

職業訓練の効果測定における脱落の影響¹⁾

坂本 和 靖

本研究では、「慶應義塾家計パネル調査(2004~2012年)」を用いて、職業訓練が賃金に与える効果に対して、サンプル脱落によるバイアスの影響を分析する。

これまでの先行研究では、職業訓練の賃金への影響を測定する場合、内生性の問題を中心に、如何にして職業訓練の効果以外のものを統制し、純粋な平均処置効果を測定するかが重要視されてきた。本研究ではその点も考慮しつつ、サンプル脱落がもたらす影響についても検証した。

職業訓練受講の内生性を考慮した Inverse Probability Weighting Regression Adjustment 推定(IP-WRA, Wooldridge 2007; 2010)と、加えて、脱落を考慮した、Inverse Probability Weightを用いて、内生性バイアスと脱落バイアスを推定すると、双方とも過大推計を招いていることが確認された。また両バイアスを比較すると、受講直後は前者の方が大きいですが、徐々に後者の方が大きくなった。

最後に、Lee Bounds 推定(Lee 2009)と IPWRA 推定とを比較すると、総じて前者の方が大きな値となった。この理由として、Lee Bounds 推定における仮定(処置の無作為な割り付け)が満たされないことによる過大推定が考えられる。

JEL Classification Codes: C40, I22, J31, J62,

1. はじめに

本稿の目的は、「慶應義塾家計パネル調査(Keio Household Panel Survey: KHPS)」を用いて、職業訓練が賃金に与える効果測定時における、サンプル脱落の影響を分析することにある。

脱落バイアスを対処するために、経済学分野の先行研究において、Inverse Probability Weighting (IPW) 法(Hirano *et al.* 2003, Busso *et al.* 2014)と Bounds 推定(Lee 2009, Angrist *et al.* 2006)の2つの推定法が用いられている。

前者は、脱落発生過程を観測可能な情報から特定できると仮定(Missing At Random: MAR)し、観測可能情報を用いて推定した、継続回答率の逆数を Weight とし、効率的な平均処置効果(Average Treatment on Treated, ATE)の推定を導く(Hirano *et al.* 2003)。現在でも脱落のバイアス補正に関するいくつかの先行研究において、この手法が用いられている(Molina and Macours 2017)。

しかしながら、IPW 法は問題がないわけではない。第一に、脱落、継続回答などの対象者の調査協力に影響し(回答するかどうかの選択方程式には影響する)、かつ結果変数(今回の場合は賃金)に影響しない(行動方程式には影響しない)「除外変数(Ex-

clusion Restriction)」の存在が重要となる。第二に、処置群と対照群との間の結果変数の分布の重複が良好であるときに IPW 法が効果的であるが、重なりが不十分である場合にはうまく機能しない点がある(Busso *et al.* 2014)。

また Bounds 推定は、一般的な点識別(点推定)が当初に厳しい仮定をかけて、それを徐々に緩めていくのとは反対に、Bounds 推定などの部分識別は最低限の仮定をかけて、そこから徐々に仮定を追加していく形をとっている。この推定では、脱落に影響する重要な変数に関する情報が得られるという仮定が緩和され、除外変数を必要としない。誰が脱落するのかについて、それほど厳しくない別の仮定の下で、関心のある処置効果の潜在的な Bounds 推定を行う。さらに共変量を用いることで、より狭い Bound 幅の推定が可能となる(Tauchmann 2013, Ksoll *et al.* 2014)。

本研究では、職業訓練の平均処置効果推定における脱落による推定バイアスを確認するべく、IPW 法による推定結果と、Bounds 推定を用いた計測結果の違いを確認する。

次節以降では、第2節では先行研究について概観し、第3節ではデータについての説明及び、職業訓練と継続回答との関係を確認し、第4節では、Pro-

bit Model を用いて、KHPS における継続回答する対象者の特徴を整理する。そして、第5節以降では、職業訓練の賃金に対する平均処置効果を推定する。第5節では、職業訓練受講の内生性を考慮した、Inverse Probability Weighting Regression Adjustment (IPWRA) 推定と、第4節で推定した継続回答率の予測値を逆数とし、脱落を考慮したモデルと考慮しなかったモデルを比較することで、職業訓練の平均処置効果の内生バイアスと脱落バイアスの測定を行った。第6節では、Lee Bounds 推定による分析結果を、最後に、第7節では主要な結論について論じる。

2. 先行研究

KHPS を用いた脱落や回答拒否に関する研究は、調査開始初期において、多く見られた(宮内他 2005, 2006, 直井 2007, McKenzie *et al.* 2007 など)。先行研究の分析結果によれば、有配偶の継続回答率は高い、特定項目(貯蓄・負債・収入・支出などの自由回答項目)に無回答があったものは翌年以降脱落する確率が高い、前年に無業や臨時雇用であるものほど、継続回答率や項目回答確率が低い(Item Non-response)という結果を得ている。

KHPS の特徴として挙げられるのは、通常の調査データと合わせて、調査員確認票データを提供されている点である。質問票には、訪問日時、接触状況²⁾、対象者または配偶者の調査への協力状況³⁾、回答拒否の理由⁴⁾、拒否の程度⁵⁾などが記載され、詳細な調査が行われている。

一般に調査員はランダムに配置されるため(Random Assignment)、調査員の属性は、調査対象者の行動とは独立と考えられる。調査員の行動は、調査対象者が調査に参加するか否かに強い影響を持つことが知られている(Bates 2003)。つまり、この調査員確認票データの情報は、調査対象者の回答を回収できるかどうかに関与するものの、調査対象者自身の家計行動(調査票で尋ねられている就業行動など)とは関係がないため、適切な除外変数であるといえる。

直井(2007)では、この調査員確認票の情報を活用し、Heckman Model により、脱落を考慮した住居移動行動を分析している。ここで、第一段階の選択方程式の被説明変数(継続回答=1, それ以外=0)に影響するが、第二段階の行動方程式の被説明変数には影響しない除外制約を満たす変数として、調査員

訪問回数、訪問曜日への配慮、初回訪問月を利用している⁶⁾。本研究でも、これらの調査員による継続回答を促す行動(Intensive Follow-up)を活用した分析を行う。

本研究では、脱落バイアスを考える対象として、職業訓練が賃金に与える処置効果推定に着目した。原(2014)によれば、人的資本理論において、職業訓練などの職業能力開発は、就業に有益な能力やスキルを習得するための学習活動であり、修得後の就業者の生産性を向上させるための投資行動でもあることから、人的資本投資とも呼ばれる。人的資本投資に関する実証分析の意義として、以下の点が挙げられる。日本では、少子高齢化による労働人口の減少に伴い、一人当たり労働力生産性を高めることは大きな課題であり、これからの日本経済の成長を考える上で、効果的な能力開発は大きな役割を担うと考えられる。職業訓練が賃金をはじめ就業などに与える影響を測定した多くの先行研究が存在する(Lalonde 1986, Bloom *et al.* 1997, Abadie *et al.* 2002)。

本研究における職業訓練とは、KHPS の調査票情報から得られる、会社から派遣される「教育訓練」と、就業者自ら参加する「自己啓発」の2つのものと定義する(詳細は3節参照)。

これら職業訓練が賃金や就業行動への影響を測定する上で、常に問題とされるのが内生性の影響である。職業訓練の効果を測定するため、受講者(処置群)と非受講者(対照群)の賃金や就業率の差の単純な計測を試みても、果たしてそれが職業訓練だけによる効果なのかを識別することが難しい。そもそも就業者個人が訓練に参加するかどうか(または企業側が訓練を受けさせたいかどうか)という選択に対して、就業者個人の能力や意欲の違いなど観測不可能な影響が含まれてしまい、純粋に職業訓練だけの効果を抽出することは難しい。先行研究では、如何にして、観察不可能な影響を統制し、職業訓練による賃金・就業率への効果を測定することに対して労力が注ぎ込まれてきた。

既に多くの先行研究において、個人の特性を考慮したとしても、職業訓練は賃金向上の効果があるという結果を得ている(黒澤 2005, 戸田・樋口 2005, 川口 2008, 市村 2010)。

日本国内の研究においては、一部の研究では、Propensity Score Matching 法(Rosenbaum and Rubin 1983)を用いて、処置効果の測定を行っている。吉田(2004)は、「消費生活に関するパネル調査」

を用いて、この方法により、通学・通信講座の受講は、4年後以降の収入を上昇することを明らかにした。小林・佐藤(2013)では、KHPSを用いて、同方法により、自己啓発が就業者の無業化の抑制、および再就職促進させる効果のみならず、自己啓発実施の3,4年後後に賃金が上昇する傾向を検証した。これらの研究では、自己啓発などの効果がすぐに現れるものではなく、長期的な観察が必要としている。

翻って、職業訓練などのプログラム参加の処置効果と脱落との関係について考えた場合、もし、職業訓練参加の有無が調査継続回答率に影響を与えていると(例えば、プログラム参加が回答者の負担を与えることで、アンケート調査の途中脱落を促すものや、プログラム参加者の属性が継続回答率と関係するなど)、経時的に比較した場合のプログラム処置効果に影響を与えかねない。この場合、例えプログラムがランダムに施行されていたとしても、脱落とプログラムが関連しているなら、処置効果にバイアスが生じる可能性が高い(Zamarro, *et al.* 2018)。第5節では、その点に着目しながら、選択バイアスと脱落バイアスの双方を考慮した推定を行いたい。

このような脱落バイアスに対する対応方法はいくつかあるが、特に経済学においては、Heckman(1976, 1979)による脱落と選択のバイアスをパラメトリックに補正する方法は実証研究における標準的な手順となっている。しかしながら、このアプローチでは厳密な仮定として、選択方程式において、当該方程式の被説明変数には影響するが、行動方程式の被説明変数には影響しない変数、「除外変数」を必要としており、その取得は困難であるとされている(Puhani 2000; Grisdal 2001)。MARを想定した、IPW法による補正やHeckman Modelなどが成立するためには、脱落バイアスから生じる処置群と対照群との間の重要な違いをコントロールするのに十分な観測可能な情報が利用可能であるかどうか依存する。

また近年では、除外変数に依存しない、Bounds推定量の利用が増えている。潜在的なバイアスについて点推定値を補正するのではなく選択メカニズムに対して弱い仮定を置き、真の処置効果が入るBound幅を推定する。そのような推定量がHorowitz and Manski(2000)やLee(2009)のBounds推定である。

これらの推定では選択メカニズムについてほとんど仮定も置かない。脱落をはじめとするMissing

Data問題に対して、Bounds推定を適用させたHorowitz and Manski(2000)では、Missing Dataがある場合における処置効果のBound、閾値(最小値、最大値)を導いた。彼らのBounds推定を発展させ、選択メカニズムに対して仮説を追加し、より幅の狭いBoundの設定する推定を考案したのがLee(2009)の推定であった。第6節ではこの推定を行い、第5節の推定結果と比較したい。

3. 職業訓練と継続回答

本節では、ここで扱うKHPSデータに関する基礎的情報、調査票から得られる職業訓練の定義、職業訓練と継続回答との関係についてみていく。

KHPSは、2004年1月から調査が始まり、年に一回、同一個人を対象としている継続調査である。初回は調査開始時点で満20~69歳男女約4,000名が、2007年には1,400名、2012年には1,000名がそれぞれ調査対象者に加わっている。サンプリング方法は、層化2段無作為抽出法を採用している⁷⁾。

調査項目は、構成世帯員の性別、年齢、就学就労条件のみならず、回答者の仕事や生活時間、世帯の支出、資産、住宅、健康状態に関するものまで含んだ非常に幅広いものとなっている⁸⁾。

表1は、調査開始コホート別の累積脱落率が示されている。調査開始コホート別の脱落率は、調査開始翌年の2年目12.7~17.3%、それ以降は各年約5~6%となり、累積脱落率は、調査9年目で50%を超えている。

次に、本節では、KHPSの調査票情報から、2種類の職業訓練を定義する。

第一の職業訓練、「教育訓練」⁹⁾は、過去1年間の間に、会社からの指示、会社から派遣された、教育訓練や研修、講習会、学校などの受講経験、職場からの指示で、勤務時間内における仕事を通じた上司・同僚の仕事ぶりや指示・アドバイスからの学び「On the Job Training(OJT)」、職場外での講習会・通信教育などからの学び「Off the Job Training(Off-JT)」から構成されている(2004~2008年のみの調査項目)。

第二の職業訓練、「自己啓発」¹⁰⁾¹¹⁾は、職場からの指示ではなく、自発的に技術や技能向上のために行う、大学[院]・専門学校などでの授業受講、各種講演会・セミナー・自主的勉強会への参加、通信教育の受講、テレビ、ラジオの講座や書籍を通じた自習などを指している(2004年以降継続する調査項目)。

表 1. 調査開始コホート別の累積脱落率

	KHPS2004			KHPS2007			KHPS2012		
	対象者数 (N)	継続回答率 (%)	累積脱落率 (%)	対象者数 (N)	継続回答率 (%)	累積脱落率 (%)	対象者数 (N)	継続回答率 (%)	累積脱落率 (%)
2004年	4,005	—							
2005年	3,314	82.75	17.25						
2006年	2,884	87.02	27.99						
2007年	2,636	91.40	34.18	1,419	—				
2008年	2,442	92.64	39.03	1,239	87.32	12.68			
2009年	2,280	93.37	43.07	1,130	91.20	20.37			
2010年	2,141	93.90	46.54	1,049	92.83	26.07			
2011年	2,037	95.14	49.14	976	93.04	31.22			
2012年	1,926	94.55	51.91	920	94.26	35.17	1,012	—	—

表 2. 職業訓練受講後の翌年の継続回答状況

		教育訓練			自己啓発				
		非受講者	受講者	全体	非受講者	受講者	全体		
翌年	継続回答	N	9,114	2,437	11,551	N	20,103	4,728	24,831
		%	89.0	89.7	89.1	%	90.7	90.3	90.7
	脱落	N	1,126	281	1,407	N	2,051	508	2,559
		%	11.0	10.3	10.9	%	9.3	9.7	9.3
全体		N	10,240	2,718	12,958	N	22,154	5,236	27,390
		%	100.0	100.0	100.0	%	100.0	100.0	100.0

