令和６年（行ウ）第３１号、８７号、８８号

人種等を理由とする職務質問の違法確認等請求事件

原告　モーリス・シェルトンほか

被告　国ほか

準　備　書　面　１２

（被告国準備書面(4)に対する反論のうち
大規模比較調査（第２・２項）に関するもの）

２０２５年9月１８日

東京地方裁判所民事第５１部２Ｄ係　御中

原告ら訴訟代理人弁護士　井　桁　大　介

同　　　　　　弁護士　浦　城　知　子

同　　　　　　弁護士　亀　石　倫　子

同　　　　　　弁護士　谷　口　太　規

同　　　　　　弁護士　戸　田　善　恭

同　　　　　　弁護士　西　　　愛　礼

同　　　　　　弁護士　宮　下　　　萌

原告ら復代理人弁護士　千　葉　飛　鳥

目次

はじめに....4
第１　本件調査がランダムサンプリングではなくとも本件運用の存在が裏付けられるものであること....5

１　用語の整理....5
２　本件調査は偏りが「無視しえない深刻なものである」場合に当たらないこと..6
（１）被告の主張の概要と、「無視しえないほど深刻なもの」かどうか....6
（２）本件調査の割付は適切であること....6
（３）偏りを考慮した上でもなお日本人と在留外国人で職務質問経験を受けた経験に差があると言えること....10

３　ウェブ調査の意義と信頼性について....12
４　ウェブ調査の特質と本件調査の信頼性....16
５　小括....17

第２　本件調査は第三変数について適切に考慮していること....17

１　本件調査はサンプルが異なることから多変量解析を行っていないこと....17
２　第三変数を考慮した上でも日本人と在留外国人で職務質問経験の差は顕著であること....18
（１）年齢階級について....18
（２）居住地域について....19
（３）生活習慣の違いについて調査票の項目に入れなかったとしても不当ではないこと....21

3　重回帰分析を行った結果 22

（1）日本人サンプルについて 23

（2）在留外国人サンプルについて 23

（3）二つのサンプルを統合した重回帰分析の結果 24

4　小括 26

第３　調査票の質問文について 26

1　職務質問に関する調査票の文言は適切であること 26

2　誤読された可能性のある質問は結果に影響していないこと 28

3　Q8に関する被告国の主張に誤りがあること 29

（1）Ｑ８を分析の対象とすべきでないこと 29

（2）Ｑ８を分析の対象にしたとしても被告国の解釈に誤りがあること 29

4　小括 31

第４　その他の被告国の主張に対する反論 31

1　本件調査のサンプル数は量的調査として十分なものであること 31

2　検定について 32

3　職務質問のきっかけ（Ｑ１１）について 33

4　被告国の調査義務について 35

第５　結語 36

二つのサンプルを統合した重回帰分析に用いた変数について 39

本準備書面では、被告国の準備書面⑷における主張のうち、原告が提出した大規模比較調査に対する主張である、第２「２　本件調査について」（同５頁以下）に対し、反論を行う。

はじめに

被告国は、同準備書面の第２の１及び２において、原告が提出した日本人と外国人の職務質問経験に関する大規模比較調査（以下、被告準備書面記載のとおり、「本件調査」という。）の結果が信用できないと主張する。

しかし、被告らは住民基本台帳といった網羅的なサンプルデータも、職務質問についてのデータも最も容易に入手できる立場にある。にもかかわらず、自らデータの収集も提供もせず、あるいは自ら調査した結果をもって実体面で原告に反論するのではなく、単に原告の調査手法の瑣末な技術的な点を論難するのみである。原告準備書面１１記載のとおり被告国には「現在の捜査実務においてレイシャル・プロファイリングが行われていないかどうかを量的及び質的に調査する」義務（③の義務のa)（５頁）があるのであるから、このような態度は許されない。

また、被告国が原告の調査について論難する各点は、いずれをもってしても原告調査が示す５．６倍という差が生じたことについての有効な反論にはなり得ない。サンプルの差や、第三変数の存在によって多少結果が動くことがあるとしても、５．６倍もの大きな差が被告の指摘する調査手法の問題によって生じることはあり得ない。すなわち、被告国の指摘にもかかわらず、原告による調査の結果が人種差別的な運用の兆候を示していることは揺るがない。そうであるならば、「国側に職務遂行時に十分な分析・検討等を行うことが期待されること、情報の偏在があるこ

と、被害者側でアクセスすることができないことなど」に鑑みれば、「本訴では、本件運用の存在にかかる主張、立証責任は転換され、被告らにおいて差別的な運用がなかったことを主張・立証しなければならない」（準備書面（８）・４４頁）。

被告国の主張によって本件調査の信頼性は揺らがず、本件運用の存在が否定されることはないものの、被告国の主張には多くの誤りも含まれるため、以下、被告国の主張のうち、①本件調査が厳密な意味でのランダムサンプリングではないことによる母集団の代表性の問題、②第三変数の問題、③調査票の質問内容の問題について、それぞれ反論を加えることとする。

第１　本件調査がランダムサンプリングではなくとも本件運用の存在が裏付けられるものであること

１　用語の整理

まず混乱を避けるため、用語法を整理する。被告国が「ウェブ調査」と呼称するものは、実際には「インターネットを使ったモニター調査」のことを指す。すなわち、「ウェブ上で協力者を募り、これをプールしておいて、調査を実施する際にこの事前にプールされた協力者を調査対象者として回答を求める」という方法（乙Ａ２６・６６頁）である。被告国が「ウェブ調査」を批判する要旨はランダムサンプリングでないとの点にあるため、以下の議論において「ウェブ調査」との用語は「インターネットを使ったモニター調査」であることを前提とする。

また、「回収率」について、被告国は「回答率」（被告国準備書面⑷・７頁）と表記するが、以下、ここでは「回答率」のことも「回収率」と表記する。

２　本件調査は偏りが「無視しえない深刻なものである」場合に当たらないこと

（１）被告の主張の概要と、「無視しえないほど深刻なもの」かどうか

被告国のサンプルに関する主張は、大要、本件調査のサンプルが母集団を代表しておらず、偏りがあることからその結果は信用できないというものである（被告準備書面（４）６頁以下参照)。これは、本件調査が厳密な意味でのランダムサンプリングではなく、モニターに対して回答を求めるウェブ調査であることをその理由とするようである。

しかし、ランダムサンプリング調査にしても、下記第１の３記載のとおり、回収率によっては、集まるデータは母集団の中でも回答に積極的な人たちのものとなる場合があること等から、偏りが生じうる。偏りが生じるのはランダムサンプリング調査にせよウェブ調査にせよ同様である。問題なのは、被告国が引用する乙A２６号証が指摘するとおり（６５、７７頁)、その偏りが「無視しえないほど深刻なものである」かどうかである（被告国準備書面⑷・７頁)。下記においては原告が行った調査において生じる偏りが「無視しえないほど深刻なものであるか」について一つずつ検討する。

なお、ビジネス業界含めて社会一般の経験や受けとめなどを聞くために現在行われているアンケート調査の大多数はモニターに対して回答を求めるウェブ調査の手法が用いられており、これらについて一般的に信用性がないかのように論じる被告国の主張は極論である。

