Baumgartner, Christoph; Srhoj, Stjepan; Walde, Janette

**Working Paper**

# Harmonization of product classifications: A consistent time series of economic trade activities

Working Papers in Economics and Statistics, No. 2022-11

**Provided in Cooperation with:**
Institute of Public Finance, University of Innsbruck

*Suggested Citation:* Baumgartner, Christoph; Srhoj, Stjepan; Walde, Janette (2022) : Harmonization of product classifications: A consistent time series of economic trade activities, Working Papers in Economics and Statistics, No. 2022-11, University of Innsbruck, Research Platform Empirical and Experimental Economics (eeecon), Innsbruck

This Version is available at:
https://hdl.handle.net/10419/273684

universität innsbruck

Faculty of
Economics and Statistics


# Harmonization of product classifications: A consistent time series of economic trade activities

**Christoph Baumgartner, Stjepan Srhoj and Janette Walde**

Working Papers in Economics and Statistics

2022-11

University of Innsbruck
http://uibk.ac.at/eeecon/

**University of Innsbruck**
**Working Papers in Economics and Statistics**

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
Faculty of Economics and Statistics
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 96136
E-mail: Dean-EconStat@uibk.ac.at

The most recent version of all working papers can be downloaded at
https://www.uibk.ac.at/eeecon/wopec/

For a list of recent papers see the backpages of this paper.

# Harmonization of product classifications: A consistent time series of economic trade activities

Christoph Baumgartner*, Stjepan Srhoj† & Janette Walde*

**Abstract:** Firm-product data provide information for various research questions in international trade or innovation economics. However, working with this data requires harmonizing product classifications consistently over time to avoid internal validity issues. The researcher must consider product code changes in the classification systems over the observation period. Harmonization is required because classification systems like the EU classifications Combined Nomenclature (CN) for goods or the Prodcom for the production of manufactured goods undergo several changes. We have addressed this problem and developed an approach to harmonize product codes. This approach tracks product codes from 1995 to 2022 for CN and 2001 to 2021 for Prodcom. Additional years can be conveniently updated. We provide the harmonized product codes for CN and Prodcom in the selected period's first (or last) year. Our approach is summarized in an easy-to-use R package so that researchers can consistently track product codes for their period. We demonstrate the importance of harmonization using the micro-level trade data for Croatia as a case study. Our approach facilitates working with firm-product data, allowing the analysis of important research questions.

**Keywords:** concordance, CN, HS, Prodcom, BEC

**JEL Classification:** C81, F14, O12, C55

## Acknowledgements

The authors thank the staff of the *Croatian Statistics Office* for data access in the safe room and for their data assistance. This work was supported by the Research Platform *Empirical and Experimental Economics* of the University of Innsbruck and by the *Croatian Science Foundation* within the project IP-CORONA-2020-12-1064.

*Department of Statistics, Faculty of Economics and Statistics, University of Innsbruck, Austria

†Department of Economics, Faculty of Economics, Business and Tourism, University of Split, Croatia

# 1 Introduction

Innovation and international trade are considered a major engine of economic growth (Aghion et al., 2021; Wagner, 2019). Innovation can take a variety of forms, including product, production or delivery, organizational, marketing or communication innovation (Gault, 2018). When an economist thinks of innovation, product innovations, like (autonomous) automobiles, computers, smartphones, robots, or drones, come first to mind. We focus on enabling investigations concerning product innovation and firm-product relations in general. For the purpose of analysing product innovation, studies mostly use the Community Innovation Survey (CIS)[1] data (e.g. Schubert & Tavassoli, 2020). CIS is considered one of the richest sources of firm innovation data. Researchers use CIS data to investigate characteristics of firms introducing product innovations, like research and development (R&D) expenditures, firm size, firm age, industry, or profitability (e.g., Griffith et al., 2006; Mairesse & Mohnen, 2010; Coad et al., 2016; Audretsch & Belitski, 2020). CIS data is also used to conduct policy evaluations of large public R&D grants or public procurement for innovation on the firms innovation output (Mairesse & Mohnen, 2002; Stojčić et al., 2020). Although availability of CIS data has had a major impact on the field of innovation economics, CIS data suffers from at least two weaknesses. Firstly, only a small sample of relatively larger firms is collected each cycle, and secondly, variables are collected via survey, thus, they are self-reported by the firms and can sometimes be imprecise. Furthermore, "the CIS microdata together with regional identifiers are not or at least not easily accessible to researchers" (Hauser et al., 2018, p. 50).

Transaction data from customs (Combined Nomenclature, CN) and the Prodcom data can be used as an alternative (European Commission, 2021, 2019). Research questions concerning firms' changes in productivity over time and its determinants, firms' product mix strategy, or R&D grants impact on the number of new exported products are just three examples that can be analyzed with firm-product data.[2] This data has several advantages to self-reported innovation measures from surveys. It is considered more objective, i.e., a firm either trades a particular product or does not, and whether the product is new to the firm and/or market can be identified in the time series of the product codes. Transaction and Prodcom data provide value and quantity of imported, exported, or sold products at a very detailed product level. This data enables the analysis of larger firm samples. In addition, Prodcom data includes a substantial part of all domestically traded products. Thus, insights from transaction and Prodcom data can be used to investigate the industry-level importance of innovative products. Identified new products can be classified into new-to-firm and new-to-market products, combined with the value of these sold or exported products providing additional insights into the magnitude of innovation.

Wagner (2019) classified the development of literature on international trade and firms into two major waves. The first wave focused on firm-level characteristics of ex-

[1] Link:
`https://ec.europa.eu/eurostat/web/microdata/community-innovation-survey` [Accessed May $23^{rd}$ 2021]

[2] Customs data (CN) includes information at the firm-product-market level, while Prodcom includes information at the firm-product level, we, therefore, use the term "firm-product data" to describe both datasets, although CN adds another (market) dimension.

porters, including questions such as whether exporters are more productive, larger, or paying higher wages compared to domestic firms or only importers. The second wave has enriched the first wave by using transaction data at the firm-product-market level. Research questions like "*Who trades how much of which goods (or service) with whom?*" can be investigated (Wagner, 2016, p. 216). This data enables to consider the distance between exporter and destination market (Mayer & Zignago, 2011), the type of exports (Wagner, 2019) and frequency of simultaneous exports and imports (Bernard et al., 2018).

Our study speaks to the second wave of the literature on international trade and firms (Wagner, 2019). Similarly, literature on innovation and firms can also be classified into two major waves. While the first wave of literature on innovation and firms uses survey data, the second wave uses firm-product data. Firm-product data is observed over a (long) period and therefore provides possibilities to develop product measures and measures of product innovation. By merging with firm-level data, studies can investigate possible drivers of new (export) products and innovation measures. Innovation literature is yet to embrace the transaction data to compose more objective firm-level innovation outcomes and usually neglected variables, like frequency of innovation, number of products in the firms' product mix, distances to the destination markets, product prices, profit margins, product quality, and products imported. Castellani & Fassio (2019) is one of the few innovation studies using firm-product-market level data. They document the positive effects of importing new intermediate inputs on small firms' new exports in Sweden. Such studies introduce the second wave of research in economics of innovation.

