

構造改革特区をどのように評価すべきか プログラム政策評価の計量手法からの考察

鈴木 亘*

(大阪大学大学院国際公共政策研究科助教授)

(社)日本経済研究センター副主任研究員)

1. はじめに

2002年12月の構造改革特別区域法公布により、経済構造改革や経済活性化推進の手段として特定地区において規制緩和の特例を設ける「構造改革特区」が実施されている。

構造改革特区は、現在までに4次にわたって324件もの特区申請が認定され(第1回認定117件,第2回認定47件,第3回認定72件,第4回認定計画88件),既に2003年中から異例の速さで規制緩和が実行されている。その内容は多岐にわたっており,①教育関連では,多様な公立教育カリキュラム,学校法人以外の学校設置運営,幼稚園保育所連携,大学設置校地面積基準の緩和,②農業関連においては農村地域における株式会社の農業経営,市民農園などに係わる農地の特例措置,農地取得の容認,③福祉・医療関連では,特別養護老人ホームの公設民営化・PFI方式の運営,外国人医師の参入,④国際交流・国際物流では,通関業務の24時間365日化,総合保税地域の許可案件の緩和,港湾施設の民間企業貸し付け,ビザ関連の制度緩和,⑤IT・新産業創出分野では,国立大学教員の兼業規制緩和や国立大学施設の民間開放による産学連携推進,次世代エネルギー・リサイクル関連の規制緩和等が推進されている。

この構造改革特区のユニークな点は,浜田(2004)が指摘するように,減税措置や補助金等の財政措置などを伴わない純粋な規制緩和のみの特区であるということであり,有効に機能するのであれば,極めて安価な経済活性化策といえるであろう。もう一つの特徴は,特区による規制緩和は有効に機能すれば,速やかに全国的な規制緩和を実行することが予定されているという点であり,その意味で特区は単なる地域活性化策ではなく,わが国全体に対する経済対策なのである。さらに,特区による規制緩和は,オープンな政策評価が行われ,その評価に基づいて全国展開するか否かが決定されることになっている。このように

*1970年生まれ。1994年上智大学経済学部卒。日本銀行を経て,大阪大学大学院修了(経済学博士)。大阪大学社会経済研究所助手,(社)日本経済研究センター研究員を経て,2002年より現職。(社)日本経済研究センター副主任研究員を兼任。主な論文に,「医療技術評価に対するCVM(Contingent Valuation Method)の適用可能性:サーベイ・データによるWTPとWTAの乖離要因の分析」『医療と社会』Vol. 12 No. 2,「医療需要行動のConjoint Analysis」『医療と社会』Vol. 10 No. 1,「ニコチン代替療法需要のコンジョイント分析」『日本経済研究』49などがある。連絡先:〒560 0043 大阪府豊中市待兼山町1 31 大阪大学国際公共政策研究科 suzuki@osipp.osaka-u.ac.jp

政策評価と政策運営の関係が明確化しているという点も、わが国の経済政策としては極めて異例な、画期的特徴である。

さて、規制緩和の全国展開を政策評価によって決定するということは、評価実施者が規制の生殺与奪の権を握る重要な役割を果たすということである。したがって、特区の評価は、国民全体が納得し得る客観的かつ科学的な手法が用いられ、透明かつ公正な政策評価プロセスに基づいて実施される必要がある。現在、構造改革特区の評価は構造改革特別区推進本部に設けられた「評価委員会」が統括して評価を行い、その意見を元に本部がその後の措置を決めることになっている。評価委員会は第三者的な立場を保つために、外部の学識経験者や民間事業者から選ばれた10人の委員から構成され、そのメンバーのうち3人は公募で選ばれている。評価委員会は、医療・福祉・労働部会、教育部会、農村活性部会、エネルギー・安全部会、国土・物流部会、産業振興部会、国際交流部会、地域活性部会の8つの専門部会に分かれ、各分野の政策評価を行うが、その評価のプロセスはおよそ次のようなものである¹⁾。まず、特区において実施されている規制緩和の規制所管省庁は、具体的なデータやアンケートにより規制緩和による弊害の調査を行い、その結果を本部に報告する。評価委員会は、この規制所管省庁の調査に加えて、規制の特例措置を全国展開することによる効果等について独自の調査を行い、弊害の発生について検証した上で、特段の問題が生じているか否かについて評価する。評価委員会が答申する評価意見は、

- 「ア）地域を限定することなく全国において実施」、
- 「イ）引き続き当該地域特性を有する地域に限定して適用」、
- 「ウ）規制の特例措置の廃止」という3つの判断である。

「ア）地域を限定することなく全国において実施」が適用される判断基準は、

- a) 弊害が生じないと認められる場合、
- b) 弊害が生じていても、規制の特例措置の要件、手続きを見直すことで弊害の予防等の措置が確保され、かつ、見直された予防等の措置について特区における検証を要しないと認められる場合、
- c) 弊害が生じていても比較的微少であり、規制の特例措置を全国展開した場合の効果と比較検討し、効果が著しく大きいと認められる場合

の3つである。一方、「イ）引き続き当該地域特性を有する地域に限定して適用」となる場合の判断基準は、弊害が生じていても、規制の特例措置の要件、手続きを見直すことで弊害の予防等の措置が確保され、是正又は追加された予防等の措置について特区における検証を要すると認められる場合ということであり、「ウ）規制の特例措置の廃止」となる場合は、弊害が生じており、かつ、規制の特例措置の要件、手続きを見直すことで予防等の措置を確保することが困難と認められる場合、となっている。

