

第2章

台湾貿易データにおける UN 貿易データ準拠への試み

海老原悦夫・野田容助

はじめに

貿易データを利用して国際比較あるいは分析をおこなうときに、特に東アジア諸国・地域を対象とするときには台湾の貿易データの必要性は避けて通れない。アジア経済研究所では国連 (UN) 作成の貿易データである UN 貿易データ (UN Comtrade Database 貿易データ) として報告されなくなった 1971 年以降の台湾貿易データを当研究所独自の作成方法により UN 貿易統計に準拠した形式および内容に変換し、国際比較可能なアジア経済研究所世界貿易データシステム (Ajiken Indicators of Developing economies: eXtended for Trade statistics : AID-XT) の基礎データとして作成し利用している。

本章の目的は台湾貿易データを UN Comtrade Database 貿易データの書式 (フォーマット) へ変換することの試みである。すなわち、アジア経済研究所では台湾貿易データを AID-XT 基礎データに変換して国際比較に供してきたが、最近では UN Comtrade Database 貿易データが on-line 検索により手軽にしかも安価で利用できるようになったのに伴って、台湾貿易データを貿易分類カテゴリーについても UN Comtrade Database 貿易データに準拠した内容および形式に変換する必要が生じてきたことによる。アジア経済研究所が入手している他の国の貿易統計のほとんどは、国連へ各国が報告したものであるから、変換ができれば台湾の貿易統計も他の

貿易統計と同類となり、今後の利用に便利であることが期待される。変換は書式をあわせるだけではなく、UN Comtrade Database 貿易データで使用されている国コード、数量単位にあわせてコード変換する必要がある。

本章は台湾貿易データを AID-XT 基礎データへ変換する方法の概要、UN Comtrade Database 貿易データに準拠 (UN 貿易データへの準拠) するための変換方法、変換のためのデータ処理、から構成されている。前述の「UN 貿易統計準拠」は AID-XT 基礎データへの変換のことであり、ここでいう「UN 貿易データへの準拠」とは異なる意味である。前者の変換については野田 [6] の「台湾の AID-XT 基礎データの作成と評価」が詳しい。本章は後者の変換に重点を置いており、(1) 台湾と国連のそれぞれで利用されている国コードの対応表の作成、(2) 台湾と国連の数量単位の対応表の作成、(3) 台湾貿易統計に対する国コード、数量単位の変換、(4) 商品分類コード、国コードが同じものを集計して csv 形式で出力、のプロセスから構成される。

1. 台湾の AID-XT 基礎データ作成

アジア経済研究所の AID-XT 基礎データの基礎となる貿易データとしてかつては UN Comtrade Database 貿易データと共に経済協力開発機構 (OECD) 作成による貿易データである

OECD 貿易データ (International Trade by Commodity Statistics: ITCS) を利用していたため、両者の共通分類としてのアジア研統一分類体系を利用している。この分類は UN Comtrade Database の分類とは必ずしも一致してはいない。そのため、本章の目的の必要性が生じているわけである。OECD 貿易データには最近年度を対象とした台湾貿易データが非 OECD 諸国として含まれている。

本節では野田 [6] にもとづき、台湾貿易データが使用している貿易の分類カテゴリーを示し、AID-XT 基礎データの作成方法とその評価の概要を示す。台湾の AID-XT 基礎データは商品総額を基準としたときにそれと個別商品分類の取引額の和が一致するというサムチェックという意味では整合性が保たれていることが確認できる。説明は省略しているが、個別相手国の整合性も保たれていることも確認できている。

1.1 台湾貿易データの分類カテゴリー

台湾貿易統計において詳細な商品分類で報告されている貿易統計書には台湾財政部關稅總局統計室 (Statistical Department Directorate General of Customs Ministry of Finance, The Republic of China) の発行する年報「中國進出口貿易統計年刊 (臺灣區)」(*The Trade of China, Taiwan District*) および月報「中華民國臺灣地區出口貿易統計月報」(*Monthly Statistics of Exports, The Republic of China, Taiwan District*) と「中華民國臺灣地區進口貿易統計月報」(*Monthly Statistics of Imports, The Republic of China, Taiwan District*) がある^(註1)。台湾貿易統計書をもとにして当研究所では以下の方法により UN 貿易統計に準拠した形式および内容のデータに変換して AID-XT 基礎データを作成している。

(1) 1971 年から 85 年までは磁気媒体によるデータの入手が不可能であったため、上記機関

の前身である海關總稅務統計處 (Statistical Department Inspectorate General of Customs, Taipei, Republic of China) の発行する「中國進出口貿易統計年刊 (臺灣區)」(*The Trade of China, Taiwan District*) に記載されている「Trade of China, Taiwan District」からアジア経済研究所が直接入力して磁気テープ化している^(註2)。

(2) 1986 年以降、前述した台湾の Statistical Department Directorate General of Customs Ministry of Finance から貿易統計月報 12 月と同一内容を含む磁気媒体によるデータを入手している^(註3)。

(3) 得られた台湾貿易データの磁気媒体によるデータはアジア経済研究所の独自の方法により商品分類、相手国のコードおよび取引額表示など UN 貿易統計に準拠した内容と形式を持つ AID-XT 基礎データに変換される。

台湾貿易データの商品分類は関税協力理事会 (Customs Cooperation Council) 作成の Customs Cooperation Council Nomenclature (CCCN) にもとづいて作成された Customs Imports Tariff of Republic of China (CCC) である。アジア経済研究所ではこの CCC を 1971 年から 1988 年までは UN 貿易統計および OECD 貿易統計と同時に利用して国際比較を可能にするために UN 貿易統計で使用されている SITC-R1 の基本項目である 4 桁または 5 桁に対応させている。1989 年以降、CCC は HS にもとづいて分類されるようになったため直接先頭から 6 桁レベルの分類コードを利用している。台湾貿易データには CCC のほかに SITC の分類にもとづく 4 桁コードもあるが、UN 貿易統計の商品分類が SITC の 4 桁または 5 桁を基本項目としてもっているため、4 桁の分類コードについては問題がないとしても 5 桁の分類コードでは配分問題の処理が発生してくる。しかも配分についてはその処理方法が必ずしも明確ではないため、CCC から SITC の基本項目への配分なしの対応表が必要となる。

表1 台湾貿易データにおける 1,000NewTaiwan\$から 1,000US\$への変換レート

y	import			export		
	1,000NT\$	1,000US\$	変換レート	1,000NT\$	1,000US\$	変換レート
1989	1385272906	52265326	26.50474	1745104214	66303952	26.31976
1990	1471803505	54716004	25.91628	1802781059	67214446	26.82133
1991	1690772977	62860545	26.89721	2040784970	76178309	26.78958
1992	1816294219	72006794	25.22393	2047961998	81470250	25.13754
1993	2034746748	77061203	26.40430	2234350533	85091458	26.25822
1994	2261650459	85349194	26.49879	2456010977	93048783	26.39487
1995	2742850423	103550044	26.48826	2949579951	111658800	26.41601
1996	2815119319	102370021	27.49945	3176624196	115942064	27.39838
1997	3276094156	114424665	28.63101	3481685084	122080673	28.51954
1998	3503568890	104665312	33.47402	3693268791	110582293	33.39837
1999	3576415487	110689863	32.31023	3917446187	121590945	32.21824
2000	4368695887	140010636	31.20260	4616301200	148320561	31.12381
2001	3619429516	107237394	33.75156	4137742390	122866261	33.67680
2002	3893680914	112530093	34.60124	4507506099	130596781	34.51468
2003	4383722834	127248508	34.44996	4952475993	144179522	34.34926
2004	5627586457	167889814	33.51951	5817800332	174014163	33.43294
2005*	5877164292	182614393	32.18353	6374493858	198431654	32.12601

(出所) 1,000US\$は 1,000US\$を単位とする取引総額であり、Directorate General of Budget, Accounting and Statistics Executive Yuan, Republic of China 発行による *Statistical Yearbook of the Republic of China* の中の Commerce and External trade から輸入は Import by regions and countries (areas) の total、輸出は Export by regions and countries (areas) の total、1,000NT\$は NewTaiwan\$を単位とする取引総額であり、AID-XT 基礎データにもとづき著者作成。

(注) 2004 年以前では輸出の欄は輸出と再輸入、輸入の欄は輸入と再輸出の変換レートをそれぞれ表示している。2004 年の総額は Gross の輸出入である。*Statistical Yearbook of the Republic of China* 2005 によれば、2006 年 1 月から台湾貿易統計の作成については UN の International Merchandise Trade Statistics Compiler's Manual, 2004 を採用している。そのため 2005 年についての輸出は Gross 輸出の 6084207901 と再輸出の 290285957 の合計である 6374493858 として表わされている。同じように輸入は輸入と再輸入の合計が計上されている。

アジア経済研究所で独自に作成した対応表は CCC の一番細かい分類に対して SITC-R1 へと対応する配分構造のない統合のみの対応関係から構成される^(注4)。1980 年の CCC の桁数は 7 桁、81 年から 87 年までは 8 桁である。

