Open Econ Rev (2017) 28:795–803
DOI 10.1007/s11079-017-9442-3

COMMENT

EU Enlargement and the New Goods Margin in Austrian Trade: Comment

Ignacio del Rosal[1]

Published online: 14 July 2017

Abstract This paper comments on the article 'EU Enlargement and the New Goods Margin in Austrian Trade' by John T. Dalton (Open Econ Rev 28(1):61–78, 2017). Dalton measures the new goods margin of trade between Austria and the new entrants to the EU in 2004, following the methodology proposed by Timothy Kehoe and Kim Ruhl (J Polit Econ 121(2):358–392, 2013). The present paper argues that the author overlooked the implicit requirement of using a consistent product classification, leading to overestimation of the new goods margin. After replicating Dalton's results and recalculating the trade share of the new goods margin using a consistent product classification, this paper shows the magnitude of the overestimation (up to 20 or 30 decimal points of the trade share).

Keywords Extensive margin · International trade · Trade liberalization · Austria · EU · Product classification

JEL Classifications F10 · F13 · F14 · F15

In a recent paper in this *Review*, John T. Dalton (2017) measures the extensive margin of trade between Austria and the ten new entrants to the European Union (EU) in 2004. He follows the definition and methodology proposed by Kehoe and Ruhl (2013) (hereinafter, KR), and the extensive margin thus refers to the set of least treaded goods, i.e. goods with zero trade as well as goods with trade in small, but positive values. The

This comment refers to the article available at doi:10.1007/s11079-016-9399-7.

An author's reply to this comment is available at doi:10.1007/s11079-017-9443-2.

✉ Ignacio del Rosal
irosal@uniovi.es

[1] Department of Applied Economics, University of Oviedo, Avenida del Cristo, s.n, 33006 Oviedo, Spain

least traded goods set is defined at the beginning of the sample period and then its trade share is observed over time, paying special attention to the final level it reaches. On average, the least traded goods margin accounts for 56% of the bilateral trade flows between Austria and the ten new Members States at the end of the sample period.

Dalton did a good job of applying the KR methodology. However, as I explain in Section 1, a key requirement implicit in the KR methodology is that the product classification has to be consistent throughout the study period, and this is not the case in Dalton (2017). If the changes in the product classification are not controlled for, part of the trade share associated with least traded goods will be spurious, a pure artefact of the shifts between codes of the product classification. In some cases, the least traded products margin can be notably inflated (up to 20 or 30 decimal points of the trade share).

This paper comprises three sections. In Section 1, I replicate the results obtained by Dalton (2017), clarifying what he actually did and focusing on the problem of changes in the product classification. In Section 2, I recalculate the new products margin employing a consistent product classification, thus showing the magnitude of the overestimation in Dalton (2017). Finally, Section 3 concludes the paper.

1 Replication of Dalton's (2017) Results

Very briefly, the basics of the KR methodology are the following. Bilateral trade flow values are studied by good. Trade flows will be reported according to some product classification, so a good will be a code of the product classification at its maximum level of disaggregation. The goods are ordered by their traded value at the beginning of the study period. To avoid dependence on the choice of the base year, the traded value is averaged over 3 years. Subsequently, the ordered products are cumulated into 10 bins, each representing 10% of the total value of exports from country m to country n in the base year. These 10 sets of goods do not change throughout the sample period. The key to the KR methodology is to observe the evolution and final level of the trade share of the first bin, i.e. the set of least traded goods. This is the new goods margin, comprising both goods with zero trade value and goods with a small, but positive trade value. The least traded goods set comprises the vast majority of goods (codes) of the product classification. The ordering of goods means that it successively takes fewer goods to fill a bin with 10% of the traded value. These subsequent bins contain goods with a significant volume of trade in the base year and can be understood as the intensive margin of trade, although KR do not pay them much attention.

Dalton (2017) applies the KR methodology to the trade between Austria and the ten new entrants to the EU in 2004. He uses Eurostat's Comext data (see Eurostat 2016). Comext provides trade data at the 8-digit Combined Nomenclature (CN) level, a European product classification based on the international 6-digit Harmonized System (HS). The 8-digit CN classification cannot be used due to its continuous and numerous changes, so the 6-digit HS, used by Dalton, is arguably the best choice. The HS product classification comprises approximately 5000 categories at the 6-digit level. The sample period is 1999–2009, therefore goods are ordered according to their average trade over 1999–2001, but the goods that comprise each of the 10 bins are defined according to 1999 data. Dalton first analyses exports flows from Austria to the ten EU entrants and then repeats the analysis with imports flows to Austria. He pays special attention to four of its border countries, namely the Czech Republic, Hungary, Slovakia and Slovenia.