*教育訓練の質問項目(2004～2008年)

*自己啓発の質問項目(2004～2012年)

それでは、本分析において重要である、職業訓練と継続回答との関係を確認する。表 2 は、訓練受講翌年の回答状況を受講者・非受講者ごとで比較したものであるが、教育訓練、自己啓発ともに、大きな違いがみられない。もう少し詳しくみるため、2004年～2008年各年における教育訓練の受講者と非受講者のその後の継続回答率を表章した(表 3)¹²⁾。

まず、教育訓練をみると、2004、2005、および2008年の受講・非受講者を比較すると、受講者の継続回答率の方が若干高く、2006、2007年の受講・非受講者の比較では非受講者の継続回答率が高くなる。

次に、自己啓発では、2004年の受講・非受講者の比較では、受講者の継続回答率が高いが、それ以降では、非受講者の回答率が高くなるか、両者の差がなくなる。2007年以降に新しく調査参加したサンプルでも同様に、調査前半の方で、受講者の継続回答率が高いが、後半では、その差がなくなっていくか、非受講者の継続回答率が高くなっていく。長期的に職業訓練参加と継続回答との関係を見ると、

教育訓練、自己啓発ともに同じような傾向(調査前半は受講者の方が継続回答率が高い)が確認された。

人的資本論の考え方にに基づき、職業訓練が就業者の能力および賃金の向上を促すと考える場合、受講者の回答率が高いことは、職業訓練による賃金への影響である処置効果が過大になり、逆に非受講者の回答率が高いことは、処置効果が過小に推定される可能性がある。ここでは、職業訓練と継続回答との間の関係が明確とはならなかったが、第 5 節において、IPW を用い、脱落による影響を考慮した場合、処理効果への影響を推定したい。

4. 継続回答の規定要因

本節では、職業訓練による賃金への効果測定の前に、調査対象者の継続回答の規定要因に関する分析を行う。

調査対象者が脱落するかどうかのメカニズムは、大きく分けて、3種類のパターンがある(岩崎 2002)。
① Missing Completely At Random (MCAR) : 脱落が完全にランダムな場合、② Missing At Random

表 3. 職業訓練受講後の継続回答率

教育訓練(2004年調査からの調査対象者)

単位：%

受講年	2004年受講			2005年受講			2006年受講			2007年受講			2008年受講		
	(a)非受講者 3450人	(b)受講者 555人	差 (b)-(a)	(a)非受講者 2806人	(b)受講者 508人	差 (b)-(a)	(a)非受講者 2445人	(b)受講者 439人	差 (b)-(a)	(a)非受講者 2247人	(b)受講者 389人	差 (b)-(a)	(a)非受講者 2085人	(b)受講者 357人	差 (b)-(a)
2004年	100.0	100.0	—	—	—	—	—	—	—	—	—	—	—	—	—
2005年	82.2	86.1	3.9	100.0	100.0	—	—	—	—	—	—	—	—	—	—
2006年	71.3	76.4	5.1	86.6	89.2	2.5	100.0	100.0	—	—	—	—	—	—	—
2007年	65.1	70.3	5.2	79.3	81.1	1.8	91.6	90.2	-1.4	100.0	100.0	—	—	—	—
2008年	60.2	65.6	5.4	73.4	75.4	2.0	85.1	82.5	-2.6	92.9	91.3	-1.6	100.0	100.0	—
2009年	56.1	61.8	5.7	68.4	71.1	2.7	79.4	77.0	-2.4	86.8	84.8	-1.9	93.3	93.8	0.6
2010年	52.6	59.1	6.5	64.4	65.7	1.4	74.7	71.5	-3.2	81.4	80.2	-1.2	87.5	88.8	1.3
2011年	49.9	56.9	7.1	61.2	63.2	2.0	70.9	69.0	-1.9	77.7	74.8	-2.9	83.1	85.4	2.4
2012年	47.2	53.7	6.5	57.8	59.8	2.0	67.0	65.8	-1.1	73.5	70.4	-3.1	78.7	80.1	1.5

教育訓練(2007年調査からの調査対象者)

受講年	2007年受講			2008年受講		
	(a)非受講者 1162人	(b)受講者 257人	差 (b)-(a)	(a)非受講者 1029人	(b)受講者 210人	差 (b)-(a)
2007年	100.0	100.0	0.0	—	—	—
2008年	87.8	85.2	-2.6	100.0	100.0	0.0
2009年	79.8	79.0	-0.8	90.8	93.3	2.6
2010年	74.0	73.5	-0.5	84.5	85.2	0.7
2011年	68.7	69.3	0.6	78.1	81.9	3.8
2012年	65.0	64.2	-0.8	73.5	78.1	4.6

教育訓練の設問あり

自己啓発(2004年調査からの調査対象者)

単位：%

受講年	2004年受講			2005年受講			2006年受講			2007年受講			2008年受講		
	(a)非受講者 2925人	(b)受講者 1080人	差 (b)-(a)	(a)非受講者 2569人	(b)受講者 745人	差 (b)-(a)	(a)非受講者 2343人	(b)受講者 541人	差 (b)-(a)	(a)非受講者 2181人	(b)受講者 455人	差 (b)-(a)	(a)非受講者 2081人	(b)受講者 361人	差 (b)-(a)
2004年	100.0	100.0	—	—	—	—	—	—	—	—	—	—	—	—	—
2005年	82.2	85.9	3.7	100.0	100.0	—	—	—	—	—	—	—	—	—	—
2006年	71.3	76.6	5.3	87.2	86.4	-0.8	100.0	100.0	—	—	—	—	—	—	—
2007年	65.1	70.9	5.8	79.8	78.5	-1.3	91.2	92.2	1.0	100.0	100.0	—	—	—	—
2008年	60.2	66.6	6.3	74.2	71.8	-2.4	84.8	84.1	-0.7	92.9	91.4	-1.5	100.0	100.0	—
2009年	56.1	61.9	5.8	69.6	66.2	-3.4	79.2	78.6	-0.6	86.9	84.4	-2.5	93.5	92.5	-1.0
2010年	52.6	57.3	4.8	65.6	61.1	-4.6	74.4	73.4	-1.1	81.7	78.9	-2.8	87.8	87.0	-0.8
2011年	49.9	55.2	5.3	62.4	58.3	-4.1	71.1	68.6	-2.5	77.9	74.5	-3.3	83.4	83.7	0.3
2012年	47.2	52.9	5.7	59.1	54.6	-4.5	67.1	65.2	-1.9	73.5	71.0	-2.5	78.7	80.1	1.4

自己啓発(2007年調査からの調査対象者)

受講年	2007年受講			2008年受講			2009年受講			2010年受講			2011年受講		
	(a)非受講者 1058人	(b)受講者 361人	差 (b)-(a)	(a)非受講者 985人	(b)受講者 254人	差 (b)-(a)	(a)非受講者 907人	(b)受講者 223人	差 (b)-(a)	(a)非受講者 868人	(b)受講者 181人	差 (b)-(a)	(a)非受講者 817人	(b)受講者 159人	差 (b)-(a)
2007年	100.0	100.0	—	—	—	—	—	—	—	—	—	—	—	—	—
2008年	87.2	87.5	0.3	100.0	100.0	—	—	—	—	—	—	—	—	—	—
2009年	79.2	80.9	1.7	90.8	92.9	2.2	100.0	100.0	—	—	—	—	—	—	—
2010年	73.1	76.5	3.4	84.2	86.6	2.5	92.8	92.8	0.0	100.0	100.0	—	—	—	—
2011年	67.7	72.0	4.3	78.4	80.3	1.9	86.1	87.4	1.3	93.0	93.4	0.4	100.0	100.0	—
2012年	63.8	67.9	4.1	74.0	75.2	1.2	81.7	80.3	-1.4	87.8	87.3	-0.5	95.1	89.9	-5.2

自己啓発の設問あり

(MAR)：脱落は観察可能な変数にのみ依存する場合、③ Non-ignorable, Missing At No Random (MANR)：脱落が、脱落時点以降の観測不可能な変数に依存している場合の3つである¹³⁾。

本節では脱落過程は観測可能な変数にのみ依存している(MAR)と仮定する。ここで、継続回答(脱落の逆の事象)の規定要因を探る理由は、後述するIPW法において、Weightとして、継続回答率の逆数を用いるためであり、脱落を考慮した推定の準備といえる。

具体的には、継続回答ダミー s_{it} (継続回答=1, 脱落=0, t 年度) を被説明変数、前年 ($t-1$ 年度) の個人的属性、経済的属性などの情報 v_{it-1} 、調査員調査票から得られる調査プロセスに関する変数 z_{it} を説明変数とし、Random Effect Probit Model による分析を行った。

$$P(s_{it} = 1 | v_{it-1}, z_{it}, \text{if } s_{it-1} = 1)$$

説明変数として用いる変数 v は、個人的属性として、年齢、年齢の2乗項、配偶関係(既婚=1, それ以外=0)、学歴(中学卒, 高校卒[リファレンスグループ], 専門学校卒, 短大卒, 大学卒以上)、性別(男性=1, 女性=0)、子ども(18歳未満)がいない、健康状態(よい, まあよい, ふつう[リファレンスグループ], あまりよくない, よくない)、過去1年間の出来事(世帯変動なし[リファレンスグループ], 出産, 世帯員死亡, 新婚, 離婚, 別居, 実家からの独立, 実家に戻る, それ以外の居住変更)、居住都市規模(政令指定都市, その他の市[リファレンスグループ], 町村)、持家居住(持家=1, それ以外=0)、地域ブロック(北海道, 東北, 関東[リファレンスグループ], 中部, 近畿, 中国, 四国, 九州)、調査年ダミーを用いている。また経済的属性として、就業(就業=1, それ以外=0)、世帯年収(対数値)を利用した。

加えて、除外変数として、調査プロセスに関する変数 $z^{14)}$ を活用する。具体的には各年度の初回訪問月、休日訪問の有無、訪問回数を用いている。各変数を継続回答に対して影響を与える変数として用いた理由は以下の通りである。KHPSは毎年1月末から3月にかけて調査が行われているが、調査時期の遅れは、回収時期が定まっている状況下では回収率の低下を招く。また、調査員調査票では、初回訪問から最終訪問まで、対象者宅を訪問した日時が確認可能のため、訪問日すべてが平日であったか、一度でも休日や祝日であったかを確認できる。一般に勤

労者世帯では、平日にコンタクトをとることが難しいため、休日・祭日に訪問しなければ、接触できない可能性が高い(Campanelli *et al.* 1997, Lynn *et al.* 2002)。最後に、訪問回数が多いほど、なかなか対象者と会えないことを示しており、回数が多いほど、回収確率が下がると考えられる。これは訪問だけではなく、電話による連絡でも同様のことがいえる(Stoop *et al.* 2010)。

推定結果(表5)をみると、年齢が高い、健康状態が良好、家族に子ども(18歳未満)がいない対象者は、継続回答する傾向が確認できる。逆に、対象者の就学年数が短く、1年以内に別居を経験している、世帯年収が高い対象者ほど、継続回答せずに脱落する傾向がある。また、調査プロセスに関する影響では、訪問回数が多いほど脱落する結果となった。

加えて、職業訓練受講経験との関係を確認すると(表6)、教育訓練、自己啓発ともに受講後1~3年は正の効果が見られるが、統計的に有意ではなかった。自己啓発は受講4年後、継続回答に対し、負に有意な影響が確認された。自発的な学習に時間を割くことで、調査回答が負担となり、回答しなくなるということが考えられる。

5. IPWRA 推定による平均処置効果推定

本節では、職業訓練受講の内生性を考慮した上で、賃金への影響をみるために、Wooldridge(2007, 2010)によるIPW-Regression-Adjustment(IPWRA)推定量(別名 Wooldridge's "Double-Robust" 推定量)を用いて分析する¹⁵⁾。

この推定量は"Double-Robust" (二重に堅牢)と言われるように、「職業訓練を行うかどうかを示す、Propensity Scoreを計算する際に利用するモデル(選択方程式)が正しい」、あるいは「賃金などの被説明変数(目的変数)を説明する回帰モデル(行動方程式)が正しい」のいずれかが成立すれば、目的変数の周辺期待値や因果効果の一致推定量を得られる(Scharfstein, *et al.* 1999)。

具体的な推定作業では、以下の三段階の手続きを行っている。第一段階で、処置割り当て(職業訓練参加)を被説明変数としたProbit Modelを推定し、職業訓練参加確率、Propensity Scoreを計算し、第二段階で、その逆数をWeightとした加重回帰モデルを参加者、非参加者それぞれで推定し、最後に、両モデルから推定される賃金の差を平均処置効果(ATE)として求めている。

表 4. 記述統計量(継続回答推定モデル)