（２）本件調査の割付は適切であること

本件調査では、年齢や性別等の不均衡を調整するために、対象者選別に割当法を

用いている（甲７７・「提言　Web 調査の有効な学術的活用を目指して」・２頁）ことは、準備書面６（８乃至１１頁）記載のとおりである。

被告国は、本件調査の割付けの比率が実態と乖離している旨主張している（被告国準備書面(4)・８頁）が、かかる反論は本件調査の割付の合理性をまったく理解していない。

ア　国籍ベースの割付は適切ではないこと

まず、被告国は、国籍ベースで割付をすることを前提に、「例えば、フィリピン国籍者の割合が約３０％に上っているのに対し、本件調査での回答者に占める同国籍者の割合は約６％であるなど、大きく乖離している。」と主張する（被告国準備書面(4)・８乃至９頁）。

しかしながら、本件調査で参照としているのは「国籍」ではなく、**「地域」**（アジア、ヨーロッパ、アフリカ等）である。本件調査が国籍ベースではなく地域ベースで割付を行った理由は２つある。一つには、本件調査の趣旨がそもそも「日本国籍者」と「外国籍者」の比較にあるわけではなく、「警察官から見た『日本人と思われる人』と警察官から見た『外国人と思われる人』の職務質問の経験について、有意な差があるか否か」を検証することによるものである（準備書面６・９頁）。すなわち、本件調査は、「外国人と思われる人」か否かの比較調査である以上、国籍の別よりもむしろ見た目により関係する民族的ルーツに比重が置かれるべきであるから、国籍をベースとする割付よりも地域をベースとする割付を行う方が適切なことは明らかである。

二つ目には、本件調査の在留外国人の調査対象者は４２２名であるところ、国籍ベースで割付を行った場合、在留比率が低い出身国の在留外国人はほとんど０としてカウントすることになり、一部の国出身の人たちの職務質問経験を聞く結果にな

ってしまい、不適切であるからである。

したがって、国籍別で割付をすることを前提とした被告国の反論は的を射ていない。

イ　在留資格別の割付を行っていないことに問題はないこと

また、被告国は、本件調査が在留資格別に割付を行っていないことをもって本件調査が信頼できない旨主張する（被告国準備書面(4)・９頁）が、この点も的を射ていない。

第一に、本件調査の主眼は「警察官から見た『日本人と思われる人』と警察官から見た『外国人と思われる人』の比較」であり、在留資格の有無は当然見た目に影響するものではないから、在留資格別に割付を行わないことは何ら不合理なものではない。また、本件調査は、在留資格という概念がない日本国籍者と、そうでない者の比較調査であるところ、「そうでない者（＝外国籍者)」のみ有している在留資格別で割付したところで意味があるものではない。また、本件調査の調査対象者は４２２名であり、これを在留資格別で割付けた場合、在留資格によってはサンプル数が極端に少なくなってしまう可能性があり、この点からも不適切である。

したがって、在留資格別で割付を行っていないことに関する被告国の反論も的を射ていない。

ウ　永住者ベースで割付を行うことが本件調査の趣旨からすれば適切であること

また、被告国は、本件調査が永住者ベースで割付を行ったことについて、「在留外国人の約４分の１（２８．１％）にすぎない永住者（特別永住者を除く。）をベースに割付を行っている点で、実態と乖離している」旨主張する（被告国準備書面(4)・９頁)。

しかしながら、この指摘も本件調査の趣旨をまったく理解していないことを示している。原告準備書面６記載のとおり、本件調査の趣旨は、「警察官から見た『日本人と思われる人』と警察官から見た『外国人と思われる人』の過去５年間における職務質問経験の比較」であり、日本での居住が過去５年に満たない在留外国人を含めると、比較の条件に齟齬が生じる（準備書面６・１０頁）。本来であれば、在留期間が５年以上の在留外国人の出身地域を割付のベースとすべきところ、これに該当する出入国在留管理庁の統計は存在しない。そのため、便宜上、永住許可ガイドラインにおいて、原則１０年間の日本在住、または特例として定住者資格で５年以上滞在していることが永住許可の要件を満たすとされていることを考慮し、５年以上は確実に日本に居住しているはずの永住者の出身地域をベースとして割付を行ったものである（準備書面６・１０頁）。このような調整や工夫をしなければ、「過去５年間に日本人と在留外国人の比較」は適切に行えないのであり、永住者ベースで割付を行ったことは適切である。

エ　小括

以上のとおり、本件調査の趣旨が「警察官から見た『日本人と思われる人』と警察官から見た『外国人と思われる人』の過去５年間における職務質問経験の比較」であることからすれば、本件における割付はこの趣旨に沿った適切な割付であり、被告国の反論はいずれも的を射ておらず、不合理である。

（３）偏りを考慮した上でもなお日本人と在留外国人で職務質問経験を受けた経験に差があると言えること

ア　本件調査の結果はウェブ調査の偏りを考慮したとしてもなお顕著な差があると言わざるを得ないこと

被告国は、モニターを利用したウェブ調査が偏る可能性を抽象的に主張するが、“モニター調査の偏りを考慮したとしても本件において調査結果が示していること”をあえて無視している。

例えば、被告国が引用する乙A２６号証においては、中国人への好感度について、「登録モニター型ウェブ調査において自己選択バイアスを生み出す交絡要因を特定し、これをコントロールしても、なお登録モニター型のウェブ調査が旧来的な方法に基づく社会調査との間に示す差異が残っている」と記載されている（７７頁）。

しかしながら、図２（７６頁）の分布を見ても、本件調査のような約５．６倍の差が大きく縮まるようなものとは到底言えない。約５．６倍という差はランダムサンプリングであろうとモニター調査であろうと日本国籍者と外国籍者の間に職務質問経験の差があるという結論を覆すものにはならないのである。曹意見書（甲２６号証）においても、日本人サンプルと在留外国人サンプルの職務質問経験の差があまりにも大きいことから、「性別や年齢、居住地域といった回答者の属性を考慮しつつ精査したところ、かなりの確度で信頼できる結果と言える。」と結論付けている（同９頁）。

本件調査の分析は、モニター型のウェブ調査であることを前提としたうえで偏りの可能性を十分に考慮し、それでもなお日本人サンプルと在留外国人サンプルとの間の職務質問経験の顕著な差は動かし難いことを示したものであり、単にモニター

型のウェブ調査であるという抽象的な理由でこれを批判する被告国の反論は有効ではない。

イ　モニター型の調査であることから偏りが生じる可能性が大きい場面とは言えないこと

また、本件調査は、モニター型の調査ゆえに偏りが生じるような場面とは異なる。例えば、「インターネットの利用」に関する調査であったとすれば、モニター型のウェブ調査には、モニターがインターネットに親しんでいる人たちの中から選ばれている以上、偏りが生じる可能性が大きいと言える。

しかしながら、「職務質問の経験」に関する調査の場合、住民基本台帳を用いたランダムサンプリング調査とウェブ調査の間で差が生じる可能性は低い。常識的に言って調査会社のモニターが、職務質問の経験を基準として採用されていることなどあり得ないからである。

