

貿易統計における不整合要因分析手法の開発

広島大学 学生会員 ○中川 潤哉

1. 背景

近年、経済活動のグローバル化が進み物流システムが複雑になりつつある。このシステムを把握する際に貿易統計は有用であるが、貿易統計には不整合問題が存在し、現状貿易統計の情報が最大限に活用されていない。不整合問題とは、輸出統計の報告値と輸入統計の報告値が一致しない問題であり、貿易統計上の多数のデータで不整合が確認されている。

Fig.1 と Fig.2 は 2017 年の各地域の主要国間における乗用車の取引重量を輸出統計と輸入統計ごとにまとめたもので、先進国の取引でも大きな不整合があることが分かる。この分野の既往研究では不整合問題の評価⁽¹⁾や、貿易統計の取引数量の欠測を補完する手法の開発⁽²⁾がなされているが、不整合問題の要因についての知見や考察は少ない。これは不整合問題に対して、従来の要因分析を行っても、多数の要因が絡み合う複雑な因果関係を解明するのは困難だからである。そこで本研究は従来とは異なるアプローチを取ることによって、不整合要因が貿易統計の不整合問題に及ぼす影響を定量的に評価できる要因分析手法を提案する。

2. 不整合要因分析手法の概要

Fig.3 は本研究で提案する不整合要因分析手法(本分析法)のフローで、Table.1 は本分析法のモデル式の一覧である。本分析法は従来の要因分析と比べて、統計データから因果関係を熟慮せずとも、ある要因が統計に与える影響をダイレクトに評価・分析できる点が最大の特徴であり、それが可能なのは不整合要因の修正方法による所が大きい。本分析法では不整合要因だと考えられるサンプルを欠測値と捉えることで、統計データの欠測処理と同じ要領でその要因を修正したデータを獲得していく。これによって特定の不整合要因が無かった場合の仮想的なデータが得られる。そのデータと元データを比較すること

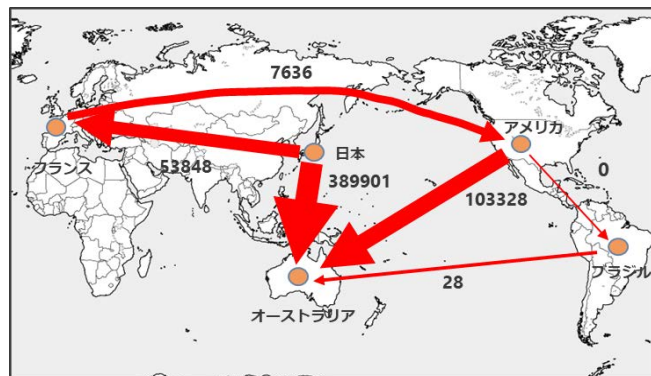
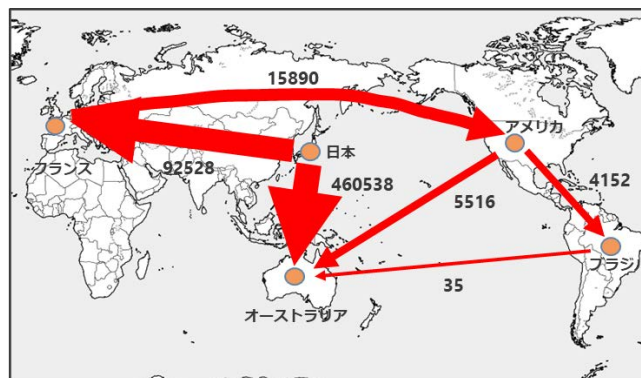
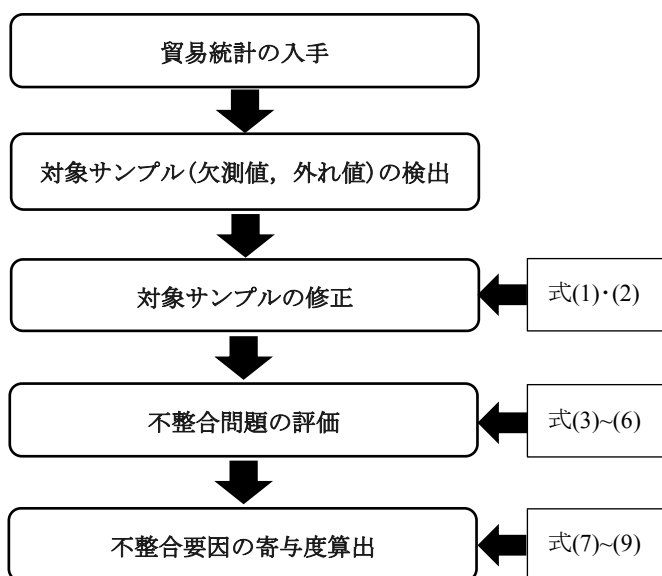
Fig.1 2017 年輸出統計の乗用車取引重量(単位 10³kg)Fig.2 2017 年輸入統計の乗用車取引重量(単位:10³kg)

Fig.3 不整合要因分析手法のフロー

キーワード 貿易統計, 不整合問題, 欠測値, 外れ値, 多重代入法, 級内相関係数(ICC)

