

Title

A harmonized and spatially-explicit dataset for the European Union's €61 billion in Common Agricultural Policy payments to farmers for 2015

Authors

Kimberly A. Nicholas^{1*}, Frida Villemoes¹, Edmund Lehsten¹, Mark V. Brady^{2,3}, and Murray W. Scown^{1,4}

Affiliations

1. Lund University Centre for Sustainability Studies, Lund, Sweden
2. Agrifood Economics Centre, Department of Economics, Swedish University of Agricultural Sciences (SLU), Sweden
3. Centre for Environmental and Climate Research (CEC), Lund University, Sweden
4. Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, The Netherlands

*corresponding author: Kimberly Nicholas (Kimberly.nicholas@lucsus.lu.se). Postal address: Box 170, Lund University Centre for Sustainability Studies, 22100 Lund, Sweden.

Keywords: agricultural subsidies, public spending, transparency, rural development, income support

Abstract

The Common Agricultural Policy (CAP) is the largest budget item in the European Union, but Member States report spending in a variety of formats and languages, hampering holistic analysis. Here we have made the payment data centered around the year 2015 spatially-explicit at the county or province (NUTS3) level, and aligned payment measures reported between Member States, by translating, harmonizing, and geolocating 16,443,532 million payments originally reported by EU Member States and compiled by the Open Data Foundation Germany through its project Farmsubsidy.org. We have assembled the first database to our knowledge to report individual farm-level CAP payments by standardized CAP funding measures and geolocation. This dataset allows in-depth analysis of over €60 billion in public spending from 2015 by purpose and location for the first time. Further, this data is representative of the distribution of annual payments dictated by the CAP program for 2014 to 2020. These data are of interest to researchers, policymakers, NGOs, and journalists alike for evaluating the distribution and impacts of CAP spending.

Background & Summary

Goals and motivation

Agriculture provides essential food and livelihoods for people, but land use change, primarily driven by agriculture, also causes the majority of global biodiversity loss¹ and 23% of climate heating². The current food system is criticized for harming both planetary and personal health, recognizing the urgent need to transform to healthy and sustainable food systems.^{3,4} Agricultural subsidies globally were recently shown to total over \$700 billion (€640 billion)⁵, with many reinforcing harmful practices.

The European Union (EU) has pledged to be a global leader in sustainable agriculture, including making the “Farm to Fork” sustainable agriculture strategy a cornerstone of the European Green Deal⁶. Currently the principal policy for European agriculture is the Common Agricultural Policy (CAP), which uses the EU budget to support farmers making a reasonable living, with the aims of improving agricultural productivity, ensuring a stable supply of affordable food, and supporting rural development, climate action, and sustainable resource management.⁷

However, the CAP has faced wide-ranging critiques, including for increasing income inequalities and for under-resourcing goals for rural development and environmental protection compared with increasing production.^{8,9} The CAP is currently under reform for 2021-2027¹⁰; the European Commission has communicated that the future CAP should evolve in line with the Sustainable Development Goals¹¹.

Currently, despite EU principles and laws on transparency in reporting public spending, basic gaps in data reporting by Member States and a lack of centralised data make it extremely difficult to get an overview of CAP spending at a finer level than the published national summaries. The goal of the present study is to present the first

harmonized and spatially-explicit database of CAP spending for 2015, enabling analysis of the purpose and location of spending for the first time, and facilitating future analysis of the social and environmental benefits of this spending, for example in relation to CAP and sustainability goals.

Study design and data generated

In this study, we formatted, translated, and harmonized CAP spending data across EU Member States by creating a “Rosetta Stone” to align measure names reported between countries (called “scheme” in the raw data and our code) to a standardized list. This spending averages €58.2 billion annually over the current program period 2014-2020¹². The raw payments data were originally reported by EU Member States, and scraped from 27 different reporting websites by the Open Knowledge Foundation Germany. They average over 600,000 records each (range: 8,600 records reported for Malta, to 3,235,524 records reported for Romania), where each line represents a payout amount to a given recipient under a given measure. We performed language translation and alignment with the purpose of the measure using both machine translation and native speakers, consulting local agricultural experts, and extensive data formatting and processing.

In brief, the workflow proceeded in two stages using a Python script (Figure 1). First, we processed raw data files reporting CAP payments for each country from 2015 (obtained from farmsubsidy.org) to create a “translated” version of the country file. The translated version included additional columns appended to assign each row to a standardized measure name, amount of spending in Euros, and a NUTS3 region (in the EU’s Nomenclature of Territorial Units for Statistical Analysis) for spatial analysis. These “translated” datafiles are suitable for individual country analysis or detailed analysis of particular measures across the EU. Second, we aggregated all of the “translated” country files to produce the “condensed year” file, which contains a total amount of spending for each CAP measure and NUTS3 region in the EU, suitable for broader-scale analyses and aggregation. Code to reproduce the full dataset is available on GitHub.

Potential reuse value

Given high public interest in this data during this time of CAP reform, as well as extensive and ongoing calls for increased transparency of CAP spending, and the need to use public resources wisely in pursuing urgent social goals such as sustainable food production, rapidly reduced climate pollution and enhanced natural carbon sinks, and biodiversity conservation, we believe the reuse potential for these data are high. Despite the huge amounts of CAP spending, lack of suitable data at the appropriate time and scale is hindering effective evaluation of CAP measures in relation to their goals¹³. In addition to the research community working on CAP, other likely users include non-governmental organizations working with the Sustainable Development Goals, environmental stewardship,¹⁴ and other policy goals, journalists reporting on public spending and government oversight, as well as the EU Member States themselves. We hope that this data harmonization effort can be carried forward to support more transparent and harmonized reporting by Member States in the coming

CAP spending period 2021-2027 to support ongoing analysis and collaboration towards achieving Europe's policy goals for sustainable agriculture.

Methods

1. Data Sources

1.1. CAP Spending Reporting by EU Member States

The CAP spending data originally came from EU Member States, who are obligated to report spending to comply with the EU's principle of transparency, including financial regulations adopted in 2012 to publish information on recipients of EU funds, and a 2013 regulation with specific obligations for publishing CAP payment recipients⁸.

Specifically, in Article 111 of Regulation (EU) No 1306/2013¹⁵, Member States are required to report the following information on a single website for at least two years following publication: payment beneficiaries (first and last names of individuals, or full legal name of associations or companies); the municipality where the beneficiary is registered (and postal code "where available"); the amounts of payment corresponding to each measure; and "the nature and description of the measures" for both EU and Member State contributions.

Regarding data accessibility, the European Commission states that *"anyone who receives EU funding under the CAP is included on a publicly available list. This list is designed to promote transparency and trust in EU funding measures. However, the rules still need to strike a balance between the rights of the public to know how their money has been spent and the rights of individuals to protect their personal data. The data is therefore limited in its scope and only available for a set period. [...] It is available from the 31 May of the year after the payments were made. This information will be publicly available for two years after its publication."*¹⁶

Some CAP beneficiaries are exempted from reporting requirements when the benefits they receive are small, or may be anonymized in some cases. For example, Article 112 of Regulation (EU) No 1306/2013¹⁷ exempts reporting requirements for beneficiaries who receive less than €1250 per year. Additionally, some countries give codes to recipients (although some codes consist of recipient names and locations) or anonymize their data in certain cases, e.g., if there are fewer than 10 beneficiaries within a municipality, Denmark does not publish the municipality and postal code of anonymous beneficiaries.¹⁸ A few Member States list open data reuse policies on their websites. For example, the UK states that the CAP data are under an Open Government License for public sector information, where users are "encouraged to use and re-use, free to copy, publish, distribute and transmit the information."¹⁹

The European Commission maintains a webpage²⁰ with links to each country's CAP payments reporting website, where they state, *"To ensure full transparency, EU countries publish information relating to the beneficiaries of all common agricultural*

policy (CAP) payments on their national websites.” Each Member State currently maintains its own database for reporting CAP spending, all of which use different formats, with no universal standard for the “nature and description of measures.” A few Member States have made it possible in recent years to directly download the full data from current years (e.g., Spain, UK, Hungary, and Portugal), but most national transparency websites only allow searching by specific recipients, measure name, or locations, or using pulldown menus, with results reported in html or displayed as only a few entries at a time (e.g., UK and Sweden). Thus, it is very difficult to get a comprehensive overview of where CAP spending went and for what it was intended.

1.2. Open Knowledge Foundation Germany & Farmsubsidy.org

Because each Member State has their own webpage for reporting CAP spending data, and each uses a different format, obtaining the raw data in a format suitable for analysis presents a significant challenge. We used raw CAP payment data curated by FarmSubsidy.org, which is a project of the Open Knowledge Foundation Germany, a non-profit organization working on transparency of public money. The Farmsubsidy.org project was started in December 2005 by three journalists. Over the years, project members have worked to “file freedom of information requests on a national level and to clean, compile and present the obtained data on the new Farmsubsidy.org website.”²¹ The stated aim of the Farmsubsidy.org project is to “...obtain detailed data relating to payments and recipients of farm subsidies in every EU member state and make this data available in a way that is useful to European citizens.”²²

The Open Knowledge Foundation works to extract the data currently reported separately by Member States and make the raw data files available in one place. To do so, they have built data scrapers for each country website, which are released under open license with the intention to be maintained by the community.²³ The code for the farmsubsidy.org scrapers are available on GitHub; most were last updated between 2016-2018, although Cyprus was added in late 2019²⁴.

Farmsubsidy.org states that they publish the data exactly as published by national governments: “Ultimately, the data available on this site is only as good as the data we have received from the governments. We do not change the data we have received from governments so if you think you have identified an error in the data you should notify the relevant government agency and, if possible, let us know too.”²⁵

The Farmsubsidy.org project discloses their funding on their website, which states that most project funding was received between 2006-2010, with the last funding listed as €5,000 in 2014. Their website states “Currently this project is not funded, but maintained by volunteers.”²⁶

1.3. Downloading data and data content

We built on the work already undertaken by farmsubsidy.org to use the data they had scraped and made available. We downloaded the raw CAP payment data for all

available Member States and years from FarmSubsidy.org on July 15, 2019, using the Linux command

`$wget -r https://data.farmsubsidy.org/latest/`. These data can be searched on the farmsubsidy.org website by address given or by amount, but not by standardized measure (scheme) name or geolocation, because these features are not reported in the original data.

Each raw data file from farmsubsidy.org listed information potentially including recipient name and address, amount paid, currency, year, and CAP measure under which the payment was made. Not all Member States report all data. Header names were evaluated manually to be sure to include all relevant data (for example, Romania lists measure names (what other countries call “scheme” in their raw data) under the column heading “scheme_2”). See Table 1 for an overview of the content of the raw data files of CAP payments.

2. Data Availability by Country and Year

2.1. Countries and years included for CAP payments

Data on farmsubsidy.org were available for a range of years that varied by country, though most countries had data available for 2014-2017. We built our database using data from 2015; where this was not available, we used the closest available year (2014 for Denmark, and 2016 for Bulgaria and Sweden). Additionally, we present data from 2016 for the Czech Republic, because the 2015 data contained only about 5% of the data volume and 20% of the payment totals than the previous and following year, and was confirmed to be more than five times smaller than the Eurostat payment reported for 2015. Finland had data only for 2004-2013, during the previous CAP period, and was therefore not included. The data for Italy reports only “Total” for each payment rather than listing specific measures, so it was not possible to classify CAP payments in Italy by measure. Note that the CAP financial year runs from 16 October to 15 October, with the payments published the following year (i.e., data published May 2015 would be regarding 16 October 2013 to 15 October 2014)²⁷. Most Member States listed the payment year in their raw data, which was repeated in the file name.

