

家計消費の経済分析

(11) 所得過程の推計

阿部修人

一橋大学経済研究所

平成 30 年 10 月 16 日

1 導入

家計が所得や選好に対し様々なショックにさらされていたとしても、効用関数が時間に関して加法に分離可能である限り、消費のオイラー方程式は単純な条件付き期待値を用いて描写可能であり、所得や選好が具体的にどのような確率過程に従うかを考慮する必要はなかった。各期の効用関数が CRRA であるとき、消費のオイラー方程式は、今期と来期の消費の関係を表わす式であるが、各期の消費の水準がどのような値になるべきかについては、我々はオイラー方程式から学ぶことが出来ず、No Ponzi Game 条件等の境界条件を用い、動学最適化問題を完全に解き切らねばならない。オイラー方程式は最適消費経路の必要条件の一部に過ぎないのである。

もしも、分析者の関心が不確実性が存在することによる消費の減少や、将来流動性制約にバインドすることを考慮にいた消費計画にあるのであれば、オイラー方程式のみでなく、消費の水準も含め、最適消費経路全てを導かねばならない。そのためには、家計がどのような不確実性に直面しているかを明確にする必要がある。具体的には、所得や健康状態、選好の変化等の確率過程を定式化した上で、動学最適化問題を解かねばならないのである。

所得の不確実性の推計は家計消費行動に興味を持つマクロ経済学者や教育投資のリターンに関心を持つ労働経済学者等、非常に多くの研究者の関心事であり、膨大な量の論文が書かれている。古くは、50 年以上前に Friedman and Kuznets (1954) が小規模のサンプルを用いて分析を行ったものまで遡ることが可能であるが、本格的な分析は、スウェーデンの家計パネルデータを用いた Hause (1980) および PSID やアメリカ合衆国の Ph.D. 取得者 1 万人強のパネルデータを用いた Lillard and Wills (1978) および Lillard and Weiss (1979) から始まる。所得過程の推計の際、特に次の二つの点が重視されてきた。第一の問題は、家計所得変動の持続性である。恒常所得仮説に従うと、人々は一時的な所得ショックに対しては消費を変化させず、一方、恒常的な

ショックに対しては消費を変化させる。したがって、消費研究の視点では、所得変動に占める恒常的要因の推計は極めて大きな問題となっている。第二の論点は、観察される家計間の所得格差は、どの程度が労働市場参入時点で決定される、家計固定効果によるものであるか、という問題である。例えば、能力のある家計は、年率7%で所得が成長し、一方能力のない家計所得は年率1%でしか成長しないと仮定すると、自由に貸借可能であれば、所得分散は年齢とともに増加しても、消費分散は変化しない。将来の所得経路を家計は正確に予測可能であるためである。

所得過程に関しては膨大な量の研究があり、Meghir and Pistaferri (2010)のような優れたサーベイがある。そのため、本章では、所得過程推計の論点を整理するとともに、日本における所得過程の推計結果を主に紹介する。

2 所得分散の加齢効果と所得過程

図1は、アメリカ合衆国における家計所得(対数値)の年齢・分散プロファイルを示している¹。図1から直ちに、対数家計所得の分散は年齢とともに増加していることが分かる。この所得分散は異なる家計間の分散であり、同一家計の生涯を通じた所得変化の分散ではないことに注意する必要があるが、もしも、各家計が idiosyncratic な所得リスクに直面しており、所得の実現値の違いは、そのようなリスクの帰結であると仮定すると、所得分散の年齢とともに増加するという事実は、所得のの不確実性に直面しており、

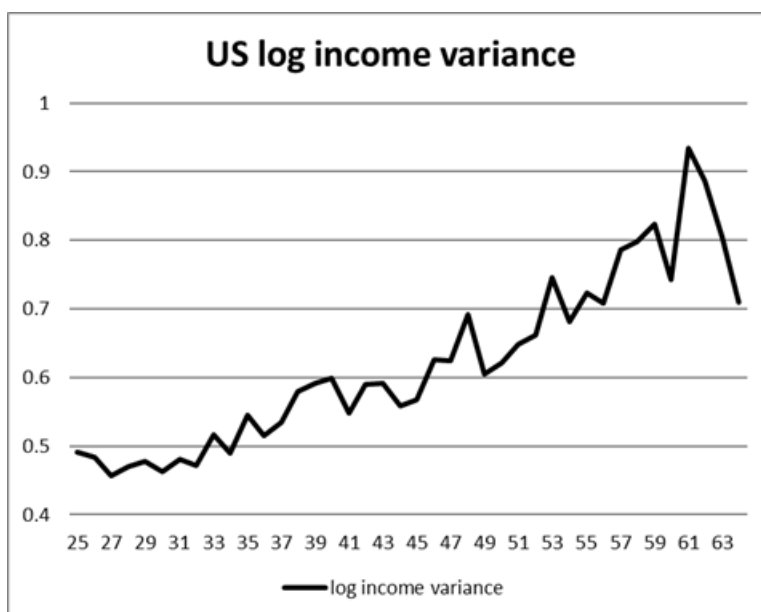
家計 i が h 歳の時、 t 年に受け取る所得(あるいは賃金)の対数値を $y_{i,h,t}$ とすると、多くの研究では下記のような所得過程が推計されている。

$$\begin{aligned} y_{i,h,t} &= X_{i,h,t}\beta + \alpha_i h_{i,t} + \varepsilon_{i,h,t} + \omega_{i,h,t}, \\ \omega_{i,h,t} &= \rho \omega_{i,h-1,t-1} + \zeta_{i,h,t}, \end{aligned} \quad (1)$$

$X_{i,h,t}$ は観察可能であり、かつ家計間で共通な効果をもつ所得の説明要因。 α_i は所得成長率の家計間格差である。 $\varepsilon_{i,h,t}$ は系列相関をもたない誤差項であり、一時的な所得ショック、あるいは所得データの計測誤差に対応する。 $\omega_{i,h,t}$ は一次の自己回帰に従う持続的($\rho = 1$ のときは恒久的)所得変動であり、 $\zeta_{i,h,t}$ は系列相関がなく、かつ、 $\varepsilon_{i,h,t}$ およびそのラグ項とも相関をもたない、持続的(恒久的)所得ショックである。なお、ここでは $\omega_{i,h,t}$ が一次の自己回帰に従うと仮定しているが、より高次のラグを含めることも、あるいは移動平均で定式化されることもある²。

¹データは Consumer Expenditure Survey(CEX)。1980年から2007年までの対数家計所得を年ダミーや家計構成人数ダミーに回帰し、その残差の分散を年齢ごとに計算しプロットしたもの。横軸は世帯主の年齢である。なお、年齢間の平均値の違いは、年齢ごとの対数分散であるため、分散の推計値には影響を与えないことに注意する必要がある。

²ショック項の分散が変動することを許容する分析もある。詳しくは Meghir and Pistaferri (2004) を参照せよ。



今、単純化のため観察可能な変数の効果 $X_{i,h,t}\beta$ を無視し、さらに α_i の分散 σ_α^2 がゼロで $\rho = 1$ を仮定しよう³。すると、

$$y_{i,h,t} = \alpha h_{i,t} + \varepsilon_{i,h,t} + \omega_{i,h,t}. \quad (2)$$

さらに、 ε および ζ の分散が一定であると仮定し、同一時点における年齢別の分散を計算すると、両辺の同一時点における年齢別の分散を計算すると、 α は確率変数ではなく、かつ、 $h_{i,t}$ の値も同一なので、

$$\text{Var}(y_{i,h,t}) = \sigma_\varepsilon^2 + \text{Var}(\omega_t), \quad (3)$$

$$\text{Var}(y_{i,h-1,t-1}) = \sigma_\varepsilon^2 + \text{Var}(\omega_{t-1}). \quad (4)$$

ところで、

$$\text{Var}(\omega_t) = \text{Var}(\omega_{t-1}) + \sigma_\zeta^2. \quad (5)$$

したがって

$$\text{Var}(y_{i,h,t}) = \text{Var}(y_{i,h-1,t-1}) + \sigma_\zeta^2, \quad (6)$$

