

第6章

SITC-R1 に変換された貿易統計基礎データ に基づく輸出単価指数の作成

黒子正人

はじめに

2002 年度にアジア経済研究所が作成した貿易指数には、以下のような問題点や課題があった^(注1)。

- (1) 報告国別・IO24 分類別にまず基準年を定め、さらに報告国別の共通の基準年を定める方法を採用したため、報告国により基準年がばらばらである。
- (2) 不規則な指数の変動への対処として、採用品目決定期間と指数計算期間を一致させる、新しい採用品目決定方法を開発するなどが必要である。
- (3) SITC-R1、SITC-R2、SITC-R3 の改訂版 (Revision)ごとに異なる指数連となっており、それらを接続できていない。
- (4) 各国別に算出された指数の国際比較の方法の開発が必要である。
- (5) 各国政府が発行している貿易指数との相関が低い^(注2)。
- (6) IO24 部門分類別になっているため、一般機械、電気機械が同じ分類に入ってしまったっておりこれらを個別に分析できない。

今回の 2003 年度研究会ではこれらの問題点を少しでも解消することが目標となった。そこで、以下の作業をおこなった。

- (1) まず IO24 部門分類ではなく、SITC 各改訂版別に SITC 中分類 (上 2 桁) で集計して指数を作成した。
- (2) 次に今回野田によって SITC-R1 ベースに変換・結合された AID-XT 基礎データを入力元として指数を作成し、またその他の作成方法の細かな見直しを行った。

これらのうち（２）の作業について以下に実際の手順とその結果の一端を報告する。

１．入力元データと作成した指数

今回の指数を作成するにあたり、入力元として利用したデータは、野田によって SITC-R1 ベースに変換・結合された AID-XT 基礎データである。これは、SITC-R2 と SITC-R3 の商品分類をそれぞれ SITC-R1 に変換して、結果として AID-XT の全てのデータを SITC-R1 に時系列に接続したものである。入力元としてこのデータを利用することにより自動的に指数の全ての年次をひとつの時系列に接続することができた。貿易指数は表 1 の条件で作成した。

表 1 今回の指数作成の条件

項目	条件
報告国	参考文献 [1] で木下が分析した 4 報告国（日本、韓国、台湾、米国）に限る。
輸出入区分	輸出のみとする。
相手国	世界のみとする。
指数の種類	SITC-R1 の中分類（先頭 2 桁）別指数とそれをウェイト集計した総合指数。いずれも単価指数。ラスパイレス、パーシェ、フィッシャー各指数を算出する。
基準年	1965 年から始まる 5 年ごとの報告年（1965, 1970, 1975 ... , 2000 年）を基準年として指数を作成し、それらを結合して 1965 年を 100 とした指数連を作成する。

