Intergenerational Correlations of Skills

Emiko Usui
Nagoya University
and
Tsunao Okumura
Yokohama National University

Questions

Individual's skills are multidimensional

- Cognitive skills (reasoning, mathematics, verbal,...)
- People skills (talking, dealing-with-people skills, prefer to work for the good of people,...)
- Motor skills (spatial and form perception, finger dexterity,...)

1. **Social skills correlate across generations?**
   - studies find that social skills affect labor market outcomes
   - Kuhn and Weinberger, 2005; Borghans et al., 2006; Goldsmith et al., 1998
2. **Overall skills correlate across generations?**
3. **Children benefit from choosing similar occupations to their parents?**

Problems

Often we lack data on parents

- use the Dictionary of Occupational Titles (DOT) to proxy for the various skill dimensions of parents

How to measure the closeness of multidimensional parent-child skills

- look at the *cosine* of the parent-child skill vectors

Main Results

Theory:

Model of intergenerational skill transfer

- technical skill and social skill
- cost savings for parent and child who choose similar skills

Empirics:

1. Social skills correlate across generations
   - father-son pairs; mother-daughter pairs.
2. Overall skills positively correlate between father-son pairs
3. Wage premium is paid to children who work in similar occupations to their fathers
   - wage premium for white men; wage penalty for black men
   - blacks hurt by father's poor skill acquisition

Theoretical Model of Intergenerational Skill Transfer

Becker and Tomes (1976): univariate Human Capital transfer
Laband and Lentz (1983): occupational following

⇩

Multidimensional skill transfer

Model

Utility of the family: $\max_{\{T_p, S_p, T^*, S^*, x\}} \beta u_P(con_P) + (1-\beta) u_C(con_C)$

- con_p: consumption
- β: weight of parent's utility in the family welfare

Two dimension: technical and social skill

- parent invests in own skill: (T_p, S_p)
- child's skill: $(T_c, S_c) = (xT_p + T^*, xS_p + S^*)$
  - impart "x" portion of parent skill at zero cost
  - invest in skill vector (T^*, S^*)

Cost of acquiring skills: $\gamma_p(T_p, S_p)$ and $\gamma_C(T^*, S^*)$; concave

Wage: $\omega_p(T_p, S_p)$ and $\omega_C(T_C, S_C)$; convex

Parent's Decision

Altruistic parent maximizes the utility of the family:

$$\max_{\{T_p, S_p, T^*, S^*, x\}} \beta u_p(con_p) + (1-\beta) u_C(con_C)$$

subject to: $(T_C, S_C) = (xT_p + T^*, xS_p + S^*)$

$con_p + R \le \omega_p(T_p, S_p) - \gamma_p(T_p, S_p)$

$con_C \le R + \omega_C(T_C, S_C) - \gamma_C(T^*, S^*)$

– R is transfer of savings

FOC:

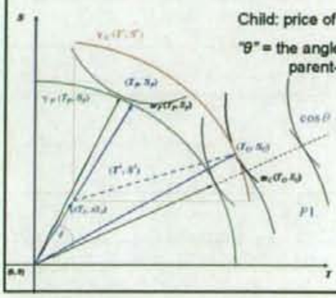
$$\frac{\partial \omega_p(T_p, S_p)}{\partial T_p} + x\frac{\partial \omega_C(xT_p + T^*, xS_p + S^*)}{\partial T_C} = \frac{\partial \gamma_p(T_p, S_p)}{\partial T_p}$$

$$\frac{\partial \omega_p(T_p, S_p)}{\partial S_p} + x\frac{\partial \omega_C(xT_p + T^*, xS_p + S^*)}{\partial S_C} = \frac{\partial \gamma_p(T_p, S_p)}{\partial S_p}$$

$$\frac{\partial \omega_C(xT_p + T^*, xS_p + S^*)}{\partial T_C} = \frac{\partial \gamma_C(T^*, S^*)}{\partial T^*}$$

$$\frac{\partial \omega_C(xT_p + T^*, xS_p + S^*)}{\partial S_C} = \frac{\partial \gamma_C(T^*, S^*)}{\partial S^*}$$

Illustration



Child: price of T-skill is greater

"θ" = the angle between the parent-child skill vector

$$\cos\theta = \frac{T_pT_C + S_pS_C}{\sqrt{T_p^2 + S_p^2}\sqrt{T_C^2 + S_C^2}}$$

P1. $d\cos\theta / dx \ge 0$

Wage & Cost Functions

To obtain explicit equilibrium of skill vectors, let:

$\omega_p(T_p, S_p) = w_{1p}T_p + w_2S_p$

$\omega_C(T_C, S_C) = w_{1C}T_C + w_2S_C$

$\gamma_p(T_p, S_p) = a_pT_p^2 + bS_p^2$

$\gamma_C(T^*, S^*) = a_CT^{*2} + bS^{*2}$

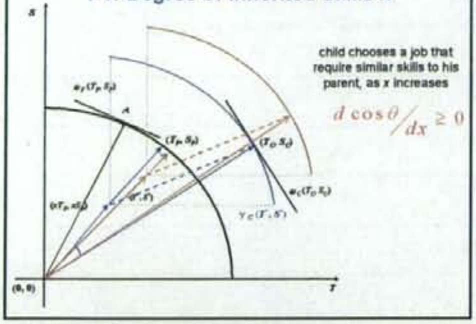
Solution: $(T_p, S_p) = \frac{1}{2}\left(\frac{w_{1p}}{a_p}\left(1 + x\frac{w_{1C}}{w_{1p}}\right), \frac{w_2}{b}(1+x)\right)$

$(T_C, S_C) = \frac{1}{2}\left(\frac{w_{1p}}{a_p}x\left(1 + x\frac{w_{1C}}{w_{1p}}\right) + \frac{w_{1C}}{a_C}, \frac{w_2}{b}(1+x) + \frac{w_2}{b}\right)$

Let "θ" be the angle between the parent-child skill vector

$$\cos\theta = \frac{T_pT_C + S_pS_C}{\sqrt{T_p^2 + S_p^2}\sqrt{T_C^2 + S_C^2}}$$

P1: Degree of Inherited Skills x



P2: Cost of acquiring T skills

Suppose $a_P = a_C = a$.

