

# 社会的・経済的要因と健康との因果性に対する諸考察 ——「社会保障実態調査」および「国民生活基礎調査」を用いた実証分析——

野 口 晴 子

## I はじめに

社会構造と社会の構成員たる個々人の「健康」<sup>1)</sup>との関連性については、産業革命による工業化・都市化の社会システムへの影響が顕在化する19世紀初頭以降2世紀以上にわたり、公衆衛生学を中心に、社会学・心理学・経済学など幅広い学問分野において膨大な数の理論的・実証的研究が蓄積されている。したがって、本庄(2007)が指摘するように、昨今とみに衆目を集めている、社会(心理)的・経済的格差と「健康」、そして、社会構造の主要な因子である個人または世帯における、ジェンダー、人種、家族関係、ソーシャル・ネットワーク、ソーシャル・キャピタル、教育、所得、資産、就労状況などの社会経済的状況(socioeconomic status: SES)と「健康」というイシューは、学問的にも政策的にも「古くて新しい」課題であるといえよう<sup>2)</sup>(Antonovsky(1967); Berkman and Kawachi(2000); 橋本(2005))。

近・現代を通じた研究の結果、SESと「健康」との間に有意に正の相関が存在することについては、おおむね研究者の間でコンセンサスが得られている。例えば、所得や資産、教育水準などが低くSESが不利な状況にある(すなわち、社会に格差が存在し、社会階層が低い)場合は、SESが有利な状況(高い社会階層)にある場合に比べて健康状態が悪く、慢性疾患になる確率や死亡率が高い傾向にあるといった関係性である。

しかし、そうした事実認知(fact-finding)を超えて、多岐にわたるSESがそれぞれどのようなメカニズムを通じて人々の「健康」と相互に関連し合うのか、また、その影響の有意性や大きさはどの程度のものなのかについては、実証的な結論が得られていない。近年、特に米国を中心としたパネルデータや開発経済学における社会的実験(social experiment)の蓄積、そして、計量経済学の理論とコンピュータによる計算能力の著しい発達を背景に、経済学を中心とした実証研究によって取り組まれてきたのが、SESと「健康」との複雑な因果関係(causality、もしくは、gradient)を識別し、紐解こうとする試みである(Cutler, Lleras-Muney, Vogl(2008))。こうした試みは、政策の費用対効果を検討する際に極めて重要である。なぜならば、国民の健康状態を改善するための政策として、所得移転、公教育の充実、公的な医療(健康)プログラムの整備など複数の選択肢がある場合、SESと「健康」との因果性の定量的検証は、どの施策に社会的資源や財源を投入するのが最も有効か、という問いに科学的エビデンスを提供するからである。

ここでは、SESと「健康」の関係を検証する場合に問題となる点を3つ指摘する。第1に、分析の基となるデータが、レセプトやカルテといった医療や介護の専門家など第三者の診断による客観情報ではなく、調査対象者本人(あるいは、要介護者や障がい者の場合は世帯構成員など)から自記式か面接法によるアンケートによって情報を得る自己申告型である場合、回答に何ら

かのバイアスがかかると、それが原因となって測定誤差 (measurement error) が生ずる可能性があるという点である。第2に、個人の「健康」をどう測るか、「何」をもって「健康」を示す尺度とするのか、という問題である。自己申告型の情報に基づく限りにおいて、個人の「健康」の正確な測定は不可能であり、どの健康尺度を用いようと測定誤差の問題を解消することは困難であることから (Bound, Brown, Mathiowetz (2001)), 健康尺度自体に対する丁寧な検証が必要となる。第3に、SESと「健康」との内生性 (endogeneity) の問題である (Adams, Hurd, McFadden, et al. (2003); Smith (2007) ほか)。個々人の健康状態はSESから影響を受けると同時に、健康状態がSESに影響を与えることもありうる。前段で述べた施策について、仮に、所得と健康の主たる因果関係が後者 (健康状態の改善 (悪化) が所得の向上 (下落) につながる) であった場合、いくら所得移転を積極的に行ったとしても国民の健康状態の改善にはつながらないだろう。教育や医療 (健康) プログラムの場合も同様である。

本稿の主たる目的は、上記3つの問題点に着目し、2007年に実施された、社会保障・人口問題基本調査「社会保障実態調査」(国立社会保障・人口問題研究所)と「国民生活基礎調査」(厚生労働省)の個票を用いて、SESと「健康」との因果性について実証分析を行う際、自己申告型のクロスセクションデータを用いることの可能性と限界について多角的・包括的に検証することにある。因果性の調整後、本稿の結果から得られる政策的含意は、国民 (成人) の健康状態の改善が主たる目的の施策としては、単に「ばらまき」型の所得移転やソーシャル・ネットワークの構築を目的とした地域政策は有効な施策とはいえず、むしろ、就労を促進し就労機会を増やすような雇用政策に対する社会的資源の投入や、教育の充実、持続可能な健康保険制度、健康行動の促進などの非経済的要因が重要であるという結果であった。

本稿の構成は以下の通りである。次節では、

データについての説明と、測定誤差および健康尺度について検討する。第3節では、SESが「健康」に与える効果を推定するための分析方法を、第4節では推定結果を概観する。最終節において、結果の考察と今後の課題について述べる。

## II データの特性、測定誤差と健康尺度

### 1 データ

本稿の分析に用いるデータは、2007年に実施された、社会保障・人口問題基本調査「社会保障実態調査 (以下、「実態調査」と略する) (国立社会保障・人口問題研究所) と「国民生活基礎調査 (以下、「基礎調査」と略する) (厚生労働省) の個票である。「実態調査」は、全国の世帯主および20~69歳の世帯員を対象としている。調査対象者は、同年実施の「基礎調査」において平成17年国勢調査区から層化無作為抽出した5,440調査地区内から無作為に選ばれた299調査地区の住民であるため、両者は突合可能である。「実態調査」における世帯票配布数 (調査客体世帯数) は 15,782 票で、うち回収数は10,766票 (有効回答率: 68.2%) であった。また、世帯主と20~69歳の世帯員を対象とした個人票の配布数は20,689票で、有効回答数は17,188票 (有効回答率: 83.1%) であった (国立社会保障・人口問題研究所 (2009))<sup>3)</sup>。

「基礎調査」と「実態調査」は、各自自治体における調査担当者や回答者への負担と回収率を考慮し、両者の突合に必要な基本属性 (性別・年齢・家族構成) 以外について原則重複がないように設計されている。したがって、「実態調査」では、「基礎調査」にはない、学歴、過去のさまざまな時点における両親との同別居や暮らし向き、過去と現在との生活水準・健康状態・仕事の状況・収入の比較、家族間での所得移転など、SESにかかわる質問項目が多岐にわたって調査されている。本稿では、そうした「実態調査」の特性を活かすため、「実態調査」の調査対象となった世帯員のみを分析対象とした。他方、調査対象者の現在の健康状態と健康にかかわる行

動（喫煙や健診の有無）については、「基礎調査（健康票）」で収集された詳細なデータを突合させて活用した<sup>4)</sup>。

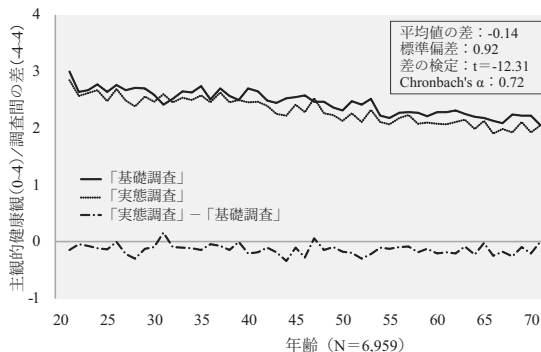
## 2 データにおける測定誤差の可能性

本稿で用いるデータは配票自計の自己申告型調査であることから、厳密に言えば、目的変数となる「健康」についてであれ、説明変数となるSESに関する指標であれ、すべての質問項目について、回答にバイアスがかかる可能性を回避することはできない（Bound, Brown, Mathiowetz (2001)）。ここでは、そうした可能性について、「実態調査」と「基礎調査（健康票）」において同じ質問が繰り返された、主観的健康観（self-rated health: SRH）について考察することにする。

SRHについては、追跡調査による将来の死亡率や特定疾病の罹患率などの客観的健康指標と強い相関を持っていることが数多くの研究によって報告されている（Adler, Boyce, Chesney et al (1994); Idler and Kasl (1995); McCallum, Shadbolt, Wang (1994); Gerdtham, Johannesson, Lundberg et al. (1999) ほか多数）。したがって、SRHが個人の健康状態を的確に反映している可能性がある一方で、調査対象者の性別、年齢、学歴、所得、職業など外的属性（Crossley and Kennedy (2002)）や楽観的か

悲観的かといった内的属性（Bound (1991); Bound et al. (1999)）、さらには、質問内容や提示された選択肢に対する捉え方（自分の「健康」を、自覚症状、障がい、診断結果など「何」に重点を置いて回答するか、健康状態が「ふつう」とはどういうことか、など）（Lokshin and Ravallion (2008)）によって評価が異なり、個人間で比較可能ではないという指摘もある。また、調査の実施方法（自記式か聞き取り方式か）（Tourangeau and Smith (1996); Grootendorst, Feeny, and Furlong (1997)）や、同じ質問の繰り返しによる調査目的に対する学習効果が、回答に影響を与えたとの報告もある（Tversky and Kahneman (1998)）。

SRHについて、「実態調査」と「基礎調査（健康票）」ではそれぞれ、「あなたの現在の健康状態についておたずねします」、「あなたの現在の健康状態はいかがですか」というほぼ同じ質問を行い、両者とも「よい」、「まあよい」、「ふつう」、「あまりよくない」、「よくない」という5段階の選択肢が同じ順番で提示されている。両者の実施日は、「実態調査」が2007年7月1日、「基礎調査」は「実態調査」より約3週間前の6月7日となっている。ある一定期間をあけて同じ質問を繰り返すこうした調査方法は、再テスト法（test-retest study）と呼ばれ、指標の妥当性や回答の信頼性を測定する際に頻繁に用いられる