すなわち、所得の分散は、恒常所得ショックの分散 σ_ζ^2 の分だけ、図1のように加齢とともに増加していくことになる。

³ 計量分析の歳には $X_{i,h,t}\beta$ の効果は、対数所得を説明変数に回帰した残差を用いることで取り除くことが可能である。

再び、(1) 式に戻り、今度は $\sigma_\alpha^2 > 0, \rho = 0$ を仮定しよう。前と同様に、 $X_{i,h,t}\beta$ を無視し、各ショック項の分散が一定であると仮定すると、

$$\text{Var}(y_{i,h,t}) = h^2\sigma_\alpha^2 + \sigma_\varepsilon^2 + \sigma_\zeta^2, \quad (7)$$

$$\text{Var}(y_{i,h,t}) = (2h - 1)\sigma_\alpha^2 + \text{Var}(y_{i,h,t-1}), \quad (8)$$

となり、 $h > 1$ であることから、加齢とともに所得分散は増加していくが、それは、所得ショックによるものではなく、家計間で所得成長率が異なるためである。

(6) 式と (8) 式は、どちらも図 1 で観察される、加齢に伴う所得分散の増加を説明することが可能である。しかし、その経済学的含意は大きく異なる。(6) 式の場合、所得分散の増加は、ex-ante で同質な家計に発生する、保険の存在しない各家計間で独立の idiosyncratic な所得ショックによるものである。したがって、もしも恒常的所得ショックに対する保険が存在すれば、経済厚生を改善することが可能である。一方、(8) 式の場合は、所得格差の増加は生まれた時から決まっている個人の能力の差により生じているものであり、不確実性やリスクとは関係ない⁴。所得格差の拡大は、能力という固定された初期資源配分の結果を反映しているのみであり、政策により厚生を改善することはできないのである。

3 所得成長率の家計間異質性の識別

3.1 所得データのみを用いる推計

スウェーデンの家計パネルデータを用いた Hause (1980)、PSID を用いた Lillard and Wills (1978) およびアメリカの PH.D. 保有者のデータを用いた Lillard and Weiss (1979) 等、初期の所得過程の推計結果は、 σ_α^2 の値を有意に推計しており、かつ、Lillard and Wills (1978) によると、クロスセクション方向で観察される所得分散の 75% が固定効果により説明されるという結果を報告している。また、 ρ の値は小さく、所得の誤差項は定常であり、 $\rho = 1$ は棄却されている。

いま、 $\sigma_\alpha^2 > 0$ および、 $\rho = 1$ を仮定しよう。このとき、MaCurdy (1982) は、PSID を用い、 σ_α^2 はゼロであるという結果を得ている。(1) 式に戻り、 $X_{i,h,t}$ を無視、あるいは $X_{i,h,t}$ に対数所得を回帰した残差を用いて分析すると仮定しよう。このとき、所得の階差を用いると、

$$\Delta y_{i,h,t} = \alpha_i + \Delta \varepsilon_{i,h,t} + \zeta_{i,h,t}. \quad (9)$$

⁴通常の固定効果分析では、階差あるいは固定効果変換を行うことで家計間異質性の水準に与える影響を除去可能であるが、ここでの家計間異質性は水準ではなく、傾きに与える影響であることに注意する必要がある。

ここで、ショック項同士は無相関で、かつ分散が一定であると仮定し、 $\Delta y_{i,h,t}$ の一次の自己共分散を計算すると、

$$Cov(\Delta y_{i,h,t}, \Delta y_{i,h-1,t-1}) = Cov(\alpha_i + \Delta \varepsilon_{i,h,t}, \alpha_i + \Delta \varepsilon_{i,h-1,t-1}), \quad (10)$$

$$= \sigma_\alpha^2 - \sigma_\varepsilon^2. \quad (11)$$

二期以上離れると、

$$Cov(\Delta y_{i,h,t}, \Delta y_{i,h-s,t-s}) = \sigma_\alpha^2 \quad \text{where } s > 2. \quad (12)$$

したがって、所得変化率の自己共分散は、二期以上離れると定数に収束していく。実際には、一時ショックや持続ショックの誤差項に系列相関が発生している可能性があり、自己共分散はより複雑になる可能性があるが、長期的には σ_α^2 に収束していくはずである。MaCurdy (1982) は、所得の自己相関係数を報告し、一期で-0.32、二期で-0.09、三期で-0.01 と二期以降では非常に小さく、かつ統計的には三期以降は全てゼロを棄却できないことから、 σ_α^2 はゼロであるという結論を導いている。そして、 σ_α^2 をゼロと仮定して推計する場合、 ρ の推計値は 0.975 となり、 $\rho = 1$ は棄却できず、所得過程はランダムウォークと $\sigma_\alpha^2 = 0$ を仮定した確率過程により描写可能であるとしている。同様の結果は、より一般的な確率過程の可能性を調べた Abowd and Card (1989)、労働者・企業のマッチングデータを用いた Topel and Ward (1992) でも得られている。MaCurdy (1982) や Abowd and Card (1989) の結果は長らく標準的なものとみなされ、Carroll 等による一連の予備的貯蓄モデルのカリブレーションにおいても、所得過程は一時ショックとランダムウォークに従う恒常ショックの二種類に分離された形で定式化されている。

近年、Güvenen (2009) は、この結論を批判している。もしも、真の所得過程が $\rho < 1$ かつ、 $\sigma_\alpha^2 > 0$ であるならば、所得変化率の自己共分散は

$$Cov(\Delta y_{i,h,t}, \Delta y_{i,h-s,t-s}) = \sigma_\alpha^2 + \rho^{s-1} \left[\frac{\rho - 1}{\rho + 1} \right] \sigma_\zeta^2, \quad (13)$$

となることを指摘している。 $\rho < 1$ である限り、右辺第二項は負の値となる。もしも、家計パネルの期間を十分に長く取ることが出来ない場合、所得変化の自己共分散は σ_α^2 よりもかなり小さな値となり、統計的にゼロを棄却できない可能性が高くなる。Güvenen (2009) は、シミュレーションを行い、PSID の所得変化の自己相関から σ_α^2 の検証を行うのは不可能であると論じている。さらに、 $\sigma_\alpha^2 > 0$ を許容し所得過程を推計すると、Lillard and Willis (1978) と同様に、所得変化率の家計間異質性は極めて重要であり、一方、 ρ は 0.82 となり、1 よりも小さくなることを報告している。

Hryshko (2009) は、真の所得過程が $\sigma_\alpha^2 = 0, \rho = 1$ であると仮定し、さらに、AR(1) に従う (自己相関係数は 0.5 および 0.25 と仮定) 一時的ショックを導入したシステムを仮定し、所得系列をシミュレートした上で、Güvenen

(2009) のモデルに従って σ_α^2 と ρ を推計すると、 σ_α^2 は有意な正の値となり、一方、 ρ は 1 よりもかなり小さな値となることを報告している。さらに、PSID を用い、様々なスペシフィケーション下で所得過程を推計すると、 $\sigma_\alpha^2 = 0, \rho = 1$ を支持する結果が多く得られると報告している。

家計間で所得成長率が異なり、なおかつ恒常所得ショックが存在するケースも考えられるが、どちらか一方を仮定すると、両方とも同様の効果を所得分散に与えるため、他方の重要性は低下していく。両者の識別は困難であり、特に、所得データそのものに計測誤差があり、それが系列相関を持つ場合は、さらに両者の識別は困難になる。

3.2 消費データを用いる推計

所得データを用いて、所得過程の推計を行うことには限界がある。具体的には、所得データに含まれる計測誤差と一時的な所得変動を識別できないのである。本章の最初で、所得過程は家計消費の動学最適化問題を具体的に解くために必要であると議論したが、その議論を逆転し、家計消費行動がよくわかっているならば、家計消費の動向から所得に含まれる不確実性を推計することも可能である。

第 5 章では、Hall (1978) による恒常所得モデルに基づき、所得と消費の間の下記の関係を導いた。

$$\theta_t \Delta c_t = \frac{r}{1+r} \sum_{s=0}^{T-t} \frac{(E_t - E_{t-1}) y_{t+s}}{(1+r)^s}, \quad (14)$$

$$\text{where } \theta_t = \left(1 - \frac{1}{(1+r)^{T-t+1}} \right). \quad (15)$$

