Bio-refineries - A solution to the EU sustainable development challenges

by Xinqi Zhu, Maria Vrachioli, Baldoni Edoardo, Robert M'Barek, and Johannes Sauer

Bio-refineries – A solution to the EU sustainable development challenges

Authors: Xinqi Zhu (TUM); Maria Vrachioli (TUM); Baldoni Edoardo; Robert M'Barek (JRC); Johannes Sauer (TUM)

Abstract

The European Commission aims to achieve a climate-neutral economy by 2050 and retain the employment growth rate while reducing fossil-based production activities for each member state. Bio-refineries are considered to contribute towards this goal of the EU. This study carries out an in-depth investigation on the impact of bio-refineries on the employment rate, and other socio-economic indicators. This study contributes to the field of bio-refinery research in two ways. First, the difference-in-difference (DiD) methodological framework is applied for the first time to analyse the impact of bio-refineries on employment in a local context. In addition, a novel approach of incorporating non-binary treatment effects into the DiD model is used to provide a robust estimation of the marginal economic effects of bio-refineries. Second, this study uses a unique regional level dataset provided by EU JRC and the EH H2020 BioMonitor project to examine the impact of the bio-refinery industry on local employment. This dataset covers multiple European Member states and enables us to account for regional characteristics to assess the effects of the bio-refineries sector on local employment. Based on our research hypothesis, previous studies might have overestimated the impact of bio-refineries on local employment by disregarding the substitution effect between fossil- and bio-based employment.

Key words: Bio-refinery, Employment, Difference-in-Difference, Bioeconomy

1. Introduction

In 2019, the European Commission announced the “European Green Deal” with an ambitious goal to be climate neutral by 2050, which means the amount of greenhouse gas emitted to the environment should be equal to the amount of greenhouse gas absorbed by plants or buried underground by carbon removal techniques. This raised the concern of slowing down the economic growth for the EU member states, as they may have to constrain their highly polluting industries, which are energy-intensive and fossil-based. For example, the industries of steel, chemicals and cement are indispensable to Europe’s economy, as they supply several key value chains, hence, the main source of employment and economic growth. According to the European Green Deal, the European Commission claimed that decarbonisation and modernization of these sectors are essential. A potential plan is to replace fossil-based inputs in these industries with bio-based inputs. The new bio-based industries, which referred to as bio-refinery, will take up the jobs of fossil-based industries and produce in a more environmentally friendly and sustainable way. Bio-refineries is known as “an overall concept of a processing plant where biomass feedstocks are converted and extracted into a spectrum of valuable products” or “the conversion of biomass feedstock into a host of valuable chemicals and energy with minimal waste and emissions” by the European Commission (EU bioeconomy glossary item). Given this definition, it is widely belief that bio-refineries helps to promote economic growth in the rural region (Lehtonen and Okkonen, 2013). This is because it develops new feedstock processing chains for various biomass. For example, scattered wood chips which previously treated as waste can now produce energy and materials (Bailey et al., 2011). This new value chain creates new employment opportunity since it requires new technology, new supply chain, and rising demand in feedstock (Thornley et al., 2013). However, given the different supply chains for fossil-based and bio-based productions, the employment growth of the two types of production cannot be assessed by a trivial substitution. Thus the key research question for this study is how establishing and operating a bio-refinery will influence employment at the regional level.

The researches tackling social issues such as employment and economic growth with the bio-refineries are emerging recently. The mainstream of these studies focuses on the employment generated in the rural region as researchers found that many bio-refineries are located in the rural area (Heijman et al., 2019). This location allows the bio-refinery to reduce its production cost by lowering transport cost and feedstock price. For this reason, the establishment of bio-refinery has a significant impact on the local region (Cambero and Sowlati, 2016). Studies that try to investigate this impact begins with theoretical models estimating the potential employment impact of the bio-refinery. These studies apply either the input-output model or linear programming model to simulate the impact of bio-refinery on regional employment. These studies concluded that despite the job created by a single bio-refinery might be moderate (around 50-300 direct jobs), the indirect employment generated by the facility can be 10-100 times of the direct jobs. The variation of employees per bio-refinery depends on a few factors. First, the processing technology the firm chooses. Second, the type of biomass feedstock processed. Third, the different primary concern of the industry, either maximizing economic, environmental or social impact. Forth, the input constraints facing by the firm given various regional characteristic. For this reason, the result of these studies cannot be generalize given they already preset the aforementioned factors. This study aims to develop a generalised framework to assess the overall impact of bio-refineries on regional employment in the EU using a different-in-different (DID) method.

The next section will provide an overview of the development of bio-refinery sectors in the EU. It is followed by a literature review on the impact of bio-refineries on employment in Section 3. The summary statistic for the variables utilized in the analysis is presented in Section 4. The methodology and findings are revealed in Section 5. Last but not least, Section 6 discusses the results. Then, Section 7 concludes the study.