These previous ideas and existing research, although scarce, has motivated us to harmonize transaction and Prodcom data in order to allow econometric analysis using firm-product data consistently over time. Van Beveren et al. (2012) already suggested an approach to harmonize product codes and provided Stata code for enabling researchers to prepare their data accordingly for further analyses. Unfortunately, these files are no longer publicly available. We enhanced their approach. Both the entire history of the product codes and the harmonized product code for further analyses dependent on the period of interest are provided. In addition, we provide so-called *utilize functions* to count the number of added, kept, and dropped products of each firm from one year to the next, together with the broad economic categories (BEC) classification (United Nations, 2016). Furthermore, all necessary computations are comprised in an R-package that can be downloaded and conveniently employed. Therefore, we invite and enable researchers to channel and intensify their research efforts in the second waves of literature on both, international trade and firms, as well as innovation and firms. The importance of harmonization is demonstrated by applying our approach on Croatian data. We obtain a similar substantial discrepancy between the raw time series and the harmonized time series as Van Beveren et al. (2012) for Belgian firm-product data and Link (2020) demonstrated for ifo Business Survey (IBS) data.

In the remainder of the manuscript, we describe the classification systems, present the idea of harmonization, and demonstrate the importance of harmonizing product codes by applying the approach to firm-product data from Croatia. Finally, we

Table 1: Example of a CN8 code for 2020

| Classification system | Digits | Code | Description |
|---|---|---|---|
| HS | 2 | 84 | Chapter 84 - Nuclear reactors, boilers, machinery and mechanical appliances; parts thereof |
| HS | 4 | 8421 | Centrifuges, including centrifugal dryers; filtering or purifying machinery and apparatus, for liquids or gases |
| HS | 6 | 842129 | Machinery and apparatus for filtering or purifying liquids (excl. such machinery and apparatus for water and other beverages, oil or petrol-filters for internal combustion engines and artificial kidneys) |
| CN | 8 | 84212920 | Made of fluoropolymers and with filter or purifier membrane thickness not exceeding 140 microns |

Source: European Commission (2021)

summarize the main features of the approach.

# 2 Harmonization of product codes

Harmonization is based on classification lists and concordance tables provided by Eurostat. There are several classification systems used in the European Union (EU) for trade and production. Although the systems are designed to be similar, there are important differences between the two systems, both at a given time and across several years.

## 2.1 Combined Nomenclature

EU international trade statistics records the value and quantity of products traded between EU member states, and from EU member states to non-EU countries. At the customs, products must be classified according to the Combined Nomenclature, which has 8-digits and thus denoted CN8. Its first six digits match the classification system of the (international) Harmonized System (HS), which is maintained by the World Customs Organization (WCO) WCO (2022). The European classification system CN8 is an extension of the HS6 classification system, analogous to the 10-digit extensions (HS10) used in the USA. Table 1 provides an example for the different product codes and their number of digits. The CN classification was developed to meet the requirements of both the Common Customs Tariffs and the EU's external trade statistics. In intra-EU trade statistics the CN is also used (Directorate-General for Taxation and Customs Union, 2022).

The HS undergoes periodic revisions. Between 1988 and 2022, it was updated five times (in 1996, 2002, 2007, 2012, and 2017). Revision years for HS6 also tend to be years of substantial changes in the CN8 classification. Moreover, there are annual updates of the CN8 classification, such that from one year to the next a product may have a different CN8 code. Such updates are motivated by policy, development of technology, or statistical requirements.

## 2.2 Prodcom classification

Firms within the EU must report their industrial production and services in the Prodcom survey, which are specified on the Prodcom list. Although Prodcom regulation is EU-based, this firm-product data is obtained by the National Statistics Institutes of the member states. If the member state considers Prodcom reporting as an administrative burden for the firms, they are able to alleviate the reporting requirement. However, the information in Prodcom covers at least 90% of national production within each sector defined by NACE 4d.

The Prodcom list was developed to measure production in the EU Member States and to allow a comparison between production and external trade statistics (Eurostat, 2006a). Therefore, Prodcom classification is closely related to the Combined Nomenclature classification. Similar to the CN classification, also the Prodcom system is an extension of other systems. The first four digits are taken from the statistical classification of economic activities in the European Community (NACE), while the digits 5 and 6 are in line with the Classification of Products by Activity (CPA). The last two digits then make up the actual Prodcom code (cf. Table 2).

Table 2: Example of a Prodcom code for 2020

| Classification system | Digits | Code | Description |
|---|---|---|---|
| NACE | 4 | 0710 | Mining of iron ores |
| CPA | 6 | 071010 | Iron ores |
| Prodcom | 8 | 07101020 | Iron ores and concentrates; agglomerated (excluding roasted iron pyrites) |

Source: European Commission (2019)

As with CN, the Prodcom codes are also subject to annual changes. These changes also include newly established Prodcom codes or definitely dropped codes. Thus, the coverage may change. Products may be covered by Prodcom codes in one year, but not covered by any Prodcom code in another year. It is impossible to keep track of the amount or value of these codes over time, so they need to be dropped from the data when harmonizing.

## 2.3 Classification by Broad Economic Categories

The broad economic categories (BEC) system classifies products into broad economic categories used internationally. This classification was introduced in 1971 and has undergone several revisions, with the latest revision number 5 in 2016. Today, its development and maintenance are done by the United Nations Statistics Division and is used worldwide. Since the most recent revision, BEC codes have eight categories and contain up to six digits instead of three as they used to. The leading digit indicates the main economic class of the product. The second digit distinguishes between goods and services within the eight main categories. Moreover, BEC codes can be classified into three basic classes defined by the System of National Accounts (SNA), which focus on the end-use of the product. The classes are called “capital

Table 3: Example of a BEC code for Rev. 5

| Classification system | Digits | Code | Description |
|---|---|---|---|
| BEC | 1 | 6 | ICT, media, computers, business and financial services |
| BEC | 2 | 61 | Goods |
| BEC | 3 | 611 | Intermediate Consumption |
| BEC | 4 | 6112 | Processed |
| BEC | 6 | 611220 | Specified |

Source: United Nations (2016)

goods", "intermediate goods" and "consumption goods" in revision 4 of BEC, while they are called "intermediate consumption", "gross fixed capital formation" and "final consumption" since revision 5. The SNA classification is the third digit in the BEC codes. The last three digits are the "processing dimension", the "specification dimension" and finally, the "durability dimension" (for an example, see Table 3).

Concordance files between HS6 and BEC Rev. 4 exist for 1996, 2002, and 2007. For 2012 and 2017, there exists a concordance between HS6 and BEC Rev. 5. Therefore, we provide BEC codes from Rev. 4 until 2011 and BEC codes from Rev. 5 thereafter. Figure 1 illustrates the connection as well as the availability of concordance files between the classification systems described in this paper [3]. Recently (simultaneously to our approach[4]) Duprez & Magerman (2022) provide concordance files for CN8 and PC8 on github for 2002 to 2014 [5]. We used nevertheless the original concordance files provided by the RAMON server.

## 2.4 Harmonization

Van Beveren et al. (2012) developed an approach to harmonize CN8, HS6, and PC8 product codes. The "concordance", as they call it, was done from 1988 to 2010 for CN8 and from 1993 to 2010 for PC8, respectively. The coding was implemented in Stata and was also available for download. However, the Stata files are no longer publicly available to the best of our knowledge.

