Tjalling C. Koopmans Research Institute

Discussion Paper Series nr: 05-24

Empirical Relevance of the Hillman Condition for Revealed Comparative Advantage: 10 Stylized Facts

Jeroen Hinloopen
Charles van Marrewijk

Universiteit Utrecht

Utrecht School of Economics

Tjalling C. Koopmans Research Institute
Utrecht School of Economics
Utrecht University

Vredenburg 138
3511 BG Utrecht
The Netherlands
telephone +31 30 253 9800
fax +31 30 253 7373
website www.koopmansinstitute.uu.nl

The Tjalling C. Koopmans Institute is the research institute and research school of Utrecht School of Economics.
It was founded in 2003, and named after Professor Tjalling C. Koopmans, Dutch-born Nobel Prize laureate in economics of 1975.

In the discussion papers series the Koopmans Institute publishes results of ongoing research for early dissemination of research results, and to enhance discussion with colleagues.

Please send any comments and suggestions on the Koopmans institute, or this series to M.Damhuis@econ.uu.nl

ontwerp voorblad: WRIK Utrecht

How to reach the authors
Please direct all correspondence to the first author.

Jeroen Hinloopen
Department of Economics
University of Amsterdam
Roeterstraat 11
1018 WB Amsterdam
The Netherlands
E-mail: J.Hinloopen@uva.nl

Charles van Marrewijk
Department of Economics
Erasmus University Rotterdam
Burgemeester Oudlaan 50
3062 PA Rotterdam
The Netherlands
E-mail: vanmarrewijk@few.eur.nl

Utrecht School of Economics
Tjalling C. Koopmans Research Institute
Discussion Paper Series 05-24

Empirical Relevance of the Hillman Condition for Revealed Comparative Advantage: 10 Stylized Facts

Jeroen Hinloopen[a]
Charles van Marrewijk[b]

[a]Department of Economics
University of Amsterdam

[b]Department of Economics
Erasmus University Rotterdam

February 2005

Abstract
The theoretically necessary and sufficient condition for the correspondence between 'revealed'comparative advantage and pre-trade relative prices derived by Hillman (1980) is analyzed empricially for virtually all countries of the world over an extended period of time. This yields 10 stylized facts, including that (i) violations of the Hillman condition are small as a share of the number of observations, but quite substantial as a share of the value of world exports, (ii) violations occur relatively frequent in the period 1970 – 1984 while they hardly ever occur in the period 1985 – 1997, and (iii) violations occur foremost in primary product and natural-resource intensive sectors, for sectors in countries in Africa, the Middle East, Latin America, and Eastern Europe. The condition appears also to be useful for identifying erroneous trade flow classifications.

Keywords: Balassa index, Hillman condition, comparative advantage

JEL classification: C81, D43, F11, F20

Acknowledgements
Thanks are due to Joe Francois for useful comments. Any errors are, of course, our own.

1. Introduction

The concept of 'revealed' comparative advantage, introduced by Liesner (1958) but refined and popularized by Balassa (1965) with his concomitant index, is widely used empirically to identify structural trade-related patterns accross countries. Porter (1990) uses it to identify strong sectoral clusters, Amiti's (1999) analysis of specialization patterns in Europe is based on it, Bojnec (2001) employs it for his study of Central and Eastern European aggricultural trade, Fertö and Hubbard (2003) analyze with it the competitiveness of Hungarian agri-food sectors, Hinloopen and Van Marrewijk (2004) use it for their analysis of the dynamics of Chinese comparative advantage, and Svaleryd and Vlachos (2005) examine to what extent a related index value is instrumental for explaining a country's level of financial development. The dynamics of (the distribution of) the Balassa index as such are considered in Proudman and Redding (1998, 2000) and Hinloopen and Van Marrewijk (2001).

The theoretical basis for the Balassa index is provided by Hillman (1980), who diagrammatically developed a necessary and sufficient condition for the correspondence between the Balassa index and pre-trade relative prices in cross-country sector comparisons, the so-called Hillman condition. As Hillman notes (1980, p. 320): "Whether this condition obtains is a matter for empirical investigation".

Because the Hillman condition can be easily verified empirically it is rather surprising that it is ignored by the large majority of empirical studies on revealed comparative advantage that have appeared since. The only empirical investigation to date of the Hillman condition as such is the study of Marchese and Nadal De Simone (1989), who analyze the exports of 118 developing countries at the 1-, 3-, 4-, and 5-digit level of sector aggregation. They conclude that in the year 1985 (the only year considered by Marchese and Nadal De Simone, 1989) the Hillman condition does not hold for about 9.5 percent of the value of exports of their group of developing countries. The sole empirical study into comparative advantage that mentions explicitly to include only those sectors that meet the Hillman condition is Hinloopen and Van Marrewijk (2001). They find that the Hillman condition does not hold for about 0.5 percent of the number of observations in their sample, which corresponds to about 7.0 percent of the value of exports.

In this paper the empirical relevance of the Hillman condition is thoroughly investigated using a comprehensive dataset consisting of annual recordings on bilateral trade flows for 1,056 4-digit sectors, 183 countries, and 28 years, yielding a total of slightly less than 18.4 million positive

observations (see also Feenstra, 2000). This dataset allows for an investigation of the Hillman index for virtually all countries of the world, over an extended period of time, and for four different levels of sector aggregation (Appendix A contains a detailed description of the dataset). The empirical relevance of the Hillman condition can thus be established.

As the dataset represents a large part of recent international trade the empirical findings are presented as stylized facts. Among these are the observation that violations of the Hillman condition are small as a share of the number of observations, but often represent a disproportionally large value of trade. Including these observations in studies into (the dynamics of) revealed comparative advantage could thus yield quite inaccurate inferences. Further, two periods can be distinguished as to the severity with which the Hillman condition is violated. From 1970 through 1984 violations happen relatively frequent and they represent a significant fraction of the value of total trade, while from 1985 onwards violations hardly ever occur and represent an insignificant fraction of total trade value. Hence 1985, the year analyzed by Nadal and De Simone (1989), is not representative for the extent to which violations of the Hillman condition is empirically relevant. Also, violations do not occur randomly across sectors or countries. They occur foremost in sectors producing primary products or that are natural-resource intensive, and for sectors that are located in countries in Africa, the Middle East, Eastern Europe, and Latin America.

