

UWARUNKOWANIA POMIARU WSPÓLZALEŻNOŚCI CECH

TADEUSZ GRABIŃSKI

Katedra Finansów
Uniwersytet Ekonomiczny w Krakowie
PL 31-510 Kraków, ul. Rakowicka 27
email: tg@ae.krakow.pl

Praca przedstawiona na posiedzeniu Komisji Nauk Ekonomicznych PAN w dniu 20 stycznia 2009 r. przez autora.

ABSTRACT

Tadeusz Grabiński, *Problems of measuring statistical dependence of variables*, Folia Oeconomica Cracoviensia, 2008–2009, 49–50: 91–107.

The paper presents propositions of modification in correlation and regression linear analysis. It is a result of problems in interpretation of statistical dependence measures in case where data are referring to aggregate units (e.g. countries). The paper shows results of research made at the beginning of 2000 for 81 countries by R. Lynn and T. Vanhanen who show that IQ is one of important factors contributing to differences in wealth of nations and rates of economic growth.

KEY WORDS — SŁOWA KLUCZOWE

measures of dependence, analysis of correlation and regression, wealth of nation, IQ
miary współzależności, analiza korelacji i regresji, bogactwo narodów, iloraz inteligencji

1. WPROWADZENIE

Podstawą badań statystycznych są informacje opisujące analizowane elementy zbiorowości (próby) z punktu widzenia wybranej cechy (jednej lub więcej). W każdej analizie należy zdefiniować jednostki badania oraz określić cechy charakteryzujące te jednostki.

Wyróżnia się jednostki elementarne oraz agregatowe. Trudno podać precyzyjne rozróżnienie tych pojęć, gdyż zależy ono od kontekstu

analizy. Przykładem jednostek elementarnych jest osoba, firma, produkt, gmina i w tym kontekście odpowiednimi jednostkami agregatowymi są: gospodarstwa domowe dla osób; branża dla firm; region, województwa, powiaty dla gmin itp.

To rozróżnienie nie zawsze jest brane pod uwagę w analizach ekonometrycznych i może prowadzić do kłopotów w interpretacji uzyskanych wyników. W wielu przypadkach mamy do czynienia z jednostkami agregatowymi, które traktujemy z jednakową wagą, podczas gdy nie zawsze jest to uzasadnione wielkością tych jednostek.

Dodatkowym elementem, na który należy zwrócić szczególną uwagę w badaniach są miana (jednostki pomiarowe) analizowanych cech. Mogą to być miana naturalne (zł, kg, osoby) lub złożone o charakterze wskaźników (dochód narodowy na osobę, zarobek tygodniowy na 1 zatrudnionego).

Przedmiotem niniejszego artykułu jest przedstawienie problemów związanych z istotą jednostek analizy i ich mian w przypadku pomiaru współzależności cech. Rozważania ograniczone zostaną do najprostszych metod oceny współzależności — liniowych współczynników korelacji i regresji jednej zmiennej, ale można je rozszerzyć na bardziej zaawansowane miary.

Dla ułatwienia przyjmuje się najprostszy przypadek analizy, w którym mamy do czynienia z danymi statycznymi. Punktem wyjścia rozważań są więc informacje liczbowe (skala ilorazowa) dotyczące n -elementowego zbioru jednostek, opisujące je z punktu widzenia dwóch cech X i Y w wybranym momencie czasowym:

$$\begin{aligned} \{X\} &= [x_1, x_2, \dots, x_n] \\ \{Y\} &= [y_1, y_2, \dots, y_n]. \end{aligned}$$

2. WSPÓLZALEŻNOŚĆ POMIĘDZY BOGACTWEM A INTELIGENCJĄ W SKALI MAKROEKONOMICZNEJ

Dla zilustrowania trudności z jakimi można się spotkać przy analizie współzależności przytoczono wyniki badań R. Lynna oraz T. Vanhanena (2002) nad bogactwem narodów. Przyczyny bogactwa i ubóstwa są przedmiotem zainteresowania wielu badaczy (Landes, 1998), począwszy od pracy Adama Smitha z 1776 roku *Badania nad naturą i przyczynami bogactwa narodów*, poprzez teorie wskazujące na takie determinanty bogactwa narodów jak: klimat (Monteskiusz, 1748; Landes, 2000), kultura i religia (Weber, 1994), położenie geograficzne (Diamond, 1998), system gospodarczy (Olson, 1996), rozwój technologii (Solow, 19956), kapitał ludzki i inteligencja (Herrnstein i Murray, 1996; Lynn i Vanhanen, 2002).

Zależność pomiędzy bogactwem a inteligencją można badać w skali mikroekonomicznej, gdzie jednostkami analizy są osoby. W pracy R. J. Herrnste-

ina i Ch. Murraya (1994) wskazywano, że inteligencja jest dobrym predyktorem sukcesu zawodowego, a niski poziom IQ jest jedną z głównych przyczyn plag społecznych — bezrobocia, przestępczości, narkomanii, samotnego macierzyństwa. Drugim możliwym ujęciem jest analiza makroekonomiczna, w której jednostkami analizy są regiony lub kraje. W pracy R. Lynna i T. Vanhanena (2002) postawiono tezę, że zamożność w skali makroekonomicznej zależy od inteligencji narodu i jest z nią wysoko oraz dodatnio skorelowana.

Zamożność krajów opisano wielkością produktu narodowego brutto (GDP) na osobę w dolarach na podstawie danych dla roku 1998. Analizę oparto na informacjach odnoszących się do 81 krajów, dla których udało się zebrać w miarę porównywalne dane (główne chodziło o trudności w zebraniu odpowiednio dużej liczby informacji opisujących stopień inteligencji). Cecha GDP jest często wykorzystywana w wielu analizach i pomimo pewnych zastrzeżeń co do adekwatności miernika GDP jako dobrego wskaźnika bogactwa na ogół jest akceptowana w wielu opracowaniach.

gorzej przedstawia się sprawa pomiaru inteligencji w skali makroekonomicznej (dla mieszkańców danego kraju). R. Lynn i T. Vanhanen założyli, że inteligencja nie jest równoznaczna z wykształceniem, wiedzą, wynalazczością czy też mądrością i w swoich badaniach mierzyli ją średnim wynikiem testu inteligencji (iloraz inteligencji — IQ) z wielu indywidualnych pomiarów wy-