継続回答		20,180	0.9257185	0.2622349	0	1
年齢		20,180	49.22354	13.24861	20	77
年齢2乗項		20,180	2598.474	1304.182	400	5929
男性		20,180	0.4906838	0.4999256	0	1
有配偶		20,180	0.781219	0.41343	0	1
子どもいない		20,180	0.3996531	0.4898392	0	1
健康状態	よい	20,180	0.2028246	0.4021129	0	1
	まあよい	20,180	0.2782953	0.4481707	0	1
	(ふつう)	20,180	0.399108	0.4897272	0	1
	あまりよくない	20,180	0.0990585	0.2987479	0	1
	よくない	20,180	0.0113479	0.1059229	0	1
学歴	中学校卒	20,180	0.1093657	0.3121052	0	1
	(高校卒)	20,180	0.4742319	0.4993479	0	1
	専門学校卒	20,180	0.0430129	0.2028911	0	1
	短大卒	20,180	0.1452428	0.3523542	0	1
	大学卒以上	20,180	0.2281467	0.4196481	0	1
過去1年間の出来事	世帯変動なし	20,180	0.8529732	0.3541414	0	1
	出産	20,180	0.0201189	0.1404106	0	1
	世帯員死亡	20,180	0.0030723	0.0553449	0	1
	新婚	20,180	0.0060456	0.0775199	0	1
	離婚	20,180	0.0028741	0.0535352	0	1
	別居	20,180	0.0017344	0.0416109	0	1
	実家からの独立	20,180	0.0065411	0.0806143	0	1
	実家に戻る	20,180	0.0027255	0.0521361	0	1
	それ以外の居住変更	20,180	0.038553	0.1925319	0	1
居住都市規模	政令指定都市	20,180	0.26333	0.4404508	0	1
	(その他の市)	20,180	0.6160059	0.4863685	0	1
	町村	20,180	0.1206145	0.3256868	0	1
持家		20,180	0.7909316	0.4066534	0	1
就業		20,180	0.7448464	0.4359583	0	1
地域ブロックダミー	北海道	20,180	0.0461843	0.2098893	0	1
	東北	20,180	0.0636769	0.2441825	0	1
	関東	20,180	0.3285927	0.4697132	0	1
	中部	20,180	0.1736868	0.3788493	0	1
	近畿	20,180	0.1878593	0.3906094	0	1
	中国	20,180	0.0556987	0.2293447	0	1
	四国	20,180	0.0324579	0.1772172	0	1
	九州・沖縄	20,180	0.1118434	0.3151815	0	1
世帯年収(対数値)		20,180	6.299799	0.6371734	0	7.711997
初回訪問月(1月)		20,180	0.188553	0.3911628	0	1
休日訪問の有無		20,180	0.6156591	0.4864511	0	1
訪問回数		20,180	2.827453	1.078999	1	11
Weight		20,180	1.835091	2.882812	1.01861	32.74721

本稿ではさらに、継続回答率による加重回帰調整を行っている¹⁶⁾。結果変数である労働収入(対数値)を y 、 k 個の説明変数、回帰係数を β 、誤差を u とし、 n 個の観測値に対する、次のような線形モデルを考える。 u は正規分布 $N(0, \sigma^2)$ に従い、互いに独立と仮定する。

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + u_i = x_i \beta + u_i$$

$$\beta = (\beta_1, \beta_2, \dots, \beta_k)^{17)}$$

Regression-Adjustment 推定では、対照群の回帰

モデルを利用し、処置群における処置なしの潜在的な結果変数(Y^0)を補完し、処置群の回帰モデルを使用して、対照群の処置ありの潜在的な結果変数(Y^1)を補完する。

T_i は処置変数($i=1, \dots, N$)であり、 $T_i=1$ の場合は処置群、 $T_i=0$ の場合は対照群であることを示す。さらに、 β_0 は対照群の回帰係数、 β_1 は処置群の回帰係数である。

$w_i, i=1, \dots, N$ はWeight(継続回答率の逆数)、

表 5. 継続回答推定モデル

	限界効果	頑健的 標準誤差	z 値		
年齢	0.007	(0.001)	4.92	***	
年齢 2 乗項	0.000	(0.000)	-4.37	***	
男性	-0.006	(0.005)	-1.25		
有配偶	0.008	(0.007)	1.16		
子どもいない	0.028	(0.006)	4.75	***	
健康状態	よい	0.012	(0.006)	2.01	**
	まあよい	0.009	(0.005)	1.6	
	(ふつう)				
	あまりよくない	-0.011	(0.007)	-1.46	
	よくない	-0.055	(0.017)	-3.26	***
学歴	中学校卒	-0.017	(0.008)	-2.16	**
	(高校卒)				
	専門学校卒	0.000	(0.012)	0.01	
	短大卒	-0.002	(0.007)	-0.25	
	大学卒以上	0.003	(0.006)	0.44	
過去 1 年間の出来事 (世帯変動なし)	出産	0.013	(0.016)	0.81	
	世帯員死亡	0.003	(0.035)	0.09	
	新婚	0.012	(0.026)	0.46	
	離婚	-0.052	(0.033)	-1.56	
	別居	-0.088	(0.039)	-2.26	**
	実家からの独立	0.003	(0.025)	0.13	
	実家に戻る	-0.023	(0.036)	-0.64	
	それ以外の居住変更	-0.007	(0.011)	-0.61	
居住都市規模	政令指定都市	0.004	(0.006)	0.71	
	(その他の市)				
	町村	-0.012	(0.007)	-1.6	
持家	0.006	(0.006)	1.06		
就業	0.006	(0.006)	1.03		
世帯年収(万円)	-0.010	(0.004)	-2.47	**	
初回訪問月(1 月)	0.000	(0.006)	-0.04		
休日訪問の有無	-0.003	(0.005)	-0.76		
訪問回数	-0.014	(0.002)	-7.04	***	
地域ブロックダミー		Yes			
調査年ダミー		Yes			
定数項	0.982	(0.282)	3.48	***	
標本数		20,180			
グループ数		4,475			
Wald chi2		294.31			
Prob > chi2		0.00			
疑似尤度		-5147.67			

注) *** : 1%, ** : 5%, * : 10% 水準で棄却.

$W_i = \sum w_i$ は Weight の合計値(継続回答者の人数), $W_0 = \sum_i (1 - T_i) w_i$ および $W_1 = \sum_i T_i w_i$ はそれぞれ対照群, 処置群の各群の人数とする. 処置群, 対照群の潜在的結果, Y^1 および Y^0 の期待値の推定量は, 次のように定義される.

$$\hat{\eta}^1 = E[Y^1] = \sum_{i=1}^N w_i T_i y_i + \frac{1 - \hat{p}}{W_0} \sum_{i=1}^N w_i (1 - T_i) x_i \tilde{\beta}_1$$

$$= \hat{p} \bar{y}_1 + (1 - \hat{p}) \bar{x}_0 \tilde{\beta}_1$$

$$\hat{\eta}^0 = E[Y^0]$$

$$= \frac{\hat{p}}{W_1} \sum_{i=1}^N w_i x_i T_i \tilde{\beta}_0 + \frac{1 - \hat{p}}{W_0} \sum_{i=1}^N w_i (1 - T_i) y_i$$

$$= \hat{p} \bar{x}_1 \tilde{\beta}_0 + (1 - \hat{p}) \bar{y}_0$$

$\hat{p} = W_1 / W$ は処置群の相対的な割合を示している. ここでは, \bar{y}_1 および \bar{x}_1 はそれぞれ処置群におけ

表 6. 継続回答モデル(教育訓練, 自己啓発の影響)

教育訓練	限界効果	頑健的 標準誤差	z 値		疑似尤度	N
(受講 1 年後)	0.0021	0.0071	0.30		-2644.5	9,884
(受講 2 年後)	0.0037	0.0074	0.49		-2644.5	9,183
(受講 3 年後)	0.0016	0.0067	0.24		-1832.3	8,679
(受講 4 年後)	-0.0061	0.0068	-0.90		-1229.8	6,363
(受講 5 年後)	0.0089	1.9703	0.00		-781.8	4,161
自己啓発	限界効果	頑健的 標準誤差	z 値		疑似尤度	N
(受講 1 年後)	0.0066	0.0054	1.22		-4721.8	18,896
(受講 2 年後)	0.0053	0.0056	0.94		-3421.9	15,141
(受講 3 年後)	0.0003	0.0056	0.05		-2527.1	11,790
(受講 4 年後)	-0.0103	0.0060	-1.72	*	-1786.8	8,646
(受講 5 年後)	-0.0115	0.0073	-1.58		-1115.1	5,660

注) ***: 1%, **: 5%, *: 10% 水準で棄却。

その他の説明変数として、表 4 のモデルと同じ説明変数を利用している。

る結果変数, 説明変数の平均値, \bar{y}_0 および \bar{x}_0 はそれぞれ対照群における結果変数, 説明変数の平均値を示す¹⁸⁾。

潜在的な結果の平均値に基づき, 平均処置効果 (ATE) の推定値は次のように定義される。

$$ATE = \eta^1 - \eta^0$$

処置関数モデル(職業訓練受講モデル)では, 説明変数として, 性別(男性=1, 女性=0), 年齢, 年齢の 2 乗項, 学歴(中学卒, 高校卒[リファレンスグループ], 専門学校卒, 短大卒, 大学卒以上), 子ども(18 歳)いない, 就業月数, 就業月数の 2 乗項, 職種(採掘・製造・建築・保守・運搬などの作業者, 販売従事者, サービス職従事者[リファレンスグループ], 管理的職種, 事務従事者, 運輸・通信従事者, 保安職業従事者, 専門的・技術的職業従事者, その他)を用いる。

また, 賃金関数モデルでは, 被説明変数として労働による月収, 年収(対数値) $\ln w$ を, 説明変数として, 賃金に影響する説明変数群 x を用いる。

説明変数群は, 性別(男性=1, 女性=0), 年齢, 年齢の 2 乗項, 学歴(中学卒, 高校卒[リファレンスグループ], 専門学校卒, 短大卒, 大学卒以上), 就業月数, 就業月数の 2 乗項, 就業先規模(0~29 人, 30~99 人, 100~499 人[リファレンスグループ], 500 人以上, 官公庁), 構成されている。

これに加えて, IPW 法を用いて, 脱落による推定バイアスの調整を行った(Fitzgerald, Gottschalk, and Moffitt 1998a, 1998b, Wooldridge 2002)。具体的

な IPW の算出方法として, 前節で行ったように, 継続回答ダミーを被説明変数とした Probit Model から予測値を得て, その予測値の逆数を Weight とし(翌年継続回答者には 1/予測値, 翌年脱落者には 1/(1-予測値)), 最後に, 目的としている推定関数に算出した Weight を用いる。

以下では, 脱落バイアスを考慮した継続回答率の逆数を IPW とした Weighted モデルと考慮しない Unweighted モデルを比較し, どのようなバイアスが発生しているかを検証する。

まず, 処置関数モデル(教育訓練)の推定結果をみると, 受講後経過年数で多少の違いはあるものの, 企業による教育訓練は, (高校卒者と比べて)短大卒, 大学卒以上など高学歴者ほど, (就業先規模が 100~499 人の企業就業者と比べて)就業先規模が 500 人以上の企業や官公庁に勤めている者ほど, (サービス職従事者と比べて)運転従事者, 車掌, 通信士などの運輸・通信従事者ほど受講している(表 8-1)。また, 自己啓発では, 同様に高学歴者, 運輸・通信従事者ほど受講しているが, 異なる点としては, 女性ほど受講する性別による影響が確認される(表 8-2)。

次に, 賃金関数モデルの推定結果をみると, (女性に比べて)男性, 年齢が高い, 就業先規模が 500 人以上の企業や官公庁に勤めている, 就業月数が高い, 学歴が高いほど賃金が高くなる傾向がある(表 9, 10)。これら賃金関数モデルの推定結果から得られる受講者, 非受講者の賃金関数から得られる賃金

表 7. 記述統計量(教育訓練受講 1 年後の平均処置効果推定)

		N	平均	標準偏差	最小値	最大値
男性		8,450	0.5964497	0.4906383	0	1
年齢		8,450	46.01763	11.85652	21	75
年齢 2 乗項		8,450	2258.183	1100.898	441	5625
子どもいない		8,450	0.4363314	0.4959591	0	1
就業先規模	1~29 人	8,450	0.4104142	0.491938	0	1
	30~99 人	8,450	0.127574	0.3336346	0	1
	100~499 人	8,450	0.1757396	0.380621	0	1
	500 人以上	8,450	0.2166864	0.4120115	0	1
	官公庁	8,450	0.0618935	0.2409763	0	1
	就業月数	8,450	128.9933	137.8706	0	680
	就業月数 2 乗項	8,450	35645.32	61420.81	0	462400
学歴	中学校卒	8,450	0.0887574	0.2844101	0	1
	高校卒	8,450	0.4629586	0.4986556	0	1
	専門学校卒	8,450	0.043432	0.2038395	0	1
	短大卒	8,450	0.1405917	0.3476205	0	1
	大学卒以上	8,450	0.2642604	0.4409647	0	1
居住都市規模	大都市	8,450	0.2647337	0.4412174	0	1
	その他の市	8,450	0.6039053	0.4891135	0	1
	町村	8,450	0.1312426	0.3376855	0	1
職種	作業者	8,242	0.1941276	0.3955516	0	1
	販売従事者	8,242	0.1526329	0.3596551	0	1
	(サービス職従事者)	8,242	0.1430478	0.3501428	0	1
	管理的職種	8,242	0.0589663	0.2355759	0	1
	事務従事者	8,242	0.1748362	0.37985	0	1
	運輸・通信従事者	8,242	0.2193642	0.413841	0	1
	保安食従事者	8,242	0.0185635	0.1349854	0	1
	その他	8,242	0.0384615	0.1278588	0	1
地域ブロックダミー	北海道	8,450	0.0398817	0.1956927	0	1
	東北	8,450	0.0669822	0.250006	0	1
	関東	8,450	0.3332544	0.4714045	0	1
	中部	8,450	0.1778698	0.3824258	0	1
	近畿	8,450	0.1843787	0.3878157	0	1
	中国	8,450	0.0518343	0.2217055	0	1
	四国	8,450	0.0319527	0.1758845	0	1
	九州・沖縄	8,450	0.1138462	0.3176431	0	1
月収(対数値)		8,450	12.24012	0.8267428	6.909	14.457
教育訓練受講		6,443	0.2129443	0.4094204	0	1
自己啓発受講		8,417	0.2093382	0.4068604	0	1
調査員訪問回数(中央値以上)		8,450	0.5325444	0.4989693	0	1
健康状態(よい、まあよい)		8,403	0.5330239	0.4989379	0	1

推定値の差から平均処理効果を計算している。これらの表では、脱落による影響を考慮していない Unweighted モデル(表 9-1, 10-1) と考慮した Weight モデル(表 9-2, 10-2) として表章している。

表 11 には、表 9, 10 で得られた平均処置効果推定の IPWRA 推定値(Weighted, Unweighted) とともに、内生性を考慮せず、受講者と非受講者の平均賃金の単純な差(Unmatched)を併記した。

ここでは、Unmatched と Unweighted の推定値の差は「内生性バイアス」を、Unweighted と Weighted との差は「脱落バイアス」と考える。内