出所) 平成19年『社会保障実態調査』（国立社会保障・人口問題研究所）および『国民生活基礎調査』（厚生労働省）を基にして筆者推計。

図1-1 年齢1歳階級別の主観的健康観（男性）

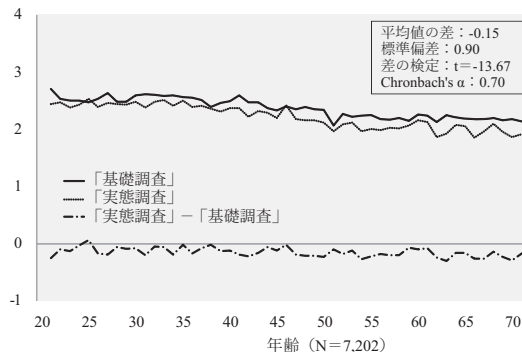


図1-2 年齢1歳階級別の主観的健康観（女性）

方法である<sup>5)</sup>。

図1-1と図1-2はそれぞれ、上記5段階の選択肢を、「よい」を4点、「まあよい」を3点、「ふつう」を2点、「あまりよくない」を1点、「よくない」を0点として点数化し、「基礎調査」と「実態調査」のそれぞれについて、男女別に年齢1歳階級別の点数と、2つの調査の差（「実態調査」－「基礎調査」）の平均値を示したグラフである。2つの図をみると、男女ともに年齢に関係なく「基礎調査」で回答したSRHが、わずか3週間後の「実態調査」において有意に悪化していることがわかる（男性：-0.14 (std dev : 0.92)；女性：-0.15 (std dev : 0.90)）。さらに、調査対象者の回答の一貫性を検証するため、クロンバックの $\alpha$ 係数 (Cronbach's alpha) を計算してみると、男性が0.72、女性が0.70と、信頼性の目安とされている0.8より若干低い結果であった。双方のデータには、死亡や特定疾病に対する診断などの客観的指標がないため、「現在の健康」に対するこうした回答の「ブレ」が、どの程度正確に個人の真の「健康」の変化を反映しているのかについては検証することができない。したがって、ここでは、こうした回答の変化に対する潜在的な原因（バイアス）についていくつか言及しておく。

まず、Lokshin and Ravallion (2008) が指摘するように、質問内容に対する調査対象者の捉え方の問題である。仮に、「現在の健康状態」における「現在」というタイム・フレームに対する人々の受け取り方が比較的短期間であるとすると、「実態調査」と「基礎調査」との回答の違いがあったとしても不思議はない。なぜならば、1日のうちですら、人は気分や体調などが変化することがあるからである。しかし、それだけでは、なぜSRHが全体的に下方に向かったのかを説明することはできない。紙幅の関係上、回帰分析結果の詳細については省くが、SRHの2つの調査の差（「実態調査」－「基礎調査」）を従属変数、「基礎調査」時点での個人属性を説明変数とした単純回帰分析の結果<sup>6)</sup>、ストレスがある場合の相談先の存在、高学歴、仕事あり、高い貯

蓄水準、医療保険（国民保険と被用者保険）、喫煙習慣なしが、SRHを改善する一方で、「基礎調査」のSRHと年齢の高さ、配偶者あり、世帯人員数、ストレスの存在が1%の有意水準で、3週間後のSRHの下方傾向に寄与していることがわかった。こうした結果は、SRHが個人の外的属性や健康行動に依存するとした、多くの先行研究と整合的である。第3に、Tversky and Kahneman (1998) が指摘するように、比較的短期間内でのSRHに関する質問を繰り返すことによって、「社会保障制度や厚生労働行政の企画立案のための基礎資料を作成する」という双方の調査に共通する目的への学習効果が表れた可能性も考えられる。社会保障政策に対する期待であれ、批判的な姿勢であれ、調査対象者のそうした認知や理解が、SRHを低めに申告させる結果につながったのかもしれない。あるいは、前段の回帰分析の結果から、「基礎調査」でSRHを高めに申告した人の方がより下方に修正される傾向があることから、単に、質問の繰り返しそのものが回答を禁欲的にさせたのかもしれない。いずれにしても、この場合は、調査対象者の「認知」や「記憶」が回答にバイアスをかけていることになる。

上記の結果から、本稿で用いるSRHにおいても測定誤算の存在することが示唆された。昨今、社会疫学を中心として、調査対象者に仮想的な質問をしたり (Salomon, Tandon, Murray et al.), 外生的な近隣環境のデータを収集したりすることによって (Pruitt, Jeffe, Yan et al. (2010)), SRHにおけるこうした測定誤差を調整し、比較可能な標準化尺度を作成しようとする試みが行われている。こうした試みについては、データの制約もあり、本稿での分析範囲を超えるため、今後の課題とするが、本稿では、SRHを単独で用いず、いくつかの指標を組み合わせることによって「健康」に対する総合指標を作成することとし、次節においてその検証を行う。

### 3 健康尺度－「健康」をどう測り、どう検証するか

個人の「健康」を分析の対象とする場合、常にディベートの対象となるのが個人の「健康」をどう測るかという尺度の問題である。本稿では、EuroQOL (EQ-5D) (Drummond, Sculpher, Torrance et al. (2005))<sup>7)</sup>と類似した項目を組み合わせて健康尺度を作成したNish, Hashimoto, Noguchi et al. (2011)に基づき、前節において考察を加えたSRHに、「基礎調査(健康票)」で質問された、以下の3つを合わせて尺度を作成する<sup>8)</sup>。①心理的健康観(6-item scale of nonspecific psychological distress: 以下、「K6」と略する)、②健康上の問題が日常生活に与える影響(instrumental activities of daily living: 以下、「IADL」と略する)、③自覚症状、の3指標である。

K6とは、精神疾患に対する効率的なスクリーニング法としてKessler, Andrews, Colpe et al. (2002)によって提案された尺度である(川上(2004))。「基礎調査(健康票)」では、過去1カ月の間に、「神経過敏に感じましたか」、「絶望的だと感じましたか」、「そわそわ、落ち着かなく感じましたか」、「気分が沈み込んで、何が起こっても気が晴れないように感じましたか」、「何をするのも骨折りだと感じましたか」、「自分は価値のない人間だと感じましたか」について、「いつも」、「たいてい」、「ときどき」、「少しだけ」、「まったくない」という5つの選択肢を与えて回答を求めている。ここではまず、各質問項目について、この5段階の選択肢を、「まったくない」を4点、「少しだけ」を3点、「ときどき」を2点、「たいてい」を1点、「いつも」を0点として点数化し6項目の合計点を算出する(最小値が0、最大値が24)。さらに、各調査対象者についてこの合計点を項目数の6で除した値(最小値0、最大値4)を本稿では心理的健康指標として用いる。

次に、IADLについては、「あなたは現在、健康上の問題で日常生活に何か影響がありますか」に対し「ある」と「ない」のいずれかを選択させた上で、「ある」と回答した者に対してどのよ

うなことに影響があるかを質問している。影響がある事柄は5つの選択肢からの複数選択で、「日常生活動作」、「外出」、「仕事、家事、学業」、「運動」、「その他」であるが、「その他」以外について影響のある場合を0点、ない場合を1点として合計点を求め点数化した(したがって、最小値は全項目に対して影響がある場合の0、最大値は、影響がない、もしくは、影響がある場合で「その他」のみを選択した場合の4)。

自覚症状については、「ここ数日、病気やけがなどで体の具合の悪いところがありますか」に対し「ある」と「ない」を選択させ、「ある」と回答した者に対して、42項目にわたる詳細な症状をきいている<sup>9)</sup>。調査対象者を20～40歳未満、40～65歳未満、65歳以上の3群に分けて、これら42項目について主成分分析を行い、第1主成分から第3主成分を求めた。年齢ごとの分析を行ったのは、年齢によって自覚症状の分布が大きく異なるからである。さらに、pattern matrixを求め、第1～第3の主成分について最も正の相関が高かった自覚症状を上から順番に4つ抽出し、各調査対象者に対し、当該自覚症状が「ある」場合をそれぞれ0点、「ない」場合を1点として合計を求め、3つの主成分ごとに点数化した(最小値は、4つすべての自覚症状がある場合の0、最大値は全く自覚症状がない場合の4)<sup>10)</sup>。

以上、SRH (0-4)、K6 (0-4)、IADL (0-4)、自覚症状(第1～第3主成分ごとに0-4、したがって、総計では0-12)の指標を1つにまとめ、健康状態が最も良い場合が24点、最も悪い場合が0点となる総合健康指標を作成した。次に、主観的な健康指標を組み合わせて作成したこうした総合健康指標が、どの程度個人の客観的な健康を正確に予測しているのかについて、Receiver operating characteristic (ROC) analysisの手法を用いて検証する。ROC分析は、患者が本当にある特定の疾患にかかっているかどうかについての判断が、診断(健康)尺度によってどの程度正確に行われたかを図式的に評価することができる統計手法で、公衆衛生学の分野において広く活用されている(Hsieh and Turnbull

(1996), Greinera M, Pfeifferb D, and Smithe RD (2000) ほか多数)。

ROC分析については、Zou KH and Zhou XHA (2001) に非常に簡潔にまとめられているので、その一部をここでは紹介することにする。仮に、ある指標(T)が連続変数だとして、 $T \leq t$  の場合は患者がある病気にかかっておらず、 $T > t$  の場合は病気にかかっていると診断されたとする。後日、こうした診断結果とは独立して、個々の患者が本当に病気にかかっていたか（実際には非罹患がm人、罹患がn人、したがって、患者数は $N=m+n$ ）どうかかわれば、図2-1のような $2 \times 2$ のマトリックスを描くことができる。マトリックスには、縦に診断結果、横に真の罹患状態が配置されている。自明のことであるが、診断結果も真の結果も陰性ならば「陰性(真)」, 診断結果は陰性であったが、真の結果は陽性である場合は「陰性(偽)」, 診断結果は陽性であったが、真の結果は陰性である場合は「陽性(偽)」, 診断結果も真の結果も陽性ならば「陽性(真)」となる。

ROC分析では、まず、この指標の正確性を評価するため、実際は罹患していなかった患者の診断による陽性率（「陽性(偽)」）を特異度(specificity), 実際は罹患していた患者の診断による陰性率（「陰性(偽)」）を感度(sensitivity)と定義する。したがって、前者は $P(T \leq t | D=0)$ , 後者は、 $P(T > t | D=1)$ で示

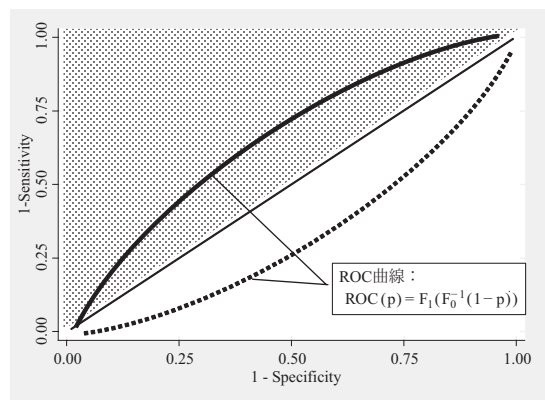
診断	ゴールドスタンダード(真の結果)		総数
	陰性 (D=0)	陽性 (D=1)	
陰性 Negative: T(+)	陰性(真)	陰性(偽)	# $\{T > t\}$ 診断結果 "+"
陽性 Positive: T(-)	陽性(偽)	陽性(真)	# $\{T \leq t\}$ 診断結果 "-"
総数	m	n	$N=m+n$