An important by-product from restricting the analysis to those observations meeting the Hillman condition is that observations based on erronously classified trade flows are identified by the Hillman condition and consequently can be dismissed. If only for this screening property of applying the Hillman condition it is recommended that it is checked always in empirical studies into revealed comparative advantage.

The next section briefly discusses the Balassa index of revealed comparative advantage and the concomitant Hillman condition. Section 3 contains some preliminaries of the Hillman condition in relation to data aggregation and trade flow classifications. All cases violating the Hillman condition are subsequently presented in Section 4. Section 5 concludes.

2 The Balassa index and the Hillman condition

Since it is hard to gauge the importance of a sector without a frame of reference, Balassa (1965) introduced normalized export shares as an indicator of revealed comparative advantage:

(1) $$BI_{i,t}^{j} = \frac{X_{i,t}^{j} / X_{i,t}}{X_{t}^{j} / X_{t}}, \quad i \in I, j \in J ,$$

where $X_{i,t}^{j}$ are country *i*'s exports in sector *j* during period *t*, *I* is a group of reference countries, *J* is the bundle of products exported by country *i*, with $X_{t}^{j} = \sum_{i} X_{i,t}^{j}$, $X_{t}^{j} = \sum_{i} X_{i,t}^{j}$, and $X_{t} = \sum_{i} \sum_{j} X_{i,t}^{j}$. If $BI_{i,t}^{j} > 1$ country *i* is said to have a revealed comparative advantage in the production of commodity *j* in time period *t* as its export share for product *j* is larger than the concomittant export share in the group of reference countries *I*. This group may vary, as indeed it does in the studies refered to in the Introduction, and is most often determined by the largest set of reference countries for which reliable data are available.

Hillman (1980) examines the correspondence between the Balassa index and pre-trade relative prices in cross-country comparisons for a specific sector under homothetic preferences by forming a Hicksian composite commodity for all other sectors. As the concomitant transformation of the Balassa index has to be monotonic, Hillman's condition can be interpreted as a monotonicity condition for scaling a country's exports by a measure of its (sector) size. In particular, the Hillman condition is that:

(2) $$1 - \frac{X_{i,t}^{j}}{X_{t}^{j}} > \frac{X_{i,t}^{j}}{X_{i,t}} \left(1 - \frac{X_{i,t}}{X_{t}} \right).$$

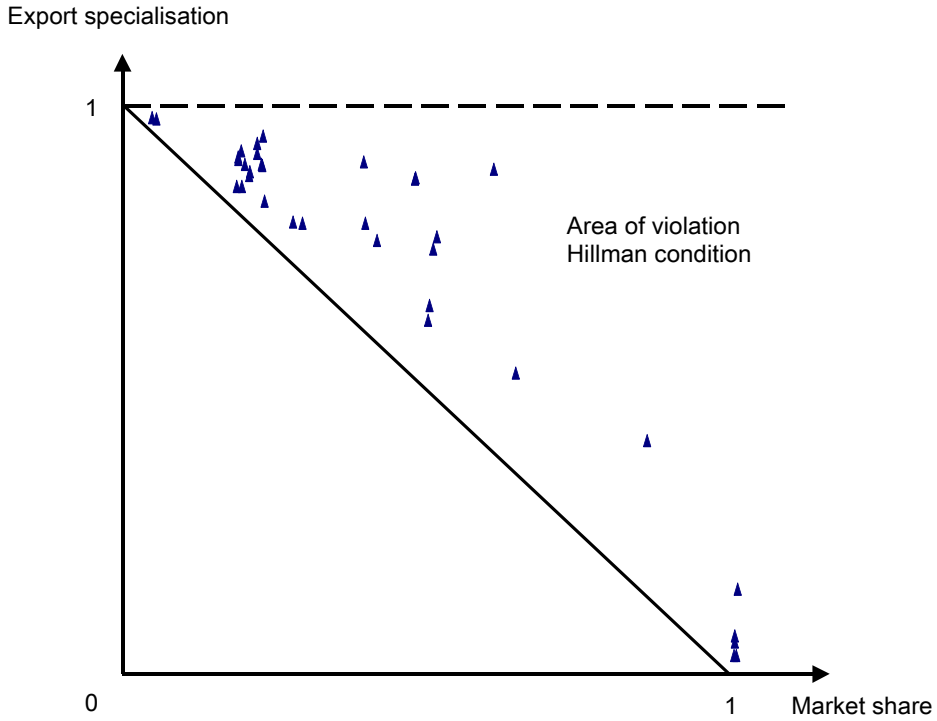
Condition (2) must be met for the value of the Balassa index (1) to be in concordance with pre-trade relative prices, that is, to ensure the Balassa index to increase if $X_{i,t}^{j}$ increases.

The Hillman condition (2) consists of three parts that all have a distinct economic interpretation:

- *market share*, as measured by $X_{i,t}^{j} / X_{t}^{j}$, that is, the share of a country's exports in a particular sector relative to the total exports in that sector of the group of reference countries;
- *degree of export specialisation*, as measured by $X_{i,t}^{j} / X_{i,t}$, that is, the share of a country's exports in a particular sector relative to that country's total exports;
- *country size*, as measured by $X_{i,t} / X_{t}$, that is, the share of a country's total exports relative to total exports of the group of reference countries.

As Hillman (1980) notes, violations of (2) readily obtain in case a country exports one commodity only (in which case $X_{i,t}^{j} = X_{i,t}$) or when a country is the sole supplier (in which case $X_{i,t}^{j} = X_{t}^{j}$). In general the Hillman condition is violated if a country has a significantly high market share in the supply of the particular commodity in combination with a significantly high degree of export specialization. The condition is somewhat less stringent for large countries. Figure 1 divides the market share – export specialisation space for an infintely small country size according to the Hillman condition being violated or not. For larger country sizes the dividing line would shift counter-clockwise outwards. In case of violations an increase in a country's exports in a particular sector increases this sector's export share in world trade more than that it increases the sector's national export share. As a result the Balassa index drops in value, which contradicts the notion of revealed comparative advantage.