この特区の評価プロセスの特徴は、「特段の問題が生じていないと判断されたものについては、速やかに全国規模の規制改革につなげる」(「経済財政運営と構造改革に関する基本方針2003」(平成15年6月27日閣議決定))とあるように、全国展開が「特段の問題がなければ」適用されるデフォルト(既定値)となっていることである。通常、予算措置の伴う政策の評価を行う場合には、政策の「有効性」を立証しなければならないのが普通であるが、逆に、弊害がなければ肯定するという方針になっているということは異例であり、極めて規制緩和推進に好都合な制度である²⁾。そして、特段の問題となる規制緩和の弊害を立証する責任は、驚くべきことに、規制所管省庁の側にある³⁾。医療裁判や公害裁判で見られるように、

1) 「構造改革特区の評価方針について(案)」平成15年12月18日評価委員会資料、「構造改革特別区域基本方針」等による。

2) これは規制緩和が基本的に予算を伴わないということが、一つの正当化理由であると思われる。もちろん、有効性については評価委員会が行う独自の調査が検証するので、有効性を評価しないということではない。

立証責任を負う立場のものは一般的に不利であるから、これも規制緩和に極めて好都合な枠組みである。さらに、評価の期間も、16年度の上期には第一回の評価が行われるという短さであり、この点も規制緩和に積極的に働くものと思われる。

ところで、そのような評価の核心となる評価手法は一体どのようなものなのであろうか。科学的で客観的な評価が期待できるのであろうか。残念ながら、現在、評価委員会がホームページ上⁴⁾で公開している評価手法の資料や計画を見る限り、筆者は不安を覚えざるを得ない。現在、各専門部会ごとに①16年度上半期の調査計画(案)⁵⁾、②規制所管省庁の調査計画⁶⁾、および評価委員会の独自調査となると考えられる③評価調査手法の例(イメージ)(民間調査会社作成中資料)⁷⁾が入手可能である。これらは、まだ最終的なものではなく、実際に変更される余地があるだろうが、総じて見て、プログラム政策評価の標準的手法からあまりにかけ離れており、以下の節で述べるように問題が多いと思われる。例えば、通常、政策評価を行うためには政策が実施された地域のみならず、比較対照として政策が実施されていない地域の調査が必要であるが、調査計画は特区のみを対象としているようである。また、政策評価に用いる指標も、規制緩和が実施された件数等の実施指標が中心であり、政策目標を図るべき成果指標が設定されていない場合もある。さらに、成果指標が設定されている場合にも、満足度調査等の極めて客観性の乏しい指標で評価しようとしている事業も数多く見受けられる。評価委員会が行う独自調査は、規制緩和を全国展開することによる効果を評価することになっているが、ある特定地域における計測結果が全国においても適用可能であるかどうかを判断する為には、後述のように、一般均衡効果などの様々な手法的課題を克服しなければならない。しかしながら、そのような課題に対処しようとしているとは少なくとも上記資料を見る限り思えない。さらに、規制緩和の評価に当たって政治的に難しい点は、ほとんどの場合において、規制緩和によって既得権益を奪われる対象が存在することである。その場合、規制緩和によって既得権益者がこうむる損失と、新たに発生する便益を比較する必要があると思われるが、そのような損失と便益を直接比較するための調査手法も考慮されていないようである。実際に、「ア)地域を限定することなく全国において実施」が適用される判断基準のうち、「c)弊害が生じていても比較的微小であり、規制の特例措置を全国展開した場合の効果と比較検討し、効果が著しく大きい」という基準が当てはめられるケースも多いと想像されるが、比較検討をどのように行うのか不明であり、実効性が懸念される。

さて、本稿はそのような問題意識の下、経済学で用いられているプログラム政策評価の計量手法の観点から、特区の評価をどのように設計すべきか考察することにする。以下、2節において標準的なプログラム政策評価の手法を紹介した上で、3節では特区評価に固有の課題を考察し、どのような調査を設計すべきかを論じる。4節は、具体的に各専門部会で公開されている3つの事業の評価例をケーススタディーとして紹介し、その問題点と、改善点を論じる。5節は結語である。

3) 「構造改革特別区域基本方針」には、評価の具体的方法として「規制所管省庁は、規制の特例措置の適用状況の調査に当たって、特区において適用された規制の特例措置による弊害の発生の有無に基づき、全国展開により発生する弊害について立証責任を有するものとする。」とされている。

4) <http://www.kantei.go.jp/jp/singi/kouzou2/hyouka.html>

5) 各部会第4回会合資料。

6) 各部会第4回会合資料。

7) 各部会第3回会合資料。

2. プログラム政策評価の標準的手法⁸⁾

2.1 Difference Estimatorと内生性

ある施策の評価を考えた場合、理想的な方法は「実験(experiment)」を行って評価を行うことである。実際、医学や心理学や自然科学では実験を行って施策の評価が行われている。例えば、医学における新薬の評価では、実験する患者に対して、新薬と効果の無いプラセボ(placebo)と言う擬似薬が無作為に配られ、その効果の差を検証することになる。無作為配布の方法は、二重盲検試験(double blind method⁹⁾)と呼ばれる厳密なものであり、患者はもちろん、配布者にもどれが新薬かプラセボかがわからないようにしている。経済学においてはこのような厳密な実験を行うことは困難ではあるが、それでも米国においては、著名なNegative Income Tax Experimentやランドによる医療保険が医療需要行動に与える影響を調査したRand Health Insurance Experiment、少人数教育や補助金の効果を調査したSTAR(Student Teacher Achievement Ratio) Projectなど、しばしば社会実験(Social Experiment)と呼ばれる無作為な対象者に対する大規模な実験調査が行われている。