ここで、いま、観察される対数所得が下記のような、恒常的所得変動 y_t^p と恒常ショック v_t 、一時的ショック u_t 、および計測誤差 ω_t で描写可能であるとする。

$$y_t = \alpha + y_t^p + u_t + \omega_t, \quad (16)$$

ただし、各ショック項と計測誤差に系列相関はなく、かつ互いに無相関であり、期待値はゼロであるとする。

$$y_t^p = y_{t-1}^p + v_t, \quad (17)$$

すると

$$y_{t+s} = y_{t-1} + u_{t+s} - u_{t-1} + \sum_{\tau=t}^{t+s} v_\tau + \sum_{\tau=t}^{t+s} \omega_\tau. \quad (18)$$

ところで、

$$E_{t-1} y_{t+s} = y_{t-1} - u_{t-1} \quad (19)$$

したがって、

$$\begin{aligned}\theta_t \Delta c_t &= \frac{r}{1+r} \sum_{s=0}^{T-t} \frac{(E_t - E_{t-1}) y_{t+s}}{(1+r)^s} \\ &= \frac{r}{1+r} u_t + \theta_t v_t.\end{aligned}\quad (20)$$

$$\Delta c_t = v_t + \frac{r\theta_t^{-1}}{1+r} u_t.\quad (21)$$

消費変化率の分散は

$$Var(\Delta c_t) = Var(v_t) + \left(\frac{r\theta_t^{-1}}{1+r}\right)^2 Var(u_t).\quad (22)$$

消費変化率と所得変化率の共分散を計算すると

$$\Delta y_t = \Delta u_t + \Delta \omega_t + v_t,\quad (23)$$

$$Cov(\Delta c_t, \Delta y_t) = Var(v_t) + \frac{r\theta_t^{-1}}{1+r} Var(u_t).\quad (24)$$

となる。特に、共分散を用いる場合、消費データに計測誤差が含まれていても、それが系列相関がなく、真の消費データや所得と無相関である限り、共分散の値に影響を与えないため、信頼性が高くなる。

所得と消費の変化率の共分散情報を用いるには、所得と消費の両方のデータがあるパネルデータが必要である。Hall and Mishkin (1982) 等が用いている PSID は、長い間食料支出しか記録しておらず、構造モデルの推計には必ずしも向いていない。多くの分析は、(1) イタリア銀行が行っている Bank of Italy Survey of Household Income and Wealth (SHIW) を用いるか (Pistaferri (2001))、(2) FES 等のリピーティドクロスセクションからコホートデータを作成し、疑似パネルを作成する (Blundell and Preston (1999))、(3) CEX に含まれる全消費データを、PSID の食料支出等の情報を用いてマージして使用する (Blundell, et al. (2008)) の三種類のうちどれかを選択している。

イギリスの FES を用いた Blundell and Preston (1999) は、イギリス家計が直面する一時的な所得ショックが 1968 年から 1992 年まで増加しているという結果を得ている。また、Blundell, et al. (2008) は、アメリカにおいて、所得分散と消費分散の乖離が大きくなった理由として、所得ショックに対する保険が充実したという Kruger and Perri (2006) の推測を否定し、所得変動に占める一時的ショックの割合が増加し、所得変動そのものは大きくなっていても、恒常的所得変動は増加していないためであるという結果を得ている。一方、Guvenen and Smith (2009) は、Blundell, et al. (2008) の手法を用い PSID と CEX のマージを行い、長期の消費・所得パネルデータを作成

した上で、Guvenen (2009) 同様、所得ショックの持続性は低い ($\rho = 0.76$) という結果を得ている⁵。

上記で展開したモデルは、Hall (1978) の恒常所得モデルに準拠し、恒常所得変動に対しては一切の保険が存在しないと仮定していたが、恒常所得変動に対する保険の程度を推計することも可能である。無限視野のとき、(24) 式は

$$Cov(\Delta c_t, \Delta y_t) = Var(v_t), \quad (25)$$

となる。いま、一般的に所得ショック x_t に対する保険の程度の指標として、

$$\phi(x_t) = 1 - \frac{Cov(\Delta c_t, x_t)}{Var(x_t)}, \quad (26)$$

を考えよう。 x_t が恒常ショック ($x_t = v_t$) のとき、恒常所得モデルでは $\phi = 0$ となり、恒常ショックに対する保険は存在しない。また、 x_t が一時ショック ($x_t = u_t$) のときは、 $\phi = 1$ となり、完全な保険が存在する。

いま、所得データに含まれる計測誤差 ω_t を無視すると、

$$\Delta y_t = \Delta u_t + v_t, \quad (27)$$

この共分散をとると、

$$Cov(\Delta y_{t+1}, \Delta y_t) = -Var(u_t). \quad (28)$$

また、

$$Cov(\Delta c_t, \Delta y_{t+1}) = -Cov(\Delta c_t, u_t). \quad (29)$$

恒常ショックに関しては、下記のような和を考える。

$$\begin{aligned} \Delta y_{t+1} + \Delta y_t + \Delta y_{t-1} &= \Delta u_{t+1} + \Delta u_t + \Delta u_{t-1} + v_{t+1} + v_t + v_{t-1}, \\ &= u_{t+1} - u_{t-2} + v_{t+1} + v_t + v_{t-1}, \end{aligned} \quad (30)$$

すると、恒常ショックの分散は

$$Cov(\Delta y_t, \Delta y_{t+1} + \Delta y_t + \Delta y_{t-1}) = Var(v_t), \quad (31)$$

で与えられる。したがって、一時ショックに対する保険の程度としては

$$\phi(u_t) = 1 - \frac{Cov(\Delta c_t, \Delta y_{t+1})}{Cov(\Delta y_{t+1}, \Delta y_t)}, \quad (32)$$

恒常ショックに対しては

$$\phi(v_t) = 1 - \frac{Cov(\Delta c_t, \Delta y_t, \Delta y_{t+1} + \Delta y_t + \Delta y_{t-1})}{Cov(\Delta y_t, \Delta y_{t+1} + \Delta y_t + \Delta y_{t-1})}, \quad (33)$$

⁵労働市場参入時点での固定効果の重要性を指摘した近年の研究に Huget et al. (2008) がある。

で保険の程度を計測可能である。Blundell et al. (2008) は、上述のようにアメリカの PSID と CEX をマッチングさせたデータを用い、一時ショックに対する保険は 95%、恒常ショックに対する保険は 36% 存在するという結果を得ている。Kaplan and Violante (2010) は、第 6 章で議論したような、CRRA 型効用関数と流動性制約を伴う予備的貯蓄モデルをシミュレートし、Blundell et al. (2008) の結果が再現されるか否かを検証した。無限視野の確実性等価モデル (各期の効用関数が二次式) では、一時ショックに対する保険は完全で、恒常ショックに対する保険はゼロとなるが、より一般的なモデルでは、解析的に消費と所得の共分散を求めることは不可能であり、保険がどの程度存在するかはパラメータの値に依存してくる。Kaplan and Violante (2010) は所得過程として所得が i.i.d. ショックとランダムウォーク (恒常ショック) により構成されると仮定し、シミュレーションを行った。その結果、一時ショックに対してはほぼ完全な保険が観察されたが、恒常ショックに対する保険は 23% と、Blundell et al. (2008) よりも低い値とならしている。所得と消費の相関を用いる分析を厳密に進めるためには、各家計が所得変化をいつの時点で認識したかを知る必要がある。これまでの分析では、所得変化はマルチンゲールとなっており、その変化は全く予期不可能とされていたが、Feigenbaum and Li (2008) 及び Primiceri and Rens (2009) は、家計単位の過去の所得等の情報を用い、将来の家計所得を予測し、実現値との乖離を用いて家計の直面するリスクとみなし、所得リスクの推計を行っている。

消費データを用いることで、所得データでは観察不可能な情報を用いることの利点は確かに大きく、また、所得の恒常的変動を求める際に過剰識別に由来するという推計上の利点も無視できない。とはいえ、消費と所得両方のデータを高精度に記録している長期のパネルデータは少なく、未だ、推計結果も分析者の間で一致をみていない。Güvenen (2009) のような、家計間の所得変化率の異質性を強調する立場と、所得に占める恒常的ショックの重要性を強調する立場の論争は今後も継続していくと思われる。