出所) Zou KH and Zhou XHA (2001) "Receiver operating characteristics (ROC) analysis". *Statistics in Epidemiology Report 1*: pp.1&4-12, を筆者が翻訳し、若干修正を加えた。

図2-1 Receiver operating characteristic (ROC) analysisの概念図

すことができるが、これらの確率はこの指標の診断の目安となっているtの値に依存する。そこで、想定されるすべてのt値について、横軸に $1 - P(T \leq t | D=0)$  (1-specificity, つまりは、「陰性(真)」の確率)を、縦軸に $1 - P(T > t | D=1)$  (1-sensitivity, つまりは、「陽性(真)」の確率)をとって、ROC曲線を描く。ROC曲線は、真の罹患状況が $D=d$ である患者の診断尺度(T)の分布が $F_d$ で示されるとき(つまり、 $D=0$ の場合は $F_0$ ,  $D=1$ の場合は $F_1$ ),  $ROC(p) = F_1(F_0^{-1}(1-p))$ で示すことができる。ここで、pは想定されるtの値に依存する「陽性(偽)」率、すなわち、 $P(T \leq t | D=0)$ である。図2-2は、このROC曲線の概念図を示している。(0,0)と(1,1)を結ぶ対角線よりも左上(斜部分)では、特異度と感度がともに高く、したがって、ROC曲線が実線で示された方の診断尺度は相対的に高い正確性を有していることになる。他方、点線の方は、ROC曲線が対角線の下方に位置しており、診断尺度の正確性が相対的に低いことを示している。これを数値化すると、ROC曲線と横軸・縦軸で囲まれた(RC曲線からみて下方)部分の面積(RC Area)を計算し、ROC Area面積が0.5を超えていれば、一定程度の正確性が確保されていることになる。

では、この手法を用いて、本稿で用いる健康



出所) 筆者作成。

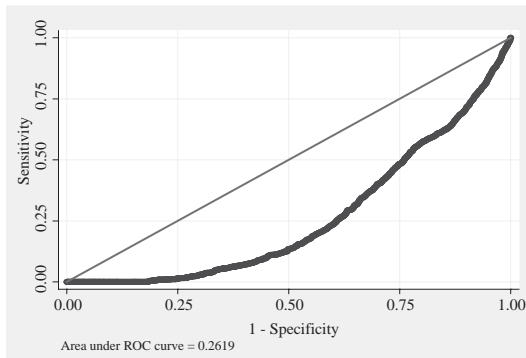
図2-2 Receiver operating characteristic (ROC) curveの概念図

尺度の正確性を、自己申告された代表的な特定疾患（高血圧症、高脂血症、糖尿病、精神疾患・こころの病・三大疾病）の罹患状況について検証してみることにする。まず、図3は、性・年齢によって標準化された総合健康指標およびK6と上記6つの特定疾患に対する罹患状況に対するROC曲線を示している<sup>11)</sup>。総合健康指標では、ROC曲線が、精神疾患・こころの病でほぼ対角線上に重なっている以外、残りの5疾病については対角線よりも右下に位置しており、正確性が担保されているとはいいがたい。他方、K6では、ROC曲線が、精神疾患・こころの病でほぼ対角線上に重なっている以外は、対角線よりも左上に位置していることから、これらの疾病の予測正確性がおおむね高いと考えられる。以上の分析から、K6がうつ指標としてばかりではなく、

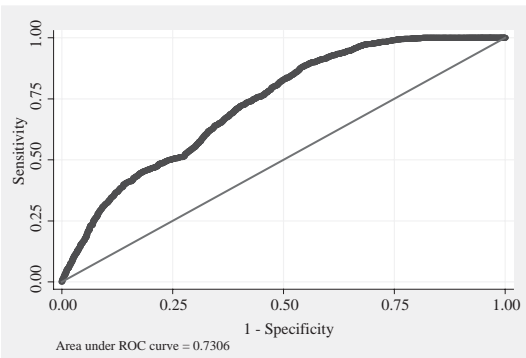
慢性疾患や三大疾病などについても優れた健康尺度であることがわかるが、K6と比べ総合健康指標の予測能力が低かった原因として、精神疾患・こころの病以外が高齢者に特有の疾患であった可能性がある。したがって、全年齢と20～40歳未満、40～65歳未満、65歳以上の年齢グループ別にROC Areaを推計した結果が表1である。

表1をみると、全般的に、総合健康指標は、SRHやIADLと比べると若干正確性が高いが、K6ほどは高くないことがわかる。年齢グループ別に分けると、精神疾患以外で、年齢が65歳未満よりも、年齢が65歳以上のグループにおける総合健康指標の正確性が高まっていることがわかる。SRHとIADLについても同様であるが、逆にK6の正確性が65歳以上で落ちている。また、20～40歳未満の三大疾病については、総合健康

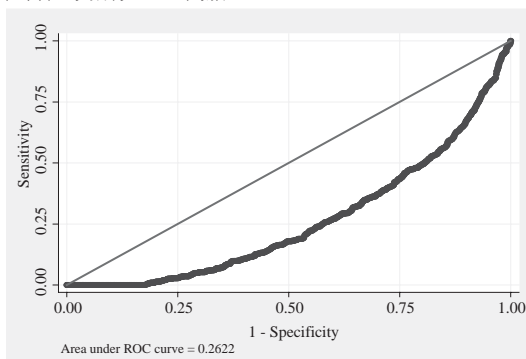
総合健康指標versus高血圧症



心理的健康指標(K6)versus高血圧症



総合健康指標versus高脂血症



心理的健康指標(K6)versus高脂血症

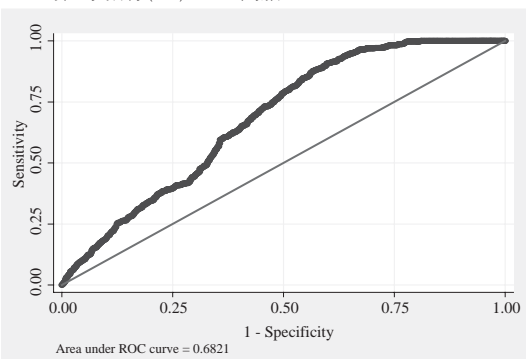
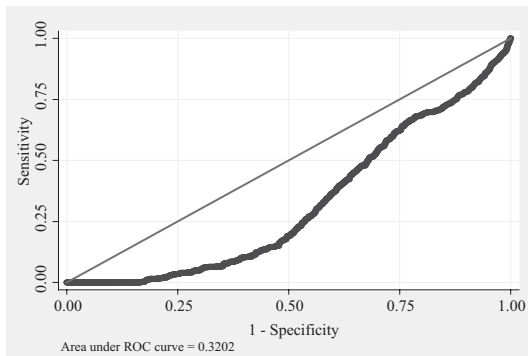
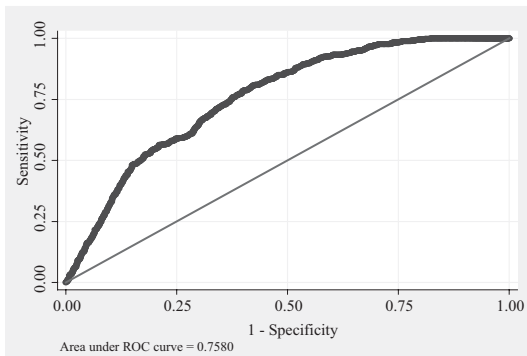


図3 性・年齢によって標準化された総合健康指標及び心理的健康指標（K6）と代表的な5特定疾患の有訴率についてのNonparametric receiver operating characteristics (ROC) analysis

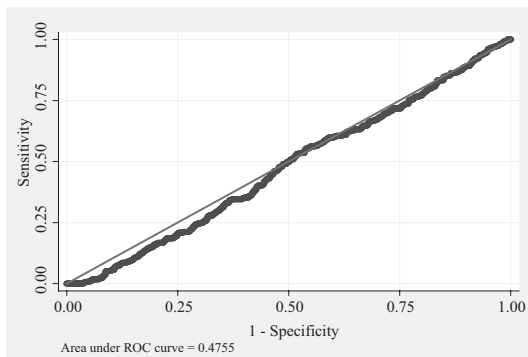
総合健康指標versus糖尿病



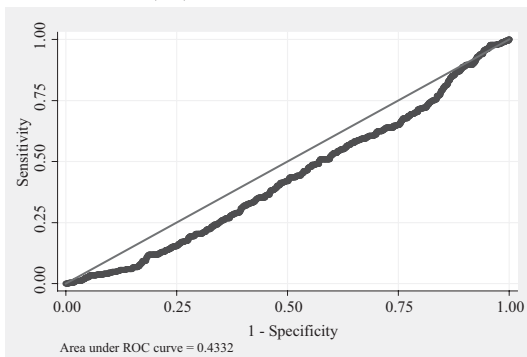
心理的健康指標(K6)versus糖尿病



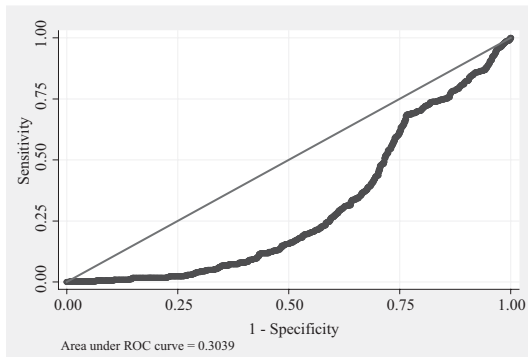
総合健康指標versus精神疾患・こころの病



心理的健康指標(K6)versus精神疾患・こころの病



総合健康指標versus三大疾病



心理的健康指標(K6)versus三大疾病

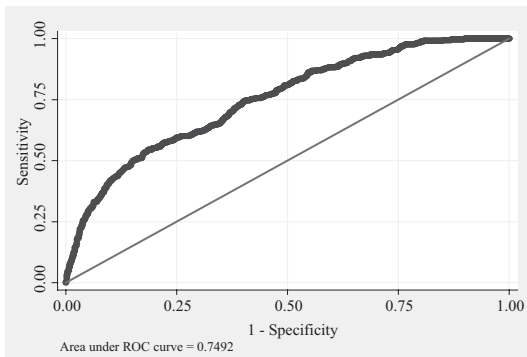


図3 続き

指標はほぼ0.5となっており、まずまずの正確性であった。本稿では、総合健康指標と、代表的な特定疾患に対する予測能力の比較的高かったK6を「健康」尺度とし、SESとの因果関係の検証を行う。上記の結果から明らかなように、総合健康指標については年齢グループにより正確