In contrast to Van Beveren et al. (2012) who provided previously harmonized product codes, our approach provides the entire history of each product code (CN8, HS6, PC8) and the BEC classification with a flexible way to extend the covered period by adding future concordance files. The complete harmonization approach is available as an R package, called *harmonizer*[6]. Harmonization of CN8 codes is

[3] A list of all changes is annually provided by *Eurostat* on the EU server RAMON: `https://ec.europa.eu/eurostat/ramon`

[4] We published our R package *harmonizer* for the first time December $8^{th}$ 2021; Link: `https://cran.r-project.org/src/contrib/Archive/harmonizer/`

[5] Link: `https://github.com/glennmagerman/Concordances_CN_PC` [Accessed May $18^{th}$ 2022] These concordance files can conveniently be included by the researchers in our R package if requested.

[6] Downloadable via CRAN: `https://CRAN.R-project.org/package=harmonizer`

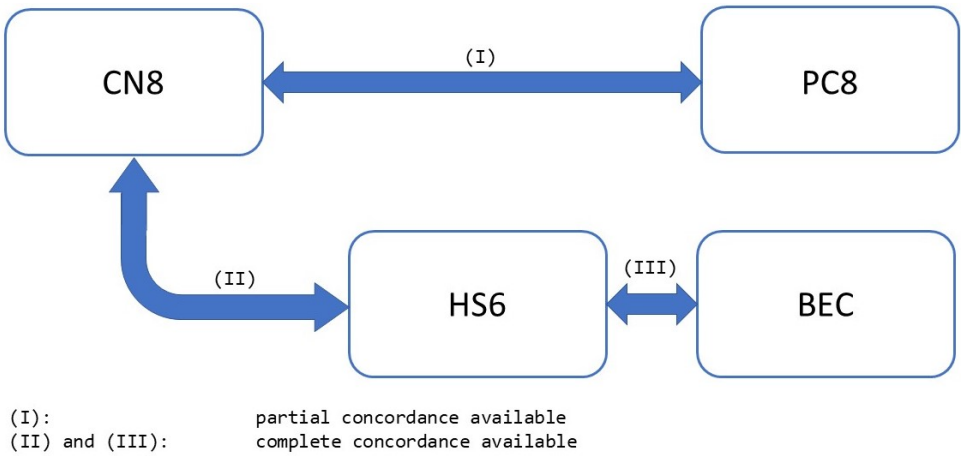


Figure 1: Connections between the different classification systems

possible from 1995 to 2022 and PC8 from 2008 to 2021, respectively. However, new concordance data is added regularly in the package by the maintainer and can even be conveniently added by the users themselves. Also, usage of product codes and concordance files provided by National Statistics Institutes is possible. We identified this feature as important because in our application, the statistics office keeps track of some national product codes, and thus harmonization is possible for these products too.

The basic idea behind harmonization is to keep track of every product code during a specific period. In the simplest case, a code does not change during the examined period, i.e., no harmonization is needed. In any other case, all changes which are associated with a specific code need to be considered. There are different product code changes: 1 to 1, 1 to many, many to 1, many to many, 1 to none, none to 1. The last two changes indicate that a code was dropped or a new code was created, respectively. Those two changes are only possible for PC8 and not for CN8 classification. A 1 to many or many to 1 change can occur if two or more codes are split or merged. It is also possible that a "mixture" of changes is present, e.g., a code can merge with another product code but remain the same afterwards in terms of notation. Harmonization aims to establish consistent time series of product codes. Therefore, the product codes and their changes are observed over time, and the final product code is recorded in the period's first/last year in a separate column [7].

Figure 2 visualizes possible product code histories. Diamonds in the figure indicate that a change of codes occurs in a given year. In case *(I)* code $A$ does not change and thus the CN8 code is recorded as CN8plus code. Case *(II)* shows a renaming of the product code, and thus product $B$ becomes simply $C$. The remaining lines show more complex changes. Cases *(III)* and *(IV)* demonstrate a merge of code $D$ and $E$ in 2011. They both will be associated with family *f1* in CN8plus. Family *f2*

[7]The columns are named CN8plus or PC8plus respectively in the package. For more technical documentation on how the harmonization works, it is recommended to read the *vignette* of the R Package `https://cran.r-project.org/web/packages/harmonizer/vignettes/harmonizer.html`

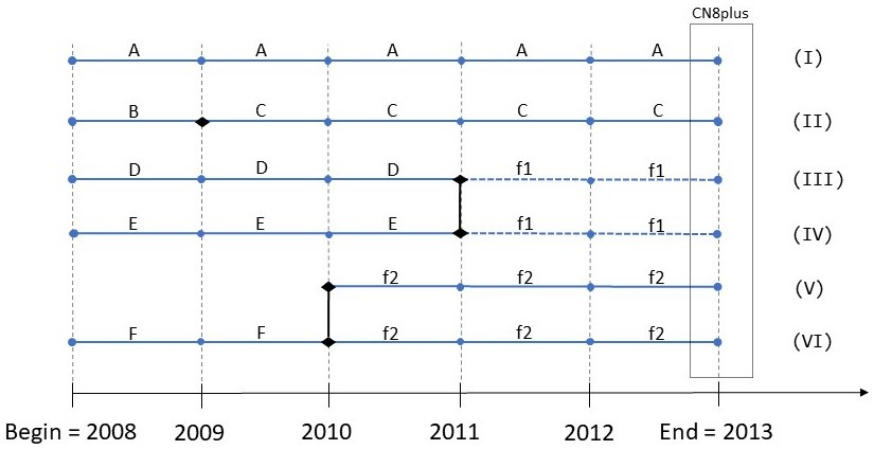


Figure 2: Example for CN8 code changes

in cases of *(V)* and *(VI)* result from a split of code *F* in 2010. Both codes have the CN8plus classification *f2*. One should note that a mixture of these cases is possible, resulting in another family code. An application of this example, using real CN8 codes, can be found in the Appendix (cf. Figure A.2).

All changes above may be valid cases for CN8, HS6, and PC8 product codes' changes over time. However, PC8 codes can also be generated or dropped within the period of interest. If changes in these classifications are not taken correctly into account, this leads to erroneous entries and exits of products, price bias, or incorrect price and quantity indices. A visual illustration for PC8 codes can also be found in the Appendix (cf. Figure A.1).

# 3 Application

We apply harmonization on firm-product-market data of Croatia in the period from 2008 to 2016. For demonstration purpose Table 4 shows one possible (artificial) entry of raw data at firm level.

Table 4: Artifical example of firm-product-market level data, classified by the CN8

| FirmID | Year | CN8 | Market | Unit Type | Unit amount | Export Value |
|---|---|---|---|---|---|---|
| 10356 | 2020 | 84181080 | AT | Number of items | 17 | 15,000 |

Table 5 provides the export value of all goods captured by CN8 codes. Columns (2) and (3) show the result without harmonization, while columns (4) to (7) take the harmonization into account. Almost 20% of the total export value averaged across all years is associated with family product codes. If we do not harmonize the product codes, amount and value of products will be incorrectly attributed to individual product codes, while actually, the codes should be interpreted in the context of a family of product codes (cf. Figure 2). A direct consequence of this misinterpretation can be seen in Table 6.