It is unclear whether all Member States report years consistently (most Member States do not clarify how years are reported). We assume that data files stating a year of 2015 refer to the majority year the payment was made (as stated by Germany on their transparency website, where searching 2018 is stated to apply to payments made from October 16, 2017 to October 15, 2018)²⁸. Thus, we believe the majority of our data report spending undertaken from October 2014-October 2015 and reported in spring 2016.

2.2. National vs. EU measures

Although Article 111 of Regulation (EU) No 1306/2013²⁹ requires Member States to report payment amounts and “the nature and description of the measures” for both EU and Member State contributions, we did not find a consistent system for distinguishing EU and Member State contributions reported in the payment data. For

the purposes of this paper we follow the EU terminology used in the legislation and refer to the various payments that Member States are permitted to make under the CAP as *measures* (which elsewhere are referred to more generally as policy instruments or payment schemes depending on the literature).

As a brief explanation for distinguishing European and Member State financing under the CAP, the CAP is divided into two programmes known as “Pillars”: Pillar I, entirely financed by the EU³⁰ through the European Agricultural Guarantee Fund (EAGF), directs the majority of the CAP budget to support direct payments to farmers (71.3% of CAP spending for 2014-2020) and market measures (4.3%).³¹ Pillar II³² finances the remaining 24.4% of CAP spending through the European Agricultural Fund for Rural Development (EAFRD). Pillar II programs support rural development and environmental measures. They are co-financed by EU funds and regional or national funds.³³

The European Commission states that funding reported under Pillar II includes both EU and national sources: “For funding from the European Agricultural Fund for Rural Development (EAFRD), the amounts published include both the money received from EU funds and from the EU country. This means that the reported amount reflects the total amount of public expenditure on the project.”³⁴

A few countries listed the funding source in the name of the measures (e.g., Hungary listed National, EAGF, or EAFRD), but most did not. Some reported only EU-funded payments; e.g., Denmark stated “National aid schemes without EU funds are not published”³⁵.

The farmsubsidy.org data scraped the raw CAP payment data exactly as reported by Member States³⁶, which “may also contain non-EU national subsidies which are sometimes in the same database” (Stefan Wehrmeyer, personal communication).

In the absence of any information about national vs. EU funded support in the raw data, we assumed all measures listed were from EU funding, and matched all that we could to the standardized list of EU measures. We classified any measures listed to have national support (either in the measure name, or from information from the reporting agency) with the ID code National, rather than matching to a specific CAP measure (ca. 1% of total payments).

3. Payment values

3.1. Negative payment values

We included all payment values exactly as reported, including negative payment values, which comprised a small proportion of payments in a few countries. The German CAP payment agency Bundesanstalt für Landwirtschaft und Ernährung (Federal Agency for Agriculture and Food) states that positive amounts in the transparency portal refer to payments from the EU budget to beneficiaries, whereas when goods from a public intervention are sold, the price paid by the buyer for the goods appears as a negative amount, representing revenue for the EU budget.³⁷

For nearly all countries, negative payments were non-existent or very small (less than 1% of total payments for the year). For instance, for the largest payment recipient, France, negative payments were about 0.4% of total payments in 2015). The only countries to have such payments be more than 2% were Cyprus (2.7%) and Greece (13%, dominated by one very large negative payment of nearly €200 million). (See file “missing_money_percentages” on our GitHub.) We checked the current data on the Greek transparency website³⁸ and found 12,024 negative values reported from 2017, so we assume these payments were valid.

3.2. Payments between countries

In the first step of our Python code, payments are summed by country that reported receiving the payment (as the EU reports). But two countries (Belgium and the Netherlands) report payments to other countries within their national transparency reporting, perhaps indicating that the landowner lives abroad. In the second condensation step of our data processing (**Figure 1**), our Python code attributes all payments to the beneficiary country listed.

For example, payments listed in the raw data reported by Belgium as having a recipient in France would be added to our translated file for France, rather than Belgium. However, these international payments were only reported by Belgium and the Netherlands, and in both cases they were a small percent of the total country payment (about €2.6 million for Belgium, less than half of 1% of their total payments received; and about €13,000 for the Netherlands in 2015, compared with their total payments received of over €1 billion). See the code for “money_movement.py” on our GitHub.

4. Standardizing location and currency

4.1. Location

To obtain a standardized geolocation for each payment entry, we extracted postal codes given in the raw CAP payment data, and matched these to NUTS3 regions, developed and maintained by the EU and generally corresponding to the finest-scale level above municipalities, such as counties or provinces depending on country nomenclature. For example, NUTS3 units in France correspond to 101 Departments.³⁹ We used the 2013 version of NUTS3, which was appropriate for our 2015 data.⁴⁰ We downloaded conversion files from postal codes to NUTS3 regions using the postal code to 2013 NUTS3 regions correspondence table available from Eurostat individually for each country⁴¹ (also found on our GitHub) and used the Python code to extract country and postal codes and match them to NUTS3 regions.

For data where postal codes in the raw data did not match to the available NUTS3 files, we manually matched them where possible by determining the classification system for each country and looking at maps of postal codes and NUTS3 regions to determine the coding pattern between postal codes, which are generally a subset of NUTS3 regions (see Python code, and descriptions in the column “How to match postal codes to NUTS” in Table 1). For most countries, this reduced the amount of

payments that could not be geolocated to less than 2% of the total payments. However, Sweden's postal code numbering system⁴² does not follow NUTS3 borders⁴³, so 19.1% of payments in Sweden could not be linked with a NUTS3 region.

For ten countries (Bulgaria, Czech Republic, Estonia, Greece, Ireland, Lithuania, Luxemburg, Latvia, Romania, and Slovenia), postal codes were not given in the raw data. These countries did include a "recipient location" (usually city) but it was deemed too time-consuming to look up postal codes or otherwise couple them to NUTS3 regions (this could not be automated since shapefiles of postal codes are unfortunately proprietary). Luxembourg, however, consists of only one NUTS3 region, so all payments could be allocated to that NUTS3, leaving nine countries where we could not allocate payments to NUTS3 regions (see Table 1). These nine countries with payments only allocated to the national (NUTS0) level rather than finer NUTS3 spatial level represent about €9 billion in total payments (about 15% of total payments in our 2015 dataset).

4.2. Currency

For countries that did not state a currency of reported payments, we assumed payments were made in Euros (cross-checked with payment totals reported by the EU). Seven countries reported payments in a currency other than the euro. These payments were converted to euros using the average conversion rate for the year of payment. We used currency exchange rates reported by the European Central Bank Euro Reference Exchange⁴⁴ to obtain the average value of the given currency for the given year. Values were extracted in (currency) to euro, for January 1 to December 31 for the year of interest. All values are given in euros for the year reported (i.e., they are not converted further to a standardized year). Currency conversion rates are listed in the Python code.

Poland did not list a currency for its payments in the 2015 raw data, but its transparency website stated that payments are shown in PLN, and the total reported payments in the raw data for 2015 totaled 27 billion, whereas the true value was known to be around €6 billion, so we converted the reported payments from Poland to euro assuming they were reported in zloty.

5. *Standardizing Payment Measures: Creating the "Rosetta Stone"*

To standardize measure names across countries, we created a master "Rosetta Stone" file (Table 2) where we aligned names for the CAP payment measures given by Member States in their national language to a common English language standard given by the "Description of Measures" published by DG AGRI, sometimes cited as Ares (2018),⁴⁵ which lists 102 individual measures, drawn from ten different pieces of underlying regulations (Table 3). In total, there are 27 individual measures within the three chapters of Pillar I of the CAP, and 75 measures across the six chapters of Pillar II, that were active at some stage during the 2014-2020 program period (Table 3).

We digitized the PDF version of “Description of Measures” into a spreadsheet to consistently align measure names with associated legislation and names in national languages. This spreadsheet formed the basis of the “Rosetta Stone” we used to match measure descriptions across countries.

We added additional classification information compiled during our research to aid with both matching of measure names (our creation of the Rosetta Stone document), as original country scheme names sometimes contained reference to e.g., rural development program measure numbers, and to understand the purpose of each measure for further analysis. (See Table 4 for an overview of the structure of the Rosetta Stone document and the source of our classifications). After this meta-data, the Rosetta Stone document consists of columns for each country with original scheme name reported, translated scheme name (where relevant), and notes on the matching criteria.

Following DG AGRI, we used the measure identifier (a combination of a Roman numeral and Arabic numeral ID) in the first column of the tables in “Description of Measures” to uniquely identify each of the 102 possible measures for the analysis, which is reproduced in the first column of Table 2. This ID cannot be directly related to legislation, but consists of two or three parts: the short heading for each section in “Description of Measures” (I, II, III, IV/A, V/B, VI/A, VI/B, VI/C and VI/D), and a row number (i.e., 1, 2, 3,...) or sub-table number and row number (i.e., 1.1, 1.2, ...), such that III.2 is the ID for the second measure listed in section III and V/B.1.10 is the tenth measure in sub-table one of table V/B. Each measure is linked to the underlying legislation via the columns immediately following the ID: Regulation, Title, Chapter, Section, Article where relevant.

Matching the reported measure name to the master list was straightforward for the 14 countries that used the measure ID to report their measures (as noted in “Includes standard numbers of Measure Descriptions” column in Table 1). However, the remaining 13 countries did not use a standardized identifier for measure names (instead using a wide variety of short descriptions in local languages and/or numerical codes with a variety of meanings). This lack of standardization means it is impossible to use the raw data reported by Member States for cross-country analysis, which is why we undertook the harmonization to make the Rosetta Stone. To take a simple example, different countries reported the following names for a measure, all of which we matched to the first measure (I.1, Single Payment Scheme):

“Single payment scheme - title III”

“I.1 - Aide unique découplée à la surface (DPU)”

“I.1 Καθεστώς ενιαίας ενίσχυσης – τίτλος III (ΕΓΤΕ)”

“I.1”

“Guarantee Fund direct support: I.1, Single payment scheme - title III”

More details about the format and examples for each country reporting their measures are given in Table 1.

For those countries whose measures were not reported according to the measure identifier, we identified unique measure names from national languages and manually matched reported country measure names to the 102 individual measures of DG AGRI

using a combination of machine translation, native language speakers, and input from national agriculture experts, as we describe in detail in the next sections.

5.1. Extracting unique measure names

We extracted all unique measure names for each raw country file for analysis in our database (N= 27, all of the EU-28 from 2015, except Finland who reported no data for 2015). We then manually matched measure names from countries to the master list of measures in the Rosetta Stone, using the measure numbers and/or descriptions given as follows.

5.2. Translation and uncertainty assessment

To match measure names used by each country, we created three columns in the Rosetta Stone document for each country: original country scheme name or code (pasted exactly from the raw data files), translated country scheme name (for non-English measure names, this contained the best translation of the original scheme name into English where relevant, preferring native speaker translations over automatic ones), and Notes justifying the basis of the match.

For measure names listed in a language other than English, we first used Google Translate to translate an Excel file containing all measure names from the original language to English, and then manually matched as many measures as possible using the translated names and combination of searching and manual matching with the master measure names.