性が異なることから、本来であれば年齢グループ別の分析をすべきところであるが、正確性の比較的高かった高齢者層に限定するとサンプル数が著しく減ってしまい、robustな結果が得られなかったため、本稿では全年齢を対象とした結果のみ考察を行う。



表1 性・年齢によって標準化された健康指標と代表的な5特定疾患の有訴率についての  
Nonparametric receiver operating characteristic (ROC) analysis<sup>1)</sup>

	総合健康指標 (N=12,702)	ROC Area (Std. Err)		
		主観的健康観 (SRH) (N=14,161)	心理的 健康指標 (K6) (N=13,188)	生活障害 (IADL) (N=14,117)
<b>個別疾患</b>				
<u>全年齢</u>				
高血圧症	0.266 (0.006)	0.202 (0.005)	0.731 (0.007)	0.203 (0.005)
高脂血症	0.266 (0.010)	0.219 (0.008)	0.682 (0.009)	0.219 (0.008)
糖尿病	0.325 (0.010)	0.243 (0.008)	0.758 (0.010)	0.245 (0.008)
精神疾患・こころの病	0.474 (0.017)	0.506 (0.017)	0.433 (0.017)	0.505 (0.017)
三大疾病	0.308 (0.011)	0.232 (0.010)	0.749 (0.013)	0.234 (0.010)
<u>20～40歳未満<sup>2)</sup></u>				
精神疾患・こころの病	0.353 (0.028)	0.353 (0.028)	0.468 (0.029)	0.353 (0.028)
三大疾病	0.499 (0.105)	0.489 (0.105)	0.679 (0.087)	0.491 (0.105)
<u>40～65歳未満</u>				
高血圧症	0.381 (0.010)	0.315 (0.009)	0.638 (0.010)	0.308 (0.009)
高脂血症	0.378 (0.016)	0.330 (0.014)	0.596 (0.014)	0.326 (0.014)
糖尿病	0.458 (0.015)	0.371 (0.014)	0.692 (0.015)	0.359 (0.014)
精神疾患・こころの病	0.487 (0.025)	0.519 (0.025)	0.429 (0.022)	0.524 (0.025)
三大疾病	0.413 (0.019)	0.356 (0.018)	0.622 (0.022)	0.350 (0.018)
<u>65歳以上</u>				
高血圧症	0.491 (0.015)	0.472 (0.015)	0.518 (0.015)	0.472 (0.015)
高脂血症	0.405 (0.021)	0.451 (0.021)	0.462 (0.022)	0.461 (0.022)
糖尿病	0.515 (0.024)	0.464 (0.023)	0.548 (0.022)	0.460 (0.023)
精神疾患・こころの病	0.440 (0.050)	0.558 (0.051)	0.388 (0.054)	0.568 (0.052)
三大疾病	0.580 (0.021)	0.422 (0.023)	0.644 (0.024)	0.409 (0.023)

注) 1) 総合健康指標 (0-28), 主観的健康指標 (0-4), 心理的健康指標 (0-4), IADL (0-4) については, 性と年齢によって標準化された値を用いて推計した。自覚症状 (0-12) については結果を省略する。

2) 分析対象となったサンプルについて, 20～40歳未満において高血圧症, 高脂血症, 糖尿病に関する有訴は無かった。

出所) 平成19年『国民生活基礎調査』(厚生労働省)を基にして筆者推計。

### III 分析の方法

#### 1 IV法

SESが健康に与える効果を定量的に検証する場合に常に問題となるのが、両者の複雑な内生性である<sup>12)</sup>。冒頭で述べたように、昨今、とりわけ欧米において詳細な健康尺度を含むパネルデータや社会実験・自然実験に基づく情報が収集され、クロスセクションデータに基づく静学的モデルでは解消することの難しかったSESと健康の因果性を、推計モデルの動学化や無作為抽出化によって識別しようとする研究の蓄積が進んでいる。そうした先行研究については、Cutler, Lleras-Muney, Vogl (2008) が包括的なレビューを行っている。しかしながら、本稿で用いる「実態調査」と「基礎調査」はいずれも定点情報であるため、パネルや実験における分析手法は使うことができない。

したがって、ここでは、操作変数 (instrumental variable: 以下、「IV」と略する) 法による調整を行う。まず、総合健康指標とK6のそれぞれについて下記のような推定式を想定する。

$$h_i = \alpha_0 + \alpha_1 Z_i + \alpha_2 x_i + u_i \quad (1)$$

推定式(1)で、 $h_i$  は  $i$  番目の調査対象者の総合健康指標、あるいは、K6を示す。 $Z_i$  は、操作変数によって調整するSESで、現在の就労状況、所得、貯蓄、ソーシャル・ネットワーク (ストレスがある場合の相談先) を示す。 $x_i$  は、性別、年齢、配偶状態、同居家族人数、そのほかのSES (学歴<sup>13)</sup>、持ち家、医療保険、居住地域属性 (大都市ダミー)、ストレス、および、健康行動 (喫煙、健診) を表す。 $\alpha_0$  は定数項、 $\alpha_1$ 、 $\alpha_2$  は推定パラメータ、 $u_i$  は誤差項を示している。(1)において問題となるのは、SESの諸変数 ( $Z_i$ ) に関する内生性の問題である。例えば、健康状態が良好な者ほど、労働の生産性が高かったり、活動的な日常生活を送ったりして、安定した職

に就き、所得や貯蓄などが高く、幅広いソーシャル・ネットワークに恵まれている可能性がある。本稿のデータでは、労働生産性や日常生活における活動記録などは、 $x_i$  では調整することができず、 $u_i$  として吸収されてしまう。したがって、SESの健康に対する効果を、もともとの健康状態の良さが起因している要因から識別する必要がある。仮に、推定式(1)を最小二乗法で推計すると、 $Z_i$  と  $u_i$  は相関をもつことから、 $Z_i$  の推定パラメータ  $\alpha_2$  にバイアスがかかることになる。

ここでは、こうした内生性を有する「観察」データから実験的な変動を抽出し、サンプルを疑似的にランダム化する計量手法として、IV法を用いる。IV ( $V_i$ ) の選定に当たっては、次の2つの仮定が鍵となる。1つは、 $E(u_i|V_i) = 0$  であり、これは、 $V_i$  は外生であり  $h_i$  とは直接的な相関が無いことを意味している。2つめは、 $E(Z_i|V_i) \neq E(Z_i)$  である。これは、 $V_i$  が  $h_i$  とは無相関であるが、SES ( $Z_i$ ) とは相関があると仮定している。この仮定は  $E(x_i|V_i=1) = E(x_i|V_i=0)$  と、セレクション・バイアスを疑似的に調整する条件として書きかえることもできる。本稿では、操作変数として、最後に通った学校の卒業年とその前年における日経平均株価、GDP、および、失業率のデータを用いる。つまり、こうした過去における外的なマクロの経済状況は、 $Z_i$  とは相関があるが、 $h_i$  とは直接的には無相関であると想定する。回帰分析は、第1段階で、 $Z_i^* = \beta_0 + \beta_1 V_i + \beta_2 x_i + \epsilon_i$  を推定し、第2段階で推定された  $Z_i^*$  を説明変数として(1)に投入する二段階推定法を行う<sup>14)</sup>。

#### 2 GMM推定法

しかしながら、IV法の最大の問題点の1つであるが、これら4つのSESは相互に強い相関があると考えられ、これらを同時に調整すると多重共線性の問題が起こる可能性がある。したがって、現在の就労状況、所得、貯蓄、ソーシャル・ネットワークのそれぞれについて、別個に回帰分析を行わなければならない、内生変数が複数存在する場合の因果性の調整モデルとしては不完

全であるといわざるをえない。ここでは、こうしたIV法の問題点を補完するため、地域（調査単位地区）内（within-geographic area）における調査対象者の属性のバリエーションではなく、地域間（between-geographic areas）におけるバリエーションをIVとする一般化モーメント推定法（General Method of Moments：以下、「GMM」と略する）による2段階推定を行った（Staiger and McClellan（2000））。

$$h_{ij} = \gamma_1 Z_{ij} + \gamma_2 x_{ij} + u_j + \varepsilon_{ij} \quad (2)$$

推定式(2)は、第j番目の地域における第i番目の調査対象者の総合健康指標（あるいは、K6）が、各地域における個々のSESとそれ以外の個人属性に依存していることを示している。 $\gamma_1$ と $\gamma_2$ は推定されるパラメータである。Staiger and McClellan（2000）では、誤差項を地域間（ $u_j$ ）と調査対象者間（ $\varepsilon_{ij}$ ）におけるSESとそれ以外の属性のバリエーションに起因する2つの部分に区別して定義している。ここでの定義は、誤差項が地域間（between-geographic areas:  $u_j$ ）と調査対象者間（within-geographic area:  $\varepsilon_{ij}$ ）における属性のバリエーションに起因していることとらえる。ここでは、 $u_j$ と $Z_{ij}$ 、および、 $\varepsilon_{ij}$ と $x_{ij}$ とは相関がないが、 $\varepsilon_{ij}$ と $Z_{ij}$ は相関があると仮定する。

次に、このモデルでは、第j番目の地域に居住する第i番目の調査対象者のSES（ $Z_{ij}$ ）は、個々の属性と地域属性によって説明される内生変数であるから、

$$Z_{ij} = \delta x_{ij} + \tau_j + v_{ij} \quad (3)$$

となり、推定式(3)において、 $\tau_j$ は第j番目の地域のSESのそれぞれに関する密集度を示す固定効果（geographic area fixed effect）である。つまり、これは、一種の近隣効果（neighborhood effect）を表していると考えられる。(3)

においても、調査対象者の個人属性  $x_{ij}$  は、誤差項  $v_{ij}$  とは相関がないと仮定する。推定式(2)に(3)を代入することによって、推定式(4)が求められる

$$\begin{aligned} h_{ij} &= \rho x_{ij} + \phi_j + \omega_{ij} \\ \text{where } \rho &= \delta \gamma_1 + \gamma_2; \\ \phi_j &= \gamma_1 \tau_j + u_j; \\ \omega_{ij} &= \gamma_1 v_{ij} + \varepsilon_{ij} \end{aligned} \quad (4)$$