また、滕ほかが行った研究に基づく論文「在日外国人を対象とするWeb調査の可能性と課題」によれば、モニター調査について、意識項目の中の「被差別経験」は確率標本調査と大きな差が見られなかったという指摘がなされている（甲７８・「在日外国人を対象とするWeb調査の可能性と課題」・１６頁）。差別に関する体験に基づき調査会社のモニターが選ばれることも通常あり得ない以上、当然のことと思われるし、本件でも同様のことが言える。

ウ　「答えたい人バイアス」について

また、仮にモニター型の調査とランダムサンプリングの間に職務質問経験の有無について偏りが生じたとしても、「答えたい人バイアス」（被告国準備書面⑷・７頁・松本意見書（乙Ａ２７・２頁））は日本人サンプルにも在留外国人サンプルにも

生じ得るから、比較の条件を可能なかぎり同じに近づけた場合に日本人サンプルには「答えたい人バイアス」が生ぜず、在留外国人サンプルには「答えたい人バイアス」が生じるとは考え難い。

すなわち、答えたい人バイアスはどちらのサンプルにも生じ得るのであるから、他の条件を可能なかぎり同じとしたときは、職務質問経験の有無についてどちらかのサンプルのみが「答えたい人バイアス」を持っていると考えるのは想定し難い。

したがって、本件調査が他の属性を考慮したうえで比較調査を行っていることからすれば、「日本人サンプル」と「在留外国人サンプル」の間で職務質問経験の有無について「無視できないほどの偏り」が生じるとはいえない。

３　ウェブ調査の意義と信頼性について

上記のとおり、被告国は、ウェブ調査の偏りの可能性についてのみ抽象的に主張し、「どのように偏っているか」について議論しないが、その実、本件において「無視しえない深刻な」偏りが生じることはあり得ないのである。

このことに加えて、「ウェブ調査の偏り」なるものは程度の問題であって、今やウェブ調査についてその用い方に注意を払いながらも基本的にはその意義と信頼性を認めるのが一般的である。

上記のとおりウェブ調査はビジネス界では当然のように日々使われているし、近年ではアカデミアにおいても多数活用されている。例えば、日本学術会議が公表した「提言　Web 調査の有効な学術的活用を目指して」においては、「無作為標本による従来型の社会調査を用いてきた社会科学者も Web 調査を用いるようになってきている」点や、「無作為標本を用いない Web 調査は学術的に意味がないという単純な

議論をすることはできない」点が指摘されている[1]（甲７７・ⅱ頁）。

無作為標本による従来型の社会調査とウェブ調査には一長一短がある。前者は、「母集団の代表性を確保する標本設計には優れているが、調査票の回収率の確保に難がある」（甲７７・１頁）。他方、後者は、「有効回答数の確保は容易であるが、標本の代表性に難がある」（甲７７・１頁）。このように、無作為標本による従来型の社会調査であっても回収率が低い場合、回収標本に偏りが生じる可能性があるのであり、「もはや『Web 調査は無作為抽出標本を用いていないので利用すべきではない。』という主張をすることは難しくなっている」（甲７７・２頁）ほど、ウェブ調査はアカデミアにおいても活用されている。

この点、被告国は、「ウェブ調査は回答率が低く、これが偏りをもたらす可能性も指摘される」旨主張し、吉村治正著の「ウェブ調査はなぜ偏るのか―２つの実験的ウェブ調査から―」（乙A２６・６８頁）を引用する（被告国準備書面⑷・７頁）。本論文に記載されていること自体は誤りではないが、被告国の解釈に誤りがある。通常のランダムサンプリング調査と異なり、ウェブ調査は「回収率」（被告国の言い方では「回答率」（被告国準備書面⑷・７頁）という概念がない。何故なら、調査者が調査会社に調査実施を依頼する際に予め必要な標本数を伝え、必要な標本数を確保することを前提にウェブ調査を実施するからである。そのため、通常のウェブ調査では「回収率」という概念は問題とならず、「回収率」が問題となるのは、通常はランダムサンプリング調査である。乙A２６号証は、そのような状況を前提として、ある種の実験的な結果として、郵送調査とウェブ調査の回収率を示

[1] ここでの「Web 調査」は本書面における「ウェブ調査」と同義である。また、「無作為標本による従来型の社会調査」は、ランダムサンプリング調査と同義である。

したものと位置付けられる。乙A２６号証が６８頁において引用する石田ほか著「信頼できるインターネット調査法の確立に向けて」（甲７９）は、郵送調査とウェブ調査を比較したものであるが、本訴訟で問題となっているのはランダムサンプリング調査とウェブ（モニター）調査の比較であるため、郵送モニターとWEBモニター調査を比較しても無意味である。

実際、下記の表１（甲７９・表３・調査設計詳細・７頁）のとおり、①郵送ランダムの回収率が２１．５％なのに対し、③WEBモニターAの回収率は４５．０％、④WEBモニターBの回収率は７３．４％と、むしろWEBモニターの回収率の方が高くなっている[2]。実際、乙A２６号証においても、通常のランダムサンプリング調査である住民基本台帳から抽出した標本調査ではウェブ調査よりも回収率が低かった旨記載されている（６８頁）。

（表１）

表3　調査設計詳細

調査名	配信数	回収数	回収率	回収締め切り方法	催促の有無	依頼主
①郵送ランダム	3500	753	21.5%	調査期間終了後、予定より多く回収したセルは無作為抽出	10/16に督促はがき発送	開示
②郵送モニター	1340	1108	82.7%		10/16に督促はがき発送	非開示
③WEBモニターA	2666	1201	45.0%		督促なし。回収数未達のセルのみ追加配信10/15に実施	
④WEBモニターB	2005	1472	73.4%	先着順受付	督促は1回(10/16) 配信方法は回収数をみながら期間中分割配信	
⑤WEBモニターC	29995	2088	7.0%	調査期間終了後、予定より多く回収したセルは無作為抽出	10/15に1回	

2　④WEBモニターCは、懸賞メーリングリストのひとつのサービスとして行われたもので（甲７９・５頁）、通常のウェブ調査とは異なるから、本訴訟における比較の対象として適切ではない。

なお、乙A２６号証（６８頁）が引用しているNHK放送文化研究所が行った比較実験は、無作為標本によるランダムサンプリングにおいて郵送で回答を求めたものとWEBで回答を求めたものとを比較したものであり、これも本訴訟で問題となっているランダムサンプリング調査とウェブ（モニター）調査を比較したものでもない。

以上のとおり、被告国の「ウェブ調査は回答率が低く、これが偏りをもたらしている可能性も指摘される」という主張は、あたかもモニター登録のウェブ調査の回収率がランダムサンプリング調査よりも低いかのような誤解を与える主張であり、誤りである。

なお、被告国は「本件調査は、本件調査会社が確保した調査対象者の数及び本件調査における回答率がどの程度であったかについても明らかにして」いないとも指摘する（被告国準備書面⑷・７頁）。しかし、モニター調査では回収率という概念が異なり、「従来の調査と同じように調査の質の判断に使うことは難しい」（甲７７・１２頁）。また、被告国は準備書面⑷の「（母数となる調査対象者の数が不明である以上、回答率を明らかにすることはできない）、回答率の低さがもたらす弊害、すなわち、回答する者が警察による職務質問そのものに興味を持つ者等に偏り得るといった問題がある。」という記載（７乃至８頁）について、これは「母数」と「母集団」の概念を取り違えており、また、「回答率」（準備書面⑷・７頁）（回収率）の低さの問題とモニター調査における対象者の偏りの問題を混同している。回収率（被告国で言うところの「回答率」（被告国準備書面⑷・７頁））の低さが問題となるのは、むしろ従来型のランダムサンプリング調査である。