Tabela 1

Współczynniki inteligencji (IQ), produkt narodowy brutto GDP (w \$ na osobę) oraz liczba ludności (w tys.) w wybranych krajach (dane za różne lata w okresie 1996–2002)

Kraj	IQ	GDP/os. (\$)	Ludność (tys)	Kraj	IQ	GDP/os. (\$)	Ludność (tys)
Argentyna	96	12 013	38 337	Lebanon	86	4 326	3 731
Australia	98	22 452	19 767	Malaysia	92	8 137	23 105
Austria	102	23 166	8 163	Marshall Islands	84	3 000	56
Barbados	78	12 001	277	Mexico	87	7 704	103 718
Belgium	100	23 223	10 331	Morocco	85	3 305	31 754
Brazil	87	6 625	183 960	Nepal	78	1 157	26 470
Bulgaria	93	4 809	7 588	Netherlands	102	22 176	16 223
Canada	97	23 582	31 889	NewZealand	100	17 288	3 960
China	100	3 105	1 291 496	Nigeria	67	795	131 728
Colombia	89	6 006	41 802	Norway	98	26 342	4 555
Congo (Brazz)	73	995	3 413	Peru	90	4 282	27 275
Congo (Zaire)	65	822	56 886	Philippines	86	3 555	86 752

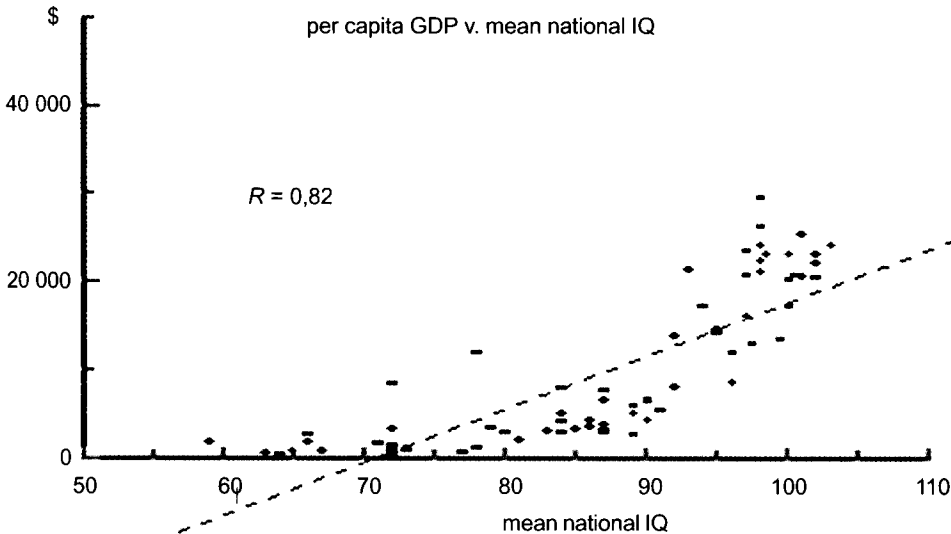
Kraj	IQ	GDP/os. (\$)	Ludność (tys)	Kraj	IQ	GDP/os. (\$)	Ludność (tys)
Croatia	90	6 749	4 498	Poland	99	7 619	38 603
Cuba	85	3 967	11 247	Portugal	95	14 701	10 480
Czech Republic	97	12 362	10 251	Puerto Rico	84	8 000	3 878
Denmark	98	24 218	5 394	Qatar	78	20 987	725
Ecuador	80	3 003	13 074	Romania	94	5 648	22 380
Egypt	83	3 041	74 761	Russia	96	6 460	144 308
Equatorial Guinea	59	1 817	536	Samoa (Western)	87	3 832	204
Ethiopia	63	574	70 366	Sierra Leone	64	458	5 571
Fiji	84	4 231	869	Singapore	103	24 210	4 277
Finland	97	20 847	5 204	Slovakia	96	9 699	5 416
France	98	21 175	62 171	Slovenia	95	14 293	2 012
Germany	102	22 169	82 398	South Africa	72	8 488	46 567
Ghana	71	1 735	21 111	Spain	97	16 212	40 217
Greece	92	13 943	10 626	Sudan	72	1 394	36 593
Guatemala	79	3 505	11 725	Suriname	89	5 161	450
Guinea	66	1 782	8 756	Sweden	101	20 659	8 970
Hong Kong	107	20 763	6 810	Switzerland	101	25 512	7 408
Hungary	99	10 232	10 058	Taiwan	104	13 000	22 543
India	81	2 077	1 057 504	Tanzania	72	480	36 199
Indonesia	89	2 651	223 070	Thailand	91	5 456	63 271
Iran	84	5 121	63 988	Tonga	87	3 000	108
Iraq	87	3 197	24 683	Turkey	90	6 422	68 109
Ireland	93	21 482	3 924	U. Kingdom	100	20 336	60 095
Israel	94	17 301	6 492	Uganda	73	1 074	26 322
Italy	102	20 585	57 998	United States	98	29 605	290 343
Jamaica	72	3 389	2 689	Uruguay	96	8 623	3 387
Japan	105	23 257	127 358	Zambia	77	719	10 800
Kenya	72	980	33 042	Zimbabwe	66	2 669	11 816
Korea, South	106	13 478	47 657				

Źródło: kol. 2-3 — Lynn i Vanhanen, 2002, kol. 4 — <http://www.census.gov/ipc/www/idb/tables.html>.

konanych na mieszkańcach poszczególnych krajów. W sposób bezpośredni udało się im zebrać informacje z 81 krajów (zapewne o różnym stopniu reprezentatywności). W niektórych krajach (Polska, Holandia, Belgia) dane te były gromadzone w ramach widowisk telewizyjnych w stylu: *Narodowy test inteligencji Polaków* (TVN, 2003, 300 tys. osób). R. Lynn i T. Vanhanen oszacowali także średnie ilorazy inteligencji dla pozostałych 104 krajów (dla których dysponowali informacjami o dochodzie narodowym) na podstawie średnich ilorazów inteligencji w sąsiednich krajach, jednakże dane te nie zostały szerzej wykorzystane.