生性バイアスは教育訓練、自己啓発ともに全てのケースで過大に推計されていたことを、また脱落バイアスは(教育訓練の 1 年後を除き)こちらも過大に推定されていたことを示している。両バイアスを比較すると、受講後の経過年数が 3 年以内では、内生性バイアスの方が大きいものの、4 年以降では脱落バイアスの方が大きくなっている。第 3 節では記述統計量から職業訓練と継続回答の関係をみたものの、明確な傾向を把握するまでには至らなかったが、ここでは、脱落による過大推定が発生していることから、職業訓練受講と継続回答との間に正の相関があ

表 8-1. 処置関数モデル(教育訓練受講)

被説明変数	教育訓練 1年前受講		教育訓練 2年前受講		教育訓練 3年前受講		教育訓練 4年前受講		教育訓練 5年前受講						
	係数	z	係数	z	係数	z	係数	z	係数	z					
男性	0.036	(0.063)	0.58	-0.040	(0.074)	-0.54	-0.017	(0.073)	-0.24	-0.082	(0.085)	-0.97	-0.159	(0.098)	-1.62
年齢	0.016	(0.017)	0.91	0.016	(0.019)	0.83	0.032	(0.019)	1.7	-0.008	(0.024)	-0.35	0.018	(0.029)	0.62
年齢 2 乗項	0.000	(0.000)	-1.42	0.000	(0.000)	-1.35	0.000	(0.000)	-2.26	0.000	(0.000)	-0.06	0.000	(0.000)	-0.81
子どもがいない	-0.015	(0.062)	-0.24	0.026	(0.065)	0.4	0.095	(0.069)	1.37	0.000	(0.082)	0	-0.039	(0.089)	-0.45
就業先規模	-0.256	(0.076)	-3.37	-0.493	(0.082)	-6.02	-0.558	(0.088)	-6.38	-0.516	(0.106)	-4.85	-0.658	(0.130)	-5.07
1~29人	0.023	(0.096)	0.24	-0.247	(0.105)	-2.36	-0.371	(0.107)	-3.47	-0.156	(0.138)	-1.13	-0.323	(0.152)	-2.13
30~99人															
(100~499人)															
500人以上	0.366	(0.078)	4.7	0.087	(0.086)	1.01	0.189	(0.093)	2.03	0.202	(0.112)	1.8	0.060	(0.138)	0.44
官公庁	0.321	(0.117)	2.73	0.076	(0.141)	0.54	0.277	(0.148)	1.88	0.237	(0.168)	1.41	0.120	(0.196)	0.61
就業月数	0.000	(0.001)	-0.18	0.002	(0.001)	2.18	0.001	(0.001)	1.37	0.001	(0.001)	0.85	0.004	(0.001)	3.18
就業月数 2 乗項	0.000	(0.000)	0.16	0.000	(0.000)	-1.19	0.000	(0.000)	-1.06	0.000	(0.000)	-0.77	0.000	(0.000)	-2.79
学歴	-0.225	(0.119)	-1.89	-0.382	(0.089)	-4.3	-0.246	(0.122)	-2.01	-0.359	(0.120)	-2.98	-0.283	(0.141)	-2
中学校卒															
(高校卒)															
専門学校卒	0.138	(0.118)	1.17	0.358	(0.165)	2.17	0.361	(0.162)	2.23	0.435	(0.196)	2.22	0.413	(0.286)	1.44
短大卒	0.188	(0.086)	2.19	0.198	(0.097)	2.04	0.086	(0.101)	0.85	0.116	(0.123)	0.94	-0.141	(0.120)	-1.18
大学卒以上	0.161	(0.072)	2.24	0.216	(0.083)	2.59	0.111	(0.087)	1.29	0.086	(0.097)	0.88	0.111	(0.117)	0.95
作業者	-0.116	(0.106)	-1.1	-0.154	(0.122)	-1.26	-0.185	(0.121)	-1.53	0.102	(0.151)	0.67	0.037	(0.156)	0.24
販売従事者	-0.164	(0.100)	-1.64	-0.029	(0.121)	-0.24	0.008	(0.127)	0.06	0.083	(0.147)	0.56	0.177	(0.157)	1.12
(サービス従事者)															
管理的職種	0.042	(0.147)	0.29	-0.083	(0.168)	-0.49	0.106	(0.168)	0.63	-0.053	(0.190)	-0.28	0.063	(0.223)	0.28
事務従事者	-0.203	(0.105)	-1.93	-0.144	(0.118)	-1.22	-0.173	(0.129)	-1.35	-0.158	(0.145)	-1.08	0.048	(0.160)	0.3
運輸・通信従事者	0.338	(0.101)	3.35	0.315	(0.116)	2.71	0.329	(0.125)	2.64	0.330	(0.140)	2.36	0.399	(0.146)	2.74
保安食従事者	-0.178	(0.133)	-1.35	-0.009	(0.166)	-0.06	-0.272	(0.164)	-1.65	0.043	(0.208)	0.21	-0.015	(0.163)	-0.09
その他	0.224	(0.182)	1.23	-0.017	(0.183)	-0.09	-0.218	(0.228)	-0.96	-0.093	(0.254)	-0.37	-0.080	(0.230)	-0.35
定数項	-1.001	(0.361)	-2.77	-0.950	(0.396)	-2.4	-1.177	(0.406)	-2.9	-0.339	(0.517)	-0.66	-1.097	(0.644)	-1.7
N			8,450			7,701			7,129			5,111			

注) ***: 1%, **: 5%, *: 10%水準で棄却。

表 8-2. 処置関数モデル(自己啓発受講)

被説明変数	自己啓発 1年前受講 受講=1, 非受講=0		自己啓発 2年前受講 受講=1, 非受講=0		自己啓発 3年前受講 受講=1, 非受講=0		教育訓練 4年前受講 受講=1, 非受講=0		教育訓練 5年前受講 受講=1, 非受講=0											
	係数	z	係数	z	係数	z	係数	z	係数	z										
男性	-0.106	(0.062)	-1.69	*	-0.212	(0.069)	-3.09	***	-0.246	(0.084)	-2.94	***	-0.377	(0.098)	-3.84	***	-0.320	(0.128)	-2.5	**
年齢	0.021	(0.016)	1.34		-0.003	(0.017)	-0.17		-0.010	(0.023)	-0.45		-0.026	(0.027)	-0.95		-0.068	(0.040)	-1.7	*
年齢 2 乗項	0.000	(0.000)	-2.28	**	0.000	(0.000)	-0.53		0.000	(0.000)	-0.14		0.000	(0.000)	0.66		0.001	(0.000)	1.73	*
子どもがいない	-0.141	(0.057)	-2.46	**	-0.052	(0.065)	-0.8		-0.129	(0.075)	-1.73	*	-0.093	(0.090)	-1.03		-0.049	(0.109)	-0.45	
就業先規模	0.199	(0.073)	2.74	***	-0.010	(0.084)	-0.12		0.104	(0.096)	1.08		0.125	(0.122)	1.02		-0.137	(0.155)	-0.88	
1~29人	0.107	(0.096)	1.11		-0.050	(0.114)	-0.44		0.021	(0.118)	0.18		-0.129	(0.136)	-0.95		-0.482	(0.166)	-2.9	***
30~99人																				
(100~499人)																				
500人以上	0.143	(0.077)	1.86	*	0.073	(0.092)	0.8		0.126	(0.109)	1.16		0.187	(0.124)	1.51		0.088	(0.172)	0.51	
官公庁	-0.031	(0.125)	-0.25		-0.248	(0.124)	-1.99		0.179	(0.154)	1.16		-0.187	(0.171)	-1.09		0.141	(0.242)	0.58	
就業月数	-0.001	(0.001)	-1.62		-0.001	(0.001)	-1.01		0.000	(0.001)	-0.52		0.001	(0.001)	0.56		0.002	(0.001)	1.76	*
就業月数 2 乗項	0.000	(0.000)	1.39		0.000	(0.000)	1.01		0.000	(0.000)	0.52		0.000	(0.000)	-0.99		0.000	(0.000)	-1.86	*
学歴	-0.084	(0.127)	-0.66		-0.420	(0.095)	-4.42	***	-0.418	(0.110)	-3.8	***	-0.382	(0.137)	-2.79	***	-0.579	(0.181)	-3.2	***
中学校卒																				
(高校卒)																				
専門学校卒	0.227	(0.117)	1.94	*	0.335	(0.148)	2.27	**	0.066	(0.140)	0.47		0.233	(0.241)	0.97		0.057	(0.181)	0.31	
短大卒	0.222	(0.078)	2.86	***	0.097	(0.089)	1.09		0.036	(0.104)	0.34		0.230	(0.125)	1.84	*	-0.222	(0.148)	-1.5	
大学卒以上	0.516	(0.070)	7.35	***	0.531	(0.082)	6.5	***	0.388	(0.099)	3.9	***	0.595	(0.114)	5.2	***	0.257	(0.151)	1.69	*
作業者	-0.184	(0.106)	-1.73	*	-0.216	(0.115)	-1.87	*	-0.208	(0.129)	-1.61		0.146	(0.167)	0.88		0.283	(0.195)	1.45	
販売従事者	-0.120	(0.099)	-1.21		-0.224	(0.113)	-1.98	**	-0.078	(0.131)	-0.6		-0.045	(0.155)	-0.29		-0.281	(0.180)	-1.57	
(サービス職従事者)																				
管理的職種	0.191	(0.144)	1.32		0.070	(0.157)	0.45		0.280	(0.203)	1.38		0.581	(0.222)	2.61	***	0.391	(0.294)	1.33	
事務従事者	-0.048	(0.097)	-0.5		-0.130	(0.116)	-1.12		-0.081	(0.132)	-0.61		0.062	(0.169)	0.37		0.142	(0.199)	0.71	
運輸・通信従事者	0.332	(0.092)	3.59	***	0.359	(0.108)	3.32	***	0.449	(0.127)	3.55	***	0.544	(0.153)	3.57	***	0.435	(0.178)	2.44	**
保安食従事者	0.225	(0.175)	1.29		0.376	(0.180)	2.09		0.595	(0.245)	2.43	**	0.694	(0.209)	3.33	***	1.241	(0.490)	2.53	**
その他	0.082	(0.205)	0.4		-0.032	(0.193)	-0.17		0.124	(0.261)	0.48		0.587	(0.305)	1.92	*	-0.133	(0.382)	-0.35	
定数項	-1.133	(0.340)	-3.33	***	-0.496	(0.372)	-1.33		-0.326	(0.479)	-0.68		-0.254	(0.571)	-0.44		0.585	(0.827)	0.71	
標本数			10,240				7,863				5,803				3,847					2084

注) ***: 1%, **: 5%, *: 10%水準で発現。

表 9-1. 賃金関数と教育訓練の月収(対数値)に対する平均処理効果(1~5年後, Unweighted モデル)

被説明変数：月収(対数値)	1年後		2年後		3年後		4年後		5年後	
	非受講者 係数 頂層的 標準誤差	受講者 係数 頂層的 標準誤差								
男性	0.900 (0.017)	0.889 (0.041)	0.880 (0.018)	0.821 (0.039)	0.889 (0.018)	0.810 (0.038)	0.878 (0.021)	0.789 (0.047)	0.850 (0.026)	0.824 (0.055)
年齢	0.070 (0.006)	0.069 (0.014)	0.071 (0.006)	0.073 (0.014)	0.074 (0.006)	0.080 (0.014)	0.068 (0.007)	0.094 (0.016)	0.056 (0.009)	0.103 (0.023)
年齢 2 乗項	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)	-0.001 (0.000)
就業先規模	-0.071 (0.021)	0.057 (0.050)	-0.084 (0.023)	0.099 (0.049)	-0.074 (0.023)	0.035 (0.048)	-0.076 (0.026)	0.003 (0.054)	-0.077 (0.032)	0.129 (0.068)
30~99 人 (100~499 人)	0.010 (0.026)	0.082 (0.056)	-0.009 (0.028)	0.102 (0.054)	-0.008 (0.028)	-0.028 (0.057)	0.023 (0.032)	-0.065 (0.069)	0.044 (0.037)	0.052 (0.075)
500 人以上	0.111 (0.023)	0.123 (0.043)	0.117 (0.024)	0.135 (0.043)	0.128 (0.024)	0.049 (0.042)	0.131 (0.028)	0.040 (0.048)	0.118 (0.035)	0.140 (0.061)
官公庁	0.181 (0.031)	0.218 (0.056)	0.185 (0.034)	0.234 (0.067)	0.228 (0.034)	0.245 (0.056)	0.240 (0.039)	0.176 (0.068)	0.236 (0.046)	0.285 (0.077)
就業月数	0.002 (0.000)	0.000 (0.000)	0.001 (0.000)	0.000 (0.001)	0.001 (0.000)	0.001 (0.000)	0.001 (0.000)	0.000 (0.000)	0.002 (0.000)	0.002 (0.001)
就業月数 2 乗項	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
学歴	-0.005 (0.027)	0.129 (0.046)	-0.004 (0.030)	-0.102 (0.080)	0.018 (0.030)	-0.015 (0.060)	0.043 (0.033)	-0.026 (0.086)	0.044 (0.041)	0.043 (0.086)
(高校卒)										
専門学校卒	0.128 (0.039)	-0.020 (0.078)	0.176 (0.038)	-0.051 (0.083)	0.162 (0.041)	-0.064 (0.096)	0.146 (0.053)	-0.112 (0.098)	0.175 (0.063)	-0.116 (0.123)
短大卒	0.084 (0.025)	0.253 (0.049)	0.066 (0.027)	0.179 (0.052)	0.057 (0.027)	0.215 (0.049)	0.109 (0.031)	0.213 (0.039)	0.112 (0.039)	0.301 (0.061)
大学卒以上	0.240 (0.019)	0.287 (0.040)	0.269 (0.020)	0.270 (0.040)	0.273 (0.020)	0.256 (0.042)	0.284 (0.024)	0.279 (0.045)	0.284 (0.029)	0.255 (0.065)
大都市	0.026 (0.018)	0.034 (0.054)	0.026 (0.019)	-0.028 (0.046)	0.018 (0.018)	0.026 (0.043)	0.011 (0.021)	0.041 (0.049)	0.018 (0.026)	0.053 (0.060)
(その他の市)										
町村	-0.020 (0.023)	0.000 (0.060)	0.010 (0.026)	-0.126 (0.067)	-0.016 (0.027)	-0.089 (0.066)	-0.019 (0.031)	-0.136 (0.069)	0.019 (0.037)	-0.133 (0.066)
地域プロックダミ-	Yes	Yes								
調査年ダミ-	Yes	Yes								
定数項	10.114 (0.122)	10.212 (0.275)	10.128 (0.131)	10.239 (0.302)	10.104 (0.136)	10.112 (0.299)	10.243 (0.166)	9.843 (0.340)	10.412 (0.206)	9.324 (0.498)
月収(対数値)推定値 Y(0), Y(1)	12.2284	12.3065	12.2391	12.3154	12.2403	12.3328	12.2509	12.3352	12.2522	12.3338
平均処理効果 ATE Y(1)-Y(0)	0.0781***	0.0763***	0.0763***	0.0763***	0.0925**	0.0843***	0.0843***	0.0843***	0.1016***	0.1016***
N	6646	1804	6054	1647	5583	1546	4038	1073	2560	667