We distinguished the following classifications for the level of certainty of the match between the national measure name with the master measure list, in order of greatest to least certainty:

Match: all information present, both words and numbers where present, was a strong match to the master list. In other words, both words and numbers matched the master list (if the measure consisted of both words and numbers), words matched the master list (if the measure name consisted only of words) and numbers match (if the measure name consisted only of numbers).

Words but not numbers match: when words from a national measure matched with the master list, but either did not contain numbers present in the master list, or contained numbers that differed from the master list. Where possible to identify, the meaning of the numbers was noted (e.g., when they referred to underlying legislation, article numbers, national measures, or other references).

Unique but imperfect match: the best available match, although some information between the original and master list was inconsistent.

Best of multiple possible matches: more than one match was possible, but enough information was given to support matching with a particular measure.

No match: it was not possible to align the original measure name with a master measure name, because the description given did not fit any of the 102 master measure descriptions from DG AGRI. All non-matches were double-checked against the raw data to ensure they had been imported properly, and that errors (e.g., measure names such as “-” or measure names that cut off in mid-sentence) had indeed been present in the original data and not introduced during our import or analysis.

Note that some measures were placed in both the category “Words but not numbers match” and “Best of multiple possible matches,” as they were matched with the best of a number of options, but the numbers did not match.

We used the full dataset (all levels of match certainty) for analysis, but report the certainty level of matches by country and measure name for most of the measures in Table 5 in case others wish to have a more stringent cutoff.

Note that most countries either had more measure names in their transparency reporting data than the 102 present in the master list (e.g., Hungary listed 274 unique measure names), or listed more than one national measure that was ultimately matched to the same “Description of Measures” name. In such cases, an overflow row was created in the Rosetta Stone, proceeding in alphabetical order by country, with one row per additional measure name. Thus, the first 102 elements of the Rosetta Stone document (rows 2-103, following the header) contain the first instance of a match for an original language scheme name across all countries, with potentially many matches across many countries for each row. Subsequent matches within a country start at row 103 in the Rosetta Stone and continue to row 795, thus there were 692 individual measures added after the main data, where each row lists only one country scheme name, and abbreviated information in the first columns to show the match with the master scheme name.

5.3. Native speaker assistance

For many countries, a substantial fraction of measures remained unmatched at this stage, and we sought the help of native language speakers to improve upon the translations suggested by Google Translate. In many cases native speaker assistance made it possible to make a successful match to align reported measures with the master list.

Native speakers were recruited by the first author on a volunteer basis in an email explaining the purpose of the study, requesting their help in translating into English the measure names that remained unmatched (typically 10-30 short phrases, requiring less than an hour of work), and promising acknowledgement in and a copy of the resulting scientific publication in exchange for their assistance. One native speaker (of Hungarian) was paid as a research assistant for approximately 10 hours of her time in researching Hungarian law as well as providing translations, which helped place many additional measures. We gratefully acknowledge the contributions of all of the native speakers in helping to create this dataset; please see the full list in the Acknowledgements.

5.4. Searching national agency websites

For all countries with measures still not matched to the master list at this stage, the first author searched national agency web pages for any additional information that might assist with matching measure names. In a few cases, it was possible to triangulate information that now appears on national transparency websites to assist in making further matches (e.g., when original language and English translations of measure names were both available), or to download a full list of measure names in their original language with enough information to make a match. These were noted in the relevant country “Notes” column of the Rosetta Stone file.

5.5. Contact with national offices and country experts

After following the procedure above, where any non-matches for measure names remained, the first author sought contact with national agricultural offices in charge of administering CAP payments. Contact information was found from following the links from the official Member States CAP transparency pages, as well as independent research, including seeking contact via the official payment agency Twitter account where available. These details were not always easy to find; to facilitate future contact for others, public email, webpage, and Twitter accounts are listed in Table 1.

The majority of contact attempts with national agricultural agencies to clarify measure names received no reply (contact by email and/or Twitter to the official contact listed for France, Romania, Hungary, Sweden, Lithuania, Bulgaria, and Croatia). Phone contact was successfully made with the UK and Germany, but attempts to follow up with the relevant expert were not successful to receive further information. (However, contact was successfully made with a CAP expert at an NGO in Germany, who provided additional translations and confirmed matches of measure names to the standard EU list in the Rosetta Stone) (Christian Rehmer, personal communication, 2 March 2020). Twitter contact was successfully made with the agricultural agency in Ireland⁴⁶, but the promised answer from the appropriate section was never received. In several cases, no contact information was readily available (for example, the Austrian CAP payments database website⁴⁷ was made by a marketing company; the government agency administering the payment was not clearly listed).

Replies to requests for measure information were received from five countries: Estonia, Latvia, Czech Republic, the Netherlands, and Germany. Representatives from the agricultural agencies in Czech Republic and the Netherlands sent a full list of native measure names aligned with the master list upon request, which enabled complete measure name matching. A representative from Latvia sent a list of measures aligned which enabled matching of all but eight measures (five of which translated as “Action” plus a number). Representatives from the agricultural agency and ministry in Germany provided helpful responses confirming our matches and clarifying the validity of funding periods for different measures. A representative from Estonia sent a match with the 2007-2013 CAP, which was not possible to translate to the current 2014-2020 CAP. Requests to do so did not receive a further reply.

5.6. Expiration and ambiguity of measure names

Complicating the analysis, four of the Regulations underlying the 102 measures expired during the 2014-2020 CAP period, but remained valid for payments through 2015⁴⁸, so Member States in 2015 used a mix of old and new measure terminology in reporting payments.

In Pillar I, Regulation EC 73/2005 (underlying the seven measures related to direct payments in Pillar I for 2007-2013, starting with Roman numeral I) expired in 2013 and was replaced by Regulation 1307/2013 (now with ten measures in Pillar I, starting with Roman numeral II). This means that for example, a country reporting a measure by the name “single area payment scheme” might refer to either I.2 or II.2, both of which use that terminology.

For Pillar II, Regulation EC 1698/2005 (underlying the 46 rural development measures in Pillar II starting with V/B) expired in 2013 and was replaced by the 25 Pillar II measures starting with IV/A associated with Regulation EU 1305/2013. The single measure on information and promotion VI/B.1 (Regulation EC 3/2008) was repealed and replaced by the same measure under Regulation EU 1144/2014 with ID VI/A.1.

Sometimes countries reported a measure using only a short, ambiguous label. For example, three measures relate to “advisory services”: IV/A.2, “Advisory services, farm management and farm relief services”; V/B.1.4, “Use of advisory services by farmers and forest holders”; and V/B.1.5, “Setting up of farm management, farm relief and farm advisory services.” When countries listed a measure using only a short descriptor such as “Advisory services,” it is impossible to say with certainty which measure was indicated.

See correspondence given by EU Regulation in Table 6 between measures in V/B. and IV/A, as well as our notes on additional possible matches when short or ambiguous wording was used by Member States to report measures instead of standardized numbers. Given the overlapping intents and names between measures, it is probably most accurate to combine payments within similar measures for analysis by broad purpose (e.g., using Table 6), rather than focusing in detail on individual measures.

Finally, one of the two measures making up measure VI/C.1, the POESI measure, expired in 2013, but the measure remained active under the same ID (VI/C.1) with the subsequent regulation.

5.7. Grouping of measures

At the broadest level, we distinguished between measures in Pillar I (the first 27 listed in Table 2) and Pillar II (the last 75 measures), by a column indicating the corresponding funding source for the two Pillars (European Agricultural Guarantee Fund (EAGF) for Pillar I and European Agricultural Fund for Rural Development (EAFRD) for Pillar II) as the local-language version of this acronym was often used in measure names reported by Member States. For further analysis of measure purpose, please see measures identified as income support to farmers and as “environmental

payments,” which encompasses all CAP measures which state in the measure wording the intention to principally benefit nature, the environment, climate, or to promote sustainable farming, see Scown et al. (in review), Table S2⁹.

5.8. Finalizing matches

To finalize the Rosetta Stone was an iterative process involving updating the placement of measure name matches, running the Python code, examining remaining errors and non-matches reported, and repeating until the only remaining non-matches were genuine. Measure names identified as valid non-matches from the translation exercise (where no appropriate match could be found to the 102 measures) were left out of the Rosetta Stone. All other measure names were matched to the most appropriate measure. The Rosetta Stone contains the most current and up-to-date data justifying translations and matches in the Notes column for each country. Some measure names returned as non-matches appeared identical to existing measures already placed, but were added as additional rows in the Rosetta Stone, with measure names pasted exactly as they appeared, in order to obtain a match. Note that all measure names were converted to strings (since there were some measure names that consisted only of numbers) and that spaces and punctuation marks were stripped from the beginning and end of the string to facilitate matching.

Data Records

As described above, the original raw data records of CAP payments by country were downloaded from farmsubsidy.org on 15 July 2019. Instructions for doing so and for setting up the file structure are found in the “readme” file on our GitHub, <https://github.com/kanicholas/CAP-farm-payments>.

We have created two sets of data records using the methods described above, both of which reside in the “Output” folder (see Figure 1). First, we have created a “Translated country file” for each of the raw country data files. This maintains the original raw data for each country (every payment reported in the raw transparency data, e.g., over 3 million records for Poland), and appends to it additional columns to facilitate analysis, including standardizing to translated measure name, adding country and year (if missing), converting currency to euro, and adding the NUTS3 region (Figure 1). Note that these files still include personal identifying information of recipients where such data was originally reported by Member States.

These 27 translated country data records are then further processed to create the “condensed year” data record, which shows the payment amounts spent on each CAP measure for each NUTS3 region for the given year. For convenience, the resulting condensed files from running the workflow described in Figure 1 are provided in Table 7, Table 8, and Table 9 for 2014, 2015 and 2016, respectively. These three files are themselves the input to the technical validation, as described below.

Technical Validation

Original data validation

Because the transparency legislation only requires data to be available for two years, it was no longer possible to download the original data from the Member States for 2015 for validation; we thus rely on the accuracy of the data scraped and stored by Farmsubsidy.org.

We are encouraged that we were able to allocate the vast majority of CAP payments in our dataset both to a geolocation using NUTS3 (83% of payments (15% from countries that did not report postcodes, plus an additional 2% from locations we were unable to match to postal codes) and to a CAP payment measure (90.5% of our total payments). Please see R code on the GitHub to read in the translated and condensed files (allocating payments to NUTS3 regions by CAP measure for all Member States, using the years noted above) and perform the validation analysis and produce the figures and tables described in this section.

CAP payment validation

It is difficult to validate our dataset against official EU or Member State data reported for CAP spending because there is no coherent, centralized source of reporting that covers all CAP spending, and there is no publicly available database of payments by measure, or by location finer than Member State. Below we describe the current state of public data reporting on CAP spending. In brief, we were not able to identify a publicly available source of information against which to validate our data in detail. We hope that the publication of our dataset spurs greater inquiry and transparency for the Member States and EU to report this data in a directly usable format (broken down by year, unique CAP measure ID, measure name, and location including postal code).

Reporting of CAP Spending Across the EU

The EU reports annual spending in its expenditure and revenue data under Section 2, “Sustainable Growth: Natural Resources.”⁴⁹ Spending is reported at the broad category of either Pillar I (European Agricultural Guarantee Fund, line item 2.0.1) or Pillar II (European Agricultural Fund for Rural Development, line item 2.0.2), but a finer breakdown by the 102 measures under these broad categories is not available.