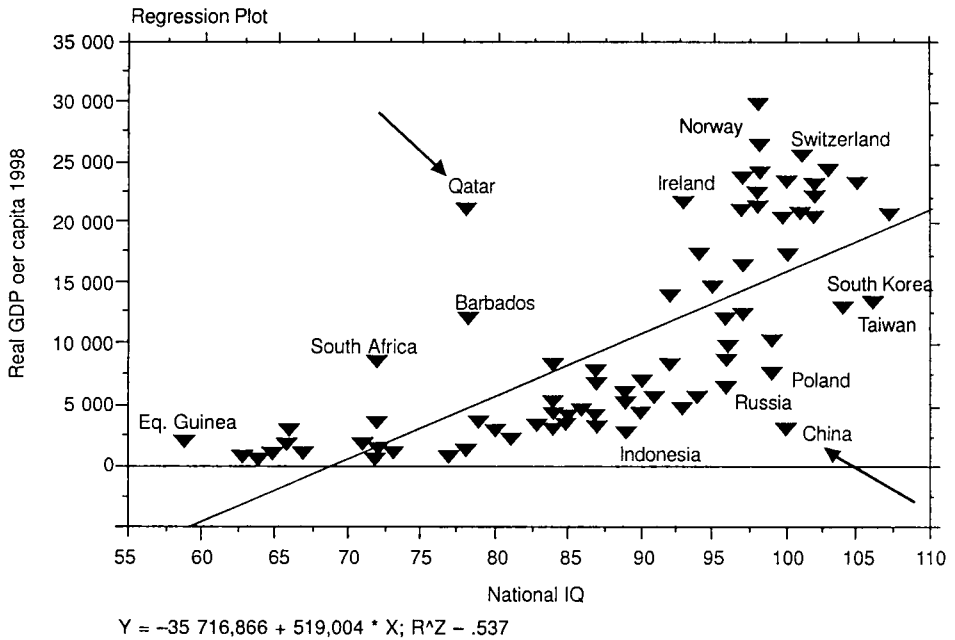
W tabeli 1 przytoczono wartości zmiennych stanowiących podstawę badań R. Lynna i T. Vanhanena. W ostatniej kolumnie przytoczono dodatkowo liczbę mieszkańców dla poszczególnych krajów, gdyż zadaniem niniejszego artykułu jest pokazanie jak się mogą zmienić wyniki analiz w przypadku uwzględnienia dodatkowych informacji (w tym przypadku liczby ludności) o badanych obiektach.

Na rycinach 1–2 przedstawiono rezultaty analizy R. Lynna i T. Vanhanena. Są to diagramy rozrzutu pokazujące zależność wyrażającą się w tym, że kraje zamożne zamieszkałe są na ogół przez mieszkańców z większymi ilorazami inteligencji. I odwrotnie, społeczeństwa krajów niezamożnych zazwyczaj charakteryzują się mniejszymi ilorazami inteligencji. Rycina 1 zawiera oryginalny diagram z pracy R. Lynna i T. Vanhanena natomiast rycina 2 — wykres zmodyfikowany przytoczony w <http://www.amren.com/ar/2002/12/>.



Źródło: <http://www.lagriffedulion.f2s.com/sft.htm>.

Ryc. 1. Oryginalny diagram rozrzutu pomiędzy GDP oraz IQ wraz z liniową funkcją regresji



Źródło: <http://www.amren.com/ar/2002/12/> *The Global Bell Curve. Uncommon sense on wealth and poverty*, reviewed by Thomas Jackson

Ryc. 2. Poprawiony diagram rozrzutu pomiędzy GDP oraz IQ dla 81 krajów wraz z liniową funkcją regresji

Różnica pomiędzy tymi rycinami wyraża się w tym, że rycina 2 odpowiada dokładnie danym zamieszczonym w tabeli 1, natomiast w rycinie 1 nie uwzględniono kilku krajów, m.in. Chin i Kataru, dla których wartości badanych zmiennych wyraźnie odbiegały od linowej linii regresji. Dlatego też współczynniki korelacji pomiędzy GDP a IQ przytoczone w obydwóch rysunkach różnią się od siebie. Na rycinie 1 (bez Chin i Kataru) współczynnik korelacji przyjmuje wartość 0,82, natomiast na rycinie 2 (z Chinami i Kataru) — tylko 0,73. Trudno się oprzeć wrażeniu, że ta operacja mogła być spowodowana chęcią wykazania większej siły zależności. Taki proceder prowokuje do pytania: czy przypadkiem nie dokonywano innych manipulacji danych dla potwierdzenia tezy o wysokim poziomie współzależności pomiędzy inteligencją a bogactwem?

Zależność widoczną na rycinie 2 opisano za pomocą liniowej funkcji regresji o postaci:

$$\text{GDP} = -35\,717 + 519 * \text{IQ} \quad (R^2 = 0,537).$$

Parametry tej funkcji można zinterpretować następująco. Parametr znajdujący się przy zmiennej IQ (519 \$ GDP/os/rok) wskazuje na „wartość” jednego punktu ilorazu inteligencji (wartość kapitału intelektualnego). Przy wzroście

średniego ilorazu inteligencji mieszkańców danego kraju o 1 punkt IQ można się spodziewać wzrostu zamożności o 519 \$ na 1 osobę w ciągu 1 roku. Jako że inteligencję „produkuje” system oświatowy kraju (szkolnictwo) można stąd oszacować wartość produktu tworzonoego przez ten system, np. w Polsce podniesienie średniego ilorazu inteligencji o 1 punkt IQ może skutkować wzrostem GDP o:

$$[519 \$ * 38 \text{ mln mieszkańców}] \approx 20 \text{ mld } \$/\text{rok} \approx 60 \text{ mld } \text{zł}/\text{rok}.$$