Table 5: Exports of firms in Croatia

| Years | Without Harmonization | | With Harmonization | | | |
|---|---|---|---|---|---|---|
| | Export value | # of codes | Export value in families | Export value in families [%] | # of families | # of final CN8plus codes |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| 2008 | 9448.5 | 6434 | 1584.5 | 16.77 | 458 | 5547 |
| 2009 | 7565.1 | 6214 | 1333.0 | 17.62 | 448 | 5378 |
| 2010 | 8786.7 | 6106 | 1695.8 | 19.30 | 443 | 5347 |
| 2011 | 9459.5 | 5994 | 1904.4 | 20.13 | 427 | 5338 |
| 2012 | 9564.3 | 6235 | 2067.9 | 21.62 | 436 | 5557 |
| 2013 | 9389.9 | 6388 | 2062.6 | 21.97 | 457 | 5630 |
| 2014 | 10112.2 | 6681 | 2095.7 | 20.72 | 471 | 5789 |
| 2015 | 11266.7 | 6845 | 2205.6 | 19.58 | 482 | 5889 |
| 2016 | 11896.9 | 6941 | 2421.1 | 20.35 | 485 | 5924 |
| Average | 9721.1 | 6426.4 | 1930.1 | 19.8 | 456.3 | 5599.9 |

Note: Columns (2) and (3) represent the export value and the number (#) of product codes if no harmonization is applied. Columns (4) to (8) summarize harmonized data. Columns (4) and (5) show the absolute or the relative value respectively by all product codes that are associated with a family of product codes, while column (7) shows the number of different families of product codes. Column (8) represents the final number of different harmonized codes in each year. All values are in million €.

Table 6 shows all added and dropped products in terms of export, i.e., firms started or stopped trading specific products. From 2008 to 2016, added or dropped products are counted from one year to the following. Thus, the firm had to exist in both years. Columns (2) to (5) represent the trading data without harmonization, while columns (7) to (11) summarize harmonized data. The percentages below the values denote the relative amount of added or dropped products for all products in the market.

Comparing the value of added products, columns (4) and (9), reveals the strong influence of the missing harmonization. One notices that the average amount of sold new-to-firm products without harmonization is overestimated by 50%. Without harmonizing the data, firms are interpreted as introducing more products than they do or being more innovative than they are. A similar pattern occurs if one looks at the value of dropped products, i.e., columns (6) and (11). By looking at non harmonized data, one would overestimate the amount of dropped products of a firm even though some were not dropped at all. Similar findings are obtained for imports (cf. Tables A.1 and A.2).

Table 6: Added and Dropped Export Products

| | Without Harmonization | | | | | With Harmonization | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Year | # of firms | Added Products | Value of Added Products | Dropped Products | Value of Dropped Products | # of firms | Added Products | Value of Added Products | Dropped Products | Value of Dropped Products |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| 2008-2009 | 5922 | 29353 | 418.7 | 36346 | 628.9 | 5791 | 27875 | 321.0 | 34619 | 518.1 |
| | | 43.90% | 4.50% | 54.3% | 6.70% | | 42.6% | 3.4% | 52.90% | 5.6% |
| 2009-2010 | 6478 | 31901 | 1214.9 | 31762 | 673.6 | 6321 | 30214 | 822.4 | 29874 | 374.5 |
| | | 52.70% | 16.30% | 52.5% | 9.10% | | 51.2% | 11.10% | 50.70% | 5.0% |
| 2010-2011 | 6748 | 31947 | 874.4 | 32664 | 815.8 | 6586 | 30326 | 657.5 | 30919 | 557.4 |
| | | 51.90% | 10.20% | 53.1% | 9.50% | | 50.5% | 7.70% | 51.50% | 6.5% |
| 2011-2012 | 7288 | 44599 | 1835.0 | 30187 | 2039.8 | 7105 | 41938 | 415.5 | 27841 | 808.1 |
| | | 74.40% | 20.00% | 50.4% | 22.30% | | 71.7% | 4.50% | 47.60 | 8.80% |
| 2012-2013 | 6988 | 37582 | 684.4 | 37893 | 615.6 | 6822 | 36107 | 504.1 | 36755 | 512.5 |
| | | 52.10% | 7.30% | 52.6% | 6.50% | | 51.3% | 5.30% | 52.20% | 5.40% |
| 2013-2014 | 5891 | 38975 | 646.7 | 34787 | 619.1 | 5738 | 36785 | 569.1 | 33491 | 573.2 |
| | | 57.80% | 7.00% | 51.6% | 6.70% | | 56.1% | 6.20% | 51.10% | 6.20% |
| 2014-2015 | 6296 | 46749 | 687.1 | 32621 | 397.0 | 6109 | 44116 | 644.2 | 31291 | 366.6 |
| | | 62.20% | 6.90% | 43.4% | 4.00% | | 61.2% | 6.50% | 43.40% | 3.7% |
| 2015-2016 | 6638 | 41776 | 666.0 | 40426 | 764.1 | 6432 | 39736 | 623.3 | 38058 | 693.5 |
| | | 46.00% | 6.00% | 44.5% | 6.90% | | 45.8% | 5.70% | 43.80% | 6.30% |
| Average | 6531.1 | 37860.3 | 878.4 | 34585.8 | 819.2 | 6363 | 35887.1 | 569.6 | 32856 | 550.5 |
| | | 55.1% | 9.8% | 50.3% | 8.96% | | 53.8% | 6.3% | 49.2% | 5.9% |

Note: Columns (2) to (6) show analyses using no harmonization. Every firm that exists in both years of the period of column (1) is included. Columns (3) and (4) show the number (#) and the corresponding value of added products. The percentage below refers to the number of all existing products and values respectively. Columns (5) and (6) follow the same idea but using dropped products. Columns (7) to (11) can be interpreted analogously, but correct for the issue of harmonization. All values are in million €.

We also investigated the firm sales data from products classified by Prodcom codes in Croatia. Table 7 shows the difference between harmonized and non-harmonized data. The raw data without any harmonization, which can be found in columns (2) and (3) of Table 7, may also include possible typos and national codes. Within the Prodcom classification system, every country can include codes that are only valid for the nation itself. These codes are not used internationally and can therefore only be harmonized with national concordance files. However, if concordance files are available, it is possible to use them within our package since the use of custom trade data, as well as custom concordance files, is possible. Nevertheless, the harmonized results shown here do not include these national codes, since we have not yet had access to the national concordance files of Croatia.