したがって、ウェブ調査であることをもって本件調査が信頼できないという被告国の主張は認められない。

４　ウェブ調査の特質と本件調査の信頼性

ウェブ調査（モニター調査）にはその調査内容ややり方次第で、ランダムサンプリングより偏りが生じることはあり得る。しかし、重要なことは、「ウェブ調査はどのようなときに偏る可能性があるのか」、「偏りを軽減するために調査設計上工夫できることは何か」、「偏りを考慮した上で最低限言えることは何か」を検討したうえで実査を行い、分析することである。

被告国が提出した乙Ａ２６号証においても、調査そのものについて、標本に関するもの以外で生じる誤差（非標本誤差）について、「理論上は想定されていないが現実には発生してしまう様々な問題、これらをなくすことは現実問題として不可能といわざるをえず、それゆえに非標本誤差の生み出す偏りを根本的に解消する方法もない。社会調査に携わる者に現状でできるのは、非標本誤差の発生を自らの経験から予想し、その予防策を講ずることで、正確な値に近づいただろうという期待を抱くだけである」点を指摘している（７８頁）。そして、標本の問題についても、「旧来的な社会調査法が極めて深刻な問題に直面していること」（回収率の低下）を指摘し、ウェブ調査は自己選択による誤差が生じることを踏まえたうえでも、「ウェブ調査がその打開策の1つになり得るかもしれないこと」を認め、「社会学がウェブ調査を用いて研究を進めるためには、その新しく生じる問題をひとつひとつ具体的に把握し、その対応策あるいは解決法を模索していかねばならない」と述べている。すなわち、ウェブ調査の意義や有効性そのものを否定したものではなく、他のあらゆる調査手法と同様に生じる誤差について留意した上で、対策をとることの重要性を指摘しているのである。

日本学術会議が公表した前掲「提言　Web 調査の有効な学術的活用を目指して」においても、「登録モニターからサンプルを構成した際の偏りをどう評価するか」

という点が１つの焦点である点を指摘している（甲７７・１３頁)。「調査会社の活動が基本的に営利を目的とした活動として行われることから、企業秘密等の保持など、超えられない壁」があり（甲７７・１４頁)、本件調査でも登録モニターの情報は開示されていない。しかしながら、本件調査では上記第１の２（１）で記載したとおり、「警察官から見た『日本人と思われる人』と警察官から見た『外国人と思われる人』の過去５年間における職務質問経験の比較」という観点から割付を十分に考慮及び工夫したものであり、「日本人と外国人といった複数の対象の比較という観点からは、より正確な調査結果が期待できる」（甲２６・３頁）し、またその出された結果の顕著さからしても、「警察官から見た『日本人と思われる人』と警察官から見た『外国人と思われる人」』において職務質問の経験に有意な差が生まれている。

５　小括

　したがって、本件調査がウェブ調査であることをもって「信用できない」という被告国の主張は不当である。職務質問経験について、他の属性の偏りを考慮したうえでなお動かし難いほど大きな差がある本件調査の結果からすれば、本件調査は偏りが「無視しえない深刻なものである」場合に当たらず、本件調査はランダムサンプリングではなくとも本件運用の存在を裏付けるに十分なものである。

第２　本件調査は第三変数について適切に考慮していること

１　本件調査はサンプルが異なることから多変量解析を行っていないこと

　被告国は、「居住地域や年齢、生活習慣の違いといった職務質問に至る事由に関連すると考えられる様々な交絡因子が存在するにもかかわらず、これに対する適切

な配慮がされていないことから、これに対する適切な配慮がされていないことから、本件調査の分析手法は不適切である。」旨主張する（被告国準備書面⑷・１１頁）。

まずそもそも、日本国籍者サンプルと外国籍者サンプルにおいて職務質問を受けた経験及び回数について顕著な差（経験の有無の差については約５．６倍）が認められた（原告準備書面６・１１～１３頁）。曹意見書においても同様の指摘がなされている（甲２６・５頁）。下記のとおり、原告は第三変数の影響を考慮しているが、考慮していない変数が何らかの影響を与えた可能性がゼロではないとしても、本件のような圧倒的な差がある場合には、本件運用の存在という結論に影響しない。

２　第三変数を考慮した上でも日本人と在留外国人で職務質問経験の差は顕著であること

原告は、準備書面６にて既に第三変数を考慮したうえで日本人と在留外国人との間でなお職務質問経験に顕著な差があることを示している。

準備書面６記載のとおり、性別は「日本人調査及び在留外国人調査の両方で男女比がほぼ同等に割り付けられているため、性別について日本人サンプル及び在留外国人サンプルとの間で差があるとはいえず、性別の偏りによって調査結果に差が出たとは言えない。」（１４乃至１５頁）。年齢階級と居住地域については、下記のとおりである。

（１）年齢階級について

年齢階級については、日本人サンプルに比して偏りがあったとしても、準備書面

６及び曹意見書記載のとおり、「『職務質問経験を有する外国人は日本人と比較して顕著に多いという結果は揺るがない。』」（準備書面６・１６頁、甲２６・８頁）。何故なら、全ての年齢階級において日本人サンプルよりも在留外国人サンプルの方が職務質問経験を有する割合が著しく高いからである（準備書面６・１５乃至１６頁・甲２６・６乃至８頁）。被告国の主張（被告国準備書面⑷・１２乃至１３頁）は、年齢階級の偏りを考慮したうえでもなお差があると言えるという原告の主張に対する反論として成立していない。

（２）居住地域について

居住地域についても、当該偏りを考慮したとしても日本人サンプルと在留外国人サンプルとの間に職務質問経験を有する者の間に顕著な差があるということは既出のとおりである（準備書面６・１６乃至１８頁、甲２５・１１頁）。

日本人サンプルが関東圏に限定されているのは、「外国人人口が日本人人口に比べて圧倒的に少ないことから回答者（調査モニター）を確保するのが相対的に難しく、居住地域などの限定をしづらいといった調査実務上の事情があったと推測される」と曹意見書でも記載されているとおり（甲２６・８頁）、調査設計上の限界があったことを理由とするものであり、恣意的にサンプルが抽出されたということからではない。日本人サンプルを関東圏在住者に限定したのは、「在留外国人に関東圏在住者が多いことから、あわせて関東圏在住者とした」（甲２５・６頁）のであり、比較の条件を揃える観点からあえて関東圏在住者に限定したものである。日本人サンプルのみ全国の人口比率に揃えた場合は、むしろ偏りが生じる可能性があり、不適切である。そのため、日本人サンプルを関東圏在住者に限定したのは適切

な割付であったと言える。

また、被告国は、居住地域が関東圏ということだけでは広範に過ぎること及び在留外国人の居住地域は２３区内が多いことから、それらを考慮していない点は適切ではない旨主張する（被告国準備書面⑷・１２頁）。しかしながら、被告国の主張は、都市部の方が郊外よりも職務質問を受けやすい地域であることが前提となっているが、その前提が誤りである。『パトロール活動・１１０番通報等に関する世論調査』（甲３・１４頁）が示すように、そもそも都市規模別の違いはほとんど見られず、「都市規模別にみると、どの都市規模でも、１回でも質問を受けたことがあるという人が約６％、１回もないという人が約９４％」となっている。