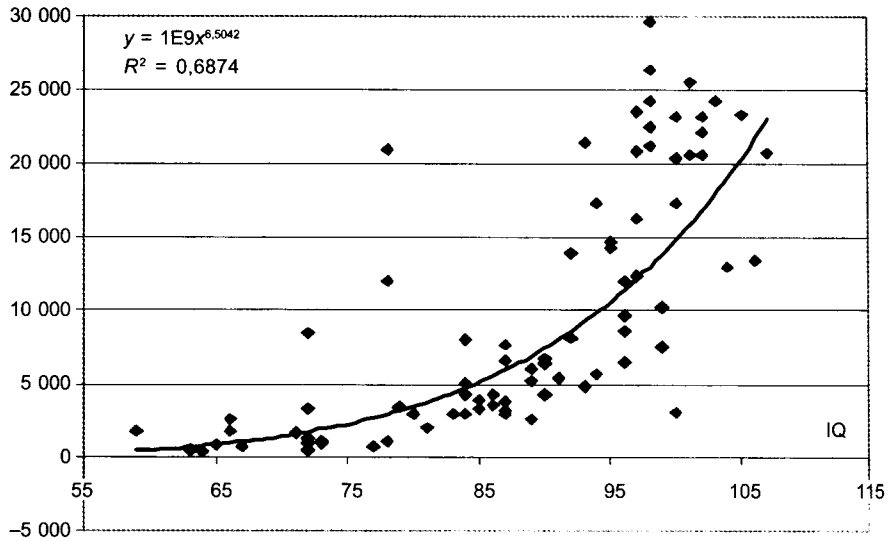
Oczywiście jest to rachunek obarczony znacznym błędem, którego wielkość można oszacować na poziomie $1 - 0,537 = 0,473$, czyli na ok. 47%. Jest zrozumiałe, że wzrost produktu narodowego zależy od wielu innych czynników, a poziom inteligencji (o ile w ogóle) jest tylko jednym z nich. Tym niemniej, nawet uwzględniając ten wysoki brak precyzji, sektory oświaty, szkolnictwa wyższego i nauki mają pewne podstawy do twierdzenia, że wartość ich działalności (po udowodnieniu, że iloraz inteligencji społeczeństwa w ciągu roku rzeczywiście wzrósł o 1 pkt IQ) kształtuje się co najmniej na poziomie ok. 30 mld zł w skali roku (tzn. 50% z 60 mld zł).

Interpretacja wyrazu wolnego jest bardziej abstrakcyjna. Formalnie jest to wielkość GDP w przypadku „zerowej” inteligencji mieszkańców danego kraju. W tym przypadku oznacza to, że do każdego „mieszkańca” należałoby doliczyć ponad 35 tys. \$ GDP na rok (ujemna wartość wyrazu wolnego funkcji regresji) aby społeczeństwo takie mogło w ogóle funkcjonować.

Bliższa analiza rycin 1 i 2 wskazuje, że liniowa funkcja regresji nie oddaje zbyt poprawnie empirycznego rozrzutu punktów. W szczególności dla krajów o niskich ilorazach inteligencji teoretyczne wielkości GDP są ujemne, podczas gdy faktyczne wielkości w tych krajach są dodatnie. Dlatego też na podstawie tych samych danych oszacowano dwie inne funkcje regresji: potęgową oraz wykładniczą. Wyniki w postaci diagramów rozrzutu z dopasowanymi funkcjami regresji przedstawiono na rycinach 3 i 4.

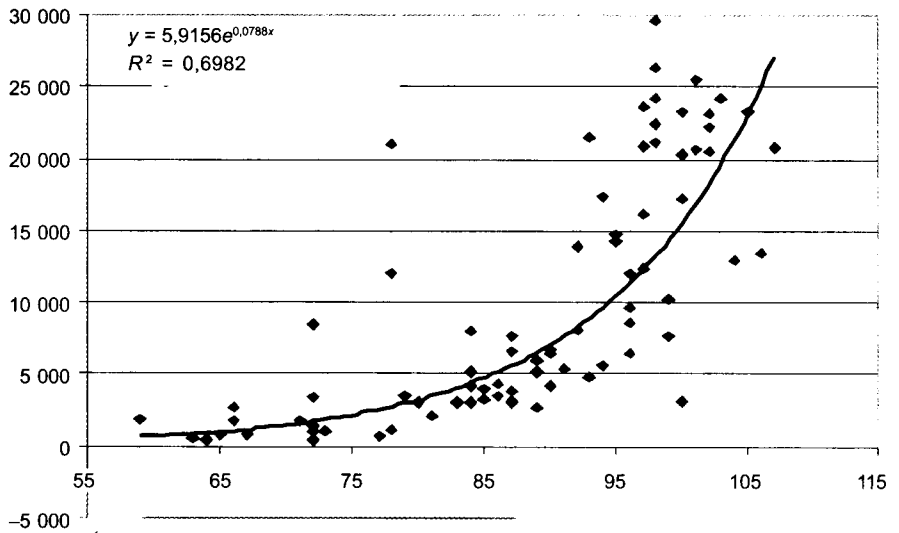
Jak można zauważyć obydwie funkcje są lepiej dopasowane do danych empirycznych, jakkolwiek funkcja wykładnicza jest nieco „lepsza” ($R^2 = 0,698$) od funkcji potęgowej ($R^2 = 0,687$). W przypadku funkcji liniowej współczynnik determinacji był mniejszy i wynosił tylko $R^2 = 0,537$. Parametry oszacowanych funkcji regresji zebrano w tabeli 2. Obok parametrów strukturalnych podano tu odpowiadające im statystyki Studenta, miary dopasowania (współczynniki korelacji i determinacji, średni błąd dopasowania i współczynnik zmienności resztowej), jak również teoretyczne wielkości GDP dla Polski wynikające z oszacowanych funkcji regresji oraz względne wielkości różnicy pomiędzy teoretycznymi a faktyczną (7619 \$/os/rok) wielkością GDP dla Polski.

Ten ostatni parametr można zinterpretować jako tę wielkość produktu narodowego brutto, która powinna uzyskać Polska w relacji do średniego ilorazu inteligencji IQ mieszkańców naszego kraju (oczywiście w sytuacji, gdyby



Źródło: obliczenia własne.

Ryc. 3. Potęgowa funkcja regresji pomiędzy GDP a IQ (81 krajów)



Źródło: obliczenia własne.

Ryc. 4. Wykładnicza funkcja regresji pomiędzy GDP a IQ (81 krajów)

inteligencja była jedynym czynnikiem determinującym poziom GDP). Jak się okazuje, faktyczna wielkość GDP (7619 \$/os/rok) jest o połowę mniejsza niż wielkości teoretyczne wynikające z oszacowanych funkcji regresji. Można to traktować jako szacunek wielkości niewykorzystanego ekonomicznie potencjału tkwiącego w kapitale ludzkim społeczeństwa Polski.