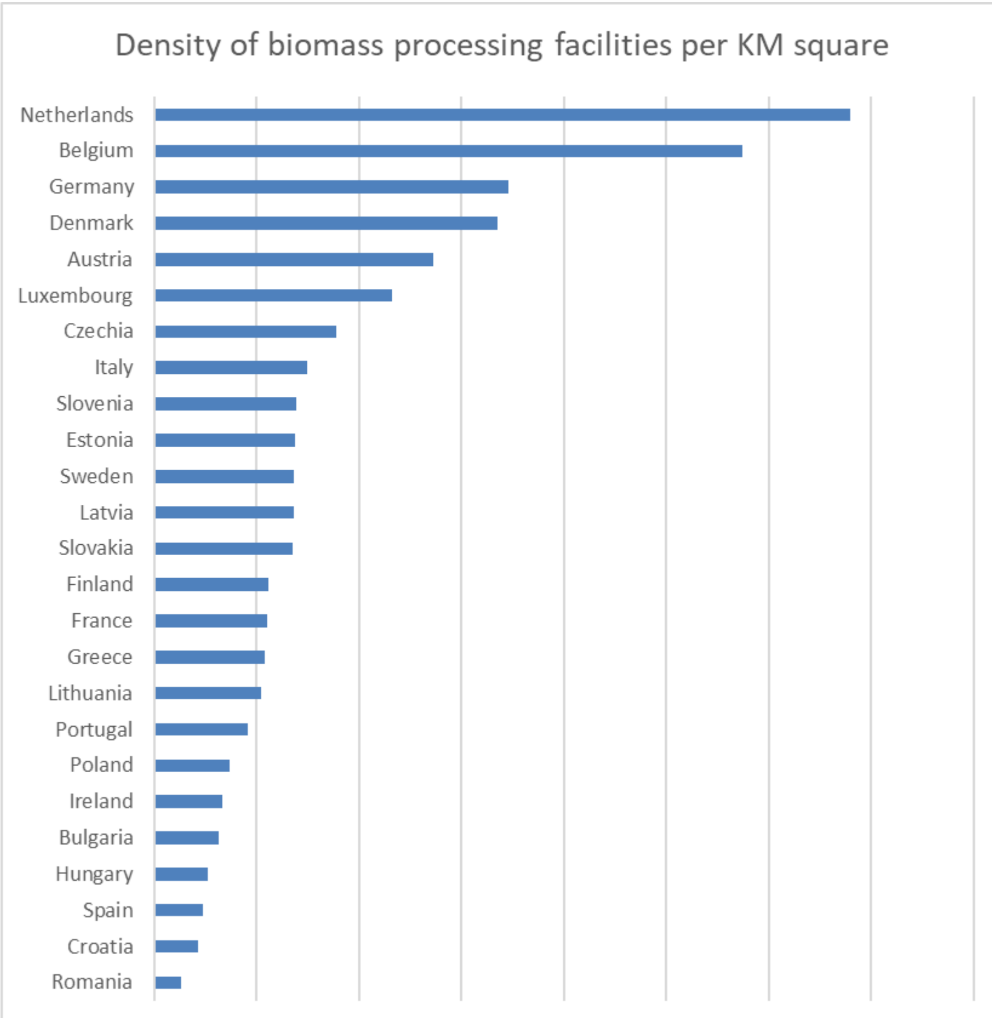
2. Bio-refineries in the EU

According to the EU bioeconomy job and wealth data from JRC DataM, the EU bioeconomy sector employed 3.53 million workers in the EU. This includes bio-based textiles, wood products and furniture, paper, bio-based chemicals, pharmaceuticals, plastic and rubber, liquid biofuel and bio-based electricity. There are 2362 biomass processing facilities identified among all EU member states producing pulps and paper, bio-methane (used mainly as the feedstock for biogas production), starch and sugar, bio-based chemical, timber, liquid biofuels, and composites and fibres. Some facilities produce multiple products. The majority of them (92.5%) are at the commercial scale and the remain 5.7% are pilot/demonstration plants. R&D facilities occupy 1.8%. Most of these facilities were located in Germany (617), France (353), and Sweden (307). In terms of the density of biomass processing facilities, the Netherland, Belgium, and Germany rank top three within the EU.

Figure 1

Source: JRC DataM Bio-based Industry
(https://datam.jrc.ec.europa.eu/datam/mashup/BIOBASED_INDUSTRY/index.html)

Figure 2



Source: Own calculation based on JRC DataM Bio-based Industry and Country size

3. Literature review on bio-refineries impact on employment

Heijman et al. (2019) investigated the bio-refinery impact on rural employment in Hungary using a regional input-output model (NUTS3 level). The bio-refinery they investigated located at one hundred kilometers from Budapest. It was constructed in 2010-2011 which process corn into bioethanol, animal feed, corn oil and other bio-based materials. The feedstock for this bio-refinery sourced from the nearby region Fejér and Tolna which is the major corn growing region. The major economic activities in this region is farming. They found that the operation of the bio-refinery helps to create employment opportunities not only through direct employment at the facility but also create and maintain jobs in farming and service industries. According to their study, the bio-refinery employed 172 people directly and created more than 5000 jobs indirectly. Hence, they believe bio-refineries can make a significant contribution to the rural development.

Thornley et al. (2008) estimates the potential impact of EU bio-refinery on employment in 5 EU member states including Germany, Netherland, Poland, Spain and United Kingdom using a process synthesis methodology. They focus on the biomass processing facilities that process two types of biomass, straw and softwood. In their analysis, they identified 27 process options and estimates their employment impact from its design stage to construction stage to operation stage. They also measured the direct employment and indirect employment (in the paper this was defined as induced employment) generated during these stages. The amount of employment required for each biomass processing option was measured in the unit of man year. Process 19, which convert straw into surfactants, 2,5-Furandicarboxylic acid (FDCA), and dry lignin product, generates the highest man year which can goes up to almost 70 thousand man years per facility. However, it doesn't provide the highest rate of return among the processing option examine by the author. Despite, the economically viable cases do not generate as much man year as process 19, they still generate around 25-50 thousand man years per facility. Given the straw and softwood biomass availability, the author estimated that 24 bio-refineries can be established in the target region they selected which in total generate 1.4 million man years of employment. Roughly 1/3 to ½ of these employment was created for plant operation which can be considered as long term employment. Hence, they concluded that bio-refineries can make a significant contribution to employment generation.

Bailey et al. (2011) projected the impact of establishing 6 lignocellulosic bio-refineries each with an annual capacity of 189 dam3 on employment in the US using IMPLAN[1], an input-output model. They estimated that 2666 new jobs will be created in Alabama which 891 of them sourcing from logging sector, 1217 indirect jobs sourcing from other sectors that has a connection with the logging sector. While more people being employment, 588 more jobs is created to satisfy this increasing demand in the worker's need such as food and services. The author also did a comparative analysis by moving location of the envisage bio-refineries to regions characterized by both abundant timber resources and persistent rural poverty. They concluded that regions with monotony economic activities enjoys a greater benefit in output, employment, income, and indirect business tax compare to regions with a more diverse economic activity.

Cambero and Sowlati (2016) developed a multi-objective mixed integer linear programming model to quantify the potential social, economic, and environmental impacts of constructing forest-based bio-refineries to the interior region of British Columbia, Canada. They found that by optimizing different objectives, either net present value, GHG emission saving, social benefit, or total job creation, the number of jobs created also varies. Maximizing total job creation generates the highest amount of new jobs (239 jobs), this is followed by maximizing social benefit (238 jobs), GHG emission saving (203 jobs), and net present value (82 jobs). The majority of the created jobs are for entry-level jobs such as logging machinery operators, heavy equipment operators, and truck drivers. Since the entry-level jobs retains high unemployment rate in British Columbia, the authors believe jobs created by the bio-refineries have great significance to the region.