Figure 1 Area of violations in market share – export specialisation space*



* The demarcation line is for infinitely small countries; the observations correspond to the 4-digit observations in Table 4, see Section 4.

3. Prelimenary considerations

Before discussing in detail the characteristics of all observations violating the Hillman condition, the impact of sector aggregation and trade flow classifications on the Hillman be violated or not need to be addressed.

3.1 Data aggregation

In Section 4 violations of the Hillman condition are reported for different degrees of sector aggregation as data aggregation affects the likelihood of violation. At lower levels of aggregation, where more sectors are identified, it becomes "easier" in principle for a country to realize a large market share in a specific sector. This tends to increase the likelihood that the Hillman condition is violated. On the other hand, the degree of export specialisation falls at lower levels of aggregation, which tends to decrease the likelihood of violation. The net result of these two forces is an empirical matter.

Table 1 Aggregation and share of exports not satisfying Hillman's condition, 1970 - 1997

	1-digit	2-digit	3-digit	4-digit
Number of observations	73	79	88	35
Share of number of observatons (%)	0.2148	0.0477	0.0176	0.0051
Value of exports (billion US $)	1319	1263	1291	127
Share of value of exports (%)	2.8802	3.4266	2.6672	0.5287
Average market share	0.0758	0.1395	0.1965	0.3988
Average export specialisation	0.9809	0.9278	0.8881	0.7319
Average country size	0.0128	0.0132	0.0132	0.0073
Share of total trade covered (%)	100.00	99.67	99.46	60.39

Table 1 summarizes the violations of the Hillman condition for different levels of data aggregation. As a share of the number of observations violations are rather insignificant; as a share of the value of total exports violations are significant. These findings suggest that in practise violations of the Hillman condition do not happen often, but when they do occur that they involve (very) large trade flows. These observations lead to the first empirical regularity:

Stylized fact 1

Violations of the Hillman condition occur in 0 – 0.25 percent of all cases; these violations represent 0 – 3.5 percent of total trade.

Table 1 also shows that sector aggregation matters. The average market share of violation cases increases with more refined sector definitions while at the same time both average export specialisation and, to a lesser extend, country size decrease. Although increasing market shares and decreasing country sizes enhance the probability of violation, going through the 1-, 2-, 3-, and 4-digit sector aggregations shows that these effects are on average more than corrected for by the concomitant reduction in export specialisations. That is:

Stylized fact 2

The higher is the degree of sector aggregation, the higher is the probability that the Hillman condition is violated.

At the same time the value of total trade involved in the violaton cases is about the same for the 1-, 2-, and 3-digit levels of sector aggregation. The reduction in trade value represented by all observations violating the Hillman condition at the 4-digit sector aggregation level is attributable to the reduced coverage of total trade. Accordingly:

Stylized fact 3

The value of trade represented by all cases violating the Hillman condition is hardly affected by the level of sector aggregation.

Stylized Fact 1 indicates that the set of observations violating the Hillman condition represents a substantial part of total trade. Indeed, 1% of total trade corresponds to more than 12 billion US $. Stylized Fact 3 reveals that this value is hardly affected by the level of sector aggregation. No matter at which level of sector aggregation revealed comparative advantage is examined, the group of observations violating the Hillman condition remains equally important as to the value of trade they represent. Checking for the validity of the Hillman condition and dismissing those observations not passing the test thus seems to be an obvious routine to be used under all circumstances.

The reduction in coverage of total trade at the 4-digit level of sector aggregation is due to yet another problem: erroneous trade flow classifications. Indeed, erroneous data aggregation is a problem in applied research if it remains unnoticed. For empirical studies into revealed

comparative advantage the Hillman condition appears to be an effective screening device for detecting these errors.

3.2 Trade flow classifications

An important advantage of analyzing trade flows in general and comparative advantage in particular at lower, more detailed levels of aggregation is the increased coherence and homogeneity of the specific markets analyzed, and therefore the more precise identification of revealed comparative advantage. An important disadvantage is that some part of all trade is not specified at lower levels of aggregation, such that a lower share of total trade is represented by the data. Table 1 shows that here this is especially relevant at the 4-digit level of sector aggregation which represents only some 60% of total trade.

Identifying and subsequently ignoring the remaining 40% is important however. At the 4-digit level raw data could contain trade flows effectively classified at the 3-digit level. For the dataset used here this occurs for instance for category 752A/X 'automatic data processing machines & units thereof' which could refer to trade flows in any of the more detailed true 4-digit SITC categories 7521, 7522, 7523, 7524, 7525 or 7528. Similar problems apply to data classifications at the 2-digit and 3-digit levels of aggregation (see Feenstra (2000) for further details). In all these cases export flows are inflated, possibly to a very large extend, yielding artificially high values of the Balassa index.

As an illustration the Hillman condition is re-examined for all countries, sectors, and years at the 3-digit level, this time including the 1-digit and 2-digit 'aggregates' that are reported at the 3-digit level in the raw data. This yields in total 188 violations of the Hillman condition (compared to 88 violations in case only 'true' 3-digit sectors are considered), 108 of which are attributable to erroneous data classification. The latter are listed in Table 2.

Table 2 Violations of the Hillman condition in case of erroneous sector aggregation, 1970 – 1997.

code	description	Country	Years
010	Meat and meat preparations	Hungary	88, 89, 91, 94
020	Dairy products and birds eggs	Hungary	78-83
040	Cereals and cereal preparations	Hungary	73, 76, 81, 92

050	Vegetables and fruit	Hungary	71
100	Beverages and tobacco	Sri Lanka	74
110	Beverages	Hungary	93, 94
200	Crude materials, inedible, except fuels	China	70-72, 74-78, 80, 81
		Austria	93
300	Mineral fuels, lubricants and related materials	Czechoslovakia	76, 78
		Guinea-Bissau	80
		Austria	93
320	Coal, coke and briquettes	Former USSR	80
		Hungary	95
330	Petroleum, petroleum products and related material	Former USSR	86
400	Animal and vegetable oils, fats, and waxes	China	70-76
		Czechoslovakia	78
420	Fixed vegetable oils and fats	Hungary	92, 94, 95
500	Chemicals and related products n.e.s.	China	71, 75
		Papua N. Guin.	79, 80
		Austria	93
600	Manufactured goods classified chiefly by material	Guinea-Bissau	79
		Germany	82-87, 89-92, 94, 96, 97
700	Machinery and transport equipment	Zaire	76
		Guinea-Bissau	79
		Germany	82-87, 90-92, 94, 96
790	Other transport equipment	Hungary	94
800	Miscellaneous manufactured articles	Ireland	70-72
		Neth Antilles	78
		Guinea-Bissau	79
		Germany	82-87, 90-92, 94, 96, 97
840	Articles of apparel and clothing accessories	Mauritius	96
900	Commodities & trans. not classified elsewhere	Ireland	70
		Czechoslovakia	78, 79
		Germany	90-94, 96, 97