貨物の国別分類である相手国について 1970 年代頃の台湾貿易統計書にはそれがコード化されていない。そのため、入力された貿易統計データは相手国の名称がそのまま相手国コードの代わりに利用されているため、この名称にアジア研統一国コードを直接対応させて対応表を作成している。また、使用されている相手国の国コードは地域の範囲としては国・関税地域と世界計のみをカバーしており、OECD 貿易統計および UN 貿易統計のように地域ブロック、経済ブ

ロックは存在しない。磁気媒体のデータとして得られた 1986 年以降の貿易統計データでは相手国コードは 3 桁の数字で表示され、さらに、1993 年以降になると ISO の 2 桁レベルのアルファベット表示と 3 桁の数字の併記で表されるように変更されている。

台湾貿易データの取引額表示は現地通貨である NewTaiwan\$を用いて表わされている。この表示を UN 貿易統計で共通に用いられる 1,000US\$に変換する輸出入別の為替レートは表 1 に示されている。1971 年から 1988 年までは変換レートとして貿易統計書の最初に記載されている輸出入別、月別レートを単純に平均した値を利用している^(注5)。しかし、貿易統計書と AID-XT 基礎データとの比較において一致す

表2 HSのSub-headingと対応しない台湾の貿易商品分類コードとSITC-R3への対応

HS description	SITC-R3 description
140430 Sphagnum moss	29299 Vegetable materials, nes
291040 Octamethyl or other organic (phenyl or other organic functional groups) cyclotetrasiloxane	51614 Propylene oxide
330750 Cleaning solution for contactlens	55359 Toilet preparations, nes
540152 ...	99999 Good not classified by kind
710392 Rubies, sapphires and emeralds, otherwise worked	66739 Oth precious stones, worked
732189 Other apparatus, domestic use, not electrically operated, of iron or steel	69732 Domestic stoves, space heaters
989900 Special transactions not classed by kind	9310 Special Transactions

(出所) 台湾貿易データとAID-XT基礎データの商品分類改訂第1版の商品名マスターから筆者作成。

る有効桁数が小さいためそのずれを無くすように1989年以降はDirectorate General of Budget, Accounting and Statistics Executive Yuan, Republic of China 発行による *Statistical Yearbook of the Republic of China* にある *Commerce and External trade* を参照して1,000US\$の値を直接利用している。表1には1989年以降について1,000 New Taiwan\$と1,000 US\$の総額がそれぞれ示されている。

貿易統計に計上する輸出入貨物の価格表示については輸出は本船甲板渡し (FOB:Free on Board)、輸入は運賃保険料渡し (CIF:Cost, Insurance and Freight) である。取引額表示で特に注意すべきことは、貿易統計書年報に報告されている輸出と輸入の額は台湾の定義に従えばネットによる輸出入のみで計上であり、表示されている輸出額は輸出から再輸入を差し引いた額、また、同じく輸入額は輸入から再輸出を差し引いた額としていることである。そのため再輸出の方が輸入よりも取引額が上回っている場合や、再輸入が輸出よりも上回っている場合にはその商品はマイナスで表示される。ところが、磁気媒体により得られた86年以降のデータは貿易統計書の月報をもとにして作成されているため、輸出入のほかに再輸出入も含み必要に応じてグロスあるいはネットによる利用が可能となっている。ただし、1988年以前のAID-XT基

礎データはSITC-R1系列の連続性を保つようにネットによる表示に合わせて作成されており、再輸出入のデータは存在していない。HS系列に変更された1989年以降は輸出入に加えて再輸出入データが存在する。

商品分類の上位桁レベル分類コードへの積み上げの取引額および数量は下位の桁レベル分類コードのそれを合計することによって得る。レートの変換で使用される数字の有効桁数は4であり、丸めの誤差をできるだけ小さくするようにするため、データ処理にさいしては上位桁レベルの積み上げは台湾の通貨である1,000 New Taiwan\$のままで行い、最後にSITCの詳細分類コードおよびその集計された取引額に対して変換レートを乗じて1,000US\$表示の取引額へと変換している。

1.2 AID-XT 基礎データ作成

本節では具体例として商品分類がHSにもとづいて分類されている1989年から2002年までのAID-XT基礎データの作成を紹介する。それ以外の年についても同様な処理過程を必要とする。台湾貿易データからAID-XT基礎データを作成するための処理過程の概要は以下の通りである。

[1] 台湾財政部關稅總局統計室から台湾貿易

データの磁気媒体に記憶されたデータを購入する。このデータを台湾貿易データの（原）オリジナルデータと呼び、原データと同じ内容のものをコピーして正・副を作成し、保管する。AID-XT 基礎データ作成処理のためにオリジナルデータは年ごとに保存される。

[2] 年、輸出入区分ごとにオリジナルデータに含まれている分類カテゴリーの中のすべての商品分類、相手国、数量単位のそれぞれを抽出する。商品分類、相手国、数量単位はそれぞれ作業用ファイルとして一時的に保存される。

[3] 相手国は台湾貿易データの相手国とアジア研統一国コードとの対応表との比較をおこない、対応しないものがあれば新規に対応関係を検討し、対応表に追加する。数量単位についても同様な処理が必要であり、完成した数量単位の対応表は年ごとに保存される。

[4] オリジナルデータの商品分類コードの桁数を先頭から6桁目までを利用してHSにもとづく詳細分類コードを作成する^(注4)。HSの商品分類表には存在しないが台湾貿易データでは使用されている分類コードがあればそれは台湾固有の商品分類コードであるので明らかにしておく必要がある。表2に示されたHSの分類コードが台湾固有の分類コードである。

[5] 相手国をアジア研統一国コードへ変換し、同時に数量単位をアジア研統一数量単位へと変換する。データフォーマットをAID-XT基礎データの形式に揃える。

[6] 年、輸出入区分、商品分類、相手国、数量単位ごとに取引額と数量を合計する。この段階で取引額を1,000NewTaiwan\$とするAID-XT基礎データが完成する。完成したデータは年ごとに保存される。

[7] 取引額を1,000NewTaiwan\$から1,000US\$へと変換する。このデータがAID-XT基礎データである。完成したデータは年ごとに保存される。

[8] 作成された1,000NewTaiwan\$と1,000US\$のそれぞれのAID-XT基礎データに対して整合性評価表を作成する。整合性が取れていると確認できなければその原因を検討し、可能な限り補正する。確認できたところでAID-XT基礎データは完成する。

このAID-XT基礎データを作成するのに先立って1,000NewTaiwan\$から1,000US\$への変換レートを追加しておかなければならない。変換レートは表1に示されているように1989年以降は*Statistical Yearbook of the Republic of China 2003*に記載されている総額から直接1,000US\$へと変換するように処理方法が変更されている。処理ステップの[2]を年ごとに実行するのはデータ処理用のプログラムであるが、ステップの[3]は手作業でおこなう。[5]から[8]まではデータ処理プログラムで実行できる。

台湾貿易データの数量は11桁レベルの商品分類コードCCCに対して割振られているためそのコードをもとにしたときはそれほど大きくないが、それをHSの6桁レベル分類コードへ集計した数量は相対的に大きくなることに注意すべきである。結果としてUN貿易統計やOECD貿易統計の他の報告国の数量単位の使用頻度と比較してもmetric tonよりもkgの表示が極端に多くなっている。UN貿易統計で使用されている商品分類に対する数量単位に一致させる等、他の報告国との間で横並びに数量単位を調整することは今後の検討課題として残されている。

また、商品総額における数量については詳細分類コードの数量の合計と比較しても大幅に値が異なっているため、意味がない数字と見なして数量単位を空白にして数量を0とする。AID-XT基礎データは原則として取引額と数量がともに0となるデータは削除している。この原則に従うとある相手国について詳細分類コードの取引額が0で数量が0ではないものがあり、

表3 アジア経済研究所作成による台湾の AID-XT 基礎データの整合性 (1,000 New Taiwan \$)