表 9-2. 教育訓練の月収(対数値)に対する平均処理効果(1-5年後, Weighted モデル)

被説明変数: 月収(対数値)	1年後		2年後		3年後		4年後		5年後	
	非受講者	受講者								
男性	0.887 (0.031) ***	0.831 (0.054) ***	0.920 (0.035) ***	0.786 (0.049) ***	0.896 (0.035) ***	0.805 (0.062) ***	0.915 (0.042) ***	0.679 (0.057) ***	0.832 (0.054) ***	0.879 (0.076) ***
年齢	0.072 (0.009) ***	0.084 (0.016) ***	0.081 (0.011) ***	0.077 (0.017) ***	0.079 (0.011) ***	0.092 (0.020) ***	0.071 (0.014) ***	0.088 (0.021) ***	0.081 (0.016) ***	0.115 (0.031) ***
年齢 2 乗項	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***
就業先規模	-0.089 (0.034) ***	0.091 (0.070) ***	-0.116 (0.036) ***	0.121 (0.063) ***	-0.085 (0.039) ***	0.135 (0.090) ***	-0.060 (0.055) ***	-0.025 (0.073) ***	-0.075 (0.077) ***	0.000 (0.085) ***
30~99 人	-0.010 (0.041)	0.237 (0.087) ***	-0.026 (0.051) ***	0.150 (0.063) ***	-0.025 (0.049) ***	0.059 (0.088) ***	0.069 (0.053) ***	-0.163 (0.101) ***	0.076 (0.075) ***	-0.136 (0.083) ***
100~499 人			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
500 人以上	0.087 (0.037) ***	0.183 (0.059) ***	0.127 (0.040) ***	0.207 (0.058) ***	0.102 (0.044) ***	0.184 (0.077) ***	0.158 (0.055) ***	0.063 (0.070) ***	0.163 (0.069) ***	0.176 (0.086) ***
官公庁	0.162 (0.053) ***	0.236 (0.079) ***	0.192 (0.056) ***	0.293 (0.084) ***	0.296 (0.066) ***	0.301 (0.084) ***	0.264 (0.060) ***	0.237 (0.083) ***	0.221 (0.072) ***	0.233 (0.122) ***
就業月数	0.002 (0.000) ***	0.000 (0.001) ***	0.001 (0.000) ***	-0.001 (0.001) ***	0.001 (0.000) ***	0.000 (0.001) ***	0.001 (0.000) ***	0.000 (0.001) ***	0.002 (0.000) ***	0.001 (0.001) ***
就業月数 2 乗項	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***
学歴	-0.066 (0.054)	0.173 (0.143)	-0.025 (0.062)	-0.128 (0.082) ***	0.041 (0.051) ***	-0.114 (0.080) ***	0.076 (0.058) ***	-0.058 (0.065) ***	0.180 (0.071) ***	0.150 (0.118) ***
(高専卒)			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
専門卒	0.006 (0.101)	-0.113 (0.116)	0.151 (0.087) ***	-0.283 (0.151) *	0.032 (0.107) ***	-0.091 (0.120) ***	0.133 (0.088) ***	-0.440 (0.196) **	0.086 (0.121) ***	-0.437 (0.178) ***
短大卒	0.038 (0.041)	0.251 (0.071) ***	0.059 (0.046) ***	0.126 (0.070) *	0.044 (0.044) ***	0.170 (0.058) ***	0.179 (0.048) ***	0.066 (0.107) ***	0.195 (0.064) ***	0.404 (0.096) ***
大学卒以上	0.211 (0.036) ***	0.273 (0.067) ***	0.229 (0.039) ***	0.218 (0.057) ***	0.238 (0.041) ***	0.192 (0.055) ***	0.236 (0.049) ***	0.233 (0.069) ***	0.257 (0.065) ***	0.145 (0.065) ***
大都市	0.010 (0.034)	0.021 (0.058)	0.035 (0.036) ***	-0.081 (0.058) ***	-0.005 (0.036) ***	0.002 (0.064) ***	-0.091 (0.044) **	0.070 (0.082) ***	-0.068 (0.053) ***	-0.157 (0.091) *
(その他の市)			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
町村	0.019 (0.038)	-0.115 (0.074)	0.036 (0.047)	-0.098 (0.070) ***	-0.009 (0.048) ***	-0.063 (0.111) ***	-0.033 (0.046) ***	-0.200 (0.094) **	0.022 (0.058) ***	-0.274 (0.077) ***
地域プロクダダミ	Yes									
調査年ダミ	Yes									
定数項	10.114 (0.203) ***	9.891 (0.320) ***	9.910 (0.257) ***	10.177 (0.362) ***	10.065 (0.241) ***	9.763 (0.478) ***	10.157 (0.321) ***	10.138 (0.476) ***	10.344 (0.372) ***	9.326 (0.750) ***
月収(対数値)推定値 Y(0), Y(1)	12.22	12.36	12.26	12.33	12.26	12.35	12.29	12.34	12.26	12.29
平均処理効果 ATE Y(1)-Y(0)		0.1319***		0.0706**		0.0822***		0.0554		0.037
N	6646	1804	6654	1647	5583	1546	4038	1073	2560	667

(注) ***: 1%, **: 5%, *: 10%水準で棄却。

表 10-1. 自己啓発の月収(対数値)に対する平均処理効果(1~5年後, Unweighted モデル)

被説明変数：月収(対数値)	1年後		2年後		3年後		4年後		5年後	
	非受講者	受講者								
男性	0.833 (0.015) ***	0.784 (0.035) ***	0.914 (0.018) ***	0.779 (0.038) ***	0.914 (0.020) ***	0.740 (0.040) ***	0.910 (0.025) ***	0.737 (0.047) ***	0.845 (0.035) ***	0.814 (0.059) ***
年齢	0.065 (0.005) ***	0.114 (0.013) ***	0.066 (0.006) ***	0.111 (0.014) ***	0.068 (0.007) ***	0.114 (0.014) ***	0.070 (0.009) ***	0.114 (0.015) ***	0.065 (0.013) ***	0.117 (0.021) ***
年齢2乗項	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***
就業先規模	-0.063 (0.019) ***	-0.058 (0.052) ***	-0.041 (0.022) *	-0.103 (0.050) **	-0.043 (0.025) **	-0.150 (0.055) ***	-0.028 (0.030) *	-0.089 (0.063) *	-0.064 (0.040) *	-0.033 (0.080) *
30~99人	0.007 (0.023)	0.019 (0.056)	0.014 (0.027)	0.008 (0.055)	0.011 (0.030)	-0.084 (0.064)	0.060 (0.035) *	-0.075 (0.073)	0.061 (0.044)	-0.062 (0.088)
100~499人										
500人以上	0.078 (0.019) ***	0.175 (0.049) ***	0.105 (0.023) ***	0.165 (0.047) ***	0.110 (0.026) ***	0.088 (0.050) *	0.147 (0.031) ***	0.109 (0.060) *	0.115 (0.043) ***	0.033 (0.089)
官公庁	0.239 (0.027) ***	0.274 (0.075) ***	0.244 (0.032) ***	0.205 (0.080) ***	0.257 (0.036) ***	0.172 (0.090) *	0.271 (0.045) ***	0.241 (0.105) **	0.264 (0.057) ***	0.207 (0.112) *
就業月数	0.001 (0.000) ***	0.001 (0.000) ***	0.001 (0.000) ***	0.001 (0.000) ***	0.001 (0.000) ***	0.001 (0.000) ***	0.001 (0.000) ***	0.001 (0.000) ***	0.002 (0.000) ***	0.004 (0.001) ***
就業月数2乗項	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***
学歴	0.000 (0.025)	0.007 (0.079)	-0.002 (0.028)	0.030 (0.070)	0.009 (0.032)	0.087 (0.072)	0.010 (0.038)	0.126 (0.081)	0.049 (0.056)	0.113 (0.106)
(諸校本)										
専門学校卒	0.158 (0.035) ***	0.003 (0.085)	0.179 (0.038) ***	-0.078 (0.096)	0.189 (0.044) ***	-0.012 (0.094)	0.180 (0.059) ***	-0.074 (0.111)	0.174 (0.071) **	-0.184 (0.158)
短大卒	0.093 (0.022) ***	0.072 (0.053)	0.089 (0.026) ***	0.094 (0.052) *	0.108 (0.028) ***	0.055 (0.049)	0.130 (0.034) ***	0.037 (0.068)	0.099 (0.046)	0.107 (0.081)
大学卒以上	0.266 (0.017) ***	0.232 (0.035) ***	0.256 (0.020) ***	0.269 (0.038) ***	0.247 (0.024) ***	0.287 (0.043) ***	0.237 (0.030) ***	0.271 (0.048) ***	0.230 (0.039) ***	0.256 (0.063) ***
居住都市規模	0.026 (0.015) *	-0.084 (0.042) **	0.029 (0.018) **	-0.130 (0.042) **	0.035 (0.020) **	-0.061 (0.045)	0.031 (0.025)	0.011 (0.050)	0.038 (0.033)	0.054 (0.064)
(その他の市)										
町村	-0.034 (0.022)	0.079 (0.059)	-0.039 (0.026)	0.085 (0.068)	-0.050 (0.030) *	0.052 (0.074)	-0.081 (0.035) **	-0.006 (0.069)	-0.034 (0.047)	0.059 (0.081)
地域プロダクティビティ	Yes									
調査年ダミー	Yes									
定数項	10.212 (0.106) ***	9.361 (0.292) ***	10.256 (0.126) ***	9.497 (0.299) ***	10.236 (0.153) ***	9.503 (0.315) ***	10.179 (0.200) ***	9.491 (0.347) ***	10.214 (0.281) ***	9.029 (0.502) ***
年収(対数値)推定値 Y(0), Y(1)	12.030	12.304	12.219	12.3269	12.212	12.357	12.226	12.332	12.2235	12.350
平均処理効果ATE Y(1)-Y(0)	0.1064***	0.112***	0.1415***	0.11267***	0.11267***	0.11267***	0.11267***	0.11267***	0.1115***	0.1115***
N	8174	2066	6251	1612	4574	1229	2993	854	1612	472

(注) ***: 1%, **: 5%, *: 10%水準で棄却。

表 10-2. 自己啓発の月収(対数値)に対する平均処理効果(1~5年後, Weighted モデル)

被説明変数：月収(対数値)	1年後		2年後		3年後		4年後		5年後	
	非受講者	受講者								
男性	0.917 (0.025) ***	0.751 (0.049) ***	0.879 (0.030) ***	0.883 (0.054) ***	0.916 (0.032) ***	0.790 (0.073) ***	0.877 (0.041) ***	0.808 (0.081) ***	0.838 (0.046) ***	0.998 (0.059) ***
年齢	0.069 (0.008) ***	0.121 (0.016) ***	0.075 (0.010) ***	0.149 (0.017) ***	0.059 (0.011) ***	0.150 (0.020) ***	0.072 (0.013) ***	0.131 (0.024) ***	0.073 (0.018) ***	0.164 (0.032) ***
年齢2乗項	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.001 (0.000) ***	-0.002 (0.000) ***	-0.001 (0.000) ***	-0.002 (0.000) ***	-0.001 (0.000) ***	-0.002 (0.000) ***	-0.001 (0.000) ***	-0.002 (0.000) ***
就業先規模	-0.088 (0.022) ***	-0.065 (0.077) *	-0.034 (0.036) *	-0.090 (0.063) *	-0.062 (0.042) *	-0.102 (0.110) **	-0.073 (0.061) *	-0.071 (0.122) *	-0.144 (0.075) *	0.098 (0.137) *
30~59人	-0.002 (0.036)	0.143 (0.076) *	0.020 (0.047)	0.008 (0.065)	0.029 (0.049)	-0.017 (0.107)	-0.034 (0.065)	0.030 (0.129)	0.026 (0.076)	0.021 (0.132)
100~499人			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
500人以上	0.082 (0.034) **	0.260 (0.067) ***	0.137 (0.039) ***	0.228 (0.058) **	0.107 (0.047) **	0.236 (0.087) ***	0.130 (0.053) **	0.227 (0.109) **	0.082 (0.073)	0.129 (0.107)
官公庁	0.255 (0.055) ***	0.342 (0.082) ***	0.301 (0.088) ***	0.173 (0.106) *	0.312 (0.066) **	0.259 (0.116) **	0.265 (0.084) **	0.364 (0.155) **	0.240 (0.098) **	0.238 (0.162) **
就業年数	0.001 (0.000) ***	0.002 (0.000) ***	0.001 (0.000) ***	-0.001 (0.001) ***	0.001 (0.000) ***	0.000 (0.000) ***	0.001 (0.000) ***	0.000 (0.000) ***	0.002 (0.001) ***	0.003 (0.001) ***
就業年数2乗項	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***	0.000 (0.000) ***
学歴	-0.018 (0.046)	0.114 (0.151)	0.040 (0.050)	-0.036 (0.090)	0.044 (0.051)	0.122 (0.084)	0.106 (0.061) *	0.127 (0.100)	0.193 (0.080) **	0.157 (0.105)
(高専卒)			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
専門学校卒	0.017 (0.091)	-0.080 (0.136)	-0.004 (0.106)	0.130 (0.148)	0.048 (0.112)	-0.014 (0.088)	0.098 (0.153)	-0.235 (0.144)	0.010 (0.146)	-0.165 (0.170)
短大卒	0.095 (0.037) **	0.075 (0.071)	0.064 (0.043)	0.298 (0.066) ***	0.114 (0.049) **	0.133 (0.086)	0.170 (0.055) **	0.092 (0.097)	0.145 (0.069)	0.167 (0.090)
大学卒以上	0.249 (0.031) ***	0.172 (0.055) ***	0.255 (0.033) ***	0.275 (0.063) ***	0.219 (0.039) **	0.214 (0.076) **	0.207 (0.052) **	0.283 (0.101) ***	0.216 (0.064) **	0.301 (0.081) ***
大都市	-0.002 (0.028)	-0.115 (0.069) *	0.001 (0.031)	-0.146 (0.065) **	-0.019 (0.036)	-0.131 (0.073) *	-0.010 (0.043)	-0.162 (0.077) **	-0.049 (0.052)	-0.091 (0.084)
(その他の市)			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
町村	-0.007 (0.038)	0.041 (0.070)	-0.066 (0.044)	0.204 (0.097) **	-0.066 (0.051)	0.158 (0.155)	-0.135 (0.064) **	0.001 (0.110)	-0.039 (0.071)	0.140 (0.118)
地域プログラマー	Yes									
調査年ダミー	Yes									
定数項	10.295 (0.182) ***	9.209 (0.371) ***	10.111 (0.225) ***	8.610 (0.369) ***	10.485 (0.239) ***	8.693 (0.491) ***	10.240 (0.287) ***	9.123 (0.631) ***	10.146 (0.411) ***	7.795 (0.813) ***
年収(対数値)推定値 Y(0), Y(1)	12.2251	12.2913	12.2496	12.3375	12.2465	12.3327	12.2539	12.2758	12.2710	12.3469
平均処理効果ATE Y(1)-Y(0)	0.0662***		0.0873***		0.0862**		0.0220		0.0789*	
N	8174	2966	6251	1612	4574	1229	2993	854	1612	472