For 85 out of these 108 cases (or 79 percent) the violation of the Hillman condition is the result of the respective country having a reported 'monopoly'. For all cases the market share is at least 92.6 percent. In the 1980s and 1990s, for example, Germany frequently is the only country classifying products at the 'miscellaneous' 1-digit level, the categories '600', '700', '800', and '900' in Table 2, resulting in an artificial monopoly. Similarly, while most other countries take the trouble to identify if the exported 'dairy products and birdseggs' are either 'milk and cream', 'butter', 'cheese and curd', or 'eggs and yolks, fresh, dried, or otherwise preserved', Hungary simply lists them as 'dairy products and birdseggs'. Although not leading to a monopoly for Hungary in the years 78-83, the Hillman condition does pick up this classification problem, as it does for Hungary's classification of sectors '010', '040', and '050'. Clearly, the Hillman condition is most useful in identifying observations based on erroneous trade flow classifications. Therefore:

Stylized fact 4

The Hillman condition is an effective screening device for identifying observations of revealed comparative advantage that are based on erroneously classified trade flows.

At the same time, because of erroneous data classifications the number of true violations of the Hillman condition dropped from 88 to 80. This is due to the effect of artificially enlarged trade flows on all computed Balassa index values (1) and the concomitant Hillman condition (2). For eight cases this means that they are not identified as violating the Hillman condition when aggregate trade flows are erroneously classified at the 3-digit level, while they are identified as such when restricting the analysis to true 3-digit level trade flows only. These observations are listed in Table 3 and give rise to the following:

Stylized fact 5

The Hillman condition suffers from a masking effect in that mild violations remain undetected if groteske violations are present.

Table 3 Violations of the Hillman condition that are masked due to erreneous trade flow classifications, 3 digit

Code	Country	Year	Market share	Export specialization	Country size
Erroneous data classification					

333	Iraq	71
333	Iraq	81
999	Former USSR	71
999	Former USSR	72
999	Former USSR	76
999	Former USSR	79
999	East Germany	70
999	East Germany	71
	Correct data classification	
333	Iraq	71
333	Iraq	81
999	Former USSR	71
999	Former USSR	72
999	Former USSR	76
999	Former USSR	79
999	East Germany	70
999	East Germany	71

Stylized Facts 4 and 5 jointly imply a natural research sequence. Given any dataset compute for all observations the Balassa index and the related Hillman condition. Examine the so identified observations not passing the Hillman test and dismiss those observations that are suspect of pertaining to erroneous data classifications. Re-calculate the Balassa index and the concomitant Hillman condition for all remaining sectors, whereby it is important to include for total country trade flows observations that are exclusively recorded at higher levels of sector aggregation in order not to underestimate these cumulatives.

Indeed, all calculations of the remainder of this paper are performed only at the appropriate level of sector aggregation whereby country's true total trade flows in any given year are used, thus including trade flows classified exclusively at higher levels of aggregation. The same then applies for the calculation of total world trade.

4. Empirical violations of the Hillman condition

For all 165 sample countries the Hillman condition is verified for all 28 sample years at the four different levels of sector aggregation. In what follows the concomitant empirical regularities are

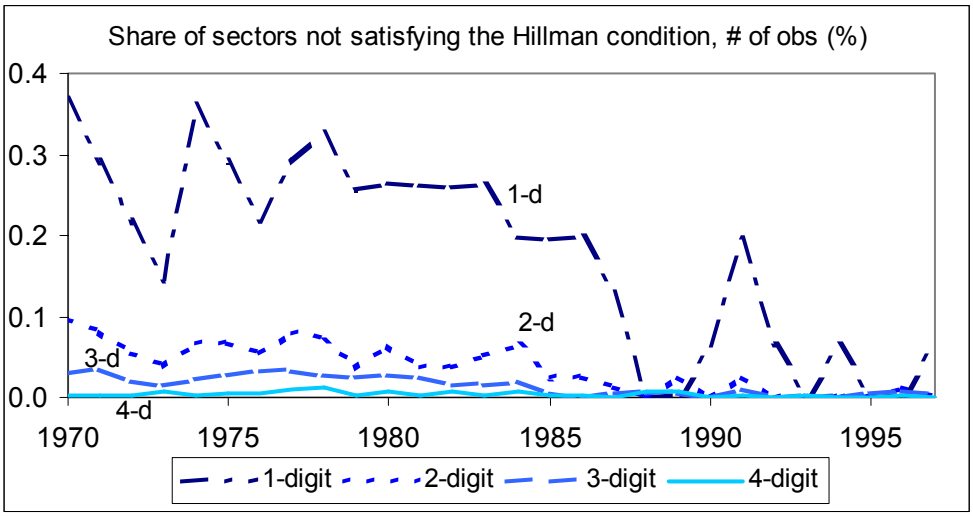
ordered along three dimensions: time (Section 4.1), sectors (Section 4.2), and countries (Section 4.3).