4 日本における所得過程の推計

4.1 導入

本セクションは、日本家計の消費・貯蓄行動を動学モデルを解く際に必要な、家計の直面する所得の動学的な確率構造に関して、限定的ではあるが、分析を試みるものである。その際に、特に二つの目標を設定している。第一は、一般の研究者に利用可能な唯一の長期に渡る日本の家計パネルデータである家計経済研究所のデータを用い、所得および労働時間のデータに含まれる真の情報と測定誤差を識別することである。より具体的には、経験年数や年齢などをコントロールした後の所得変動のどの程度が恒常的なものであり、

どの程度が一時的な変動によるものかを計算する。第二の目標は、アメリカの PSID 等のデータを用いた先行研究である Abowd and Card (1989) により得られた共分散行列と同様の行列を計算し、日米比較を通じて日本家計の特徴を明らかにすることである。主要結果は下記の通りである。

(1) 日本の所得変化率の分散はアメリカの 1/3 程度である。

(2) 所得変化率の一階のラグとの共分散はマイナスであり、分散の 1/2 弱である。

(3) 二期以上離れると、相関はほぼ消滅する。

(4) 所得と労働時間の相関は極めて小さく、多くの場合有意に検出されない。

(2) と (3) はアメリカでの先行研究とほぼ同じであり、(1) と (4) が大きく異なる。これは、日本の家計所得過程はアメリカと同様の傾向を有しているが、分散の大きさの点で相違があることを示している。また、所得および労働時間を Random Walk と一時的変動の混合モデルにより近似し、推計した場合は MA(2) モデルよりも良好な Fit を得ることが出来た。また、同時に観察される所得分散の 2/3 程度が一時的変動であるという結果を得た。なお、データセットの詳細および推計の技術的側面の説明はすべて Appendix で説明している。

4.2 データ

本節では、財団法人家計経済研究所による「消費生活に関するパネル調査」を用いる。この調査における対象者は若年女性であり、収入・支出・貯蓄、就業行動、家族関係などの諸側面から若年女性の生活実態を明らかにすることが調査目的としてあげられている。調査対象として選ばれた若年女性には、学校卒業後就職し、結婚・出産を迎えるといったように世帯変動が大きく、または、シングルとしての生活を継続するといったように個人のライフスタイルが多様な年齢にあるという点に大きな特徴がある。また、先に挙げた生活実態の把握という点に加え、世代の違い、ライフステージの移行過程での変化、特別な出来事に直面しての変化に着目することによって、国民生活に関わる適切な提言を行うという点も調査の大きな目的である⁶。

次に、調査の具体的内容について以下に述べる。第 1 回調査が行われたのは 1993 年で、その時点における満 24 歳から 34 歳までの女性が調査対象となっている⁷。第 1 回調査の完了数は 1500 票（コーホート A）で、1997 年調査（第 5 回調査）で満 24 歳から 27 歳までのサンプル 500 票（コーホート B）が追加されている。我々が使用したデータは、1993 年（第 1 回調査）から 2002 年調査（第 10 回調査）までの計 10 回分である。10 年間のサンプル

⁶調査の目的については、財団法人家計経済研究所のホームページから引用した。
URL:<http://www.kakeiken.or.jp/research/aboutpanel.html>

⁷調査対象は日本全国で、調査時期は毎年 10 月となっている。層化 2 段無作為抽出法で抽出された世帯に対し、留置法による調査が行われている。

サイズの推移については表 1 に示した。ここで、パネルデータを利用する際には、脱落データの問題について留意する必要がある。「消費生活に関するパネル調査」を用いて脱落サンプルの研究を行った坂本（2003）は、脱落サンプルには以下のような特徴があると報告している。それは、結婚予定者や新婚者といった、大きなイベント前後における脱落傾向が強く見られるという点であり、また、有配偶者については、本人、夫の収入変化額が正に大きいほど脱落する傾向があるという点である。さらに坂本（2003）は、脱落サンプルによる推計バイアスの影響について、世帯消費の推計を例に考察しており、脱落サンプルによる影響は観察されないと報告している。

調査の全体像は上に述べた通りであるが、これ以降、本章の分析で対象とする世帯、諸変数について説明を行う。はじめに勤労所得⁸、年間労働時間⁹の共分散構造を求める際に対象とした世帯の選択方法について述べ、次に Abowd and Card(1989) に即して分散共分散行列を推計する前段階での推定（First Stage の推定）について説明を行う。

対象とした世帯は、夫に関してプラスの勤労所得、プラスの労働時間¹⁰が確認される世帯である。これにより、配偶者のいないサンプルや、夫が勤労所得以外の所得を得ているサンプルなどが除かれる¹¹。これらの選択によって、サンプルサイズは以下ようになる。初年度における有配偶者率は 66.8%（1002/1500 世帯）である¹²。さらに、有配偶者のうち、初年度で夫が就業している世帯は 99.2%（994/1002 世帯）で、このうちの 84.6%（841/994 世帯）について夫が勤労者である。

次に、First Stage の推定について以下に説明する。Abowd and Card（1989）では、世帯間、世帯内での就業経験の相違をコントロールするために、勤労所得（ Y_{it} ）と労働時間（ L_{it} ）の対数差分を被説明変数とする以下の推定を第一段階で行なっている。

$$\Delta \ln Y_{it} = X_{it}\beta' + v_{it} \quad (34)$$

$$\Delta \ln L_{it} = X_{it}\beta' + \mu_{it} \quad (35)$$

$$\Delta g_{it} = \hat{v}_{it}, \Delta h_{it} = \hat{\mu}_{it} \quad (36)$$

ここで得られた残差 Δg_{it} 、 Δh_{it} が経験、年齢などでコントロールされた

⁸夫の税込み勤労所得を指す。また、2000 年を基準年とした消費者物価指数（総合）を用いて実質化を行った。

消費者物価指数データ出所：<http://www.stat.go.jp/data/cpi/2004np/zuhyou/a009-1.xls>

⁹年間労働時間を直接回答させる設問が存在しないため、平日の 1 日あたり労働時間、1 日あたりサービス残業時間、年間労働日数、の設問から（ - ）× として算出した。ちなみに、賃金の支払われていないサービス残業時間は含まない。

¹⁰夫が勤労者ではなく事業所得等を得ている世帯では、労働時間についての設問が存在しないため、労働時間を把握することができない。

¹¹この他にも異常値と思われるサンプルを取り除いた。異常値の取り扱いについては、Appendix1 を参照のこと。

¹²ちなみに、第 10 回調査での有配偶者比率は 81.5%。（パネル B を含まない。パネル B を含んだ場合は 77.5%。）

勤労所得、労働時間の変動分 (Experience Adjusted Changes) であり、本章における分散共分散行列の推定もこれらを用いて行なわれる。

説明変数 X_{it} として Abowd and Card(1989) は、年次ダミー、年齢と就学年数から求めた就業年数¹³を用いているが、本章ではこれに加え家族構成、同居の親に所得がある世帯を識別するダミー変数を用いた。

これまでに挙げた変数について、各年において欠損値がある世帯を取り除いた結果、推定に用いたサンプルサイズは表1のようになる。ちなみに、サンプル期間すべてを通じて当該変数に欠損値のない世帯を対象とした場合は、サンプル世帯数 262 のバランスパネルデータになる。First Stage の推定に用いた変数について基本統計量を表2に示した。年間勤労所得の平均値は 530 万円である。年間労働時間の平均値は 2326 時間で、例えば、1 年間に 52 週、週休 2 日で働いたとすると、1 日あたり 8.9 時間労働したことになる。世帯主年齢の平均値は 36 歳で、就業年数の平均値は約 15 年となっている。家族構成については、夫婦と子のみという世帯が最も多く (55 %)、次いで親と同居している世帯が多い (35 %)。また、同居している親が働いている世帯の割合は 11 % で、親と同居している世帯のうち、約 3 分の 1 の世帯について親に所得があることがわかる。また、バランスパネルデータについての基本統計量を表3に示した。アンバランスパネルデータと比較した場合、年間勤労所得は 578 万円でアンバランスパネルデータよりも高く、年間労働時間は 2300 時間と若干短い。また、家族構成について、バランスパネルデータに該当する世帯では、夫婦のみの世帯の割合が 5 % と低くなっている。