Tabela 2

Parametry liniowej, potęgowej i wykładniczej funkcji regresji pomiędzy GDP a IQ

Parametr	Funkcja liniowa	Funkcja potęgowa	Funkcja wykładnicza
Wyraz wolny [a]	-35 717	1,4545E-09	5,9156
Statystyka Studenta [t(a)]	-7,40	-9,22	3,42
Współczynnik kierunkowy [b]	519,0	6,504	0,0788
Statystyka Studenta [t(b)]	9,57	13,18	13,52
Współczynnik determinacji [R ²]	0,537	0,687	0,698
Współczynnik korelacji [R]	0,732	0,829	0,836
Średni błąd dopasowania [s]	5 813 \$	5 672 \$	5 665 \$
Współczynnik. zmienności resztowej [V]	57,8%	56,4%	56,3%
Wartość teoretyczna GDP dla Polski	15 664 \$	13 893 \$	14 391 \$
Względna różnica między teoretycznym a faktycznym GDP dla Polski	-51,4%	-45,2%	-47,1%

Źródło: obliczenia własne.

Oczywiście interpretacja tych parametrów jest dyskusyjna i trudno ją rekomendować. Pokazuje tylko formalne możliwości wykorzystania klasycznych modeli regresji.

3. MODYFIKACJE ANALIZOWANYCH ZMIENNYCH

W przedstawionej powyżej analizie wykorzystano standardowe podejście, w którym rozpatruje się zależności korelacyjno-regresyjne. Na pierwszy rzut oka jednak widać, że dyskusyjne jest identyczne traktowanie w ocenie współzależności takich krajów jak Chiny (1,3 mld mieszkańców, czyli 1/4 mieszkańców całej badanej zbiorowości) oraz np. Wysp Marshalla (50 tys. mieszkańców) czy też Wysp Tonga (100 tys. mieszkańców), które to państwa mogłyby być traktowane jako niewielkie miasta. Tymczasem na diagramach rozrzutu czy też w obliczeniach wszystkie obiekty (kraje) mają takie samo znaczenie (wagę).

W tabeli 3 podano skumulowane wartości (w procentach) trzech zmiennych — liczby ludności [LUD] oraz iloczynów liczby ludności przez średnie ilorazy inteligencji [IQ*LUD] a także przez wskaźniki produktu narodowego brutto [GDP*LUD]. Kraje uporządkowane są według rosnącej liczby ludności,

Skumulowana liczba ludności (LUD), kapitału intelektualnego (IQ*LUD)
oraz kapitału ekonomicznego (GDP*LUD)

Lp.	Kraj	LUD (tys.)	IQ*LUD/ 100	GDP*LUD/ 1000	% CUM LUD	% CUM (IQ*LUD)	% CUM (GDP*LUD)
1	Marshall Islands	56	47	169	0,00	0,00	0,00
2	Tonga	108	94	325	0,00	0,00	0,00
3	Samoa (Western)	204	177	781	0,01	0,01	0,00
4	Barbados	277	216	3 321	0,01	0,01	0,03
5	Suriname	450	400	2 322	0,02	0,03	0,01
6	Equatorial Guinea	536	316	974	0,03	0,02	0,01
7	Qatar	725	566	15 218	0,05	0,04	0,14
8	Fiji	869	730	3 676	0,06	0,05	0,05
9	Slovenia	2 012	1 911	28 752	0,10	0,10	0,42
10	Jamaica	2 689	1 936	9 114	0,15	0,14	0,09
11	Uruguay	3 387	3 252	29 209	0,22	0,33	0,50
12	Congo (Brazz)	3 413	2 491	3 396	0,28	0,19	0,04
13	Lebanon	3 731	3 209	16 140	0,36	0,26	0,22
14	Puerto Rico	3 878	3 257	31 023	0,43	0,40	0,75
15	Ireland	3 924	3 649	84 296	0,51	0,55	2,62
16	New Zealand	3 960	3 960	68 465	0,58	0,64	2,17
17	Singapore	4 277	4 405	103 541	0,67	0,82	3,18
18	Croatia	4 498	4 048	30 356	0,75	0,72	0,59
19	Norway	4 555	4 464	119 998	0,84	0,91	5,00
20	Finland	5 204	5 048	108 496	0,94	1,02	4,05
21	Denmark	5 394	5 286	130 635	1,04	1,25	6,04
22	Slovakia	5 416	5 200	52 534	1,15	1,13	1,99
23	Sierra Leone	5 571	3 565	2 551	1,26	0,47	0,02
24	Israel	6 492	6 102	112 317	1,38	1,50	4,35
25	Hong Kong	6 810	7 286	141 391	1,51	1,81	6,43
26	Switzerland	7 408	7 482	189 001	1,66	1,97	8,78
27	Bulgaria	7 588	7 057	36 493	1,80	1,65	1,03
28	Austria	8 163	8 326	189 096	1,96	2,49	9,30
29	Guinea	8 756	5 779	15 602	2,13	1,37	0,18
30	Sweden	8 970	9 060	185 318	2,30	2,68	7,76

Lp.	Kraj	LUD (tys.)	IQ*LUD/ 100	GDP*LUD/ 1000	% CUM LUD	% CUM (IQ*LUD)	% CUM (GDP*LUD)
31	Hungary	10 058	9 957	102 911	2,50	3,93	2,90
32	Czech Republic	10 251	9 944	126 724	2,69	3,51	5,69
33	Belgium	10 331	10 331	239 913	2,89	4,16	10,57
34	Portugal	10 480	9 956	154 066	3,10	3,72	7,25
35	Greece	10 626	9 776	148 158	3,30	3,30	6,83
36	Zambia	10 800	8 316	7 765	3,51	2,31	0,07
37	Cuba	11 247	9 560	44 616	3,73	3,09	1,58
38	Guatemala	11 725	9 263	41 097	3,95	2,88	1,46
39	Zimbabwe	11 816	7 798	31 536	4,18	2,13	0,84
40	Ecuador	13 074	10 459	39 261	4,43	4,38	1,23
41	Netherlands	16 223	16 548	359 767	4,75	5,05	16,51
42	Australia	19 767	19 371	443 798	5,13	5,88	19,99
43	Ghana	21 111	14 989	36 628	5,53	4,70	1,13
44	Romania	22 380	21 037	126 404	5,97	6,77	5,34
45	Taiwan	22 543	23 445	293 065	6,40	8,19	12,06
46	Malaysia	23 105	21 257	188 006	6,85	7,23	8,27
47	Iraq	24 683	21 474	78 913	7,32	7,69	2,39
48	Uganda	26 322	19 215	28 270	7,83	5,47	0,35
49	Nepal	26 470	20 646	30 625	8,34	6,32	0,67
50	Peru	27 275	24 548	116 792	8,87	9,22	4,67
51	Morocco	31 754	26 991	104 948	9,48	10,92	3,75
52	Canada	31 889	30 933	752 014	10,10	11,58	28,43
53	Kenya	33 042	23 790	32 381	10,73	8,70	0,93
54	Tanzania	36 199	26 063	17 375	11,43	9,78	0,27
55	Sudan	36 593	26 347	51 010	12,14	10,34	1,84
56	Argentina	38 337	36 803	460 541	12,88	13,09	21,24
57	Poland	38 603	38 217	294 115	13,62	15,49	12,86
58	Spain	40 217	39 011	652 005	14,40	16,33	26,38
59	Colombia	41 802	37 204	251 062	15,21	14,68	11,26
60	South Africa	46 567	33 528	395 258	16,10	12,30	17,59
61	Korea, South	47 657	50 516	642 316	17,02	18,36	24,60
62	Congo (Zaire)	56 886	36 976	46 760	18,12	13,88	1,71