連絡先 〒739-0046 広島県東広島市鏡山 1 丁目 4-1 広島大学工学研究科

T E L 082-424-7505

で不整合要因に対して定量的評価をしていくのが本分析法の大きな流れである。なお本研究では不整合要因を①欠測値、②外れ値、③その他の要因の3つに分類し不整合問題の要因分析を行う。そして元データ、各要因修正後のデータ、と両者修正後のデータ間で不整合評価の結果がどれだけ変化したかによって、各要因の不整合問題への寄与度を算出していく。

次に各過程の詳細について説明する。はじめに貿易統計のデータを入手し、欠測値と外れ値の検出を行う。本研究での欠測値は貿易統計の取引数量の項目で記録が未回答のものである。外れ値は取引数量報告値が0であるもの、又は輸入統計と輸出統計の数量報告値に100倍以上の差があり、かつ取引金額報告値から算出される数量推定値と数量報告値に100以上の差があるものとした。外れ値の検出には一般的に使用されている外れ値検定を利用していないが、これでは報告値が外れ値ではないとされ、不整合問題の評価と要因分析の際に有意差が見られなかったためである。

欠測値と外れ値の検出後、これらの修正を行う。両者の修正方法は欠測補完手法の1つである多重代入法を利用する⁽⁴⁾。一般的な欠測補完方法と比べて多重代入法は、欠測値の不確実性をより反映させる事ができる利点がある。本分析法では目的変数を乗用車の取引数量の自然対数、説明変数を取引重量と取引金額の自然対数として回帰分析を複数回実行し、100個の代入データを作成する。次に式(1)から代入データを統合して補完データを作成する。補完データのパラメータは100個の代入データ間の算術平均である。そして取引数量の修正値は、補完パラメータを利用した式(2)から算出する。

その次に不整合問題の評価を行う。本分析法では級内相関係数(ICC)という統計的手法を用いて評価する⁽⁶⁾。ICCとは近年医療統計の分野で注目されており、ある検者がある被検者を複数回測定したときの測定値の一致度を示す場合に利用される信頼性の指標である⁽⁶⁾。ICCには(a)Case1:同一検者による検者内信頼性、(b)Case2:複数検者における検者間信頼性、(c)Case3:検者間の一貫性を考慮した複数検者における検者間信頼性、の3タイプが存在するが、本分析法では検者を輸出統計と輸入統計とし、被験者を

Table.1 本分析法のモデル式の説明

モデル式	式の説明
$(1)\bar{\alpha}_i = \left(\frac{1}{M} \sum_{m=1}^M \alpha_{k,m} \right)$	$\bar{\alpha}_i$:補完パラメータ m:代入データ番号 k:説明変数番号 $\alpha_{k,m}$:代入データのパラメータ
$(2)Y_i = \exp\{\bar{\alpha}_0 + \bar{\alpha}_1 \ln(X_{1,i}) + \bar{\alpha}_2 \ln(X_{2,i})\}$	Y_i :取引数量修正値 i:OD ペア番号 $X_{1,i}$:取引金額報告値 $X_{2,i}$:取引重量報告値
$(3)ICC = \frac{BMS - EMS}{BMS + (k-1)EMS + k(JMS - EMS)/n}$	i:OD ペア番号 j:貿易データの種類
$(4)BMS = \frac{k \sum_{i=1}^n (\bar{X}_i - \bar{X})^2}{n-1}$	BMS:被検者間の平均平方和
$(5)JMS = \frac{n \sum_{j=1}^k (\bar{X}_j - \bar{X})^2}{k-1}$	JMS:検者間の平均平方和
$(6)EMS = \frac{\sum_{i=1}^n \sum_{j=1}^k (X_{ij} - \bar{X}_i - \bar{X}_j + \bar{X})^2}{(n-1)(k-1)}$	EMS:誤差の平均平方和 X_{ij} :取引数量報告値
$(7)MV = B - I \times \frac{ M-I }{ M-I + O-I }$	MV:欠測値の寄与度 OL:外れ値の寄与度 TO:その他の要因の寄与度
$(8)OL = B - I \times \frac{ O-I }{ M-I + O-I }$	I:元データの ICC M:欠測値修正後の ICC O:外れ値修正後の ICC
$(9)TO = B - 1 $	B:両者修正後の ICC

両者の OD ペアとすることで、輸出統計と輸入統計の一致度の評価を目的としているので、Case2 を不整合評価の指標とした。ICC は式(3)から式(6)によって算出される。

最後に不整合要因の寄与度を算出する。本研究では寄与度を式(7)から式(9)によって求める。これらは不整合評価の指標である ICC が、①欠測値のみ修正した場合、②外れ値のみ修正した場合、③両方を修正した場合のそれぞれでどれだけ1に近づいたかによって各要因の寄与度を算出している。

以上が本分析法の一連の過程である。なお本研究では不整合要因を欠測値と外れ値の2つだけにして、例えば途上国の貿易統計が不整合問題に及ぼす影響を分析する際などにも本分析法を利用でき、汎用性がある点が特徴的である。