Table 7: Products classified by Prodcom in Croatia

| | Without Harmonization | | With Harmonization | | | |
|---|---|---|---|---|---|---|
| Years | Sales value | # of codes | Value in families | Value in families [%] | # of families | # of final PC8plus codes |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| 2008 | 18349.5 | 1746 | 699.5 | 3.81 | 60 | 1607 |
| 2009 | 16005.5 | 1704 | 560.5 | 3.50 | 61 | 1583 |
| 2010 | 16467.2 | 1651 | 558.1 | 3.39 | 59 | 1536 |
| 2011 | 17238.8 | 1647 | 622.4 | 3.61 | 60 | 1532 |
| 2012 | 17286.5 | 1615 | 579.1 | 3.35 | 57 | 1513 |
| 2013 | 16110.2 | 1603 | 528.0 | 3.28 | 58 | 1510 |
| 2014 | 15837.7 | 1637 | 523.4 | 3.31 | 61 | 1540 |
| 2015 | 16178.9 | 1623 | 533.2 | 3.30 | 62 | 1528 |
| 2016 | 15940.7 | 1581 | 539.5 | 3.38 | 61 | 1490 |
| Average | 16601.7 | 1645.2 | 571.5 | 3.4 | 59.9 | 1537.7 |

Note: Columns (2) and (3) represent the export value and the number (#) of product codes if no harmonization is applied. Columns (4) to (8) present a harmonization for internationally used PC8 codes (i.e., national codes excluded). Columns (4) and (5) show the absolute respectively the relative value of all codes that are associated with a family, while column (7) shows the number of families. Column (8) gives the final number of harmonized codes in a certain year. All values are in million €.

Similar findings to Table 6 providing dropped and added products for Prodcom codes can be found in the Appendix in Table A.3. Valid identification of the changes in the firms' product mix allows developing innovation measures and therewith enhancing the second wave of innovation and firm literature. Identified new products can be classified into new-to-firm and new-to-market products, and can also be combined with the value of these products sold or exported — thus providing additional insights into the magnitude of innovation. Researchers can work with the harmonized data to not only identify new-to-firm products but combining this data with firm-level geographical data and identify new-to-firm-region or new-to-firm-national market products. In addition, with the export or sales value, the importance of these new products can be quantified.

# 4 Conclusion

Similarly to the second wave of research on firms and international trade (Wagner, 2019), the early second wave of research on firms and innovation uses transaction data at firm-product or even firm-product-market level (Castellani & Fassio, 2019). Such data can be used for many purposes, for example, to develop innovation measures at various regional levels or to identify characteristics of innovative or high growth firms. In policy evaluation, for example, it can be used for evaluation of R&D grants or export boosting policies to measure direct and indirect policy outcomes and mechanisms (Srhoj et al., 2020; Van Biesebroeck et al., 2016). Even when such data is used for policy evaluation, there are several studies that do not harmonize their data (for examples: Srhoj et al., 2020), which can hamper the validity of the findings.

The application of our approach on Croatian firm-product-market data demonstrated the importance of product code harmonization. In our application, the average export value of new products in the firms' product mix is overestimated by 50% using not harmonized data. The extent of the significance of harmonizing is in line with the application of Van Beveren et al. (2012) on the data of Belgium. The R package *harmonizer* comprises all efforts of harmonization and provides the possibility for computing the number of added, kept, or dropped products for each firm. In addition, updating harmonization is conveniently possible as new concordance files can be provided as input files. This allows even to harmonize sales data of firms that report volume and quantity of products classified by national product codes.

With our approach of flexible harmonization of trade and sales data for specific periods, new and dropped products can be counted reliably, changes in export and import values can be computed, price indices derived, and new products distinguished between new-to-firm products or new-to-market products. The *harmonizer* package can be conveniently applied. We hope our contribution is going to push forward the research in innovation economics by enabling research with the appropriate data.

# References


Aghion, P., Antonin, C., & Bunel, S. (2021). *The power of creative destruction: Economic upheaval and the wealth of nations*. Harvard University Press. Retrieved 2022-05-21, from https://doi.org/10.4159/9780674258686 doi: doi:10.4159/9780674258686

Audretsch, D. B., & Belitski, M. (2020). The role of R&D and knowledge spillovers in innovation and productivity. *European Economic Review*, *123*, 103391. Retrieved from https://www.sciencedirect.com/science/article/pii/S0014292120300234 doi: https://doi.org/10.1016/j.euroecorev.2020.103391

Bernard, A. B., Blanchard, E. J., Van Beveren, I., & Vandenbussche, H. (2018, 02). Carry-Along Trade. *The Review of Economic Studies*, *86*(2), 526-563. Retrieved from https://doi.org/10.1093/restud/rdy006 doi: 10.1093/restud/rdy006

Castellani, D., & Fassio, C. (2019). From new imported inputs to new exported products. Firm-level evidence from Sweden. *Research Policy*. doi: 10.1016/j.respol.2018.08.021

Coad, A., Segarra, A., & Teruel, M. (2016). Innovation and firm growth: Does firm age play a role? *Research Policy*, *45*(2), 387-400. Retrieved from https://www.sciencedirect.com/science/article/pii/S0048733315001687 doi: https://doi.org/10.1016/j.respol.2015.10.015

Directorate-General for Taxation and Customs Union. (2022, May). *The Combined Nomenclature.* Retrieved from https://ec.europa.eu/taxation_customs/business/calculation-customs-duties/customs-tariff/combined-nomenclature_en ([Online; accessed 20. May 2022])

Duprez, C., & Magerman, G. (2022, March). *Correspondences of eu product classifications* (Working Paper). Retrieved from https://static1.squarespace.com/static/55e85d72e4b0146280523def/t/62388fb92cba1d26fa8408f6/1647873979840/concordances_live.pdf

European Commission. (2019, November). Prodcom list. *Official Journal of the European Union*. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019R1933&qid=1653055165284

European Commission. (2021, November). Combined Nomenclature. *Official Journal of the European Union*, *64*. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2021:414:FULL&from=EN

Gault, F. (2018). Defining and measuring innovation in all sectors of the economy. *Research Policy*, *47*(3), 617-622. Retrieved from https://www.sciencedirect.com/science/article/pii/S0048733318300076 doi: https://doi.org/10.1016/j.respol.2018.01.007

Griffith, R., Huergo, E., Mairesse, J., & Peters, B. (2006, 12). Innovation and Productivity Across Four European Countries. *Oxford Review of Economic Policy*, *22*(4), 483-498. Retrieved from https://doi.org/10.1093/oxrep/grj028 doi: 10.1093/oxrep/grj028

Hauser, C., Siller, M., Schatzer, T., Walde, J., & Tappeiner, G. (2018). Measuring regional innovation: A critical inspection of the ability of single indicators to shape technological change. *Technological Forecasting and Social Change*, *129*, 43-55. Retrieved from https://www.sciencedirect.com/science/article/pii/S0040162517300604 doi: https://doi.org/10.1016/j.techfore.2017.10.019

Link, S. (2020). Harmonization of the ifo business survey's micro data. *Jahrbücher für Nationalökonomie und Statistik*, *240*(4), 543–555. Retrieved 2022-05-20, from https://doi.org/10.1515/jbnst-2019-0042 doi: doi:10.1515/jbnst-2019-0042

Mairesse, J., & Mohnen, P. (2002). Accounting for innovation and measuring innovativeness: An illustrative framework and an application. *American Economic Review*, *92*(2), 226-230. Retrieved from https://www.aeaweb.org/articles?id=10.1257/000282802320189302 doi: 10.1257/000282802320189302

Mairesse, J., & Mohnen, P. (2010). Chapter 26 - using innovation surveys for econometric analysis. In B. H. Hall & N. Rosenberg (Eds.), *Handbook of the economics of innovation, volume 2* (Vol. 2, p. 1129-1155). North-Holland. Retrieved from https://www.sciencedirect.com/science/article/pii/S0169721810020101 doi: https://doi.org/10.1016/S0169-7218(10)02010-1