4.1. Empirical violations of the Hillman condition over time

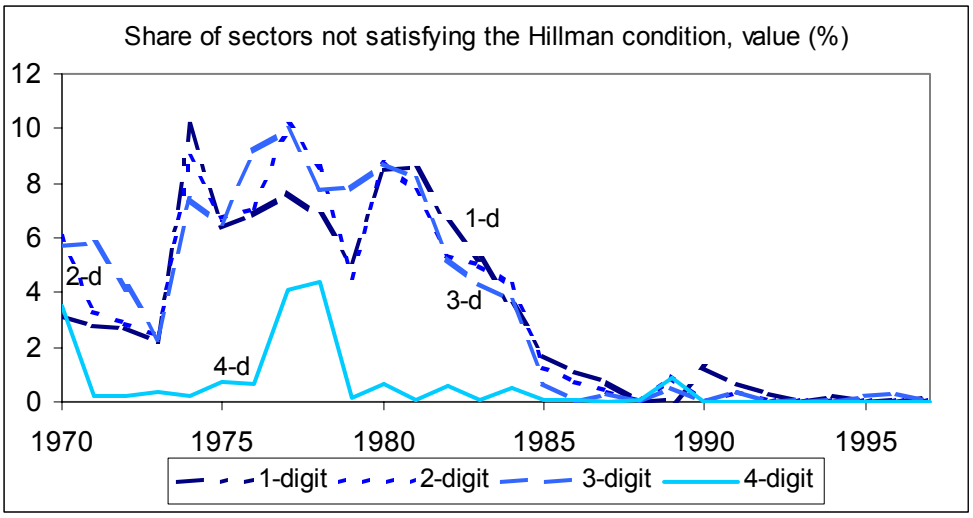
Table A2 in Appendix 3 lists for each sample year and all four levels of sector aggregations the observations that violate the Hillman condition as a fraction of the total (annual) number of observations and of total (annual) trade. The annual average violations correspond well to the sample totals in Table 1, both in terms of the number of violations and in terms of the value of trade represented by these observations.

Figure 2 Development over time of violations of the Hillman condition, 1970 – 1997.

Panel a



Panel b



Considering the evolution of violations over time reveals quite a different picture however. Figure 2 illustrates this development for the four different levels of aggregation, both as a share of the number of observations (panel *a*) and as a share of total trade (panel *b*). For all years the number of observations violating the Hillman condition is small (as a fraction of all annual observations never to exceed 0.4 percent in any given year), and decreasing over time. This reduction is even more pronounced in panel *b*; violations of the Hillman condition are very valuable in the period 1970-84, with a peak of 10.14 percent at the 1-digit level in 1974, to become much less important and to virtually disappear in the period 1988 - 1997. Hence:

Stylized fact 6

Concerning violations of the Hillman condition over time two periods can be distnguished: (i) 1970 – 1984, during which violations occur relatively frequent and represent a substantial fraction of total trade, and (ii) 1985 – 1997, during which violations hardly ever occur and represent an insignificant fraction of total trade.

It follows that the single year (1985) investigated by Marchese and Nadal De Simone (1989) is not representative for the empirical relevance of the Hillman condition. Violations are much more pronounced in the 15 years preceeding the year 1985, while in the following years their importance gradually vanished. It also means that studies into revealed comparative advantage that ignored the Hillman condition are more prone to errors if they include observations from the seventies until the mid-eighties of the last century.

4.2. Empirical violations of the Hillman condition across sectors

Table 4 lists all observations violating the Hillman condition for the four different levels of sector aggregation. At the 1-digit level of sector aggregation violation of the Hillman condition occurs in two sectors only: "Food and live animals chiefly for food" (SITC code 0), and "Mineral fuels, lubricants and related materials" (SITC code 3). Going from this 1-digit level to the 2-digit level of sector aggregation the "non-ferrous metals" sector (SITC code 68) joins the group of violating sectors. At the 3-digit level four additional sectors are included: "Fertilizers, crude" (SITC code 271), "Radio-active and associated materials" (SITC code 286), "Copper" (SITC code 524), and "Ships, boats and floating structures"(SITC code 793). The latter two sectors leave the basket of violators again if sectors are considered at the 4-digit level of aggregation.

It thus appears that more detailed sector definitions encompass the group of sectors violating the Hillman condition at higher levels of sector aggregation (the narrowing scope of violating sectors when going from the 3-digit level of sector aggregation to the 4-digit level is related to the concomitant reduction in trade flow coverage). At the same time, at lower levels of sector aggregation violations occur in sectors that are not identified as violators at higher sector aggregation levels. Accordingly:

Stylized fact 7

The correlation between sectors violation the Hillman condition across levels of sector aggregation is asymmetric; violations at lower levels of sector aggregation are likely to occur at higher levels as well, while violations at higher levels of sector aggregation need not to occur at lower levels.

Table 4 Observations not satisfying the Hillman condition, 1970 - 1997

SITC code	Description	Country	Years
	1-digit		
0	Food and live animals chiefly for food	St Pierre Miqu	79
3	Mineral fuels, lubricants and related materials	Algeria	79-87, 91, 92, 94, 97
		Libya	70-83, 85, 86
		Venezuela	70, 74
		Kuwait	70-72, 74
		Qatar	70, 71
		Saudi Arabia	70-84, 90
		Iran	74-78, 83
		Oman	75
		Iraq	77, 78, 80, 85-87
		Un Arab Em	78

		Nigeria	81, 82, 84, 91
		Paraguay	91
	2-digit		
06	Sugar, sugar preparations and honey	Cuba	75-78
33	Petroleum, petroleum products and related material	Libya	70-81, 83, 86
		Venezuela	70
		Kuwait	70-72, 74
		Qatar	70, 71
		Saudi Arabia	70-84
		Iraq	71, 77, 78, 80, 83-89
		Iran	74-78, 83, 84
		Oman	75
		Nigeria	81, 82, 84, 85, 91
35	Electric current	Paraguay	91
68	Non-ferrous metals	Zambia	70-74
93	Special transactions & commod., not class. to kind	South Africa	80
99	Non-identified products	Former USSR	70, 77, 78
		Zimbabwe	79
		Romania	80, 82, 84
		East Germany	89
		Reunion	96
	3-digit		
061	Sugar and honey	Cuba	75-78
271	Fertilizers, crude	Morocco	74
286	Ores and concentrates of uranium and thorium	Niger	78-80, 81
333	Petrol. oils & crude oils obt. from bitumin. minerals	Libya	70-77, 80, 81
		Qatar	70, 71, 76
		Saudi Arabia	70-84
		Iran	74-78, 83, 84
		Oman	75
		Iraq	71, 77-81
		Nigeria	79, 81-85, 87, 91
		Kuwait	95, 96
351	Electric current	Paraguay	91
524	Radio-active and associated materials	Niger	88
682	Copper	Zambia	70-77
793	Ships, boats and floating structures	Reunion	97
931	Special transactions & commod., not class. to kind	South Africa	80
999	Non-identified products	Former USSR	70-72, 76-79
		Zimbabwe	79
		Romania	80, 82, 84
		East Germany	70, 71, 89
		Reunion	96
	4-digit		
0611	Sugars, beet and cane, raw, solid	Cuba	75-77, 78
2479	Pitprops, poles, piling, posts & other wood in rough	Indonesia	73
2814	Roasted iron pyrites, whether or not agglomerated	Brazil	86, 88
		Papua N. Guin.	93
2873	Aluminium ores and concentrates (includ.alumina)	Guinea	78, 82-85, 91
		Jamaica	80, 81
3359	Petroleum oil prep & residues nes	Neth Antilles	88, 89
6821	Copper and copper alloys, refined or not, unwrought	Zambia	71-78