(注) ***: 1%, **: 5%, *: 10%水準で棄却。

表 11. 教育訓練・自己啓発の平均処理効果比較表
自己啓発

教育訓練

		N	ATE	(頑健的) 標準誤差	t 値 or z 値
受講 1 年後	(1) Unmatched	8,450	0.1928	0.0218	8.82 ***
	(2) Unweighted	8,450	0.0781	0.0207	3.78 ***
	(3) Weighted	8,450	0.1320	0.0294	4.49 ***
内生性バイアス(1)-(3)			0.1147		
	脱落バイアス(3)-(2)		-0.0539		
受講 2 年後	(1) Unmatched	7,701	0.1995	0.0228	8.76 ***
	(2) Unweighted	7,701	0.0763	0.0198	3.85 ***
	(3) Weighted	7,701	0.0706	0.0277	2.55 **
内生性バイアス(1)-(3)			0.1232		
	脱落バイアス(3)-(2)		0.0057		
受講 3 年後	(1) Unmatched	7,219	0.2271	0.0233	9.76 ***
	(2) Unweighted	7,219	0.0925	0.0196	4.73 ***
	(3) Weighted	7,219	0.0823	0.0313	2.62 *
内生性バイアス(1)-(3)			0.1448		
	脱落バイアス(3)-(2)		0.0103		
受講 4 年後	(1) Unmatched	5,111	0.2243	0.0275	8.16 ***
	(2) Unweighted	5,111	0.0844	0.0236	3.57 ***
	(3) Weighted	5,111	0.0554	0.0378	1.47
内生性バイアス(1)-(3)			0.1399		
	脱落バイアス(3)-(2)		0.0290		
受講 5 年後	(1) Unmatched	3,227	0.1894	0.0348	5.44 ***
	(2) Unweighted	3,227	0.1016	0.0279	3.64 ***
	(3) Weighted	3,227	0.0370	0.0438	0.85
内生性バイアス(1)-(3)			0.0878		
	脱落バイアス(3)-(2)		0.0646		

		N	ATE	(頑健的) 標準誤差	t 値 or z 値
受講 1 年後	(1) Unmatched	10,240	0.1768	0.0205	-8.65 ***
	(2) Unweighted	10,240	0.1064	0.0183	5.80 ***
	(3) Weighted	10,240	0.0663	0.0304	2.18 **
内生性バイアス(1)-(3)			0.0705		
	脱落バイアス(3)-(2)		0.0401		
受講 2 年後	(1) Unmatched	7,863	0.1790	0.0229	7.80 ***
	(2) Unweighted	7,863	0.1120	0.0191	5.87 ***
	(3) Weighted	7,863	0.0879	0.0299	2.94 ***
内生性バイアス(1)-(3)			0.0670		
	脱落バイアス(3)-(2)		0.0241		
受講 3 年後	(1) Unmatched	5,803	0.2075	0.0261	7.94 ***
	(2) Unweighted	5,803	0.1415	0.0210	6.75 ***
	(3) Weighted	5,803	0.0862	0.0390	2.21 **
内生性バイアス(1)-(3)			0.0659		
	脱落バイアス(3)-(2)		0.0553		
受講 4 年後	(1) Unmatched	3,847	0.1886	0.0312	6.04 ***
	(2) Unweighted	3,847	0.1127	0.0238	4.74 ***
	(3) Weighted	3,847	0.0220	0.0427	0.52
内生性バイアス(1)-(3)			0.0759		
	脱落バイアス(3)-(2)		0.0907		
受講 5 年後	(1) Unmatched	2,084	0.1279	0.0424	3.02 ***
	(2) Unweighted	2,084	0.1115	0.0313	3.56 ***
	(3) Weighted	2,084	0.0789	0.0474	1.66 ***
内生性バイアス(1)-(3)			0.0165		
	脱落バイアス(3)-(2)		0.0325		

注) ***: 1%, **: 5%, *: 10%水準で棄却。

Unmatched は t 値, 標準誤差, それ以外は z 値, 頑健的標準誤差。

注) ***: 1%, **: 5%, *: 10%水準で棄却。

Unmatched は t 値, 標準誤差, それ以外は z 値, 頑健的標準誤差。

表 12. Sample Selection Indicator

Type		$T = 0$	$T = 1$
g_{11}	$S_0 = 1, S_1 = 1$	Observed	Observed
g_{01}	$S_0 = 0, S_1 = 1$	Not Observed	Observed
g_{10}	$S_0 = 1, S_1 = 0$	Observed	Not Observed
g_{00}	$S_0 = 0, S_1 = 0$	Not Observed	Not Observed

引用) DiNard *et al.*(2006)p. 16.

ることが推察される。これが職業訓練への参加行動と回答行動との間に直接的な関係があるのか、あるいは、訓練参加者と継続回答者が持つ属性が類似したことによる影響なのかは特定できていない。

6. Bounds 推定による処置効果推定¹⁹⁾

本節では、Bounds 推定を用いて、脱落を考慮した処置効果の測定結果についてみたい。まずは、脱落をはじめとする Missing Data 問題を Bounds 推定に適用させた Horowitz and Manski(2000) の Bounds 推定について紹介する。処置効果の Upper Bounds と Lower Bounds はそれぞれ、以下の通りとなる。

$$\begin{aligned}\bar{\theta}_M &= E[y|T=1]P[S=1|T=1] \\ &\quad + \bar{Y}(1-P[S=1|T=1]) \\ &\quad - E[y|T=0]P[S=1|T=0] \\ &\quad + \underline{Y}(1-P[S=1|T=1]) \\ \underline{\theta}_M &= E[y|T=1]P[S=1|T=1] \\ &\quad + \underline{Y}(1-P[S=1|T=1]) \\ &\quad - E[y|T=0]P[S=1|T=0] \\ &\quad + \bar{Y}(1-P[S=1|T=1])\end{aligned}$$

$P[S=1|T=1]$: 処置群(受講者)の継続回答率
 $P[S=1|T=0]$: 対照群(非受講者)の継続回答率
 S : 継続回答ダミー ($S=1$: 継続回答(観測可能)
 $S=0$: 脱落(観測不可能) \underline{Y} : 結果変数の最小値,
 \bar{Y} : 結果変数の最大値

上記の推定では、範囲かなり広がる可能性があるものの、Bounds の設定に対する仮定が少ないため、ベンチマークとする(Worst-Case Scenario)。Horowitz and Manski(2000)の Bounds 推定を発展させ、選択過程に対して新たな仮説を追加し、より限定的な Bounds 推定を考案したのが Lee(2009)の推定であった(Lee's Bounds)。

ここでは、Sample Selection Indicator g を用いて、処置群において観測できるもの ($S_1=1$) と観測できないもの ($S_1=0$)、対照群において観測できるもの ($S_0=1$) と観測できないもの ($S_0=0$) とし、対象者を以下の4つのグループに分類する(表 12)。

① g_{11} : いずれの群でも観測できる ($S_1=1, S_0=1$),
 ② g_{01} : 処置群であれば観測できるが、対照群では観測できない ($S_1=1, S_0=0$), ③ g_{10} : 対照群であれば観測できるが、処置群では観測できない ($S_1=0, S_0=1$), ④ g_{00} : いずれの群でも観測できない ($S_1=0, S_0=0$)。④の分類が回答拒否や脱落にあたる。

観察できるかどうか(継続回答するかどうか)を式で示すと、 $S=S_1T+S_0(1-T)$ となる。これは、調査対象者は処置群に割り振られ、 $S_1=1$ であるか、対照群に割り振られ、 $S_0=1$ であるなら、観測できるということの意味する。

Lee(2009)では、以下の2つの仮定を設定し、常に観察できるタイプ g_{11} の対象者の処置効果の Bounds を設定する。

(1) Random Assignment of Treatment: 処置はランダムに割り振られる

(2) Monotonicity: 唯一処置の割り当てが、調査対象者が継続回答するか、脱落するかに影響する

観察可能な継続回答者の平均処置効果 θ (ATE) のための Bounds の設定方法は以下の通りである。

$S_1 > S_0$ と仮定した場合、処置群の継続回答率は対照群の継続回答率より高い(翻せば、処置群の脱落率は、対照群と比べると低い)ことを示す。

(Monotonicity の仮定の下で)観測された処置群の個人の情報は、 g_{11} (常に観測できる)と g_{01} (処置群なら観察できるが、対照群なら観察できないもの)の組み合わせである。

この場合、Lee(2009)による Bounds は、以下のように示される。観察された個体の割合が両方の群に対して同等となるようにサンプルを除外する。

$$\begin{aligned}\bar{\theta}_L &\equiv E[Y|T=1, S=1, Y \geq y_{(\rho_0)}] \\ &\quad - E[Y|T=0, S=1] \\ \underline{\theta}_L &\equiv E[Y|T=1, S=1, Y \leq y_{(1-\rho_0)}] \\ &\quad - E[Y|T=0, S=1] \\ y_{(\rho_0)} &\equiv G_{S=1, T=1}^{-1}(\rho_0) \\ \rho_0: \text{除外基準 } \rho_0 &\equiv \frac{P[S=1|T=1] - P[S=1|T=0]}{P[S=1|T=1]} \\ G_{S=1, T=1}^{-1}: T=1, S=1 \text{ 時の逆累積分布関数}\end{aligned}$$

例えば、継続回答率が処置群は70%、対照群は60%であるとする。ここでは、推定された継続回答率が高い群の対象者数を(余分なものとして)取り除く。

$$\begin{aligned} \text{除外基準は, } & \frac{P[S=1|T=1]-P[S=1|T=0]}{P[S=1|T=1]} \\ = & \frac{0.7-0.6}{0.6} = \frac{0.1}{0.6} = 0.1666\cdots = 16.7\% \end{aligned}$$

Upper Bound $\bar{\theta}_L$ の計算方法の手順は以下の通りである。

(1) 対照群の平均値を計算

(2) 処置群から下位 16.7% より低い値を削除したのちに、処置群の平均を計算

(3) (1), (2) で計算した両平均の差をとり、 $\bar{\theta}_L$ を得る

Lower Bound θ_L の計算方法の手順は以下の通りである

(1)' 処置群から上位 83.3% より高い値を削除した後、処置群の平均を計算

(2)' 対照群の平均と、(1)' で作成した対照群の平均の差をとり、 θ_L を得る

ここから、Lee (2009) の Bounds 推定では、回答継続あるいは脱落に対して、説明力がある説明変数を用いて、Bounds の幅を狭くする。具体的には、継続回答確率を操作するようなデータ収集方法を利用する。DiNarido *et al.* (2006) では、その代理変数として、調査員による調査者宅への訪問回数、平日だけではなく休日にも訪問するなど回収努力を挙げている。前述した通り、KHPS では、調査手順などがわかる調査員確認票が利用可能であるため、本研究でもこれを用いた分析を行う。第 4 節で継続回答に対する規定要因を検証した際に、統計的に有意だった、訪問回数と共に、子ども (18 歳未満) がいるかどうか、健康状態 (1: よい, まあよい, 0: ふつう, あまりよくない, よくない) を変数として利用する。ここで扱う変数は全てダミー変数などにすることで、サンプルを分割し、グループ毎の平均値が、更にはグループごとの加重平均が計算できる。

単純化のために、回答継続に影響する変数 Z をダミー変数とし、Bounds 幅を狭めることができる変数として想定する。 $Z=1$ であるなら、調査員によって、調査回収のために努力が払われた、 $Z=0$ であるなら、調査員によって、努力がなされなかったとする。