A broad comparison between our data and EU reported spending at the Pillar level confirms very close agreement. Our data total €38.9 billion for Pillar I and €16.0 billion for Pillar II for the years used (centered around 2015, with four countries using data from 2014 or 2016 as noted above). For the same countries and years, Eurostat reports €40.0 billion spending in Pillar I and the Commission reports €15.6 billion in spending for Pillar II. These totals are broadly in line with the budget and spending during the 2014-2020 CAP period. In our dataset, we were able to identify only a handful of payments made under measures from national as opposed to EU funding

(totaling €0.62 billion, roughly 1% of our payment total) (Table 10), which could partly explain the difference between our total and that from the Commission.

Eurostat administers data on “subsidies on production” (item code 25000) in their “Economic accounts for agriculture by NUTS 2 regions” (Table 11). However, after repeated requests to Eurostat support during 2018 and 2019 it remained unclear to us what these data actually represented in relation to CAP spending. The Eurostat “subsidies on production” total just over €47 billion for 2015 at the NUTS 0 (Member State) level, but many NUTS 2 regions and even several Member States contain no data in this table.

We compared the payments reported by Eurostat for each Member State with our data, finding generally good agreement, though the payments in our dataset were generally slightly higher than those reported by Eurostat, with the exception of Denmark (Figure 2; Table 12). Our data generally show the expected pattern that Pillar I comprises the majority (approximately three-quarters or more) of the total CAP spending, with the notable exceptions of Austria, Hungary, and Poland, where about half of total spending in our data came from Pillar II.

Reporting of CAP Spending by Member States

Distinguishing the purpose (measure) of reported funding is particularly difficult for Pillar II. Each Member State submits one or more Rural Development Programs (there were a total of 118 Rural Development Programs for the 2014-2020 CAP⁵⁰), which are available as individual PDF files⁵¹ and not in a database format that would enable detailed analysis of spending. Some of the countries that receive the most CAP funding have a large number of different rural development programs for different regions (30 in France, 23 in Italy, 19 in Spain, and 15 in Germany),⁵² making it very difficult to examine rural development spending by country.

The European Network for Rural Development provides spending breakdowns for a subset of measures within Pillar II, but these are reported as aggregated over four to seven years in individual PDF reports for Member States where available, for example for individual measures from 2007-2011⁵³ or for percentage of spending by measure for each Member State for 2014-2020.⁵⁴

Member States tend to report their spending over the full CAP program period, making it difficult to get an independent report of spending for a specific year. For example, the country factsheets list total spending by Member State, including total finances available for the full funding period, but do not distinguish which measures and in what way it is used. For example, Sweden’s Rural Development Program 2014-2020⁵⁵ states that Sweden will use €4.3 billion of public money from 2014-2020 for rural development, of which they state €1.8 billion is from the EU budget, €2.5 billion is national co-funding, but a further breakdown of Pillar II spending is not given.

Financing for rural development consists of co-financed money from the European Agricultural Fund for Rural Development (EAFRD, €99.6 billion for the 2014-2020 multiannual financial framework, averaging €14.2 billion per year over the seven-year spending period) and regional or national public funds varying by region and measure

(consisting of €50.9 billion in regional or national co-funding, and a further €10.7 billion of purely national funding, totaling €61.6 billion)⁵⁶, or an average of €8.8 billion per year. Thus on average across the seven years, €23 billion in total is spent on rural development, with about two-thirds coming from the EAFRD and one-third from regional or national co-funding.

The four Member States who receive the most funding from EAFRD across the seven years of the 2014-2020 CAP are France (€11.4 billion), Italy (€10.4 billion), Germany (€9.4 billion) and Poland (€8.7 billion).⁵⁷

In response to a request via a web form, an Information Systems Support Officer of the European Commission Directorate General for Regional and Urban Policy sent data on 2015 expenditure under Pillar II, broken down by Member State and measure (Yves Durinck, personal communication, 14 May 2020). Overall agreement for Member State spending under Pillar II was generally reasonably close, though our data were notably higher than that from the Commission for Austria, Spain, France, and Poland (Figure 3). Here the Commission reported Pillar II spending in Italy of about €1.7 billion, which would fit reasonably well as about 30% of our total of €5.8 billion (which we were unable to analyze by measure as the raw data reported only the measure name “Total”). Note that data for this figure come only from 2015 from the Commission, but from 2014 or 2016 for four countries in our dataset as noted above.

We analyzed agreement between our data for Pillar II spending and the data sent by the Commission for a set of 832 specific measures within countries, finding generally good agreement (Figure 4; Table 13).

At the NUTS2 level, we were able to compare total CAP payments for 148 NUTS2 regions that were shared between our data and Eurostat, finding generally good agreement (Figure 5).

From validating our dataset against these external sources, we identified anomalously low payments for the Czech Republic in 2015 (leading us to realize the 2015 raw data was only about 5% as long as the files before and after, and we decided to use 2016 which had a full dataset). We note our data for Romania are about a third lower than the payments from Eurostat, but Eurostat data also report 2015 as a year of substantially lower payments to Romania compared with 2014 or 2016, so we elected to keep our 2015 data. See all country-specific notes in the “Errors and Uncertainties” column in Table 1.

Usage Notes

Missing data and uncertainties

Overall, we were able to successfully match most payments from most countries both to the measure they supported (90.5% of payments) and their spatially explicit geographic location within a country (83% of payments; most of the unmatched were in countries that did not report postal codes or other location information). Note that the vast majority of countries where measures were listed were able to be matched, as Italy listed no measures and comprised nearly 10% of total payments in our database.

See Python code “error_percentages.py” for euro amount of payments unmatched, and “error_list.py” for the names of reported measures unmatched, on our GitHub.

A country-level breakdown of the percent of total funding within the year studied matched to both measure and location is shown in Table 1. For many countries, less than 1% of payments were unmatched to either a measure or location.

In terms of country-level measure matching, more than 98% of all payments within a country were successfully matched to a measure except in Greece (7.9% of payments unmatched to measure, due to high payments to six measure names not matched to the master list), Latvia (4% unmatched), and Denmark (4% unmatched to measure, due to errors in measure names); see Table 1. Besides the nine countries who did not report postal codes and were therefore 100% unmatched to a NUTS3 region, the only countries with over 2% of payments unmatched to location were Sweden (19.1% unmatched to NUTS3, due to non-overlap between postal codes and NUTS3 regions in Sweden); the Netherlands (13.4%), France (4.7%); Slovakia (4.4%); Italy (4.0%); and Malta (3.6%).

Individual Member States tend to report the total amount of funding received for the whole CAP period 2014-2020 (not by year), and do not consistently distinguish EU from Member state funding in their payment reporting. Going forward, it would be helpful to clearly distinguish funding source in all payments reported.

Data users should take care to understand the assumptions made about distinguishing national vs. EU measures, the certainty of measure matches to the master list based on the potential for multiple matches with the name used for the measure and/or the existence of multiple measures dealing with the same topic, and other uncertainties described above. Nonetheless, we believe this dataset represents a substantial step forward in transparency of EU budget spending.

A number of countries had specific errors or issues with their measure formatting that required special processing or analysis. In brief, common errors included double entry of both total payments and subtotals for the same recipient (Latvia); ambiguous entries, such as article numbers that could apply to multiple measures (Denmark); reporting of old/expired measures under the previous CAP (e.g., Estonia); measure names not reported at all (Italy); and repeated entries for the same measure name, with variations in punctuation or spelling (e.g., Denmark, Romania), among others. See details about how country-specific issues were handled in the Python code and in Table 1, column “Errors and Uncertainties in Raw Data”.

Recommendations for future improvement in CAP data reporting

We have undertaken to harmonize existing data reported by individual Member States on public spending on the Common Agricultural Policy. The difficulty we encountered in doing so, and the resulting uncertainty in the data, demonstrates the need for common-sense reforms to streamline data reporting and curation. Here we echo previous calls⁵⁸ for EU governments to release data “according to a common format so that it is possible to analyse the data in a meaningful way across the European Union.” Below we offer recommendations to make the CAP payment data reported by Member States more standardized in order to facilitate analysis and

transparency, from most fundamental (following existing regulations and good practices in data curation) to those that would require updates or changes in current practices, but would offer substantial benefits to public transparency. These recommended improvements to data reporting would enable detailed analyses of CAP spending, which could inform and improve the overall performance of the CAP.

Most fundamentally, we recommend that future Member State reporting of CAP spending require the following documentation be reported through each country's data transparency portal:

- Member State payment reporting should include a column with the standardized EU-wide measure name and associated unique identifier for each measure alongside the national measure name to minimize errors.
- A dictionary file that lists the measure names reported in the native language/format, referenced to a standard list of measure names and unique identifiers provided by DG Agri. (The Rosetta Stone document provided in this paper (Table 2) could serve as an initial template to do so, though it would need to be updated for the measures adopted in the 2021-2027 CAP).
- A document listing the names and briefly describing any additional national measures reported in the payment data (those not included in the standardized EU measures).
- Clearly specify the name of the agency and department responsible for curating the payment data as well as a contact person, and provide their contact information for questions about the dataset.
- Provide meta-data for the whole dataset, including an explanation of negative values.

To facilitate analysis and minimize errors, the **format of data reported** should always include the following elements and consistently follow data curation conventions:

- In addition to any data preferred to report by the reporting Member State, CAP reporting should always include the following data: country name, year, standardized EU measure name and identifier, recipient identifier (name OR recipient ID, with unique European identifier for legal entities), recipient postal code, payment amount, currency used to report payment amount (if data were collected in a currency other than Euros and converted to Euros, the exchange rates used to convert to Euros should be reported in the meta-data).
- A column to distinguish the source of payments made (EU or Member State funds and their proportional contributions).
- Identifying data (measure name, ID, recipient name, etc.) should be completely filled in for each and every row of payments reported (not left blank under headings assumed to be carried down until a new entry appears, which hinders analysis). Where values are zero or not applicable, appropriate codes should be used.
- Measure names should be reported from a standardized list or pull-down menu, not entered by hand, to reduce the frequency of duplicate and erroneous measure names.
- When numbering is used as part of measure names, it should be done using unique consecutive numbers (such as 001, 002, ...). Sometimes current numbering conventions yield non-sequential sort orders, such as measures in Denmark which mix numbers at the start with subsequent names; sorting them yields 1, 10, 11, 12, ... 2, 21.
- Validation should be carried out by reporting agency data responsible before reporting data to minimize errors. That is, the data should be totaled and reconciled with official statistics to ensure all payments have been accounted

for, and standardized, unique measure names and ID numbers should be used to ensure every payment reported can be uniquely associated with its recipient and purpose, and to avoid errors in measure names.

Additional suggestions to improve data reporting (beyond current mandates)

- Payment data would be much more usable in spatial format if postal codes were required to be reported (current regulation requires postal codes to be reported “where available”). Because they encode geospatial information, postal codes are much easier to link to open-source spatial databases that allow meaningful analysis than alternative geospatial identifiers like municipality.
- Member State reports would be much more useful if they directly reported NUTS3 region (and stated which NUTS3 version was used in the meta-data), since the documents available from Eurostat did not always convert postcodes to NUTS3 regions with high accuracy. This would enable detailed spatial analyses of CAP payments against other agricultural statistical information from Eurostat.
- Member States should make it possible to directly download the full year’s payment data from their websites (as a few countries have already done).
- A unique European identifier for legal entities would help match recipients across countries and avoid duplicate entries. (The Open Knowledge Foundation Germany notes⁵⁹ that the US and Mexican governments publish unique recipient ID codes that allow tracking the same recipient over different years and different datasets.) Currently, all farms/farmers have an individual ID number, but the system depends on the individual Member State, and there is no common system in the EU (Ronald Hiessrich, Federal Ministry of Food and Agriculture Germany, personal communication, 8 May 2020). A unique identifier would also help address privacy concerns.