Serrano-Hernandez et. al (2017) developed a mixed integer linear programming model to determine the optimal location for establishing a biofuel facility within Navarre, Spain. Their model considered that the facility owner might choose the location that optimized either on economics objective or environmental

---

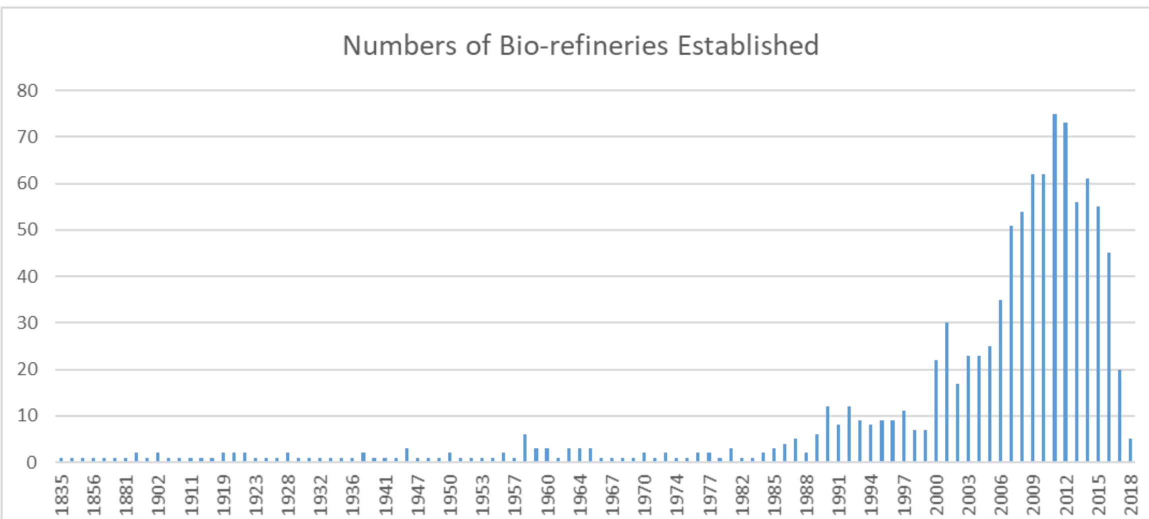
[1] Economic Impact Analysis for Planning (IMPAN) is an sophisticated economic input-output model use to evaluate the impact of forest management on rural development. URL: https://blog.implan.com/what-is-implan

objective. Economic objectives include purchase, transport, and stock cost. Environmental objective is to minimize the distance to transfer feedstock. This study reveals the decision making process on how the location of bio-refinery has been determined. For this reason, this study compares the possibility to established a bio-refinery in the region based on the aforementioned economy and environmental criteria.

4. Data

This study used a novel bio-refinery survey data conducted by a EU bioeconomy project named BioMonitor. We investigated the bio-refinery sector development across all EU member states. We used the information about the year of establishment of bio-refineries within Europe from JRC. There are 1005 biorefineries in the database, which produce bioenergy, bio-methane, biochemical, biofuel, and bio-based components. We excluded the sugar, and pulp and paper industries in this case study since these conventional bio-refineries do not initially aim to replace fossil-based production. Many of them have also been established before the concept of bioeconomy emerged.

Figure 3: Biorefineries



Source: JRC owned database

Figure 3 shows the year of establishment of all the bio-refineries that produced bioenergy, bio-methane, biochemical, biofuel, and bio-based components within our database. Before the concept of bioeconomy was introduced and established, there were a few bio-refineries that were established. Hence, their impact on regional employment considered as very limited. The number of bio-refineries established has prompted after 2000 and reached its maximum in 2011 where 74 bio-refineries were established only that year. For this reason, we are more interested in the employment impact introduced by the bio-refineries after 2000. This also coincides with the period of interest for the BioMonitor project.

In the aim of isolating the causal effect of bio-refinery establishing on regional employment, we also control for other regional characteristics that might influence the regional employment. This includes population, disposable income, education, business birth rate and survival rate. These variables were selected based on a literature review that selected variables that may have an explanatory power on employment change (Biag and Lucifora, 2008; Fritsch and Schindele, 2011; Torp, 1994).

We also collected three additional variables on patents, motorways and railways since they are key determinants of where a bio-refinery can be located (Serrano-Hernandez et al., 2017). These variables were used for matching purposes in order to account for the possibility that some regions have a higher possibility to accommodate the needs of a bio-refinery establishement.

All the aforementioned variables were collected from the Eurostat at the NUTS 2/3 level. Table 1 presents a detailed description of these variables.

Table 1: Definition of Variables

Variable name	Unit	Description
Employment	Thousands of employees	Number of people employed for all NACE activities
Bio-refinery	Number	Number of bio-refineries established based on their year of foundation
Population	Thousands of people	Average annual population
Disposable income	EURO	Purchasing power index, per inhabitant
Education	Percentage	Percentage of economically active population with tertiary education (ISCED level 5-8)
Patent	Number	Number of patent applications per million inhabitants
Business birth rate	Number	Number of births of enterprises in t
Business survival rate	Number	Number of enterprises newly born in t-3 having survived to t
Motorways	Kilometres	Kilometres per thousand square kilometres
Railway	Kilometres	Kilometres per thousand square kilometres

Table 2 presents the descriptive statistics of the raw dataset. For each year from 2000 to 2018, we have a range of 192 to 240 observations. The mean employment for each year ranges from 753.88 to 857.41 thousand people. The mean number of bio-refinery establishments per year ranges from 0.01 to 0.25. This number was calculated by dividing the number of bio-refineries established in that year by the total number of observations. The mean population in our data for each year ranges from 1682.19 to 1851.12. The mean disposable income for each year ranges from 10745.6 to 15932.5, while the mean population with tertiary education ranges from 9% to 13%. The number of patent applications per million inhabitants stays between 103.92 to 129.46 applications per year with data not being available after 2012. The number of newly introduced bio-refineries per year is between 646.03 to 952.88 for the 2000-2018 period. The number of enterprises with at least a 3-year survival rate varies from 443.44 to 591.47. The data for business birth and survival rates were only available from 2008 onwards. The mean motorways kilometres per thousand square kilometres ranges from 27.24 to 33.35, while the mean railways kilometres per thousand square kilometres ranges from 55.8 to 79.23.

Due to the amount of data missing in our database, we aggregated NUTS 3 level data into NUTS 2 level. We then further filled the missing data with multiple imputation methods using Stata. Details of the multiple imputations can be found in the Appendix of this report. We imputed each observation 20 times and took the average of the results. The cleaned dataset contains 3421 observations, which covers 222 NUTS2 regions from 2000 to 2018. This is an unbalanced panel dataset since some regions do not have data available for each of the years within the period of interest. The maximum number of observations for each region is 19 with an average of 15.4. Table 3 presents the descriptive statistics after cleaning the dataset using the multiple imputation method.