y	x_{TW}	e	d_1	d_2	d_3	d_4	d_5	d_6	$e_c + e_{c,p}$	$e_p + e_{c,p}$
(import)										
1989	1385272906	0	0.0000	0	0	0	0	4761	0	0.000
1990	1471803505	0	0.0000	0	0	0	0	4720	0	0.000
1991	1690772977	0	0.0000	0	0	0	0	4735	0	0.000
1992	1816294219	0	0.0000	0	0	0	0	4772	0	0.000
1993	2034746749	0	0.0000	0	0	0	0	4764	0	0.000
1994	2262650459	0	0.0000	0	0	0	0	4753	0	0.000
1995	2742850423	0	0.0000	0	0	0	1	4764	0	0.000
1996	2815119319	0	0.0000	0	0	0	1	4776	0	0.000
1997	3276094156	0	0.0000	0	0	0	0	5073	0	0.000
1998	3503568890	0	0.0000	0	0	0	0	4869	0	0.000
1999	3576415487	0	0.0000	0	0	0	0	4867	0	0.000
2000	4368695887	0	0.0000	0	0	0	0	4858	0	0.000
2001	3619429516	0	0.0000	0	0	0	0	4836	0	0.000
2002	3893680914	0	0.0000	0	0	0	0	4857	0	0.000
(export)										
1989	1745104214	0	0.0000	0	0	0	0	4344	0	0.000
1990	1802781059	0	0.0000	0	0	0	0	4300	0	0.000
1991	2040784970	0	0.0000	0	0	0	0	4320	0	0.000
1992	2047961998	0	0.0000	0	0	0	0	4322	0	0.000
1993	2234350533	0	0.0000	0	0	0	0	4306	0	0.000
1994	2456010977	0	0.0000	0	0	0	1	4451	0	0.000
1995	2949579951	0	0.0000	0	0	0	0	4487	0	0.000
1996	3176624196	0	0.0000	0	0	0	0	4508	0	0.000
1997	3481685084	0	0.0000	0	0	0	0	4803	0	0.000
1998	3693268791	0	0.0000	0	0	0	0	4625	0	0.000
1999	3917446187	0	0.0000	0	0	0	0	4564	0	0.000
2000	4616301200	0	0.0000	0	0	0	0	4488	0	0.000
2001	4137742390	0	0.0000	0	0	0	0	4491	0	0.000
2002	4507506099	0	0.0000	0	0	0	0	4562	0	0.000
(re-export)										
1989	3426229	0	0.0000	0	0	0	0	1262	0	0.000
1990	5636717	0	0.0000	0	0	0	0	1472	0	0.000
1991	10264021	0	0.0000	0	0	0	0	2014	0	0.000
1992	16391226	0	0.0000	0	0	0	0	2134	0	0.000
1993	25598008	0	0.0000	0	0	0	0	2321	0	0.000
1994	33021516	0	0.0000	0	0	0	0	2409	0	0.000
1995	44594943	0	0.0000	0	0	0	0	2474	0	0.000
1996	44909277	0	0.0000	0	0	0	0	2509	0	0.000
1997	59804837	0	0.0000	0	0	0	0	2671	0	0.000
1998	67204587	0	0.0000	0	0	0	0	2552	0	0.000
1999	68927649	0	0.0000	0	0	0	0	2594	0	0.000
2000	112983488	0	0.0000	0	0	0	0	2582	0	0.000
2001	116541960	0	0.0000	0	0	0	0	2679	0	0.000
2002	162897575	0	0.0000	0	0	0	0	2765	0	0.000
(re-import)										
1989	1303466	0	0.0000	0	0	0	0	868	0	0.000
1990	1885061	0	0.0000	0	0	0	0	1168	0	0.000
1991	7562089	0	0.0000	0	0	0	0	1698	0	0.000
1992	8723999	0	0.0000	0	0	0	0	1760	0	0.000
1993	8747683	0	0.0000	0	0	0	0	1804	0	0.000
1994	9244478	0	0.0000	0	0	0	0	1861	0	0.000
1995	12243880	0	0.0000	0	0	0	0	1975	0	0.000
1996	15189305	0	0.0000	0	0	0	0	2066	0	0.000
1997	15100176	0	0.0000	0	0	0	0	2223	0	0.000
1998	18939275	0	0.0000	0	0	0	0	2249	0	0.000
1999	16372742	0	0.0000	0	0	0	0	2223	0	0.000
2000	22530731	0	0.0000	0	0	0	0	2190	0	0.000
2001	24752437	0	0.0000	0	0	0	0	2172	0	0.000
2002	24733928	0	0.0000	0	0	0	0	2155	0	0.000

(出所) 野田容助「台湾の AID-XT 基礎データ作成と評価」の表4にもとづき著者作成。

(注) 紙面の都合により年は1989年から2002年までを表示している。

しかも商品総額が単位未満の0となると、この相手国の商品総額は取引額と数量がともに0となり、削除の対象となる。このことにより詳細分類コードがあるのにその商品合計がないという相手国ができてくる。この矛盾を避けるため、商品総額については NewTaiwan\$ から 1,000US\$ への変換過程において生じた取引額と数量がともに0となるデータは詳細分類コードが存在する場合には削除しないことにする。

1.3 AID-XT 基礎データの整合性評価

台湾のAID-XT基礎データにおいて取引金額が 1,000NewTaiwan\$ の整合性の評価は表3に示されている。アジア経済研究所では商品分類における整合性は商品総額を評価基準として、詳細分類コード *mdcc* の取引額をすべて合計すると商品総額に一致することを整合性の評価としている。貿易統計データにおける取引額にもとづく整合性の評価および補正については野田・深尾 [7] の「貿易マトリクス作成における整合性の評価—新および旧 AID-XT 基礎データにもとづいて—」に説明されている。これによると、完全に整合性の取れた貿易マトリクスの取引額表が存在するとき、相手国による誤差は $e_p(\bullet) + e_{c,p} = 0$ 、商品分類コードによる誤差は $e_c(\bullet) + e_{c,p} = 0$ 、総合誤差は $e = 0$ となる。

表3によれば1989年から2002年までは総合誤差がすべて0であり、サムチェックにもとづく誤差はないことがわかる。この表において注意すべきことは、1994年の輸出と1995年および1996年の輸入に *mdcc* としてそれぞれ1個の4桁レベル分類コードが含まれていることである。アジア経済研究所では商品分類の体系あるいは桁レベルの分類コードとの関係とは別に、実際に得られた貿易統計データにおいて取引金額がゼロでない商品分類コードに対して階層的に構成された分類コードの中で下位の階層の分

類コードを持たないものを詳細分類コード (most detail classification code: *mdcc*) と呼んでいる。1994年の輸出の4桁レベル分類コードは8501であり、1995年および1996年の輸入のそれは8409である。

相手国の評価については台湾の貿易統計書の「中華民國臺灣地區出口貿易統計月報」 (*Monthly Statistics of Exports, The Republic of China, Taiwan District*) と「中華民國臺灣地區進口貿易統計月報」 (*Monthly Statistics of Imports, The Republic of China, Taiwan District*) に記載されている相手国総額とAID-XT基礎データの相手国別商品総額を比較することで可能となる。取引金額が同一の1,000NewTaiwan\$表記であるためこの比較は直接的におこなえる。本章では比較結果を明示的に示していないが個別相手国について整数の下1桁まで一致していることを確認できる。

2. UN 貿易データ準拠への変換

本節では台湾貿易データを UN Comtrade Database 貿易データに準拠するために必要とされる両者の分類カテゴリーの対応関係の作成を示す。分類カテゴリーの中で商品分類についてはAID-XT基礎データ作成のときに処理されているので本節では特に相手国となる国コードと数量単位の対応関係が対象となる。

2.1 国コードの対応表

アジア経済研究所は以前に、①台湾統計データで使用されている台湾定義の国コードとアジア研統一国コードの対応表を年毎に作成していた。前節における1.2のAID-XT基礎データ作成で述べた [2] と [3] を一表にまとめ、さらに②UN Comtrade Database貿易データで使用されている国連定義の国コードとアジア研統一国コードと国

図1 国コード対応表の例

(1) 台湾定義の国コードとアジア統一国コードの関係

001	137590		*	...	
003	127760		*		

(2) UN 定義の国コードが付けられた (1) の関係

	Bhutan		001	137590	45064	*	...	
	Brunei		003	127760	45096			

(出所) 著者作成。

名の対応表、③台湾定義の国コードと国名を参照して、「台湾定義国コード」→「アジア統一国コード」→「国連定義国コード」、というような変換表を作成したい。一度表ができてしまえば実際は台湾から国連へ直接変換することになる。まず既に作成されていた台湾とアジア研のそれぞれの国コードの表(①)を見てみよう。例えば、1989年の一部は以下ようになる。

twn.c ide.c iso

001 137590 BT
 003 127760 BN
 004 137540 MM
 014 117840 HK

ここで、表題で表わされている *twn.c* は台湾定義の国コード、*ide.c* はアジア統一国コード、*iso* は国際標準化機構 (International Organization for Standardization: ISO) のアルファベット表示の2桁国コードをそれぞれ表わす。最初の行の「001 137590 BT」は国名のブータン (Bhutan) を表わしている。台湾の国コードと国連のそれとの変換の軸になるのは *ide.c* の137590である。1989年から2003年までの表をマージして図1の(1)に示されているような表を作成した。表作成の手順は、上記①のデータを読み取り、*twn.c* と *ide.c* を表に埋め込んでいく。①のファイルは1989年から2003年までであるのだが、たとえば①の1989年のファイルの最初のデータ

レコードである「001 137590 BT」から001と137590を表に編集する。1989年のファイルを参照しているので、星取り表の欄の89年のカラムに※印を入れる。これによって編集のあと国コードが何年に使用されたかがわかる。表のどこに編集されるかは図1の(1)を参照。①のファイルの1989年から2003年までについて編集作業を行って図1の(1)のような表を作成した。

次に②のファイルを読み取って国連の国コードと国名を表に埋め込んでいく。②のファイルの内容の一部を表示してみると次のようである。

<i>n_un.c</i>	<i>p_un.c</i>	<i>ide.c</i>	<i>un.c desc</i>
45408	45408	117180	KOREA D P R
45496	45496	117280	MONGOLIA
41392	41392	117310	JAPAN
45344	45344	117840	HONG KONG
45064	45064	137590	BHUTAN
45096	45096	127760	BRUNEI

ここで、*n_un.c*は新国連国コード、*p_un.c*は旧国連国コード、*un.c desc*は国連定義の国名を表わす。この表を参照して137590のコードの国名がBHUTANであり、対応する国連定義の国コードが45064であることがわかるのでその国名と国連コードを表に加える。こうして図1の(2)の表ができあがる。