次に、勤労所得と年間労働時間の推移について表4に示した¹⁴。上段がレベルの推移で、下段は対数差分の推移を表している。年間勤労所得平均値は 1998 年、1999 年で若干減少するものの、サンプル期間を通じて緩やかに増加していることがわかる。一方、年間労働時間は年によって減少もしくは増加しており、一貫した傾向はみられない。バランスパネルデータについての勤労所得と年間労働時間の推移は表5に示した。勤労所得はサンプル期間を通じて一貫して上昇しているが、労働時間については、アンバランスパネルデータ同様、一貫した上昇あるいは減少傾向はみられない。これら二つの表から、バランスパネルデータとアンバランスパネルデータでは、年間勤労所得の平均値を比べた場合、前者の方が常に上回っていることがわかる。この背後には、脱落せずに全サンプル期間中観察される世帯を用いたサンプルセットと、そうでない世帯のサンプルセットに何らかの違いが存在する可能性は否めない。ただし、対数差分に目を移した場合、これらのデータセット間で系統だった違いは観察されない。

¹³本稿では、就業年数を以下のように算出した。調査初年度において、学校卒業以降の通算就労期間に関する設問がある。これ以降、各調査年度において世帯主の就労が確認される場合は先の通算就労期間に 1 を加算し、各年における就業年数とした。

¹⁴第 1 回調査では年間労働時間に関する設問が存在しない。

4.3 予備的分析

共分散構造を用いた詳細なモデル分析の前に、本節では前節で説明したデータを用い(1) 様々な回帰分析、および(2) 共分散行列の観察を通じて、日本家計の所得・労働時間の特徴に関して考察する。

4.3.1 回帰分析

表6および表7は対数家計所得および対数労働時間を自己ラグ等に回帰した結果である。残念ながら、回帰分析の結果はあまり好ましいものとは言えない。被説明変数のラグを含む固定効果推定は、ラグ項に関しては負のバイアスだが、通常のOLSに関しては正のバイアスが発生することが知られている。事実、表6における固定効果モデルの所得ラグ項の係数は0.131であり、OLS推定量に比べてはるかに低い。また、変量効果モデルの決定係数の分解から、家計所得分散の大部分が家計間格差であることがわかる¹⁵。

表7のArellano and Bondによる推定量では、一階のラグ項は有意であるか、有意であっても極めて小さい値となっている。ラグ項の推定値が極めて小さいことは、このデータに含まれる一時的変動、あるいはMeasurement Errorが大きいことを強く示唆している。測定誤差が深刻な問題である場合、回帰分析の推定量は下方バイアスをもち、この影響を是正するには操作変数が必要であるが、適切な操作変数を見つけるのは極めて困難である。

単純なモデルの推計結果が思わしくない場合、とるべき道の一つは、より複雑なモデルの推計を試みることであろう。Meghir and Pistaferri (2004)のようにパネルGARCHを行う試み、あるいはAlvarez, Browning, and Ejrnae (2001)のように、非常に多くのモデルをSMDで推計し比較していくアプローチも考えられる。本章では、このようなアプローチをとらず、データの共分散構造をまず眺め、その共分散行列が示唆する若干のモデルについて推定を行うことにする。このアプローチを採用する理由の一つは、共分散構造そのものは、特定のモデルに依存せず推定することが可能であることである¹⁶。どのようなモデルを仮定するにせよ、所得・労働時間の共分散構造は極めて重要な情報であるし、まずそのパターンを眺めることは重要なステップであると考えられる。第二の理由は、所得・労働時間の共分散構造分析は、欧米において膨大な先行研究がある一方、日本において、筆者の知る限り存在しないことである。我々の用いるデータが果たして欧米のデータと比較可能であるか否かという問題は残るが、データの加工および統計手法に関しては他の先行研究とそろえることが可能であり、日本のデータの特徴を掴むことが可能

¹⁵したがって、クロスセクションデータを用い、異なる家計の所得をプールし、その家計属性を十分に考慮することなく、観察される所得分散を同一家計が直面する所得の分散とみなすアプローチは危険であることが伺える。

¹⁶無論、First Stageで経験年数等をコントロールする際には「モデル」を考えていることになるが、所得の動きそのものに関しては特に強いモデルを仮定しているわけではない。

となる。第三の理由は、データに含まれる一時的変動あるいは Measurement Error の大きさをある程度推計することが可能であることをあげることができる。モデルとデータを fit させる際、データに潜む誤差をあらかじめ知っておくことは、モデルの評価の際に絶対に必要であるためである。

4.3.2 共分散構造

表 8 は日本の勤労所得変化率および総労働時間変化率の共分散行列である。先行研究に従い、上三角部分は相関係数を、下三角部分は共分散成分を示し、括弧内の数値は共分散成分の標準誤差を示している¹⁷。まず目につくのは、所得成長率の分散が年により一定ではなく、1996-97 年の 0.0388 と 2001-02 年の 0.1053 まで大きな変化が生じていることがあげられる。標準誤差が小さいことから、検定するまでもなく、所得過程が定常ではないことがわかる¹⁸。それに比べると総労働時間変化率の分散は安定しており、0.09 から 0.1 程度である。次に、所得変化率、労働時間変化率共に、一期先との相関は負であり、共分散の絶対値は分散の半分程度となっていることがわかる。しかしながら二年以上経つと相関は急速に低下し、ときおり標準誤差が小さくなるときはあっても、二年以上離れた所得変化率、および労働時間いずれも無相関に近い。これは、所得・労働時間の変化率に関しては、家計固有效果が無視できるほど小さいことを示している。すなわち、家計間で、所得・労働時間のトレンドに関しては、Heterogeneity は小さい、ということである。また、所得と労働時間の相関をみると、1998-99 年を除くと正であるが、その共分散は極めて小さいことがわかる。

表 9 はアメリカ合衆国の PSID を用いた Abowd and Card (1989) の Table IV である。日本の共分散行列と比較して直ちにわかることは、分散、共分散のいずれも日本の所得変化率は PSID の半分から 1/3 程度しかない、すなわち日本の所得変化率の家計間格差はアメリカ合衆国に比べて著しく小さいことがわかる¹⁹。一方、労働時間に関しては、日本のほう方が分散は小さいが、所得ほど顕著な差は見られない。一期先の変化率との相関が負であり、共分散の絶対値が分散の半分弱である点は、アメリカと日本は同じである。PSID では、二期先の変化率とも弱いながらも有意な負の相関がみられるが、これは日本では明確には出てこない。すなわち、アメリカの所得・労働変化率は、日本よりも長い系列相関を持つ。三期以上先との相関はほぼゼロなので、変化率のトレンドに家計間の差はほとんどみられない、という点は日米で共通である。所得と労働時間の相関は、同一時点では有意であり、0.5 程度と極め

¹⁷この表の導出の詳細は、Appendix 2 で説明している。

¹⁸First Stage において、年ダミーに回帰しているため、年による所得のレベルシフトはコントロールされている。

¹⁹Abowd and Card (1989) と本論文では First Stage の説明変数がやや異なり、本論文のほうが多くなっているが、日本の所得成長率の分散がアメリカの先行研究結果に比べて低いという結果は First Stage の specification には依存しない。何もコントロールしない対数所得変化率のみを用いても、所得成長率の分散はアメリカに比べて著しく低くなっている。

て大きな相関が観察されている。所得・労働時間いずれも分散は一定ではない点も日米共通である。

以上、まとめると、アメリカ家計データと日本家計データの共通点は (1) 所得・労働時間いずれも一期先と負の相関があり、その共分散の絶対値は分散の半分弱である、(2) 労働時間の分散は 0.1 程度、(3) 三期以上先のデータとは相関はなく、変化率のトレンドに関しては家計間格差は無視できるものである、(4) 所得・労働時間の分散は一定ではない。一方、相違点は (1) 所得の分散・共分散いずれも日本の値はアメリカの半分以下である、(2) 二期先の負の相関は日本では観察されない、(3) 所得と労働時間の相関はアメリカでは高いが、極めて小さい。