$d\cos\theta / da < 0$ & $d\omega_C(T_C, S_C)/da < 0$ if

$$\left[x\left(\frac{w_{1C}}{w_1}\right)+\left(\frac{w_{1P}}{w_2}\right)\right]\left[x^2\left(\frac{w_{1C}}{w_2}\right)+x\left(\frac{w_{1P}}{w_2}\right)+\left(\frac{w_{1C}}{w_2}\right)\right]>\left(\frac{a}{b}\right)^2(x+1)(x^2+x+1)$$

- A smaller correlation in parent-child skills, if
  a *increase* in the cost of acquiring the T skills that pay a *higher wage* and involve a *lower cost* than the S skills

$d\cos\theta / da \geq 0$ & $d\omega_C(T_C, S_C)/da < 0$ if

$$\left[x\left(\frac{w_{1C}}{w_2}\right)+\left(\frac{w_{1P}}{w_2}\right)\right]\left[x^2\left(\frac{w_{1C}}{w_2}\right)+x\left(\frac{w_{1P}}{w_2}\right)+\left(\frac{w_{1C}}{w_2}\right)\right]\leq\left(\frac{a}{b}\right)^2(x+1)(x^2+x+1)$$

- A smaller correlation in parent-child skills, if
  a *decrease* in the cost of acquiring the T skills that pay a *lower wage* and involve a *higher cost* than the S skills

Results hold when, $w_{1P} \neq w_{1C}$

P3: Cost of acquiring T skills

Suppose $a_P \neq a_C$

If $\frac{w_{1P}}{w_{1C}} + x < (1+x)\frac{a_P}{a_C}$,

$d\cos\theta / da_P < 0 \quad d\cos\theta / da_C > 0$

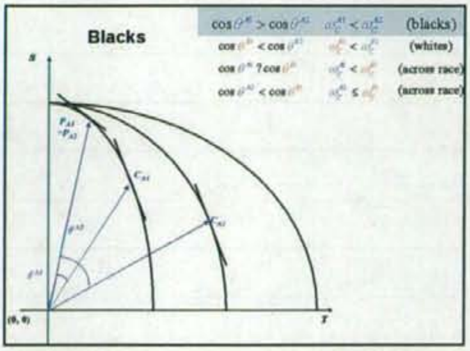
When the T skill investment is less rewarded and more costly for the parent than the child, the parent acquires less T skills (and more S skills) than the child does. In this situation, if the parent's cost of the T skill decreases, he would increase the T skill investment (and decrease the S skill investment), and thus, the parent's skill vector gets closer to the child's vector.

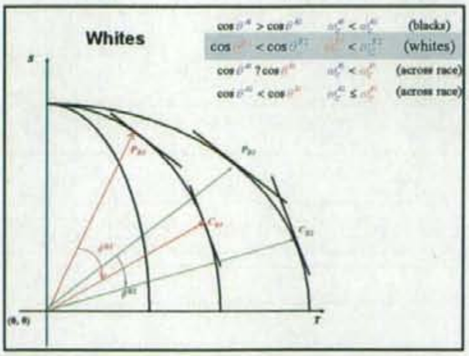
$d\omega_C(T_C, S_C)/da_P \leq 0, \; d\omega_C(T_C, S_C)/da_C \leq 0 \; (=0 \text{ if } x=0)$

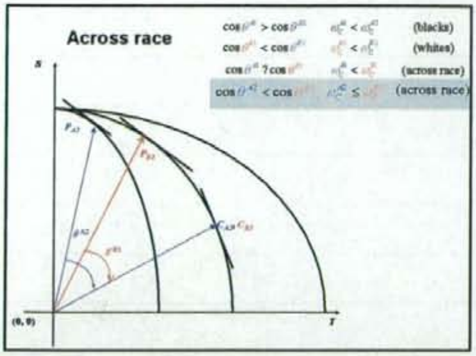
Explaining the Racial Differences

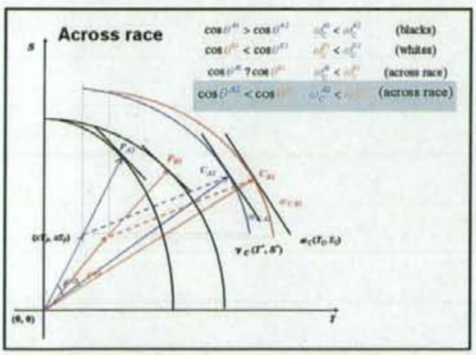
1. Wages to T skill for child are greater than those for parent. Relative wages to T-skill for parent are higher than relative cost. $w_{1C} > w_{1P}$; $\frac{w_{1P}}{w_2} > \frac{a_P^{B1}}{b}$
2. Consider Four Family Types: A1, A2, B1, and B2
   - A1, A2: blacks
     - Parents: Pay higher cost to acquire *T* skill than B1 and B2
     - Children: A2 pays less cost than A1
   - B1, B2: whites
     - Parents and children both pay lower cost to acquire T skill than A1 and A2

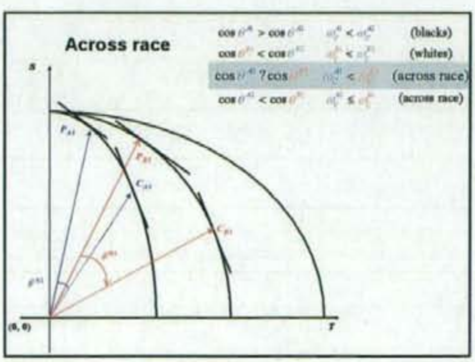
$$a_P^{A1} = a_P^{A2} > a_P^{B1} = a_C^{B1} > a_P^{B2} = a_C^{B2}$$
$$a_P^{A1} \geq a_C^{A1} > a_C^{A2} = a_C^{B1}$$






Across race
(blacks)
(whites)
(across race)
(across race)
S
T
(0, 0)


Across race
(blacks)
(whites)
(across race)
(across race)
S
T
(0, 0)


Across race
(blacks)
(whites)
(across race)
(across race)
S
T
(0, 0)

Summary of Implication

Within Racial Group

$\cos\theta^{B1} > \cos\theta^{B2}$ $\quad w_c^{B1} < w_c^{B2}$ (within blacks)
wage penalty when $\cos\theta \uparrow$