これだけではやや不安なので、さらに③台湾定義の国コードと国名の対照表も参照して国連

が定義している国名と照合して、国コードの対応が問題ないかどうかを検証してみた。

iso twn.c desc

TD 313 1700 CHAD
 TF 810 8500 FRENCH SOUTHERN TERR
 TG 376 7250 TOGO
 TH 052 7200 THAILAND
 TJ 289 7120 TAJIKISTAN
 TK 798 8480 TOKELAU

desc は国名を表わす。この処理の目的は言うまでもなく台湾定義の国コードを国連定義の国コードに変換するための対応表の作成であるのだが、アジア研統一国コードを仲立ちにしている。

そのためその仲立ちの国コードに定義がなければ、その国のコードの変換はつながらないので目的が達せられない。③を参照して国名を表に入れるときに参照されなかった国コードと国名も表に追加しておいて、国連の定義の国名と照らし合わせて、アジア研統一国コードがないため、つながらないものがあるかどうかチェックしてみた。後で述べるが *Mayotte* がそうであった。

また国連定義の国名、アジア経済研究所定義の国名、台湾定義の国名を三種類並べてみて不自然なところがないか検査して、作業ミスの防止に努めた。

さて出来上がりをみってみる。台湾が定義した国コードで、複数のそれらがアジア研統一国コードや国連で定義されているコードに変換すると同じものになる場合がある。それらはグループ分けして表示した。これは国連、アジア研統一国コードでそれぞれ分類に差があるからである。おおむね台湾の分類のほう細かいように見受けられた。表の一部を見てみよう。表をすべて表示すると横長で見づらいので、国連定義の国名と三種類のコードのみを抜き出してみた。

	<i>un.c des</i>	<i>twn.c</i>	<i>ide.c</i>	<i>un.c</i>
1	Malaysia	033	127650	45458
1	Malaysia	034	127650	45458
1	Malaysia	035	127650	45458
2	India	017	137580	45356

2 India 018 137580 45356
 2 India 019 137580 45356

表の最初の部分であるが、この例では台湾定義のコードの 033、034、035 はアジア研統一国コードでは 127650 が割り当てられており、その結果それに対応する国連のコードも 45458 になっている。③を調べても *Malaysia* の 033 しかリストになく 034 と 035 は国名が③では確認できなかった。これらについては別途アジア経済研究所が入手した台湾貿易データに関する説明資料を参照し、*Malaysia: Sabah* は 034、*Malaysia: Sarawak* は 035、であることがわかった。②を調べてみると国連の国コードではそれぞれ 45461、45457 が割り当てられている。そこでこれに従い表を修正した。しかし *India* (インド) についてはコードの 018 と 019 は *West coast* と *others* であるが、国連の統計ではそのような区別はないため足し合わせて *India* に分類した。*France* (仏) も *France* と *France:others* があるが両方とも *France* に分類した。*Spain*、*USA*、*Canada*、*Mexico* も同様の扱いとした。

Mayotte は台湾統計では国コードが 353 であるが、台湾の国コードとアジア研統一国コードの対照表には出てこなかった。②の表を仔細に調べたところ国連統計では 16174 (*Comoros (Mayotte)*) が割り当てられていることがわかった。これについては直接コードを表に追加して紐付けをした。最後のできあがりの付表 1 では *Mayotte* の行のアジア経済研究所定義の国コードが空白になっているのはそのためである。*Greenland* は台湾からアジア経済研究所へ、アジア経済研究所から国連へのコードの対応はそれぞれ

twn.c ide.c un.c
 410 → 218850
 398841 → 22304

となっていたためリンクできなかったため表に直接 410 → 22304 の対応をセットした。同様な例としては、

<i>un.c des</i>	<i>twn.c</i>	<i>un.c</i>
OTHER ASIA	098	→ 45490
OTHER EUROPE	298	→ 57568
VENEZUELA	690	→ 33862
NAURU	729	→ 72520

が存在する。Belgium Luxembourg は国連のコードで Luxembourg が個別国として定義されているのでそれを生かして、206 →53058、237 →53442、とした。Monaco はアジ研統一国コードでは France に分類されている。そのままリンクさせれば France に編入されてしまう。Monaco は独立国であるのでそれは不自然であると思われたので、別途コードをおこして 241 →90915、とした。Liechtenstein (リヒテンシュタイン) も同様の理由で Switzerland とは区別して、236→90916 とした。

わかるところは台湾定義の国コードから国連定義のそれへの変換のデータを追加してプログラムを動かして作業を行ってみたが、それでも国コード変換表にない国コードが出現した。502、528、820、710 である。これらは個々の数値をオリジナルのファイルから拾い出し、印刷物の統計書と照合してみた。それぞれ BR T/C C A (英領中央アメリカ)、Neth. T/C in C.A.(オランダ領中央アメリカ)、Origin Unidentifiable (原産国不明)、British Pacific IS (英領太平洋諸島) であることが判明した。これらについて次のような変換をすることにした。

<i>twn.c</i>	<i>un.c</i>
502	→ 90913
528	→ 90912
710	→ 90914
820	→ 90898

また、台湾国コード、アジ研統一国コード、国連国コードにおける3者の対応関係において、台湾国コードとアジ研統一国コードが存在するにもかかわらず、国連国コードが存在しないものについては新たに国連コードを作成して、3者を連結可能にしている。台湾国コードと国連国コードは存在するのにアジ研統一国コードが

存在しないときはそのままの状態を維持しているためできあがりの表の該当欄は空白となる。国コードの変換表は、どのように変換したかを示す重要なものであるから、コードの対応部分をすべて章末の付表1に紹介しておく。

2.2 数量単位

数量単位も台湾統計のそれと国連統計のそれは異なるから、コード変換をしなければならない。内容を見てみよう。

twn.u ide.u

KGM KO
TNE WO

ここで、*twn.u* は台湾定義の数量単位、*ide.u* はアジア経済研究所の数量単位である。1989年以降の台湾貿易統計書は巻頭に数量単位の一覧があるので、それらにリストされていれば調べやすい。そこにも載っていないものは、オリジナルのデータの当該のレコードと統計書を比べてその記述のなかで数量、金額、相手国が一致するものを探し出して、単位の記述があるかどうかを調べた。判明したものを変換のための表に書き加えた。この表も重要なものなので内容を表4に紹介する。

表4は紙面の都合で台湾のコードと国連のコードと単位の説明の三つをならべてある。国連の貿易統計のサンプルをみていると数量単位は数字の一桁であるが、台湾のコードを対応させるためには足りなかった。そこでやむなくアルファベットも追加して新たなコードを割り当てた。Head (頭数) なんかがいわば個数の一種であるから PC や UNIT と同じコードの5を当ててもよかったのかもしれない。数量単位の定義はあまり統一が取れておらず、PC にピリオドがついていたり、ついていなかったりしていた。ダースなどは四つもコードがある。SQ.M とか CUB.M あるいはKVA などは一目瞭然であるが、

表4 台湾貿易データで使用されている数量単位と国連の数量単位の対応表

<i>twu.u</i>	<i>un.u</i>	<i>desc</i>	<i>twu.u</i>	<i>un.u</i>	<i>desc</i>	<i>twu.u</i>	<i>un.u</i>	<i>desc</i>
M/T	9	Weight in Metric Ton	M.	4	metres			(Unit)
KG	8	Weight in KiloGrAms	M	4	metres	DOS	5	Dose 一服
KG.	8	Kilograms	SQ. M.	2	Area in square metres	BOT	5	BOTLE
MT	9	Weight in Metric Ton			立法メートル	DOZ	D	DOZEN ダース
:::	1	No Quantity (all quantities zero, standard if 0-3 digits)	MTK	2	Area in square metres	DOZ.	D	DOZEN ダース
			M2	2	square metres	DZN	D	DOZEN ダース
			M2. M.	2	square metres、	DEN	D	DOZEN ダース
KGM	8	Weight in Kilograms	CUB. M.	C	Cubic Metre、	HED	H	Head 頭数
TNE	9	Weight in Metric Ton			平方メートル	HD	H	Head
	1	空白も数量なしとする	MTQ	C	Cubic Metre	SET	S	SET セット
TND	9	Dry Metric Ton	M3	C	Cubic Metre	100'S	Q	100 Sticks
KVA	3	Electrical energy in thousands of Kilowatt-hours	M3B. M.	C	Cubic Metre	100, S	Q	100 Sticks
			PC	5	Number of items	100' S	Q	100 Sticks
NPR	6	Number of pairs	PCE	5	Number of items (Piece)	HSK	Q	100 Sticks
PR	6	Number of pairs				CS	Q	100 Sticks
PR.	6	Number of pairs	UNIT	5	Number of items (Unit)	KS	R	1000 Sticks
LTR	7	Volume in litres				KSK	R	1000 Sticks
L.	7	litre	UNT	5	Number of items (Unit)	ML.	R	1000 Sticks 千支
L	7	litre リットル				M1	R	出現せず
MTR	4	Length in metres	UNI	5	UNIT	M1.	R	MLと同じ
			NIU	5	Number of items	DMT	X	不明 92年

(出所) 著者作成。

(注) *un.u* は国連定義の数量単位 (一桁)