Lp.	Kraj	LUD (tys.)	IQ*LUD/ 100	GDP*LUD/ 1000	% CUM LUD	% CUM (IQ*LUD)	% CUM (GDP*LUD)
63	Italy	57 998	59 158	1 193 896	19,24	22,01	36,40
64	U. Kingdom	60 095	60 095	1 222 085	20,40	23,29	43,04
65	France	62 171	60 928	1 316 479	21,60	24,60	46,63
66	Thailand	63 271	57 577	345 207	22,82	20,74	15,53
67	Iran	63 988	53 750	327 684	24,06	19,51	14,59
68	Turkey	68 109	61 299	437 399	25,37	25,91	18,78
69	Ethiopia	70 366	44 331	40 390	26,73	17,28	1,34
70	Egypt	74 761	62 051	227 347	28,17	27,23	9,92
71	Germany	82 398	84 046	1 826 688	29,76	30,63	51,61
72	Philippines	86 752	74 606	308 402	31,43	28,83	13,70
73	Mexico	103 718	90 235	799 044	33,43	34,45	30,60
74	Japan	127 358	133 726	2 961 959	35,89	37,31	65,66
75	Nigeria	131 728	88 257	104 723	38,43	32,52	3,47
76	Russia	144 308	138 536	932 229	41,22	40,27	33,14
77	Brazil	183 960	160 045	1 218 734	44,77	43,70	39,72
78	Indonesia	223 070	198 532	591 358	49,07	47,95	22,85
79	United States	290 343	284 536	8 595 591	54,67	54,04	100,00
80	India	1 057 504	856 579	2 196 437	75,08	72,36	57,59
81	China	1 291 496	1 291 496	4 010 095	100,00	100,00	76,58

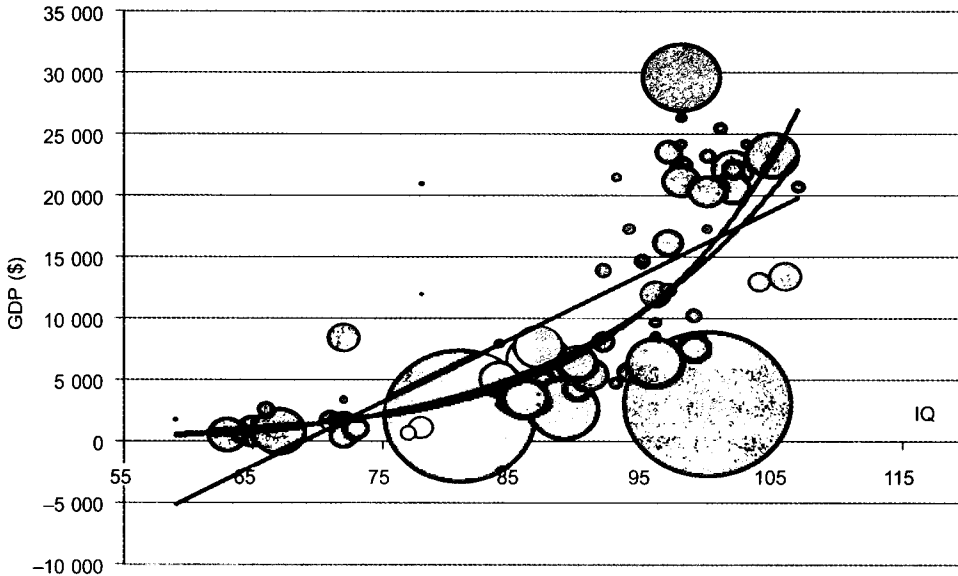
Źródło: obliczenia własne.

wskutek czego kumulanty zmiennych [IQ*LUD] i [GDP*LUD] wskazują na inną kolejność krajów. Zmienną [IQ*LUD] można interpretować jako całkowity zasób kapitału intelektualnego znajdującego się w dyspozycji danego kraju, natomiast zmienną [GDP*LUD] jako całkowity zasób produktu narodowy brutto w tym kraju.

Jak można zauważyć z tabeli 3 łączna liczba mieszkańców połowy (50%) najmniejszych krajów wyczerpuje zaledwie 5% populacji wszystkich analizowanych krajów. Podobnie jest z zasobem kapitału intelektualnego [IQ*LUD]. Z drugiej strony liczba mieszkańców samych tylko Chin odpowiada skumulowanej liczbie mieszkańców 67 krajów spośród 81, a zasób kapitału intelektualnego Chin jest taki jak kapitał 70 mniejszych krajów. Liczba mieszkańców i kapitał intelektualny dwóch największych krajów — Chin i Indii odpowiada prawie połowie zasobów tych zmiennych w całym analizowanym zbiorze krajów.

Kumulanty liczby ludności i kapitału intelektualnego praktycznie się pokrywają ($r = 0,99$) natomiast kumulanta całkowitego produktu narodowego [GDP*LUD] kształtuje się nieco inaczej niż poprzednie dwie kumulanty. Wynika to stąd, że największe udziały w zakresie [GDP*LUD] mają Stany Zjednoczone (30% ogólnej sumy), a nie Chiny (14,2%) czy Indie (7,7%). Ponadto dochodzi tu jeszcze Japonia (10% potencjału gospodarczego wobec 2,5% udziału w liczbie ludności oraz 4% udziału w zasobie kapitału intelektualnego).