$\cos\theta^{W1} < \cos\theta^{W2}$ $\quad w_c^{W1} < w_c^{W2}$ (within whites)
wage premium when $\cos\theta \uparrow$

Across Racial Groups

$\cos\theta^{B1} \; ? \; \cos\theta^{W1} < \cos\theta^{W2}$ $\quad w_c^{B1} < w_c^{W1}$ (across race)

$\cos\theta^{B2} < \cos\theta^{W1} < \cos\theta^{W2}$ $\quad w_c^{B2} \le w_c^{W1}$ (across race)

↓

$\text{Avg}(\cos\theta^{B1} \,\&\, \cos\theta^{B2}) < \text{Avg}(\cos\theta^{W1} \,\&\, \cos\theta^{W2})$ $\quad \text{Avg}(w_c^{B1} \,\&\, w_c^{B2}) < \text{Avg}(w_c^{W1} \,\&\, w_c^{W2})$

Black cosine < White cosine $\quad$ Black Wage < White Wage

Summary of Implication

1. Greater skill correlation when parent can impart more skills at zero cost to the child: $d\cos\theta / dx \ge 0$
2. Racial differences in skill correlation and wages
   - T-skill pays higher wage
   - Black parents incur larger cost in acquiring T-skill than white parents

Blacks ⇒ cos θ is smaller (greater angle)
child obtains higher wage when cos θ is smaller
(child acquires more T skill)

Whites ⇒ child obtains higher wage when cos θ is larger

Empirical Analysis

Dictionary of Occupational Titles (DOT)

- Conducted by Department of Labor
  - used for job matching applications; career guidance
  - fourth edition (1977)
  - revised fourth edition (1991)
- Classify into four broad skill categories
  - cognitive skills
  - people skills
  - motor skills
  - physical demands

Cognitive Skills in the DOT

- general reasoning ability
- mathematical ability
- language ability
- general intelligence aptitude
- verbal aptitude
- numerical aptitude
- clerical perception
- routine/creative activity
- complexity of function in relation to data

Detailed description in Table 1

People Skills in the DOT

- adaptability to dealing with people beyond giving and receiving instructions
- adaptability to situations involving the interpretation of feelings, ideas or facts in terms of personal viewpoint
- adaptability to influencing people in their opinions, attitudes or judgments about ideas or things
- a preference for activities concerned with the communication of data
- a preference for activities involving business contact with people
- a preference for working for the presumed good of people
- job requires talking
- complexity of function in relation to people

Motor Skills in the DOT

- finger dexterity
- motor coordination
- manual dexterity
- eye-hand-foot coordination
- spatial perception
- form perception
- color discrimination
- adaptability to situations requiring the attainment of setting limits, tolerances or standards
- complexity of function in relation to things

Physical Demands in the DOT

- strength requirement of jobs
- climbing
- stooping

National Longitudinal Survey of Youth

U.S. panel data on individuals
- age 14-21 when first surveyed in 1979

<u>Parent's information:</u>
- Education
- Occupation
  - Match to the DOT and proxy for parent's skills

<u>Children's information:</u>
- Cognitive skill
  - Armed Force Qualification Test (AFQT)
- Social skill: Sociability in early adulthood / sociability at age 6
  - Would you describe yourself as: extremely shy, somewhat shy, somewhat outgoing, or extremely outgoing?
- Club participation in high school
- Occupation

National Longitudinal Survey of Youth

<u>Children of the NLSY female respondents:</u>
- Social skills
  - sociability scores (age 2-6)
- Club participation

Intergenerational Link in Social Skills

Estimate the effect of the parent's DOT people skills on the child's sociability

- but DOT people skills complement DOT cognitive skills

Approach 1: Use Factor Analysis

- Perform factor analysis on the parent's DOT skills
  - Obtain factor scores
  - Label them as: cognitive factor, people factor, motor factor, physical demands factor
- Estimate the effect of the parent's people factor on the child's sociability
  - child's sociability = α_1 cog_fac + α_2 mot_fac + α_3 peo_fac + α_4 phys_fac + controls for parent's edu, child's edu & AFQT, family income

Table 6: Intergenerational Link in Social Skills Approach 1: Use Factor Analysis

	Father-Son		Father-Daughter		Mother-Son		Mother-Daughter	
	Sociability in Early Adulthood	Sociability at Age 6	Sociability in Early Adulthood	Sociability at Age 6	Sociability in Early Adulthood	Sociability at Age 6	Sociability in Early Adulthood	Sociability at Age 6
Parent's Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Factor 1: Cognitive Skill	0.015	0.015	0.050	0.074	0.098	0.107	0.102	0.074
	(0.031)	(0.030)	(0.030)	(0.029)	(0.036)	(0.036)	(0.034)	(0.033)
Factor 2: Motor Skill	0.003	0.016	-0.061	-0.075	0.012	0.046	0.016	0.047
	(0.025)	(0.024)	(0.024)	(0.023)	(0.030)	(0.029)	(0.028)	(0.027)
Factor 3: People Skill 1	0.077	0.080	0.022	0.036	0.058	0.021	0.023	0.028
	(0.025)	(0.024)	(0.024)	(0.023)	(0.030)	(0.029)	(0.029)	(0.028)
Factor 4: Physical Demands	-0.021	0.013	-0.020	-0.018	0.004	-0.028	-0.0001	-0.027
	(0.025)	(0.024)	(0.025)	(0.024)	(0.030)	(0.029)	(0.028)	(0.027)
Factor 5: People Skill 2	0.014	-0.004	-0.014	0.005	-0.016	-0.002	-0.024	-0.006
	(0.025)	(0.024)	(0.027)	(0.025)	(0.030)	(0.028)	(0.029)	(0.028)
Parent's Education	0.001	0.027	0.009	0.009	-0.013	-0.010	0.003	-0.002
	(0.009)	(0.008)	(0.009)	(0.008)	(0.013)	(0.012)	(0.013)	(0.012)
Pseudo-R^2	0.007	0.008	0.010	0.012	0.011	0.010	0.009	0.008
N	2060	2054	2186	2184	1440	1438	1567	1565