Two aspects of Dalton's results caught my attention when I read the paper. First, looking at the bar plots used to illustrate the changes in the distribution of trade value among the 10 bins (as in KR; see Fig. 1 below in this paper), I was relatively surprised with the major growth in the trade share of the new goods margin in all cases. At the same time, the bins representing the intensive margin of trade (and especially the bin comprising the top products) showed a reduction in their trade shares, approaching zero in some cases. That is, the role of the new goods margin appeared to be inflated, while the role of the intensive margin seemed to be deflated. In fact, these results reported by Dalton (2017) are different to those obtained by the same author in a previous version of the paper (Dalton 2013; not cited in the final version). However, so long as Dalton (2013) is based on trade data reported at a less disaggregated level (4-digit Standard International Trade Classification revision 2, comprising 786 codes), the discrepancy would be justified. The second aspect that caught my attention was the number of goods in the bar plots. Summing the number of goods in the 10 bins gives the 'theoretical' number of goods, 7385. Is this a plausible number of HS codes? No, it is not, and this issue made me wonder whether Dalton might have mixed different versions of the HS.

The HS product classification was developed and is maintained by the World Customs Union (WCU), and is updated every 5–6 years. In Dalton's sample period (1999–2009), three different HS versions were successively in use: the 1996 HS (comprising 5113 6-digit codes), 2002 HS (5224) and 2007 HS (5052). So the number of products reported by Dalton does not match any single HS version. Would it be the case if the three different HS versions were mixed into one? No, it would not either. The amendments introduced with each revision suppose a change in a certain number of codes, which receive a new code. Approximately 10% of the HS codes change in every HS revision (see Yu 2008). Using the correlation tables between HS editions provided by the WCU, or the update of CN codes tables provided by Eurostat through its metadata server 'RAMON',[1] the number of different HS codes comprising the 'mixed HS' can be determined, giving a figure of 5709 codes. This figure is still far from the number of goods apparently used by Dalton, 7385.

The presence of alphanumeric codes would fill the gap. Comext data reports some trade flows with alphanumeric codes. Eurostat (2016) explains that alphanumeric codes are mainly used to identify confidential or adjusted data, or when the classification of the trade flow at the maximum level of disaggregation is not possible due to some allowed simplified declaration.[2] For instance, code 85SSS7 (precisely the most important alphanumeric code in the present study case) refers to some electrical machinery that cannot be associated with a complete 6-digit HS for reasons of confidentiality. As long as these codes cannot be perfectly tracked by the researcher, the usual practice is to discard them. However, if this is not the case, the alphanumeric codes have to be counted to determine the theoretical number of goods in the exercise. Analysing Austrian import and export flows with the ten EU new entrants as a whole, there are 113 different alphanumeric codes in the period 1999–2009.

I replicate Dalton's (2017) results using the mixed HS classification without discarding alphanumeric codes. The theoretical number of goods is then 5709 + 113 = 5822. Tables 1 and 2 shows the trade share evolution of least traded products

[1] See http://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC.

[2] See also Eurostat (2015) for an assessment on the quality of Eurostat trade data, including the impact of confidentiality.

Table 1 Share of least traded goods (mixed HS)

	Austrian exports to:									
year	Cyprus	Czech Republic	Estonia	Hungary	Latvia	Lithuania	Malta	Poland	Slovakia	Slovenia
1999	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
2000	0.08	0.09	0.09	0.09	0.13	0.11	0.06	0.09	0.10	0.08
2001	0.09	0.09	0.09	0.09	0.11	0.11	0.05	0.08	0.09	0.08
2002	0.65	0.20	0.14	0.22	0.35	0.27	0.25	0.21	0.28	0.20
2003	0.55	0.25	0.17	0.28	0.36	0.26	0.27	0.24	0.29	0.26
2004	0.42	0.29	0.37	0.35	0.45	0.40	0.48	0.27	0.37	0.33
2005	0.56	0.31	0.70	0.34	0.45	0.62	0.58	0.32	0.36	0.26
2006	0.48	0.27	0.44	0.36	0.41	0.55	0.89	0.33	0.37	0.27
2007	0.52	0.33	0.57	0.41	0.42	0.56	0.64	0.38	0.37	0.29
2008	0.55	0.34	0.60	0.44	0.46	0.61	0.51	0.42	0.43	0.31
2009	0.61	0.36	0.57	0.46	0.53	0.53	0.61	0.43	0.47	0.35

for Austrian exports and imports with the ten new entrants to the EU. These tables almost identically reproduce Tables 1 and 2 in Dalton's paper. Therefore, Dalton used a mixed HS product classification (mixing three different versions of the HS, 1996, 2002 and 2007), and did not purge the alphanumeric codes. The theoretical number of products deduced from Dalton's paper (7385) does not have much sense in my opinion.