Mayer, T., & Zignago, S. (2011). *Notes on cepii's distances measures: The geodist database* (Tech. Rep.). Retrieved from http://dx.doi.org/10.2139/ssrn.1994531 doi: 10.2139/ssrn.1994531

Schubert, T., & Tavassoli, S. (2020). Product innovation and educational diversity in top and middle management teams. *Academy of Management Journal*, *63*(1), 272-294. Retrieved from https://doi.org/10.5465/amj.2017.0741 doi: 10.5465/amj.2017.0741

Srhoj, S., Vitezić, V., & Wagner, J. (2020). *Export boosting policies and firm behaviour: Review of empirical evidence around the world* (Working Paper Series in Economics No. 395). Lüneburg. Retrieved from http://hdl.handle.net/10419/234584

Stojčić, N., Srhoj, S., & Coad, A. (2020). Innovation procurement as capability-building: Evaluating innovation policies in eight central and eastern european countries. *European Economic Review*, *121*, 103330. Retrieved from https://www.sciencedirect.com/science/article/pii/S0014292119301904 doi: https://doi.org/10.1016/j.euroecorev.2019.103330

United Nations. (2016). *Classification by Broad Economic Categories Rev.5.* Retrieved from https://unstats.un.org/unsd/trade/classifications/Manual%20of%20the%20Fifth%20Revision%20of%20the%20BEC%20(Unedited).pdf ([Online; accessed 20. May 2022])

Van Beveren, I., Bernard, A. B., & Vandenbussche, H. (2012, December). *Concording eu trade and production data over time* (Working Paper No. 18604). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w18604 doi: 10.3386/w18604

Van Biesebroeck, J., Konings, J., & Volpe Martincus, C. (2016, 09). Did export promotion help firms weather the crisis? *Economic Policy*, *31*(88), 653-702. Retrieved from https://doi.org/10.1093/epolic/eiw014 doi: 10.1093/epolic/eiw014

Wagner, J. (2016). A survey of empirical studies using transaction level data on exports and imports. *Rev. World Econ.*, *152*(1), 215–225. doi: 10.1007/s10290-015-0235-8

Wagner, J. (2019). *International trade in goods: Evidence from transaction data.* World Scientific Publishing Company Pte. Limited.

WCO. (2022). *Harmonized System - Frequently asked questions.* (http://www.wcoomd.org/en/topics/nomenclature/overview/harmonized_system_faq.aspx [Online; accessed 20. May 2022])

# A Appendix

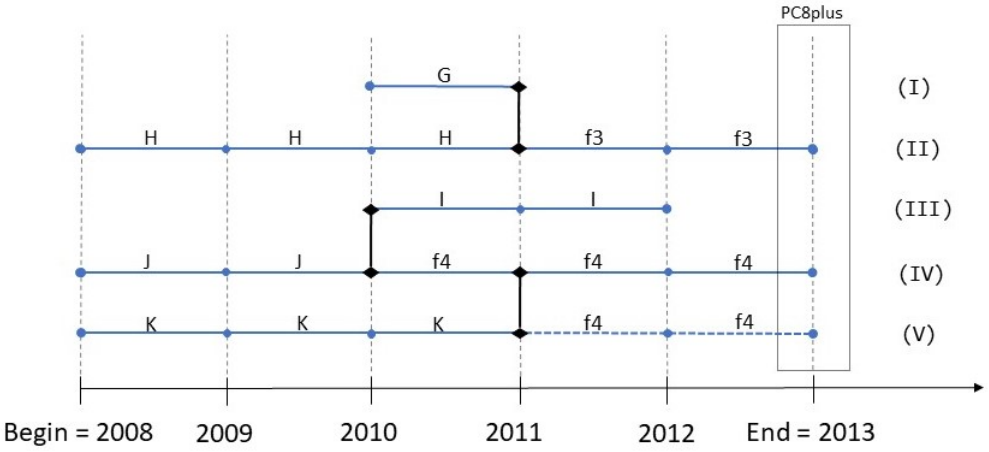


Figure A1: Example for Prodcom code changes

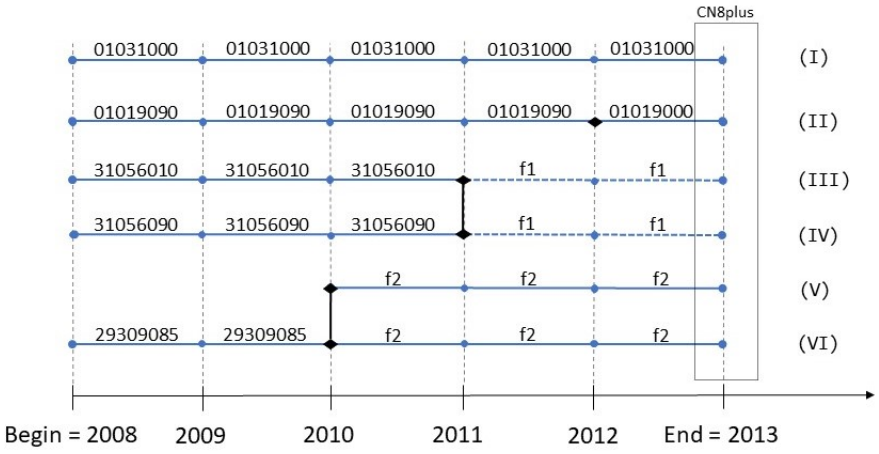


Figure A2: Example for real CN8 code changes

Table A1: Imports of firms in Croatia

| Years | Without Harmonization | | With Harmonization | | | |
|---|---|---|---|---|---|---|
| | Import value | # of codes | Import value in families | Import value in families [%] | # of families | # of final CN8+ codes |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| 2008 | 20527.8 | 8222 | 3147.3 | 15.33 | 543 | 6914 |
| 2009 | 15295.4 | 8100 | 2320.3 | 15.17 | 540 | 6869 |
| 2010 | 14934.9 | 7970 | 2406.0 | 16.11 | 527 | 6827 |
| 2011 | 16073.0 | 7829 | 2830.3 | 17.61 | 529 | 6835 |
| 2012 | 16073.6 | 7977 | 2599.6 | 16.17 | 536 | 6931 |
| 2013 | 16276.8 | 8090 | 2587.2 | 15.90 | 543 | 6981 |
| 2014 | 16906.4 | 8075 | 2943.1 | 17.41 | 531 | 6938 |
| 2015 | 18260.1 | 8080 | 3067.6 | 16.80 | 529 | 6916 |
| 2016 | 18995.4 | 8175 | 3012.4 | 15.86 | 532 | 6941 |
| Average | 17038.1 | 8057.56 | 2768.2 | 16.26 | 534.44 | 6905.78 |

Note: Columns (2) and (3) represent the import value and the number (#) of product codes if no harmonization is applied. Columns (4) to (8) summarize harmonized data. Columns (4) and (5) show the absolute or the relative value respectively by all product codes that are associated with a family of product codes, while column (7) shows the number of different families of product codes. Column (8) represents the final number of different harmonized codes in each year. All values are in million €.