9999	Non-identified products	Former USSR	70, 77, 78
		Zimbabwe	79
		Romania	80, 82, 84
		East Germany	89
		Reunion	96

A further sector classification is obtained when taking into account the related factor intensity. This leads UNCTAD/WTO to distinguish six sector categories, which are described in detail in van Marrewijk (2002): (i) primary, (ii) natural-resource intensive, (iii) unskilled-labour intensive, (iv) technology intensive, (v) human-capital intensive, and (vi) not classified. In appendix 3 the ordering of all 3-digit SITC sectors according to these six categories is listed.

Violations of the Hillman condition appear to be concentrated in two categories only: primary products and natural-resource intensive products. Considering Hillman condition (2) this comes not as a surprise. It is precisely in these two sector categories more likely for countries to enjoy a (natural) large market share of world trade and/or, to specialize exclusively in the export of these commodities. That is:

Stylized fact 8

At all levels of sector aggregation violations of the Hillman condition occur almost exclusively in primary product sectors and natural-resource intensive sectors.

Contrary to one of the implications of Stylized Fact 6, studies that did not include primary product sectors and/or natural-resource intensive sectors are not exempt from possible inclusion of erroneous observations on revealed comparative advantage. This would be the case only if these industries *are* included when calculating total trade for all sample countries.

4.3. Empirical violations of the Hillman condition across countries

Further empirical regularities can be distinguished if the sample of violations is examined along the country dimension. All countries hosting observations that violate the Hillman condition at the 2-digit level of sector aggregation are present in the group of violators at the 3-digit sector aggregation level, while 9 of the 13 countries with violations at the 1-digit sector aggregation level are home to violators at the 2-digit level of sector aggregation as well. The consequences of the reduction in trade flow coverage when considering the 4-digit level of sector aggregation is quite apparent in this context: only 7 out of 20 countries remain listed as the home country of

sectors violating the Hillman condition at the 3-digit level of sector aggregation. This drop in trade flow coverage should not blur the following:

Stylized fact 9

The correlation between countries hosting sectors that violate the Hillman condition across levels of sector aggregation is asymmetric; violations at lower levels of sector aggregation are likely to occur at higher levels as well, while violations at higher levels of sector aggregation need not to occur at lower levels.

Finally, for identifying stylized facts across groups of countries the set of sample countries needs to be ordered. For that the classification of the World Bank of all countries into 7 distinct geographical regions is used (see World Bank, 2003): East Asia and Pacific (EAP, 25 countries), Eastern Europa and Central Asia (ECA, 29 countries), North America (NAM, 3 countries), Latin America and the Caribbean (LAC, 35 countries), Middle East and North Africa (MNA, 21 countries), South Asia (SAS, 8 countries), and Sub-Sahara Africa (SSA, 44 countries). As there is a good, but not perfect, correspondence between the country labelling of the World bank and the set of sample countries used in this study, a congruence is constructed (see Appendix 1 for details).

Stylized fact 10

At all levels of sector aggregation violations of the Hillman condition occur foremost for observations involving countries in Africa (including the Middle East), and, to a lesser extend, involving countries in Latin America, the Carribean, and Eastern Europe.

Observe again that studies into revealed comparative advantage that did not include countries in Africa, the Middle East and Latin America are not flawless *a priori* as sectors from these countries contribute to total world trade and hence affect the calculated value of the Balassa index (1) and the concomitant Hillman condition (2). Stylized fact also explains the substantial difference between Marchese and Nadal de Simone (1989) and Hinloopen and Van Marrewijk (2001) as to the number of observations violating the Hillman condition. Indeed, Hinloopen and Van Marrewijk (2001) consider countries from the European Union only while Marchese and Nadal de Simone (1989) focus exclusively on a set of developing countries.

6. Conclusions

Using a comprehensive data set of annual bilateral trade flows for 1,056 4-digit SITC sectors between 183 countries for the years 1970 - 1997, the empirical relevance is examined of the necessary and sufficient condition for the correspondence between revealed comparative advantage, as measured by the Balassa index, and pre-trade relative prices. The findings on the empirical relevance of this Hillman condition are presented as stylized facts because of the exhaustive representation of the dataset.

It appears that (i) violations of the Hillman condition are small as a share of the number of observations, but often represent a disproportionally large value of trade, (ii) from 1970 through 1984 violations happen relatively frequent and they represent a significant fraction of the value of total trade, while from 1985 onwards violations hardly ever occur and represent an insignificant fraction of total trade value, and (iii) violations do not occur randomly across sectors or countries; they occur foremost in sectors producing primary products or that are natural-resource intensive, and in sectors that are located in countries in Africa, the Middle East, and Latin America.

Restricting empirical analyses into revealed comparative advantage to those industries that meet the Hillman condition has the additional advantage that it acts as a screening device for observations that are based on erronously classified trade flows.

References

Amity, M. (1999), "Specialization patterns in Europe," *Weltwirtschaftliches Archiv*, **135**: 573-593.

Balassa, B. (1965), "Trade liberalization and 'revealed' comparative advantage," *The Manchester School of Economic and Social Studies* **33**: 92-123.

Bojnec, S. (2001), "Trade and revealed comparative advantage measures: regional and central and east European agricultural trade," *Eastern European Economics*, **39**: 72-98.