The present paper would not be worth the trouble, however, if only the theoretical number of goods were questionable. The key issue is that mixing different versions of the HS product classification leads to overestimating the new goods margin. When

Table 2 Share of least traded goods (mixed HS)

	Austrian imports from:									
year	Cyprus	Czech Republic	Estonia	Hungary	Latvia	Lithuania	Malta	Poland	Slovakia	Slovenia
1999	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
2000	0.06	0.09	0.11	0.08	0.10	0.10	0.07	0.10	0.08	0.09
2001	0.13	0.10	0.11	0.08	0.07	0.18	0.03	0.10	0.09	0.08
2002	0.32	0.16	0.19	0.20	0.18	0.40	0.13	0.19	0.27	0.18
2003	0.37	0.19	0.57	0.26	0.30	0.55	0.20	0.23	0.36	0.22
2004	0.94	0.26	0.55	0.38	0.43	0.70	0.52	0.28	0.52	0.36
2005	0.88	0.24	0.65	0.40	0.58	0.58	0.90	0.34	0.53	0.24
2006	0.91	0.25	0.56	0.47	0.65	0.63	0.94	0.36	0.65	0.24
2007	0.87	0.32	0.63	0.49	0.85	0.83	0.88	0.42	0.67	0.24
2008	0.96	0.33	0.74	0.50	0.79	0.82	0.93	0.43	0.68	0.33
2009	0.85	0.34	0.79	0.47	0.80	0.82	0.61	0.56	0.66	0.37

different HS versions are merged into one, the new codes introduced in later versions will have nil associated trade in the earlier years. This means that trade flows classified using the new codes will spuriously form part of the least traded goods. This distortion is especially serious when the shifted codes include some of the most important trade flows. Let us see a specific, very clear-cut example to illustrate this point. In Austrian exports to Slovakia, 1996 HS code 271000 (which comprises several petroleum products) represents the largest traded value in the period 1999–2001. In the 2002 revision of the HS, code 271000 is split into four new codes