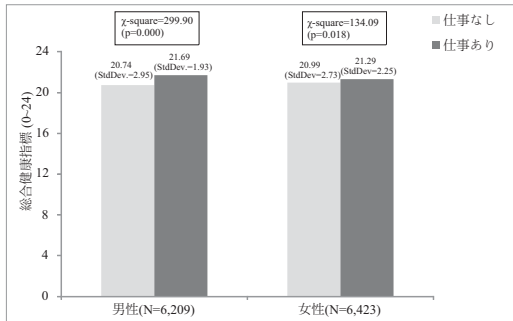
具体的な推定の手順は、まず、個票データに基づき推定された回帰式(3)と(4)から推計される誤差項の平均値を各地域について求め、地域レベルのデータを作成する。つまり、(4)における  $\phi_j = \gamma_1 \tau_j + u_j$  を求める。この式において、地域間におけるSESの固定効果（ $\tau_j$ ）は、観察できない地域間での属性のバリエーション（ $u_j$ ）からは独立であると仮定する。ここで、 $\phi_j$ と $\tau_j$ が推定誤差なしに観察可能であるならば、SESの固定効果をIVとして投入し、第j番目の地域における調査対象者の健康に対する回帰式(4)を、各地域の住民数（ $m$ ）による加重回帰分析を用いて推定することができる。つまり  $\gamma_1 = 1/m(\hat{\tau}'\hat{\tau})^{-1}(\hat{\tau}'\hat{\phi})$  である。ここで、 $\hat{\tau}_j$ には測定誤差が存在し、 $1/m(\hat{\tau}'\hat{\tau})$ は、地域間におけるSESの分散、 $E(\tau'\tau) \equiv \text{Var}(\tau_j)$ 、に対して過剰推計であると考えられる。したがって、Staiger and McClellan（2000）にならい、この推定誤差を修正するため、第1段階で求めた固定効果の推定ベクターから、地域間のSESについて、 $\hat{\tau}_j$ における推定誤差の不偏分散推定値を示す共分散行列（ $\bar{\Omega}$ ）を求め、これを過剰推計された  $1/m(\hat{\tau}'\hat{\tau})$  から差し引く（ $1/m(\hat{\tau}'\hat{\tau}) - (\bar{\Omega})$ ）によって、修正を行った。

#### IV SESの健康（総合健康指標・K6）に対する効果

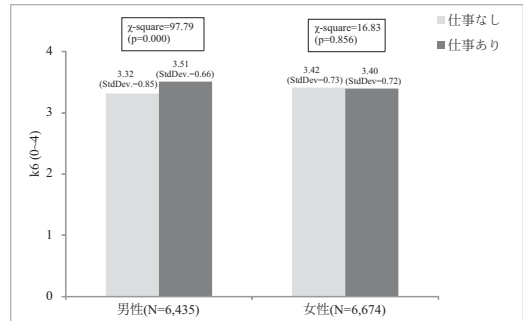
##### 1 基本統計量

本節では、分析で用いるSESと健康尺度の基本統計量について簡単に概観する。図4は、本稿

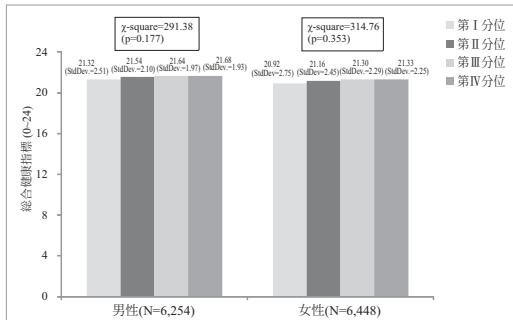
現在の就労状況versus総合健康指標



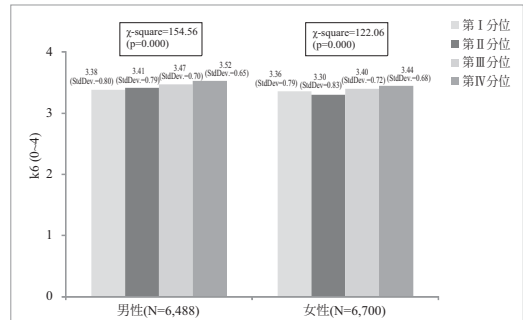
現在の就労状況versus K6



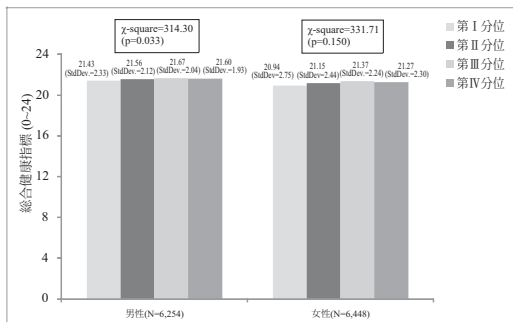
等価所得4分位versus総合健康指標



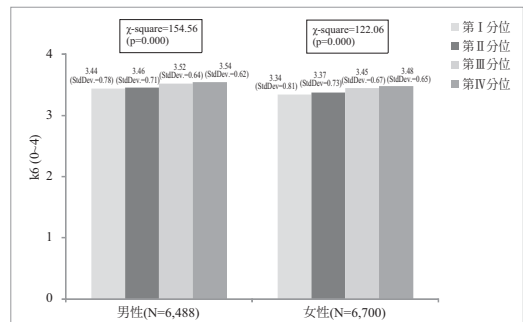
等価所得4分位versus K6



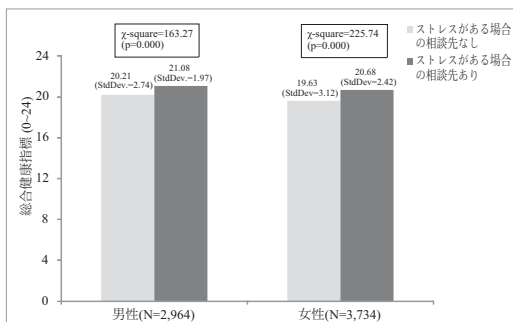
貯蓄階級4分位versus総合健康指標



貯蓄階級4分位versus K6



ソーシャル・ネットワークversus総合健康指標



ソーシャル・ネットワークversus K6

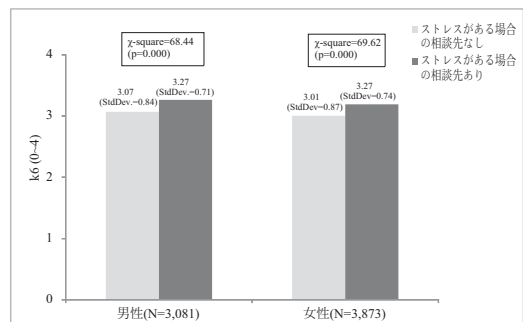
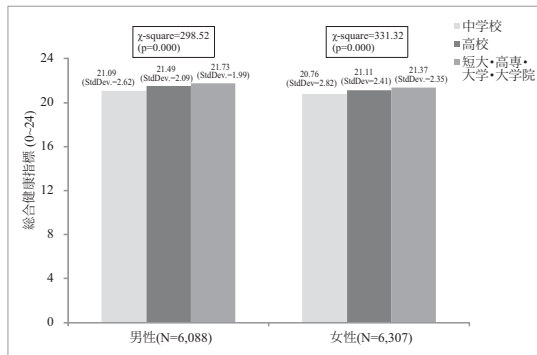


図4 性・SES別, 総合健康指標および心理的健康指標 (K6)

現在の就労状況versus総合健康指標



現在の就労状況versus K6

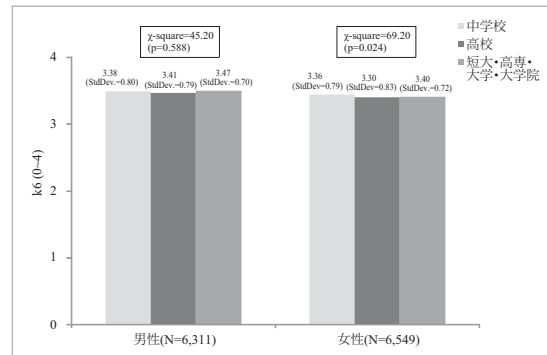


図5 性・最後に通った学校別、総合健康指標および心理的健康指標 (K6)

で分析対象とするSESごとに総合健康指標とK6の違いを男女別に表したものである。

現在の就労状況(仕事の有無)では、女性のK6以外ですべて有意な差が観察された。総合健康指標では、仕事ありと仕事なしの平均が、男性で20.74と21.69、女性で20.99と21.29となっており、 $\chi^2$ 検定の結果、男性が1%水準、女性が5%水準で仕事ありが仕事なしに比較して健康状態が良好である。また、男性だけについてであるが、仕事ありの方のK6が有意に高い。次に、等価所得の4分位別の結果をみると、男女ともに総合健康指標では所得による平均値の差は有意ではなかった。他方、K6については、男女でそれぞれ平均値が、3.38と3.36(第I分位)、3.41と3.30(第II分位)、3.47と3.40(第III分位)、3.52と3.44(第IV分位)と、女性の第II分位と第III分位で逆転がみられるものの、おおむね所得階層が高くなるほど心理的健康状態の改善傾向が認められる。さらに、貯蓄の4分位では、女性の総合健康指標以外ですべて有意な差が観察された。とりわけ、K6については、男女の平均値が、3.44と3.34(第I分位)、3.46と3.37(第II分位)、3.52と3.45(第III分位)、3.54と3.48(第IV分位)と、貯蓄が多い階層ほどK6が有意に高い傾向にあることがみてとれる。ソーシャル・ネットワークについては、ストレスのあるサンプルのみを対象として、家族・友人・職場に相

談相手がいる場合を1、いない場合を0として分析を行った。結果、総合健康指標とK6について、男女ともに、相談先のある方がない場合に比べ、有意に健康状態が良好である。

最後に、本稿では所与として扱ったが、Grossman(2006)にあるように、数多くの先行研究が対象としている学歴とSESとの関係のみをみておく。図5は、最後に通った学校別の総合健康指標とK6の違いを男女別に示している。K6について男性で有意でなかったほかは、先行研究で得られた結果同様、特に総合健康指標で、男女の平均値が、21.09と20.76(中卒)、21.49と21.11(高卒)、21.73と21.37(短大・高専・大卒以上)と、学歴が高いほど健康状態が有意に良好であった。そのほか、推計に用いたすべての変数についての基本統計量は表2に詳しい。