Intergenerational Link in Social Skills

Approach 2: Construct "Adjusted" People Skills

Obtain "adjusted" people skills that do not overlap with other DOT skills

- Project the DOT people skill onto other DOT skills, and obtain the residual error (="adjusted" people skill)
- Perform factor analysis on the parent's "adjusted" people skills
  - obtain factor scores: factor1, factor2
- Estimate the effect of the parent's people factors on the child's sociability
  - child's social skill = α_1 *people factor1* + α_2 *people factor2* + controls for parent's edu, child's edu & AFQT, family income

Table 7: Intergenerational Link in Social Skills Approach 2: Construct "Adjusted" People Skills

Father's DOT people skills are adjusted by DOT cognitive, motor skills and physical demands

	Son's Sociability in Early Adulthood		Son's Sociability at Age 6		Daughter's Sociability in Early Adulthood		Daughter's Sociability at Age 6	
		Adjusted		Adjusted		Adjusted		Adjusted
Father's var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
People Factor 1	**0.069**	**0.052**	0.032	0.009	**0.069**	-0.001	**0.094**	-0.007
	(0.028)	**(0.025)**	(0.027)	(0.024)	**(0.027)**	(0.024)	**(0.026)**	(0.023)
People Factor 2	-0.001	0.023	-0.017	0.016	-0.025	-0.025	-0.007	-0.030
	(0.025)	(0.025)	(0.024)	(0.024)	(0.027)	(0.026)	(0.026)	(0.025)
Education	-0.0001	0.000	**0.027**	**0.027**	0.012	0.009	0.013	0.011
	(0.008)	(0.009)	**(0.008)**	**(0.009)**	(0.008)	(0.009)	(0.008)	(0.009)
Verbal		0.547		0.664		-0.301		-0.312
		(0.385)		(0.372)		(0.387)		(0.374)
R^2	0.006	0.010	0.008	0.012	0.010	0.014	0.010	0.015
N	2061	2061	2055	2055	2187	2187	2185	2185

Table 5: Effect of Mother's Sociability on Child's Sociability:

NLSY females and their children

Independent Variable (Separate Regression)	*Dependent Variable* Daughter: Sociability Score (2-6 yrs. old)	Daughter: Whether Belong to a Club (10-14 yrs. old)	Son: Sociability Score (2-6 yrs. old)	Son: Whether Belong to a Club (10-14 yrs. old)
	(1)	(2)	(3)	(4)
Mother: Sociability in Early Adulthood	0.036 (0.026)	**0.052** **(0.018)**	0.039 (0.027)	0.010 (0.016)
Mother: Sociability at Age 6	**0.048** **(0.019)**	0.021 (0.013)	0.021 (0.019)	0.011 (0.012)
Mother: Number of Clubs Participated in HS	**0.031** **(0.013)**	**0.042** **0.010**	0.013 (0.013)	**0.029** **(0.009)**

Summary: Intergenerational Link in Social Skills

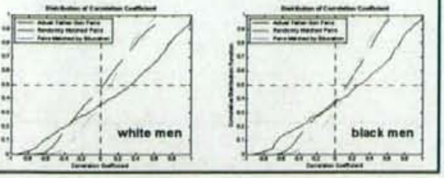
- Approach 1: Use factor analysis
  - **father-son** — **sociability in early adulthood**; sociability at age 6
- Approach 2: Construct "adjusted" people skills
  - **father-son** — **sociability in early adulthood**
- Link in sociability between NLSY female and her child
  - **mother-daughters**

Intergenerational Links in Overall Skills

Using the parent-child DOT skills compute the cosine:

Distribution of the correlation coefficient

1. Actual father-son skill pairs
2. Randomly matched "hypothetical" father-son skill pairs
3. "Hypothetical" father-son skill pairs matched by education



Racial Wage Gap

Wage premium is paid to sons who work in occupations that require similar skills as their fathers?

$$w=\beta_0 B+\beta_1 B\times r+\beta_2 W\times r+\beta_3 B\times 1[OC_c=OC_p]+\beta_4 W\times 1[OC_c=OC_p]$$

- *control for cognitive skill (AFQT)*
- *r : correlation coefficient*
- *B: black dummy, W: white dummy*
- *OC_c=OC_p: same occupation between parent & child*

- Transfer of occupationally related human capital: β_1, β_2
- Nepotism effect: β_3 , β_4

• Neal and Johnson (1996): cognitive skill explains almost all of the racial wage gap

Table 11: Effect of Father-Son Skill Correlation on Son's Earnings

		Father-Son		
		[C],[P]	[C], [M]	[C], [P], [M], [Ph]
Variables	(1)	(2)	(3)	(4)
Blacks	-0.047	-0.029	-0.025	-0.026
	(0.020)	(0.020)	(0.021)	(0.021)
Whites × Correlation Coefficient (r)		0.075	0.055	0.039
		(0.014)	(0.014)	(0.014)
Blacks × Correlation Coefficient (r)		-0.052	-0.079	-0.080
		(0.024)	(0.025)	(0.026)
Whites × 1(Same Occupation as Father)	0.075	0.014	0.030	0.043
	(0.035)	(0.036)	(0.037)	(0.037)
Blacks × 1(Same Occupation as Father)	0.066	0.099	0.117	0.117
	(0.052)	(0.054)	(0.054)	(0.054)
Father's Education	0.009	0.008	0.008	0.008
	(0.003)	(0.003)	(0.003)	(0.003)
R^2	0.260	0.265	0.264	0.263
N	17480	17480	17480	17480

Effect of Father-Son Skill Correlation on Son's Earnings

Variables	[C], [P], [M], [Ph]
Blacks	-0.032
	(0.021)
Whites × r × Level 1	0.081
	(0.029)
Blacks × r × Level 1	-0.077
	(0.062)
Whites × r × Level 2	0.048
	(0.022)
Blacks × r × Level 2	-0.038
	(0.043)
Whites × r × Level 3	0.025
	(0.032)
Blacks × r × Level 3	-0.131
	(0.054)
Whites × r × Level 4	-0.049
	(0.030)
Blacks × r × Level 4	-0.104
	(0.040)
R^2	0.298
N	17388

Interact *r* with family income level in 1978

(measure that controls for individual's family size, farm/non-farm distinction, and state of residence)