表 10 は日本の家計所得・労働の水準に関する共分散行列である。階差をとらない水準であるため、First Stage でコントロールした年齢や経験年数等を除外すれば、家計固有効果をコントロールしていない。ただちにわかるのは、所得・労働時間の両方とも長期にわたり相関が持続することである。これは、OLS と固定効果推定の差からわかるように、所得や労働のサンプル内での変動のほとんどが家計間の相違によるものであり、家計内の変動が小さいことを意味する。また、所得と労働時間の相関が正であるが、やはりかなり小さいこともわかる。

表 11 は日本の家計データをバランスデータとした所得・労働時間の変化率の共分散を計算したものである。アンバランスデータを用いた表 8 と比較すると、分散水準そのものが小さくなっていることがわかるが、その他の傾向はアンバランスデータとそれほど変わらない。

4.4 モデル分析

本節では、前節で計算した共分散行列を用い、極めて単純な Data Generating Process を仮定した上で、そのモデルの説明力を検証する。特に重視するのは、家計モデルで重要な役割を果たす所得の恒常的変動要因とデータに含まれる一時的変動の測定である。

家計所得 (経験年数や年齢、産業をコントロールした後) が Random Walk で記述できる場合、所得変動は恒常的なものであることを意味する。恒常所得仮説に従うと、所得変動が恒常である場合に限り家計は消費水準を変化させる。所得変動が一時的である場合は貯蓄のみが変動し、消費水準は変化しない。

Model 1: (Pure Random Walk)

$$g_{it} = g_{it-1} + \xi_{it} \quad (37)$$

$$h_{it} = h_{it-1} + \varphi_{it} \quad (38)$$

$$\xi_{it} : i.i.d.E(\xi_{it}) = 0, Var(\xi_{it}) = \sigma_{\xi}^2$$

$$\varphi_{it} : i.i.d.E(\varphi_{it}) = 0, Var(\varphi_{it}) = \sigma_{\varphi}^2$$

$$Cov(\xi_{it}, \varphi_{it}) = \sigma_{\varphi\xi}$$

このモデルに基づき、変化率の分散共分散行列を計算すると、対角成分および労働と所得の同時相関以外の成分は全てゼロとなる²⁰。推定するパラメータ分散成分のみで3つである。表8の共分散行列から明らかのように、このモデルの課す制約は厳しく、日本家計データとは合致しないことが予測される。

報告された家計所得が系列相関をもたない一時的変動、あるいは測定誤差 (Measurement Error) のみで構成されている場合、データの経済学的な意味は極めて乏しくなる。報告された家計所得の全てが一時的変動であるという極端な仮定をおくと、下記のようなモデルを考えることができる。

Model 2: (Pure Transitory Shock)²¹

$$g_{it} = \varepsilon_{it} \quad (39)$$

$$h_{it} = \omega_{it} \quad (40)$$

$$\varepsilon_{it} : i.i.d.E(\varepsilon_{it}) = 0, Var(\varepsilon_{it}) = \sigma_{\varepsilon}^2$$

$$\omega_{it} : i.i.d.E(\omega_{it}) = 0, Var(\omega_{it}) = \sigma_{\omega}^2$$

$$Cov(\varepsilon_{it}, \omega_{it}) = \sigma_{\varepsilon\omega}$$

このモデルに従い変化率の共分散行列を計算すると、Pure Random Walk モデルと異なり、一次の負の自己相関が発生し、その共分散の大きさは分散の半分となる。すなわち、

$$Var(\Delta g_{it}) = \sigma_{\Delta g}^2 = 2\sigma_{\varepsilon}^2 \quad (41)$$

$$Cov(\Delta g_{it}, \Delta g_{it-1}) = -\sigma_{\varepsilon}^2 \quad (42)$$

推定すべきパラメータは Model 1 と同じく3つであるが、一次の系列相関の存在を許しているところ、およびその符号が表8や表11と整合的であることから、Model 1 よりも良好な fit を期待することが出来る。しかしながら、一次の共分散の絶対値は分散の1/2よりも小さい傾向があり、このモデルでも十分でない可能性がある。

Model 1 と Model 2 の両者を結合させ、観察される所得データが恒常的要因と一時的要因の双方から構成されると仮定すると、下記のようなモデルを考えることが出来る。

²⁰具体的な Moment Conditions は Appendix 4 に記載している。

²¹本章では Abowd and Card (1989) に従い、white noise で記述される要素を Measurement Error と定義しているが、労働所得に占める一時的変動と解釈することも可能である。本章のアプローチでは両者を識別することができないため、本章では white noise の要素を Measurement Error と呼ぶことにする。

Model 3: (Mixture of Permanent and Transitory Shock)

$$g_{it}^* = g_{it-1}^* + \xi_{it} \quad (43)$$

$$h_{it}^* = h_{it-1}^* + \varphi_{it} \quad (44)$$

$$g_{it} = g_{it}^* + \varepsilon_{it} \quad (45)$$

$$h_{it} = h_{it}^* + \omega_{it} \quad (46)$$

$$\xi_{it} : i.i.d.E(\xi_{it}) = 0, Var(\xi_{it}) = \sigma_{\xi}^2$$

$$\varphi_{it} : i.i.d.E(\varphi_{it}) = 0, Var(\varphi_{it}) = \sigma_{\varphi}^2$$

$$Cov(\xi_{it}, \varphi_{it}) = \sigma_{\varphi\xi}$$

$$\varepsilon_{it} : i.i.d.E(\varepsilon_{it}) = 0, Var(\varepsilon_{it}) = \sigma_{\varepsilon}^2$$

$$\omega_{it} : i.i.d.E(\omega_{it}) = 0, Var(\omega_{it}) = \sigma_{\omega}^2$$

$$Cov(\varepsilon_{it}, \omega_{it}) = \sigma_{\varepsilon\omega}$$

このモデルの推定すべきパラメーターは6つであり、今までのモデルよりも自由度が低下する。分散と一時の自己相関との関係は Model 2 と異なり下記のようなになる。

$$Var(\Delta g_{it}) = \sigma_{\Delta g}^2 = 2\sigma_{\varepsilon}^2 + \sigma_{\xi}^2 \quad (47)$$

$$Cov(\Delta g_{it}, \Delta g_{it-1}) = -\sigma_{\varepsilon}^2 \quad (48)$$

(47) 式と (48) 式の両式から、恒常ショックと一時的ショックを識別することができる。

最後に、一時的ショックの効果にある程度の継続性があると仮定した MA モデルを考える。Abowd and Card (1989) はアメリカの様々なパネルデータを用い、Bivariate の MA(2) により、所得や労働時間をある程度描写できると主張している。

Model 4: (General MA(2))

$$g_{it} = \varepsilon_{it} + \rho_1 \varepsilon_{it-1} + \rho_2 \varepsilon_{it-2} \quad (49)$$

$$h_{it} = \omega_{it} + \gamma_1 \omega_{it-1} + \gamma_2 \omega_{it-2} \quad (50)$$

$$\varepsilon_{it} : i.i.d.E(\varepsilon_{it}) = 0, Var(\varepsilon_{it}) = \sigma_{\varepsilon}^2$$

$$\omega_{it} : i.i.d.E(\omega_{it}) = 0, Var(\omega_{it}) = \sigma_{\omega}^2$$

$$Cov(\varepsilon_{it}, \omega_{it}) = \sigma_{\varepsilon\omega}$$

$$\Delta g_{it} = \varepsilon_{it} - (1 - \rho_1) \varepsilon_{it-1} - (\rho_1 - \rho_2) \varepsilon_{it-2} - \rho_2 \varepsilon_{it-3} \quad (51)$$

$$\Delta h_{it} = \omega_{it} - (1 - \gamma_1)\omega_{it-1} - (\gamma_1 - \gamma_2)\omega_{it-2} - \gamma_2\omega_{it-3} \quad (52)$$