Table 2: Descriptive summary of raw data

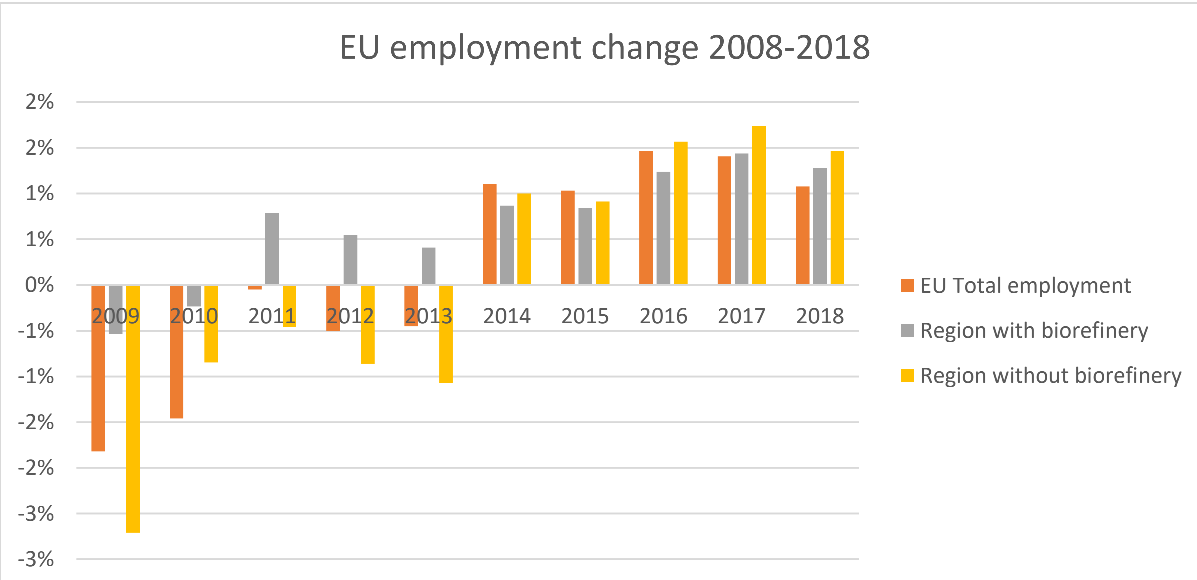
	Obs	Employment	Bio-refinery	Population	Disposable income	Education	Patent	Business birth rate	Business survive	Motorways	Railway
2000	193	753.88	0.04	1682.19	10745.6	0.09	120.05	missing	missing	29.33	56.26
2001	193	760.6	0.07	1683.82	11262.69	0.09	110.86	missing	missing	29.88	55.8
2002	193	760.16	0.02	1686.39	11567.36	0.09	110.25	missing	missing	27.24	55.88
2003	193	763.77	0.07	1692.76	11948.19	0.09	114.23	missing	missing	29.11	57.12
2004	192	770.74	0.04	1706.62	12320.83	0.09	122	missing	missing	28.5	56.67
2005	192	778.3	0.05	1713.17	12753.13	0.1	124.94	missing	missing	28.07	72.63
2006	193	789.47	0.07	1710.83	13320.21	0.1	126.82	missing	missing	29.9	71.5
2007	193	804.66	0.07	1718.24	13833.16	0.1	129.46	missing	missing	30.43	69.57
2008	193	811.06	0.11	1725.6	14380.31	0.11	125.5	792.62	505.48	32.59	79.23
2009	193	793.77	0.18	1729.8	14072.54	0.11	124.87	646.03	485.66	32.99	71.34
2010	193	786.37	0.16	1732.28	14211.92	0.11	128.58	721.56	539.3	33.35	71.61
2011	193	785.53	0.25	1733.99	14405.18	0.11	128.09	730.97	529.84	31.65	72.04
2012	196	773.94	0.22	1710.2	14328.57	0.12	103.92	702.79	443.44	31.46	70.25
2013	196	770.05	0.2	1712.05	14397.45	0.12	missing	778.22	468.55	31.59	70.12
2014	196	776.89	0.17	1714.18	14595.92	0.12	missing	894.67	515.57	31.68	66.25
2015	223	812.5	0.13	1808.46	14973.99	0.13	missing	830.3	486.63	29.88	66.1
2016	240	832.17	0.1	1844.88	14965	0.13	missing	915.76	536.67	28.23	65.3
2017	240	845.66	0.03	1847.87	15436.25	0.13	missing	918.7	591.47	27.71	66.87
2018	240	857.41	0.01	1851.12	15932.5	0.13	missing	952.88	545.54	29.53	66.02

Table 3: Descriptive summary of data processed with multiple imputations

	Obs	Employment	Bio-refinery	Population	Disposable income	Education	Patent	Business birth rate	Business survive	Motorways	Railway
2000	168	778.9	0.05	1735.22	11267.26	0.1	126.29	1138	660.31	31.95	81.72
2001	168	786.65	0.08	1737.31	11785.71	0.1	121.3	1117.55	650.28	32.93	82.59
2002	170	787.23	0.02	1737.84	12050.59	0.1	120.65	1133.78	655.65	31.46	81.88
2003	170	791	0.08	1745.04	12435.29	0.11	126.11	1149	662.08	32.63	81.41
2004	169	799.79	0.04	1759.98	12857.99	0.11	133.06	1117.25	649.46	33.2	80.45
2005	170	806.19	0.06	1764.29	13278.82	0.12	136.22	1125.58	653.38	32.6	86.39
2006	171	820.71	0.07	1768.78	13832.75	0.12	138.95	1101.52	641.68	35.15	88.29
2007	172	834.91	0.08	1774.03	14313.37	0.12	140.97	1124.48	660.58	35.57	86.06
2008	173	841.01	0.13	1778.13	14857.23	0.13	136.51	995.64	596.23	35.96	91.8
2009	174	826.68	0.2	1788.52	14503.45	0.13	132.94	909.62	587.64	35.71	86.75
2010	175	819.03	0.18	1788.86	14645.71	0.13	135.59	949.78	614.25	36.92	86.84
2011	176	819.01	0.27	1795.83	14817.05	0.13	133.81	974.43	618.06	36.29	87.54
2012	178	817.08	0.25	1801.27	14600	0.14	111.26	946.61	567.76	36.74	86.12
2013	178	812.9	0.22	1803.55	14655.62	0.15	174.36	963.71	576.19	36.89	86.04
2014	175	828.46	0.18	1825.92	14926.86	0.15	175.67	1016.28	588.71	36.88	86.16
2015	196	883.2	0.14	1961.17	15354.08	0.15	180.14	995.31	586.8	35.77	86.75
2016	212	901.49	0.11	1992.79	15328.3	0.15	181.39	1031.45	595.88	34.26	86.6
2017	214	902.02	0.03	1964.36	15788.79	0.16	189.03	1038.33	641.25	34.46	86.73
2018	212	925.77	0.01	1988.79	16436.32	0.15	199.96	1071.19	618.07	36.49	83.47

We compared the overall employment change in the EU with the regional employment change. We divided the regions into those having at least one bio-refinery established and those that have not had any bio-refineries established within the period of interest. By comparing the employment change, regions with bio-refineries perform better when overall employment decreases (less employment loss or employment growth). Regions without bio-refinery establishments have a higher employment growth when overall employment increases.