したがって、都市部の方が郊外よりも職務質問を受けやすいという仮説を前提とする被告国の主張は誤りである。

なお、準備書面１１記載のとおり、「現在の捜査実務においてレイシャル・プロファイリングが行われていないかどうかを量的及び質的に調査する」義務（③の義務のａ)（５頁）があるにもかかわらず、被告国は、何らの量的調査を行っていない。被告国は、自らランダムサンプリングによる調査を行うことが可能であり、仮にランダムサンプリングではないモニター調査であったとしても被告国であればこのような居住地域の細かな偏りについても詳細な設定ができる立場であるにもかかわらず、調査を実施する姿勢は皆無であり、都市部の方が郊外よりも職務質問を受けやすいという前提に立つのであればその仮説を裏付けるデータを示されたい。

（３）生活習慣の違いについて調査票の項目に入れなかったとしても不当ではないこと

被告国は、生活習慣の違いが交絡因子になるところ、本件調査ではこれが考慮されていない旨指摘する（被告国準備書面⑷・１３乃至１５頁）。

確かに、本件調査では、生活習慣の違いは調査票の項目に入れていない。しかしながら、そもそも交絡因子というのはどのようなものであっても入れれば良いわけでもない。一般的に、性別、年齢等の基礎項目（フェイス項目）については、調査票の項目には必ず入れる変数であり、分析にあたっては、調整変数として入れたうえで多変量解析を行うことが一般的である。しかしながら、フェイス項目以外の変数については、どのような変数を調査票に入れるかは、調査者が「どのような仮説を立てるか」に依存する。例えば、これが「ネット利用」に関する調査であれば、「普段どれくらいインターネットを利用するか」という変数は調査票の項目に入れるべきであり、分析にあたっても考慮すべきものであろう。

本件調査では、職務質問経験の有無に影響を与える変数は、日本人か在留外国人かという「警察官から見て『外国人』に見えるか否か」である仮説を立てた。そのため、フェイス項目については、職務質問経験の有無に大きな影響を与える最低限のものとした。これは仮説が、職務質問経験の有無に最も影響を与える変数は、日本人か在留外国人かという「警察官から見て『外国人』に見えるか否か」であるからであり、生活習慣の違いを調査票の項目に入れなかったとしても何ら不合理な点はない。曹意見書が性別、年齢、居住地域を比較対象としたのは、現時点では最も信頼できる先行研究として参考とした『パトロール活動・１１０番通報等に関する世論調査』（甲３）が性別、年齢、居住地域の変数が取り上げていたことから、先行研究を参照に分析したことによるものである（甲２６・５頁）。

これに比して被告国は、「生活習慣の違いは交絡因子になる」旨断言をして（被告国準備書面(4)・１３頁)、生活習慣の違いを考慮していない点を批判するが、その仮説を支えるような先行研究等は何ら提出していない。なお、被告国は、本件調査の「職務質問を受けたきっかけになった状況」に関する質問項目（甲２５・１４頁、Q１１）を引用して、「生活習慣の違いが交絡因子になる」という仮説を提示しようとしているようにも思える。しかしながら、これは職務質問を受けた後の人に対する回答であり、回答者が「普段どのような生活を送るか」という生活習慣の違いを裏付けるようなものではない。Q１１に対する回答をもって回答者がどのような生活を普段送っているかは分からないのは勿論のこと、この結果をもって「生活習慣の違いが交絡因子になる」ということは断言できるものではないし、前提とすべきことでもない。

もし仮に、被告国の「生活習慣の違いが大きな影響を与える変数となる」との主張と、本件調査の結果とを併せ考えると、それはすなわち外国籍者のサンプルの中に職務質問を受けやすい生活習慣をしている人が圧倒的に多かった可能性があるという主張になりうる。しかし、当然ながらこのような前提は成り立つはずがないし、そのような主張自体が人種差別的な発想に基づくものであろう。

３　重回帰分析を行った結果

また、多変量解析を行う必要がないほど日本国籍者と外国籍者の間で職務質問経験の有無及び回数について圧倒的な違いがあることは明らかであるし、既に他の変数の影響を考慮しているので、多変量解析の分析は必須ではない。しかし、被告国の反論のため、下記のとおり補足的に重回帰分析を行い、他の変数の影響を取り除

いたうえで日本国籍者と外国籍者の間に職務質問が行われた回数について有意な差があるかどうかの検証を行った。

（１）日本人サンプルについて

使用した変数については、別紙１のとおりである。結果は、表２のとおりであり、性別、年齢、居住地域のいずれも有意ではなかった。

（表２）
職務質問経験の回数を従属変数とする重回帰分析の結果（日本人サンプル）

	回帰係数		標準誤差
切片	0.620	***	0.158
ジェンダー（*ref.* 男性）			
女性	-0.142		0.073
年齢	-0.007		0.003
居住地域（*ref.* 東京都）			
茨城県	-0.225		0.162
栃木県	-0.313		0.193
群馬県	-0.079		0.189
埼玉県	-0.101		0.111
千葉県	-0.147		0.116
神奈川県	-0.080		0.101
調整済み決定係数	0.007		
N	501		

*=p<0.05, **=p<0.01,***=p<0.001,

（２）在留外国人サンプルについて

使用した変数については、別紙２のとおりである。結果は、表３のとおりであり、男性に比して女性の方が職務質問を受ける回数は少なく、これは０．１％水準で有意であった。年齢については、有意な差が見られなかった。国籍（地域別）については、ヨーロッパ、北アメリカ、南アメリカはアジアよりも職務質問を受ける回数が多いことについて０．１％水準で有意な差が見られ、オセアニアについては

５％水準で有意であった。在留資格については、永住者に比べて「就労関連」、「留学」は職務質問を受ける回数が多いことが０．１％水準で有意であった。

（表３）
職務質問経験の回数を従属変数とする重回帰分析の結果（在留外国人サンプル）

	回帰係数		標準誤差
切片	0.624		0.338
ジェンダー（*ref.* 男性）			
女性	-0.678	***	0.167
年齢階級（*ref.* 20代）			
30代	0.136		0.223
40代	-0.070		0.319
50代	-0.312		0.418
国籍（地域）（*ref. アジア*）			
ヨーロッパ	1.096	***	0.291
アフリカ	0.700		0.452
北アメリカ	1.022	***	0.298
南アメリカ	1.135	***	0.206
オセアニア	1.018	*	0.511
在留資格（*ref. 永住者*）			
定住者・配偶者	0.582		0.298
就労関連	1.340	***	0.242
留学	1.180	**	0.400
その他	0.472		0.489
調整済み決定係数	0.185		
N	422		

*=p<0.05, **=p<0.01,***=p<0.001,

（３）二つのサンプルを統合した重回帰分析の結果

日本国籍者と外国籍者の間で職務質問経験の有無及び回数について圧倒的な違いがあることは明らかであり、このような場合、異なる二つのサンプルを統合して重回帰分析を行うということは行わないが、被告国への「交絡因子を考慮していない」という主張への反論を行うため、日本人サンプルと在留外国人サンプルを統合