こうした実験調査の評価に際しては、まず施策の目標となる成果指標(outcome)が設定される。例えば、教育に対する施策の効果を評価するSTAR Projectでは、学力テストの点数(Stanford Achievement Test)、Randの調査では保険によって影響される受診行動(外来日数、入院日数、医療費)が設定されている。そして、調査対象を、施策(Treatment)を行うグループ(Treatment Group)と、施策を行わない比較対照グループ(Control Group)に無作為に分け、両グループの成果指標の差を比較して評価を行う。その効果の計測を行うもっとも簡単な手法は、成果指標を Y_i 、施策の実施レベルを X_i と、その他の成果を決める観察可能な要因を $W_{i,j}$ として、次のような回帰分析を行うことである。

$$Y_i = \alpha_0 + \alpha_1 X_i + \sum_j \beta_j W_{i,j} + u_i \quad (1)$$

今、例えば新たな教育プログラムという施策を実施すると考えると、 Y_i は学生の成績、 X_i は教育プログラムを実施する対象を1、実施しない対象を0とするダミー変数、 $W_{i,j}$ はもともとの成績を決めている親の所得、親の学歴、私学か公立か、塾通いの有無等が考えられる。この場合の誤差項 u_i は、成果 Y_i を決める観察不可能な要因を含んでいる。観察不可能な要因とは、この例では、例えば学生のIQ等の学習能力ややる気といったものである。もし、誤差項 u_i がConditional Mean Independence Assumption(CMIS)と呼ばれる次のような条件を満たすのであれば、 α_1 はバイアス無く推定されることになる。

$$\text{CMIS } E(u_i | X_i, W_{i,j}) = \alpha_0 + \sum_j \beta_j W_{i,j} \quad (2)$$

8) 本節は、経済学の分野で行われているプログラム政策評価の計量的な手法の初歩的な紹介を行う。教科書レベルの紹介としては、Heckman and Smith (1995)、Wooldridge (2001)、Stock and Watson (2003)等が挙げられる。邦文では、富岡淳(2004)が恐らくはじめての包括的なサーベイである。経済学で行われている狭義の政策評価ではなく、より広義の政策評価について、特に特区と絡めて解説している論文として、小野(2003)が挙げられる。

9) これは自分がプラセボか新薬のどちらが配られているかわかることにより、心理的な効果が変わり、厳密な実験にならないからである。社会実験においても、自分が効果計測対象となっていることがわかると、心理的な効果が作用してしまうバイアスが生じることが知られている(Hawthorne Effect)。

CMISは、施策以外の要因 $W_{i,j}$ をコントロールした上で u_i が X_i と関連していないということであり、 X_i が要因 $W_{i,j}$ ごとに無作為に振られていれば満たされる。これは例えば、この教育プログラムの例で言えば、私立校か公立校かという要因などが考えられる。私立校ダミー W_i は明らかに学生の成績に影響を与えているが、私立校、公立校のそれぞれで対象者を無作為に抽出していればCMISが満たされる¹⁰⁾。この時、施策の効果 (Treatment Effectもしくは、Causal Effect) は、

$$E(Y_i | X_i = 1, W_{i,j}) - E(Y_i | X_i = 0, W_{i,j}) = \tau \quad (3)$$

と定義され、OLSを用いてバイアス無く推定される。これをプログラム政策評価の計量分析では、Difference Estimatorと呼んでいる。Difference Estimatorは、逆に言えば、CMISが満たされない場合にはバイアスを持ってしまう。例えば、施策 X_i が無作為ではなく、個人の応募によって決定されるといった場合、学生のやる気や学習能力が含まれる u_i と X_i が関連すると考えられる。この時、施策の効果 τ はバイアスをもって過大もしくは過少に推定されてしまう。しかしながら、社会実験ではなく、通常、経済学が対象とするような政策プログラムの場合には、希望者に対して施策が行われるという形式が通常であり、この問題を避けることが非常に困難である。この問題は、内生性問題 (endogeneity problem) もしくはselection biasと知られている。内生性の問題は、無作為に対象者を割り当てた社会実験においても生じ得る。例えば、新教育プログラムを割り当てられたとしても、実際には転校などによりプログラム受講を拒否する家庭もあるだろうし (Partial Compliance)、やる気や能力が無くて新教育プログラムから脱落することがありえる (Attrition)。この場合には、やはり実際のプログラム受講 X_i と、誤差項 u_i が関連してしまうので、Difference Estimatorのバイアスは大きい可能性がある (Heckman and Smith, 1995)。

さて、このような内生性問題に対するもっとも教科書的な対処法は、操作変数法を用いることである。つまり、 X_i と関連し、 $E(u_i | Z_i) = 0$ となる操作変数を見つけ、それを用いて推定することにより、バイアスの無い推定が可能である。例えば、新教育プログラムのPartial ComplianceやAttritionの例では、実際の受講 X_i が u_i と関連を持ったとしても、そもそもの無作為の割り当て状況を操作変数 Z_i として用いることができることから、比較的対処は容易である。しかしながら、希望に応じたプログラム参加のような内生性問題では、操作変数を見つけること自体が非常に困難であるため、対処が難しい¹¹⁾。また、類似の方法として X_i をモデル化して同時方程式で推定するという方法もあるが、モデル設定が推定結果に大きく影響するため、適切なモデルをどう選択するかという新たな問題が生じることとなる。もう一つの対処法は、新教育プログラム希望者の中から無作為に受講者を選ぶという方法である。これであれば、プログラム参加と能力ややる気といった観察不可能な要因が無関係になるので、バイアスが生じないことになる。しかしながら、問題は、ここで得られる効果の計測結果は、「希望者」に対するものであり、希望者以外を論じることができないという点である。つまり、特区のようにプログラムを非希望者にも広げることが想定している場合には、非希望者への効果に対する情報は全く得られないことになる¹²⁾。