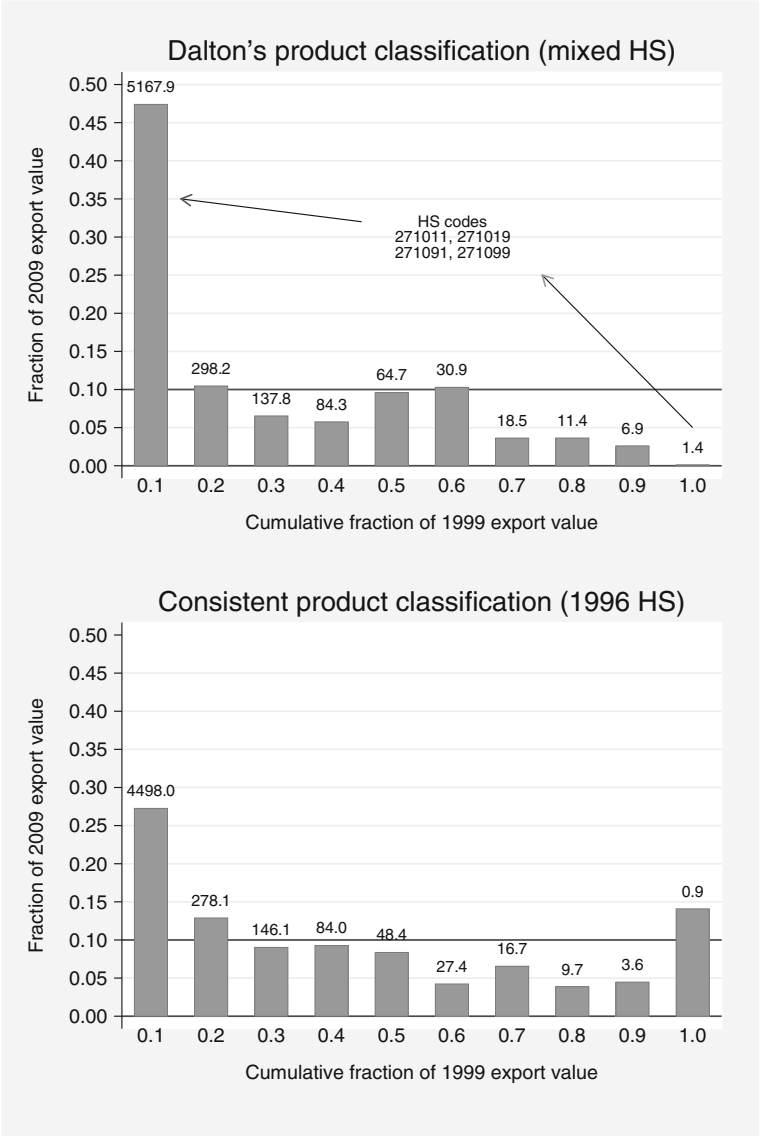

Fig. 1 Composition of exports: Austria to Slovakia (comparison of product classifications)

(271011, 271019, 271091 and 271099). When yearly Comext data (with each yearly dataset classified according to the corresponding HS version) is merged into a single dataset in order to apply the KR methodology, the four new codes will have nil value in the period 1999–2001 and will form part of the first 10% bin, whereas the former code will form part of the intensive margin of trade (in fact, it will be in the top 10% bin), and will have nil value in 2002 and afterwards. The new products margin is then spuriously inflated and the intensive margin of trade is deflated. This issue explains Fig. 3 in Dalton's paper, concerning Austrian exports to Slovakia (see my reproduction in Fig. 1).

It is also worth noting that the application of the KR methodology to the specific sample period 1999–2009 highlights the problem. Using the averaged traded valued over the period 1999–2001 is truly unfortunate, because a HS revision took place in 2002. Another averaging period would mitigate the problem, although the dysfunction of using an inconsistent product classification would still remain, of course. Furthermore, the spurious increase in the trade share of the least traded products in 2002 leads Dalton to make some comments about the timing of the impact of EU enlargement that are somewhat questionable (see Figs. 5 and 10 and the last paragraph of Section 3.1 in Dalton 2017).

2 The least Traded Goods Margin Using a Consistent Product Classification

Changes in product classifications do not have perfect solutions. In the present case, I think that the best option would be to use the conversion tables between different HS versions provided by the United Nations Statistics Division (UNSD).[3] The UNSD uses the correlation tables provided by the WCU to draw up usable conversion tables that allow the conversion of trade data reported in later versions of the HS to earlier versions. I thus convert the Comext 2002–2006 trade data reported in the 2002 HS version and the 2007–2009 data reported in the 2007 HS version into data in the homogeneous 1996 HS version. Of course, there is no need to make any conversion for the 1999–2001 data. In short, the 1996 HS version thus constitutes the product classification for the entire sample period 1999–2009. The 1996 HS version comprises 5113 6-digit codes.

[3] See http://unstats.un.org/unsd/trade/conversions/HS%20Correlation%20and%20Conversion%20tables.htm. There is another suitable option, which is the concordance procedure developed by Pierce and Schott (2012) and adapted for Eurostat's CN by Van Beveren et al. (2012). Details can be found in both papers, but the main idea is to create a consistent product classification over time by generating synthetic codes that group together the interrelated changing codes. The major advantage of the procedure is that it allows the tracking of changes in product classifications over time and the translating of trade data into a common and consistent classification. However, it does not come without a cost. If the sample period is relatively large and covers several versions of the product classification (as is the present case, which involves three editions of the HS), the procedure generates synthetic codes that will encompass many changing codes, so it may tend to artificially concentrate trade data. For this reason, my first choice is to use the UNSD conversion tables. I follow the concordance procedure in Van Beveren et al. (2012) to perform some robustness analysis, replicating the computations done in the benchmark, and the results are quite similar. These results are available upon request.

I repeat the KR methodology using Dalton's sample, i.e. Austrian exports and imports to/from the ten EU new entrants. For the reasons mentioned in the previous section, I purge the alphanumeric codes appearing in the Comext data. The average number of alphanumeric codes in Austrian exports to the group of ten countries is approximately 43, representing, on average, 4.8% of the total export value (with a clearly decreasing pattern, from 9.7 to 2.6%). For Austrian imports, the average figures are 10 alphanumeric codes and 0.08% of the total import value.