して職務質問回数を従属変数とする重回帰分析を行った。すなわち、二つのサンプルに共通する共通変数である「性別」及び「年齢階級」を投入したものである。使用した変数については別紙３のとおりである。

結果は表４のとおりである。係数が1.620とあるように、在留外国人は日本人に比して職務質問回数が圧倒的に多く（外国籍者であればあるほど、日本国籍者に比して職務質問経験を受ける回数が１．６２回多い）、これは０．１％水準で有意である。また、女性は男性に比して職務質問経験を受ける回数が少なく０．１％水準で有意であるが、標準化回帰係数は-0.140であり、日本国籍者の有無の標準化回帰係数が0.491であるのに比べると明らかに「日本人か在留外国人か」という国籍因子の方が相関が大きい。年齢についても、４０代と５０代は２０代に比べて職務質問を受ける回数が少ないが、これについても標準化回帰係数は-0.103、-0.105であるため、明らかに「日本人か在留外国人か」という国籍の方が相関は大きい。

（表４）
職務質問経験の回数を従属変数とする重回帰分析の結果（統合済みサンプル）

	回帰係数		標準誤差	標準化回帰係数
切片	0.674	***	0.114	
日本国籍の有無（*ref.* 日本人）				
在留外国人	1.620	***	0.097	0.491
ジェンダー（*ref.* 男性）				
女性	-0.460	***	0.090	-0.140
年齢階級（*ref.* 20代）				
30代	-0.043		0.122	-0.012
40代	-0.398	**	0.133	-0.103
50代	-0.444	**	0.144	-0.105
調整済み決定係数	0.306			
N	923			

*=p<0.05, **=p<0.01,***=p<0.001,

４　小括

被告国は、他の変数が影響を与えた可能性の主張をするが、他の変数を統制（他の変数の影響を考慮すること）しても、在留外国人の方が日本人よりも職務質問経験を受ける回数が多いことが裏付けられた。したがって、国の主張に合理性はない。

第３　調査票の質問文について

１　職務質問に関する調査票の文言は適切であること

被告国は、「本件調査の質問**内容**において使用されている『職質経験』という用語の定義付けについて、日本人用調査票では『職質経験』（甲２５号証２９ページ）、在留外国人用調査票（英語）では「police questioning」（甲２５号証４１ページ）とそれぞれ特段の説明もなく、その用語が使用されている。」旨指摘する（被告国準備書面(4)・１０頁）。

しかしながら、当該被告国の主張は、甲２５号証の調査票に対するミスリーディングと言わざるを得ない。本件調査では、職務質問経験の有無を尋ねた質問はＱ１０である。Ｑ１０の日本人調査の質問は、「あなたは過去５年のあいだに、日本で警察官から公共の場や屋外などで質問されたことがありますか？」というものであり（甲２５・３２頁）、「職務質問」という用語は使っていない。Ｑ１０の他にも、Ｑ１１（甲２５・３３頁）では、「警察官から質問された」、「質問されるきっかけ」、とあくまで「警察官からの質問」を尋ねており、Ｑ１３～Ｑ１５、Ｑ１７、Ｑ１９、Ｑ２０でも同様に「警察官からの質問」という記載になっている（甲２５・３４乃至３８頁）。また、Ｑ１６・Ｑ１８（甲２５・３６乃至３７頁）では「警察官からの声か

け」という文言が使われており、これも「職務質問」という文言は使われていない。

本件調査で「職質経験」という文言が使われているのは、質問<u>内容</u>とは関わらない柱書（甲２５・２９頁）のみであり、それにもかかわらず質問<u>内容</u>にて『職質経験』という用語が使われているという被告国の主張は、誤りである。調査票を十分に読んでいないのか、殊更にミスリードをしようとしているのか意図は不明であるが、少なくとも、内容面に関わる質問項目で「職務質問」という用語を使っていることを前提とする被告国の主張は明確な誤りである。

また、英語における調査票も同様に職務質問について尋ねたＱ１０は「questioned by police in Japan in public spaces or outside」との記載になっている（甲２５・４６頁）。

なお、被告国は、「『職務質問』の定義は、それ自体、曖昧である上、この点に関する文献『警察との接触経験に関する探索的研究』（乙Ａ２９号証）においても、『人々が有している職務質問のイメージがどのようなものであるかは明らかではなく、また、そこに個人差もあり得ることが考えられたため、職務質問という言葉自体は使わない』（乙Ａ２９号証２０２ページ）と指摘されている。」と主張する（被告国準備書面(4)・１０頁）。しかしながら、被告国が引用する乙Ａ２９の調査票を見てみると、「あなたはこの１年くらいの間に、以下のような状況で、警察官に呼び止められたり、自宅を訪問されたりして、質問されたことが、どのくらいありましたか。」という記載となっている（乙Ａ２９・２２３頁）。乙Ａ２９にて行われた調査は、「この調査において尋ねられている警察官との接触は、主として職務質問に関心を寄せるものであるが、職務質問以外の接触経験をも含むものとなっている」との記載どおり（２０３頁）、職務質問に限定しない警察官との「接触」を明

らかにすることを目的としている。本件調査とは調査範囲が異なるため、調査票の質問文の記載も異なっていたとしても何らおかしなことではない（もっとも、実際には上記のとおり原告の調査においても「職務質問」という言葉は質問内容に用いられていない）。

２　誤読された可能性のある質問は結果に影響していないこと

Ｑ５については、在留外国人サンプルの中で日本以外の民族的ルーツが「ない」と答えた回答した者が６０．２％いたが、Ｑ５の質問文は、「Do you have ethnic roots in a country/countries other than Japan? If you are Japanese with mixed origins, select “Yes”.」である（甲２５・４４頁）。この訳は「あなたは日本以外の国に民族的ルーツがありますか？日本と他の民族のミックスなどの場合も『ある』を選択してください。」である（甲２５・５８頁）。

しかしながら、日本語以外の質問を読んだ人は「If you are Japanese」のところを含めてこの２文をセットで読み、自身が日本国籍者ではないことから「No」を回答してしまった回答者が多発した可能性が高い。そうでなければ外国籍者を対象にしている質問と整合しない。この質問文の組み合わせゆえに誤解が生じ、外国籍者であっても在留外国人サンプルの中で日本以外の民族的ルーツが「ない」と回答してしまった者が多発したと考えられる。

もっとも、本件調査の中核であるＱ１０の質問は、「In the past 5 years, have you been questioned by police in Japan in public spaces or outside? If never, select “Never”, if you have been, select an approximate number (Select one)」であり（甲２５・４６頁）、Ｑ５と異なり、回答者に誤解を与えるような文言ではない。そして、Ｑ１０の質問は、Ｑ５で「はい」、「Yes」と答えた人だけではな

く、全員に対して行った質問であり（日本人サンプルはn=521、在留外国人サンプルはn=422のため。甲２５・１１頁）、Ｑ５は分岐していないことから、Ｑ５で誤訳による誤回答があったとしても、Ｑ１０に何ら影響を与えるものではない。

したがって、Q5の誤訳をもって、「この一事をもってしても、本件調査に信ぴょう性がないことは明らかである。」（被告国準備書面(4)・１１頁）と指摘する被告国の主張は妥当ではない。