10) もちろん、 $W_{i,j}$ に全く無関係に無作為抽出してもこの条件は満たされる。

11) また、操作変数と X_i と関連が低い場合にはかえってバイアスが生じることも知られている (Staiger and Stock)。また、これまで自然実験を利用してかなりうまく操作変数を設定したとして知られるいくつかの業績 (Angrist and Krueger (1991) や Card and Sullivan (1988)) に対しても貧弱な操作変数の問題が生じていることが指摘されている (Rosenzweig and Wolpin (2000))。

12) そのほか、観察不可能な要因と相関する観察可能な変数を使って観察不可能な要因をノンパラメトリックにコントロールするという方法も研究されており、成果を挙げつつあるが、本稿では高度なトピックスになるので触れない。詳細は Heckman, Ichimura and Todd (1997, 1998) を参照されたい。

2.2 Difference in Difference Estimator

このように、社会実験が行えないような通常の政策プログラムの評価では内生性問題に対処するのはなかなか困難であるが、通常の政策プログラムの中には、施策実施に対して個人が影響を及ぼせないような外生的な施策が行われることがある。例えば特区内の全ての人に及ぼされるプログラムはその典型的な例であるが、特区内の個人にとっては選択の余地無く選ばされるので施策は外生的である。このようなケースは、自然実験 (Natural Experiment) や疑似実験 (Quasi Experiment) と呼ばれており、特区内の対象者 (Treatment Group) と、特区外の非対象者 (Control Group) の施策開始の前後の情報を利用して、Difference in Difference Estimatorと呼ばれる方法を用いてバイアスのない効果が計測できる¹³⁾。Difference in Difference Estimatorでは、施策前後の同一個人¹⁴⁾の成果指標の変化差 (Y_i) を利用して、(4)式のような推定式を推定する。

$$Y_i = \alpha_0 + \alpha_1 X_i + \sum_j W_{i,j} + u_i \quad (4)$$

Difference in Difference Estimatorは、外生的な施策実施前後における同一個人の行動の差を、そのような施策の対象とならなかった同質なグループにおける同時期の差と比較していることになる。つまり、 α_1 は施策対象のTreatment Groupにおける成果指標の前後差 ($Y_{i,treatmentt}$) から施策非対象のControl Groupにおける成果指標の前後差 ($Y_{i,control}$) を引いたものと等しくなる ($\alpha_1 = Y_{i,treatmentt} - Y_{i,control}$)。Difference in Difference Estimatorの利点は、同一個人の異時点間の比較によってその個人の固有効果 (観察不可能な能力ややる気を含む) を除去し、また対象群と比較することによって、その時期に生じた社会全体の变化を完全に制御することができる点にある。したがって、Difference Estimatorで問題となった内生性問題に対処できるのである。しかしながら、施策実施が個人あるいは企業などの主体にとって外生的ではない場合には、 u_i が X_i と相関してこの場合のCMISが満たされなくなるので、バイアスが生じてしまう。これは施策実施の情報を事前に知って、施策前の行動を変化させている場合でもバイアスが生じることを意味する。また、 $W_{i,j}$ ごとにTreatment GroupとControl Groupが同質的であるという条件もCMISにより課されることになる。

2.3 その他の問題点

さて、上記のようにうまく自然実験を拾ってDifference in Difference Estimatorを得たり、あるいは操作変数を適切に設定してDifference Estimatorを得たりできた場合においても、その結果が特定地域以外に適用できるかという問題は別問題である。例えば、実験 (自然実験) 地域や対象が非常に特殊な地域である場合には、普遍化が困難である (Nonrepresentative sample)。その場合でも十分にサンプルが多い場合には、その中からサンプルを再抽出するBootstrap等の方法を用いて、普遍化する対象 (例えば全国) と同質なサンプルを擬似的に作って推定値を修正することができる。しかしながら、サンプルが少ない場合にはそのような修正も困難であるし、そもそも小サンプルの推定では特定地域の推定値も信用できなくなってしまう。

もうひとつ、深刻な問題は一般均衡効果 (General Equilibrium Treatment Effects) と呼ばれるもので

13) わが国におけるDifference in Difference Estimatorの利用例は最近急増しつつあるが、最近の例としては、安部(1997)、大石(2003)、大竹・山鹿(2003)、鈴木(2004)等がある。

14) 同一個人を用いずに、施策前後でrepeated cross section dataをとって比較するDifference in Difference Estimatorも存在する。