Zróznicowane wielkości krajów powodują, że uzasadnionym wydaje się zaproponowanie zmodyfikowanej formuły wyznaczania miar zależności. Na rycinie 5 przedstawiono bąbelkowy diagram rozrzutu wraz z wszystkimi trzema funkcjami regresji. Średnica poszczególnych okręgów(krajów) jest tu proporcjonalna do liczby ludności [LUD]. Wykres ten lepiej oddaje istotę rozpatrywanej zależności i sugeruje uwzględnienie wielkości obiektów w trakcie wyznaczania miar korelacji i regresji.



Źródło: obliczenia własne.

Ryc. 5. Bąbelkowy diagram rozrzutu pomiędzy GDP a IQ wraz z funkcjami regresji

4. PROPOZYCJE ZMODYFIKOWANYCH MIERNKÓW WSPÓLZALEŻNOŚCI

Zgodnie z dotychczas stosowanym podejściem w analizie zależności w większości badań empirycznych współczynnik korelacji liniowej wyznaczany jest według następującego wzoru:

$$r = \frac{\text{cov}(X, Y)}{s_X \cdot s_Y} = \frac{\sum_{i=1}^n X_i Y_i - \frac{\left(\sum_{i=1}^n X_i\right) \left(\sum_{i=1}^n Y_i\right)}{n}}{\sqrt{\left(\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}\right) \left(\sum_{i=1}^n Y_i^2 - \frac{\left(\sum_{i=1}^n Y_i\right)^2}{n}\right)}}, \quad (1)$$

gdzie $\text{cov}(X, Y)$ to kowariancja pomiędzy zmiennymi X i Y natomiast s_X, s_Y to ich odchylenia standardowe.

W omawianym przykładzie zmienną Y jest wskaźnik produktu narodowego brutto (GDP), natomiast zmienną X — średni iloraz inteligencji (IQ). Wzór (1) można przekształcić w ten sposób, że przy każdym jego elemencie pojawi się waga F_i (równa liczbie ludności dla i -tego kraju — LUD).

$$r = \frac{\sum_{i=1}^n F_i X_i Y_i - \frac{\left(\sum_{i=1}^n F_i X_i\right) \left(\sum_{i=1}^n F_i Y_i\right)}{\sum_{i=1}^n F_i}}{\sqrt{\left(\sum_{i=1}^n F_i X_i^2 - \frac{\left(\sum_{i=1}^n F_i X_i\right)^2}{\sum_{i=1}^n F_i}\right) \left(\sum_{i=1}^n F_i Y_i^2 - \frac{\left(\sum_{i=1}^n F_i Y_i\right)^2}{\sum_{i=1}^n F_i}\right)}}. \quad (2)$$

Aby przybliżyć interpretację współczynnika danego wzorem (2) należy zauważyć, że mierzy on zależność pomiędzy kapitałem intelektualnym a potencjałem gospodarczym w sytuacji, w której jednostkami obserwacji są wszystkie osoby (a nie kraje) zamieszkałe w analizowanych 81 krajach, przy czym wszyscy mieszkańcy danego kraju mają jednakowy (na średnim poziomie) iloraz inteligencji oraz są twórcami jednakowego (na średnim poziomie) produktu narodowego brutto.

Inną propozycją jest wyznaczenie współczynnika korelacji liniowej pomiędzy zmiennymi $[IQ \cdot LUD]$ oraz $[GDP \cdot LUD]$, co sprowadza się do następującej formuły:

$$r = \frac{\sum_{i=1}^n (F_i X_i)(F_i Y_i) - \frac{\left(\sum_{i=1}^n (F_i X_i)\right) \left(\sum_{i=1}^n (F_i Y_i)\right)}{n}}{\sqrt{\left(\sum_{i=1}^n (F_i X_i)^2 - \frac{\left(\sum_{i=1}^n (F_i X_i)\right)^2}{n}\right) \left(\sum_{i=1}^n (F_i Y_i)^2 - \frac{\left(\sum_{i=1}^n (F_i Y_i)\right)^2}{n}\right)}}, \quad (3)$$

względnie do formuły zmodyfikowanej, w której przez liczbę obserwacji rozumie się sumę wszystkich realizacji zmiennej wykorzystywanej w charakterze wag F :

$$r = \frac{\sum_{i=1}^n (F_i X_i)(F_i Y_i) - \frac{\left(\sum_{i=1}^n (F_i X_i)\right)\left(\sum_{i=1}^n (F_i Y_i)\right)}{\sum_{i=1}^n F_i}}{\sqrt{\left(\sum_{i=1}^n (F_i X_i)^2 - \frac{\left(\sum_{i=1}^n (F_i X_i)\right)^2}{\sum_{i=1}^n F_i}\right)\left(\sum_{i=1}^n (F_i Y_i)^2 - \frac{\left(\sum_{i=1}^n (F_i Y_i)\right)^2}{\sum_{i=1}^n F_i}\right)}} \quad (4)$$

Analogicznie do przedstawionych powyżej wzorów można podać cztery różne formuły służące do wyznaczania parametrów liniowej funkcji regresji $Y = f(X)$. Zebrano je w tabeli 4.

Tabela 4

Wzory na nieważone (1) i ważone (2)–(4) parametry liniowej funkcji regresji

Wzór	Współczynnik kierunkowy	Wyraz wolny
(1)	$b = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i\right)^2}$	$a = \frac{\sum_{i=1}^n Y_i}{n} - b \frac{\sum_{i=1}^n X_i}{n}$
(2)	$b = \frac{\sum_{i=1}^n F_i \sum_{i=1}^n F_i X_i Y_i - \sum_{i=1}^n F_i X_i \sum_{i=1}^n F_i Y_i}{\sum_{i=1}^n F_i \sum_{i=1}^n F_i X_i^2 - \left(\sum_{i=1}^n F_i X_i\right)^2}$	$a = \frac{\sum_{i=1}^n (F_i Y_i)}{\sum_{i=1}^n F_i} - b \frac{\sum_{i=1}^n (F_i X_i)}{\sum_{i=1}^n F_i}$
(3)	$b = \frac{n \sum_{i=1}^n (F_i X_i)(F_i Y_i) - \sum_{i=1}^n (F_i X_i) \sum_{i=1}^n (F_i Y_i)}{n \sum_{i=1}^n (F_i X_i)^2 - \left(\sum_{i=1}^n (F_i X_i)\right)^2}$	$a = \frac{\sum_{i=1}^n (F_i Y_i)}{n} - b \frac{\sum_{i=1}^n (F_i X_i)}{n}$
(4)	$b = \frac{\sum_{i=1}^n F_i \sum_{i=1}^n (F_i X_i)(F_i Y_i) - \sum_{i=1}^n (F_i X_i) \sum_{i=1}^n (F_i Y_i)}{\sum_{i=1}^n F_i \sum_{i=1}^n (F_i X_i)^2 - \left(\sum_{i=1}^n (F_i X_i)\right)^2}$	$a = \frac{\sum_{i=1}^n (F_i Y_i)}{\sum_{i=1}^n F_i} - b \frac{\sum_{i=1}^n (F_i X_i)}{\sum_{i=1}^n F_i}$