Figure 4: Employment change



Source: Eurostat

We also compare the demographics of bio-refinery regions with non bio-refinery regions. Bio-refinery regions have higher employment compared to non bio-refinery regions. Table 4 shows that only population and education are insignificant different between bio-refinery regions and non bio-refinery regions. Employment, disposable income, patent tendency, business birth and survival rates, motorways and railways are all significantly different.

Table 4: Region demographics

	Biorefinery region		Non-biorefienry region		
Variable name	Mean	Std. Dev.	Mean	Std. Dev.	t-test
Employment	873.9593	546.3924	809.5082	792.277	0.01
Population	1761.857	1012.939	1861.248	1759.619	0.0637
Disposable income	15550.04	4205.051	13338.75	3738.583	0
Education	221.1502	119.7349	212.1427	218.7479	0.1728
Patent	196.9443	131.8483	119.5729	99.34823	0
Business birth rate	1093.509	425.1121	1016.861	579.4171	0

Business survival rate	652.9145	250.2445	602.3811	352.0921	0
Motorways	37.87755	23.70831	33.0411	27.76143	0
Railway	114.7579	87.10307	67.62009	43.47597	0

5. Methods and Results

Our analysis begins by investigating how the establishment of bio-refineries correlates with regional employment. To do this, we first apply a multivariate linear regression model, as presented in Equation (1):

$$
\begin{aligned}
Employment = \beta_0 + \beta_1 Biorefinery + \beta_2 Population + \beta_3 Disposable\ income \\
+ \beta_4 Education + \beta_5 Busines\ birth + \beta_6 Business\ survive + \varepsilon
\end{aligned} \tag{1}
$$

With this model we can identify how changes in employment can correlate with changes in the number of bio-refinery establishments, while holding all other factors fixed. Population, income, education, business birth and survival rates are included as control variables.

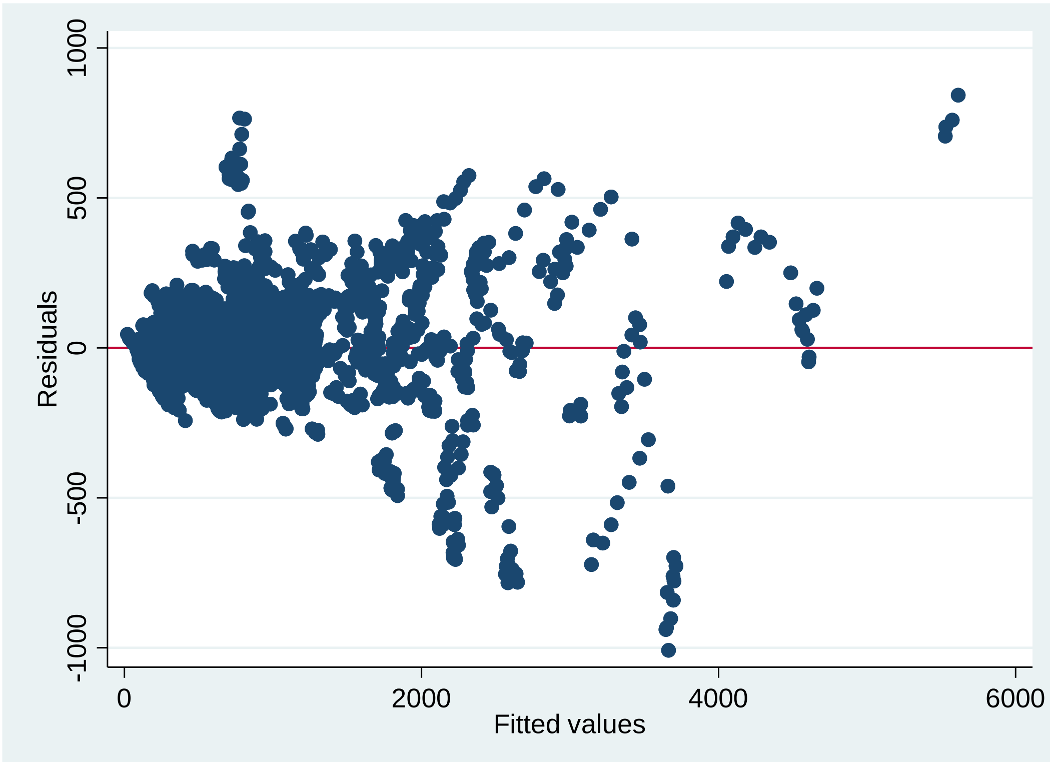
Table 5: Multivariate linear regression results

Employment	Estimated coefficient	t-test
Biorefinery	17.43***	2.95
Population	0.44***	181.06
Disposable income	0.02***	24.60
Education	218.45***	5.68
Business birth rate	-0.04***	-3.03
Business survive	0.14***	5.30
No. Obs	3421	
R^2	0.9502	

Significant level: ***=1%, **=5%, *=10%

The coefficient for bio-refinery is significant according to the regression results of Table 5 suggesting a positive correlation between employment and bio-refinery establishments.

Figure 6: Regression residuals



According to the residual plot (Figure 6), we observe that there might be a heteroscedasticity problem in our data. We also performed standard Breusch-Pagan tests for heteroscedasticity and the chi-square test statistic rejected the null hypothesis. This indicates that there is a heteroscedasticity problem in our data. This problem might source from individual heterogeneity. Hence, we applied a fixed-effects panel regression model with time dummies to account for the unobserved factors, which might influence regional employment that is invariant over time. For example, culture and labour preference (Brügger, 2009; Van de Walle, 2015). We checked whether these time-invariant are not correlated with the error term by performing the Hausman test suggesting a fixed-effects model instead of random effects model.

$$\begin{aligned} Employment_t = \alpha_i + \beta_{1t} Biorefinery + \beta_{2t} Population + \beta_{3t} Disposable\ income \\ + \beta_{4t} Education + \beta_{5t} Busines\ birth\ rate + \beta_{6t} Business\ survive \\ + \beta_{7t} Time\ dummy + \varepsilon \end{aligned} \tag{2}$$

We introduce the variable Bio-refinery in two different ways in the fixed effect model analysis. This difference is based on whether we assume there is a persistent impact of bio-refinery on employment after its establishment. If no persistent impact, the bio-refinery variable is equal to the number of bio-refineries established in year t. If we assume persistent impact, then the bio-refinery variable will be equal the sum of bio-refineries established up until year t. For instance, if a region has one bio-refinery established in 2005 and another one established in 2010, the number of bio-refineries established will be 0 before 2005; equal 1 from 2005 to 2009; equal 2 from 2010 onwards if we use the persistent impact framework.