M2 が平方メートルで M3 が立法メートルであるから、類推して M1 がメートルであるとみなすのは大間違いで、これは 1000 支であった。変わったところでは 100'S の百支 (one hundred sticks)、それからいま述べた千支である。この単位が使われている品目は紙巻タバコであった。DOS もめずらしい。これは商品分類は薬であった。

台湾貿易統計書と照らし合わせても判明しなかったのは DMT である。1992 年の統計に出現し、品目は苛性ソーダ (Sodium Hydroxide) である。オリジナルのデータを抜き出してみると主数量は MT (トン) で計上され、従数量として計上されている。DMT の数量は同じ数値のときもあれば、MT の半分くらいの値だったり、ゼロの場合もある。同じ品目で 1993 年の輸入は従数量が TND で計上されているから、想

像するに乾燥重量 (ドライトン) の可能性が高い。

台湾貿易データと UN 貿易データは共に数量は主数量 (日本貿易統計を基準としたときの第 2 数量) と従数量 (同じく第 1 数量) から構成されている。主数量はだいたい KG で計上されているが、MT などのトンの数量が計上されているものもある。国連の貿易統計は主数量が無条件に KG ということになっていて、数量単位はレコードのなかに記載はない。もし MT 表示の数量をそのまま移せば、KG と区別がつかないから、別途 MT の数量で表示されている品目コードの一覧を用意せねば、利用者が困ることになる。しかしそれもあまりに不便であろうから、悩んだすえに 1000 倍して KG に換算することとした。

一部に主数量も従数量も計上されていない

ものがある。1989年輸入で989900という品目
 がその例であるが、Special transactions not classed
 by kind、特殊取扱品とでもいうべきものであつ
 た。1991年輸入の710310 (Precious Stones) は
 統計書にはKGの数値がある。同様に92年輸出
 にある950699 (その他のスポーツ用品) につい
 ても統計書にはKGの表示があつた。なぜ電子
 ファイルのほうには数量がないのか、その理由
 は不明である。2003年には電力の輸出がわずか
 ながら実績としてあり、さすがに重量の数量(主
 数量)はなくKVAが従数量に計上されていた。
 ちなみに輸出先は中国 (mainland) である。1989
 年輸出、401210 (Retreaded tyres of rubber for other
 vehicles、再生タイヤ) は主数量がPCであり従
 数量が空欄であつたので、数値を従数量に移し、
 主数量を空欄とした。しかし統計書のほうはPC
 とKGが同一品目内で入り混じっているのは不
 可解である。

3. 変換のためのデータ処理

本節は台湾貿易データから UN Comtrade
 Database 貿易データへの変換のためのデータ処
 理手順を示す。その内訳は台湾貿易データの書
 式、国連貿易統計の形式、HS 各改訂版の確定、
 プログラミングの問題、を含む。

3.1 台湾貿易データの書式

台湾の貿易データの書式については、1977年
 から1980年まで、1981年から1988年まで、1989
 年以降の3パターンある。書式は次のとおりで
 ある。

(1) 1977年から1980年までの書式

最初の3桁の数字はそれぞれの項目の先頭の
 位置を表わし、～に続く3桁の数字はその最後
 の位置を示す。

001~008 CCCN

009~013 台湾独自の品目分類?
 014~017 国コード (4桁のもの)
 018~025 不明
 026 D2: 後述のD1と組み合わせて輸出入
 の区分を表わす。
 027~028 不明
 029~032 SITC
 033~038 従数量単位
 039~044 主数量単位 (重量)
 045~047 国コード (3桁): 国連形式への変換は
 この国コードを用いる。
 048~113 current の数値
 114~122 従数量 (輸出)
 123~131 主数量 (輸出)
 132~141 金額 (FOB)
 142~150 従数量 (輸入)
 151~159 主数量 (輸入)
 160~169 金額 (輸入、CIF)
 170~179 金額 (輸入、DUTY)
 180~181 不明
 182 D1

D1とD2の組み合わせによって次のような意味
 を表す。

D2 D1

0	1	輸出
0	2	輸入
1	1	再輸入
1	2	再輸出

再輸入の数値は輸出の欄に記載されている。台
 湾の貿易統計は前述したとおり、ネットの輸出
 はグロスの輸出から再輸入を引いたものになつ
 ており、再輸入はマイナスの輸出であると考え
 ているからかもしれない。再輸出の数値も同様
 に輸入の欄を参照する。

(2) 1981年から1988年までの書式

001~008 CCCN
 009~015 CCC (台湾の品目コード。CCCNと組み
 合わせて細分されている?)
 016~019 国コード (4桁)
 020~023 SITC
 024~032 不明
 033 D2
 034~035 CCCNのSection
 036~041 従数量単位
 042~047 主数量単位
 048~050 国コード (3桁)
 051~120 current
 121~130 従数量 (輸出)
 131~140 主数量 (輸出)
 141~150 金額 (FOB)
 151~160 従数量 (輸入)
 161~170 主数量 (輸入)

表5 UN Comtrade Database貿易データのcsv形式

Reporter Code, Trade Flow Code, Classification, Commodity Code, Partner Code, Year, Value, Netweight (KG), Quantity Unit Code, Supplementary Quantity

84, 1, "H1", "01", 0, 1998, 185479, , 0,
 84, 2, "H1", "01", 0, 1998, 137544, , 0,
 84, 2, "H1", "0102", 0, 1998, 137246, 0, 5, 450
 84, 2, "H1", "010290", 0, 1998, 137246, 0, 5, 450
 84, 1, "H1", "0105", 0, 1998, 184801, 0, 5, 280296
 84, 1, "H1", "010511", 0, 1998, 183811, 0, 5, 279796
 84, 1, "H1", "010519", 0, 1998, 990, 0, 5, 500
 84, 1, "H1", "0106", 0, 1998, 678, 0, 5, 1
 :
 84, 2, "H1", "9999AA", 0, 1998, 2790632, 516615, 8, 516615
 84, 3, "H1", "9999AA", 0, 1998, 2786632, 511178, 8, 511178
 84, 1, "H1", "TOTAL", 0, 1998, 292261248, , 0,
 84, 2, "H1", "TOTAL", 0, 1998, 162217536, , 0,
 84, 3, "H1", "TOTAL", 0, 1998, 15792413, , 0,
 84, 2, "H1", "20", 28, 1998, 192863, , 0,
 84, 2, "H1", "2009", 28, 1998, 192863, 111751, 7, 111751
 84, 2, "H1", "200911", 28, 1998, 48360, 28187, 7, 28187
 84, 2, "H1", "200919", 28, 1998, 7533, 4375, 7, 4375

(出所) UN Comtrade Database にもとづき著者作成。

(注) 報告国の国コードが 084 であるベリーズ (Belize) のデータである。欠損値は, として表わされている。商品分類と商品分類コードは文字属性が与えられており”で囲まれている。それ以外は数字属性である。

171~180 金額 (輸入、CIF)
 181~190 金額 (輸入、DUTY)
 191~192 不明
 193 D1
 (3) 1989 年以降の書式
 001~006 HS 007~011 CCC
 012~015 国コード (4 桁、または 2 文字)
 016~027 不明
 028 D : 1 は輸出、2 は輸入、4 は再輸出、
 5 は再輸入
 09~030 CCC の Section
 031~033 従数量単位
 034~036 主数量単位
 037~039 国コード (3 桁)
 040~080 Current ?
 081~090 従数量
 091~100 主数量
 101~110 金額 (CIF または FOB)
 111~119 輸入の DUTY
 120 不明
 ただし 1981 年のファイルは D2 の位置が 029 カラムになっている。

アジア経済研究所が入手している on-line 検索により得られる UN Comtrade Database 貿易データは表 5 に示されているように csv 形式に変換されている。ファイルの最初に文字による分類カテゴリーおよび統計値に関する説明記述があり、貿易データに含まれる項目にどのようなものがあるかわかるようになっている。すなわち、Reporter Code は報告国、Trade Flow Code は輸出入区分であり、1 は輸入、2 は輸出と再輸出の和、3 は再輸出、Classification は商品分類コードの体系、Commodity Code は商品分類コード、Partner Code は国コード (貿易相手国)、Year は年 (西暦)、Value は金額、Netweight (kg) は主数量 (KG)、Quantity Unit Code は従数量単位、Supplementary Quantity は従数量である。

表 5 は報告国の国コードが 084 であるベリーズ (Belize) のデータであって、商品分類が

3.2 国連貿易統計の形式

表6 台湾で使用されているHS分類コードとHS-88、HS-96、HS-02の対応関係

y	Total	H(88・96・02)	H(88)	H(88・96)	H(88・02)	H(96)	H(96・02)	H(02)	H(tw)	C
1989	4881	4415	263	194	1	0	0	0	8	88
1990	4851	4391	261	192	1	0	0	0	6	88
1991	4867	4404	259	196	1	0	0	0	7	88
1992	4886	4424	260	195	1	0	0	0	6	88
1993	4880	4418	261	198	1	0	0	0	2	88
1994	4883	4422	262	195	1	0	0	0	3	88
1995	4875	4418	262	189	1	0	1	0	4	88
1996	4893	4429	260	199	1	0	1	0	3	88
1997	5203	4423	258	193	1	20	306	0	2	mixed(1)
1998	4994	4422	43	193	1	20	313	0	2	96
1999	4984	4424	29	195	0	20	314	0	2	96
2000	4956	4414	19	195	0	19	307	0	2	96
2001	4951	4412	18	186	0	19	314	0	2	96
2002	4953	4413	17	189	0	19	313	0	2	96
2003	5163	4439	16	191	1	20	313	181	2	mixed(2)
2004	5164	4435	16	88	1	13	312	297	2	02
2005	5067	4431	18	14	1	3	312	286	2	02