このモデルで推定するパラメータは7つであり、Model 3 よりも多い。しかしながら、Model 3 と Model 4 は入れ子にはなっておらず、Model 4 の方の fit が良いかどうかは先見的には明らかではない。

表 12 は各モデルの推定結果をまとめたものである。なお、 $\chi^2(m2=0)$ とは、各モデルから導かれる Moment Conditions がゼロであるとき、その成分がゼロであるか否かを検定する統計量である。具体的には Abowd and Card (1989) の Appendix に書かれている手法に従い、ゼロ成分のみを抽出し、それに対応する Weight Matrix を作成し計算している。多くの 0 成分を含むモデルではこの統計量が計算されていないが、それは Weight Matrix に含まれる誤差のためであると考えられる。

残差二乗和の値はバランスデータとアンバランスデータで全く異なるが、4 つのモデルはいずれも同一の共分散行列を用いているため、相対的な大きさを比較することでモデルのデータへの fit の程度を知ることができる。Model 1 と Model 2 を比較すると、推定するパラメータ数はどちらも同じであるが、Model 2 の方が残差二乗和ははるかに小さい。また Model 3 と Model 4 を比較すると、Model 4 の方が推定パラメータは多いにもかかわらず、Model 3 の方が二乗和は小さい。したがって、日本のデータに関しては、一時的ショックに系列相関を仮定する MA(2) よりも、Random Walk と White Noise の一時的ショックを仮定する Model 3 の方がよりデータと適合することがわかる。

次に $\chi^2(m2=0)$ をみると Model 4 で二次の Moment がゼロとなる共分散成分は 50 個程度しかないにもかかわらず、それがゼロであることは棄却されてしまう。無論、Model 1 および Model 3 も同様に棄却されるが、Abowd and Card (1989) では MA(2) によるゼロ制約は棄却できないことが指摘されており、日本の家計データはアメリカ合衆国のデータのような単純な MA 表現では記述できないことがわかる。

各分散パラメータの標準誤差はきわめて小さいが、Appendix 3 で議論しているように、巨大な共分散行列の推定にともなう誤差のため、この結果には留保が必要である。最後に、Model 3 の結果をみると、 σ_ε と σ_ξ の大きさはほぼ同じである。これは、所得データに含まれる変動、正確には一階差分により残される変動は、2/3 程度が Measurement Error、あるいは一時的変動であることを示している。一方労働時間に関しては 40% 程度が一時的ショックあるいは一時的変動となっている。

4.5 結論

本章は、10 年にわたる日本の家計パネルデータを用い、家計勤労所得と労働時間の共分散構造を分析し、その背後の確率過程に関する考察を行った。欧米では長い歴史のある研究分野であるが、筆者の知る限り、日本に関して

は初の研究である。また、その結果をアメリカの家計データを用いた Abowd and Card (1989) の結果と比較した。その結果、日米では系列相関の長さや所得変化率の分散と一期の自己相関との関係に関しては共通の性質を有するが、日本の方が所得分散は小さく、労働時間と所得間の相関も観察されない点が異なることが明らかになった。また、モデル分析によると、日本の家計所得のデータ変動のうち、2/3 が Measurement Error、あるいは一時的ショックであることがわかった。

本章で残された課題は多いが、特に (1) Bootstrap を用いた Weight Matrix の推定、(2) 推計された所得分散を持つ、家計消費モデルへの含意、(3) 日米での所得・労働時間の相関の違いの理由、(4) データの共分散構造を再現するような、より fit の良いモデルの考察、の四点は特に重要な課題であると思われる。

5 Appendix 1 (異常値の取り扱い)

年間勤労所得と年間労働時間について、異常値として推定から除外する基準を以下のように定めた。

(1) 年間勤労所得

年間勤労所得と年間労働時間から 1 時間あたりの賃金を算出し、これが最低賃金を下回るサンプルは取り除いた。

(2) 年間労働時間

関連項目から年間労働時間を算出したため、測定誤差が大きくなる可能性は排除できない。よって、年間労働時間を以下の 2 つの方法で算出し、算出された値の差が極端に大きい世帯（ここでは、10 倍以上とした）については取り除くという方針をとった。一つ目の算出方法は、脚注 4 で述べた方法である。二つ目の方法は、(1) 1 週間の労働時間、(2) 1 週間の休日日数、(3) 年間労働日数、の設問から、 $((1) \div (7 - (2))) \times (3)$ で求めた。また、総労働時間がサービス残業時間を超える場合や、1 日あたりの労働時間が 24 時間を超えるサンプルについても取り除いた。

6 Appendix 2 (共分散行列の計算)

共分散行列の導出は、基本的には Abowd and Card (1989) に沿って計算されている。唯一異なる点は、Abowd and Card (1989) ではバランスパネルを用いているが、本章ではアンバランスパネルデータを用いていることである。そのため、共分散成分毎にサンプルサイズが異なることになり、プログラミングでは欠損値処理を各段階で行う必要がある。

観測される、年齢や経験をコントロールした所得変化率を Δg_{it} 、同労働時間変化率を Δh_{it} とし、

$$y_i = \begin{pmatrix} \Delta g_{i1} \\ \vdots \\ \Delta g_{iT} \\ \Delta h_{i1} \\ \vdots \\ \Delta h_{iT} \end{pmatrix} \quad (53)$$

と y_i を定義する。 T はデータセットが含む変化率の観測期間であり、 y_i は最大で $2T$ の要素を持つ。ここでのサンプルでは家計 i によりサンプル期間に差があるので、 y_i の次元は家計により異なることになる。 y_i の要素は回帰の残差であることに注意すると共分散行列は、バランスパネルであれば家計数 N を用いて $C = (1/N) \sum_i (y_i y_i')$ と計算することが可能である。アンバランスの場合は、各成分毎に外積を計算し、外積を計算したサンプルサイズで割ることで共分散行列を得ることが出来る。共分散行列 C の要素を縦に並べたベクトルを m とする。ただし、共分散行列は対称であるから、上(下)三角部分のみを対象にし、 $2T(2T+1)/2$ の要素が m に含まれることになる。各家計ごとに y_i の外積を計算し、その上(下)三角成分を縦に並べたものを m_i とすると、バランスパネルの場合は $m = (1/N) \sum_i m_i$ となり、アンバランスの場合は、各成分毎に観測数を計算し、それで割ることで m を得ることが出来る。分散共分散行列成分 m の共分散行列は、バランスデータの場合は

$$V = \left(\frac{1}{N} \right) \sum_i (m_i - m)(m_i - m)' \quad (54)$$

で計算可能であり、アンバランスの場合は、やはり、各成分毎に観測数で割ることになる。本章で用いている観測期間は所得が 9 期間、労働が 8 期間であり、共分散行列 C は 17×17 行列であり、上(下)三角成分は 153 存在する。したがって、共分散成分の共分散行列 V は 153×153 行列となる。

ここで、Abowd and Card (1989) が指摘しているような問題が生じる。この大きな共分散行列の精度が著しく低い、という問題である。実際、Bootstrap で計算されたものと、推定された V を比較すると、有効桁数で 3 桁程度から大きく異なってくる。Altonji and Segal (1996) は Bootstrap で推計されたものを使うべきであると主張している。また、Horowitz (1998) も Bootstrap を提唱している。