ある。例えば、ある実験（自然実験）において、職業訓練の実施がその後の就業確率や賃金を上昇するという効果が計測されたとしよう。この効果は実は近隣地域などのプログラム非参加者の就業機会を奪った結果をみているだけなのかもしれない（displacement effect）。この場合、全国での効果を見る場合には、明らかに就業を奪われた人を考慮して効果を判断する必要がある。また、ある地域で賃金が増えたとしても、全国で同様の職業訓練プログラムが実施されれば、そのような訓練を受けた人の供給が増えるので、計測されたよりも賃金上昇率はずっと低くなる可能性がある。政策評価は地域的かつ短期的で部分均衡の効果を計測している場合が多いから、このように一般均衡としてみた場合には計測された効果が意外に大きなバイアスを持っていると考えられる（Heckman, Lochner and Taber, 1998）。この場合は、応用一般均衡分析モデル（CGE）を使って評価するなどの対処法が一応考えられるが、現行のCGEには定量的な評価に耐えられるほどの精度を持たせることは困難であり、計測値を定量的に修正することは今後の学術的発展を要する課題である。また、全国適用をしようとした途端、人々の期待が変更されて、これまでの計測結果が変化するというLucus Critiqueの問題も忘れるわけにはいかない（Lucus, 1976）。

3．構造改革特区の政策評価をどのように設計すべきか

次に、前節の標準的な評価手法を特区に当てはめる場合に、どのような課題や注意点があり、どのように対処すべきであるかを考察する。

3.1 内生問題への対処

まず、特区評価に当たってもっとも大きな問題は、社会実験とは異なり、特区認定が無作為に割り当てられたものではなく、各地域の自主的申請を元に認定されているという点である。ただし、この問題は、評価の対象が個人や企業といった特区内の個々の主体である場合には、それほど問題とはならない。なぜなら、個人や企業にとっては、特区による規制緩和は個人や企業の影響が及ばない外生的な制度変更とみなすことができ、Difference in Difference Estimatorを用いることができるからである。つまり、特区外の非規制緩和地域の個人や企業とともに、特区認定前後のデータを収集して、(4)式を計測すれば良い。しかしながら、いくつか注意すべき点がある。第一に、太田市の外国語教育特区のように特区外から応募ができるようなものについては、特区外からの応募者は明らかにプログラム参加が内生変数となるために、特区内の応募者に限り計測を行うべきである。第二に特区内の個人や企業にとって、規制緩和を受けるか受けないかという選択が可能な場合の処理である。例えば、三歳未満児に掛かる幼稚園入園事業等については、特区内の人々でもその事業を利用するか否かという選択ができるが、その場合にも同様の内生問題が発生する。この場合の対処方法としては、特区内の対象者が否かという変数を操作変数として、操作変数法により推定をすることが考えられる。第三は、産業振興関係の特区などでは、規制緩和を要望する企業がそもそも存在するから特区を申請したというように、特区申請が外生的な制度変更とみなせない場合があるということである。この場合には、特区認定直後にプラントを開始したといった内生性が疑われる企業を除いて計測を行うなどの工夫が必要になると思われる。第四に、制度改正前として抽出するデータは、特区認定前の時期のデータであっても、特区認定が行われることが予想されている望ましい推定ができない。したがって、予想が形成される以前の、したがってかなり以前の長期データを収集することが望ましい。

しかしながら、特区申請が内生的であることも問題はむしろ個人や企業ではなく、港湾物流特区や空港

特区など、特区全体で一つの対象を観察するしかないような場合である。この場合は、明らかに、物流関係を自由化したいというやる気や能力がある自治体が申請をするために、「プログラム参加」が内生的となる。この場合の対処は、既に2.1で触れたように、操作変数を考えたり、プログラム参加自体を決めるモデルを同時推定する方法があるが、現実的に適切な操作変数やモデルを設定することは極めて困難であると想像される。残された方法は、特区申請をした自治体の中から、特区認定を無作為に割り当てて「認定時期」をずらし、認定された自治体と認定されなかった自治体の両者からデータを収集できるようにすることである。この場合には、やる気や能力に無関係に特区認定が行われることになるから、内生問題が生じない。もっとも、ここで計測される効果は、「やる気や能力が有る場合の」効果であって、やる気や能力がない自治体に規制緩和を適用した場合の効果つまり全国適用の効果を計測できることにはならない。したがって、できることならば、こうした特区認定をする際に、自主的な申請をしなかった自治体に対しても、特区認定を割り当てるといような措置が必要なのかもしれない。

3.2 一般均衡効果への対処

2.3で触れたように、一般均衡効果は様々なものがあるが、特区評価の場合に特に深刻なものは、近隣地区から効果を奪うDisplacement effectである。例えば、宇都宮にぎわい特区のように、大店法の規制緩和がある地区だけで行われると、それができない隣接市区町村の顧客が奪われて、特区の効果が過大に出ることが予想される。この場合、全国に規制緩和が普及した後では、Displacement effectがなくなるのであるから、その効果を除いた推計を行わなければならない。その場合、考えられるのは、顧客が及び得る近隣地区も全て含めた全域でデータを収集して、効果を計測するということである。これは特区の種類によっては近隣地域にとどまらない。例えば、港湾物流特区では、かなり離れた地域であっても他の同規模の港湾を含めるべきかもしれないし、空港特区ではもっと地理的に離れた空港をデータに採るべきかも知れない。また、この問題は、Displacement effectの除去というだけでなく、Control Groupの設定の問題でもあり得る。つまり、特区の効果を見る場合には、適切な非特区地域のデータを比較対照群(Control Group)として選ばなければならないが、通常は同規模の近隣地区が自然条件なども同質的なため、選択されやすいものと考えられる。しかしながら、Displacement effectが及び得る近隣地区を選んだ場合には、Control Groupの対照効果が非常に低くなっているという意味で、二重に特区の効果を過大に計測するバイアスが生じていることになる。このような場合、Control Groupとして近隣地区を設定することはできない。