New suggestions for reporting administration

Current legislation requires each Member State to report their payment data on their own websites. If Member States follow the recommendations above, their reported data will be much more usable. However, stronger guidance and coordination at the EU level would make the data much easier to analyze than downloading it from 27 separate websites. At a minimum, as noted above, the appropriate EU agency, such as DG AGRI, should produce a spreadsheet template with standardized measure names against which each Member State should submit a dictionary file mapping how their reported measures map to the master measure names. Centralized curation of the data would likely improve accuracy and accessibility, and to make evaluation against result and impact indicators possible. It would also facilitate analysis if the EU made geospatial data on postal codes (.shp files) open source (they are currently proprietary). It would be a great help if centrally reported data were available in data (spreadsheet) format by year and Pillar (instead of 118 separate PDF files for Rural Development Programs across 2014-2020).

Future research

For further analysis of past CAP spending, it would be helpful to expand the work done here to align measures reported during the entire 2014-2020 CAP (beyond 2015) with the standardized measure list. However, acquiring historical CAP spending data from Member States is a challenge, since they are only required to make the data

available for two years, after which most countries seem to remove older data from their transparency websites. Some additional historical data on country-level CAP payments are available from Farmsubsidy.org, although they are not able to archive all years given their current all-volunteer status. With additional resources, further research could be done to scrape and archive the data reported by Member States for the current reporting period. Going forward to the CAP starting 2021, we urge European Commission (DG Agri) and Member States to follow the recommendations above so that such extensive compilation and harmonization will not be necessary to reveal how public money is being spent.

Code Availability

Our Python and R scripts, instructions for accessing the raw data, and associated output data files are available on GitHub at: <https://github.com/kanicholas/CAP-farm-payments>.

Acknowledgements

This research would not have been possible without the hard work of many to assemble the CAP and European agricultural data. We are very grateful to the Open Knowledge Foundation for years of work in obtaining and curating the raw CAP spending data by Member State which we harmonised, standardised, and made spatially explicit here. Thank you to Stefan Wehrmeyer of the Open Knowledge Foundation for helpful correspondence.

We thank our translators Lina Vaitkunaite (Lithuanian), Markus Vihma (Estonian), Katalin Lakatos (Hungarian), Ruxandra Popovici (Romanian), Klara Winkler (German), Kristina Cekanovic (Croatian), Ivaylo Hlebarov (Bulgarian), and Juris Treiberis (Latvian) for assistance with language translations of country measures; and help from Xoco Shinbrot and Kira Treiberis via Twitter to identify a Latvian speaker.

Kind assistance to match reported payments to CAP measures in national languages was provided by: Jan Holub and Jana Ludvíková of the International Relations Unit at the State Agricultural Intervention Fund of the Czech Republic; Jens te Kampe of the Federal Office for Agriculture and Food, Germany; Ilze Palde of the Information and Statistics Division of the Information Department, Rural Support Service, Latvia; Giovanni Belloni at the Twitter account for the Netherlands Enterprise Agency (@RVO_Duurzaam) and Maaïke Rensing at Netherlands Enterprise Agency; and Liina Arumägi, Budget and Analysis Department, Agricultural Registers and Information Board of Estonia.

Thank you to Yves Durinck, Information Systems Support Officer of the European Commission Directorate General for Regional and Urban Policy for sending a very helpful breakdown of 2015 expenditure under the EAFRD by Member State and measure.

Thank you to Ronald Hießerich, of Unit 615 „EU-Agrarfunding, Responsible Authority“ of the Federal Ministry of Food and Agriculture (BMEL), Germany for helpful correspondence regarding payment periods, recipient IDs, and other technical aspects of CAP payment administration.

This research was supported by the Swedish Research Council Grant 2014-5899/E0589901.

Author contributions

Nicholas conceived the study and obtained funding for the project; hired and supervised research assistants in compiling the dataset; obtained translation assistance; supervised the coding; created the standardized measures template from EU legislation; validated the payment dataset including contacting national agricultural offices; and led writing of the paper.

Scown co-designed the study, conducted the spatial analysis and validated geolocation mapping, validated the CAP dataset against Eurostat and European Commission payment reports, and co-wrote the paper.

Brady provided expert CAP input to the study and co-wrote the paper.

Lehsten wrote the Python code to process and standardize the raw subsidies data, and contributed to writing the paper.

Villemoes conducted the initial matching of measure names to the standardized list and performed manual Google translation of measure names, and contributed to writing the paper.

Competing interests

The authors declare no conflict of interest.

Figure Legends

Figure 1. Workflow for generating the data for Common Agricultural Policy (CAP) spending by NUTS3 region within the EU and one of 102 CAP payment measures.

Input files include raw data from farmssubsidy.org; the keys-csv translation file to align country-specific measure names with standardized measure numbers, generated by the research team (this is the “Rosetta Stone” file, table 2, with meta-data about measure names removed; and files to translate postal codes to NUTS regions (downloaded from EU). Steps in the Python code are shown in green. The first phase takes in raw country data from farmssubsidy.org and outputs the translated file, where rows are matched to NUTS3 regions and CAP measures. The second phase extracts the relevant columns from each country file, and for all countries within one year, produces a condensed file of CAP spending by measure per NUTS3. See code and readme file on our GitHub.

Figure 2. Comparison of farm payments for production reported by Eurostat with the payments in our database.

Eurostat payments reported are shown in light green, with our payments shown in both medium green (Pillar I) and dark green (total payments, Pillar I + Pillar II). Payments are for the years reported in Table 1 (2015 except for 4 countries.)

Figure 3. Comparison of Pillar II payments from 2015 reported by the European Commission and the Pillar II payments in our database.

Data from the Commission were obtained by personal communication with Yves Durinck, 14 May 2020 (see Table 10). Note that EC data are all from 2015, and our data are for the years reported in Table 1 (2015 except for 4 countries.)

Figure 4. Comparison of Pillar II payments from 2015 reported by the European Commission and the Pillar II payments in our database, by measure and Member State.

Data from the Commission were obtained by personal communication with Yves Durinck, 14 May 2020 (see Table 10). Each of the 832 datapoints represents a unique PII measure in each country (see Table 10 for 55 negative values that were not plotted due to log transform scale). The grey line is a lowess curve, with 95% confidence interval shaded in light grey. Note that EC data are all from 2015, and our data are for the years reported in Table 1 (2015 except for 4 countries.)

Figure 5. Comparison of CAP payments between Eurostat and our data at the NUTS2 level.

Comparison of CAP payments from 148 NUTS2 regions that were common between our data and Eurostat. Note 3 datapoints were excluded from panel (A) with zeros for Eurostat. Also note Italy NUTS2 regions all excluded with all zeros in Eurostat. The grey line is a lowess curve, with 95% confidence interval shaded in light grey.

Tables

Table 1: CAP reporting data by EU Member State

Overview of the content of the raw data files of CAP payments, data availability, errors for each country, and report of validation data on how much of payments were matched to standardized geolocation and measure names.

Table 2: Rosetta Stone

The full document with classification of the 102 Measures for which payments are made to farmers under the European Common Agricultural Policy, harmonized with original measure names from each country matched to master legislation and unique ID. This file includes translations and alignment of measure names used for all payments made in 2015 (except Denmark, which is reported for 2014, and Sweden, Czech Republic, and Bulgaria, which are reported for 2016). Meta-data on Table 2 is found in Table 4.

Table 3: Regulations underlying the 102 measures in the 2014-2020 CAP

Regulations underlying the 102 measures comprising the 2014-2020 CAP. Note that four of the pieces of legislation underlying the 102 measures expired during the 2014-2020 CAP period, but remained valid for payments through 2015, so Member States in 2015 used a mix of old and new measure terminology in reporting payments; see Section 5.6 of text.

Table 4: Meta-data on the “Rosetta Stone” (Table 2)

Description of columns in Rosetta Stone master document used to match names for CAP Measure payments across 27 EU Member States in 2015, including data source.

Table 5: Google Translate

List of unique measure names in our original CAP payment database, and our translation of them and classification of match certainty to the standardized list of measures, based on initial work and contributions of native speakers. Subsequent matches of about 200 measure names from further analysis, information received from national farm offices, and from the decision to use 2016 data from the Czech Republic after validation illustrated a problem with the 2015 (see text) are not included. All the measure names that were successfully matched are found in the appropriate row for their country’s column in the Rosetta Stone.

Table 6. Measure Correspondence

Correlation table for rural development measures that expired and were replaced during the 2014-2020 CAP program period. The original basis of this table was digitization of Annex I in the European Commission Delegated Regulation (EU) No 807/2014 of 11 March 2014.⁶⁰ We added notes on additional measures that may be potentially ambiguous if only a short name such as “Technical Assistance” is used to identify measures in reporting payments. All countries including a standardized measure coding would eliminate the need for such conversion and ambiguities in the future.

Table 7, Table 8, and Table 9. Condensed dataset attributing CAP payments to NUTS3 regions and standardized measures for 2014, 2015, and 2016, respectively.

These files are the final output of the workflow shown in Figure 1. Please see GitHub for full documentation and raw input and output files at finer resolution.

Table 10. Comparison of our total payment data to EU reported totals

Comparisons of our payment totals in our database with Eurostat (Pillar I) and European Commission (Pillar II) for the same years, rounded to the nearest euro.

Table 11. Eurostat data on subsidies on production

Extracted 15 June 2020 from “Economic accounts for agriculture by NUTS 2 regions” and used to validate our dataset.

Table 12. Country-level comparisons between our data and EU reported data

Data used for validation of our spatially explicit and measure-specified dataset against existing reported data, grouped by Member State and CAP Pillar.

Table 13. Country-level comparisons between our data and EU reported data

Comparison between Pillar II payments reported in our database for the given year centered around 2015, and those reported by the European Commission in 2015 broken down by Measure (personal communication, Yves Durinck, European Commission Directorate General for Regional and Urban Policy, 14 May 2020). This table includes 887 unique combinations of spending by measures within country. Note that the 55 negative values listed here were not plotted in the log-transformed figure.