Źródło: obliczenia własne.

Wartości współczynników regresji oraz korelacji liniowej wyznaczone za pomocą powyższych wzorów przedstawiono w tabeli 5. Jak można zauważyć ważne miary zależności (2) dają nieco inne rezultaty aniżeli standardowe miary nieważone (1). Współczynnik korelacji zmniejszył się z 0,73 do 0,48. Również zmniejszyły się parametry linowej funkcji regresji — współczynnik kierunkowy z 519 \$ GDP/os/rok do 369 \$ GDP/os/rok i podobnie wyraz wolny, z -35717 \$ GDP do -26216 \$ GDP.

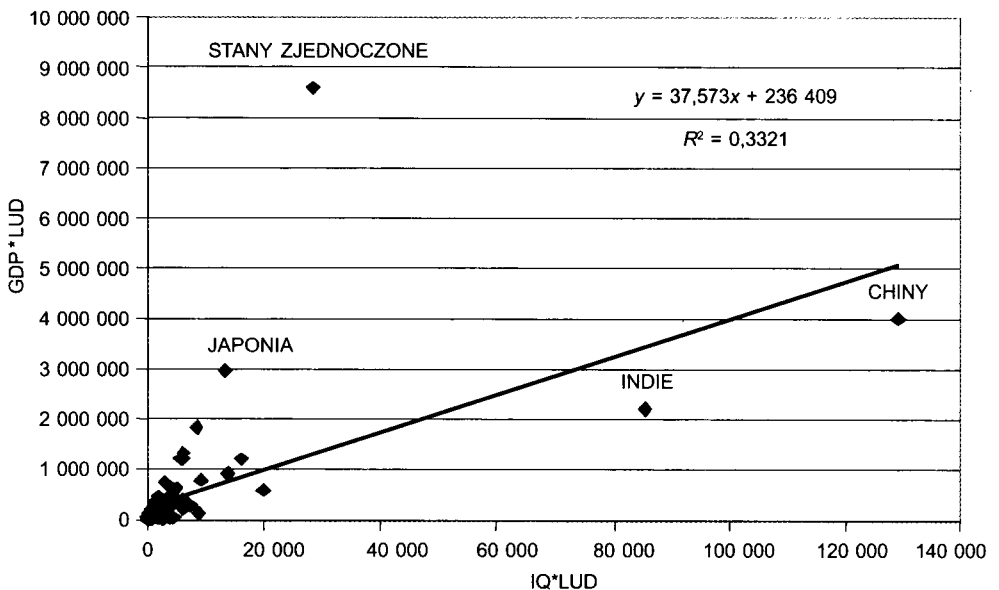
Tabela 5

Ważone parametry regresji i współczynniki korelacji pomiędzy GDP a IQ

Wzór	Współczynnik a	Współczynnik b	Współczynnik r
(1)	-35 717	519	0,733
(2)	-26 215	369	0,481
(3)	236 409	38	0,576
(4)	3 326	42	0,625

Źródło: obliczenia własne.

Miary zależności opisane za pomocą trzeciej i czwartej formuły dają odmiennie rezultaty. Jak wynika z ryciny 6 kształt diagramu rozrzutu dla zmiennych [IQ*LUD] oraz [GDP*LUD] determinują 3–4 kraje: Chiny, Stany Zjedno-



Źródło: obliczenia własne.

Ryc. 6. Diagram rozrzutu wraz z liniową funkcją regresji pomiędzy zmiennymi GDP*LUD oraz IQ*LUD

zione, Indie oraz Japonia i one decydują o ocenie kierunku oraz sile zależności pomiędzy potencjałem intelektualnym a bogactwem. Pozostałe kraje tworzą jedno skupisko zgrupowane w pobliżu początku układu współrzędnych (0, 0). „Wartość” jednostkowa kapitału intelektualnego spada tu z 369–519 \$ do 38–42 \$ GDP na 1 osobę w ciągu roku.

BIBLIOGRAFIA

- Herrnstein R., Murray Ch. 1994. *The Bell Curve: Intelligence and Class Structure in American Life*, Free Press 1994.
- [http://www.amren.com/ar/2002/12/ The Global Bell Curve. Uncommon sense on wealth and poverty., reviewed by Thomas Jackson.](http://www.amren.com/ar/2002/12/The_Global_Bell_Curve._Uncommon_sense_on_wealth_and_poverty._reviewed_by_Thomas_Jackson.)
- [http://www.census.gov/ipc/www/idb/tables.html.](http://www.census.gov/ipc/www/idb/tables.html)
- [http://www.lagriffedulion.f2s.com/sft.htm.](http://www.lagriffedulion.f2s.com/sft.htm)
- Human Development Report 2000*. 2000. United Nations Development Programme (UNDP), Oxford University Press, New York.
- Jensen A.R. 1998. *The g Factor*, Praeger, Westport (CT).
- Landes D.S. 1998. *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*, W.W. Norton & Company, New York.
- Lynn R. 2008. *The Global Bell Curve*, Augusta, GA: Washington Summit Publisher, 2008.
- Lynn R., Vanhanen T. *IQ and the Wealth of Nations*. Westport, CT Praeger, 2002.
- Lynn R., Vanhanen T. *IQ and Global Inequality*. Washington Summit Publishers (GA), Augusta.
- Rushton J.P. 2000. *Race, Evolution, and Behavior: A Life History Perspective*, Charles Darwin Research Institute (MI), Port Huron.