## 2 IV, および、GMMの推定結果

本節では、回帰分析の結果について概観する。表3-1と表3-2はそれぞれ、総合健康指標とK6に対するSESの効果を、OLS, IV, GMMにより推計した結果をまとめたものである。前節で述べたように、これらすべての回帰分析は、調査対象者の人口社会学的属性、そのほかのSES、ストレスや健康行動によって調整済みである。

まず、OLSによる単純回帰の結果を見てみると、総合健康指標、K6ともに、SESが有意にプ

表2 男女別，推定に用いる諸変数

	総数	男性	女性
	平均値 (Std. Dev.)	平均値 (Std. Dev.)	平均値 (Std. Dev.)
<u>目的変数</u>			
総合健康指標 (N=12,702)	21.37 (2.30)	21.56 (2.11)	21.18 (2.45)
K6 (N=13,188)	3.45 (0.71)	3.49 (0.69)	3.41 (0.72)
<u>本稿で調整するSES</u>			
現在の就労状態 (仕事あり)	0.73 (0.44)	0.86 (0.35)	0.61 (0.49)
所得 (千円)	4,013 (15,100)	4,224 (15,700)	3,809 (14,500)
貯蓄 (千円)	6,946 (21,000)	6,799 (20,000)	7,090 (21,800)
ソーシャル・ネットワーク (相談先あり)	0.29 (0.45)	0.22 (0.42)	0.36 (0.48)
<u>人口社会学的属性</u>			
性別 (女性)	0.51 (0.50)	0.00 (0.00)	1.00 (0.00)
年齢	48.31 (13.90)	48.31 (13.98)	48.31 (13.82)
配偶者あり	1.30 (0.46)	1.31 (0.46)	1.30 (0.46)
<u>同居家族人数</u>	2.28 (1.42)	2.24 (1.42)	2.32 (1.41)
<u>その他SES</u>			
中学校	0.13 (0.34)	0.13 (0.34)	0.13 (0.34)
高校	0.46 (0.50)	0.44 (0.50)	0.48 (0.50)
短大・高専	0.17 (0.38)	0.09 (0.29)	0.25 (0.43)
大学・大学院	0.23 (0.42)	0.33 (0.47)	0.13 (0.34)
持ち家	0.75 (0.44)	0.74 (0.44)	0.75 (0.43)
医療保険：被用者保険	0.35 (0.48)	0.34 (0.47)	0.35 (0.48)
医療保険：国民健康保険	0.63 (0.48)	0.64 (0.48)	0.62 (0.48)
医療保険：その他	0.02 (0.12)	0.02 (0.13)	0.02 (0.12)
大都市	0.23 (0.42)	0.22 (0.42)	0.23 (0.42)
<u>ストレスと健康</u>			
ストレスあり	0.50 (0.50)	0.45 (0.50)	0.55 (0.50)
喫煙なし	0.66 (0.47)	0.49 (0.50)	0.83 (0.38)
健診受診あり	1.60 (1.44)	1.54 (1.46)	1.66 (1.41)

出所) 平成19年『社会保障実態調査』(国立社会保障・人口問題研究所) および『国民生活基礎調査』(厚生労働省)を基にして筆者推計。

表3-1 総合健康指標に対する諸属性の効果<sup>(注)</sup>

	OLS			IV			GMM 係数 (Robust Std. Err.) 地区数=294
	係数 (Robust Std. Err.) N=12,249	係数 (Robust Std. Err.) N=12,311	係数 (Robust Std. Err.) N=12,311	係数 (Std. Err.) N=11,855	係数 (Std. Err.) N=11,909	係数 (Std. Err.) N=11,909	
被説明変数 総合的健康指標 (0-24点)							
現在の就労状況 (仕事あり)	0.47 (0.05)	---	---	0.66 (0.30)	---	---	0.93 (0.40)
等価所得 (自然対数)	---	0.01 (0.01)	---	---	-0.17 (0.11)	---	0.11 (0.11)
貯蓄 (自然対数)	---	---	0.01 (0.00)	---	---	-0.09 (0.05)	0.03 (0.07)
ソーシャル・ネットワーク	---	---	0.76 (0.07)	---	---	-3.66 (1.82)	-0.84 (0.55)
モデルのTest statistics							
Adj/Pseudo R <sup>2</sup>	0.21	0.21	0.21	0.22	187.03	189.57	0.07
Wald/LR chi <sup>2</sup>	198.96	199.45	198.48	201.28	206.09	209.96	5.59
操作変数の弱相関の検定							
Anderson's CC LM statistic							
p値				296.92	27.17	40.54	15.84
Cragg-Donald Wald F statistic				0.000	0.000	0.000	0.015
過剰識別制約検定				50.67	4.53	6.77	2.64
Sargan statistic				6.44	8.30	7.90	3.44
p値				0.266	0.141	0.162	0.632

表3-2 K6に対する諸属性の効果<sup>(注)</sup>

	OLS			IV			GMM 係数 (Robust Std. Err.) 地区数=294
	係数 (Robust Std. Err.) N=12,695	係数 (Robust Std. Err.) N=12,765	係数 (Robust Std. Err.) N=12,765	係数 (Std. Err.) N=12,280	係数 (Std. Err.) N=12,342	係数 (Std. Err.) N=12,342	
被説明変数 k6 (0-4点)							
現在の就労状況 (仕事あり)	0.08 (0.01)	---	---	-0.16 (0.09)	---	---	0.85 (0.37)
等価所得 (自然対数)	---	0.003 (0.002)	---	---	0.03 (0.03)	---	0.10 (0.10)
貯蓄 (自然対数)	---	---	0.004 (0.001)	---	---	0.02 (0.02)	0.02 (0.07)
ソーシャル・ネットワーク	---	---	0.18 (0.02)	---	---	0.50 (0.48)	-0.77 (0.50)
モデルのTest statistics							
Adj/Pseudo R <sup>2</sup>	0.21	0.21	0.21	0.22	214.27	209.96	0.07
Wald/LR chi <sup>2</sup>	231.72	232.68	232.63	235.16	213.11	214.19	5.59
操作変数の弱相関の検定							
Anderson's CC LM statistic							
p値				310.01	27.53	42.64	14.42
Cragg-Donald Wald F statistic				0.000	0.000	0.000	0.025
過剰識別制約検定				52.92	4.59	7.12	2.40
Sargan statistic				0.74	2.77	0.97	2.36
p値				0.981	0.755	0.965	0.798

注) 回帰分析は、表2で示した個人属性により調整済み。\*\*\*、\*\*、\* / はそれぞれ、1%、5%、10%水準で推定値が有意であることを示す。  
出所) 平成19年『社会保障実態調査』(国立社会保障・人口問題研究所) および『国民生活基礎調査』(厚生労働省) を基にして筆者推計。

ラスに影響している。総合健康指標とK6はそれぞれ、仕事がある場合はない場合に比べ0.47と0.08ポイント高く、自然対数値とした等価所得と貯蓄の1%上昇に対しては1%と0.3-0.4%、ソーシャル・ネットワークの存在については0.76と0.18ポイント、高まる傾向にあることがわかる。

しかし、前節で論じたように、こうした単純推計の結果は、内生性の問題を孕んでいる可能性が高い。したがって、SESごとにIV法を用いて内生性を調整した結果が次に示されている。

現在の就労状況では、仕事がある場合はない場合に比較して、総合健康指標は0.66ポイント高まるが、K6が0.16ポイント低下する傾向にあることがわかる。ただし、後者については、有意性はあまり高くない。他方、貯蓄のK6に対する効果が若干マイナスとなった以外では、等価所得と貯蓄など世帯が有する経済資源やソーシャル・ネットワークの存在の健康に対する効果は有意ではなかった。ここでは、操作変数として、最後に通った学校の卒業年とその前年における3

表4 IV法による学歴、健康保険、喫煙の健康指標に対する効果<sup>注)</sup>

被説明変数 総合的健康指標 (0-24点)	現在の就労状況	等価所得	貯蓄	ソーシャル・ネットワーク
	係数 (Std. Err.)	係数 (Std. Err.)	係数 (Std. Err.)	係数 (Std. Err.)
	N=11,855	N=11,909	N=11,909	N=11,909
高卒	0.184 ** (0.071)	0.268 *** (0.089)	0.354 *** (0.120)	0.145 (0.092)
短大・高専	0.274 *** (0.081)	0.385 *** (0.109)	0.526 *** (0.168)	0.265 *** (0.102)
大学・大学院	0.342 *** (0.079)	0.455 *** (0.109)	0.667 *** (0.207)	0.300 *** (0.101)
医療保険：被用者保険	0.548 *** (0.138)	0.664 *** (0.154)	0.809 *** (0.199)	0.825 *** (0.210)
医療保険：国民健康保険	0.546 *** (0.145)	0.876 *** (0.192)	0.996 *** (0.234)	0.929 *** (0.208)
喫煙なし	0.199 *** (0.045)	0.166 *** (0.045)	0.274 *** (0.078)	0.161 *** (0.053)
被説明変数 K6 (0-4点)	現在の就労状況	等価所得	貯蓄	ソーシャル・ネットワーク
	係数 (Std. Err.)	係数 (Std. Err.)	係数 (Std. Err.)	係数 (Std. Err.)
	N=12,280	N=12,342	N=12,342	N=12,342
高卒	0.002 (0.022)	-0.003 (0.025)	-0.035 (0.035)	0.016 (0.023)
短大・高専	0.027 (0.025)	0.016 (0.031)	-0.035 (0.049)	0.035 (0.025)
大学・大学院	0.048 ** (0.024)	0.038 (0.031)	-0.032 (0.061)	0.063 ** (0.025)
医療保険：被用者保険	0.135 *** (0.042)	0.116 ** (0.048)	0.070 (0.059)	0.098 * (0.055)
医療保険：国民健康保険	0.185 *** (0.044)	0.122 ** (0.061)	0.071 (0.070)	0.121 ** (0.054)
喫煙なし	0.026 * (0.014)	0.031 ** (0.013)	0.002 (0.023)	0.032 ** (0.013)