References

- ¹ IPBES (2019): Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. S. Díaz, J. Settele, E. S. Brondizio E.S., H. T. Ngo, M. Guèze, J. Agard, A. Arneth, P. Balvanera, K. A. Brauman, S. H. M. Butchart, K. M. A. Chan, L. A. Garibaldi, K. Ichii, J. Liu, S. M. Subramanian, G. F. Midgley, P. Miloslavich, Z. Molnár, D. Obura, A. Pfaff, S. Polasky, A. Purvis, J. Razzaque, B. Reyers, R. Roy Chowdhury, Y. J. Shin, I. J. Visseren-Hamakers, K. J. Willis, and C. N. Zayas (eds.). IPBES secretariat, Bonn, Germany. 56 pages. https://ipbes.net/sites/default/files/2020-02/ipbes_global_assessment_report_summary_for_policymakers_en.pdf
- ² IPCC, 2019: Summary for Policymakers. In: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, J. Malley, (eds.)]. https://www.ipcc.ch/site/assets/uploads/sites/4/2020/02/SPM_Updated-Jan20.pdf
- ³ Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S.,..., Wood, A. (2019). Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. *Lancet*, 393(10170), 447–492.
- ⁴ Recanatì, F., C. Maughan, M. Pedrotti, K. Dembska, M. Antonelli, Assessing the role of CAP for more sustainable and healthier food systems in Europe: A literature review. (2019). *Science of The Total Environment*, Volume 653, Pages 908-919, <https://doi.org/10.1016/j.scitotenv.2018.10.377>.
- ⁵ FOLU & Coalition, T. F. and L. U. *Growing Better: Ten Critical Transitions to Transform Food and Land Use*. <https://www.foodandlandusecoalition.org/wp-content/uploads/2019/09/FOLU-GrowingBetter-GlobalReport.pdf> (2019).
- ⁶ https://ec.europa.eu/info/sites/info/files/european-green-deal-communication_en.pdf
- ⁷ European Commission: The Common Agricultural Policy at a glance. https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/cap-glance_en
- ⁸ Pe'er, G. *et al.* A greener path for the EU Common Agricultural Policy. *Science* (80-.). **365**, 449–451 (2019).
- ⁹ Scown, Murray W, Mark V Brady and Kimberly A Nicholas. In review. “€24.2 billion in misspent EU agricultural subsidies could support Sustainable Development Goals.” *One Earth*.
- ¹⁰ EU. Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL establishing rules on support for strategic plans to be drawn up by Member States under the Common agricultural policy (CAP Strategic Plans). vol. 10103/1/19 (2019).
- ¹¹ EC. The Future of Food and Farming. (2017).
- ¹² European Parliament (2020), *Financing of the CAP* [Online]. European Parliament. Available: <https://www.europarl.europa.eu/factsheets/en/sheet/106/financing-of-the-cap> [Accessed 2020-05-06].
- ¹³ Andersson, A., S. Höjgård and E. Rabinowicz (2017). "Evaluation of results and adaptation of EU Rural Development Programmes." *Land Use Policy*, 67: 298-314.
- ¹⁴ WWF (2015). "A Common Environment and Rural Policy." WWF Baltic Ecoregion Programme 2010.
- ¹⁵ European Parliament and Council, 2013. REGULATION (EU) No 1306/2013 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 17 December 2013 on the financing, management and monitoring of the common agricultural policy and repealing Council Regulations (EEC) No 352/78, (EC) No 165/94, (EC) No 2799/98, (EC) No 814/2000, (EC) No 1290/2005 and (EC) No 485/2008. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32013R1306&from=en>
- ¹⁶ European Commission. Undated. "Common Agricultural Policy: Controls and Transparency." https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/financing-cap/financial-assurance_en
- ¹⁷ European Parliament and Council, 2013. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32013R1306&from=en>
- ¹⁸ Miljø-og Fødevarerministeriet Landbrugsstøtte - See text under “Hvad bliver ikke offentliggjort” (“What will not be published?”) <https://lbst.dk/tilskud-selvbetjening/soeg-i-registre/modtagere-af-estoeette/landbrugsstoeette/#c33575>

- ¹⁹ The National Archives. Undated. “Open Government License for public sector information.” <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>
- ²⁰ European Commission. Undated. “Beneficiaries of CAP funds: Beneficiaries by country.” https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/financing-cap/controls-and-transparency/beneficiaries_en
- ²¹ Open Knowledge Foundation Germany. Undated. “Farmsubsidy.org at a glance.” <https://farmsubsidy.org/about/>
- ²² Open Knowledge Foundation Germany. Farmsubsidy.org FAQs. <https://farmsubsidy.org/faq/>
- ²³ OpenSpending/Holger Drewes. 2014. Farmsubsidy.org Developer Documentation. <https://farmsubsidy.readthedocs.io/en/latest/>
- ²⁴ Stefan Wehrmeyer. Undated. “Scrapers for FarmSubsidy data in the member states.” <https://github.com/openspending/farmsubsidy-scrapers>
- ²⁵ Open Knowledge Foundation Germany. Farmsubsidy.org FAQs. <https://farmsubsidy.org/faq/>
- ²⁶ Open Knowledge Foundation Germany. Farmsubsidy.org FAQs. <https://farmsubsidy.org/faq/>
- ²⁷ See FAQs about CAP fiscal year here: <https://www.agriculture.gov.ie/agri-foodindustry/euinternationalpolicy/commonagriculturalpolicy/cap/capbeneficiariesdatabase/>
- ²⁸ Bundesanstalt für Landwirtschaft und Ernährung - Empfänger EU-Agrarfonds – Suche. <https://www.agrar-fischerei-zahlungen.de/Suche>
- ²⁹ European Parliament and Council, 2013. “REGULATION (EU) No 1306/2013 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 17 December 2013 on the financing, management and monitoring of the common agricultural policy and repealing Council Regulations (EEC) No 352/78, (EC) No 165/94, (EC) No 2799/98, (EC) No 814/2000, (EC) No 1290/2005 and (EC) No 485/2008.”
- <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32013R1306&from=en>
- ³⁰ Fact sheets on the European Union - Second pillar of the CAP: Rural Development Policy. <https://www.europarl.europa.eu/factsheets/en/sheet/110/second-pillar-of-the-cap-rural-development-policy>
- ³¹ European Parliament (2020), *Financing of the CAP* [Online]. European Parliament. Available: <https://www.europarl.europa.eu/factsheets/en/sheet/106/financing-of-the-cap> [Accessed 2020-05-06].
- ³² Pe’er, G., Lakner, S., Müller, R., Passoni, G., Bontzorlos, V., Clough, D., ..., Zinngrebe, Y. (2017). Is the CAP Fit for Purpose? An Evidence-Based Fitness-Check Assessment. German Centre for Integrative Biodiversity Research (iDiv).
- ³³ Fact sheets on the European Union - Second pillar of the CAP: rural development policy. <https://www.europarl.europa.eu/factsheets/en/sheet/110/second-pillar-of-the-cap-rural-development-policy>
- ³⁴ European Commission. “Controls and Transparency.” <https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/financing-cap/controls-and-transparency>
- ³⁵ Miljø-og Fødevarerministeriet Landbrugsstøtte - See text under “Hvad bliver ikke offentliggjort” (“What will not be published?”) <https://lbst.dk/tilskud-selvbetjening/soeg-i-registre/modtagere-af-eu-stoette/landbrugsstoette/#c33575>
- ³⁶ Open Knowledge Foundation Germany. Farmsubsidy.org FAQs. <https://farmsubsidy.org/faq/>
- ³⁷ Bundesanstalt für Landwirtschaft und Ernährung. Empfänger EU-Agrarfonds – Suche. https://www.agrar-fischerei-zahlungen.de/agrar_suche_hilfe.html
- ³⁸ Hellenic Republic Ministry of Reconstruction of Production, Environment, and Energy “CAP BENEFICIARIES PAYMENTS” <https://transpay.oepeke.gr/>
- ³⁹ Nomenclature of Territorial Units for Statistics. In Wikipedia. https://en.wikipedia.org/wiki/Nomenclature_of_Territorial_Units_for_Statistics
- ⁴⁰ Eurostat. (n.d.). *History of NUTS*. <https://ec.europa.eu/eurostat/web/nuts/history> (Eurostat).
- ⁴¹ Eurostat (n.d.). *Postcodes and NUTS*, <https://ec.europa.eu/eurostat/web/nuts/correspondence-tables/postcodes-and-nuts> (Eurostat).
- ⁴² Sweden’s postal code numbering system https://upload.wikimedia.org/wikipedia/commons/e/ee/2_digit_postcode_sweden.png
- ⁴³ Eurostat “SVERIGE - NUTS level 3” <https://ec.europa.eu/eurostat/documents/345175/7451602/2016-NUTS-3-map-SE.pdf>
- ⁴⁴ European Central Bank “Euro foreign exchange reference rates” https://www.ecb.europa.eu/stats/policy_and_exchange_rates/euro_reference_exchange_rates/html/index.en.html
- ⁴⁵ DG Agri. 2018. “Transparency/ Working Document Rev6 Ares(2018)1833750 - 04/04/2018” [http://cap-payments.defra.gov.uk/Download/Ares\(2018\)1833750-rev6-Measures_description-EN.pdf](http://cap-payments.defra.gov.uk/Download/Ares(2018)1833750-rev6-Measures_description-EN.pdf)

-
- ⁴⁶ https://twitter.com/KA_Nicholas/status/1231969619465900041?s
- ⁴⁷ Transparenzdatenbank EU “Information zur Veröffentlichung von Zahlungen im Rahmen der Gemeinsamen Agrarpolitik der EU” <https://www.transparenzdatenbank.at/>
- ⁴⁸ EAFRD (Pillar II) payments for the programming period 2007-2013 were allowed until December 31, 2015, according to Art. 71 para. 1 VO (EU) Nr. 1698/2005.
- ⁴⁹ European Commission. (2019). *EU Expenditure and Revenue, 2014-2020*. https://ec.europa.eu/budget/graphs/revenue_expenditure.html (European Commission, Brussels).
- ⁵⁰ European Commission “Rural development programmes by country” https://ec.europa.eu/agriculture/sites/agriculture/files/rural-development-2014-2020/country-files/common/number-of-rdp-per-country_en.pdf
- ⁵¹ <https://ec.europa.eu/agriculture/rural-development-2014-2020/country-files/>
- ⁵² https://ec.europa.eu/info/sites/info/files/food-farming-fisheries/key_policies/documents/number-of-rdp-per-country-2014-20_en.pdf, <https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/rural-development/country>
- ⁵³ European Network for Rural Development “Measure 323 – Conservation and upgrading of the rural heritage” https://enrd.ec.europa.eu/sites/enrd/files/assets/pdf/measure-information-sheets/2014-06-19/C_Infosheet_323.pdf
- ⁵⁴ European Network for Rural Development “Priority & Focus Area Summaries” https://enrd.ec.europa.eu/policy-in-action/rural-development-policy-figures/priority-focus-area-summaries_en
- ⁵⁵ European Commission “Sweden's Rural Development Programme for 2014-2020 approved by European Commission” https://ec.europa.eu/info/sites/info/files/food-farming-fisheries/key_policies/documents/rdp-sweden-press-summary-26-05-2015_en.pdf
- https://ec.europa.eu/agriculture/sites/agriculture/files/rural-development-2014-2020/country-files/se/factsheet_en.pdf
- ⁵⁶ https://ec.europa.eu/agriculture/sites/agriculture/files/rural-development-2014-2020/country-files/common/rdp-list_en.pdf
- https://ec.europa.eu/agriculture/sites/agriculture/files/rural-development-2014-2020/country-files/common/rdp-list_en.pdf
- ⁵⁷ European Parliament, 2019, p. 4 http://www.europarl.europa.eu/ftu/pdf/en/FTU_3.2.6.pdf
- ⁵⁸ Open Knowledge Foundation Germany. Farmsubsidy.org FAQs. <https://farmsubsidy.org/faq/>
- ⁵⁹ Open Knowledge Foundation Germany. Farmsubsidy.org FAQs. <https://farmsubsidy.org/faq/>
- ⁶⁰ Commission Delegated Regulation (EU) No 807/2014 of 11 March 2014 supplementing Regulation (EU) No 1305/2013 of the European Parliament and of the Council on support for rural development by the European Agricultural Fund for Rural Development (EAFRD) and introducing transitional provisions. Stable ELI: https://eur-lex.europa.eu/eli/reg_del/2014/807/oj