3.3 弊害と効果の大きさをどのように比較考量するか

恐らくどのような規制緩和であっても、規制緩和により被害を受ける既得権益グループが存在するのが通常であり、その意味で弊害が全く想定されない規制緩和というものほとんどありえないであろう。問題は、規制緩和の効果と弊害をどのように比較考量して、ネットの効果(正の効果と負の効果を差し引いた純効果)を評価するかである。評価委員会の資料では、「c) 弊害が生じていても比較的微小であり、規制の特例措置を全国展開した場合の効果と比較検討し、効果が著しく大きい」場合には、「ア) 地域を限定することなく全国において実施」としているが、具体的にどのように比較検討されるかは不明である。実際、各専門部会における調査計画などをみても、規制所管省庁が行う弊害の調査内容と専門部会や委託民間調査会社が行う調査内容が直接的に比較可能でない場合が多い。また、ほとんどの場合、規制緩和の効果と弊害は直接比較不能である。この場合、何らかの「総合判断」を行う必要があるが、主観的な判断

にならざるを得ず、説得力に欠けるものになる可能性が高い。

規制緩和の効果と弊害を比較可能にするには、一度、全ての効果・弊害指標を比較可能な「総合指標」に置き換えることが必要である。そのもっとも有力な候補は経済効果（不効果）としての金銭価値額である。規制緩和がコスト削減や付加価値額に反映されるものであればその効果を金銭価値にすることは可能であるし、既得権益者の損失も同様である。また、消費者や既得権益者の効用（不効用）の価値も、CVM（Contingent Valuation Method）と呼ばれる選好表明法を用いて便益（損失）額にすることが可能である。CVMとは、市場で取引されない財や仮想的なプロジェクトに関して、その金銭的な価値を回答者に表明させる手法であり、経済学の中では主に環境経済学や公共経済学・実験経済学等の分野で発展を遂げてきた。その代表的なものは、WTP（Willingness to Pay：支払い意思額）やWTA（Willingness to Accept：受け取り意思額）、CA（Conjoint Analysis）である¹⁵⁾。問題は、戦略的バイアスと呼ばれるように、金銭価値を表明するときに、その表明額が実際に政策の意思決定に反映されることがわかると、得をするものは効果を過大に表明し、損をするものは効果を過少に表明する可能性がある点である。しかしながら、例えば、特区内での選好表明額は、その特区の改廃には用いず、特区の規制緩和を他の自治体や全国に適用する際のみ用いるという制度にすれば、少なくとも特区内の既得権益者や便益者の利害には関係がなくなることから、戦略バイアスが含まれないように調査設計することは可能であると思われる。また、CVMは実際に規制緩和が行われていない地域でも調査が可能であることから、他地域で調査を行って、全国適用が行われた場合の効果をよりの確に計算することが可能である¹⁶⁾。

4．各分野の評価手法の評価

最後に、ケーススタディー的に平成16年上期に計画されているいくつかの評価計画例をとり上げ、問題点の指摘と改善策を探ることとする。既に1節で述べたように、評価委員会の資料として、①16年度上半期の調査計画（案）、②規制所管省庁の調査計画、③評価調査手法の例（イメージ）（民間調査会社作成中資料）が公開されている。①から③までの資料が全てそろっている事業のうち、「官民共同窓口の設置による職業紹介事業（医療・福祉・労働部会）、「中心市街地における商業の活性化事業（地域活性化事業）、「臨時開庁手数料の軽減による貿易促進事業」（国土物流部会）の3つをとり上げる。

4.1 官民共同窓口の設置による職業紹介事業

これは第一回の認定で認められた東京都足立区「人材ビジネスを活用した雇用創出特区」で行われる事業であり、足立区役所にハローワーク足立と民間の職業紹介業者が同時に窓口を設置し、互いに補いながら職業紹介を行うという事業である。

これに対して、評価委員会が行なう「効果等についての調査」と規制所管官庁が行う「弊害についての調査」の2つが実施される。さらに、評価委員会「効果等についての調査」は、事務局が行う定期的な「地方公共団体を通じての調査」と、必要な場合に行う「事業者・需要者・消費者に対する直接調査」の2つがある。「地方公共団体を通じての調査」は、他のほとんどの事業と共通した定型的なものであり、①進

15) CVMを解説した文献は多いが、例えば医療分野の適用をしたものとして、辻・鈴木ほか（2002）や鈴木・大日（2000）等がある。

16) これを便益移転（Benefit Transfer）可能性という。WTP等の評価では、複数のサンプル地域の調査結果から便益移転可能性がチェックされた上で、サンプル地域から計算された便益移転関数を用いて、（アンケート調査ができない）全域の便益額評価が行われる。特区の場合においても、そのようなプロセスが必要となるのかもしれない。

捗状況を示す指標，②進捗の自己評価，③具体的な成果の3点を記述することになっている。この事業では具体的に，①の進捗状況としては，官民共同窓口の利用者数とその雇用契約件数¹⁷⁾を平成16年4月時点と（それ以前の比較数値がある場合には）それ以前の数値を記述することになっている。②進捗の自己評価は，A：目的がほぼ達成された，又は順調に進んでいる，B：問題はあるが進展している，C：始まったばかりであり，まだ時間を要する，D：準備段階であり具体的な取り組みはこれから，E：弊害の発生等により止まっているの5段階評価が行われる。③具体的な成果は特に具体的な指示が無く，成果を記述するように指示がある¹⁸⁾。「事業者・需要者・消費者に対する直接調査」については「評価調査手法の例（イメージ）（民間調査会社作成中資料）」によれば，民間職業紹介事業者と求職者を対象にして，①事業を利用した求職者数と，②そのうちの就職に至った割合，③本事業に対する事業者と求職者の満足度を，平成16年第一四半期から毎年実施するということである。一方，規制所管官庁「弊害についての調査」は，民間事業者の窓口担当者，ハローワークの窓口担当者，及びそれぞれの窓口での求職者，求人者，地方公共団体へのアンケート及びヒアリング調査が行われる。調査項目は，紹介の取扱状況のほか，苦情の状況，利用者の反応，プライバシー配慮の措置，待ち時間，利用者の希望，財政負担の状況，メリットデメリット項目の列挙などとなっている。