Feenstra, R. C. (2000), *World trade flows; 1980 – 1997*, Institute of Governmental Affairs, University of California, Davis.

Feenstra, R., R. Lipsey, and H. Bowen (1997), "World trade flows, 1970 – 1992, with production and tariff data," NBER Working Paper No. 5910.

Fertö, I., and L.J. Hubbard (2003), "Revealed comparative advantage and competitiveness in Hungarian Agri-food sectors," *World Economy*, **26**: 247-259.

Hillman, A.L. (1980), "Observations on the relation between 'revealed comparative advantage' and comparative advantage as indicated by pre-trade relative prices," *Weltwirtschaftliches Archiv*, **116**: 315 – 21.

Hinloopen, J. and C. van Marrewijk (2001), "On the empirical distribution of the Balassa index", *Weltwirtschaftliches Archiv*, **137**: 1 – 35.

Hinloopen, J., and C. van Marrewijk (2004), "Dynamics of Chinese comparative advantage", Tinbergen Institute Discussion Paper, 2004-034/2.

Liesner, H.H. (1958), "The European Common Market and British industry," *Economic Journal* **68**: 302-316.

Marchese, S., and F. Nadal De Simone (1989), "Monotonicity of indices of 'revealed' comparative advantage: empirical evidence on Hillman's condition," *Weltwirtschaftliches Archiv* **125**: 158 – 167.

Marrewijk, C. van (2002), *International trade & the world economy*, Oxford: Oxford University Press.

Porter, M. (1990), *The competitive advantage of nations*, MacMillan, London.

Proudman, J. and S. Redding (1998), "Persistence and mobility in international trade," in: J. Proudman and S. Redding (eds.), *Openness and growth*, Bank of England, Chapter 2.

Proudman, J. and S. Redding (2000), "Evolving patterns of international trade," *Review of International Economics*, **8**: 373-396.

Svaleryd, H. and J. Vlachos (2005), "Financial markets, the pattern of industrial specialization and comparative advantage: evidence from OECD countries", *European Economic Review*, **49**: 113 – 144.

World Bank, 2003, *World Development Report 2003*, Washington DC: The World Bank.

Appendix 1. Data

Two separate data sets provided by the Center for International Data, University of California, Davis (CID/UCD), were merged, the first covering the years 1970 through 1993 (see Feenstra, Lipsey and Bowen, 1997) and the second covering the years 1980 through 1997 (see Feenstra, 2000). For the overlapping years, the data from the latter source are used. The data set contains bilateral trade flows between 183 trading partners, including n.e.s. (not elsewhere specified) regions for trade flows that could not be classified further than within a broad geographical region (such as "Middle East", or "North Africa"), an "Areas n.e.s." region for trade flows that cannot be attributed to any country or to any of the used broad geographical regions but that do come from a well-defined geographical region, and an "Unknown Partner" category for trade flows that could not be attributed at all due to various reasons (see Feenstra, 2000).

This leaves a sample of 165 genuine countries that are grouped in Table A1 according to the World Bank classification of world regions (see World Bank, 2003). In some cases the latter is more detailed than the sample of countries used here. For instance, The World Bank distinguishes between Bosnia, Croatia and Slovenia. Here all of these three countries fall under the heading "(former) Yugoslavia". As long as the countries that are grouped together belong to the same geographical region as identified by the World Bank the classification of the latter still applies. This turns out always to be the case. On the other hand, countries that are distinguished here but not as such in the World Bank classification are grouped according to their geographical location. This was done for the Falkland Islands, Guadeloupe, Reunion, St. Helena, and St. Pierre Miqu. Finally, the constructed database contains three entries for Yemen: Former Democratic Republic of Yemen, Former Yemen, and Yemen. The World Bank classification includes Yemen only. Needless to say that all three identified countries belong to the same geographical area (in the Table A.1 below these are not further distinguished).

Table A.1 Sample country classification according to World Bank regions

East Asia & Pasific (EAP); 25 countries		
Australia	Laos	Philippines
Brunei	Malaysia	Solomon Islands
Cambodia	Mongolia	Thailand
China	Myanmar (Burma)	South Korea
Fiji	New Zealand	Singapore
Hong Kong	North Korea	Taiwan
Indonesia (incl. Macau)		Vietnam
Japan	New Caledonia (incl. French Polynesia, and Vanuata)	
Kiribati (incl. Tonga, and Tuvalu)	Papua New Guinea	
Europe & Central Asia (ECA); 29 countries		
Albania	Greece	Portugal
Austria	Greenland	Romania
Belgium-Luxemburg	Germany	Spain
Bulgaria	Hungary	Sweden
Cyprus	Iceland	Switzerland
Czechoslovakia	Ireland	Turkey
Denmark (incl. Faroe Islands)	Italy	United Kingdom
(former) East Germany	Netherlands	(former) USSR
Finland	Norway	(former) Yugoslavia (incl. Croatia, and Slovenia)
France	Poland	
North America (NAM); 3 countries		
Bermuda	Canada	USA
Latin America & Caribbean (LAC); 35 countries		
Argentina	Ecuador	Nicaragua
Bahamas	El Salvador	Panama
Barbados	Falkland Islands	Paraguay
Belize	French Guiana	Peru
Bolivia	Guadeloupe (incl. Martinique)	St. Kitts & Nevis (incl. Dominica, Montserrat, St. Luca, St. Vincent, and Grenada)
Brazil	Guatemala	St. Pierre Miqu
Cayman Islands	Guyana	Surinam
Chile	Haiti	Trinidad & Tobago
Colombia	Honduras	Turks Caicos Islands
Costa Rica	Netherlands Antilles	Uruguay
Cuba	Jamaica	Venezuela
Dominican Republic	Mexico	
Middle East & North Africa (MNA); 21 countries		
Algeria	Iraq	Oman
Bahrain	Jordan	Qatar
Djibouti	Kuwait	Saudi Arabia
Egypt	Lebanon	Syria
Gibraltar	Libya	Tunisia
Israel	Malta	United Arab Emirates
Iran	Morocco	Yemen
South Asia (SAS); 8 countries		
Afghanistan	India	Pakistan
Bangladesh	Maldives	Sri Lanka