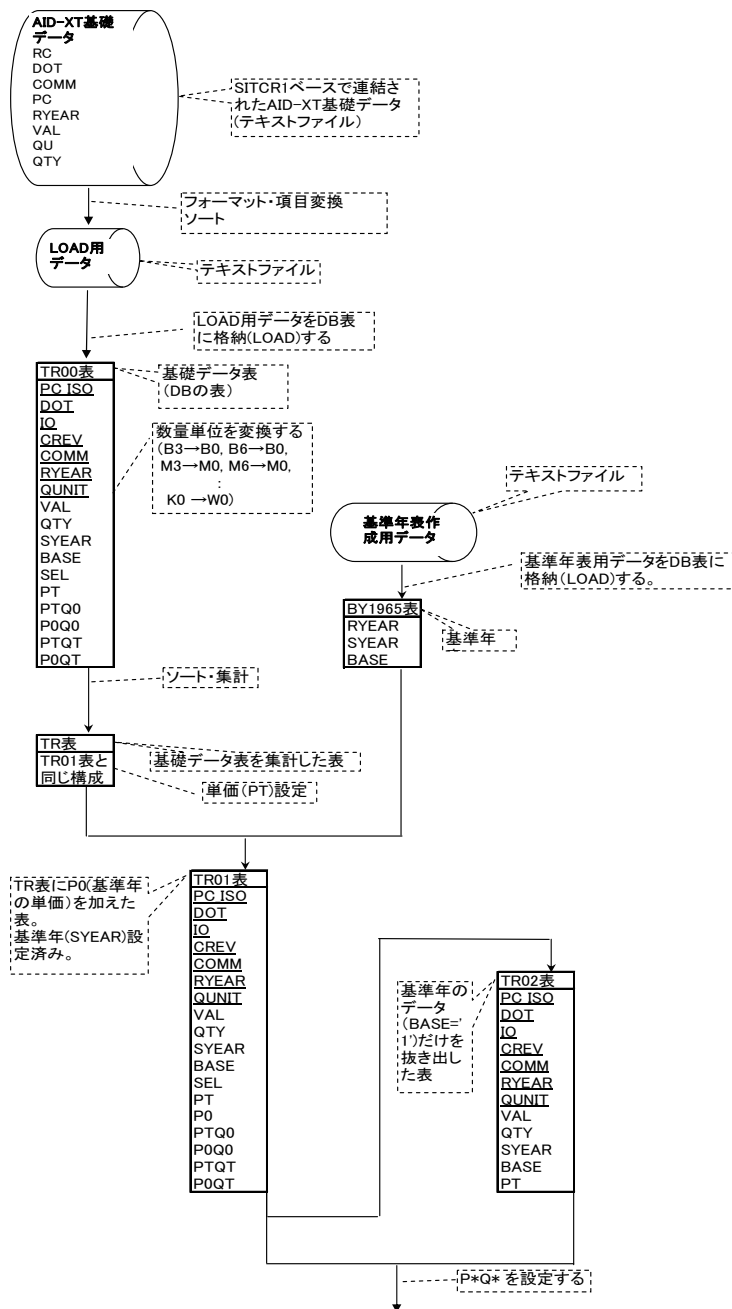
（出所）筆者作成

２．指数作成の手順

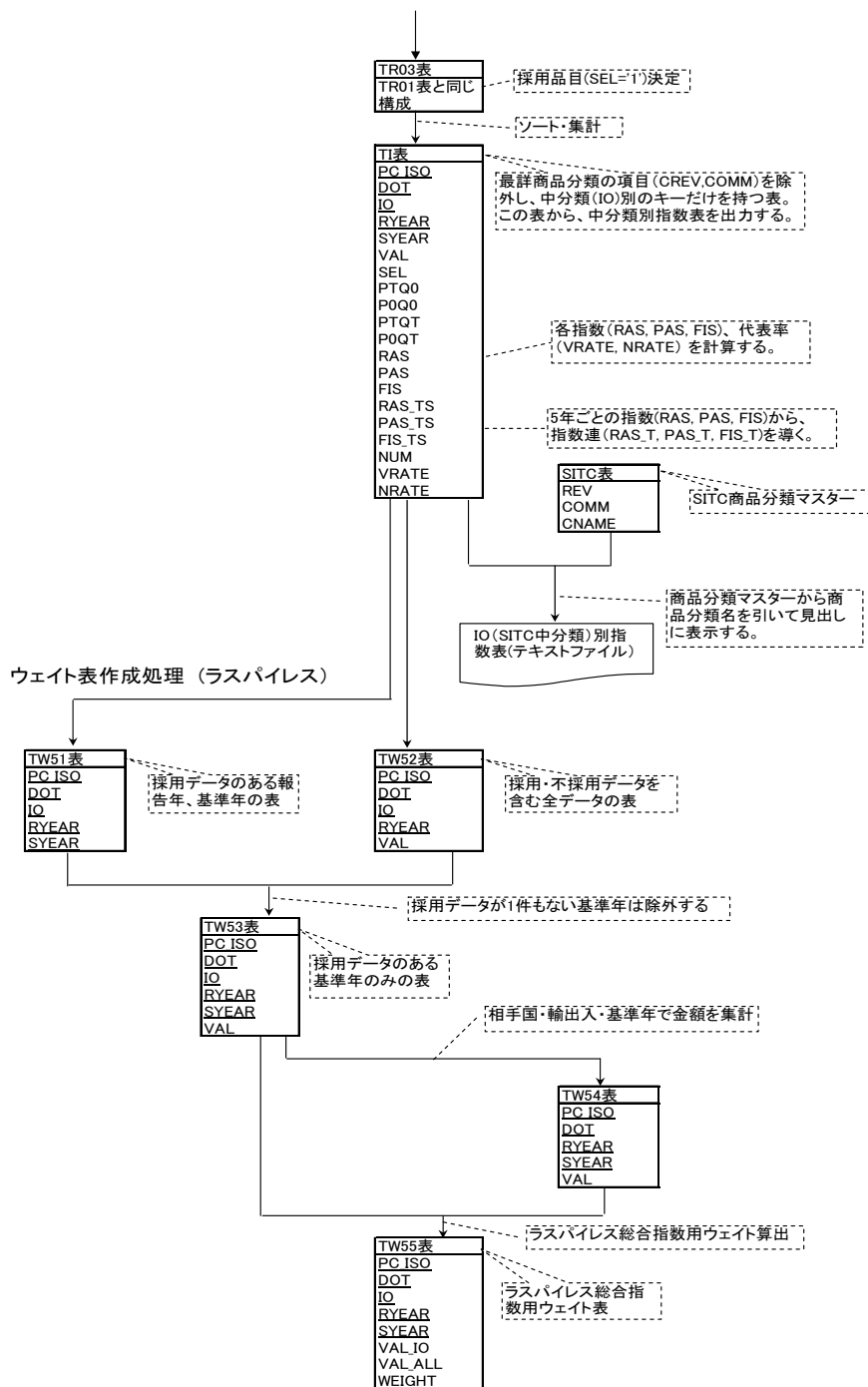
図 1 に、今回行った指数作成の実際の手順をまとめた。図 1 で使われている変数名は実際の処理で使われているものをそのまま掲載しているのでわかりにくい、主な変数名の意味は表 2 のとおりである。

図1 今回の貿易単価指数作成の手順

(SITC-R1 ベースで接続された AIDXT データに基づく)

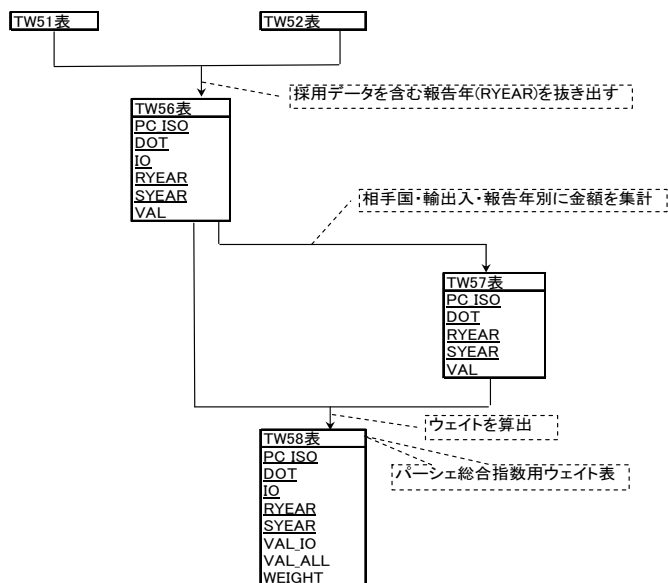


(図 1 の続き)



(図1の続き)

ウェイト表作成処理 (パーシェ)