Divide in 4 groups

- **Level 1: Most affluent**
- **Level 2: More affluent**
- **Level 3: Less affluent**
- **Level 4: Least affluent**

Effect of Father-Son Skill Correlation on Son's Earnings

Variables	[C], [P], [M], [Ph]
Blacks	-0.042
	(0.021)
Whites × r × Professional, Technical	0.207
	(0.042)
Blacks × r × Professional, Technical	0.382
	(0.182)
Whites × r × Managers, Administrators	0.140
	(0.038)
Blacks × r × Managers, Administrators	0.261
	(0.092)
Whites × r × Clerical, Sales	0.172
	(0.068)
Blacks × r × Clerical, Sales	0.022
	(0.144)
Whites × r × Craftsmen	0.011
	(0.027)
Blacks × r × Craftsmen	0.075
	(0.045)
Whites × r × Operatives	-0.079
	(0.031)
Blacks × r × Operatives	-0.130
	(0.047)
Whites × r × Service, Laborer	-0.106
	(0.036)
Blacks × r × Service, Laborer	-0.181
	(0.042)
Whites × r × Farmers	-0.127
	(0.096)
Blacks × r × Farmers	-0.262
	(0.498)
Whites × r × Protective, Armed Services	-0.007
	(0.096)
Blacks × r × Protective, Armed Services	-0.172
	(0.233)
R^2	0.276
N	17398

Interact *r* with father's occupation

Conclusions

- Present a model of intergenerational skill transfer
- Social skills correlate across generations
  - Father-son: only in early adulthood; not at age 6
  - Mother-daughter
- Overall skills correlate across generations for father-son pairs
- Wage premium paid to white sons who work in similar occupation to their fathers
  - but wage penalty for blacks

Age, Cohort, and Wave effects in Japanese Household Consumption[†]

Preliminary Draft (not to be quoted)

Seki Asano*
Graduate School of Systems and Information Engineering, University of Tsukuba

Jie Jing Wang**
Graduate School of Engineering, University of Tokyo

February 12, 2009

Abstract

Consumer behavior changes with life cycle and generation, as well as with macroeconomic shocks. We estimate Japanese demand system from the 5 waves of national surveys that cover 21 years by allowing for varying taste parameters and household equivalence scale. Out of 11 broad consumption categories, significant age effects are found in housing, transportation, leisure, education and miscellaneous expenditures. Cohort effects appear in younger age in leisure and transportation but they taper off with age. Housing is strongly affected by wave effects and show irregular movement. We propose a method to estimate equivalence scales and implement it. The results indicate that equivalence scale is roughly in between 0.2 and 0.6, a slightly less than the conventional value, and a degree of economy of family scale is decreasing with time.

Keywords: Japanese national survey of family income and expenditure, Consumer demand system, Cohort effects, Age effects, Household Equivalence Scale, Almost Ideal Demand System

---

[†] We would like to thank Masayo Wakabayashi for helpful advice on assembling expenditure and price data sets. We also thank Wakabayashi and seminar participants of the Socio Economic Research Center, Central Research Institute of Electric Power Industry, for comments and discussion on preliminary draft of this paper. Any remaining errors are ours.

[*] 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan 305-8573, asano@sk.tsukuba.ac.jp

[**] 7-3-1 Hongo, Bunkyo, Tokyo, Japan 113-8956, tt086398@mail.ecc.u-tokyo.ac.jp

1. Introduction

Households' tastes and consumption behavior change with life cycle, they are affected by shocks in macro economy and future expectation, and they vary by generation. In this study we estimate Japanese household consumer demand system from the 5 waves of National survey of family income and expenditure (NSFIE) conducted over 21 years, from 1984 to 2004. Our main focus is to examine how the parameters of various age groups and generations change, and at the same time how household equivalence scales change.

For this purpose, NSFIE has various advantages over other national surveys. First of all, it covers over 55 thousand households, while another large scale survey, annual Family Income and Expenditure Survey (FIES) covers 7 thousand. Secondly, NSFIE offers tables classified and aggregated by 47 prefectures, and by age groups. Thirdly, information on household size is available for these subgroups from which we estimate household equivalence scale. In addition, NSFIE has data on imputed rent and ratio of home owners for each subgroup[1]. Using this information we can calculate expenditures on housing for each subgroup. Buying house is one of the biggest decisions in one's life cycle and there exists close relationship between expenditure on housing and those on other commodities. In any rate, we cannot ignore housing expenditure in estimating demand system.

There are not many demand system studies that paid attention to taste change due to age and generation. Wakabayashi (2001), and Wakabayashi & Hewings (2007) [2] estimated consumer demand system by pooling 4 waves (1984-1999) of NSFIE and analyzed taste change by age. Also, using NSFIE data Maki (2006), analyzed change in expenditure shares and saving rates before and after the so-called collapse of bubble economy in 1990, and found evidence of structural change.

These studies either focused on effects of age on consumption or effects of wave on consumption and savings, but not on both. To our knowledge, however, this is the first attempt to analyze effects of three factors, age, wave and generation on consumption through estimation of demand system parameters.

The paper proceeds as follows: Section 2 describes data on expenditures and prices. In section 3, we visualize shift in expenditure shares by graphical methods, and discuss notable patterns by ages and generations. In section 4, theoretical model of consumer demand is introduced and we propose a method to estimate household equivalence scales. Section 5 presents estimation results and discusses their implications. Section 6 concludes.

2. Data

2.1 Expenditures

NSFIE is conducted every 5 years and the most recent wave was in 2004. In this study we use 5 waves starting from 1984. Number of households in each survey is approximately 55 thousand. Individual household data are not open to public for privacy protection. Only the aggregated figures are accessible. Sample households report expenditures during the 3-month period starting from September[3]. We use all households' average monthly expenditures on 11 expenditure groups, classified by 47 prefectures, and 12 age groups. Age groups are divided by 5 years. The lowest bracket is under 25 years old (exclusive), and the highest is 75 and over. We estimate demand systems from these (47 x 12) representative consumers' data, observed over 5 waves.