さて，2節3の議論からこの調査内容をみると，次の点が問題あるいは改善点として指摘できる。第一に，区役所内の共通設置窓口だけではなく，民間事業者，ハローワークを含めた足立区全体の職業紹介状況を把握すべきである。また，特区の前の比較可能なデータとしては，特区認定が予想される前の1，2年前からのデータ（足立区全体の紹介状況と成功率）を収集しなければならない。第二に，比較対照のために，足立区以外の地区についても，共通の項目の調査を行うべきである。これは近隣地区のDisplacement effectを考えると，近隣地区と近隣ではない地区の両者からデータを収集することが望ましい。第三に，満足度調査は，前後の比較や共通窓口以外との比較ができないようであればあまり意味がない。第四に，効果と弊害を金銭価値で比較可能とするために，弊害として指摘されたプライバシー侵害が行われた場合のWTA（Willingness to Accept）評価値の調査，ハローワーク適正配置ができなくなることの件費コストの増加，職業紹介成功者の収入を調査することも一案である。

4.2 中心市街地における商業の活性化事業

この事業は，宇都宮市の中心市街及びJR宇都宮西口地区において，大店立地法の手続きを簡素化し，大型店の立地促進を図るというものである。まず，評価委員会事務局が行う「地方公共団体を通じての調査」は，①進捗状況を示す指標として出店店舗数，②5段階の進捗の自己評価の2項目がある。一方，「事業者・需要者・消費者に対する直接調査」は，「評価調査手法の例（イメージ）（民間調査会社作成中資料）」によれば，商工会議所に行う①大店設置状況（名称，時期，規模，効果），②周辺生活環境への影響，③大店設置者に対する特例により設置意欲が高まったかどうかの意識調査を行うとされている。所管規制省庁の弊害調査は，周辺商工会議所，商工会等に対して，「大店立地法が大規模小売店舗の設置者に求めている周辺生活環境に対する弊害事項に関して，何か弊害がありましたか？，もし，なんらかの弊害が発生したのであれば，それは何が原因だと思いますか？，また，その発生を防止するためには，何が必要だと思いますか？（自由回答）」というものである。

17) それ以外にも効果を表す適当な指標がある場合にはそれをあげるようにとの指示がある。

18) 記入例として「（例）特例措置の活用による経済効果」とある。

これらの評価手法についてまず言えることは、評価の時期が16年上期では早すぎるのではないかということである。大店設置業者への意識調査ではあまり厳正な評価は期待できない。また、宇都宮は大規模小売店が次々に撤退したためにこの特区を申請したという経緯があるため、出発点が非常に低いところであり、大店舗が出店した場合の効果を大きく計測しやすい状況である点にも注意が必要である¹⁹⁾。さて、その上で、2節3節に照らして考えると、第一に、宇都宮以外にもControl Groupとして比較地域を選んで調査を行うべきであることが指摘できる。同様に、近隣地区のDisplacement effectが予想されるために、近隣地区の調査も別途必要である。第二に、弊害と効果を直接比較するために、大店舗の売上額や小売価格、地元商店街の売上減少額や小売価格といった客観的な指標も評価する必要があると考えられる。地元住民に対するWTPを聞くことも一案である。意識や実績店舗数などの指標では、最終的な政策目標の評価にはならないし、この場合特に重要と思われる弊害との比較考量も困難であるからである。

4.3 臨時開庁手数料の軽減による貿易促進事業

この事業は、横浜、神戸などの港湾が認定されたものであり、税関の執務時間外における通関体制の整備とあいまって、貿易流通量の促進を目指している。まず、評価委員会事務局が行う「地方公共団体を通じての調査」は、①進捗状況を示す指標として臨時開庁申請数と、②5段階の進捗の自己評価の2項目である。一方、「事業者・需要者・消費者に対する直接調査」は、「評価調査手法の例(イメージ)(民間調査会社作成中資料)」によれば、税関に対して、①特例による時間外通関の増加量、②特例によるコスト・収入・収支のバランス状況、貿易会社に対する③利用状況・利用理由、④コスト・リードタイムの変化、⑤より時間外通関を利用するための条件・意向、通関業者に対する⑥利用状況・利用理由、⑦より時間外通関を利用するための条件・意向、⑧特例によるコスト・収入・収支のバランス状況を調べるとしている。所管規制省庁の弊害調査は、手数料割引による手数料減少額のみである。

この事業の場合は、評価の対象となる指標が、通関の利用増加量・額及び貿易増加額と明確であり、それに対する弊害も手数料減少額と直接比較が可能であるため、とても評価が行いやすいと考えられる。ただし、調査対象は当該事業認定特区だけではなく、他地域の港湾も同様に調査を行う必要がある。ひょっとすると、韓国等の外国の状況も調査の対象とするのが適切なのかもしれない。また、Displacement effectは近隣に止まらず存在すると予想される。