総合指数表作成

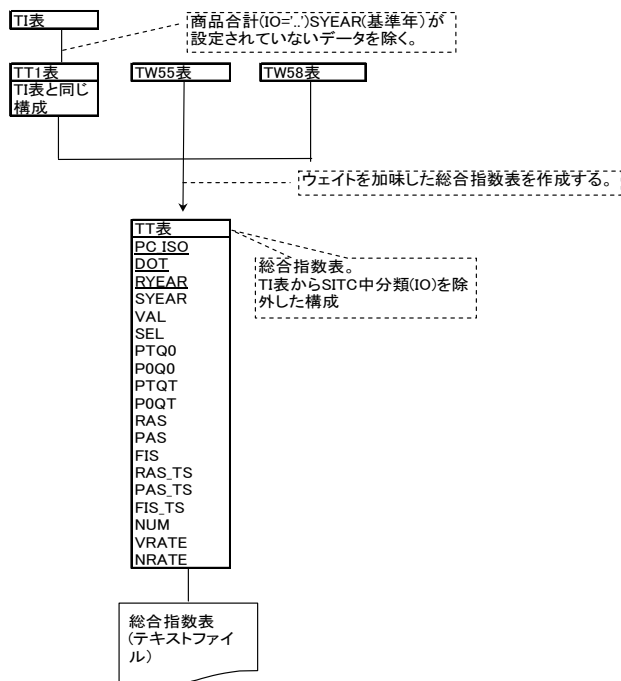


表 2 指数作成処理で使われている主な変数名

変数名	意味
PC_ISO	ISO のアルファベット 3 桁の国コードに準拠した国コードで相手国を表した国コード（例：日本の場合、"JPN"。世界は、"AAA"で表す）。今回は常に"AAA"である。
DOT	AID-XT 輸出入区分（1：輸入、2：輸出、3：再輸出、4：再輸入）。今回は常に2である。
IO	2002 年度の指数では、SITC 商品分類から変換された国際産業連関表 24 部門分類を示したが、今回はこの項目に SITC 商品分類の先頭 2 桁（中分類）を代入して、集計用項目として利用している。
CREV	SITC 商品分類の改訂版（Revision）。今回は常に 1 である。
COMM	SITC 商品分類（R2 と R3 は R1 に変換されている）。
RYEAR	報告年(1962 年から 2001 年まで)。
QUNIT	AID-XT 数量単位 ^(注3) 。 数量単位が異なると連続した指数が作れないため、以下のような 3 種類の方法で変換している。 （1）数量単位が B3 の場合、数量に 1,000 を掛けて、数量単位を B0 にする。M3, N3, P3, U3, V3, W3 の場合も同様。（2）数量単位が B6 の場合、数量に 1,000,000 を掛けて、数量単位を B0 にする。M6, P6, U6 の場合も同様。 （3）K0 は数量を 1,000 で除して、数量単位を W0 にする。
VAL	取引金額。単位は 1000US ドル。
QTY	取引数量。
SYEAR	5 年ごとの基準年。1965 年から始まる 5 年ごとの報告年（1965, 1970, 1975 ... , 2000 年）を基準年とした。ただし、報告年が基準年と同一年の場合、その一つ前の基準年を基準年とすることにより、指数をオーバーラップさせ指数連を作成できるようにした。例えば、報告年 1970 年のデータの基準年（SYEAR）は、1965 年とした ^(注4) 。
BASE	報告年が基準年と同一年の場合、'1'。
SEL	採用データの場合は、'1'。
PT	比較年 T の価格。VAL/QTY で求められる。
P0	基準年 0 の価格。(基準年 0 の VAL)/(基準年 0 の QTY)で求められる。

PTQ0	比較年 T の価格 * 基準年 0 の数量
P0Q0	基準年 0 の価格 * 基準年 0 の数量。基準年 0 の取引金額 (V0) に等しい。
PTQT	比較年 T の価格 * 比較年 T の数量。比較年 T の取引金額 (VAL) に等しい。
P0QT	基準年 0 の価格 * 比較年 T の数量
RAS	5 年ごとのラスパイレス単価指数
PAS	5 年ごとのパーシェ単価指数
FIS	5 年ごとのフィッシャー単価指数
RAS_TS	ラスパイレス単価指数 (1965 年を 100 とする)
PAS_TS	パーシェ単価指数 (1965 年を 100 とする)
FIS_TS	フィッシャー単価指数 (1965 年を 100 とする)
NUM	採用データの件数
VRATE	採用データの金額合計
NRATE	採用データの件数合計

(出所) 筆者作成

手順としては、まず AID-XT 基礎データをデータベースの表に読み込み、TR が先頭についた基礎データ表(TR**表) にまとめる。次に基礎データ表から、SITC 中分類 (IO) 別指数表 (TI 表) を作成する。次にこの TI 表からラスパイレス総合指数用ウェイト表 (TW55 表) とパーシェ総合指数用ウェイト表 (TW58 表) とを作成する。最後に TI 表と各ウェイト表から総合指数表 (TT 表) を作成する。以上が大まかな手順である。

前回の 2002 年度に作成した指数と今回作成した指数との算出方法の違いは主に以下の点である。

(1) 前回の指数算出ではアジア経済研究所が作成する国際産業連関表で使われている IO24 部門分類別に集計して貿易価格指数を算出した。今回は、この IO24 部門分類のかわりに SITC-R1 の中分類 (先頭 2 桁) 別に貿易価格指数を算出した。これにより、IO24 部門分類では同じ分類に入っているもの (例えば、一般機械、電気機械が「機械 (17)」に入っている) を別々に分析することが可能になった。

(2) 前回の指数算出では、5 年ごとの基準年を中心とした 5 年間を採用品目決定期間とし、実際に指数を作成するのは基準年から始まる 6 年間であつ

たため、採用品目決定期間と指数作成期間にずれがあり特異値を発生させる要因のひとつとなっていた。例えば 1995 年が基準年の場合、1993 年から 1997 年までの 5 年間で採用品目を決定し、1995 年から 2000 年までの 6 年間で指数を作成していた。今回は、採用品目決定期間と指数作成期間を一致させた。例えば同様に 1995 年が基準年の場合、1995 年から 2000 年までの 6 年間で採用品目決定期間であり同時に指数を作成する期間である。

(3) 基準年については、今回は報告国別・IO 分類別にまず基準年を定め、さらに報告国別の共通の基準年を定める方法を採用した。これは SITC 改訂版別・相手国別にすると場合によっては指数を作成する元データの件数がかなり少なくなることがあり、有効なデータを少しでも多くするための方法であった。今回は同じ SITC-R1 ベースの商品分類で連続した期間が取れ、また相手国も世界だけであるためこの方法を採用せず、最初の基準年を 1965 年に固定し、そこから数えて 5 年ごとの基準年 (1970、1975、・・・2000 年) を固定的に決定した。

(4) 指数を作成するための採用品目の選択基準を変更した。今回は、5 年の採用品目決定期間で 3 年以上取引実績があり、しかも 1 年あたりの単価の変化率が 5 分の 1 以上 5 倍以下の範囲内に収まっていることが基準であった。今回は、取引実績の基準は外し、指数を実際に算出する期間で 1 年あたりの単価の変化率が上記と同じ範囲内に収まっていることを基準とした。また、バスケット品目を採用品目から除外する基準を追加した。なお、バスケット品目とは、品目が既存の分類体系に分類できないような複数の商品が含まれる品目である。SITC 商品分類の末尾が 9 で終わるもの、および、先頭が 9 で始まるもの (SITC 大分類が 9 のもの。"Goods not classified by kind") をここではバスケット品目とみなす。