2.2 Classifications

NSFIE has 10 large commodity classification table, they are 1) Food, 2) Housing, 3) Fuel, light and water charges, 4) Furniture and household utensils, 5) Clothes and footwear, 6) Medical care, 7) Transportation and communication, 8) Education, 9) Reading and recreation, and 10) Other expenditures. We separate 7) into "Transportation" and "Communication", and combined "eating out", which is in 1), with 9) and call it "Leisure." The resulting 11 classifications and their abbreviations are given below. We use these abbreviations

[1] FIES does not collect imputed rent data.

[2] Asano [1997] estimated Japanese extended demand system using panel of 47 prefectures for average workers' households using FIES and labor supply data. He introduced fixed time effects, which allowed variations in intercepts, but assumed constancy of slope parameters.

[3] Thus we can expect that regional effects due to difference in climate are not large.

in table and figures. Also, we use paired parentheses () to indicate reclassified 11 commodity groups.

11 classification
(1) Food (Fd)
(2) Housing (Hs)
(3) Fuel&Water (Eg)
(4) Furniture (Fn)
(5) Clothes (Cl)
(6) Medical Care (Md)
(7) Transportation (Tr)
(8) Communication (Cm)
(9) Leisure (Ls)
(10) Education (Ed)
(11) Other (Ot)

2.3 Imputed rent

"Housing" in NSFIE does not include imputed rent for home owners. Thus, while large portion of housing expenditure for non-home owners is rent payment, for home owners it is mostly repair and maintenance costs. For this reason, home owners' housing shares are substantially less than those of non-home owners. To cope with this problem we made an adjustment for imputed rent for home owners from the figures reported in NSFIE. Namely, we use imputed rent for prefecture in each wave, and home ownership ratio in each age group to construct housing expenditure[4]. We defined housing expenditure by the following:

Housing = "Housing" in NSFIE + Imputed rent x Subgroup's Home ownership ratio

Total expenditure, which corresponds to "income" in standard microeconomic theory, is defined as the sum of expenditures on 11 commodity groups, in which housing includes imputed rent for home owners.

2.4 Prices

The data source for prices is the National Survey of Consumer Prices (NSCP). NSCP is conducted every 5 years, but timing of the survey is not the same as NSFIE. There is 2-year difference between the two surveys. We use 1982, 87, 92, 97, and 2002 NSCP data, and interpolate[5] them to create regional price data corresponding to NSFIE. In this manner, it is possible to construct price indexes of 11 commodity groups that vary by 47 prefectures and by 5 waves. We construct NSFIE survey years' price indexes by the following formula (*)[6].

Price of commodity i, in region j and year t = CPI_{it}(1995=100) x NSCP Regional Difference Index R_{ij} (*)

where subscript i,j, and t stand for commodity group i, region j, and time t, respectively.

Also, price indexes for above-mentioned reclassified groups, 1) Food, 7) Transportation, 8) Communication, 10) Leisure, should be recalculated using the results from (*). We used the following relationship (**) in obtaining prices for Food (P_{fd}) and Leisure (P_{ls})

$$P_{fd} = P_{eo}w_{eo} + P_{ei}w_{ei}$$
$$P_{ls*} = P_{eo}w_{lseo} + P_{ls}w_{ls} \qquad (**)$$

Here we omit subscripts for region, and year. P represents regional prices, and w is expenditure share. Subscripts are for the following items. fd: food including eating out, eo: eating out, ei : food excluding eating out, ls*: leisure including eating out, ls: leisure excluding eating out. It follows that w_{eo} (w_{ei}) is the share of "eating out (excluding eating out)" in "food" in the 10-large group classification ($w_{ei}+w_{eo}=1$). Similarly, w_{lseo}(w_{ls}) is the share of "eating out (excluding eating out)" in "Leisure" in 11-large group classification

[4] NSFIE gives single imputed rent figure for each prefecture in a given wave. So we use the same imputed rent for all age groups. Proportions of home owners vary by age, prefecture, and wave.
[5] In case of 2002 NSCP, we extrapolate them to obtain 2004 price data.
[6] The formula yields two price indexes, depending on interpolation made by CPI_{it} is "forward" or "backward." We use average of two figures in NSFIE survey years.

($w_{ls}+w_{lseo}=1$).

For transportation, middle-classification price indexes for "public transportation" and "automobile related expenditures" are used in calculating (*) and then we constructed weighted average by using expenditure shares. Middle-classification figures are available for communication, so we can use (*) to obtain the price index for communication.

3. Change in Expenditure Shares

It is natural to assume that ones consumption behavior change as he/she ages. For example, share of education is low for family with infants, but it increases as kids grow, and it will go down after kids become independent and leave family. Also, people in, their 30's in 1985 may not have the same taste as 30's in 2005. We examine taste variation due to age, generation, and timing of survey, by looking at parameter estimates of demand systems. We call these three effects as Age effects, Cohort effects and Wave effects.

Age effects represent taste change due to family head's age. As people age their consumption behavior changes with their life stage. Cohort effects reflect change in life style due to difference in generation. For example, tastes of post WW-II baby boomers are certainly different from those born after 1970's. Those belong to the later generation, tend to marry at older age, tend to have fewer children, and for them marriage rate is lower. Also, kinship in younger generation is getting weaker and they have stronger attachment to nuclear family. Wave effects represent factors specific to the survey year, such as business cycle and other macro economic shocks. In fact, Japanese economy is said to have 4 phases in the sample period. Early 1980's was a stable growth period, 85 to 89 was so called "Bubble economy" era, 90 to 2000 was called "lost decade", and after 2000 may be tentatively called recovery period. These macro economic factors should have effects on family expenditures.

We visualize changes in consumption patterns in Fig.1 by showing shift of expenditure shares. Fig.1 has 11 sub-figures that corresponding to 11 expenditure group. The horizontal axis is age (AGE) and vertical axis is share (W). Each sub-figures have 2 representations, **a** and **b**. The points on both **a** and **b** show identical data[7], but connecting lines are different. In **a**, points are connected for those belong to the same wave. In **b**, points are connected for the points belonging to the same (pseudo) cohort. There are 8 pseudo cohorts grouped by 5-year bracket and denoted by A~H. They are, household head was born A: after 1960, B: between 1955~59, C: 1950~54, D: 1945~49, E: 1940~44, F: 1935~39, G: 1930~34, and H: 1925~29.