この巨大な共分散行列は、共分散成分 m の標準誤差の計算および、その後のモデル推定の際の GMM ウェイト、および仮説検定の際に重要な役割を果たすことになる。

7 Appendix 3 (モデル検証)

ここでは、モデルにより導出される共分散成分がパラメター b^0 の関数 $f(b^0)$ として記述できると仮定する。すなわち

$$m = f(b^0) \quad (55)$$

という理論的な関係を考える。これは理論モデルが導き出す Moment Condition であるため、この関係式を用いる推定は一種の Moment 法と考えることが出来る²²。Amemiya (1974) の非線形操作変数推定法、および Hansen (1982) や Ogaki (1993) の GMM 等に従うと、 b^0 を推計する効率的な手法は、

$$(m - f(b^0))' V^{-1} (m - f(b^0)) \quad (56)$$

を最小にするようなパラメターを推定することである。バランスパネルの場合は

$$N (m - f(b^0))' V^{-1} (m - f(b^0)) \quad (57)$$

アンバランスパネルの場合は、各 Moment の計算に使用した観測数を N の代わりに乗じたものが χ 二乗分布に従う。なお、その際の自由度は Moment Condition の数から $f(b^0)$ の Jacobian を b の真の値で評価したものの階数を引いたものとなる。しかしながら、Abowd and Card (1989) およびその他の多くの文献が指摘するように、 V の推計値の誤差が大きく、その逆行列をウェイトにもつ上記の二次形式の値は、実際に最小化させるときに必要な凹性等を満たさないケースが多く、推計は極めて困難となる。ここでは、Abowd and Card (1989) に従い、Optimal Weight ではなく、単位行列をウェイトとする Equal Weighted Minimum Distance (EQWMD) Estimator を用いた。すなわち、最小化させるのは

$$(m - f(b^0))' (m - f(b^0)) \quad (58)$$

である²³。なお、推定パラメター b の共分散行列には Delta Method を用いた。これは漸近的に成立する分布であり、今回のように Moment Condition の数が 100 個程度の場合は、小標本バイアスが深刻な問題である可能性が存在する。したがって、 b やモデル評価の際の標準誤差に関しては留保が必要である。

²²Abowd and Card (1989) は Moment 法、あるいは GMM という言葉を用いていないが、Optimal Weight を用いる場合は Hansen (1982) による GMM となら変わることはないように思われる。

²³無論、Altonji and Segal (1996) や Horowitz (1996) が提唱している Bootstrap によるウェイト行列の推定も考慮せねばならないが、先行研究の乏しい日本では、まず最も単純な推計法を試すことにもそれなりの価値があるはずである。アンバランスデータから Bootstrap サンプリングを行いウェイト行列を推計することは将来の課題としたい。

8 Appendix 4 (Moment Conditions)

各モデルの推計で用いた Moment Conditions は全て Second Moments であり、各モデルに記載されていない Moment は全てゼロになっている。共分散成分の数は各モデルで共通であるため、記載されていない Moment Condition が少ないモデルは、その分、Second Moment がゼロであるという Moment Condition が多いことを意味することに注意せよ。

Model 1: Pure Random Walk

$$Var(\Delta g_{it}) = \sigma_{\Delta g}^2 = \sigma_{\xi}^2 \quad (59)$$

$$Var(\Delta h_{it}) = \sigma_{\Delta h}^2 = \sigma_{\varphi}^2 \quad (60)$$

$$Cov(\Delta g_{it}, h_{it}) = \sigma_{\varphi\xi} \quad (61)$$

Model 2: Pure Transitory Shock

$$Var(\Delta g_{it}) = \sigma_{\Delta g}^2 = 2\sigma_{\varepsilon}^2 \quad (62)$$

$$Var(\Delta h_{it}) = \sigma_{\Delta h}^2 = 2\sigma_{\omega}^2 \quad (63)$$

$$Cov(\Delta g_{it}, \Delta g_{it-1}) = -\sigma_{\varepsilon}^2 \quad (64)$$

$$Cov(\Delta h_{it}, \Delta h_{it-1}) = -\sigma_{\omega}^2 \quad (65)$$

$$Cov(\Delta h_{it}, \Delta g_{it}) = 2\sigma_{\varepsilon\omega} \quad (66)$$

$$Cov(\Delta h_{it}, \Delta g_{it-1}) = -\sigma_{\varepsilon\omega} \quad (67)$$

$$Cov(\Delta g_{it}, \Delta h_{it-1}) = -\sigma_{\varepsilon\omega} \quad (68)$$

Model 3: Mixture of Permanent and Transitory Shock

$$Var(\Delta g_{it}) = \sigma_{\Delta g}^2 = 2\sigma_{\varepsilon}^2 + \sigma_{\xi}^2 \quad (69)$$

$$Var(\Delta h_{it}) = \sigma_{\Delta h}^2 = 2\sigma_{\omega}^2 + \sigma_{\varphi}^2 \quad (70)$$

$$Cov(\Delta g_{it}, \Delta g_{it-1}) = -\sigma_{\varepsilon}^2 \quad (71)$$

$$Cov(\Delta h_{it}, \Delta h_{it-1}) = -\sigma_{\omega}^2 \quad (72)$$

$$Cov(\Delta h_{it}, \Delta g_{it}) = 2\sigma_{\varepsilon\omega} + \sigma_{\varphi\xi} \quad (73)$$

$$Cov(\Delta h_{it}, \Delta g_{it-1}) = -\sigma_{\varepsilon\omega} \quad (74)$$

$$Cov(\Delta g_{it}, \Delta h_{it-1}) = -\sigma_{\varepsilon\omega} \quad (75)$$

Model 4: Bivariate MA(2)

$$Var(\Delta g_{it}) = \sigma_{\Delta g}^2 = [1 + (1 - \rho_1)^2 + (\rho_1 - \rho_2)^2 + \rho_2^2] \sigma_{\varepsilon}^2 \quad (76)$$

$$Var(\Delta h_{it}) = \sigma_{\Delta h}^2 = [1 + (1 - \gamma_1)^2 + (\gamma_1 - \gamma_2)^2 + \gamma_2^2] \sigma_{\omega}^2 \quad (77)$$

$$Cov(\Delta g_{it}, \Delta g_{it-1}) = [-(1 - \rho_1) + (1 - \rho_1)(\rho_1 - \rho_2) + \rho_2(\rho_1 - \rho_2)] \sigma_{\varepsilon}^2 \quad (78)$$

$$Cov(\Delta g_{it}, \Delta g_{it-2}) = [-(\rho_1 - \rho_2) + \rho_2(1 - \rho_1)] \sigma_{\varepsilon}^2 \quad (79)$$

$$Cov(\Delta g_{it}, \Delta g_{it-3}) = -\rho_2 \sigma_{\varepsilon}^2 \quad (80)$$

$$Cov(\Delta h_{it}, \Delta h_{it-1}) = [-(1 - \gamma_1) + (1 - \gamma_1)(\gamma_1 - \gamma_2) + \gamma_2(\gamma_1 - \gamma_2)] \sigma_{\omega}^2 \quad (81)$$

$$Cov(\Delta h_{it}, \Delta h_{it-2}) = [-(\gamma_1 - \gamma_2) + \gamma_2(1 - \gamma_1)] \sigma_{\omega}^2 \quad (82)$$

$$Cov(\Delta h_{it}, \Delta h_{it-3}) = -\gamma_2 \sigma_{\omega}^2 \quad (83)$$

$$Cov(\Delta h_{it}, \Delta g_{it}) = [1 + (1 - \rho_1)(1 - \gamma_1) + (\rho_1 - \rho_2)(\gamma_1 - \gamma_2) + \rho_2(1 - \rho_1)] \sigma_{\varepsilon\omega} \quad (84)$$

$$Cov(\Delta g_{it}, \Delta h_{it-1}) = [-(1 - \rho_1) + (\rho_1 - \rho_2)(1 - \gamma_1) + \rho_2(\gamma_1 - \gamma_2)] \sigma_{\varepsilon\omega} \quad (85)$$

$$Cov(\Delta g_{it}, \Delta h_{it-2}) = [-(\rho_1 - \rho_2) + \rho_2(1 - \gamma_1)] \sigma_{\varepsilon\omega} \quad (86)$$

$$Cov(\Delta g_{it}, \Delta h_{it-3}) = -\rho_2 \sigma_{\varepsilon\omega} \quad (87)$$

$$Cov(\Delta h_{it}, \Delta g_{it-1}) = [-(1 - \gamma_1) + (\gamma_1 - \gamma_2)(1 - \rho_1) + \gamma_2(\rho_1 - \rho_2)] \sigma_{\varepsilon\omega} \quad (88)$$

$$Cov(\Delta g_{it}, \Delta h_{it-2}) = [-(\gamma_1 - \gamma_2) + \gamma_2(1 - \rho_1)] \sigma_{\varepsilon\omega} \quad (89)$$

$$Cov(\Delta g_{it}, \Delta h_{it-3}) = -\gamma_2 \sigma_{\varepsilon\omega} \quad (90)$$