(5) 総合指数を算出する時のウェイトのかけ方を変更した。昨年度は総合指数を算出する時に $P*Q*$ にウェイトをかけてそれを集計して指数を求めている。例えばラスパイレス総合指数の場合、 $\sum (\text{IO 分類別の } P0QT \times \text{IO 分類別の } Weight) / \sum (\text{IO 分類別の } P0Q0 \times \text{IO 分類別の } Weight)$ として求めている。この方法だと $P*Q*$ 自身がウェイトを含む要素であるため、二重にウェイトをかける結果になっていたと思われる。今回は、 $\sum (\text{中分類別の指数} \times \text{中分類別の } Weight)$ とし、直接中分類別の指数にウェイトをかけて集計する

ように変更した。

3. 総合指数における特異値とその原因分析

図2が作成された総合指数のグラフである。報告国により様々な形状となっている。左側が総合指数、右側が代表率を示している。図3はラスパイレス、パーシェ、フィッシャーの各指数ごとのグラフである。

総合指数が特異な値を示しているのは、図2より一見して日本の1993年ラスパイレス指数の1475、米国の1994年ラスパイレス指数の912であることがわかる。まず日本の1993年の指数について原因を分析するために、指数を作っている元の表（TR表、TI表）にさかのぼって調べる必要がある。

まず総合指数表の次の集計レベルであるSITC中分類別指数表（TI表）からSITC中分類（IO）、ラスパイレス指数連（RAS_TS）を表示させる。SITC中分類別指数が1993年の総合指数の1496より大きいものが総合指数を押し上げていると考えられるが、これに該当するのは、SITC中分類12, 00, 59, 73, 04, 29, 06, 84, 69, 33である。これらの中分類の明細にさかのぼるためにラスパイレス指数を作る構成要素であるPTQ0、PT、P0、PT/P0を表示させる。このときPTP0の逆順に表示させると、明らかに1件だけ飛びぬけて大きなPTQ0の値をもったデータがある。このデータの商品分類(COMM)は、7323 ("Lorries and tracks") である。(表3)

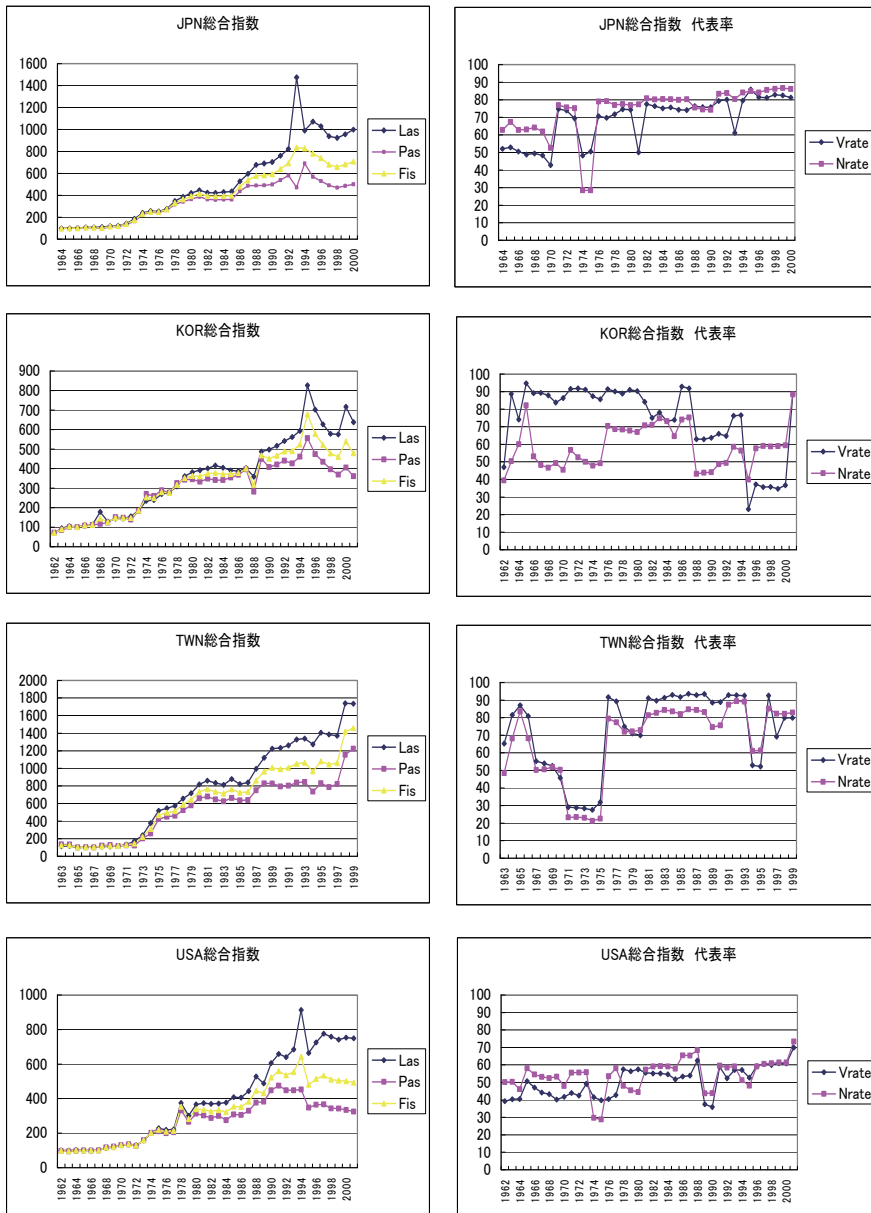
1990年から1995年までの商品分類7323に絞って見ると表4のようになり、明らかに1993年の単価PTが特異な変動を示していることがわかる。試みに商品分類7323を1993年の採用品目から除外して指数を算出したところ、総合指数の特異値が解消されるのが確認された。このような結果から1993年の商品分類7323の中に他の年とは単価の大きく異なる種類の商品が

表3 日本のラスパイレス総合指数の特異値の調査 (1)

COMM	SEL	VAL	PTQ0	PT	P0	PT/P0
7323	1	91, 223	142, 837, 293	41. 01753	8. 20413	4. 99961
73291	1	96	6, 850, 884	4. 00000	1. 32336	3. 02259
69891	1	385, 874	1, 181, 984	8. 79524	2. 59635	3. 38754
69421	1	882, 341	1, 139, 701	6. 12690	4. 02555	1. 52200
69524	1	1, 133, 631	994, 807	23. 34351	17. 57424	1. 32827
: (以下省略)						