(出所) NewTaiwan\$表示のAID-XT基礎データとHS1988、HS1996、HS2002のそれぞれの商品分類表にもとづき著者作成。

(注) Cは商品分類を表わし、HS1988、HS1996、HS2002をそれぞれ88,96,02としている。また、mixed(1)はHS1988とHS1996の過渡期、mixed(2)はHS1996とHS2002の過渡期をそれぞれ表わしている。

HS1996年度版の一部である。なお、HS1996年度版はH1で表わされる。商品分類の上位桁レベル分類コードでは各種数量単位を持つ商品が含まれているため数量に意味がなくなり、数量単位に意味を持たせていない0と数量の空白で表わされている。

3.3 HS各改訂版の確定

台湾貿易データの基礎となっている台湾貿易統計書の月報にはHSの改訂版の版番号は明示的に説明されていないため、HSについては改訂版を特定化する必要がある。アジア経済研究所ではHS1988,HS1996,HS2002の商品分類コードをそれぞれの台湾の商品分類コードと対応させ一致した個数の多いほうを改訂版として採用している。HS1988からHS2002までのそれぞれに含まれる商品分類コードの関係は8つに分

けることができる。

- (1) HS1988、HS1996、HS2002に対して共通して所属する商品分類コードをH(88・96・02)、
 - (2) HS1988にのみ所属するものをH(88)、
 - (3) HS1988とHS1996に対して共通して所属するものをH(88・96)、
 - (4) HS1988とHS2002に対して共通して所属するものをH(88・02)、
 - (5) HS1996にのみ所属するものをH(96)、
 - (6) HS1996とHS2002に対して共通して所属するものをH(96・02)、
 - (7) HS2002に対してのみ所属するものをH(02)、
 - (8) HS1988、HS1996、HS2002に対していずれにも所属しない商品分類コードは台湾貿易データ固有のものであり、これをH(tw)とする。
- 台湾貿易データで使用されているHS分類コードとHS1988、HS1996、HS2002の各改訂版の

対応関係は表6に示されている。表6において示されている数字は左から年ごとの台湾の商品分類コードの個数、H(88・96・02)に一致した分類コードの個数、H(88)およびH(96)はそれぞれHS1988とHS1996に一致した個数である。以下同様である。H(*tw*)が上述した台湾固有の商品分類コードである。UNから入手したHS各改訂版とSITC-R3の対応表は6桁レベルのSub-headingしか含まれていないため台湾の*mdcc*である2個の4桁レベル分類コードの8501と8409とは一致するものが存在しない。このことは第1節の1.3で述べたように「1994年の輸出の4桁レベル分類コードは8501であり、1995年および1996年の輸入のそれは8409である」ことがわかる。一致しないHSの分類コードは1994年の輸出の8501、1995年および1996年の輸入の8409に存在するほか、商品総額は対応表には含まれていないため一致しない。これらもH(*tw*)に含めて個数が計算されている。

結果として、1989年から1996年まではほとんどの商品分類がH(88・96・02)、H(88)、H(88・02)に所属し、88の関わる場所へ集中しているため商品分類はHS1988と判断される。1998年から2002年まではほとんどがH(88・96・02)、H(88・96)、H(96・02)に所属し、96の関わる場所へ集中しているため商品分類はHS1996と判断される。H(88)にもわずかではあるが所属しているのは台湾貿易データにおいてHS1988の一部が変更なしに利用されていると考えられる。しかし、1997年は88に関わるH(88)とH(88・02)、96に関わるH(96)とH(96・02)は共に250近くを越えて所属しているため、分類変更の過渡期と判断される。表6ではこの時期のHSの改訂版をmixed(1)として表している。2004年から2005年についてはほとんどがH(88・96・02)、H(96・02)、H(02)に所属し、02に関わる場所へ集中しているため商品分類はHS2002と判断される。同じようにして、2003年はHS1996

とHS2002の過渡期と判断され、mixed(2)としている。

台湾貿易データにおいて、1998年から2002年までのHS1996において旧改訂版のHS1988が変更なしに利用されていたのと同じように、2004年から2005年についてはHS1988およびHS1996の旧改訂版の一部が変更なしに利用されていることに注意する必要がある。というのは、変換のための対応表作成にはこれらの旧改訂版の一部を含めて作成する必要があるからである。

3.4 変換の手順

3.1で述べた貿易データの書式を調べて、次のような手順で変換することにした。左側は台湾貿易データ、右側がUN Comtrade Database貿易データのそれぞれの分類カテゴリーである。報告国は、報告国コード→UN Comtrade Databaseの台湾国コード(490)を固定でセットする。輸出入区分は、D1とD2の組み合わせに従って、

D1=1 且つ D2=0 →2 (輸出)
 D1=2 且つ D2=0 →1 (輸入)
 D1=1 且つ D2=1 →4 (再輸入)
 D1=2 且つ D2=1 →3 (再輸出)

とする。国連の輸出入区分では輸出と言うのではなく、「輸出と再輸出の合計」に2が割り当てられているので、それを輸出に用いることにした。再輸入についてはUN Comtrade Database貿易データにおいて2006年の7月以降、新たに輸出入区分に新たに加わった定義であるが、台湾のそれとの定義の関係を明らかにする時間がなかったため、今後の課題として残されている。

商品分類は、CCCNコードを品目コードの欄に転記する。1988年まではCCCN分類で作成され、1989年からHSコードになっている。CCCNは遡ればBTNに起源を持つ分類であり、これが発展してHSになったものである。そこで「商品

分類の種類 (Classification) は1988年までは“CN”を入れ、1989年以降は“HS”をいれた。UN Comtrade Database貿易データではSITCの改訂第1版から改訂第3版に対して、それぞれS1からS3、HSについては1988年度版、1996年度版、2002年度版に対してそれぞれH0、H1、H2の記号が割り振られている。CNはCCCNから2文字を取って作った商品分類体系のコードである。正確には、表5で示したようにHSについても各改訂版を対応させなければならないが、1997年と2003年において商品分類の改訂時期の新旧商品分類が混在しているため本章ではHSと一本化している。HSの各改訂版間における改訂版の識別とコード化については今後の課題として残されている。

国コード3桁は付表1の対応表で変換して5桁の国連の国コードにして国コード (貿易相手国) にいれる。台湾貿易データの入力ファイル名をみて年 (西暦) をセット。金額 (CIF またはFOB) →金額、主数量→主数量 (KG)、従数量→従数量にそれぞれ転記する。従数量単位→「2. の対応表で変換」して従数量単位に変換する。もしも主数量がなくて、従数量がKGであったならば、それを主数量 (KG) に転記して、UN 貿易データへの準拠のほうの従数量はゼロとする。主数量がトン (MT) ならば数値を1000倍してKGに換算して、UN 貿易データへの準拠のほうの主数量に記入する。

作業繰り返しの結果を踏まえて、新たに必要な国コードを追加するなど、手作業で国コード対応表を修正し、付表1を最終版とした。また数量単位についても同様に、変換に必要な数量単位を補い、表4を作成した。変換されたデータは中間結果として出力し、さらに集計計算をおこなった。その必要性について少し述べてみたい。台湾の貿易統計の国分類、品目分類は、国連のそれより細かいことはすでに述べた。Malaysia が Sarawak、Sabah とか区別されている

ものはなるべくそれを尊重したが、USA に関して、PACIFIC や OTHER はひとつに分類したので、「はじめに」の (3) の出力のファイルには、国が同じものがありうるので集計をする必要がある。中間結果のファイルをそれぞれ輸出入区分、品目 (CCCN)、国コード、数量単位で sort し、この4つのコードが同じものは集計してまとめて最終的な結果を csv 形式ファイルで出力した。使用したプログラムは FORTRAN で作成した。

3.5 プログラミングの問題

前項で述べた集計をするとオーバフローする可能性がわずかながらもあるので、金額と従数量は桁をひとつ増やして11桁整数とした。KGの数量はトンを換算したものも含むので、余裕をみて15桁に増やした。私が用いた FORTRAN の整数タイプ変数では、いずれもオーバフローするので、途中の計算では実数タイプでおこない、最後に小数点をはずして整数の形で出力する工夫をした。

この種の作業ではよほど注意を払ってもなかなかプログラミングのミスを完全に防ぐことが難しい。そこで検証のプログラムは言語を変えて COBOL を使用した。出来上がりのファイルとオリジナルのデータを読み取り、国ごとに数量と金額とを集計し、また輸出入区分ごとの全体値を集計して、両方のデータを比較し、金額と数量の照合をおこない、すべて一致することを確認した。またオリジナルのファイルの合計金額は統計書と比較してみた。ほとんどは一致しているのだがなかにはかなりの食い違いがある。いくつかの例をあげると以下ようになる。大きく違うものでは例えば、1986年輸出 (CD) については総額が1,504,348,764であるのに対して、統計書の1991年版の86年値は1,506,602,707であり、1995年版の86年値は1,507,043,655で