Table 6: Fixed effect model results

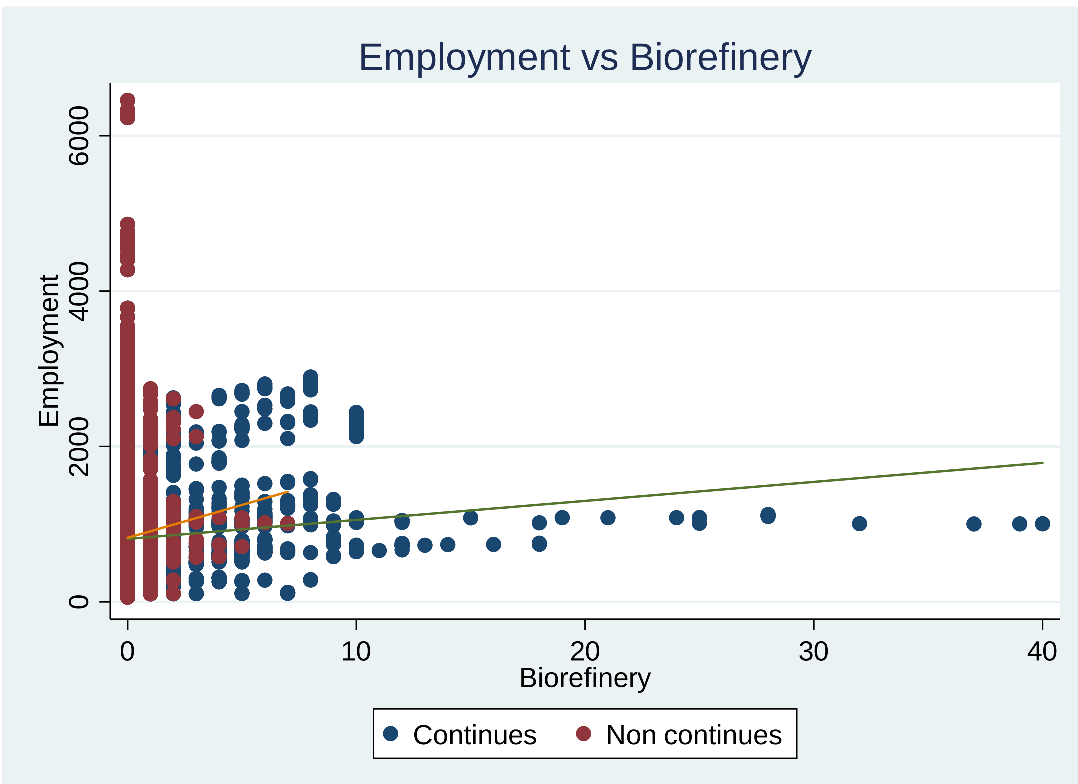
	WITHOUT TIME DUMMY		**WITHOUT PERSISTENT IMPACT**		**WITH PERSISTENT IMPACT**	
	Estimated coefficient	t-test	Estimated coefficient	t-test	Estimated coefficient	t-test
Biorefinery	2.11	1.6	6.61***	4.25	4.159***	2.81
Population	0.32***	7.98	0.35***	9.04	0.356***	9.31
Disposable income	0.01***	10.51	0.02***	8.16	0.016***	7.87
Education	2.11	-0.86	43.31*	1.75	53.865**	2.3
Business birth rate	0.32**	-2.21	-0.03***	-2.75	-0.033***	-2.72
Business survive	0.01**	2.55	0.06***	2.83	0.053***	2.77
2001			-2.67*	-1.75	-1.976	-1.37
2002			-4.99*	-1.87	-4.319*	-1.69
2003			-10.91***	-2.96	-9.598***	-2.76
2004			-14.30***	-3.02	-12.866***	-2.88
2005			-17.12***	-2.96	-15.169***	-2.8
2006			-14.48**	-2.05	-11.797*	-1.8
2007			-10.98	-1.38	-7.741	-1.05
2008			-16.58*	-1.85	-12.793	-1.55
2009			-32.35***	-3.42	-29.527***	-3.36
2010			-41.99***	-4.23	-39.761***	-4.28
2011			-47.48***	-4.34	-45.465***	-4.42
2012			-46.13***	-4.24	-45.820***	-4.41
2013			-52.13***	-4.64	-52.903***	-4.9
2014			-48.29***	-4.29	-49.844***	-4.61
2015			-46.70***	-4.04	-48.605***	-4.38
2016			-41.24***	-3.42	-43.398***	-3.75
2017			-38.29***	-3.01	-40.214***	-3.3
2018			-33.23**	-2.46	-34.552***	-2.7
No. Obs	3421		3421		3421	
Groups	222		222		222	
R2 within	0.4608		0.5045		0.5181	
R2 between	0.9570		0.9572		0.9582	
R2 overall	0.9490		0.9500		0.9508	

Significant level: ***=1%, **=5%, *=10%

With employment being a time-variant variable, without taking into account the time dummy, the bio-refinery variable does not have a significant impact on employment. Hence, we introduce time dummies to the fixed effect model. The significance of the time dummies indicates that employment increases over time. The bio-refinery variable also becomes significant. Based on the results presented in Table 6, the establishment of bio-refineries has a positive impact on regional employment overall. This includes direct employment such as staff for operating the bio-refinery and indirect employment such as farmers, workers for logistics and transportation.

The persistent impact setting has a higher R^2 compared to the without persistent impact setting. This indicates the persistent impact setting explains more variation in the model, hence we are in favour of concluding there is a persistent impact of bio-refinery on employment. Figure 7 shows graphically the positive relationship between employment and bio-refinery established.

Figure 7: Employment vs Bio-refineries



In order to interpret the causal influence of bio-refinery establishment, we adopt the PSM-DID method proposed by Heckman et al. (1997, 1998). The propensity score matching (PSM) helps to eliminate the selection bias from establishing bio-refineries in certain regions with favourable characteristics. The difference-in-difference method helps to isolate the impact of the intervention, which in this case is translated as the establishment of the bio-refinery in the region. The combination of these two methods provide a more solid methodological framework in order to study how bio-refinery establishments can impact regional employment.