注) \*\*\*/ \*\*, \*/はそれぞれ、1%、5%、10%水準で推定値が有意であることを示す。

出所) 平成19年『社会保障実態調査』(国立社会保障・人口問題研究所) および『国民生活基礎調査』(厚生労働省)を基にして筆者推計。



つのマクロ経済指標を用いたが、IV法における操作変数の弱相関の検定でも、操作変数が調査対象者の総合健康指標およびK6と有意な相関を持つことを確認することができる。また、過剰識別制約の検定結果から、これらの操作変数が健康関数の推定式(1)における誤差項( $u_1$ )と無相関であるという帰無仮説は棄却されない。したがって、このモデルにおけるIVの有効性が上記の推定については検証されたことになる。ただし、Cragg-Donald Wald F statisticが、現在の就労状況以外の回帰式であまり高くはないことから、第一段階での予測値が第二段階でのIVとして適切であるかについては、議論のあるところである。

次に、GMM推定法の結果についてみる。この推定法では、個票レベルの情報を地域ごとに集約することから、分析の対象となるサンプル数は地区数(294)に等しい。分析の方法で述べたように、この方法の利点は、IV法では個別に分析したSESを同時推定可能なことである。この結果から、総合健康指標にもK6にも有意な効果が観察されたのは、現在の就労状況であった。仕事がある場合はない場合に比べ両者をそれぞれ0.93と0.85ポイント高める傾向にあることがわかる。すべての推定において有意な結果を得た現在の就労状況について、IVやGMMの推定値がOLSよりも大きくなっている。このことは、操作変数を用いて内生性に対処したことにより、測定誤差による下方バイアスの問題が緩和されたか、健康状態が良好であるために仕事をしているという本稿の想定とは逆の因果関係の影響が軽減されたことによるものかもしれない。

最後に、本稿では内生性の調整の対象とはしなかったが、学歴、医療保険、喫煙行動と健康との関連性についてみる<sup>15)</sup>。表4は、総合健康指標とK6それぞれを従属変数とした4つのIV推定式の結果を示している。この結果から、高学歴、医療保険の存在、非喫煙者であることがrobustに総合健康指標にプラスの効果をもたらしていることがわかる。他方、K6については、

総合健康指標ほどには、効果の有意性は高くないものの、やはり似た傾向にはある。総合健康指標については、現在の就労状況の推定式における推定値の大きさがおおむねOLSの推定値と同程度であることから、所得や貯蓄などの経済資源を調整した推定式では学歴と医療保険の推定値が、また、ソーシャル・ネットワークの推定式では医療保険の推定値が、過剰に大きく推定されていることがわかる。したがって、世帯の経済資源やソーシャル・ネットワークと学歴や医療保険との間には強い正の相関があると考えられる。他方、喫煙行動については、K6の貯蓄を調整した推定式以外で、推定値の大きさも比較的安定しており、robustな結果であった。

## V 考察と今後の課題

本稿では、測定誤差、健康尺度、内生性という3つの問題点に着目し、SESと「健康」との因果性について実証分析を行う際、自己申告型のクロスセクションデータを用いることの可能性と限界について検証を行った。本稿が得た実証結果から、第1に、「実態調査」と「基礎調査(健康票)」のSRHには測定誤差の可能性があり何らかの調整が必要であること、また、健康尺度について、いくつかの尺度の組み合わせにより指標の正確性が若干改善されること、また、K6が調査対象者の「健康」を測定する上で非常に有効な指標であることが示された。第2に、単純回帰では、現在の就労状況、等価所得、貯蓄、ソーシャル・ネットワークが有意にSESを改善するという結果が得られたが、SESと健康の因果性を調整した結果、唯一robustな結果が得られたのは現在の就労状況のみであった。最後に、本稿では所与として処理した学歴、健康保険、喫煙行動については、いずれも総合健康指標に対して有意な効果が観察されたが、学歴と健康保険については世帯の経済資源との相関性が疑われ、推定結果にバイアスをかけている可能性は否めない。

本稿が得た結果から得られる政策的含意は、

国民の健康状態の改善が主たる目的の施策としては、所得移転やソーシャル・ネットワークの構築を目的とした地域政策は有効な施策とはいええず、むしろ、就労を促進し就労の機会を増やすような雇用政策に対する社会的資源の投入が極めて重要であるということである。なぜならば、世帯の経済資源やソーシャル・ネットワークについては、健康状態の改善による逆方向の因果を調整した場合、SESの健康に対するインパクトは相当程度相殺されるためである。また、教育の充実、持続可能な健康保険制度、健康行動の促進などの非経済的要因が重要であるという結果は、パネルデータを用いたSmith (2007)とも整合的ではあるが、こうした変数についても内生性が強く疑われ、クロスセクションデータを用いた本稿での調整は不十分といわざるをえない。

本稿では、定点観察データという制約の中で、できる限りSESと健康との因果性を調整することを試みたが、さまざまな課題と可能性が残されている。例えば、Pruitt, Jeffe, Yan et al. (2010) にならい、外生的な近隣情報を活用することで、SESと健康の両者に存在する測定誤差の調整を行うといった可能性が考えられる。また、本文中では触れなかったが、健康に関する質問項目における欠損値の問題も深刻である。本稿では、imputationを行わず、欠損値のまま処理したが、欠損がランダムに分布していなかった場合、推定結果にバイアスがかかる可能性は否定できない。今後はimputation技術を使った検証も行う必要があるだろう。健康尺度について、「基礎調査（健康票）」には、自己申告ではあるものの、実際の罹患情報が豊富にあり、SESとこうした罹患状況との関係についても重要なテーマである。さらに、データがパネルでないため、内生性の調整には限界があるが、データに内在する何らかの自然実験を発見し活用することも可能であろう (Nishi, Hashimoto, Noguchi et al. (2011))。「基礎調査」には存在しない数多くの貴重な情報を含む「実態調査」が広く一般の研究者の関心を集めることによって、

今後のさらなる当該分野における研究の発展が期待される。

#### 謝辞

本研究は、国立社会保障・人口問題研究所一般会計プロジェクトにより実施された2007年社会保障・人口問題基本調査「社会保障実態調査」、および、「国民生活基礎調査」（厚生労働省）に基づく。国立社会保障・人口問題研究所において開催された編集会議において、ソーシャル・ネットワークのアイデアを頂戴した小塩隆士氏、本稿の方法論についての論文を共同執筆中である西晃弘氏、ならびに、同研究所における執筆者会議において、ご助言をいただいた田宮菜奈子氏、高橋秀人氏をはじめ、出席者の方々から有益なご意見を多数頂戴した。また、日本福祉大学経済学部ワークショップにおいて、参加者の方々から貴重なコメントを頂いた。記して御礼申し上げる。本稿における残るすべての誤りは著者に帰するものである。

#### 注

- 1) 本節において、健康を「健康」と括弧付きで表記したのは、「健康」を何で測るのかという、本稿で検証するテーマにとって極めて重要な問題を孕んでいるためである。詳細については、後述する。
- 2) 社会疫学の定義やその発展の経緯については、本庄 (2007) に詳細にまとめられている。また、Antonovsky (1967) では、1960年以前の社会的・経済的諸要因 (Socio-economic status) と健康との関連性に関する先行研究の包括的レビューがなされているので、参照のこと。
- 3) 「基礎調査」(世帯票・健康票)における世帯票配布数(調査客体世帯数)は287,807票で、うち回収数は229,821票(有効回答率:79.9%)であった(厚生労働省(2008))。
- 4) 分析対象となるサンプル数については、SRH以外の健康尺度、そして健康尺度との関連性を考察するSES変数が分析によって異なるため、個別に提示する。
- 5) 「基礎調査」と「実態調査」との3週間というタイム・ラグについては、再テスト法で頻繁に用いられるタイム・フレームである。2-3週間という期間は、調査対象者が前回の回答の記憶をとどめるには長すぎ、個人属性を含め対象者

をとりまく状況や環境が大きく変わるには短すぎる期間だからである。つまり、このタイム・フレームは、再テストにおける調査対象者の回答に対し、対象者自身の「記憶」によるバイアスがかかる危険性を排除することができ、かつ、回答に対するバイアスとなる個人の外的・内的状況が一定であるという条件を満たすのに適した時間であると考えられている (Pruitt, Jeffe, Yan et al. (2010))。

- 6) 説明変数として、「基礎調査」のSRH、性別、年齢、配偶者、世帯人員数、ストレス、ストレスがある場合の相談先、学歴、就業の有無、等価所得、貯蓄、持ち家、医療保険、喫煙習慣、健康診断受診の有無、居住地域の属性(大都市)、を投入し、OLS分析を行った。回帰分析の結果については、著者による提供が可能である。
- 7) EuroQOL (EQ-5D) とは、ヨーロッパにおいて1987年以降開発されてきたHealth-related quality of life (HRQOL) スコアで、health outcomeの測定に用いられている標準化のための手法であり、健康に関する国際比較や公共政策の費用対効果分析に広く用いられている指標である。日本語版EuroQOL 開発委員会(委員長:慶應義塾大学医学部医療政策・管理学教室池上直己氏)により、正規の日本語版として認定を受けたものが存在する (Tsuchiya, Ikeda, Ikegami et al. (2002))。
- 8) 総合健康指標を作成するほかの尺度が「基礎調査(健康票)」のみで質問されているため、時間的整合性を考慮し、ここでは、「実態調査」ではなく「基礎調査(健康票)」のSRHを用いる。
- 9) 具体的な項目については、「基礎調査(健康票)」の質問票を参照のこと(厚生労働省(2008))。
- 10) 各年齢の第1主成分との相関が高かった自覚症状は、20~40歳未満と40~65歳未満で、「眠れない」、「いらいらしやすい」、「体がだるい」、「もの忘れする」、65歳以上では、「手足の動きが悪い」、「手足の関節が痛む」、「足のむくみやだるさ」、「手足のしびれ」である。第2主成分では、20~40歳未満が「きこえにくい」、「物を見づらい」、「耳なりがする」、「目のかすみ」、40~65歳未満が、「手足の動きが悪い」、「手足の関節が痛む」、「手足が冷える」、「手足のしびれ」、65歳以上が、「きこえにくい」、「物を見づらい」、「もの忘れする」、「目のかすみ」である。第3主成分では、20~40歳未満が、「月経不順・月経痛」、「手足が冷える」、「便秘」、「足のむくみやだるさ」、40~65歳未満が、「動悸」、「息切れ」、「ゼイゼイする」、「前胸部に痛みがある」、65歳以上が、「眠れない」、「いらいらしやすい」、「体がだるい」、「食欲不振」となった。
- 11) SRHとIADLについての結果は、総合健康指標よりも正確性が低かったため、ここでは省略する。
- 12) 国内外で多数の先行研究が存在する。例えば、Smith (1999); Benzeval, Taylor, and Judge (2000); Benzeval and Judge (2001); Case, Lubtsky, Paxon (2002); Meer, Miller, and Rosen (2003); Adams, Hurd, McFadden et al. (2003); Frijters, Haisken-DeNew, and Shelds (2005); 中馬・大石 (1998a); 中馬・大石 (1998b); 岩本 (2000); 大石 (2000); 泉田 (2006); 鈴木 (2007); 野口 (2008), ほか。
- 13) 学歴に関しても、子どもの頃の健康状態が良好な者の方が高い学歴を達成するかもしれない。子どもの頃の健康状態が良いのは親からの遺伝や、親の教育水準なども含め家庭環境の影響があるとするならば、現在の健康状態が良好であるのは学歴による影響ではなく、遺伝や与えられた環境によるものである。本稿のデータでは、遺伝的要素や親の教育水準などは、 $x_i$  では調整することができず、 $u_i$  として吸収されてしまう。したがって、本来であれば、学歴も内生変数として調整すべきであるが、適当な操作変数が見つからなかったため、今後の研究課題とし、ここでは学歴は所与であると仮定する。
- 14) 着目したSESのうち、現在の就労状況とソーシャル・ネットワークについては、連続変数ではなく二項変数である。また、厳密に言えば、目的変数である総合健康指標とK6にしても純粋な連続変数ではなく、本来であれば、latentモデルを採用すべきであるが、IV検定の簡易性を優先し、ここでは、すべてlinear probabilityによる推定とする。latentなモデル設定については今後の課題とする。
- 15) このほかの説明変数の結果をまとめると、総合健康指標については、年齢、配偶関係、居住区が大都市であること、ストレスが有意にマイナスの効果をもつ一方で、K6については、配偶関係のみが有意にマイナスという結果であった。また、同居家族の人数が増えるほど、総合健康指標・K6ともに改善されるという結果を得た。これらの結果については、著者による提供が可能である。