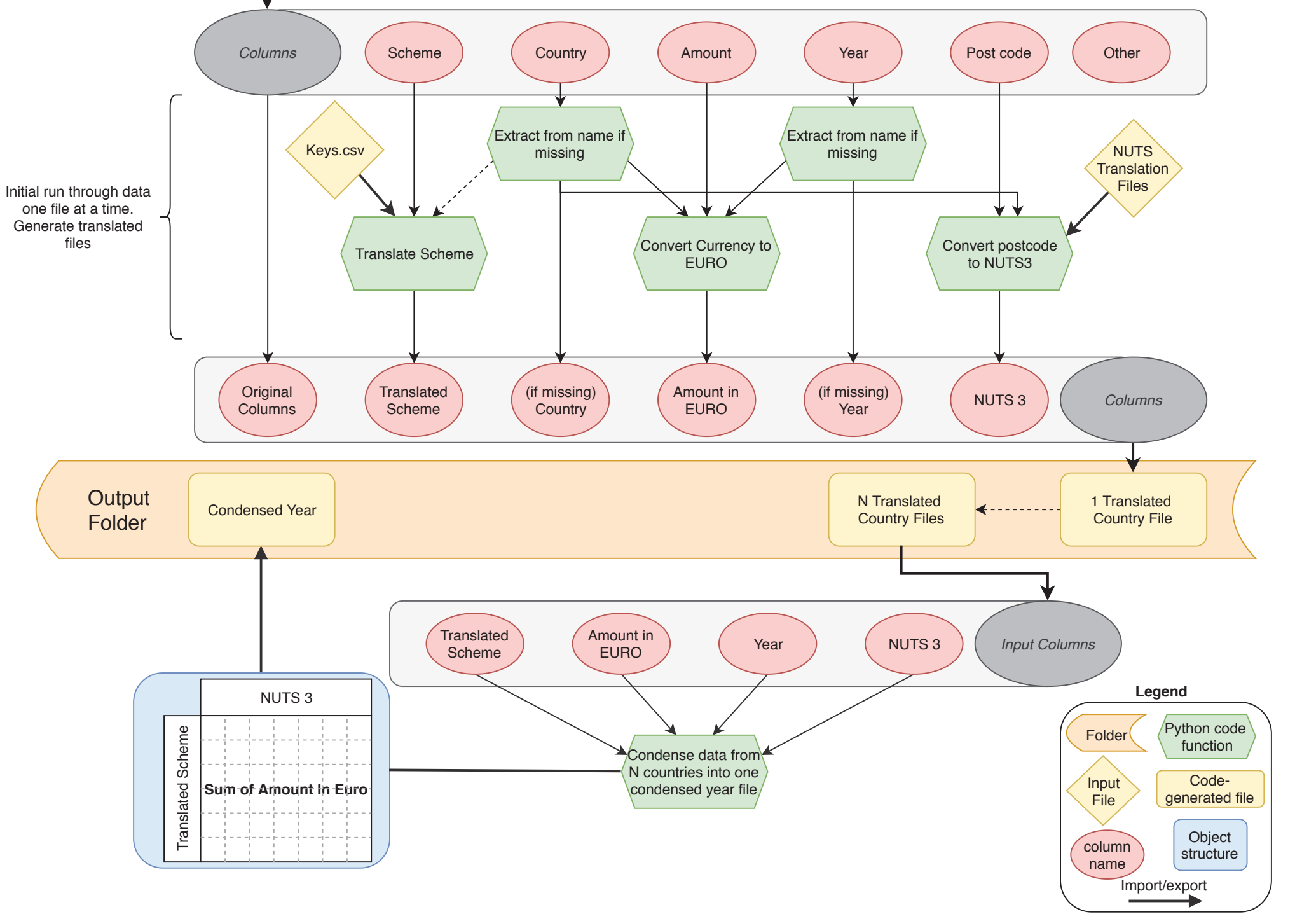


Figure 2 [Click here to access/download;Figure;Fig2_NUTS0_comparison_](#)

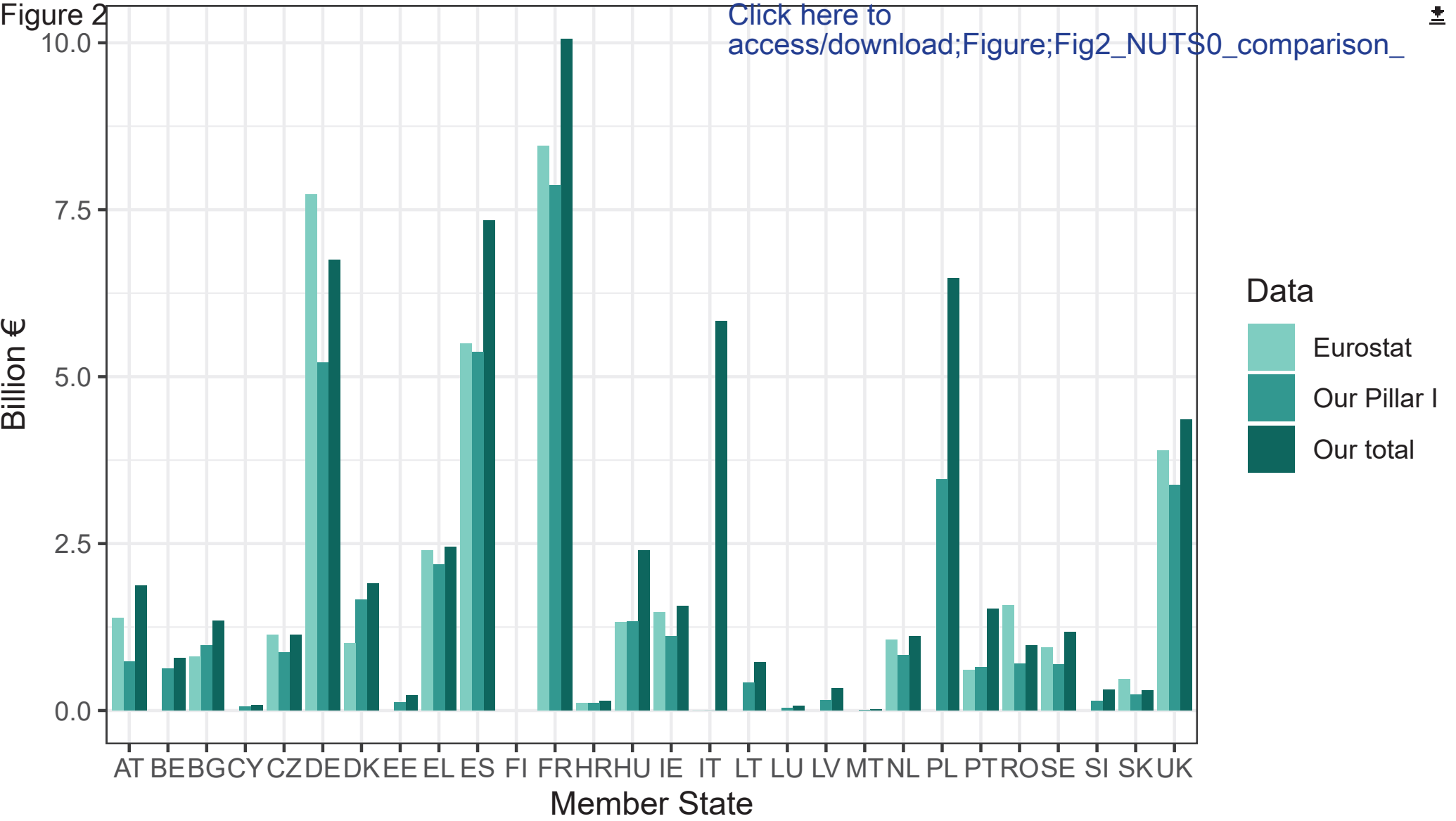


Figure 3 [Click here to access/download;Figure;Fig3_NUTS0_comparison_](#)