ある。このケースでは印刷物でも年版によって一致しない。おそらく後の年の数値は改訂されたものであろう。日本の貿易統計も月表で公表されてはいないが、磁気テープ形でさかのぼって購入したときは、数値が新しくなっていることが見いだされたことがあった。2001年輸出（CD）は4,137,742,390であるのに、2002年版の2001年の値は4,137,742,391である。このような微妙な違いは計算誤差と想像されるが、断定はできない。貿易統計は計上ミスなどの修正をすることがあるので、年版によって値が違うことがあるは不思議ではない。しかし、これら数値の違いの原因究明は今後の課題である。

おわりに

データの変換の主要な部分は国コードの変換と数量単位の変換であった。コードの体系は国連と台湾とではもともと同じではないから、機械的に置き換えればよいというわけではない。変換先、この場合は国連準拠の形式だが、そこに対応するコードがなければ補わなければならない。実際数量単位は英文字まで動員して種類を増やす必要があった。であるから出来上がったものはあくまで国連準拠のものであり、国連統計そのものにはなりえない。また作り方もバリエーションがありうるし、今回の作品はそのなかの一つにすぎない。当然どういったポリシーで作り上げたのかは今回の作業の重要な要点になる。

アウトプットは別の利用者、研究者が分析などに使うわけだから、なるべく元の情報を失わないようにして、かつ利便性も考慮しなければいけない。両方満たせればよいのだが、相反すればどこかで折り合いをつけなければならない。まして今後どのような利用のされ方がなされるのか予測も難しい。たとえば国分類は Monaco、Liechtenstein（リヒテンシュタイン）は France、

Switzerland に含ませることはせず独立の分類にしたのだが、USA なども細かい分類を同様にそのまま残せばよかったかもしれない。しかしそれでは反って不便であろうと考えてまとめてしまった。また MT を KG に換算したことも、なるべく手を加えないのをよしとするならば、それに反している。今後の利用者の批判、意見を仰ぎたい。

（注1）アジア経済研究所の所蔵する台湾貿易統計は1988年版については海關總稅務司署統計處（Statistical Department Inspectorate General of Customs, Taipei, Republic of China）の発行する「中國進出口貿易統計年刊（臺灣區）」（*The Trade of China, Taiwan District*）と「中華民國臺灣地區進出口貿易統計月報」（*Monthly Statistics of Trade, The Republic of China, Taiwan District*）である。この年までは輸出入が1冊の統計書に載っている。商品分類は関税協力理事会（CCC）の分類である CCCN にもとづいて分類されている。1989年からは輸出と輸入がそれぞれ別冊になっており、当研究所ではこの年からは月報のみ所蔵している。月報の12月号には12月の統計値と同時に1月から12月までの累計である年の合計値が併記されているため、1992年までの12月号には「(年刊)」という記載がある。1989年の統計書は「中華民國臺灣地區出口貿易統計月報（年刊）」（*Monthly Statistics of Exports, The Republic of China, Taiwan District*）と「中華民國臺灣地區進口貿易統計月報（年刊）」（*Monthly Statistics of Imports, The Republic of China, Taiwan District*）である。この年以降の商品分類はHSである。1990年以降は発行する機関の名称へ変化があり財政部關稅總局統計室（Statistical Department Directorate General of Customs Ministry of Finance, The Republic of China）となった。統計書は「中華民國臺灣地區出口貿易統計月報」（*Monthly Statistics of Exports, The Republic of China, Taiwan District*）と「中華民國臺灣地區進口貿易統計月報」（*Monthly Statistics of Imports, The Republic of China, Taiwan District*）である。

（注2）貿易統計の磁気テープ化は商品分類コードごとに相手国世界（トータルデータと呼ぶ）と個々

の相手国（ディテールデータと呼ぶ）を別々に入力することから始まる。入力後は入力されたデータと貿易統計書との整合性を確かめるためにディテールデータを商品分類コードおよび相手国ごとに合計してトータルデータと突き合わせをおこない、数量と取引額が共に一致するかどうかを検査する。この整合性を確かめるための合計による検査の一連作業をサムチェック作業という。一致しない場合には再度貿易統計書とトータルデータおよびディテールデータの両データをチェックしてミスパンチの部分を修正しサムチェック作業を完了させる。

（注3）アジア経済研究所は台湾貿易データとして、1989年までは *Statistical Department Inspectorate General of Customs* から *Monthly Statistics of Trade, The Republic of China, Taiwan District* と同一の内容の磁気テープを入手していたが、1990年に作成機関の名称の変更にもない現在の機関から入手している。台湾貿易データ磁気テープデータは輸出、再輸出、輸入、再輸入の4つの輸出入区分に分かれているほか、台北、基隆、台中、高雄の各港ごとの集計値であることが特徴である。

（注4）商品分類は商品総額も含めて各層に分けられた桁レベルの商品分類コードから成り立っており、その桁レベルにおける商品分類コードのなかで最も細かい分類コードは概念的にはSITCの各改訂版では4桁レベルあるいは5桁レベルの分類コードから構成されており、基本項目（アイテム:item）またはBasic Heading といわれる。HSの各改訂版では6桁レベルの分類コードのSub-Headingから構成されている。アジア経済研究所では商品分類の体系あるいは桁レベルの分類コードとの関係とは別に、実際に得られた貿易統計データにおいて取引金額がゼロでない商品分類コードに対して階層的に構成された分類コードの中で下位の階層の分類コードを持たないものを詳細分類コード（most detail classification code: mdcc）と呼んでいる。体系的に定義されている基本項目と実際上の分類である詳細分類コードは必ずしも同一であるとは限らないので貿易統計データを利用のさいにはその両者を混同しないように注意が必要である。

（注5）月別変換レート単純平均により求めた変換レートは、1970は(40.00, 40.10*)と表される。()の中の前者は輸入、後者は輸出である。同じように、1971(40.00, 40.10*)、1972(40.00, 40.10)、1973(38.15,

38.34)、1974(37.90, 38.10)、1975(37.95, 38.05)、1976(37.95, 38.05)、1977(37.95, 38.05)、1978(35.97, 36.07)、1979(35.97, 36.07)、1980(35.95, 36.05)、1981(36.68, 36.78)、1982(38.90, 39.00)、1983(40.01, 40.11)、1984(39.57, 39.67)、1985(39.79, 39.89)、1986(37.92, 38.02)、1987(32.01, 32.11)、1988(28.55, 28.65)である。1970年と1971年の変換レートには*が付いているが、これはこの年度の統計が1,000NewTaiwan\$で表示されていることを表している。

【参考文献】

[1] 黒子正人「貿易統計データベースにおける国コードの利用」（野田容助編『商品分類の改訂に伴う貿易統計の変換』統計資料シリーズ No.83 アジア経済研究所 2001）

[2] —— 「国コード表」（野田容助編『商品分類の改訂に伴う貿易統計の変換』統計資料シリーズ No.83 アジア経済研究所 2001）

[3] —— 「貿易商品分類SITCからIO24部門分類への変換—変換エラーデータの処理—」（野田容助編『改訂版世界貿易マトリクス—国際産業連関表24部門分類にもとづいて—』統計資料シリーズ No.84 改訂版 アジア経済研究所 2003）

[4] 野田容助『東南アジア貿易統計データ処理（サムチェック作業）について』統計参考資料74-8 No.94 アジア経済研究所 1975）

[5] —— 「世界貿易マトリクス作成における整合性の評価と補正」（野田容助編『改訂版世界貿易マトリクス—国際産業連関表24部門分類にもとづいて—』統計資料シリーズ No.84 改訂版 アジア経済研究所 2003）

[6] —— 「台湾のAID-XT基礎データ作成と評価」（野田容助編『貿易指数の作成と応用—長期時系列貿易データの推計と分析に向けて—』調査研究報告書 開発研究センター2003-IV-20 アジア経済研究所 2004）

[7] 野田容助・深尾京司「貿易マトリクス作成における整合性の評価—相手国、数量単位および数量を考慮に入れて—」（野田容助編『貿易指数の作成と応用—長期時系列貿易データの推計と分析に向けて—』調査研究報告書 開発研究センター2003-IV-20 アジア経済研究所 2004）

- [8] *Trade of China (Taiwan District) 1963~1980*
[9] *Monthly Statistics of The Republic of China Taiwan District 1981~2006*

- [10] 「表2-2 アジ研国コード推移表」(木下宗七・野田容助編『世界貿易データシステムの整備と利用』統計資料シリーズ No.67 アジア経済研究所 1995)