Bhutan	Nepal	
Sub-Saharan Africa (SSA); 44 countries		
Angola	Ghana	Rwanda
Benin	Guinea	Senegal
Burkina Faso	Guinea-Bissau (incl. Cape Verde)	Seychelles
Burundi	Kenya	Sierra Leone
Cameroon	Liberia	Somalia
Central African Republic	Madagascar	South Africa
Chad	Malawi	St. Helena
Comoros	Mali	Sudan
Congo	Mauritania	Tanzania
Cote dÍvoire	Mauritius	Togo
Democratic Republic Congo (Zaire)	Mozambique	Uganda
Equatorial Guinea	Niger	Westren Sahara
Ethiopia	Nigeria	Zambia
Gabon	Reunion	Zimbabwe
Gambia	Republic Congo	

The bilateral trade flows are decomposed into 1,249 sectors, comprising 747 genuine 4-digit sectors, based on SITC (Standard International Trade Classification), revision 2. The remaining 502 sectors refer to aggregates at the 1-, 2-, or 3-digit level, and a "Non-identified products" category. The 4-digit subset contains 60.39 % of all trade, the 3-digit subset covers 99.46 % of all trade, and the 2-digit subset comprises 99.67 % of all trade.

The data were first compiled by Statistics Canada and made available through the CID/UCD (see Feenstra, 2000). The former makes use of various sources (according to Statistics Canada 87% of all trade flows is based on independent sources of *both* imports and exports, while 98% is based on reports of at least one side of trade), yielding a rather complete coverage of world trade flows. The CID/UCD transforms the data such that trade flows for all years, all countries, and all industry groups are consistent and presented in a unified manner. Each observation in the raw data consists of four entries: importing country, exporting country, sector, and size of the trade flow (in 1,000 US $). The data are thus classified according to the importing country. This is *not* to say that the data are based on import sources only, as explained above. After merging the two separate datasets a second dataset is created by "inverting" the data, in that all trade is classified according to the exporting country.

Appendix 2. Violations of the Hillman condition

Table A.2 Annual violations of the Hillman condition, 1970 – 1997.*

	share of # of observations (%)				share of value of exports (%)			
year	1-digit	2-digit	3-digit	4-digit	1-digit	2-digit	3-digit	4-digit
1970	0.37	0.10	0.03	0.00	3.17	6.08	5.67	3.50
1971	0.29	0.08	0.03	0.00	2.77	3.31	5.79	0.24
1972	0.22	0.05	0.02	0.00	2.70	2.86	4.20	0.23
1973	0.14	0.04	0.01	0.01	2.21	2.38	2.33	0.38
1974	0.36	0.07	0.02	0.00	10.14	8.96	7.46	0.23
1975	0.29	0.07	0.03	0.01	6.37	6.75	6.53	0.71
1976	0.22	0.05	0.03	0.01	6.87	7.13	9.20	0.64
1977	0.29	0.08	0.03	0.01	7.68	10.12	9.96	4.08
1978	0.33	0.07	0.03	0.01	6.78	8.61	7.79	4.42
1979	0.26	0.04	0.03	0.00	5.09	4.53	7.81	0.12
1980	0.26	0.06	0.03	0.01	8.47	8.75	8.69	0.67
1981	0.26	0.04	0.02	0.00	8.52	7.73	8.17	0.08
1982	0.26	0.04	0.01	0.01	6.53	5.36	5.25	0.62
1983	0.26	0.05	0.01	0.00	5.20	4.96	4.34	0.04
1984	0.20	0.06	0.02	0.01	3.51	4.32	3.71	0.50
1985	0.20	0.02	0.00	0.00	1.67	1.21	0.66	0.04
1986	0.20	0.02	0.00	0.00	1.07	0.71	0.00	0.10
1987	0.13	0.01	0.00	0.00	0.71	0.40	0.29	0.00
1988	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.09
1989	0.00	0.02	0.00	0.01	0.00	0.92	0.52	0.87
1990	0.06	0.00	0.00	0.00	1.21	0.00	0.00	0.00
1991	0.20	0.02	0.01	0.00	0.68	0.36	0.36	0.03
1992	0.06	0.00	0.00	0.00	0.28	0.00	0.00	0.00
1993	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
1994	0.06	0.00	0.00	0.00	0.20	0.00	0.00	0.00
1995	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.00
1996	0.00	0.01	0.01	0.00	0.00	0.00	0.27	0.01
1997	0.07	0.00	0.00	0.00	0.25	0.00	0.00	0.00
ann. aver.	0.18	0.04	0.01	0.00	3.29	3.41	3.55	0.63

* The total number of cases not satisfying the Hillman condition in the period 1970-1997 is 73 at the 1-digit level, 79 at the 2-digit level, 88 at the 3-digit level, an 35 at the 4-digit level.

Appendix 3. Sector classifications according to factor intensity

Primary															
001	011	012	014	022	023	024	025	034	035	036	037	041	042	043	044
045	046	047	048	054	056	057	058	061	062	071	072	073	074	075	081
091	098	111	112	121	122	211	212	222	223	232	233	244	245	246	247
248	251	261	263	264	265	266	267	268	269	271	273	274	277	278	281
282	286	287	288	289	291	292	322	323	333	334	335	341	351	411	423
424	431	941													
Natural-resource intensive															
524	611	612	613	633	634	635	661	662	663	667	671	681	682	683	684
685	686	687	688	689											
Unskilled-labour intensive															
651	652	653	654	655	656	657	658	659	664	665	666	793	812	821	831
842	843	844	845	846	847	848	851	894	895						
Technology intensive															
511	512	513	514	515	516	522	523	541	562	572	582	583	584	585	591
592	598	711	712	713	714	716	718	721	722	723	724	725	726	727	728
736	737	741	742	743	744	745	749	751	752	759	764	771	772	773	774
775	776	778	792	871	872	873	874	881	882	883	884	893	951		
Human-capital intensive															
531	532	533	551	553	554	621	625	628	641	642	672	673	674	675	676
677	678	679	691	692	693	694	695	696	697	699	761	762	763	781	782
783	784	785	786	791	885	892	896	897	898	899					
Not classified															
911	931	961	971	999											