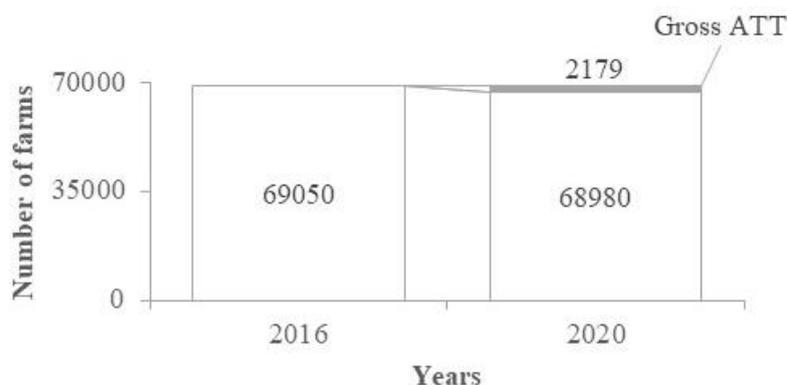


Fig. 1. Gross ATT effect from participation in food GPP on farm numbers in 2020

The calculated gross ATT effect at 2179 farms represents the number of farms that would have ceased the operation if at least one company in the respective community had not directly participated in food GPP in at least one year over the period from 2018 to 2020. In other words, the sustainability for about 3.2% of Latvian farms in 2020 is attributable solely to food GPP.

The broader discussion of the results obtained is limited by the scarcity of the research on the topic in EU. Therefore, it is not feasible to make cross-country comparisons, especially with neighbouring countries such as Estonia and Lithuania.

Conclusions

1. Coarsened exact matching (CEM) has to be considered a suitable method applied in regional studies for establishing a causal inference in longitudinal data panels using binary treatment variable.
2. The share of combined food and catering GPP in total Latvian food market stands at about mere 3.9% and it has to be considered a small and stagnant market segment.
3. While the direct farm participation in food GPP is negligible, the participation of other companies has a positive net direct impact on farm numbers.
4. The positive net average annual direct impact of participation in food GPP on farm numbers in a single rural community is around 5.1 farms.
5. The positive gross average annual impact of participation in food GPP on farm numbers in the whole country is around 545 farms.
6. Participation in a food GPP over the period from 2018 to 2020 has contributed to sustainability of 2179 farms. Considering that only about 21% of rural communities have direct participation in food GPP, a modification of the National regulation of GPP towards increased preferences for local suppliers would improve the resilience of the farming sector.

Author contributions

Conceptualization, J.H. and A.Z.; methodology, J.H.; software, J.H.; validation, J.H. and A.Z.; formal analysis, J.H. and A.Z.; J.H. and A.Z.; data curation, J.H. and A.Z.; writing - original draft preparation, J.H.; writing - review and editing, J.H. and A.Z.; visualization, J.H., Both authors have read and agreed to the published version of the manuscript.

References

- [1] Iepirkumu uzraudzības birojs (IUB) (Procurement monitoring bureau (PMB)). (In Latvian). [online] [07.05.2022]. Available at: <https://info.iub.gov.lv/lv/meklet>
- [2] Maxima mazumtirdzniecības kompass. Mazumtirdzniecības evolūcija Latvijā: No “akcijas preces” fenomena līdz mākslīgajam intelektam (Maxima retailing compass. Evolution of retailing in Latvia: from “on-sale” phenomenon to artificial intelligence). (In Latvian). [online] [25.08.2022]. Available at: https://www.maxima.lv/fileman/Uploads/2021/Mazumtirdzniecibas_kompass_septembris_2021.pdf.