5 . 結語

本稿は、16年上期から実施される特区の評価に対して、近年発展の著しいプログラム評価の計量分析の観点から、そのあり方を議論した。特区の評価は、構造改革特別区推進本部に設けられた「評価委員会」が統括して評価を行うが、極めて規制緩和に有利な枠組みとなっている。しかしながら、慎重に科学的かつ客観的な評価が実施されるかどうかは不明である。現在、公表されている資料からは、①特区の評価となる最終的な評価指標の選定が適切ではない場合がある、②特区内の指標だけではなく、比較対照となる地域の指標を集める必要がある、③Displacement effectが予想される場合には近隣地区の影響も調査する必要がある、④弊害と効果を比較考量するために、各評価指標を金銭価値額に置き換えることも一案である、などの点が指摘できる。特区は、単なる地域活性化策ではなく、全国の規制緩和を進めることが目標

19) こうした効果は、職業訓練プログラムにおいても存在しており、Ashenfelter & dipとして知られている。

となっている。その全国展開をするか否かの決定を行うのが、評価の役割であり、したがって単なる政策評価以上の重要性を帯びている。規制緩和が一部の既得権益者の意見に矮小化されること無く、国民的なコンセンサスの中で進められることを願ってやまない。そのためにも、国民全体が納得できるような、客観的かつ科学的な評価が行われる必要があるだろう。

[参考文献]

- 小野達也 (2003) 「構造改革特区が日本の経済社会を活性化する条件 社会実験としての構造改革特区」 富士通総研研究レポート 177
- 辻正次・鈴木亘・田岡文夫・手嶋正章 (2002) 「医療技術評価に対するCVM (Contingent Valuation Method) の適用可能性: サーベイ・データによるWTPとWTAの乖離要因の分析」 『医療と社会』 (財団法人医療科学研究所) Vol. 12 No. 2, pp. 107-119
- 富岡淳 (2004) 「政策評価Program Evaluationの計量経済学: サーヴェイ」 日本経済研究センター 『行政評価研究会報告書』
- 安部由起子 (1997) 「1980~1990年代の男性高齢者の労働供給と在職老齢年金制度」 『日本経済研究』 36, pp. 1-32。
- 大石亜希子 (2003) 「有配偶女性の労働供給と税制・社会保障制度」 『季刊社会保障研究』 Vol. 39 No. 32, pp. 86-300
- 鈴木亘 (2004) 「レセプトデータを用いたわが国の医療需要の分析と医療制度改革の効果に関する再検証」 『日医総研ワーキングペーパー』 97
- 大竹文雄・山鹿久木 (2003) 「在職老齢年金制度と男子高齢者の労働供給」 『選択の時代の社会保障』 清家篤・田近栄治編, 東京大学出版会
- 浜田宏一 (2004) 「特区の経済的意義」 『法と経済学機関誌』 第1巻第1号
- 八代尚宏 (2003) 『規制改革「法と経済学」からの提言』 有斐閣
- Angrist, J., 1990, "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administration Records," *American Economic Review*, 80 (3) 313-335.
- Angrist, Joshua and Krueger, Alan, B., 1991, "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106, pp. 979-1014.
- Card, David and Sullivan, Daniel 1988 "Measuring the Effect of Subsidized Training Programs on Movements in and out of Employment," *Econometrica*, 56, pp. 497-530.
- Heckman, J., 1997, "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations," *Journal of Human Resources*, 32 (3) 441-462.
- Heckman, J., 2000, "Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective," *The Quarterly Journal of Economics*, 115 (1) 45-97.
- Heckman, J., 2001a, "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture," *Journal of Political Economy*, 109 (4) 673-748.
- Heckman, J., 2001b, "Accounting for Heterogeneity, Diversity and General Equilibrium in Social Policy Evaluation," *Economic Journal*, 111, F654-699.
- Heckman, J., Lalonde, R., and Smith, J., 1999, "The Economics and Econometrics of Active Labor Market Programs," *Handbook of Labor Economics*, Volume 3, Ashenfelter, A. and D. Card, eds., Am-

sterdam: Elsevier Science.

- Heckman, J., Lochner, L. and Taber, C., 1998, " General Equilibrium Treatment Effects: A Study of Tuition Policy , " *American Economic Review* , 88 (2) 381 386.
- Heckman, J. and Smith, J., 1995, " Assessing The Case For Randomized Social Experiments , " *Journal of Economic Perspectives* , 9 (2) 85 110.
- Heckman, J.J., H. Ichimura and P. Todd., 1997, " Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program , " *Review of Economic Studies* 64, pp. 605 654.
- Heckman, J.J., H. Ichimura and P. Todd., 1998, " Matching as an Econometric Evaluation Estimator , " *Review of Economic Studies* 65, pp. 261 294.
- Lucas, R., 1976, " Econometric Policy Evaluation: A Critique , " *Carnegie Rochester Conference Series on Public Policy* , 1, 19 46.
- Staiger, D. and J.H. Stock., 1997, " Instrumental Variables Regression with Weak Instruments , " *Econometrica* , pp. 557 586.
- Stock, J.H and M. Watson., 2003, *Introduction to Econometrics* , Addison Wesley
- Rosenzweig, Mark and Wolpin, Kenneth, I., 2000, " ' Natural Experiments ' in Economics , " *Journal of Economic Literature* 38, pp. 827 874.
- Wooldridge, J., 2001, *Econometric Analysis of Cross Section and Panel Data* , Cambridge, MA: MIT Press.