３　Q8に関する被告国の主張に誤りがあること

（１）Ｑ８を分析の対象とすべきでないこと

Ｑ５の分岐となっている「見た目」に関する質問であるＱ８は、Ｑ５に連動するものであることから（Ｑ５で「はい」と答えた日本人サンプル１８人（n=18）、「Yes」と答えた在留外国人サンプル１６３人のみが回答している（甲２５・１０頁）。）、誤回答が含まれている可能性が明らかなＱ８を分析の対象とすることは適切ではない。そのため、原告準備書面６においてもＱ８を用いた分析は行っていない。したがって、Ｑ８を前提とする被告国の主張は的を射ていない。

（２）Ｑ８を分析の対象にしたとしても被告国の解釈に誤りがあること

また、仮にＱ８とＱ１０のクロス集計を行って分析する場合でも、Ｑ８に関する被告国の解釈とその主張に誤りがある。被告国は、「Ｑ１０において、職務質問をされた経験が『ない』と答えた者の割合は、Ｑ８において、①見た目によって外国人と判断されたことが『ない』と回答した者では５０％であったのに対し、②見た目によって外国人と判断されたことが「何度かある／ときどきある／頻繁にある」と回

答した者では５２．４％であって、両者にはほとんど差がな」い点指摘し（被告国準備書面⑷・１８頁）、松本意見書でも「有意な傾向差とは認定しがたい」点（乙A２７・４頁）が指摘されている。しかしながら、これは明らかなミスリーディングである。

本来比較の対象とすべきは、①見た目によって外国人と判断されたことが「ない」と回答した者と、③見た目によって外国人と判断されたことが「常にある」と回答した者であり、②見た目によって外国人と判断されたことが「何度かある／ときどきある／頻繁にある」は①と③の中間であるから、①の比較対象としては不適切であり、本来比較の対象とすべきは①と③である。そして、③の「常にある」と回答した者のうち、Q１０の職務質問をされた経験が「ない」と答えた者の割合は、４０．５％であり、①の回答者の５０．０％とは１０ポイント近く異なる。回数で見ても、①の加重平均は０．８回であったのに対し、③は１．７回と２倍近い開きがある（甲２５・１２頁）。本来の比較対象が①と③であるはずなのに、①と②を比較対象とした点は、裁判所をミスリードしようとしたものと疑わざるを得ない。

上記第３の３（１）記載のとおり、本来誤回答が含まれていた可能性があるQ８に関する調査結果は分析の対象とすべきでないが、仮にQ８を分析の対象としたとしても、「常に外国人だと判断される者」とまったくそうではない者の職務質問経験の１０ポイント近い開きがあることからすれば、この点は「外国人だという見た目に基づいて職務質問を行う」という本件運用を裏付ける方向に働く要素であり、「むしろ、在留外国人の回答者についての職務質問の経験の有無は、他人から見た目で外国人と判断された経験の有無とは関連しないとの結果が得られたとさえ評価できる」という被告国の主張は明確な誤りである（被告国準備書面⑷・１８乃至

１９頁)。

４　小括

　したがって、調査票の質問文に関する被告国の反論はいずれも的を射たものではなく、本件調査の信頼性に影響を与えるものではない。

第４　その他の被告国の主張に対する反論

　被告国は、その他本件調査に対する信用性に関する主張を縷々行うが、以下、「その他の論点」として反論を行う。

１　本件調査のサンプル数は量的調査として十分なものであること

　松本意見書は、本件調査の対象者が、日本人サンプルが５００人強、在留外国人サンプルが４００人強であることについて（乙Ａ２７・１頁)、本件調査のサンプル数が量的調査に耐えられないものであるかのような印象を与える可能性があることから、本件調査のサンプル数は量的調査として十分なものであることを念のため付言する。

　社会調査では慣例として最低サンプル数を４００とすることが多い。これは、「標本からの推定値が母集団の値から確率的にずれて出現するときの、ずれの度合い」（甲８０・『社会調査法入門』・１１９乃至１２０頁）である標本誤差を基準として算出されるサンプルサイズである。表５（甲８０の表７．１（１２０頁)）が示すとおり標本数が４００のとき、標本比率の値は、「母比率±４．９％」の範囲に０．９５の確率で出現する。この標本誤差の大きさが±１０％に収まることから、慣例として最低サンプル数を４００とすることが多いのである。

（表５）

表 7.1　標本誤差の大きさ

〔母集団において比率が 50 % であるとき
0.95 の確率で標本の比率がずれる範囲〕

標本数	標本誤差の大きさ
100	±9.8%
400	±4.9%
2,500	±2.0%
10,000	±1.0%

なお、本件調査が仮にランダムサンプリング調査でサンプリング数が約４００の場合、職務質問経験の割合が７１．１％であったとすると、母集団での職務質問経験は９５％の確率で６６．２％～７６％の間に収まると言える。

このように、本件調査は最低サンプル数を上回っており、量的調査に耐えられないようなサンプル数ではない。

２　検定について

原告準備書面６では、検定結果についても記載しているが、検定はランダムサンプリングでなければ本来行わないものであることは前提としている。そもそも、本件調査では日本人サンプルと在留外国人サンプルの差が約５．６倍もあることから、検定などせずとも顕著な差があることは明らかであるが、統計的な知見に照らしても差があることを確認したにすぎない。

また、学術調査では、ウェブ調査であったとしてもランダムサンプリングと同様に検定結果を出すことは慣例として行われている。曹意見書においても、「十分な無作為抽（ママ）でない場合も、変数間の関連の有無を述べる際には、便宜的にそれを無作為抽出標本のデータとみなして有意確率を計算することが多い」（甲２

６・５頁・脚注９）点指摘されている。

また、松本意見書の「有意差検定とは、本来、同一調査中の特定質問に関して、回答結果間の比率の相違が有意なものと言えるかどうかを判定するための指標である。にもかかわらず、調査報告書、および、意見書においては、外国人モニター調査と日本人モニター調査という質的に全く異なる別種の調査の回答結果間の判定に用いられており、社会科学の限界を超えた用いられ方がなされている」という指摘は、明確な誤りである（乙Ａ２７・４頁）。

検定には様々な種類がある。母比率の差の検定は、「２つの母集団の間で、ある特性をもっているものの割合に差があるかどうか」という問題を検証する場合は、母比率の差の検定が行われる（甲２８・２５７頁）。これは統計的に確立された手法であり、松本意見書の「社会科学の限界を超えた用いられ方がなされている」という主張は、趣旨が不明かつ明確な誤りと言わざるを得ない。

以上のとおり、「本件調査の調査結果に基づいた検定は意味をなさない」という被告国の主張は誤りである（被告国準備書面⑷・１９乃至２０頁）。

３　職務質問のきっかけ（Ｑ１１）について

被告国は、職務質問のきっかけになる質問（Ｑ１１）を取り上げ、「職務質問を受けた経験がある在留外国人の回答者の約７割」が「警職法２条１項の定める不審事由（あるいは警察法２条１項による声かけ（職務上の質問）を行うに足りる事由）によって警察官から質問を受けたものと認識していることになる」旨主張する（被告国準備書面⑷・１７乃至１８頁）。