(出所) AID-XT 基礎データにもとづき筆者作成。

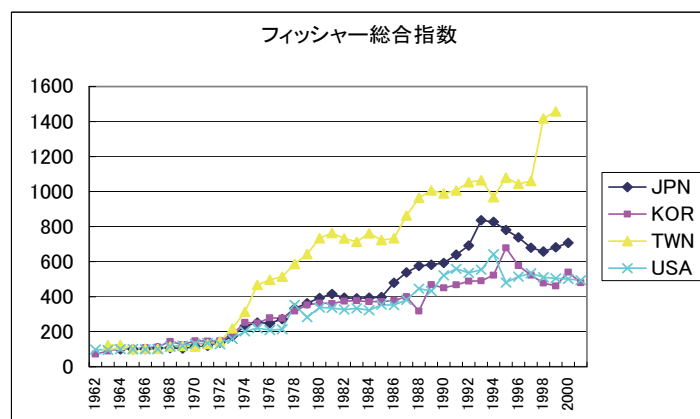
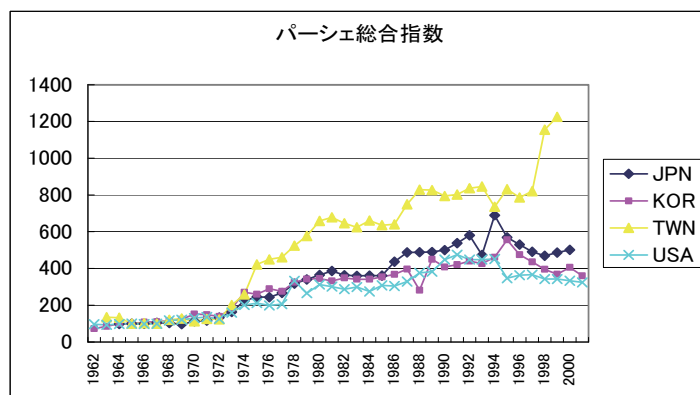
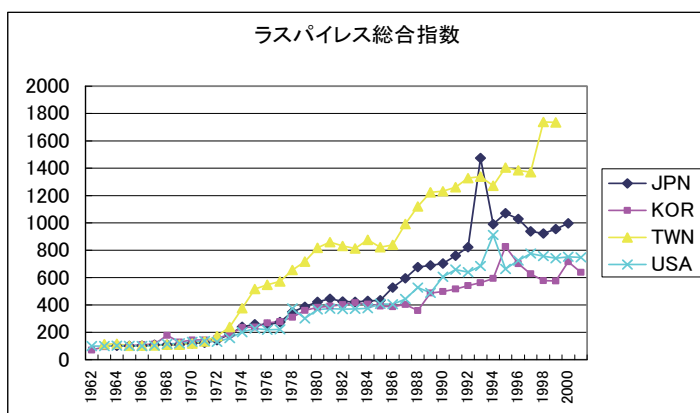
図2 総合指数と代表率



(出所) AID-XT 基礎データにもとづき筆者作成。

(注) 国・地域は JPN : 日本、KOR : 韓国、TWN : 台湾、USA : 米国、指数の算式は LAS : ラスパイレス指数、PAS : パーシェ指数、FIS : フィッシャー指数、代表率は Vrate : 金額代表率、Nrate : 件数代表率、をそれぞれ表す。

図3 総合指数



(出所) AID-XT 基礎データにもとづき筆者作成。

(注) 図2に同じ。

表 4 日本のラスパイレス総合指数の特異値の調査 (2)

COMM	SEL	RYEAR	QUNIT	VAL	PTQ0	PT	PO
7323	1	1990 NO		28,569,660	15,117,849	8.20413	4.31770
7323	1	1991 NO		30,439,397	30,912,248	8.87684	8.20413
7323	1	1992 NO		33,549,955	34,666,478	9.95491	8.20413
7323	.	1993 ..		32,390,792	.	.	.
7323	1	1993 NO		91,223	142,837,293	41.01753	8.20413
7323	1	1994 NO		31,841,353	42,144,675	12.10237	8.20413
7323	1	1995 NO		29,596,676	44,291,317	12.71881	8.20413

(出所) AID-XT 基礎データにもとづき筆者作成。

紛れ込んでいるのは明らかである。さて問題はこのような処理方法が妥当かどうかであるが、ここではこの問題には立ち入らない。

米国の 1994 年の特異値について日本の場合と同様に分析したが、原因は日本と同様に、1 件の特異データによるものであった。

次に、総合指数算出時の代表率が低いところについて、その原因を元データにさかのぼって分析した。まず、台湾の 1971 年から 1975 年にかけて 30% 前後の異常に低い金額代表率となっている。件数代表率はさらに低い。総合指数の特異値と同様に、まず、影響力をもつ中分類を調べるために以下の SQL 文を使う。

```
select io, ryear, sel, sum(val) val_sum from ti_twn where ryear = 1971 group by io,
ryear, sel order by val_sum desc
```

これにより、上位の中分類が金額の大きい順に、84, 72, 05, 89, 63, 65 と続くことがわかる。上位の中分類が 84 と 72 と 05 を合計すると、1971 年の総輸出額 (1,997,402 千 US ドル) の 4 割以上を占める。これらの中分類が総合指数に与える影響は大きいはずだが、いずれも採用区分 (SEL) は欠損値であり、採用されていない。まずこれらがなぜ採用されていないかをさらにさかのぼって、TR03 表から見てみることにする。以下の SQL 文で 1975 年以前の中分類が 84 のデータを金額の大きい順に表示させると、8414* という商品分類が上位に来ていることがわかる。そこで、この 8414* という商品分類に絞って年別の推移を以下の SQL 文で見る。

```
select comm, qunit, val, qty, ptq0, ryear, sel from tr03_twn where (ryear < 1976)
and io='84' and comm like '8414__' order by ryear, comm.
```

結果は表 5 のとおりである。表 5 において欠損値は”.”で表す。全ての年度で不採用 (SEL が欠損値) となっている。基準年である 1965、1970 年の数