For the sake of illustration, Fig. 1 replicates Fig. 3 in Dalton (2017) and repeats the same bar plot using a consistent product classification, namely 1996 HS. Figure 1 refers to Austrian exports to Slovakia, the example used in Section 1 to illustrate the problem of the shifting HS codes. As can be seen, using the mixed HS means that the top 10% bin is emptied, whereas the first 10% bin is inflated with the new 2002 HS codes explained in Section 1 and shown in Fig. 1 (but not exclusively, as there are other shifted codes that generate a similar transfer of trade value from the intensive margin bins to the new goods bin).

The general results are shown in Tables 3 and 4. These tables replicate previous Tables 1 and 2, applying the KR methodology to Austrian trade with the group of EU new entrants using the 1996 HS product classification. For ease of comparison, the companion Fig. 2 displays the resulting differences in the trade share of the least traded products in 2009. As can be seen, the share of least traded products decreases significantly when a consistent product classification is used, although the new goods margin continues to play an important role. The overestimation in Dalton's paper is more pronounced in Austrian export flows than in import flows, and is also more pronounced for the most relevant countries among the EU new entrants (the four border countries, the Czech Republic, Hungary, Slovenia and Slovakia, plus Poland). Overestimation exceeds 20 decimal points in Austrian exports to Hungary (the overestimation in Austrian exports to Slovakia being quite similar), and exceeds 30 decimal points in Austrian imports from Slovakia. Overestimation is not so high in the trade flows with

Table 3 Share of least traded goods (1996 HS)

	Austrian exports to:									
year	Cyprus	Czech Republic	Estonia	Hungary	Latvia	Lithuania	Malta	Poland	Slovakia	Slovenia
1999	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
2000	0.08	0.09	0.10	0.08	0.12	0.12	0.09	0.09	0.09	0.08
2001	0.09	0.09	0.11	0.08	0.11	0.14	0.07	0.08	0.10	0.08
2002	0.08	0.10	0.15	0.11	0.28	0.27	0.16	0.11	0.10	0.10
2003	0.24	0.11	0.31	0.13	0.31	0.29	0.22	0.15	0.12	0.13
2004	0.57	0.16	0.36	0.20	0.39	0.42	0.41	0.18	0.18	0.28
2005	0.46	0.16	0.67	0.18	0.43	0.58	0.55	0.21	0.19	0.17
2006	0.35	0.16	0.43	0.22	0.40	0.51	0.09	0.24	0.24	0.17
2007	0.33	0.18	0.50	0.23	0.42	0.54	0.61	0.26	0.24	0.17
2008	0.49	0.17	0.54	0.22	0.43	0.60	0.48	0.25	0.24	0.19
2009	0.55	0.18	0.50	0.24	0.48	0.55	0.49	0.29	0.27	0.21

Table 4 Share of least traded goods (1996 HS)

year	Austrian imports from: Cyprus	Czech Republic	Estonia	Hungary	Latvia	Lithuania	Malta	Poland	Slovakia	Slovenia
1999	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
2000	0.06	0.09	0.11	0.08	0.10	0.10	0.07	0.10	0.08	0.09
2001	0.13	0.10	0.11	0.08	0.07	0.18	0.03	0.10	0.09	0.08
2002	0.16	0.12	0.18	0.13	0.18	0.40	0.12	0.16	0.12	0.13
2003	0.11	0.13	0.57	0.21	0.29	0.54	0.20	0.19	0.22	0.17
2004	0.91	0.22	0.41	0.31	0.42	0.65	0.52	0.24	0.26	0.33
2005	0.73	0.19	0.65	0.24	0.58	0.55	0.90	0.25	0.25	0.22
2006	0.67	0.22	0.55	0.31	0.65	0.59	0.94	0.32	0.35	0.23
2007	0.86	0.24	0.58	0.32	0.82	0.80	0.88	0.34	0.36	0.22
2008	0.89	0.24	0.70	0.32	0.76	0.81	0.93	0.36	0.33	0.29
2009	0.82	0.27	0.79	0.34	0.79	0.80	0.61	0.48	0.35	0.27

Slovenia, the Czech Republic or Poland, but is still quite substantial. The problem of the shifted HS codes has a minor impact in Cyprus, Estonia, Latvia, Lithuania and Malta (in this last case, the impact is not so minor in exports), and the computation using the 1996 HS classification gives a slightly higher trade share for the least traded goods in Austrian exports to Lithuania. This subgroup of five countries plays a very