しかしながら、同質問は、職務質問を受けた者の受け止め（推測）を尋ねたものであり、実際に警察官が職務質問をしようと思ったきっかけと一致するとは限らな

い。そして、日本人サンプルの４３．３％が「まったく思い当たる理由がない」と回答しているのに対し、在留外国人の同回答が９．０％にとどまることを考える（甲２５・１４頁）と、この数字の解釈はバイアスが生じていると思われ、単純比較をすべきでない。

すなわち、在留外国人サンプルの職務質問経験は、回数からいっても「１回」と回答している者は１６．６％にとどまり（甲２５・１１頁)、過半数の者（５４．５％）が複数回の職務質問を受けている。複数回職務質問を受けた者は、職務質問を受ければ受けるほど、職務質問のトラウマが生じうる。具体的には、「自分に非があったのではないか」と過度に自分に非があると思いがちであり、実際には警職法の該当する事由があったとまでは言えないような微細な事情についても「職務質問を受けたきっかけがあった」と回答した可能性が認められる。これに対し、日本人サンプルの職務質問を受けた割合は１２．８％であり、５％の者が「１回のみ」と回答している。複数職務質問を受けた者の割合は、７．６％にとどまる。職務質問を受けた経験がない者にとって、「何故自分が職務質問をされたのか納得がいかない」と考えるのは想定できるものであり、このような被害の経験の過多が、「職務質問のきっかけとなった事情」という主観的な質問に対する回答に歪みを与えた可能性がある。そして、「警察官から目をそらした」「歩き方を変えた」などは、基本的には職務質問の不審事由に当たらず、レイシャル・プロファイリングの方便とされたことと矛盾しない。

また、日本国籍者と外国籍者との間での顕著な職務質問経験の差が、実際の不審事由の有無の差によって基礎付けられる可能性を指摘する被告国の主張は、「在留外国人の方が不審事由を行いやすい」ということを前提とするものであり、それ自体不合理かつ差別的である。

さらに、職務質問を受けた日本人サンプルのうち４３．３％の者が、不審事由がないにもかかわらず職務質問を受けたと答えているが、この認識と実際の不審事由の有無とを結びつけて考えるのだとすると、これは被告国が主張する「不審事由が認められないかぎり職務質問を行わない」という姿勢とは乖離している結果となるであろう。

Ｑ１１の回答は、過去の職務質問経験の差や弱い立場にいることによる権力組織である警察規範の内面化など個人の主観に影響を及ぼして得られた結果と考えられ、実際に不審事由の存在が警察官による質問のきっかけになったことを示すものではない。

したがって、当該被告国の主張は当たらない。

４　被告国の調査義務について

本準備書面冒頭で述べたとおり、被告国は、調査義務があるにもかかわらず、これを行わず、本件運用がなされていないことに関する立証責任を果たしていない。被告国であれば、住民基本台帳を用いたランダムサンプリング調査は容易なはずである。

また、外国人調査は、日本人を対象とした調査と異なりランダムサンプリング調査を行うことが非常に難しい。「日本の人口に占める外国人の割合は約３％に過ぎず、住民基本台帳を用いた標本抽出には多大な費用や労力、準備期間が必要となる一方で、回収率は低い水準にとどまっている」点も指摘されている（甲７８・１４頁）。そのような中で、在留外国人を対象としたウェブ調査も注目されはじめている。被告国は、「ウェブ調査が優れているのは飽くまでも『利便性、迅速性、廉価

性』であって、ウェブ調査によって得られた結果を一般化」することは相当ではない旨主張する（被告国準備書面⑷・２０頁）が、本件調査をランダムサンプリング調査で行った場合、回収率が大変低くなる可能性も予測されたことから、ウェブ調査を行ったことはむしろ適切であったと言える。

被告国は、自ら調査を行わないばかりか、自らが引用している文献の読み方も解釈が誤っているものが散見され、自ら引用する「文献の一部のみを切り取って引用」しているのは、むしろ被告国の方である（被告国準備書面⑷・２１頁)。

第５　結語

以上述べてきたとおり、被告国の本件調査に対する反論はいずれも認められず、本件調査は本件運用を裏付けるものである。

以上

別紙１

使用した変数について（日本人サンプル）

日本人サンプルの分析に使った変数は、従属変数であるＱ１０（職務質問経験）とフェイス項目である性別（Ｑ１）、年齢（Q2)、居住地域（Ｑ２－１（Q2S1)）である。

Ｑ１０の回答は、「ない」を０、「１回」を１、「２回」を２、「３～５回程度」を４、「６～９回程度」を７．５、「１０回以上」を１０とする連続変数とみなした。「回数はわからない・覚えていない」(n=20）は欠損値として処理した。

Ｑ１（性別）の選択肢は「男性」、「女性」、「その他答えたくない」であるが、「その他答えたくない」の回答者はいなかったため、「男性」を参照カテゴリーとする二値変数（男性か女性かの二値とする変数）とした。

Ｑ２（年齢）は連続変数とした。

Ｑ２－１（Q2S1)（居住地域）は、関東に限定されているため、回答は、「茨城県」、「栃木県」、「埼玉県」、「千葉県」、「東京都」、「神奈川県」である。居住地域は、「東京都」を参照カテゴリーするカテゴリカル変数（質的変数）とした。

別紙２

使用した変数について（外国人サンプル）

在留外国人サンプルの分析に使った変数は、従属変数であるＱ１０（職務質問経験）とフェイス項目である性別（Ｑ１）、年齢階級（Q2)、国籍（地域別）（S2)、在留資格（Q4）である。

Ｑ１０の回答は、「ない」を０、「１回」を１、「２回」を２、「３～５回程度」を４、「６～９回程度」を７．５、「１０回以上」を１０とする連続変数とみなした。「回数はわからない・覚えていない」の回答はなかった。

Ｑ１（性別）の選択肢は「男性」、「女性」、「その他答えたくない」であるが、「その他答えたくない」の回答者はいなかったため、「男性」を参照カテゴリーとする二値変数とした。

Ｑ２（年齢）は「２０代」を参照カテゴリーとするカテゴリカル変数とした。

S2（国籍（地域別））は、アジア（東南アジア）を参照カテゴリーとするカテゴリカル変数とした。

Q4（在留資格）は、永住者を参照カテゴリーとするカテゴリカル変数とした。

別紙３

二つのサンプルを統合した重回帰分析に用いた変数について

分析に使った変数は、従属変数であるＱ１０（職務質問経験)、独立変数である国籍（日本国籍か否か）並びに共通変数である性別及び年齢階級である。

Ｑ１０の回答は、「ない」を０、「１回」を１、「２回」を２、「３～５回程度」を４、「６～９回程度」を７．５、「１０回以上」を１０とする連続変数とみなした。「回数はわからない・覚えていない」(n=20）は欠損値として処理した。

国籍については、「日本人」を参照カテゴリーとするカテゴリカル変数とした。

性別の選択肢は「男性」、「女性」、「その他答えたくない」であるが、「その他答えたくない」の回答者はいなかったため、「男性」を参照カテゴリーとする二値変数とした。

年齢について、日本人サンプルは、年齢が連続変数となっていたため、在留外国人サンプルと合わせて「２０代」､」「３０代」、「４０代」、「５０代」とカテゴリカル変数に変換し、「２０代」を参照カテゴリーとした。