量単位（と数量）が欠損しており単価が得られないために不採用となり代表率を低下させていることがわかる。商品分類(COMM)に着目すれば、1960年代は、1967、1968年は5桁分類、それ以外の年は4桁分類である。1972年以降は5桁分類である。数量単位(QUNIT)に着目すれば、1960年代は数量単位が”N0”だったのが、1970年代に”W0”に変わってしまっているものがある。

(84142) これらの商品分類と数量単位の不連続性により、1960年代と1970年代との間に断層が生じてしまっていることがわかる。

代表率の落ち込みの事例として、次に韓国の1995年から2000年を分析する。台湾のときと同じように、以下のSQL文で例として1995年のどの中分類が多数を占めるかを調べる。

表5 日本のラスパイレス総合指数の特異値の調査 (3)

COMM	QUNIT	VAL	QTY	PTQ0	RYEAR	SEL
8414	..	846	0	.	1963	.
8414	..	1,789	0	.	1964	.
8414	..	2,763	0	.	1965	.
8414	..	4,439	0	.	1966	.
84142	N0	813	8,385,708	.	1967	.
84143	W0	417	187	.	1967	.
84144	W0	3,597	507	.	1967	.
84142	N0	977	8,114,700	.	1968	.
84143	W0	1,532	645	.	1968	.
84144	W0	5,903	859	.	1968	.
8414	..	10,778	0	.	1969	.
8414	..	11,309	0	.	1970	.
8414	..	12,366	0	.	1971	.
84141	W0	7,292	1,080	.	1972	.
84142	W0	4,559	980	.	1972	.
84143	W0	22,392	3,475	.	1972	.
84144	W0	261,121	41,391	.	1972	.
84145	W0	52,811	8,641	.	1972	.
84141	W0	9,561	1,310	.	1973	.
84142	W0	8,167	1,407	.	1973	.
84143	W0	75,046	9,418	.	1973	.
84144	W0	407,389	53,957	.	1973	.
84145	W0	501	66	.	1973	.
84141	W0	11,838	1,244	.	1974	.
84142	W0	11,859	1,609	.	1974	.
84143	W0	133,312	12,928	.	1974	.
84144	W0	450,587	52,165	.	1974	.
84145	W0	318	43	.	1974	.
84141	W0	10,222	1,213	.	1975	.
84142	..	14,155	0	.	1975	.
84143	W0	144,101	15,231	.	1975	.
84144	W0	452,420	60,007	.	1975	.
84145	..	481	0	.	1975	.

(出所) AID-XT 基礎データにもとづき筆者作成。

```
select io, ryear, sum(val) val_sum from ti_kor where ryear = 1995 group by io, ryear
order by val_sum desc
```

すると、SITC 中分類が上位 3 つの 72、73、71 だけで取引金額合計の 50%以上を占めていることがわかる。これらのうち、最も金額の大きいのは、以下の SQL 文より、中分類が 7143、7293、7353 であることがわかる。

```
select comm, ryear, sel, qunit, sum(val) val, sum(ptq0) ptq0 from tr03_kor where
( ryear >=1994 and ryear <=2001) and ( io='71' or io='72' or io='73' ) group by
comm, ryear, sel, qunit order by val desc, comm, ryear
```

以下の SQL 文により、これらの商品の 1990 年以降の明細を個別に見る。

```
select comm, ryear, sel, qunit, val, ptq0 from tr03_kor where ryear >=1990 and
( comm='7143' or comm='7293' or comm='7353' ) order by comm, ryear
```

1995 年だけが数量単位が"N0"で、他の年が"W0"である。数量単位の不連続のために 1996 年から 2000 年までが不採用になっているのがわかる。

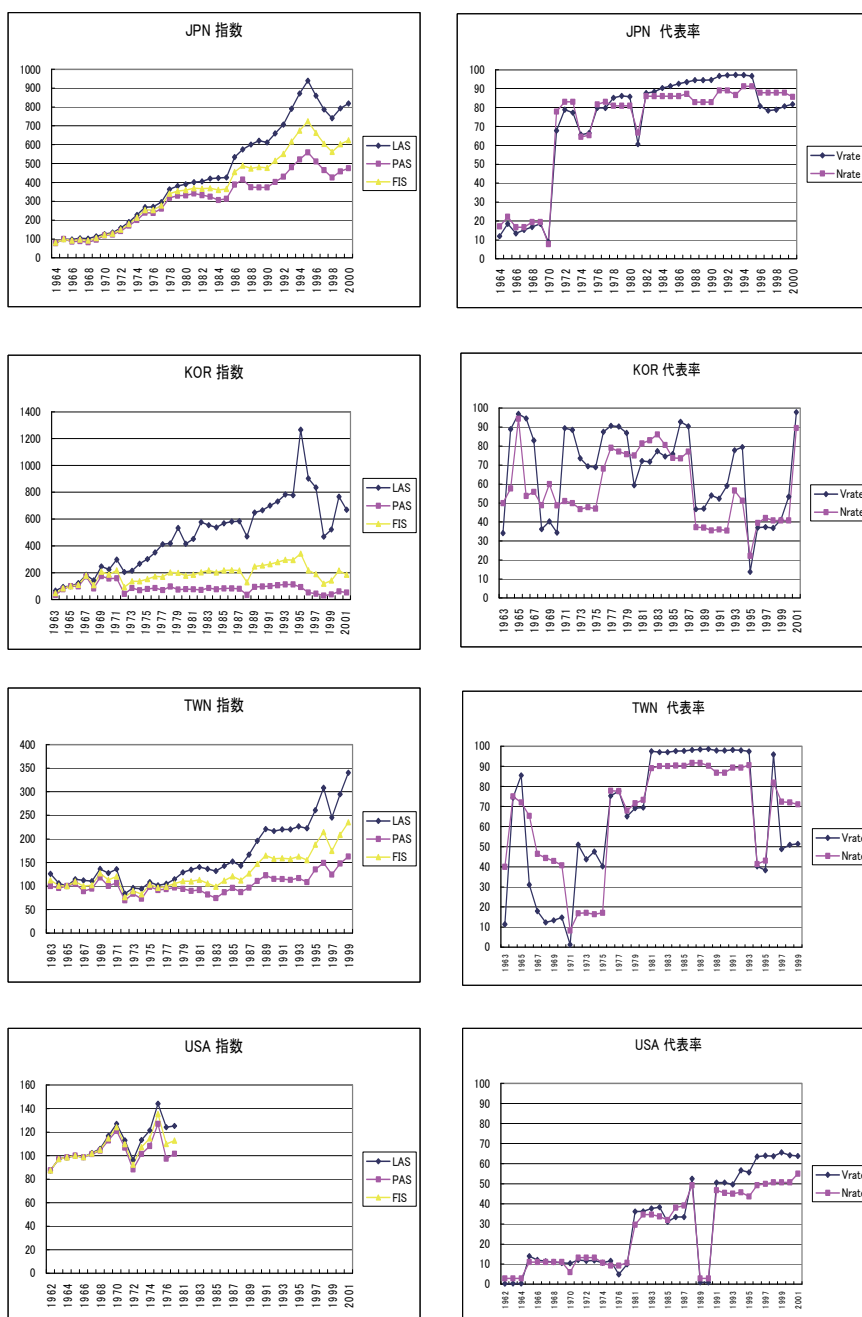
以上の総合指数における特異値の原因をまとめると、(1) 1 件もしくは数件の特異なデータによるもの、(2) 数量単位や商品分類の不連続性によるもの、に分かれると考えられる。(1) については方法を決定して個別の補正をすることによりある程度解決が図れると考えられるが、(2) については国別の元データ全体の構造にかかわる問題であり簡単に解決できない問題である。