付表1 国コード対応表

<i>un.c desc</i>	<i>tw.n.c</i>	<i>ide.c</i>	<i>un.c</i>	<i>un.c desc</i>	<i>tw.n.c</i>	<i>ide.c</i>	<i>un.c</i>
MAINLAND	090	117140	45156	DENMARK	213	213860	53208
NORTH KOREA	094	117180	45408	NORWAY	245	213880	55579
JAPAN	023	117310	41392	SWEDEN	257	213940	53752
HONGKONG	014	117840	45344	ICELAND	232	214140	55352
KOREA	026	118140	45410	GREENLAND	410	218850	22304
RYUKYU	046	118260	45647	BELGIUM	206	223641	53058
MACAU	032	118360	45446	LUXEMBOURG	237	223642	53442
VIET NAM	093	127240	45704	FRANCE	220	223660	53251
MALAYSIA	033	127650	45458	FRANCE: MEDITERRANEAN			
MALAYSIA: SABAH	034	127650	45461		221	223660	53251
MALAYSIA: SARAWAK				FRANCE: OTHERS	222	223660	53251
	035	127650	45457	MONACO	241	223660	90915
SINGAPORE	049	127680	45702	MAYOTTE	353		16174
BRUNEI	003	127760	45096	GERMANY	225	223680	53276
CAMBODIA	091	127980	45116	NETHERLANDS	244	223760	53528
LAOS	092	128040	45418	U KINGDOM	264	223840	53826
INDONESIA	020	128080	45360	LIECHTENSTN	236	224040	55757
PHILIPPINES	040	128160	45608	SWITZERLAND	258	224040	55757
THAILAND	052	128180	45764	IRELAND	216	224160	53372
MYANMAR	004	137540	45104	ANDORRA	202	224341	57020
SRI LANKA	008	137560	45144	GERMANY: EAST	224	224660	56278
MALDIVES	725	137561	45462	HOLY SEE	230	233740	53381
INDIA	017	137580	45356	ITALY	233	233740	53381
INDIA: WEST COAST				SAN MARINO	253	233740	53381
	018	137580	45356	MALTA	240	233741	57470
INDIA: OTHERS	019	137580	45356	PORTUGAL	249	233980	53620
BHUTAN	001	137590	45064	SPAIN	255	234340	53724
PAKISTAN	039	137641	45586	SPAIN: OTHERS	256	234340	53724
BANGLADESH	038	137642	45050	FINLAND	219	244360	53246
AFGHANISTAN	102	137940	45004	U. S. S. R.	291	244540	68810
NEPAL	037	138240	45524	POLAND	248	244740	56616
TURKEY	138	144280	44792	AUSTRIA	203	253960	53040
JORDAN	116	144980	44400	CZECHOSLOVAK	210	254640	56200
CYPRUS	108	145040	44196	CZECH REP.	209	254641	56203
ISRAEL	113	145140	41376	SLOVAKIA	211	254642	56703
IRAN	111	145160	44364	HUNGARY	231	254680	56348
IRAQ	112	145180	44368	GREECE	226	264260	53300
LEBANON	122	145340	44422	SLOVENIA	273	264381	59705
SYRIAN ARAB REP.				CROACIA	272	264382	59191
	135	145360	44760	BOSNIA	271	264384	59070
SOUTH YEMEN	136	154840	44720	MACEDONIA	268	264386	59807
BAHRAIN	105	154860	44048	YUGOSLAVIA	270	264387	59891
KUWAIT	119	154880	44414	ALBANIA	201	264560	56008
QATAR	131	154940	44634	BULGARIA	208	264580	56100
SAUDI ARABIA	134	155240	44682	ROMANIA	251	264760	56642
YEMEN	137	155260	44886	ARMENIA	286	274541	46051
YEMEN ARAB REPUBLIC				AZERBAIJAN	287	274542	46031
	141	155260	44886	GEORGIA	285	274543	46268
OMAN	128	155282	44512	KAZAKHSTAN	283	274544	46398
U ARAB E	139	155283	44784	KYRGYZSTAN	290	274545	46417
OTH M NEAR E	198	195590	44879	TAJIKISTAN	289	274546	46762
OTH ASIAN	098	198340	45490	TRUKMENISTAN	288	274547	46795

UZBEKISTAN	284	274548	46860	PARAGUAY	660	453260	33600
BELARUS	282	274549	58112	URUGUAY	680	453340	33858
ESTONIA	294	274550	58233	F T/C C A	515	493390	33473
LATVIA	295	274551	58428	VIRGIN ISLANDS (BRITISH)			
LITHUANIA	293	274552	58440		598	493390	33473
MOLDOVA REP.	292	274553	58498	OTH S A	698	493390	33473
RUSSIA	280	274554	58643	LIBYAN ARAB	342	515060	13434
UKRAINE	281	274555	58804	EGYPT	303	515380	13818
OTH EUROP	298	294791	57568	SUDAN	372	515440	13736
U. S. A.	431	392410	21841	W. SAHARA	371	516863	13732
USA: GULF-HOUSTON				ALGERIA	301	516940	13012
	432	392410	21841	MOROCCO	349	516960	13504
USA: PACIFIC	433	392410	21841	TUNISIA	377	516980	13788
USA: OTHERS	434	392410	21841	GAMBIA	329	525860	16270
PUERTO RICO	533	392410	21841	GHANA	330	525880	16288
CANADA	401	392510	21124	NIGERIA	355	525960	16566
CANADA: PACIFIC	402	392510	21124	SIERRA LEONE	368	526040	16694
CANADA: OTHERS	403	392510	21124	BENIN	318	526160	16204
COSTA RICA	505	412880	34188	COTE D'IVOIRE	334	526180	16384
EL SALVADOR	512	413040	34222	MALI	347	526240	16466
GUATEMALA	518	413060	34320	SENEGAL	367	526250	16686
HONDURAS	522	413140	34340	MAURITANIA	348	526260	16478
BELIZE	507	413141	36084	NIGER	354	526280	16562
MEXICO	421	413160	33484	BURKINA FASO	381	526340	16854
MEXICO: PACIFIC	422	413160	33484	GUINEA	331	526360	16324
MEXICO: OTHERS	423	413160	33484	TOGO	376	526660	16768
NICARAGUA	529	413180	34558	CAPE VERDE	310	526841	16132
PANAMA	532	413240	36590	GUINEA B	332	526842	16624
CUBA	591	422940	35192	EQ GUINEA	320	526861	14226
DOMINICAN REP.	509	422960	35214	LIBERIA	341	526880	16430
HAITI	521	423080	35332	C AFRICAN REP.	312	536440	14140
DOMINICA	508	423395	35212	CHAD	313	536460	14148
BR T/C C A	502	428611	90913	CONGO	314	536480	14178
JAMAICA	525	428640	35388	GABON	328	536540	14266
TRINIDAD T	536	428660	35780	CAMEROON	311	536560	14120
US T/P C A	539	428664	35850	CONGO, THE DEMOCRATIC REP.			
ANTIGUA B	506	428666	35028		315	536740	16180
NEVIS	527	428667	35658	ZAIRE	384	536740	16180
ST KITTS AND NEVIS				BURUNDI	309	536751	16108
	537	428667	35658	RWANDA	362	536752	16646
ST LUCIA	534	428668	35662	ANGOLA	302	536760	16024
ST VINCENT	535	428669	35670	SAO TOME P	365	537011	16678
ANGUILLA	504	428670	35660	ETHIOPIA	321	545460	16230
GRENADA	519	428671	35308	ERITREA	322	545461	16232
MONTSERRAT	526	428673	35500	SOMALIA	369	545480	16706
BAHAMAS	503	428680	35044	DJIBOUTI	324	545540	16262
BARBADOS	501	428740	35052	TANZANIA	375	545742	16834
COLOMBIA	621	432860	33170	KENYA	337	545760	16404
VENEZUELA	690	433340	33862	UGANDA	378	545780	16800
GUYANA	650	438780	36328	SEYCHELLES	307	547044	16690
SURINAM	670	438960	36740	SOUTH AFRICA	370	555680	11710
FRENCH GUIANA	640	439070	36254	MAURITIUS	350	555940	16480
BOLIVIA	610	442760	33068	ZAMBIA	385	555982	16894
CHILE	620	442840	33152	ZIMBABWE	363	555984	16716
ECUADOR	630	442980	33218	MALAWI	346	555986	16454
PERU	661	443280	33604	MADAGASCAR	345	556580	16450
ARGENTINA	600	452740	33032	COMOROS	316	556650	16174
BRAZIL	611	452780	33076	MOZAMBIQUE	351	556780	16508

BOTSWANA	306	557041	16072	TUVALU	747	669341	72798
LESOTHO	340	557042	16426	KIRIBATI	720	669350	72296
SWAZILAND	374	557043	16748	SOLOMON ISL.	740	669530	72090
NAMIBIA	352	557045	16516	NAURU	729	669540	72520
AUSTRALIA	700	615640	71036	TOKELAU	798	669580	72162
NEW ZEALAND	731	615660	71554	OTH AFRICAN	398	799999	90837
PAP N GUINEA	730	619261	72598	OTH N A	498	799999	90837
SAMOA	755	629350	72882	FRENCH SOUTHERN TERR			
FIJI	715	639240	72242		810	799999	90837
TONGA	745	639351	72776	UniDentifiAble	820		90898
VANUATU	734	639541	72548	528 Neth.T/C in C.A.			
TAHITI	744	649550	72258		528		90912
NEW CALEDONIA	733	649570	72540	British PACific IS			
A SAMOA	701	659575	72582		710		90914
US T/P P A	750	659575	72582				

(出所) 著者作成。

〈注〉 MAYOTTE(353)、UniDentifiAble(820)、Neth.T/C in C.A.(528)、British PACific IS(710)は対応するアジア統一コードがなかったため、その欄が空白となる。