Age effects can be seen by the shape of curves in Fig.1.*-a and Fig.1.*-b (* = 1,.,11). Wave effects are found by shift of curves in Fig.1.*-a, and Cohort effects can be captured by shift of curves in Fig.1.*-b. We look at shifting patterns of shares and discuss some notable patterns for individual expenditure groups.

Movement of food shares show presence of all three effects, age, wave, and cohort effects. In Fig.1.1-a curves shift downward with time which indicate improvement in standard of living (wave effect). Also, Fig.1.1-b shows that curves move down for younger cohorts (cohort effect), and the age profiles are such that it starts out higher for 20's and steadily goes down by age and it go up a little at age over 60 (age effect).

For housing (HS) wave effects are observed in Fig.1.2-a. On the other hand, age profiles of curves in Fig.1.2-a, and Fig.1.2-b show presence of age effects, but in Fig.1.2-b curves do not shift much by cohort, thus cohort effects are weak.

Fuel and Lights (EG) does not show presence of three effects.

Furniture (Fn) and clothes (Cl) show similar patterns. They go down slightly with age (age effects), and between generations they tend to shift down for younger cohorts (cohort effects). Younger people tend to be more interested in fashion and clothes, and spend more money on furniture and electric products, but such tendency recedes with age.

For medical care (Md) we cannot observe wave and cohort effects but age effects exists. Fig.1.6-b shows that the shares of Md start rising at age 55-60. Higher share at younger age may be attributed to medical expense for infants.

[7] To be precise, data points not belonging to cohorts A to H are excluded in Fig.1.*-b, but they are in Fig1.*-a.

Transportation shows clear age effects. Shares peak at late 20's and go down gradually with age. Also, 2004 shares are higher than other waves.

Communication (Cm) appears to have all three effects. Fig.1.8-a shows that the curves stay close to each other until 1994 but they move up in 99 and 2004 for younger cohorts. These are due to widespread use of internet and cellular phone in recent years and resulting growth in expenditures. As expected, increase in shares is higher for younger cohorts (Fig.1.8-b).

Leisure (Ls), education (Ed) , and other expenditures (Ot) show clear age effects. Also, in LS and Ot neither cohort effects nor wave effects are observed.

Leisure (Ls) has two peaks in 30~35 and 60~70. Peak in 60-70 is reflection of strong demand for travel and recreation by the post retirement generation. People in these age groups have both money and time. Decline in 35~55 is attributed to the fact that households heads being too busy in their work places, and so have no time for leisure.

Education (Ed) shows age effects which move like a slightly distorted mirror image of leisure. Age profiles have peak values 6% to 8% at age 40 to 50, and go down sharply to 1%~2% in late 50'. We note that despite steady decline of Japanese birth rate and number of children per family, shares are increasing around peak age in recent years (Fig.1.10-a). After age 55 and over, both cohort effects and wave effects disappear.

Other expenditures (Ot) show distinct age profile, and variations by age are large. Shares stay around 20% until early 40's but jump up 30% in 50's then go down to 20% in 60 and over. Ot includes social expenditures, such as ceremonial expenses, remittance, such as money sent to the family members not living together, and other miscellaneous expenditures[8].

4. Model

4.1 Basic AIDS

The basic model employed in this study is Almost Ideal demand system by Deaton and Muellbauer (1980). The outline of the model is given below. The log cost function in AI demand system is,

$$ln e(\mathbf{p},u) = \alpha_0 + \Sigma_i \alpha_i ln p_i + 1/2\ \Sigma_i \Sigma_j \gamma_{ij} ln p_i ln p_j + u\beta_0 \Pi_i p_i^{\beta i} \qquad (1)$$

where, **p** is the price vector, subscripts i,j represent *i*th and *j*th good respectably, p_i's are individual prices, and u is the utility level which takes value between 0 (subsistence) and 1(bliss). α_0 is the log subsistence expenditure, and β_0 is the log expenditure at the bliss, when all the prices are normalized at one.

Partial differentiation of (1) by log prices will yield share equations in (2).

$$w_i = \alpha_i + \Sigma_j \gamma_{ij} ln p_j + \beta_j ln(Y/P) \qquad (2)$$

where, w_i: expenditure share of the *i*th good, Y is the total expenditure, and P is the price index given by,

$$ln P = \alpha_0 + \Sigma_i \alpha_i ln p_i + (1/2) \Sigma_i \Sigma_j \gamma_{ij} ln p_i ln p_j. \qquad (3)$$

From equation (1) we can see P is the minimum expenditure to achieve subsistence. Expenditure (or income) elasticity is given by $E_i = 1 + \beta_i/W_i$

In order for the system to be consistent with the demand theory, there are the following restrictions.

(i) $\Sigma_i \alpha_i = 1,\ \Sigma_i \beta_i = 0,\ \Sigma_i \gamma_{ij} = 0$ (Additivity)

Expenditure shares add up to 1 for any prices and expenditure level.

[8] Average shares within Ot are 21% (social expenses), 32% (remittance) and 47%(other). Other expense in Ot includes, hair cut, hair care, cigarettes, etc.

(ii) $\Sigma_j\gamma_{ij} = 0$ (Homogeneity)

Share equation is homogeneous of degree zero in prices and total expenditure.

(iii) $\gamma_{ij} = \gamma_{ji}$ (Symmetry)

This condition assures symmetry of the Hicks substitution matrix[9].

It is a common and very popular practice to estimate the system by replacing lnP by the Stone price index given by (4),

$$lnP^{*} = \Sigma_i w_i lnp_i \tag{4}$$

Then the unrestricted version of the system can be estimated by equation by equation OLS. This approach is called LAIDS. It is needless to say that it is desirable to estimate the system by full non-linear fashion, and use the results obtained by nonlinear estimation for welfare analyses. Because parameters obtained from the LAIDS cannot be interpreted as the parameters of expenditure function and so the connection between LAIDS and indirect utility function is unclear[10]. In this study we estimate the full nonlinear system.