4. 中分類別指数における特異値とその原因分析

総合指数が若干の不規則な変動を除けばおおむね安定した指数連となっているのに比べて中分類別指数では指数の変動や代表率の低さにおいて問題が多い。図 4、5、6 は、機械類 (SITC-R1 の 71、72 および 73) について指数と代表率を国別にグラフにしたものである。

日本では、電気機械 (72) の 1995 年以降のパーシェ指数の落ち込みが目を引き (図 5)。総合指数で行ったように元データにさかのぼって調べると、1 件の特異データによるものであることがわかる。SITC 7293 (Transistors, Valves, Tubes etc) の 1995 年データであるが、単価が 1990 年の基準年に大きく増加したため、PT/P0 の値が $0.0538=1/19$ となっている。5 年間なので閾値は $0.04=1/25$ だが、これにかろうじて収まって採用データとなっている。POQT

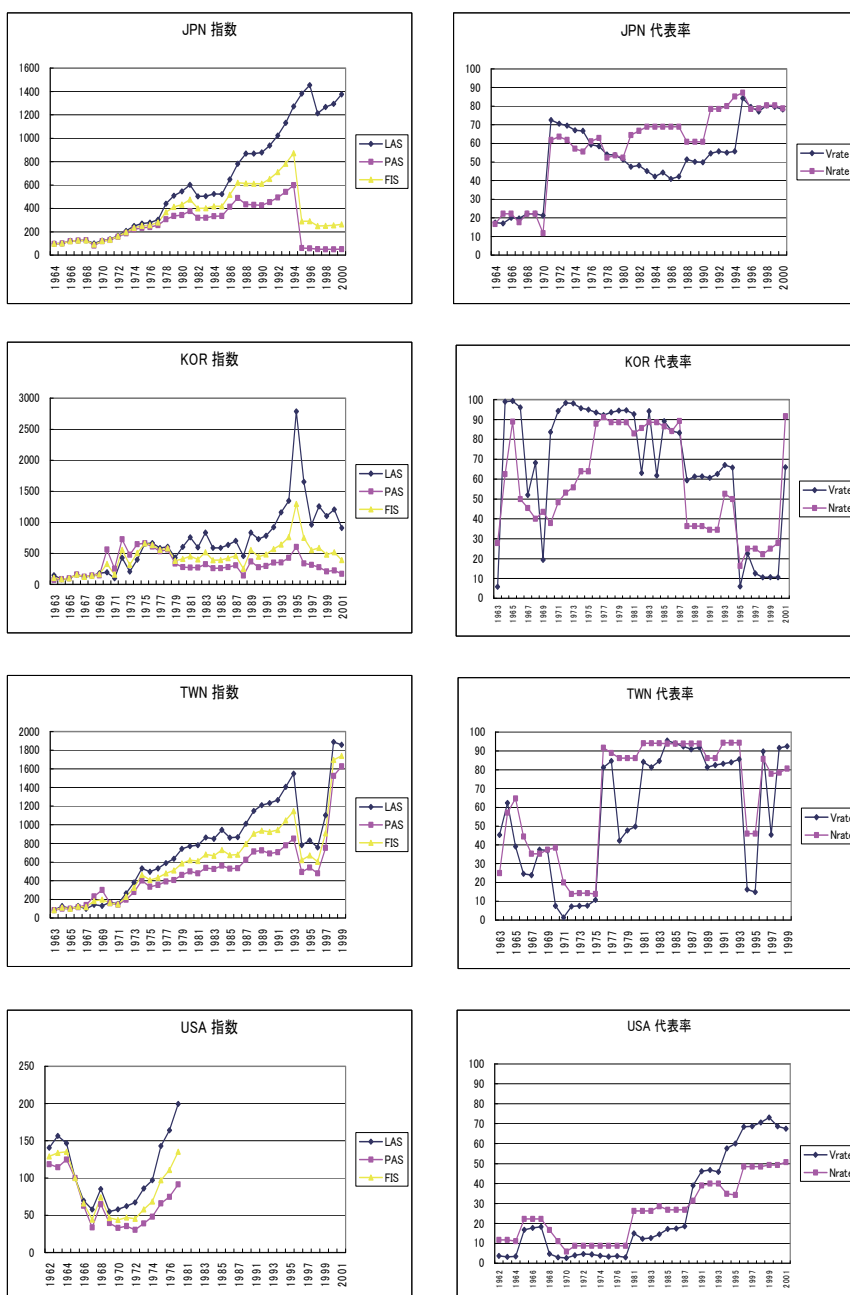
図4 中分類別指数 (SITC-R1 71 一般機械)