- [3] Lursoft. Blogs (Lursoft. Blog). (in Latvian). [Online] [07.05.2022]. Available at: <https://blog.lursoft.lv/2019/06/13/47-edinasanas-sektora-stradajoso-uznemumu-pedeja-pusgada-laika-bijis-nodoklu-parads/>
- [4] Simanovska J., Ķibilda N. When is Green Really Green? A Latvian Case Study on Updating of Food Supply Green Public Procurement. *Environmental and Climate Technologies* 24(3), 2020, pp. 513-520.
- [5] Zvaigzne A., Krivašonoka I., Kotāne I. Factors Influencing Local Food Sales Through Green Public Procurement in Rezekne Municipality. *Selected Papers from the 17th International Conference on Reliability and Statistics in Transportation and Communication, RelStat'17, Riga, Latvia. Springer International Publishing, 2017, pp. 327-337.*
- [6] Pelša I. Green public procurement: case study of latvian municipalities. *Economics and Business*, 33(1), 2019, pp.207-221.
- [7] Mensah C., Karriem A. Harnessing Public Food Procurement for Sustainable Rural Livelihoods in South Africa through the National School Nutrition Programme: A Qualitative Assessment of Contributions and Challenges. *Sustainability*, 13(24), 2021, p.13838.
- [8] Cervantes-Zapana M., Yagüe J.L., De Nicolás V.L., Ramirez A. Benefits of public procurement from family farming in Latin-AMERICAN countries: Identification and prioritization. *Journal of Cleaner Production*, 277, 2020, p.123466.
- [9] Eurostat. Main farm indicators by agricultural area, type and economic size of the farm, share of consumed production, legal status of the holding and NUTS2 region. [Online] [07.05.2022]. Available at: https://ec.europa.eu/eurostat/databrowser/view/ef_m_farmleg/default/table?lang=en
- [10] Centrālā statistikas pārvalde. Teritoriju ekonomiskie rādītāji reģionos, novados, pilsētās, pagastos (atbilstoši robežām 2023. gada sākumā), apkaimēs un blīvi apdzīvotās teritorijās (eksperimentālā statistika) (Central statistical office. Economic indicators of territories in regions, counties, cities, parishes (corresponding to the boundaries at the beginning of 2023), neighborhoods and densely populated areas (experimental statistics)) (in Latvian). [Online] [07.05.2022]. Available at: <https://data.gov.lv/dati/lv/dataset/tea>
- [11] Centrālā statistikas pārvalde. 2001., 2010. un 2020. gada lauksaimniecības skaitīšanas. (Central statistical office. 2001, 2010 and 2020 Censuses of Agriculture) (in Latvian). [Online] [07.05.2022]. <https://stat.gov.lv/lv/statistikas-temas/noz/lauksaimn/cits/2975-2001-2010-un-2020-gada-lauksaimniecibas-skaitisanas>
- [12] Rosenbaum P.R., Rubin D.B. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70, 1983, pp. 41-55.
- [13] King G, Nielsen R. Why propensity scores should not be used for matching. *Political analysis*. 27(4), 2019, 99. pp. 435-54.
- [14] Iacus S.M, King G., Porro G. Causal inference without balance checking: coarsened exact matching. *Political Analysis* 20(1), 2011, pp. 1-24.
- [15] Iacus S.M, King G., Porro G. Multivariate Matching Methods That Are Monotonic Imbalance Bounding. *Journal of the American Statistical Association*, Volume 106, issue 493, 2011, pp. 345-361.
- [16] Black B., Lalkiya P., Lerner J. The Trouble with Coarsened Exact Matching. *SSRN Electronic Journal*. 10.2139/ssrn.3694749, 2020. 46 p.
- [17] Guy, D., Karp, I., Wilk, P., Chin, J., Rodrigues, G. Propensity score matching versus coarsened exact matching in observational comparative effectiveness research. *Journal of Comparative Effectiveness Research*, 10(11), 2021, pp.939-951.
- [18] Iacus S.M, King G, Porro G. Matching for causal inference without balance checking. Available at SSRN 1152391, 2008.
- [19] Scott D.W. Scott's rule. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4), 2010, pp. 497-502.