Table A2: Added and Dropped Import Products

| Year | Without Harmonization | | | | | With Harmonization | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | # of firms | Added Products | Value of Added Products | Dropped Products | Value of Dropped Products | # of firms | Added Products | Value of Added Products | Dropped Products | Value of Dropped Products |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| 2008-2009 | 15881 | 167909 | 1360.2 | 207749 | 2234.4 | 15662 | 159089 | 1188.3 | 197906 | 1999.3 |
| | | 38.00% | 6.70% | 47.10% | 11.00% | | 36.60% | 5.90% | 45.50% | 9.9% |
| 2009-2010 | 16403 | 171341 | 2099.1 | 188505 | 2547.2 | 16171 | 160777 | 1366.6 | 176265 | 1756.4 |
| | | 41.90% | 13.90% | 46.10% | 16.90% | | 40.00% | 9.10% | 43.80% | 11.6% |
| 2010-2011 | 16645 | 178434 | 1405.0 | 175674 | 1556 | 16421 | 167633 | 1083.3 | 163835 | 1203.3 |
| | | 45.20% | 9.50% | 44.50% | 10.60% | | 43.20% | 7.40% | 42.20% | 8.2% |
| 2011-2012 | 17827 | 260135 | 2617.5 | 169092 | 2470.7 | 17572 | 245834 | 1377.9 | 156074 | 1089.3 |
| | | 65.30% | 16.70% | 42.40% | 15.70% | | 62.70% | 8.80% | 39.80% | 6.9% |
| 2012-2013 | 17082 | 178683 | 1429.3 | 230035 | 1411 | 16842 | 172798 | 1230.8 | 224343 | 1217.4 |
| | | 36.10% | 9.00% | 46.50% | 8.90% | | 35.50% | 7.80% | 46.10% | 7.7% |
| 2013-2014 | 10409 | 139553 | 1448.9 | 181060 | 1273.9 | 10274 | 133469 | 1222.4 | 175524.0 | 1119.2 |
| | | 36.10% | 9.10% | 46.80% | 8.00% | | 35.20% | 7.70% | 46.30% | 7.1% |
| 2014-2015 | 10647 | 160713 | 1242.5 | 134545.0 | 1083.9 | 10514 | 151742.0 | 1080.6 | 129748.0 | 973.7 |
| | | 45.80% | 7.40% | 38.30% | 6.50% | | 44.30% | 6.50% | 37.90% | 5.8% |
| 2015-2016 | 11100 | 140076 | 1412.2 | 145586.0 | 1171.7 | 10944 | 134651.0 | 1302.1 | 138448.0 | 1074.0 |
| | | 37.40% | 7.90% | 38.90% | 6.50% | | 37.10% | 7.20% | 38.20% | 6.0% |
| Average | 14499.3 | 174605.5 | 1626.8 | 179030.8 | 1718.6 | 14300 | 165749.1 | 1231.5 | 170267.9 | 1304.1 |
| | | 43.2% | 10.0% | 43.83% | 10.51% | | 41.8% | 7.6% | 42.48% | 7.9% |

Note: Columns (2) to (6) show analyses using no harmonization. Every firm that exists in both years of the period of column (1) is included. Columns (3) and (4) show the number (#) and the corresponding value of added products. The percentage below refer to the number of all existing products and value respectively. Column (5) and (6) follow the same idea but using dropped products. Columns (7) to (11) can be interpreted analogously, but correct for the issue of harmonization. All values are in million €.

Table A3: Added and Dropped Prodcom Products

| | Without Concordance | | | | | With Concordance | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Year | # of firms | Added Products | Value of Added Products | Dropped Products | Value of Dropped Products | # of firms | Added Products | Value of Added Products | Dropped Products | Value of Dropped Products |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| 2008-2009 | 3147 | 878 | 966.2 | 945 | 795.1 | 2927 | 783 | 842.1 | 822 | 612.2 |
| | | 9.70% | 5.40% | 10.40% | 4.50% | | 9.1% | 4.7% | 9.60% | 3.40% |
| 2009-2010 | 3010 | 549 | 630.7 | 648 | 446.7 | 2794 | 415 | 505.2 | 488 | 330.5 |
| | | 6.40% | 4.10% | 7.60% | 2.90% | | 5.1% | 3.3% | 6.00% | 2.10% |
| 2010-2011 | 3003 | 493 | 855.6 | 559 | 697.6 | 2787 | 403 | 550.6 | 444 | 456.7 |
| | | 5.90% | 5.30% | 6.60% | 4.30% | | 5.0% | 3.4% | 5.60% | 2.80% |
| 2011-2012 | 2973 | 669 | 1206.8 | 662 | 698.3 | 2763 | 417 | 960.9 | 414 | 420.1 |
| | | 8.20% | 7.20% | 8.10% | 4.20% | | 5.4% | 5.8% | 5.40% | 2.50% |
| 2012-2013 | 2797 | 531 | 1006.0 | 580 | 574.2 | 2592 | 475 | 928.2 | 485 | 464.0 |
| | | 6.90% | 5.90% | 7.50% | 3.40% | | 6.5% | 5.5% | 6.60% | 2.70% |
| 2013-2014 | 2873 | 569 | 892.7 | 555 | 1091.7 | 2659 | 529 | 792.5 | 496 | 991.9 |
| | | 7.30% | 5.60% | 7.20% | 6.80% | | 7.2% | 5.0% | 6.70% | 6.20% |
| 2014-2015 | 3007 | 546 | 1110.7 | 569 | 1000.5 | 2759 | 457 | 500.6 | 493 | 737.2 |
| | | 6.90% | 7.10% | 7.20% | 6.40% | | 6.1% | 3.2% | 6.60% | 4.70% |
| 2015-2016 | 2995 | 852 | 1335.3 | 916 | 1061.8 | 2737 | 450 | 311.9 | 492 | 356.3 |
| | | 10.80% | 8.30% | 11.60% | 6.60% | | 6.0% | 1.9% | 6.60 | 2.20% |
| Average | 2975.6 | 635.9 | 1000.5 | 679.3 | 795.7 | 2752.3 | 491.1 | 674.0 | 516.8 | 546.1 |
| | | 7.76% | 6.1% | 8.3% | 4.9% | | 6.3% | 4.1% | 6.6% | 3.3% |

Note: Columns (2) to (6) show analyses using no harmonization. Every firm that exists in both years of the period of column (1) is included. Columns (3) and (4) show the number (#) and the corresponding value of added products. The percentage below refers to the number of all existing products and value, respectively. Columns (5) and (6) follow the same idea, but use dropped products. Columns (7) to (11) can be interpreted analogously, but correct for the issue of harmonization. All values are in million €.

**University of Innsbruck - Working Papers in Economics and Statistics**
**Recent Papers** can be accessed on the following webpage:

https://www.uibk.ac.at/eeecon/wopec/

2022-11 **Christoph Baumgartner, Stjepan Srhoj and Janette Walde:** Harmonization of product classifications: A consistent time series of economic trade activities

2022-10 **Katharina Momsen, Markus Ohndorf:** Seller Opportunism in Credence Good Markets – The Role of Market Conditions

2022-09 **Christoph Huber, Michael Kirchler:** Experiments in Finance – A Survey of Historical Trends

2022-08 **Tri Vi Dang, Xiaoxi Liu, Florian Morath:** Taxation, Information Acquisition, and Trade in Decentralized Markets: Theory and Test

2022-07 **Christoph Huber, Christian König-Kersting:** Experimenting with Financial Professionals

2022-06 **Martin Gächter, Martin Geiger, Elias Hasler:** On the structural determinants of growth-at-risk

2022-05 **Katharina Momsen, Sebastian O. Schneider:** Motivated Reasoning, Information Avoidance, and Default Bias

2022-04 **Silvia Angerer, Daniela Glätzle-Rützler, Philipp Lergetporer, Thomas Rittmannsberger:** How does the vaccine approval procedure affect COVID-19 vaccination intentions?