4.2 Household Equivalence Scale

In order to incorporate household size into the basic model, we go back to the expenditure function. We denote household size by n. An obvious and the simplest way to introduce household size is to replace Y by per capita expenditure, Y/n in (1)

$$w_i = \alpha_i + \Sigma_j\gamma_{ij} lnp_j + \beta_j ln(Y/nP) \tag{5}$$

A more general version of (5) is

$$w_i = \alpha_i + \Sigma_j\gamma_{ij} lnp_j + \beta_j ln(Y/n^{\delta}P) \tag{6}$$

The parameter δ is called equivalence elasticity and represents the degree of economy of scale. δ should be between 0 and 1. if δ=1 there is no economy of scale, and only the per capita expenditure matters for family members' welfare, and if δ = 0, only the total family expenditure matters regardless of family size. Thus Y/n^{δ} is interpreted as real income adjusted to family size. What would be the value of δ? Sure enough δ can't be zero, and it is unlikely to be one. Common practice is to assume δ = 0.50~0.72 and it is widely used in many governmental and international agency's (such as OECD) reports in evaluating real income level and constructing inequality measures.[11] In this study we try to obtain δ's through estimation of demand system.

With a little abuse of notation, one of the candidates for log cost (or expenditure) function which yield share equation in (6) is (7).

$$lne(\mathbf{p},u;n) = \alpha_0 + \Sigma_i\alpha_i lnp_i + 1/2\, \Sigma_i\Sigma_j\gamma_{ij} lnp_i lnp_j + u\beta_0 \Pi_i p_i^{\beta_i} + \delta ln\, n \tag{7}$$

where e(**p**,u;n) is the expenditure function of a family with household size n, and u should be interpreted as the average utility level of the family members.

(7) has very appealing feature as money metric utility function. Moving $\delta ln\, n$ to the left hand side of (7), $lne(\mathbf{p},u;n) - \delta ln\, n$ is a monotonic transformation of u, and may be interpreted as the log of real income. This hope, however, had been convincingly shattered by Pollak and Wales (1978). Deaton (1997) gives comprehensive summary and demonstrated that there are other candidates for the expenditure functions that yield the identical share equations as (6) [12], and emphasized that "we cannot infer welfare from behavior". However, backward reasoning is possible. If we are employing Y/n^{δ} in measuring family welfare, the

[9] See Deaton and Muellbauer (1980) for the Hicks substitution matrix and negativity condition.

[10] See Chen (1998).

[11] Alternative approach is to define real per capita income as Y/ϕ(n), and assign ϕ(1) = 1, ϕ(2) = ϕ(1) + d_2, ϕ(3) = ϕ(2) + d_3, etc., increments d_j's are less than 1 and decreasing in j. See OECD (2008).

[12] See Deaton (1997) chapter 4, in particular pp.268-269.

demand system consistent with such assumption is (6) and it is worth estimating them.

Share equation in (6) is rewritten as

$$w_i = \alpha_i + \Sigma_j\gamma_{ij} ln p_j + \beta_i ln(Y/P) - \beta_i \delta lm \tag{8}$$

For estimation purpose we rewrite the share equation in (8) by adding disturbance term ε_i,

$$w_i = \alpha_i + \Sigma_j\gamma_{ij} ln p_j + \beta_i ln(Y/P) - \beta_i \delta lm + \varepsilon_i \tag{9}$$

where ε_i's have zero expectation $E(\varepsilon_i) = 0$, are independent across samples, and $Var(\varepsilon_i) = \sigma_i^2$, and $Cov(\varepsilon_i, \varepsilon_j) = \sigma_{ij}$. $(i \neq j)$.

4.3 Estimation

Nonlinear estimation of (9) with P defined in (3) will yield household equivalence scale estimate, along with other parameters. Note that there is an additional restriction in (9). That is, coefficient vector $\beta = (\beta_1, \beta_2, .., \beta_{11})'$ on real expenditure term $ln(Y/P)$ and those on family size lm are proportional, and have opposite sign. We call it proportionality condition.

For estimation purpose, we regroup 12 age groups in the original NSFIE to 5 groups. These 5 groups are, group 1) family head is in 20's, 2) in 30's, 3) in 40's, 4) in 50's, 5) 60's[13]. We assume that families in each of these groups have the same taste within the same wave. We estimate 25 systems (5 age groups times 5 waves) by allowing for taste variation by age and wave. The typical sample size in each group is 94 (47 prefectures x 2 age groups in the original NSFIE) [14]. Thus, typically we have 940 data points to estimate 76 parameters[15].

We estimate (3) and (9) with three restrictions (homogeneity, symmetry, and proportionality). Also, we estimated the system without proportionality restriction. That is, the coefficients on lm are not restricted to be proportional to β_i's. In both cases the systems are estimated by the maximum likelihood with normality assumption.

5. Estimation Results

We examine taste changes due to three factors, age, generation, and wave. As mentioned earlier, we estimated 25 sets of demand systems using data obtained from 5 age groups over 5 waves. Taste change can occur in the form of shifts in $\alpha_i, \beta_i, \gamma_{ij}$, and δ_i. In order to get an intuition on how tastes change, however, we focus our attention mainly on β values which capture response of expenditures to income change.

5.1 Shift of β's

Table 1 reports estimates and their standard errors of β's. Brown (blue) shaded figure indicates that the coefficient is significantly positive (negative) at the 2.5% significance level. As mentioned earlier, positive (negative) β indicates that the corresponding category is luxury (necessary) good. To get a good glimpse of how they change with age and wave, it is more informative to see the graphs that show shift of β's for 11 expenditure categories.

Fig.2 shows estimate of β's for each expenditure group. Horizontal axis is age and vertical axis is point estimates of β's. Waves are indicated by separate colors, and estimates obtained in the same wave are connected by solid line. So there are five lines in each sub-figure. If these five lines are horizontal and stay close, we may interpret that the taste change (in the form of shift in β) did not take place for that expenditure group.

Expenditure groups that did not show notable taste change are Eg, Cl, Md, and Cm and their β values are

---

[13] We excluded age groups over 70 from estimation.

[14] Actual sample size vary due to missing values. .

[15] There are 11 shares for each sample, of which one of 11 is determined if other 10 shares are given. Thus the number of data points on the LHS is 94x10 = 940. The number of parameters is 76 = 10 α_i's + 10 β_i's + 55 γ_{ij}'s (with symmetry and adding up restriction) + 1 (δ).