The key element on performing the PSM-DID method is to identify the clear cut-off for the pre-treatment period and post-treatment period. Different regions have their first bio-refinery established in different years. In order to identify the pre-treatment period and post-treatment period, we first calculate the PSM score for each region from 2000 to 2018 and get an average of the score for each region. These PSM scores indicate the probability of establishing bio-refineries in the region. Then we match the regions based on their PSM scores. We use nearest neighbour matching to match the treated and control region. Each control region will now have a corresponding treated region. This helps us to identify the pre-treatment period and post-treatment period. For example, if region A is a treated region that has its first biorefinery established in 2010, the pre-treatment period will be before 2010 and the post-treatment period will be 2010 and after. If region B is the corresponding control region for region A, it will have the same pre and post-treatment period.

Next we perform the kernel PSM-DID estimation following the approach of Heckman et al. (1997, 1998). We first estimate the kernel PSM using the logit model, as given in Equation (3):

$$\begin{aligned} \mathrm{P(Treated)} = \alpha &+ \beta_{1it} Population + \beta_{2it} Disposable\ income + \beta_{3it} Education \\ &+ \beta_{4it} Patent + \beta_{5it} Busines\ birth\ rate + \beta_{6i} Business\ survive \\ &+ \beta_{7i} Motorway + \beta_{8it} Railway + \varepsilon \end{aligned} \tag{3}$$

where P(Treated) is the propensity score estimated.

Figure 8: Sample standardized bias

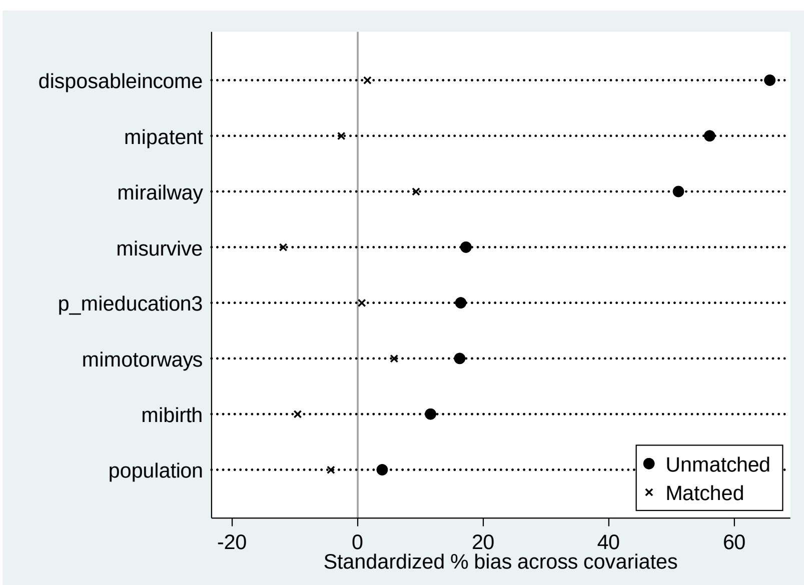


Table 7: Matching result

	Treated mean	**Control mean**	**Bias (%)**	**t-value**
Population	1856.8	1916.5	-4.3	-0.81
Disposable income	16111	16050	1.5	0.33
Education	0.13752	0.1371	0.6	0.14
Patent	198.01	201.14	-2.6	-0.5
Birth	1088.4	1137.3	-9.6	-1.86
Survive	659.31	695.65	-11.9	-2.07
Motorway	37.692	36.224	5.8	1.21
Railway	108.84	102.43	9.3	1.67

Next, we test the balance of variance between the treated and control group using a two-sample t-test on pre-treatment period. Table 8 shows that there is no significant difference between the treated and control mean. This result justifies that the kernel matching helps to balance the treated and control groups.

Table 8: Balancing test

	Treated mean	**Control mean**	**Difference**	**t-value**
Employment	756.83	741.43	-15.401	0.38
Population	1618.76	1524.81	-93.944	1.14
Disposable income	14000	14000	7.914	0.03
Education	0.13	0.13	0.005	1.62
Birth	1135.53	1103.18	-32.358	1.6
Survive	651.83	635.81	-16.02	1.24

Next, we perform the DID to identify the impact of bio-refinery establishment on employment:

$$\mathrm{Y_{it}} = \beta_0 + \beta_1 T + \beta_2 D + \beta_3 DT + \beta_4 X_{it} + \varepsilon_{it} \tag{4}$$

where Y is the employment in region i and period t, T is the time variable where T=0 if pre-treatment, T=1 if post-treatment. D is a dummy variable where D=1 if the treatment is active, otherwise D=0. X are other control variables. We are interested in $\widehat{\beta_3}$ which measures the average treatment effect on the treated (ATT).

Table 9: Difference in Difference result

	Pre-treatment difference			Post-treatment difference		
	Difference	S. Err.	t-value	Difference	S. Err.	t-value
Control (C)	756.829			715.214		
Treated (T)	741.429			902.775		
Diff (T-C)	-15.401	38.161	-0.4	187.561***	29.692	6.32
Diff-in-Diff ($\widehat{\beta_3}$)	202.962***	48.351	4.2			

The DID result shows that there is a significant difference in employment after bio-refinery(-ies) have established. The employment is no significant different in the pre-treatment period. The ATT scores are significant positive revealing a positive impact of bio-refinery establishment on employment.

6. Discussion

We discovered that regions with bio-refineries have higher employment based on the data we gathered. Between the bio-refinery region and the non-bio-refinery region, however, there are considerable disparities in disposable income, patent tendency, business birth and survival rates, highways, and railways. Since these factors might also impact regional employment, we adopt the kernel matching to eliminated the impact from the aforementioned variables on regional employment. This allows us to isolate the impact of establishing bio-refineries from the other factors which may also impact employment. The DID result shows that regions with bio-refineries on average have higher employment compare to regions without bio-refineries. This confirms the belief that promoting bio-refineries helps to provide jobs and business opportunity while tackling environmental issues at the same time. With the aim to achieve climate-neutral by 2050, promoting bio-refineries should focus on highly polluted fossil-based industries (i.e. steel, chemicals, and cement). These industries are indispensable to Europe's economy, as they supply several key value chains and currently contributes to a large proportion of Europe's GDP and employment.

7. Conclusion

Bio-refineries have been seen as a nostrum for the EU to achieve a climate-neutral economy and maintain employment growth while lowing fossil-fuel based production activities. In order to take the decisive step on promoting bio-refineries, policy makers would like to understand the economic, environmental and social impact of bio-refineries. This study investigates the impact of establishing bio-refineries on regional employment using the PSM-DID methods instead of the input-output method adopted by the previous studies. This is because the input-output model focuses on job creation by establishing a new bio-refinery. It leaves the potential negative impact of bio-refinery on the employment of other sectors unattended, e.g. its fossil-based counterpart. Hence, it cannot provide a comprehensive picture of the employment change when establishing new bio-refineries.