Lee's Bounds の計算は以下の通りである。

$$\begin{aligned} \bar{\theta}_L^{Z=1 \text{ or } 0} = & \\ E[Y|T=1, S=1, Z=1 \text{ or } 0, Y \geq y_{(p_0)}^{Z=1 \text{ or } 0}] & \end{aligned}$$

$$-E[Y|T=0, S=1, Z=1 \text{ or } 0]$$

$$\theta_L^{Z=1 \text{ or } 0} =$$

$$E[Y|T=1, S=1, Y \leq y_{(1-p_0)}^{Z=1 \text{ or } 0}, Z=1 \text{ or } 0]$$

$$-E[Y|T=0, S=1, Z=1]$$

$$y_{(p_0)}^{Z=1} = G_{S=1, T=1, Z=1}^{-1}(p_0),$$

$$y_{(p_0)}^{Z=0} = G_{S=1, T=1, Z=0}^{-1}(p_0)$$

$$p_0 =$$

$$\frac{P[S=1|T=1, Z=1]-P[S=1|T=0, Z=1]}{P[S=1|T=1, Z=1]}$$

全対象の処置効果の条件付きの各 Bound は以下の式のように示すことができる。

$$q_1 \bar{\theta}_L^{Z=1} + (1-q_1) \bar{\theta}_L^{Z=0}$$

$$q_1 \theta_L^{Z=1} + (1-q_1) \theta_L^{Z=0}$$

$$q_1 = P(Z=1|T=1)$$

これは、前者は条件を課していない Upper Bound $\bar{\theta}_L$ より小さく、後者は条件を課していない Lower Bound θ_L よりも大きいいため、Bounds の幅が無条件のものよりも狭くなる。

次に、Lee (2009) の Bounds 推定の結果について説明する。表 13-1, 2 では、条件を課さない Worst-Case Scenario と条件付きの Lee's Bounds の推定結果を示した。前述の通り、説明変数には、調査員の訪問回数 (平均回数以上=1, それ以外=0)、子ども (18 歳未満) がいるかどうか、健康状態 (よい・まあよい=1, ふつう・あまりよくない・よくない=0) を用いる。推定結果をみると、自己啓発の受講 5 年後を除き、教育訓練、自己啓発ともに、Lower Bound, Upper Bound ともに 0 を上回っており、賃金を上昇させる効果があることが確認される。

また、説明変数による制限をかけることで、平均処置効果 (ATE) の Bounds 幅が狭まり、教育訓練は約 4~9% 減少、自己啓発は 1~2 年後は予想とは逆に幅が広がり、3 年後以降は幅が約 1~9% 減少していることが確認された。

これらの推定値と、5 節で推定した IPWRA 推定値とを比較すると (図 1, 2)、教育訓練の IPWRA 推定値の多くは Bounds 幅には含まれておらず、自己啓発の IPWRA 推定値では、Unweighted モデルのみが Bounds 幅に収まっているが、脱落バイアスを考慮した Weighted モデルでは Bounds の下限値よりも低くなっている。総じて、IPWRA 推定による推定値と比べて部分識別による推定値は大きな値をとっている。

この結果の背景には、Bounds 推定における処置はランダムに割り振られるという仮定、内生性を考

表 13-1. Bounds 推定(教育訓練)

教育訓練		係数	標準誤差	z 値		標本数	除外されない 標本数	除外基準
1 年後	Lower	0.108	(0.023)	4.77	***	9440	8450	0.0606
	Upper	0.317	(0.023)	13.84	***			
2 年後	Lower	0.119	(0.024)	5.04	***	8769	7708	0.0622
	Upper	0.323	(0.024)	13.36	***			
3 年後	Lower	0.141	(0.024)	5.82	***	8296	7142	0.0737
	Upper	0.365	(0.025)	14.82	***			
4 年後	Lower	0.135	(0.029)	4.67	***	6064	5119	0.0622
	Upper	0.340	(0.030)	11.16	***			
5 年後	Lower	0.085	(0.038)	2.22	**	3916	3235	0.0813
	Upper	0.347	(0.038)	9.04	***			

注) *** : 1%, ** : 5%, * : 10% 水準で棄却.

教育訓練		係数	標準誤差	z 値		標本数	除外されない 標本数	除外基準
1 年後	Lower	0.107	(0.024)	4.46	***	9440	8450	0.0606
	Upper	0.307	(0.024)	12.97	***			
2 年後	Lower	0.116	(0.026)	4.52	***	8769	7708	0.0622
	Upper	0.301	(0.026)	11.6	***			
3 年後	Lower	0.130	(0.026)	5.09	***	8296	7142	0.0737
	Upper	0.344	(0.026)	13.47	***			
4 年後	Lower	0.133	(0.033)	4.02	***	6064	5119	0.0622
	Upper	0.326	(0.034)	9.69	***			
5 年後	Lower	0.057	(0.048)	1.2	**	3916	3235	0.0813
	Upper	0.304	(0.046)	6.62	***			

注) *** : 1%, ** : 5%, * : 10% 水準で棄却.

慮していないという影響が考えられる。しかしながら、自己啓発の平均処置効果だけをみると、IPWRA 推定値の一部が Bounds 内に収まっていることから、上記の仮定を逸脱した場合であっても、内生性を考慮した点識別に基づく推定値を部分的に Bounds 内に包摂しうる。

7. 結論

本研究では、職業訓練が賃金に与える平均処置効果推定における内生性バイアスと脱落バイアスをそれぞれ測定した。過去の先行研究では、内生性に関する影響を考慮した分析が多く存在したが、処置である職業訓練と脱落との関係について論じられているものは少ない。

職業訓練と継続回答との関係を確認すると、調査当初における受講・非受講者の比較では、受講者の継続回答率が高いが、それ以降では、非受講者の回答率が高くなるか、両者の差がなくなる。職業訓練が就業者の能力および賃金の向上を促すと想定する

場合、受講者の継続回答率が高いことは、職業訓練による賃金への影響である処置効果が過大に推定される可能性がある。ここでは、脱落による影響を考慮すべく、継続回答率の逆数を Weight として用いて推定を、また、職業訓練受講の内生性を考慮すべく、IPWRA 推定を併せて行った。

その結果、職業訓練の受講者と非受講者との単純な平均賃金差の測定値と比べて、IPWRA 推定値は小さく、内生性バイアスによる過大推定が確認された。加えて、IPWRA 推定値を脱落を考慮した Weighted モデルと考慮しない Unweighted モデルとを比較したところ、前者の推定値と比べて後者の方が高くなっており、ここでは脱落バイアスによる過大推定が確認された。両バイアスを比較したところ、受講後の年数を経るごとに、内生性バイアスより脱落バイアスの方が大きくなった。

また合わせて、脱落過程に対する厳しい仮定を置かない、平均処置効果の Bounds 推定結果をみると、ほとんどのケースで、教育訓練、自己啓発ともに平

表 13-2. Bounds 推定(自己啓発)

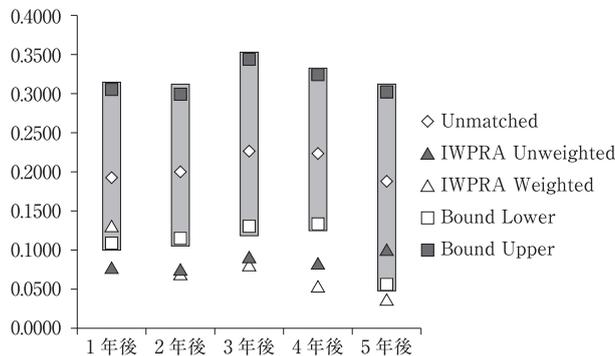
自己啓発		係数	標準誤差	z 値		標本数	除外されない 標本数	除外基準
1 年後	Lower	0.115	(0.023)	5.01	***	11413	10240	0.043
	Upper	0.281	(0.024)	11.68	***			
2 年後	Lower	0.113	(0.026)	4.37	***	8965	7877	0.0448
	Upper	0.275	(0.027)	10.25	***			
3 年後	Lower	0.130	(0.029)	4.48	***	6789	5814	0.0537
	Upper	0.323	(0.031)	10.53	***			
4 年後	Lower	0.085	(0.035)	2.42	**	4593	3855	0.0706
	Upper	0.328	(0.036)	9.18	***			
5 年後	Lower	-0.013	(0.049)	-0.26		2558	2091	0.0918
	Upper	0.298	(0.050)	5.91	***			

注) ***: 1%, **: 5%, *: 10% 水準で棄却。

自己啓発		係数	標準誤差	z 値		標本数	除外されない 標本数	除外基準
1 年後	Lower	0.104	(0.024)	4.25	***	11413	10240	0.043
	Upper	0.277	(0.025)	11.15	***			
2 年後	Lower	0.098	(0.029)	3.42	***	8965	7877	0.0448
	Upper	0.264	(0.030)	8.83	***			
3 年後	Lower	0.126	(0.033)	3.85	***	6789	5814	0.0537
	Upper	0.302	(0.033)	9.13	***			
4 年後	Lower	0.089	(0.040)	2.19	**	4593	3855	0.0706
	Upper	0.318	(0.038)	8.3	***			
5 年後	Lower	-0.026	(0.056)	-0.47		2558	2091	0.0918
	Upper	0.281	(0.053)	5.29	***			

注) ***: 1%, **: 5%, *: 10% 水準で棄却。

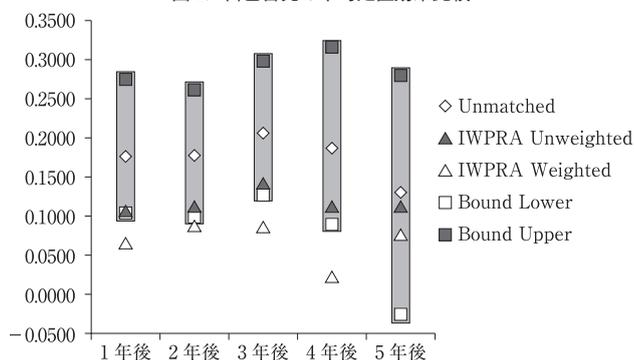
図 1. 教育訓練の平均処置効果比較



均処置効果の Bounds がゼロ，負値を含んでおらず，賃金を押し上げる効果があることが確認された。また，この Bounds 推定値と前述の IPWRA 推定値とを比較したところ，教育訓練の IPWRA 推定値の多くは Bound 内には含まれておらず，自己啓発の IPWRA 推定値では，Unweighted モデルのみが Bound 内に収まっているが，脱落バイアスを考慮

した Weighted モデルでは Bounds の下限値よりも低く，総じて，IPWRA 推定による推定と比べ Bounds 推定は大きな値となった。この原因の一つに，Bounds 推定における仮定(処置はランダムに割り振られる)があり，内生性を考慮していないことで過大推定が発生していると考えられる。しかし，自己啓発の平均処置効果では，IPWRA 推定値の一

図2. 自己啓発の平均処置効果比較



部が Bounds 内に収まっており、上記の仮定を逸脱した場合であっても、内生性を考慮した点識別に基づく推定値を部分的に Bounds 内に包摂する結果を得た。

本来であれば、本稿で行った各種の補正方法の実用性やパフォーマンスを理解するために、実データに基づく分析だけでなく、人為的に割り当てをシミュレーションし、様々な条件下における脱落による影響を測定すべきであるが、それは別稿に譲りたい。

(群馬大学社会情報学部)

注

1) この研究は、科学研究費補助金基盤研究C「世帯内の資源配分および不平等に関する実証研究」(15K03502)、「社会的規範と世帯内分配に関する実証分析」(18K01643)の支援を受けている。「慶應義塾パネル調査(Keio Household Panel Survey)」のデータの利用については、慶應義塾大学経済研究所パネルデータ設計・解析センターの許可をいただいている。

なお、一橋大学経済研究所定例研究会での本論文の報告に対し、討論者の坂口尚文氏(慶應義塾大学)、同研究会参加者から多くの示唆に富んだコメントと助言をいただいた。記して感謝したい。本論文の内容および誤りについては、著者に帰すものである。

2) 「インターホンで話した」、「本人も家族にも会えなかった」、「本人・配偶者を除く、その他の家族に会えた」、「配偶者に会えた」、「本人に会えた」など。

3) 対象者から調査票を有効票として回収した。調査票を対象者から回収したが、記入状況が悪く、欠票となった。調査票を受け取ってもらえなかった(調査拒否された)。調査拒否以外の理由で調査票を渡せなかった。

4) 「肉体的または精神的な健康状態が良くないから」、「忙しくて記入している時間がない」、「調査項目の量が多すぎて記入するのが大変」、「一般のアンケート類に対して不信感があるから」、「本調査の調査票に対して不信感があるから」、「前回と同様の調査項目で、

もう十分に協力したからこれ以上やりたくない」、「調査項目が難しく理解しづらいから」、「離婚・死別・失業など、プライバシーにかかわる不愉快な質問があり答えたくない」、「家族の事情(災害など含む)のために答える気にならない」、「家族が反対するから」、「漠然とした理由、やりたくない、迷惑、面倒とを感じるから」、「挨拶状やその後の報告などを受け取っていないから」、「謝礼が少ないから」など。

5) 「非常に強い」「強い」「弱い」「別の機会ならば、受けてもらえそうだった」

6) それぞれ継続回答に対する影響としては、訪問回数が多いほど、調査員が回答者に会えない(負の影響)、調査員の訪問日が休日であるほど回答者に会える(正の影響)、初回訪問月が早いほど、回答者に余裕をもって回答してもらえる(正の影響)などが考えられる。

7) 第1段階では地方(8ブロック)・都市階級(3区分)により24層に層化し、各層に住民基本台帳人口の人口割合に合わせ標本を配分し、その上で、1つの調査地点あたりの抽出数を10程度(2007、2012年では5程度)として、調査地点数を決定する。調査地点は、抽出単位として、国勢調査の調査区を使用している。第2段階では、選定された調査地点の住民基本台帳を抽出台帳として、調査対象適格者を対象に、指定された起番号、抽出間隔に基づき1調査地点について、約10人(2007、2012年は5人)を抽出している。

また、予備対象を事前に用意し、予定された標本サイズ(4,000、1,400、1,000)を確保する。予備対象は、正規対象と同じ調査区内に居住し、同じ性別、同じ年齢区分から選ばれている。