2022-03 **Robert Böhm, Cornelia Betsch, Yana Litovsky, Philipp Sprengholz, Noel Brewer, Gretchen Chapman, Julie Leask, George Loewenstein, Martha Scherzer, Cass R. Sunstein, Michael Kirchler:** Crowdsourcing interventions to promote uptake of COVID-19 booster vaccines

2022-02 **Matthias Stefan, Martin Holmén, Felix Holzmeister, Michael Kirchler, Erik Wengström:** You can't always get what you want–An experiment on finance professionals' decisions for others

2022-01 **Toman Barsbai, Andreas Steinmayr, Christoph Winter:** Immigrating into a Recession: Evidence from Family Migrants to the U.S.

2021-32 **Fanny Dellinger:** Housing Support Policies and Refugees' Labor Market Integration in Austria

2021-31 **Albert J. Menkveld, Anna Dreber, Felix Holzmeister, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Sebastian Neusüss, Michael Razen, Utz Weitzel and et al:** Non-Standard Errors

2021-30 **Toman Barsbai, Victoria Licuanan, Andreas Steinmayr, Erwin Tiongson, Dean Yang:** Information and Immigrant Settlement

2021-29 **Natalie Struwe, Esther Blanco, James M. Walker:** Competition Among Public Good Providers for Donor Rewards

2021-28 **Stjepan Srhoj, Melko Dragojević:** Public procurement and supplier job creation: Insights from auctions

2021-27 **Rudolf Kerschbamer, Regine Oexl:** The effect of random shocks on reciprocal behavior in dynamic principal-agent settings

2021-26 **Glenn E. Dutcher, Regine Oexl, Dmitry Ryvkin, Tim Salmon:** Competitive versus cooperative incentives in team production with heterogeneous agents

2021-25 **Anita Gantner, Regine Oexl:** Respecting Entitlements in Legislative Bargaining - A Matter of Preference or Necessity?

2021-24 **Silvia Angerer, E. Glenn Dutcher, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** The formation of risk preferences throughsmall-scale events

2021-23 **Stjepan Srhoj, Dejan Kovač, Jacob N. Shapiro, Randall K. Filer:** The Impact of Delay: Evidence from Formal Out-of-Court Restructuring

2021-22 **Octavio Fernández-Amador, Joseph F. Francois, Doris A. Oberdabernig, Patrick Tomberger:** Energy footprints and the international trade network: A new dataset. Is the European Union doing it better?

2021-21 **Felix Holzmeister, Jürgen Huber, Michael Kirchler, Rene Schwaiger:** Nudging Debtors to Pay Their Debt: Two Randomized Controlled Trials

2021-20 **Daniel Müller, Elisabeth Gsottbauer:** Why Do People Demand Rent Control?

2021-19 **Alexandra Baier, Loukas Balafoutas, Tarek Jaber-Lopez:** Ostracism and Theft in Heterogeneous Groups

2021-18 **Zvonimir Bašić, Parampreet C. Bindra, Daniela Glätzle-Rützler, Angelo Romano, Matthias Sutter, Claudia Zoller:** The roots of cooperation

2021-17 **Silvia Angerer, Jana Bolvashenkova, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Children's patience and school-track choices several years later: Linking experimental and field data

2021-16 **Daniel Gründler, Eric Mayer, Johann Scharler:** Monetary Policy Announcements, Information Schocks, and Exchange Rate Dynamics

2021-15 **Sebastian Bachler, Felix Holzmeister, Michael Razen, Matthias Stefan:** The Impact of Presentation Format and Choice Architecture on Portfolio Allocations: Experimental Evidence

2021-14 **Jeppe Christoffersen, Felix Holzmeister, Thomas Plenborg:** What is Risk to Managers?

2021-13 **Silvia Angerer, Daniela Glätzle-Rützler, Christian Waibel:** Trust in health care credence goods: Experimental evidence on framing andsubject pool effects

2021-12 **Rene Schwaiger, Laura Hueber:** Do MTurkers Exhibit Myopic Loss Aversion?

2021-11 **Felix Holzmeister, Christoph Huber, Stefan Palan:** A Critical Perspective on the Conceptualization of Risk in Behavioral and Experimental Finance

2021-10 **Michael Razen, Alexander Kupfer:** Can increased tax transparency curb corporate tax avoidance?

2021-09 **Changxia Ke, Florian Morath, Anthony Newell, Lionel Page:** Too big to prevail: The paradox of power in coalition formation

2021-08 **Marco Haan, Pim Heijnen, Martin Obradovits:** Competition with List Prices

2021-07 **Martin Dufwenberg, Olof Johansson-Stenman, Michael Kirchler, Florian Lindner, Rene Schwaiger:** Mean Markets or Kind Commerce?

2021-06 **Christoph Huber, Jürgen Huber, and Michael Kirchler:** Volatility Shocks and Investment Behavior

2021-05 **Max Breitenlechner, Georgios Georgiadis, Ben Schumann:** What goes around comes around: How large are spillbacks from US monetary policy?

2021-04 **Utz Weitzel, Michael Kirchler:** The Banker's Oath And Financial Advice

2021-03 **Martin Holmen, Felix Holzmeister, Michael Kirchler, Matthias Stefan, Erik Wengström:** Economic Preferences and Personality Traits Among Finance Professionals and the General Population

2021-02 **Christian König-Kersting:** On the Robustness of Social Norm Elicitation

2021-01 **Laura Hueber, Rene Schwaiger:** Debiasing Through Experience Sampling: The Case of Myopic Loss Aversion.

University of Innsbruck

Working Papers in Economics and Statistics

2022-11

Christoph Baumgartner, Stjepan Srhoj and Janette Walde

# Harmonization of product classifications: A consistent time series of economic trade activities

**Abstract**
Firm-product data provide information for various research questions in international trade or innovation economics. However, working with this data requires harmonizing product classifications consistently over time to avoid internal validity issues. The researcher must consider product code changes in the classi- fication systems over the observation period. Harmonization is required because classification systems like the EU classifications Combined Nomenclature (CN) for goods or the Prodcom for the production of manufactured goods undergo several changes. We have addressed this problem and developed an approach to harmonize product codes. This approach tracks product codes from 1995 to 2022 for CN and 2001 to 2021 for Prodcom. Additional years can be conveniently updated. We provide the harmonized product codes for CN and Prodcom in the selected period's first (or last) year. Our approach is summarized in an easy-to-use R package so that researchers can consistently track product codes for their period. We demonstrate the importance of harmonization using the micro-level trade data for Croatia as a case study. Our approach facilitates working with firm-product data, allowing the analysis of important research questions.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)