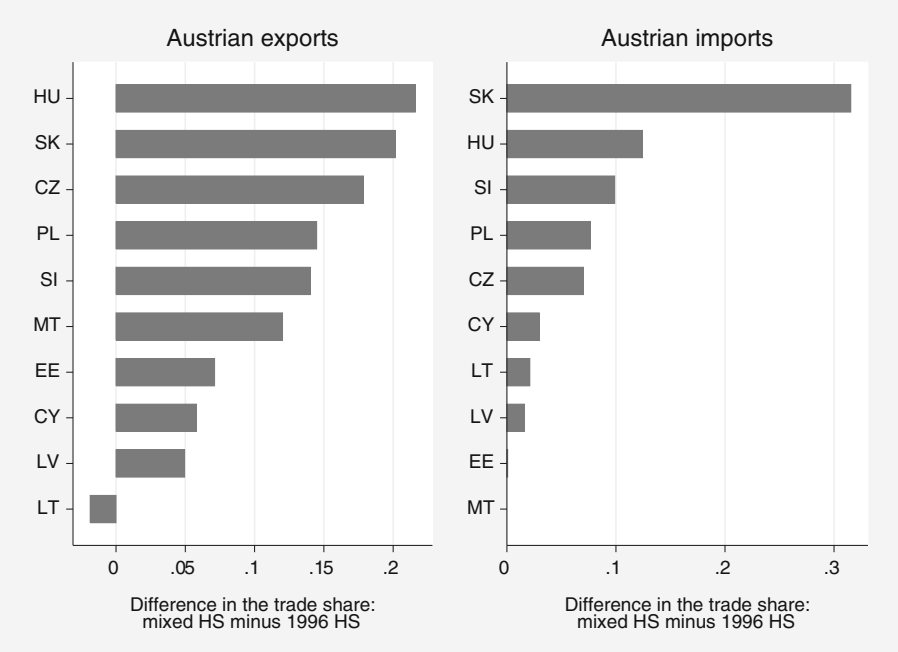

Fig. 2 Overestimation of the trade share of least traded products: mixed HS minus 1996 HS

marginal role in Austrian trade, so it is reasonable for the share of the least traded goods to reach such high levels.[4]

3 Conclusion

This short paper comments on the results reported by Dalton (2017) concerning the role of the least traded goods in Austrian trade with the group of the ten new entrants to the EU in 2004. Although Dalton correctly applied the Kehoe and Ruhl (2013) methodology with Comext trade data, he overlooked the implicit requirement of using a consistent product classification. This problem leads to significant overestimation of the trade share of the least traded goods. After converting the Comext data into a consistent product classification, I show the magnitude of this overestimation.

In my opinion, product classification issues can be cumbersome and are not sufficiently addressed in many empirical papers using trade data at the product level. In this respect, Dalton's paper is an illustrative example of the problems that an inconsistent product classification can generate.

Acknowledgements Financial support from the Consejería de Educación, Cultura y Deporte del Principado de Asturias is acknowledged.

References

Dalton JT (2013) EU enlargement and the new goods margin in Austrian trade, MPRA paper no. 50353
Dalton JT (2017) EU enlargement and the new goods margin in Austrian trade. Open Econ Rev 28(1):61–78
Kehoe TJ, Ruhl KJ (2013) How important is the new goods margin in international trade? J Polit Econ 121(2): 358–392
Eurostat (2015) Quality report on European statistics on international trade in goods. Publications Office of the European Union, Luxembourg
Eurostat (2016) User guide on European statistics on international trade in goods. Publications Office of the European Union, Luxembourg
Pierce JR, Schott PK (2012) Concording US Harmonized System categories over time. J Off Stat 28(1):53–68
Van Beveren I, Bernard AB, Vandenbussche H (2012) "Concording EU trade and production data over time", NBER working paper no. 18604
Yu D (2008) "the Harmonized System - amendments and their impact on WTO members' schedules", World Trade Organization, economic research and statistics division staff working paper ERSD-2008-02

[4] Averaging over the sample period, the group of the ten new entrants to the EU represents 6.2% of Austrian exports and 5.4% of Austrian imports. These shares are also quite stable. The four border countries plus Poland suppose, on average, 96.8% of Austrian exports to the ten EU new entrants and 98.9% of its imports.