8) 詳しくは、慶應義塾大学パネルデータ設計・解析センターのホームページを参照されたい。http://www.pdrc.keio.ac.jp/

9) KHPSでは当該項目について、受講の有無、受講期間、費用負担者について尋ねている(2004~2008年)。「過去1年間の間に、あなたの今お勤めの会社からの指示により、あるいは、会社から派遣されて、教育訓練や研修、講習会、学校の授業を受けましたか。」「過去1年間に、すべてあわせて何日ぐらい教育訓練や研修、講習会に参加しましたか。」「その教育訓練や研修、講習会の費用を主に負担したのは誰ですか。」

10) KHPSでは当該項目について、受講の有無、

受講方法、受講時間および費用について尋ねている。「あなたは昨年2月から現在までの1年間の間に、自分の意志で仕事にかかわる技術や能力の向上のための取り組み(例えば、学校に通う、講座を受講する、自分で勉強する、など)をしましたか。」「どのような方法で学びましたか。」「あなたはそれらの方法で学ぶために、1ヶ月の間で平均どのくらいの時間を使用しましたか。また、1ヶ月の間で平均どのくらいの費用がかかりましたか。」

11) KHPS では、さらに自己啓発については、②A: 自己啓発を行った期間(時間/月)、②B: 自己啓発にかけた費用(万円/月)、②C: 自己啓発の受講密度を把握するために、参加・受講した種類をみることができるが、本研究では扱わない。

12) 「教育訓練」に関する設問が2004~2008年と限定されているため、調査前半における教育訓練受講者を対象としている。

13) a_i は個人 i が脱落するかどうかの指標を、 ϕ は a の分布を特徴づける既知または未知のパラメータを示すものとする。ここで、 $a_i = j$ は j 時点で脱落し、 $j-1$ 時点以前の変数 ($y_{i1}, \dots, y_{i,j-1}$) は観測可能だが、時点以降の変数 (y_{ij}, \dots, y_{iT}) は観測不可能であることを意味する。

① Missing Completely At Random (MCAR): 脱落が完全にランダムな場合、過去 ($j-1$ 時以前)、また脱落が発生する j 時点、それ以降 ($j+1$ 時以降) の変数にも依存しない。脱落は無視可能 (Ignorable) となる。

$$Pr(a_i = j | y_{i1}, \dots, y_{iT}; \phi) = Pr(a_i = j | \phi) \\ \text{for all } y_{i1}, \dots, y_{iT}$$

② Missing At Random (MAR): 脱落は観察可能な変数 ($y_{i1}, \dots, y_{i,j-1}$) にも依存する。

$Pr(a_i = j | y_{i1}, \dots, y_{iT}; \phi) = Pr(a_i = j | y_{i1}, \dots, y_{i,j-1}; \phi)$
 j 時点で脱落するかどうかは、 $j-1$ 以前の観測値から推測可能となる。

③ Missing Not At Random (MNAR): 脱落は、脱落時点以降の値 (y_{ij}, \dots, y_{iT}) にも依存する。

$Pr(a_i = j | y_{i1}, \dots, y_{iT}; \phi) = Pr(a_i = j | y_{ij}, \dots, y_{iT}; \phi)$
 j 時点で脱落するかどうかは、 $j-1$ 以前の観測値から推測不可能となる。

14) MAR の考え方にに基づき、継続回答率に影響するのが過去の情報とするならば、 $z_{i,t-1}$ と表記すべきところだが、ここでは、調査直前、調査年と同年の情報を利用していることから、 z_{it} とした。

15) IPWRA の説明については、星野(2009)、Wooldridge(2010)を参考にした。

16) 本節の説明は、Jann(2019)に依拠している。

17) 個体 i の β の影響値は以下のように計算できる。

$$\lambda_i^\beta = w_i \frac{N}{W} \mathbf{Q} x_i (y_i - x_i \beta) \sqrt{\frac{N-1}{N-k}}$$

$$\mathbf{Q} = (E[X'X])^{-1} = \left(\frac{1}{W} \sum_{i=1}^N w_i x_i x_i' \right)^{-1}$$

$X = (x_1', x_2', \dots, x_k')$ は説明変数の行列、 $w_i, i=1, \dots, N$ は Weight (継続回答確率の逆数) をそれぞれ示している。

18) これらの変数の影響関数は次のように記述できる。

$$\lambda_i^{\bar{y}^1} = w_i T_i \frac{N}{W_1} (1 - \hat{p}) \bar{x}_0 Q_1 x_i' (y_i - x_i \bar{\beta}_1) \\ + w_i T_i \frac{N}{W_1} \hat{p} (y_i - \bar{y}_1) \\ + w_i (1 - T_i) \frac{N}{W_0} (1 - \hat{p}) (x_i - \bar{x}_i) \bar{\beta}_1 \\ + w_i \frac{N}{W} (\bar{y}_1 - \bar{x}_0 \bar{\beta}_1) (T_i - \hat{p})$$

$$\lambda_i^{\bar{y}^0} = w_i (1 - T_i) \frac{N}{W_0} \hat{p} \bar{x}_0 Q_0 x_i' (y_i - x_i \bar{\beta}_0) \\ + w_i (1 - T_i) \frac{N}{W_0} (1 - \hat{p}) (y_i - \bar{y}_0) \\ + w_i T_i \frac{N}{W_1} \hat{p} (x_i - \bar{x}_i) \bar{\beta}_0 \\ + w_i \frac{N}{W} (\bar{x}_i \bar{\beta}_0 - \bar{y}_0) (T_i - \hat{p})$$

Q_0, Q_1 はそれぞれ対照群と処置群の Q を指す。

ATT は、次の式から得られる。

$$\lambda_i^{ATE} = \lambda_i^{\bar{y}^1} - \lambda_i^{\bar{y}^0}$$

19) 本節の Lee's Bound 推定の説明は、DiNardo *et al.* (2006) を参考にしている。

参考文献

- 原ひろみ(2014)『職業能力開発の経済分析』勁草書房。
星野崇宏(2009)『調査観察データの統計科学』岩波書店。
市村英彦(2010)「ミクロ実証分析の進展と今後の展望」, 日本経済学会編『日本経済学会75年史』有斐閣, pp. 289-361。
岩崎学(2002)『不完全データの統計解析』エコノミスト社。
川口大司(2008)「労働政策評価の計量経済学」『日本労働研究雑誌』No. 579, pp. 16-28。
小林徹・佐藤一磨(2013)「自己啓発の実施と再就職・失業・賃金」, 瀬古美喜, 照山博司, 山本勲, 樋口美雄, 慶應-京大連携グローバルCOE編著『日本の家計行動のダイナミズム[IX] 家計パネルデータからみた市場の質』pp. 85-116。
黒澤昌子(2005)「積極労働政策の評価レビュー」『ファイナンシャル・レビュー』77号, pp. 197-220
McKenzie, C. R., 直井道生, 宮内環, 木曾研介(2007)「労働市場における個人行動とサンプル脱落問題」, 樋口美雄, 瀬古美喜, 慶應義塾大学経商連携21世紀COE『日本の家計行動のダイナミズム[III]』慶應義塾大学出版会, pp. 13-75。
宮内環, C. R. McKenzie, 木村正一(2005)「回答行動の分析——調査と拒否の選択行動」, 樋口美雄, 慶應義塾大学経商連携21世紀COE『日本の家計行動のダイナミズム[I]』慶應義塾大学出版会, pp. 43-91。
宮内環, C. R. McKenzie, 木村正一(2006)「パネルデータ継続と回答行動の分析」, 樋口美雄, 慶應義塾大学経商連携21世紀COE『日本の家計行動のダイナミズム[II]』pp. 9-52。
直井道生(2007)「家計の住居移動行動とサンプル脱落問題」, 樋口美雄, 瀬古美喜『日本の家計行動のダイナミズム[III]』pp. 77-98。
奥村綱雄(2018)『部分識別入門 計量経済学の革新的

- アプローチ』日本評論社。
- 戸田淳仁・樋口美雄(2005)「企業による教育訓練とその役割の変化」, 樋口美雄・児玉俊洋・阿部正浩編著『労働市場設計の経済分析——マッチング機能の強化に向けて』東洋経済新報社, pp. 251-281.
- 吉田恵子(2004)「自己啓発が賃金に及ぼす効果の実証分析」『日本労働研究雑誌』No. 532, pp. 40-53.
- Abadie, Alberto, Joshua Angrist and Guido Imbens (2002) “Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings,” *Econometrica*, Vol. 70, No. 1, pp. 91-117.
- Angrist, Joshua, Eric Bettinger and Michael Kremer (2006) “Long-term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia,” *American Economic Review*, Vol. 96, No. 3, pp. 847-862.
- Bates, Nancy (2003) “Contact Histories in Personal Visit Surveys: The Survey of Income and Program Participation (SIPP) Methods Panel,” *Proceedings of the Section Research Methods*, American Statistical Association, pp. 7-14.
- Bloom, Howard S., Jarry L. Orr, Stephen H. Bell, Geroge Cave, Fred Doolittle, Winston Lin and Johannes M. Bos (1997) “The Benefits and Costs of JTPA Title II-A Programs: Key Findings from the National Job Training Partnership Act Study,” *Journal of Human Resources*, Vol. 32, No. 3, pp. 549-576.
- Busso, Matias, John DiNardo and Justin McCrary (2014) “New Evidence on the Finite Sample Properties of Propensity Score Reweighting and Matching Estimators,” *Review of Economics and Statistics*, Vol. 96, No. 5, pp. 885-897.
- Campanelli Pamela C., Patrick Sturgis and Susan Purdon (1997) “Can You Hear Me Knocking: An investigation into the Impact of Interviewers on Survey Response Rates,” The Survey Methods Centre at SCPR, London.
- DiNardo, John, Justin McCrary and Lisa Sanbonmatsu (2006) “Constructive Proposals for Dealing with Attrition: An Empirical Example,” University of Michigan Working Paper.
- Fitzgerald, John, Peter Gottschalk and Robert Moffitt, 1998a. “An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics,” *Journal of Human Resources*, Vol. 33, No. 2, pp. 251-299.
- Fitzgerald, John, Peter Gottschalk and Robert Moffitt, 1998b. “An Analysis of the Impact of Sample Attrition on the Second Generation of Respondents in the Michigan Panel Study of Income Dynamics,” *Journal of Human Resources*, Vol. 33, No. 2, pp. 300-344.
- Grasdal, Astrid (2001) “The Performance of Sample Selection Estimators to Control for Attrition Bias,” *Health Economics*, Vol. 10, No. 5, pp. 385-398.
- Heckman, James J. (1976) “The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models,” *Annals of Economic and Social Measurement*, Vol. 5, No. 4, pp. 475-492.
- Heckman, James J. (1979) “Sample Selection as a Specification Error,” *Econometrica*, Vol. 47, No. 1, pp. 153-161.
- Hirano, Keisuke, Guido W. Imbens and Geert Ridder (2003) “Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score,” *Econometrica*, Vol. 71, No. 4, pp. 1161-1189.
- Horowitz L. Joel and Charles F. Manski (2000) “Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data,” *Journal of the American Statistical Association*, Vol. 95, No. 449, pp. 77-84.
- Jann, Ben (2019) “Influence Functions for Linear Regression (With an Application To Regression Adjustment),” University of Bern Social Sciences Working Paper No. 32.
- Ksoll, Christopher, Jenny Aker, Danielle Miller, Karla C. Perez-Mendoza and Susan L. Smalley (2014) “Learning Without Teachers? A Randomized Experiment of a Mobile Phone-based Adult Education Program in Los Angeles,” Center for Global Development Working Paper 368.
- LaLonde, Robert (1986) “Evaluating the Econometric Evaluations of Training Programs with Experimental Data,” *American Economic Review*, Vol. 76, No. 4, pp. 604-620.
- Lee, David (2009) “Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects,” *Review of Economic Studies*, Vol. 76, No. 3, pp. 1071-1102.
- Lynn, Peter, Patrick Sturgis, Paul Clarke and Martin, Jean (2002) “The Effects of Extended Interviewer Efforts on Nonresponse Bias, in *Survey Nonresponse* eds. Groves, Robert M., Dillman, Don A., Eltinge, John L. and Roderick, J. A. Little, Wiley & sons, Inc, New York, pp. 135-148.
- Molina, Millan Teresa and Karen Macours (2017) “Attrition in Randomized Control Trials: Using Tracking Information to Correct Bias,” IZA Discussion Paper, No. 10711.
- Puhani, Patrick, A. (2000) “The Heckman Correction for Sample Selection and Its Critique,” *Journal of Economic Surveys*, Vol. 14, No. 1, pp. 53-68.
- Rosenbaum Paul R. and Donald B. Rubin (1983) “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, Vol. 70, No. 1, pp. 41-55.
- Scharfstein, O. Daniel, Andrea Rotnitzky and James M. Robins (1999) “Rejoinder,” *Journal of the American Statistical Association*, Vol. 94, No. 448, pp. 1135-1146.
- Stoop, Ineke, Jaak Billiet Achim Koch and Rory Fitzgerald (2010) *Improving Survey Response: Lessons learned from the European Social Survey*, John

- Wiley & Sons, Ltd.
- Tauchmann, Harald (2014) "Lee's Treatment Effect Bounds for Non-random Sample Selection — an Implementation in Stata," *The Stata Journal*, Vol. 14, No. 4, pp. 884–894.
- Wooldridge, Jeffery (2002) "Inverse Probability Weighted M-estimators for Sample Selection," *Portuguese Economic Journal*, Vol. 1, No. 2, pp. 117–139.
- Wooldridge, Jeffery (2007) "Inverse Probability Weighted Estimation for General Missing Data Problems," *Journal of Econometrics*, Vol. 141, No. 2, pp. 1281–1301.
- Wooldridge, Jeffery (2010) *Econometric Analysis of Cross Section and Panel Data*, 2nd ed., MIT Press.
- Zamarro, Gema, Cheng, Albert, Shakeel, M. Danish and Hitt, Collin (2018) "Comparing and validating measures of non-cognitive traits: Performance task measures and self-reports from a nationally representative internet panel," *Journal of Behavioral and Experimental Economics*, Elsevier, Vol. 72(C), pp. 51–60.