(出所) AID-XT 基礎データにもとづき筆者作成。

(注) 図2に同じ。

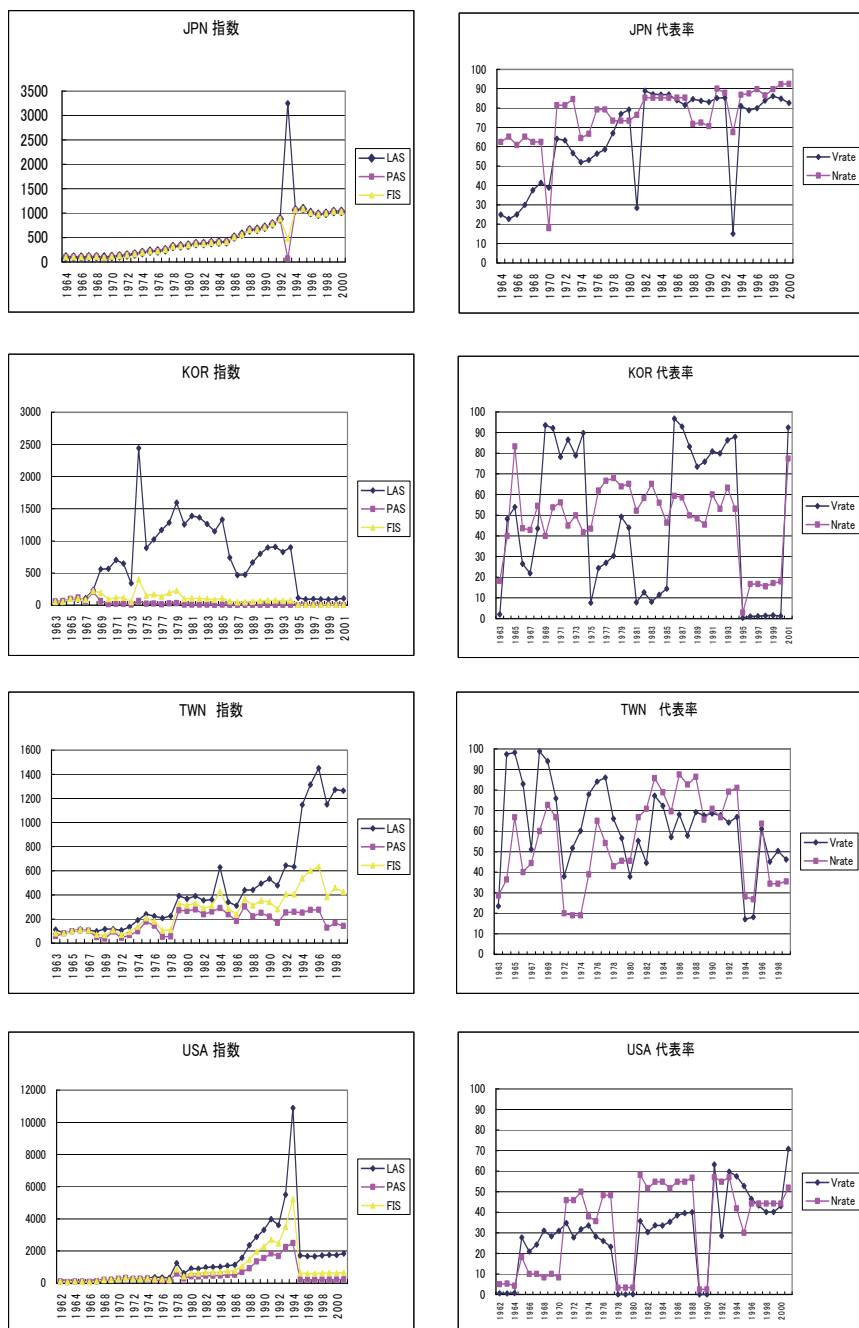
図5 中分類別指数 (SITC-R1 72 電気機械)



(出所) AID-XT 基礎データにもとづき筆者作成。

(注) 図2に同じ。

図6 中分類別指数 (SITC-R1 73 輸送機械)



(出所) AID-XT 基礎データにもとづき筆者作成。

(注) 図2に同じ。

の値が SITC 中分類 72 のなかで、92%を占め、指数全体に大きな影響を与えていることも原因である。また、輸送機械（73）の 1993 年のラスパイレス指数の特異値は総合指数でも特異値となった 7323 の特異データの影響である（図 6）。

韓国の電気機械（72）の 1995 年のラスパイレス指数の特異値と 1995 年から 2000 年までの代表率の落ち込みは、1 件の特異データおよび 1995 年だけ主なデータの数量単位が N0 に変わっていることによると考えられる（図 5）。

米国の一般機械（71）、電気機械（72）は 1978 年以降連続した指数を作成できていない。原因は、1977 年以前の元データにほとんど数量単位がなく、あった場合も不連続であること、商品分類が不連続であることである（図 4、5）。これは台湾の総合指数で見たときと同じ状況である。米国の元データの全体の構造にかかわる問題であり、指数作成の段階で解決するのは難しい。

おわりに

今回作成した指数で明らかになった問題点を解決するための今後の課題として、以下があげられる。

（1）より粒度の細かい元データ、たとえば HS に基づく国連貿易統計（COMTRADE）や各国政府が作成したオリジナルの貿易統計などによって指数を作成する。これにより指数を作成できる時系列は短くなるが、機械類などの個別分析用の指数として、よりきめ細かな、かつ精度の高い分析を可能にする。

（2）連鎖指数を作成する。基準年と比較年との間の単価や品質の違いをより少なくして単価の乖離による指数の変動を抑える。

（3）今回作成した SITC R1 により接続された長期時系列データに基づく指数連の精度を高め、変動を解消する方法を開発する。

（注 1）参考文献 [2] 33 頁参照。

（注 2）参考文献 [1] 参照。

（注 3）参考文献 [3] 6 頁参照。

(注4) 参考文献 [4] 71 頁参照。

【参考文献】

[1] 木下宗七「類別貿易物価指数の算出と算出結果の特徴について－台湾、韓国、アメリカ、日本の比較分析－」(野田容助編『貿易指数の作成と応用－東アジア諸国・地域を中心として－』統計資料シリーズ第 87 集、アジア経済研究所、2003 年)

[2] 黒子正人「IO24 部門分類による貿易単価指数の推計－貿易指数データベースの作成－」(野田容助編『貿易指数の作成と応用－東アジア諸国・地域を中心として－』統計資料シリーズ第 87 集、アジア経済研究所、2003 年)

[3] ———「世界貿易マトリクスの作成に伴う諸問題－貿易指数の推計に向けて－」(野田容助編『世界貿易マトリクスの作成と評価－貿易指数の推計に向けて－』調査研究報告書 2001-III-12、アジア経済研究所、2002 年)

[4] 野田容助「貿易価格指数の算式とその作成方法」(野田容助編『世界貿易マトリクスの作成と評価－貿易指数の推計に向けて－』調査研究報告書 2001-III-12、アジア経済研究所、2002 年)