3. 不整合要因分析のケーススタディ

本研究ではケーススタディとして貿易統計の特定の品目に対して本分析法を適用した。今回は「品目番号(HS コード)8703：定員が9人以下の人員の輸送用に設計した自動車」を対象に、国連貿易統計⁽⁵⁾を用いて、2001年から2017年まで年別に分析を行った。

乗用車貿易統計における各不整合要因の寄与度とそれらの比率を Fig.4 と Fig.5 へ年別に示す。多くの年で外れ値が不整合問題の大きな要因を占めていることが分かった。2008年、2011年、2017年では欠測値の寄与度が外れ値のそれよりも大きかったが、これらの年は欠測値の OD ペアが占める割合が多かった年であり、相対的に外れ値の寄与度が小さい結果となった。また2014年と2017年はその他の要因の寄与度が全体に占める割合が高かった。これは修正前の元データの ICC が1にかなり近いことや、外れ値として検出された OD ペアが他年度よりも少なかったことより、元データと修正データの ICC がほとんど変化しなかったためである。なお本分析法における外れ値の基準は明確な根拠を持っておらず、また取引数量推定値算出の際に利用した金額データの値の正確性について考慮されていないので、基準に偏りが生じる場合が考えられる。あくまで不整合要因の一例として外れ値を挙げたが、公平な外れ値の基準作成が今後必要である。

4. 結論

本研究は、貿易統計の不整合問題の要因分析手法を開発し、ケーススタディとして乗用車貿易統計を対象に、数量データ内の欠測値と外れ値の存在が不整合問題にどれだけ影響を与えているかを検証した。欠測値と外れ値の修正の際には多重代入法を使用し、修正前後で不整合問題の評価指標値が改善されたことから、欠測値と外れ値が不整合問題に大きく寄与していることを確認した。今後の課題として他品目での検証と、公平な外れ値の基準作成、さらに不整合要因をより詳細に定義・分類し、不整合要因の主となるものを明らかにすることが挙げられる。また不整合問題の要因分析から発展して、不整合問題の解決法を開発することが今後求められる。

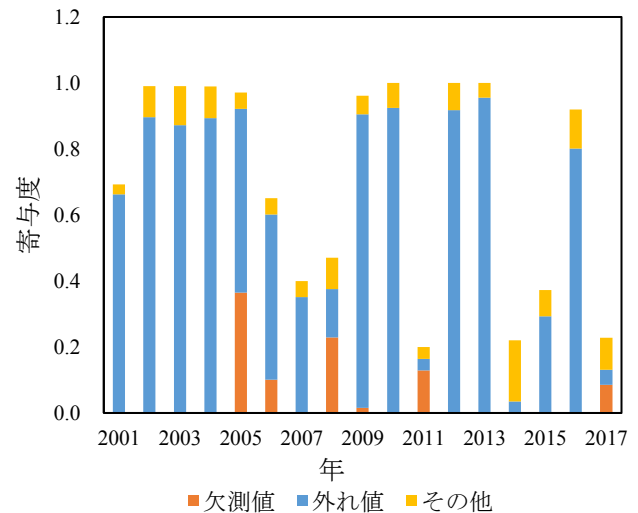


Fig.4 乗用車貿易統計の不整合寄与度

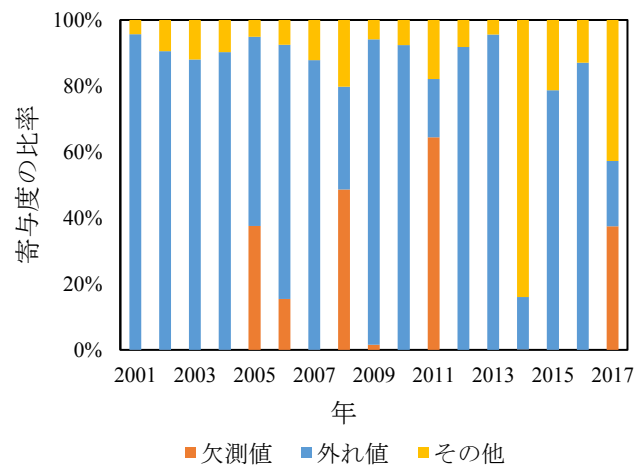


Fig.5 乗用車貿易統計の不整合寄与度の比率

参考文献

- (1)小坂浩之:貿易統計の活用に関する基礎的分析, 海上技術安全研究所報告第14巻第4号特集号, 2016.
- (2)J.Farhan:Overview of missing physical commodity trade data and its imputation using data augmentation, Transportation Research Part C 54, pp1-14, 2015.
- (3)Kohei endo, Masaaki Fuse:Uncertainly Analysis of Global Reuse Monitoring, procedia CIRP, pp172-176, 2017.
- (4)Rubin D.B : Multiple Imputation for Nonresponse in Surveys, John Wiley&Sons, 1987.
- (5) UN Comtrade <<https://comtrade.un.org/data/>>
- (6) 谷 浩明: 評価の信頼性, 理学療法科学, 12 巻, 3 号, p. 113-120, 1997