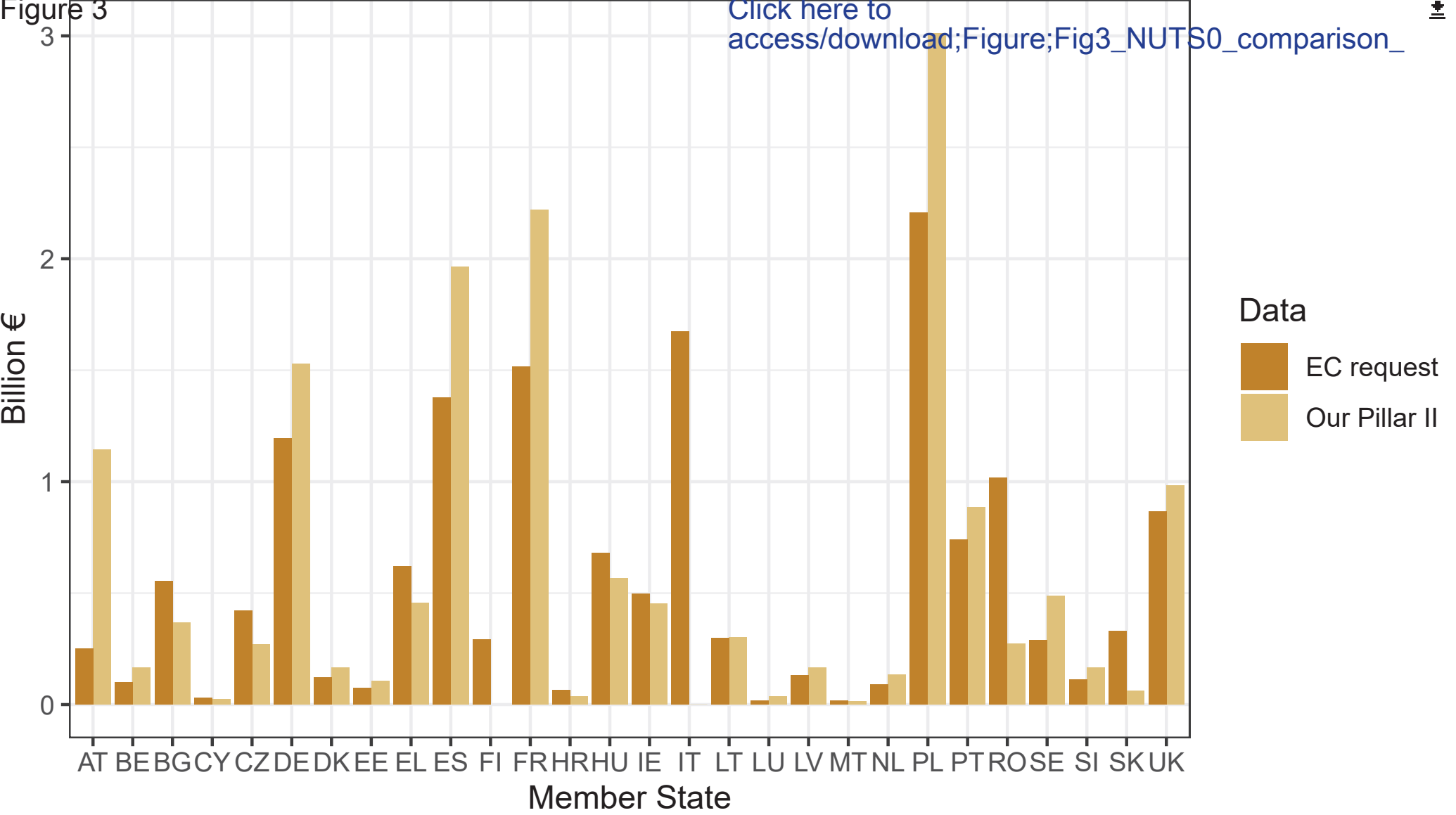


Figure 4

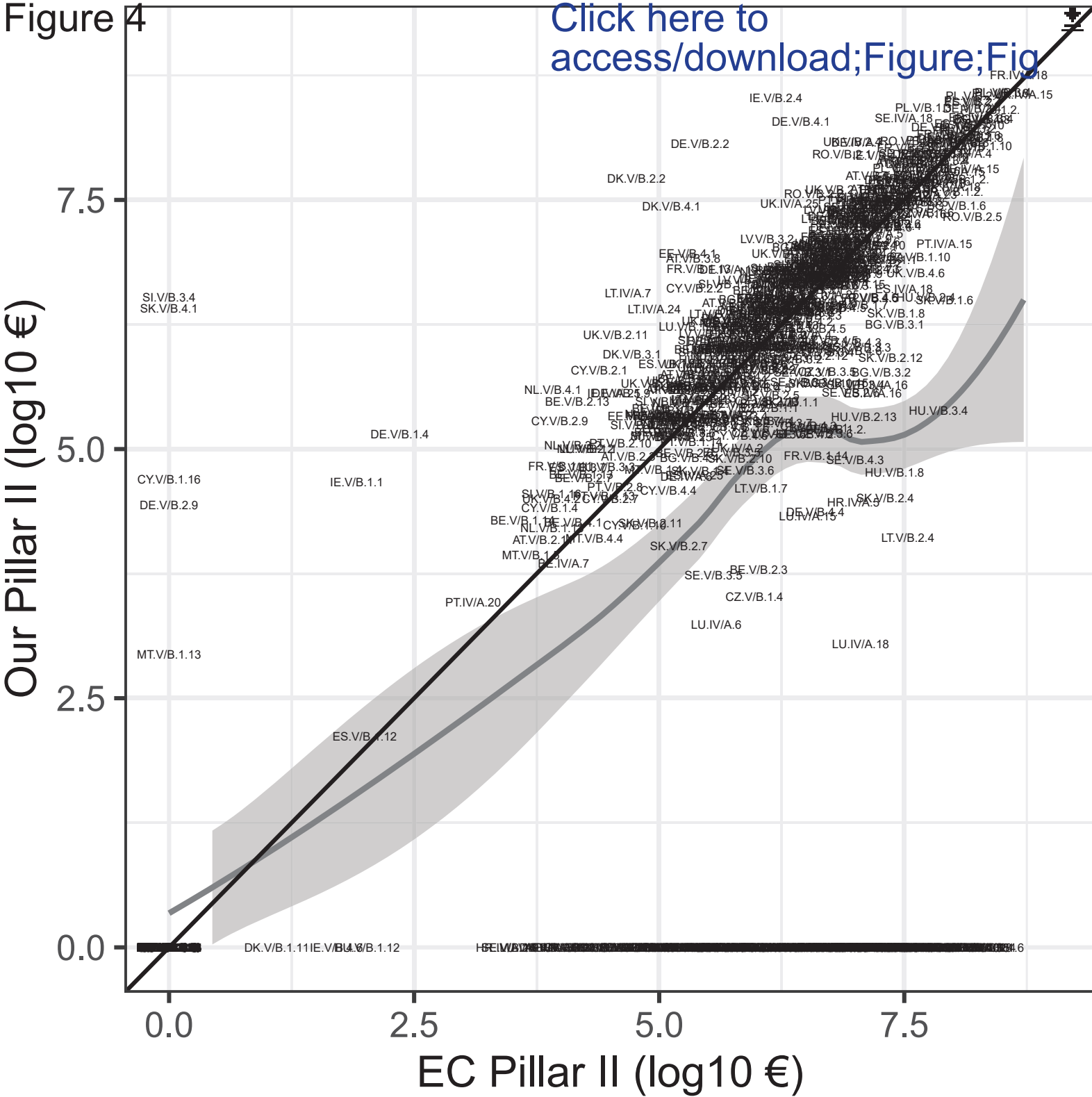
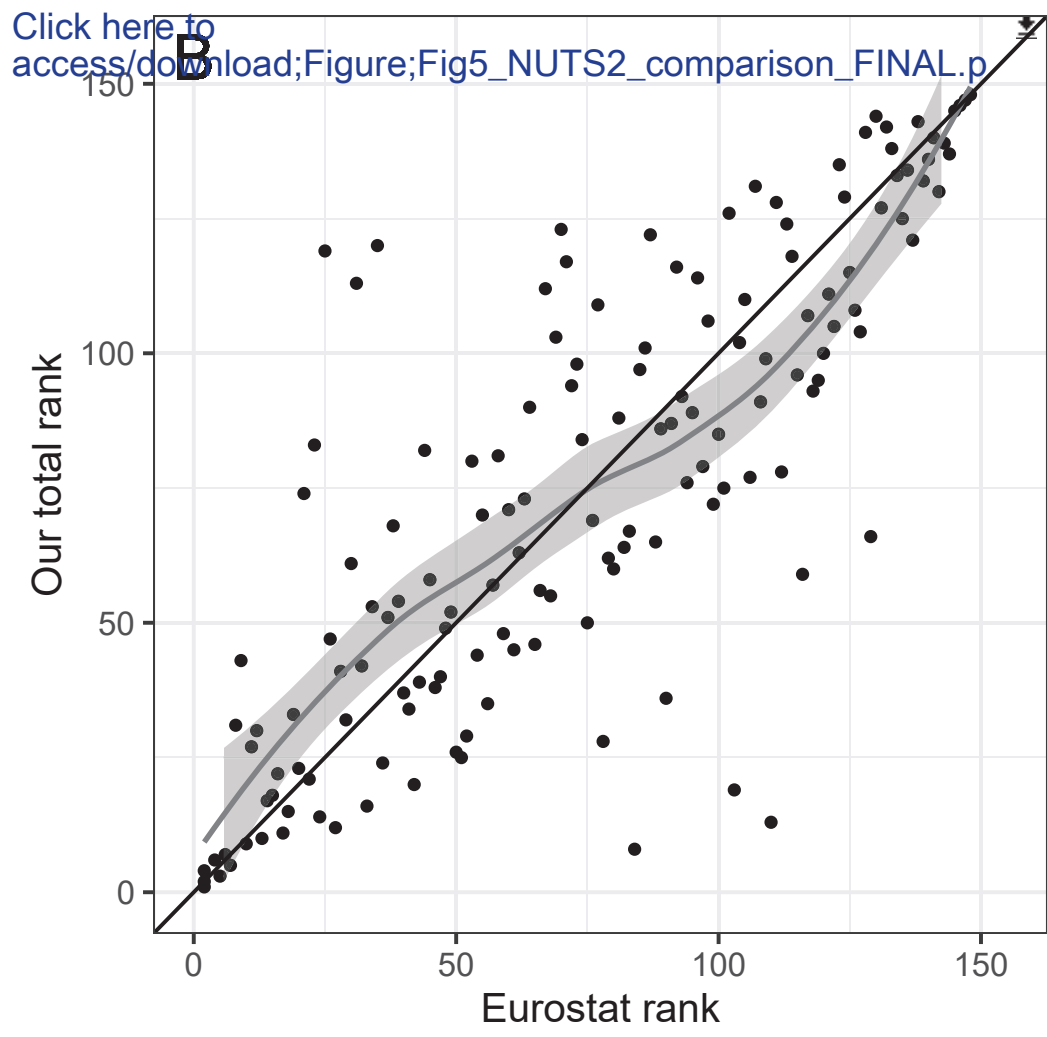
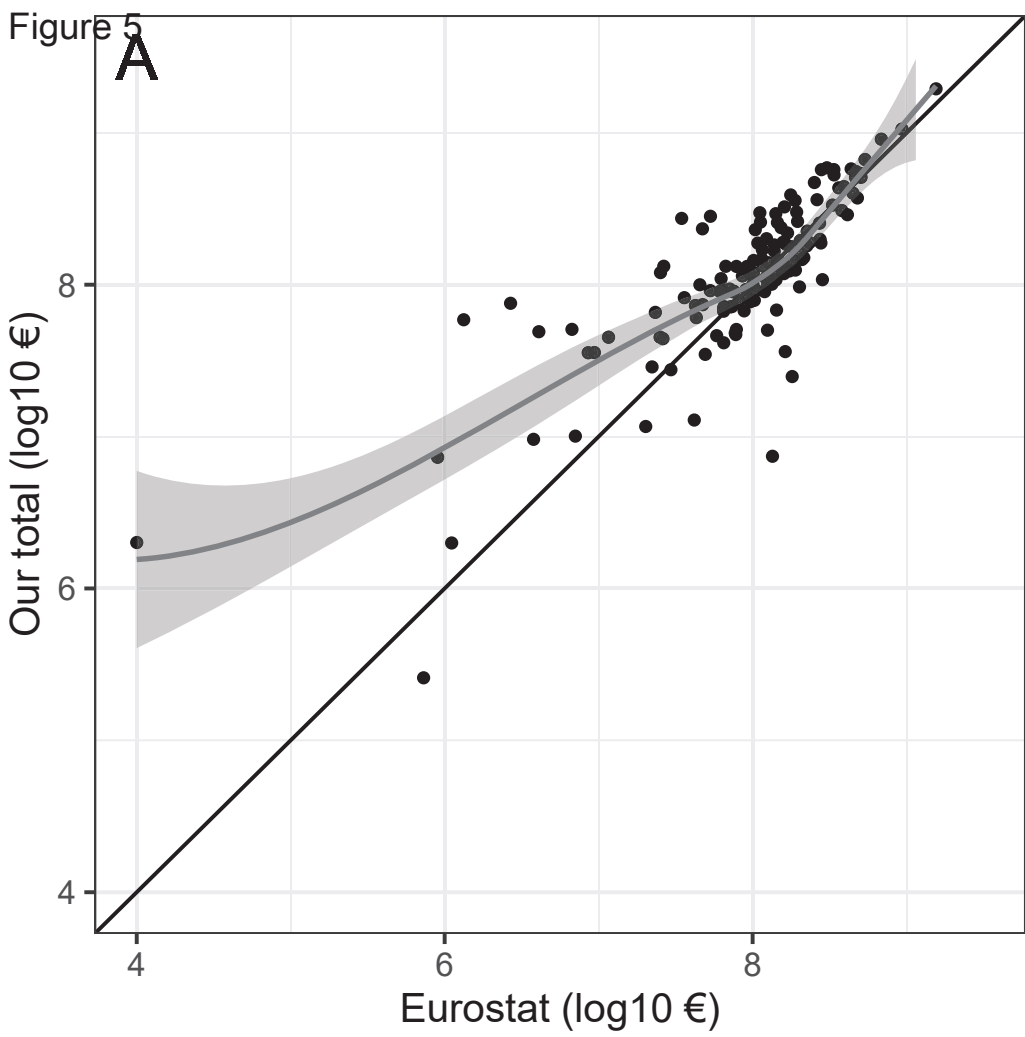


Figure 5



[Click here to access/download;Figure;Fig5_NUTS2_comparison_FINAL.p](#)