The findings of this study show that the establishment of bio-refineries has a positive impact on regional employment. However, because of data limitations, it is impossible to determine if some fossil-based employment have been replaced by bio-based employment. This can be done when there is a breakdown of the fossil-based jobs and bio-based jobs in the regional statistics. Despite this, encouraging bio-refineries contributes to the EU's sustainable development by creating more jobs and economic prospects.

References

1. Aragón, Fernando M., Juan Pablo Rud, and Gerhard Toews. "Resource shocks, employment, and gender: evidence from the collapse of the UK coal industry." Labour Economics 52 (2018): 54-67.

2. Bailey, Conner, Janice F. Dyer, and Larry Teeter. "Assessing the rural development potential of lignocellulosic biofuels in Alabama." Biomass and Bioenergy 35.4 (2011): 1408-1417.

3. Biagi, Federico, and Claudio Lucifora. "Demographic and education effects on unemployment in Europe." Labour Economics 15.5 (2008): 1076-1101.

4. Bitler, Marianne P., and Christopher S. Carpenter. "Health insurance mandates, mammography, and breast cancer diagnoses." American Economic Journal: Economic Policy 8.3 (2016): 39-68.

5. Blundell, Richard, Mike Brewer, and Andrew Shephard. "Evaluating the labour market impact of Working Families' Tax Credit using difference-in-differences." (2005).

6. Brügger, Beatrix, Rafael Lalive, and Josef Zweimüller. "Does culture affect unemployment? Evidence from the Röstigraben." (2009).

7. Cambero, Claudia, and Taraneh Sowlati. "Incorporating social benefits in multi-objective optimization of forest-based bioenergy and biofuel supply chains." Applied Energy 178 (2016): 721-735.

8. European Commission. Directorate-General for Research and Innovation. Innovating for sustainable growth: A bioeconomy for Europe. Publications Office of the European Union. (2012)

9. European Commission. "A European strategy for plastics in a circular economy." (2018)

10. Fritsch, Michael, and Yvonne Schindele. "The contribution of new businesses to regional employment—an empirical analysis." Economic Geography 87.2 (2011): 153-180.

11. Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. "Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme." The review of economic studies 64.4 (1997): 605-654.

12. Heckman, James J., Hidehiko Ichimura, and Petra Todd. "Matching as an econometric evaluation estimator." The review of economic studies 65.2 (1998): 261-294.

13. Heijman, Wim, Zoltán Szabó, and Esther Veldhuizen. "The contribution of biorefineries to rural development: the case of employment in Hungary." Studies in Agricultural Economics 121.1316-2019-1168 (2019).

14. Lehtonen, Olli, and Lasse Okkonen. "Regional socio-economic impacts of decentralised bioeconomy: a case of Suutela wooden village, Finland." Environment, development and sustainability 15.1 (2013): 245-256.

15. Lindorfer, J., M. Lettner, F. Hesser, K. Fazeni, D. Rosenfeld, B. Annevelink, and M. Mandl. "Technical, Economic and Environmental Assessment of Biorefinery Concepts." Developing a practical approach for characterization. IEA (International Energy Agency). Bioenergy: Task 42, no. 2019. (2019)

16. Parisi, Claudia; Baldoni, Edoardo; M'barek, Robert; European Commission, Joint Research Centre: Bio-based industry and biorefineries. European Commission, Joint Research Centre (JRC) (2020) [Dataset] PID: http://data.europa.eu/89h/ee438b10-7723-4435-9f5e-806ab63faf37.

17. Patacchini, Eleonora, and Yves Zenou. "Spatial dependence in local unemployment rates." Journal of Economic Geography 7.2 (2007): 169-191.

18. Petrick, Martin, and Patrick Zier. "Regional employment impacts of Common Agricultural Policy measures in Eastern Germany: a difference‐in‐differences approach." Agricultural Economics 42.2 (2011): 183-193.

19. Ronzon, Tévécia; Piotrowski, Stephan; M'barek, Robert; Carus, Michael; Tamošiūnas, Saulius (2020): Jobs and wealth in the EU bioeconomy / JRC - Bioeconomics. European Commission, Joint Research Centre (JRC) [Dataset] PID: http://data.europa.eu/89h/7d7d5481-2d02-4b36-8e79-697b04fa4278

20. Serrano-Hernandez, Adrian, et al. "Determining an optimal area to locate a biorefinery under economic and environmental criteria." Transportation Research Procedia 22 (2017): 95-104.

21. Thornley, Patricia, John Rogers, and Ye Huang. "Quantification of employment from biomass power plants." Renewable Energy 33.8 (2008): 1922-1927.

22. Thornley, Patricia, Katie Chong, and Tony Bridgwater. "European biorefineries: Implications for land, trade and employment." Environmental science & policy 37 (2014): 255-265.

23. Torp, Hege. "The impact of training on employment: Assessing a Norwegian labour market programme." The Scandinavian Journal of Economics (1994): 531-550.

24. Van de Walle, Steven, Bram Steijn, and Sebastian Jilke. "Extrinsic motivation, PSM and labour market characteristics: A multilevel model of public sector employment preference in 26 countries." International Review of Administrative Sciences 81.4 (2015): 833-855.

25. Wing, Coady, Kosali Simon, and Ricardo A. Bello-Gomez. "Designing difference in difference studies: best practices for public health policy research." Annual review of public health 39 (2018).

Appendix – Multiple imputation

Multiple imputation is a general solution to the issue of missing data that can be found in several widely used statistical sets. It aims to account for uncertainty about missing data by producing many possible imputed data sets and integrating the results from each of them appropriately. In this study, we imputed control variables education, business birth rate, business survival rate, patent, motorway, and railway. These variables are imputed with a linear regression model. We perform the imputation 20 times and calculated the average value of the imputed missing values.

The linear regression for education:

$$\text{Education} = \text{Training} + \text{Education Participation} + \text{Disposable Income}$$

The linear regression for patent:

$$\text{Patent} = \text{Biochemistry patent} + \text{Vegetable oil patent} + \text{Education}$$

The linear regression for business birth rate:

$$\text{Business birth rate} = \text{Patent} + \text{Biorefinery} + \text{Education}$$

The linear regression for business survival rate:

$$\text{Business survival rate} = \text{Population} + \text{Disposable Income} + \text{Business birth rate}$$

The linear regression for motorway:

$$\text{Motorway} = \text{Population} + \text{Disposable Income} + \text{Railway} + \text{Land}$$

The linear regression for railway:

$$\text{Railway} = \text{Population} + \text{Disposable Income} + \text{Motorway} + \text{Land}$$