#### 参考文献

- 泉田信行 (2006) 「生活習慣病罹患と労働時間の関係」2006年日本経済学会秋季大会報告論文。  
 岩本康志 (2000) 「健康と所得」国立社会保障・人口問題研究所編『家族・世帯の変容と生活保障

- 機能』東京大学出版会。
- 大石亜希子 (2000) 「高齢者の就業決定における健康要因の影響」『日本労働研究雑誌』481, pp.51-62。
- 川上憲人 (2004) 「心の健康問題と対策基盤の実態に関する研究」総括研究報告書, 平成14年度厚生労働科学研究費補助金 (厚生労働科学特別研究事業)。
- 厚生労働省 (2008) 「平成19年国民生活基礎調査: 調査の概要」<http://www.mhlw.go.jp/toukei/list/20-19-1a.html>
- 国立社会保障・人口問題研究所 (2009) 「2007年社会保障・人口問題基本調査 社会保障実態調査: 結果の概要」<http://www.ipss.go.jp/ss-seikatsu/j/jittai2007/janda/jittai2007.pdf>
- 鈴木 亘 (2007) 「肥満と長時間労働」2007年日本経済学会秋季大会報告論文。
- 中馬宏之・大石亜希子 (1998a) 「高齢者の就業決定における健康要因の影響・『国民生活基礎調査』による分析」『高齢社会における社会保障体制の再構築に関する理論的研究事業の調査研究報告書II』長寿社会開発センター。
- (1998b) 「高齢者の就業決定における健康要因の影響・『高齢者就業実態調査』による分析」『高齢社会における社会保障体制の再構築に関する理論的研究事業の調査研究報告書II』長寿社会開発センター。
- 野口晴子 (2008) 「世帯の経済資源が産出・育児期における女性の心理的健康に与える影響について: 「消費生活に関するパネル調査」を用いた実証分析」『経済研究』59(3), pp.209-227。
- 橋本秀樹 (2005) 「社会疫学」青山英康監修, 川上憲人・甲田茂樹編『今日の疫学』pp.318-327, 医学書院。
- 本庄かおり (2007) 「社会疫学の発展」『保健医療科学』*Journal of the National Institute of Public Health* 56(2), pp.99-105。
- Adams P, Hurd MD, McFadden D, Merrill A, and Ribeiro T (2003) "Health, Wealth, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status". *Journal of Econometrics* 112, pp.3-56.
- Adler NE, Boyce T, Chesney MA, Cohen S, Folkman S, Kahn RL, and Syme SL (1994) "Socioeconomic Status and Health: the Challenge of the Gradient". *American Psychologist* 49(1), pp.15-24.
- Antonovsky A (1967) "Social Class, Life Expectancy, and Overall Mortality". *Milbank Memorial Fund Quarterly* XLV, pp.31-73.
- Berkman LF and Kawachi I (2000) "A Historical Framework for Social Epidemiology". In: Berkman LF and Kawachi, I (eds). *Social epidemiology*, pp.3-12, Oxford university press.
- Benzeval M, Taylor J, and Judge K (2000) "Evidence on the Relationship between Low Income and Poor Health. Is the Government Doing Enough?" *Fiscal Studies* 21, pp.371-399.
- Benzeval M and Judge K. (2001) "Income and Health: The Time Dimension". *Social Science and Medicine* 52, pp.1371-1390.
- Bound J (1991) "Self-Reported versus Objective Measures of Health in Retirement Model". *Journal of Human Resources* 26, pp.107-137.
- Bound J Schoenbaum M, Stinebrickner TR, and Waidmann T (1999) "The Dynamic Effects of Health on the Labour Force Transitions of Older Workers". *Labour Economics*, 6, pp.179-202.
- Bound J, Brown C, and Mathiowetz N (2001) "Measurement Error in Survey Data". In: Heckman JJ and Leamer EE (eds). *Handbook of Econometrics* 5, Chapter 59, pp.3705-3843, Elsevier.
- Case A, Lubstsky D, Paxson C (2002) "The Lasting Impact of Childhood Health and Circumstances". *Center for Health and Wellbeing Discussion Paper*, Princeton.
- Crossley TF and Kennedy S (2002) "The Reliability of Self-Assessed Health Status". *Journal of Health Economics* 21, pp.643-658.
- Cutler DM, Lleras-Muney A, and Vogl T. (2008) "Socioeconomic Status and Health: Dimensions and Mechanisms". *NBER Working Papers* 14333.
- Drummond M, Sculpher M, Torrance G, O'Brien B and Stoddart G (2005). "Methods for the Economic Evaluation of Health Care Programmes". Oxford University Press: Oxford.
- Frijters P, Haisken-DeNew JP, and Shelds MA (2005) "The Causal Effect of Income on Health: Evidence from German Reunification". *Journal of Health Economics* 24, pp.997-1017.
- Gerdtham UG, Johannesson M, Lundberg L, and Isacson D (1999) "A Note on Validating Wagstaff and Van Doorslaer's Health Measure in the Analysis of Inequality in Health". *Journal of Health Economics* 18, pp.117-224.
- Greinera M, Pfeiffer D, and Smithe RD (2000) "Principles and Practical Application of the Receiver-Operating Characteristic Analysis for Diagnostic Tests". *Preventive Veterinary*

- Medicine* 45, pp.23-41.
- Grootendorst P, Feeny D, and Furlong W (1997) "Does It Matter Whom and How You Ask? Inter and Intra-Rater Agreement in the Ontario Health Survey". *Journal of Clinical Epidemiology* 50, pp.127-136.
- Grossman M. (2006) "Education and Nonmarket Outcomes". In: Hanushek E and Welch F (eds). *Handbook of the Economics of Education* 1, Chapter 10, pp.578-633, Elsevier.
- Hsieh F and Turnbull BW (1996) "Non Parametric and Semiparametric Estimation of the Receiver Operating Characteristics Curve". *The Annals of Statistics* 24(1), pp.25-40.
- Idler EL and Kasl SV (1995) "Self-ratings of Health: Do They Also Predict Change in Functional Ability?". *Journal of Gerontology* 50B, S344-353.
- Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand S-LT, Walters EE, and Zaslavsky A. (2002) "Short Screening Scales to Monitor Population Prevalences and Trends in Nonspecific Psychological Distress". *Psychological Medicine* 32, 959-76.
- Lokshin M and Ravallion M (2008) "Testing for an Economic Gradient in Health Status Using Subjective Data". *Health Economics* 17(11), pp.1237-1259.
- McCallum J, Shadbolt B, Wang D (1994) "Self-rated Health and Survival: 7 Years Follow-up Study of Australian Elderly". *American Journal of Public Health* 84(7), pp.1100-1105.
- Meer J, Miller D, and Rosen H (2003) "Exploring the Health-Wealth Nexus". *Journal of Health Economics* 22, pp.713-730.
- Nishi A, Hashimoto H, Noguchi, H, Tamiya N (2011) "Does Coinsurance Rate Reduction Reach Poor Older People? Japan's Health Insurance Natural Experiment". *The Global Health Metrics & Evaluation: Controversies, Innovation, Accountability conference*, March 14-16, 2011 in Seattle, WA, USA.
- Pruitt SL, Jeffe DB, Yan Y, and Schootman M (2010) "Reliability of Perceived Neighborhood Conditions and the Effects of Measurement Error on Self-Rated Health across Urban and Rural Neighborhoods". *Journal of Epidemiol Community Health* [Epub ahead of print] doi:10.1136/jech.2009.103325.
- Salomon JA, Tandon A, Murray CJL, and World Health Survey Pilot Study Collaborating Group (2004) "Comparability of Self Rated Health: Cross Sectional Multi-Country Survey Using Anchoring Vignettes". *British Medical Journal*; 328:258 doi:10.1136/bmj.37963.691632.44.
- Smith JP (1999) "Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status". *The Journal of Economic Perspective*, 13(2), pp.145-166.
- Staiger D and McClellan (2000) "Comparing the Quality of Health Care Providers". In Alan Garber (ed). *Frontiers in Health Policy Research* 3, pp.113-136, the MIT Press.
- Tourangeau R and Smith TW (1996) "Asking Sensitive Question the Impact of Data Collection Mode, Question Format, and Question Context". *Public Opinion Quarterly* 60, pp.275-304.
- Tsuchiya A, Ikeda S, Ikegami N, Nishimura S, Sakai I, Fukuda T, Hamashima C, Hisashige A and Tamura M (2002) "Estimating an EQ-5D Population Value Set: The Case of Japan". *Health Economics* 11(4), pp.341-353.
- Zou KH and Zhou XHA (2001) "Receiver Operating Characteristics (ROC) Analysis". *Statistics in Epidemiology Report* 1, pp.1&4-12.

(のぐち・はるこ 国立社会保障・人口問題研究所  
社会保障基礎理論研究部第2室長)